

The Development and Application of Approaches to Discover Siglec Ligands

by

Edward N. Schmidt

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

University of Alberta

© Edward N. Schmidt, 2024

Abstract

Across all domains of life, the surface of all cells is coated in a dense array of carbohydrates appended to proteins and lipids, collectively referred to as glycans. The interaction between glycans and glycan-binding proteins, also known as lectins, are at the interface of nearly every biological interaction at the cellular level. Unlike other biological macromolecules (nucleic acids and proteins), glycan biosynthesis is not template driven and glycans can be branched polymers, resulting in heterogenous structures that are difficult to predict, and which are very structurally complicated. Glycans terminating with the monosaccharide sialic acid are capable of engaging with the Siglec lectin family. Siglecs facilitate many cellular functions, however, all Siglec functions are initiated by engaging with a glycan terminating with sialic acid. In humans, Siglecs are a family of fifteen cell surface receptors, which are generally found on immune cells. The immunomodulatory properties of Siglecs motivates a better understanding of their physiological and pathophysiological roles. Therefore, the first step in understanding the biology of a Siglec starts with understanding the ligands of that Siglec. Nevertheless, the complexity and diversity of glycans coupled with multiple Siglecs often being expressed by a single cell makes deconvoluting Siglec-glycan interactions challenging. Through the development and optimization of approaches that enable the systematic description of Siglec ligands a better understanding of the roles of Siglecs in health and disease can be established. Insights gained by describing Siglec ligands can then be leveraged to investigate the function of the Siglec and potentially apply these discoveries to develop novel therapeutics towards the treatment of many pathologies such as cancer, bacterial/viral infection, autoimmune disease, neurodegenerative disease and many others.

In Chapter 2, improvements with respect to versatility, activity, yield, and storage of a previously established soluble version of a Siglec referred to as a Siglec-Fc-chimera were made. The versatility of this construct was demonstrated by its use in many different approaches for describing Siglec-ligands including a cell-based glycan array, bead assay, ELISA, and mass spectrometry-based assay. Using these different approaches many insights were made with respect to Siglec ligands. Two examples of new insights are: the relationship between *cis* and *trans* binding of Siglecs on primary human immune cells and the glycan ligands of CD33. It is noteworthy that the tools and approaches developed in this Chapter were extensively applied in the later Chapters.

In Chapter 3, a novel liposomal nanoparticle formulation was developed for optimal engagement of Siglecs by glycolipids. Using this formulation, the human and murine Siglec family was screened against a panel of glycolipids. During this interrogation, it was revealed that glycolipids are ligands for Siglec-6. Additionally, the glycolipid binding profiles between human and murine Siglecs were compared and the appropriateness of murine models to study humans Siglec-glycolipid interactions was assessed.

Building on the work from Chapter 3, the ability of Siglec-6 to bind glycolipids was systematically dissected using a panel of synthetic glycolipids referred to as neoglycolipids coupled with mutagenesis studies. These studies revealed that a solvent exposed tryptophan residue is critical for Siglec-6 engagement of sialosides when presented from a liposome. Additionally, Siglec-6-liposome binding was probed on genuine human cells and tissues demonstrating that glycolipid bearing liposomes can target Siglec-6 in physiologically relevant contexts. A possible biological role was ascribed to Siglec-6 when it was found that Siglec-6 can facilitate the internalization of natural nanoparticles, known as extracellular vesicles, in a glycolipid-dependant manner. In summary, the work described above demonstrates how systematic description of Siglec ligands leads to the development of probes which can be applied to learn more about the biological roles Siglecs play in health and disease.

Preface

This thesis is submitted for the degree of Doctor of Philosophy at the University of Alberta, Department of Chemistry. It represents a summary of the research completed by Edward N. Schmidt under the supervision of Dr. Matthew S. Macauley from 2019-2024. The contributions from coauthors and collaborators for this thesis are summarized below.

In Chapter 2, the introduction was adapted from a published book chapter Schmidt E.N., Jung J., and Macauley M.S.; *Flow cytometry-based detection of Siglec ligands; Carbohydrate-Protein Interactions: Methods and Protocols*, 181-193, Springer US, **2023**. I worked closely with Dr. Jaesoo Jung to develop and validate Siglec-Fc chimeras produced from CHO cells. Lentiviruses were produced by Dr. Jaesoo Jung. Dr. Jaesoo Jung and Christopher D. St. Laurent generated knock out cell lines. Mass spectrometry experiments were performed by Dr. Elena N. Kitova and Dr. Duong T. Bui under the supervision of Prof. John S. Klassen. Human spleen samples were provided by Prof. Lori J. West. All experiments involving human sample collection were approved by the human research ethics board (HREB) biomedical panel at the University of Alberta.

In Chapter 3, large portions of the results section were published as Schmidt E.N. *et al.*, *Siglec-6 mediates the uptake of extracellular vesicles through a noncanonical glycolipid binding pocket*, *Nat Commun*, **2023** and Schmidt E.N. *et al.*, *Dissecting the abilities of murine Siglecs to interact with gangliosides*, *J. Biol. Chem.*, **2024**. In this chapter, I worked with Dr. Jaesoo Jung to develop and validate cell lines expressing full length human Siglecs. Dr. Jaesoo Jung developed CHO cell lines which express soluble murine Siglecs. Dr. Mirat Sojitra under the supervision of Prof. Ratmir Derda performed LiGA binding experiments. Mass spectrometry experiments were performed by Dr. Ling Han, Dr. Duong T. Bui and Dr. Elena N. Kitova under the supervision of Prof. John S. Klassen. Xue Yan Guo assisted in performing glycolipid binding assays.

In Chapter 4, large portions of the results section were published as Schmidt E.N. *et al.*, *Siglec-6 mediates the uptake of extracellular vesicles through a noncanonical glycolipid binding pocket*, *Nat Commun*, **2023**. Neoglycolipids were synthesized and characterized by Dr. Maju Joe and Fahima Mozaneh. Dr. Kelli A. McCord generated B cell lines transduced with Siglec-6 as well as performed imaging flow cytometry. Prof. Elisa Fadda identified a solvent exposed tryptophan residue in Siglec-6. Jasmine Nguyen and Dr. Dimitra Lamprinaki under the supervision of Prof. Meghan R. Riddell and Prof. Lara K. Mahal respectively performed experiments with human placental explants. Human spleen samples and LAD2 cells were provided by Prof. Lori J. West and Prof. Marianne Kulka, respectively. Extracellular vesicles were isolated and characterized from human blood by Dr. Dimitra Lamprinaki under the supervision of Prof. Lara K. Mahal. β 1-4GalNT1^{-/-} N2a cells were produced by Amanda R. Kryslar and Christopher R. Cromwell under the supervision of Prof. Basil P. Hubbard. EVs from β 1-4GalNT1^{-/-} N2a cells were isolated by John Monyror under the supervision of Prof. Simonetta Sipione. All experiments involving human blood samples and placental sample collection were approved by the HREB biomedical panel at the University of Alberta.

Acknowledgments

I would like to start this section off with a quote from Eric Church:

*And I hope they know I never would've made it this far on my own
'Cause where would we all be without those
Fathers and mothers, sisters and brothers
The friends I've made, the long lost lovers
I wouldn't be who I am today
If not for those I've loved along the way*

To my supervisor, Prof. Matthew S. Macauley, I thank you for the opportunity to work in your lab and support during my PhD studies. You were always available to lend an ear to help with a problem whether it was related to the lab or otherwise. No matter the time or day your support was always felt. I truly feel I have become a better scientist and person during my time in your lab. To my committee members, Prof. Lara K. Mahal, and Prof. Christopher W. Cairo, thank you for all the support over the years and always pushing me to be the best version of myself. I would also like to thank Prof. Michael S. Boyce for making the journey to Edmonton to be my external examiner and to Prof. Joanne Lemieux for being a part of the examining committee. I would also like to acknowledge the funding agencies (GlycoNET, NSERC, and Alberta Innovates) who supported me during my PhD studies. Additionally, I would like to acknowledge that the University of Alberta is situated on Treaty 6 territory, traditional lands of First Nations and Métis people. I would also like to thank Prof. Geoffrey P. Horsman for the opportunity you provided me to work in your lab which gave me the confidence I needed to apply to graduate school. I also think it is very important to acknowledge all the scientists who came before. Scientific progress is not always fast but thanks to the efforts of the generations of scientists who came before, I was able to hit the ground running. We do truly "...Stand on the shoulders of giants"-Isaac Newton

I would like to dedicate this thesis to the memory of Maria Kalvaitis. When my loving parents introduced me to Maria, I did not appreciate what they did for me at the time. Now with everything I write, present, or read I think about the countless hours she worked with me and how she changed my life. Without her, this thesis would not exist. Thank you Maria. I would also like

to acknowledge my family. My loving parents Ed and Gail Schmidt (néé Billings) whose unconditional support always made me feel like nothing was out of reach as long as I did not give up. My dear sister Allie Schmidt, who has never been anything but supportive of my dreams and I needed her support many times and will continue to need her support in the future. To my grandparents, Rita (néé Pletz) and Ed Schmidt, Bruce (rest in peace) and Anne (néé Peters) Billings I always felt like I was at home when I heard your voices over the phone, which, was something I really needed from time to time. I would also like to thank Andreas Arth (Uncle Hoops). Thank you for the opportunity to work with you while also working in the lab.

To John and Pam Catricala, the experiences, laughs, high and lows we shared are some of the best memories of my life. I learned so much from you and I apply the lessons I learned from you in and outside of the lab every day. To my friends who always stayed in touch after we moved apart which made moving across the country feel not so far. Thanks to Grant Trenwith, Ryan Poetker, Dana Sowa, Joel Mitchell, Cole Williamson, Renee Hordyk, Joanne McKay, Mr. Watson (Warren Bradly Watson), and Alena Pratasouskaya.

I consider myself beyond lucky to have met the people I did through my time at the University of Alberta. While I knew Zachary W. Schroeder before moving to Alberta, you have always been a great friend, and I felt much better about moving west when I heard you were headed that way too. To Kelli, A McCord, I do not know how to thank you enough for everything you did for me in and outside the lab. From the recruitment weekend to your last day, you were always there making the lab a better place to be. There were some BIG curves in the road during our degrees, but thanks to you I made it through. For all the laughs and all the cries, you will always be one of my best friends and I look forward for what the future holds for us. To Xue Yan Guo, you were the first person who I ever mentored, and I know that I was not always the best mentor. Thank you for your patients and all your help with all the projects we worked on and helping me become a better person. I hope we can remain friends for a long time. To Christopher D. St. Laurent, I hope one day you'll miss me! Until then, I want to thank you for all your help and

support during my degree. Whenever I had a problem, you were always there to give advice which always made me feel at home in the lab. To Susmita Sarkar, I know you'll miss me 😊. Thank you for everything and you always made the lab a better place and were always ready to lend a hand even at 9 pm on a Friday. To Fahima Mozaneh and Jhon R. Enterina, thank you for all your help and your company. To Santiago Tijaro Bulla and Dhanraj Kumawat thanks for being with me from the start.

To Jaesoo Jung, I will always remember how you were the first one of the group to introduce yourself to me and how when I was scared of the autoclave, you came with me every time until I was comfortable, we'll keep how many times that was between us. I would also like to acknowledge Mike Barteski and Ryan Lewis. I cannot quantify the help and support you gave me over the years. Thank you for all your help and friendship. To Dr. Kei Takahashi-Yamashiro, thank you for working next to me and answering my ridiculous questions, I always enjoyed talking with you. To Dr. Maju Joe thank you for supporting so much of my academic goals and for all the chats, I always felt good after chatting with you. To Dr. Ghazaleh Eskandari-Sedighi, thank you for all the wisdom and friendship. You made the lab a better place to be.

Last but certainly not least, I would like to thank Zeinab Jame-Chenarboo. I do not have the words to describe what you being in my life means to me, but I'll try. I never knew support like I did when you came into my life. It has been quite the ride with ups and downs, tears, and laughs but I look forward to seeing where our next chapter takes us. Thank you for everything.

Table of Contents:

Abstract	ii
Preface	iv
Acknowledgments	vi
Table of Figures	xvi
Table of Tables	xxiv
List of Abbreviations	xxv
Chapter 1: Glycobiology and the Siglecs	1
1.1: Introduction to Glycobiology	2
1.1.1: Glycans as biological polymers	2
1.1.2: Glycans as a family of biological macromolecules	2
1.2: Types of Glycans	5
1.2.1: Building blocks and glycosyltransferases	5
1.2.2: <i>N</i> -glycans	9
1.2.3: <i>O</i> -glycans	11
1.2.3.1: <i>O</i> -GalNAc glycans	11
1.2.3.2: <i>O</i> -Mannose glycans	13
1.2.3.3: <i>O</i> -Fucose glycans	13
1.2.3.4: <i>O</i> -Xylose glycans	14
1.2.3.5: <i>O</i> -GlcNAc glycans	14
1.2.4: Glycolipids	15
1.2.5: GPI-anchored proteins	16

1.2.6: C-mannose	17
1.2.7: Extensions and capping of glycans	17
1.2.8: Sialic acid as a cap for glycans	21
1.2.9: Sialic acid as a self-associated molecular pattern.....	24
1.3: The Siglec family of cell surface receptors.....	26
1.3.1: Siglecs across species	27
1.4: Siglec ligand binding.....	29
1.4.1 <i>cis vs trans</i> binding.....	29
1.4.2 Siglec-ligand interactions are driven by non-covalent interactions.....	29
1.4.3: Relative strength of Siglec-glycan interactions.....	31
1.5: Siglecs in biological systems.....	33
1.5.1: Siglec functions	33
1.6: Current state of Siglec ligands.....	35
1.6.1 Approaches to describe Siglec Ligands	35
1.6.2 Plate-based glycan arrays.....	35
1.6.3 Cell-based glycan arrays.....	36
1.6.4 Nanoparticle-based glycan arrays.....	37
1.6.5 Mass spectrometry-based glycan arrays.....	38
1.7: Thesis statement and objectives	41
1.7.1: Chapter 2: The versatility of Siglec-Fcs.....	42
1.7.2: Chapter 3: Liposomes as tools to study Siglec ligands	42
1.7.3: Chapter 4: Exploring the ligands of Siglec-6	43
Chapter 2: Siglec-Fcs as Versatile Tools to Study Siglec Ligands.....	45
2.1: Acknowledgements	45
2.2: Introduction.....	46

2.2.1: Importance of describing Siglec ligands.....	46
2.2.2: Siglec-Fc chimeric proteins	46
2.2.3: Challenges with the previous versions of the Siglec-Fc	47
2.2.4: Flow cytometry	49
2.2.5: Siglec-Fc protein production.....	52
2.2.6: Siglec-Fc expression	52
2.2.7: Siglec-Fc purification	53
2.2.8: Approaches for modulating cellular glycans.....	54
2.2.8.1: Enzymatic digest of cellular glycans	55
2.2.8.2: Introduction of exogenous glycan modifying enzymes	56
2.2.8.3: Knock-out of glyco-genes	58
2.3: Results.....	59
2.3.1: Cell-based glycan array.....	59
2.3.1.1: Comparison of Strep-Tactin and anti-hIgG as a secondary	59
2.3.1.2: Enzymatic treatment of cell glycans.....	60
2.3.1.3: Overexpression of glycan modifying enzymes.....	61
2.3.4: Applying engineered cells to describe Siglec ligands.....	63
2.3.4.1: Describing the ligands of hCD33	64
2.3.5 Probing Siglec ligands on primary cells.....	66
2.3.5.1 Strep-Tactin is a superior secondary when profiling Siglec ligands on primary cells.....	66
2.3.5.2 Siglec expression on primary cells.....	67
2.3.5.3 Siglec ligand expression on primary cells	69
2.3.6: Other Siglec-Fc-based approaches.....	70
2.3.6.1: Siglec-Fc ELISA.....	70
2.3.6.2: Siglec-Fc bead assay.....	71
2.3.6.3: Siglec-Fc mass-spectrometry-based assay	73
2.3.7: Improving the Siglec-Fc.....	76
2.3.7.1: Improving Siglec-Fc activity through amino acid selection	76

2.3.7.2: Improving Siglec-Fc expression and activity	77
2.4: Discussion.....	81
2.4.1: Siglec-Fcs as tool to describe Siglec ligands	81
2.4.2: Siglec expression by primary splenocytes cells	81
2.4.3: Siglec ligands expressed by primary splenocytes.....	82
2.4.4: Relationship between <i>cis</i> and <i>trans</i> Siglec ligands	83
2.4.5: Siglec-Fcs combine with engineered cell lines for describing Siglec-ligands.....	84
2.4.6: Improvements to the Siglec-Fc chimeras	86
2.5: Conclusion	88
2.6: Methods	88
Chapter 3: Glycolipid bearing liposomes as tools for Discovering Siglec Ligands	96
3.1: Acknowledgements	96
3.2: Introduction.....	97
3.2.1: Glycolipids as Siglec ligands	97
3.2.2 Chemical characteristics of gangliosides	98
3.2.3: Biosynthesis of gangliosides	99
3.2.4: Roles of Siglec-ganglioside interactions in disease	104
3.2.5: Murine models to study human physiology and pathophysiology	104
3.3: Results.....	107
3.3.1: Cell-based Siglec-ganglioside binding assay	107
3.3.2 Optimization of liposome formulation for Siglec binding.....	108
3.3.3: Ganglioside crowding impedes Siglec-ganglioside interactions.....	113
3.3.4 Human Siglec-ganglioside interactions	118
3.3.5 Development and validation of Siglec expressing CHO cells.....	118
3.3.6: Interrogating human Siglecs with ganglioside liposomes.....	119
3.3.7: Unimodal binding profiles may be a general Siglec phenomenon	121

3.3.8 Mouse Siglec-ganglioside interactions	122
3.3.8.1 Profiling mouse Siglec-ganglioside interactions within the context of a bilayer	122
3.3.8.2 Profiling mouse Siglec-ganglioside interactions outside the context of a bilayer	123
3.3.8.3: Re-optimization of liposome formulation does not reveal GM1a-mSiglec-1 interaction.	124
3.3.8.4: Murine Siglec-1 prefers terminally linked $\alpha(2\rightarrow3)$ linked gangliosides.....	126
3.4: Discussion.....	130
3.4.1: Siglec-ganglioside interactions.....	130
3.4.2: Liposome formulation effects Siglec-ganglioside binding.....	130
3.4.3: Siglec-ganglioside binding partners	131
3.4.4: Mouse models are ideal for studying human Siglec-ganglioside interactions.....	135
3.4.5: GM1a is a poor ligand for mSiglec-1	135
3.5: Conclusion	137
3.6: Methods	138
Chapter 4: Exploring the ligands of and functions of Siglec-6	143
4.1: Acknowledgements	143
4.2: Introduction.....	144
4.2.0: Overview of Siglec-6	144
4.2.1: Siglec-6 structure	144
4.2.2: Siglec-6 expression	145
4.3: Results.....	148
4.3.1: Siglec-6 ligand presentation	148
4.3.1.0: Siglec-6-ganglioside binding outside the context of a bilayer	148
4.3.1.1: Siglec-6-Fc does not bind in cell-based glycan array	149
4.3.1.2: Bilayer presentation rescues Siglec-6 ganglioside binding	151
4.3.2: Probing Siglec-6 with synthetic glycolipids.....	153
4.3.2.1: Siglec-6 binding is dependent on oligosaccharide presentation	153

4.3.2.2: Optimizing neoglycolipid liposomes	155
4.3.3: Identifying which domain(s) of Siglec-6 are important for glycolipid recognition	160
4.3.3.1: Identifying which domain(s) of Siglec-6 are important for glycolipid recognition.....	160
4.3.3.2: Mutagenesis analysis of Siglec-6.....	162
4.3.3.3: Bilayer binding residues in Siglec-6	163
4.3.4: Targeting physiological Siglec-6 ⁺ cells and tissues <i>ex vivo</i>	166
4.3.4.0: Targeting Siglec-6 on memory B cells	166
4.3.4.1: Targeting Siglec-6 on mast cells	168
4.3.4.2: Targeting Siglec-6 on placental syncytiotrophoblast.....	169
4.3.5: Siglec-6 internalizes neoglycolipid liposomes	170
4.3.6: Siglec-6 binding to extracellular vesicles.....	173
4.3.6.0: Siglec-6 binds extracellular vesicles	173
4.3.6.1: Siglec-6 binds extracellular independent of canonical arginine residue	174
4.3.6.2: Siglec-6 binds extracellular through $\alpha(2\rightarrow3)$ gangliosides.....	175
4.3.5.3: Siglec-6 internalizes Extracellular Vesicles.....	177
4.4: Discussion.....	179
4.4.0: Bilayer presentation is important for glycolipid recognition by Siglec-6.....	179
4.4.1: Siglec-6 recognizes $\alpha(2\rightarrow3)$ linked glycolipids	180
4.4.2: Trp127 is critical for Siglec-6 binding to glycolipid bearing liposomes	180
4.4.3: Targeting Siglec-6 under physiological conditions	183
4.4.4: Biological roles of Siglec-6	184
4.5: Conclusion	184
4.5: Methods	185
Chapter 5: Conclusion and Future Directions	191
References	200
Appendices	225

Appendix A2:	225
Appendix A3:	236
Appendix A4:	243

Table of Figures

Figure 1.1: Comparison of biological macromolecule families	3
Figure 1.2: Proposed mechanism of interconversion of isomers of glucose	4
Figure 1.3: Chemical structures of the nine monosaccharide building blocks in mammals	6
Figure 1.4: Chemical structures of the nine monosaccharide donors	7
Figure 1.5: Proposed general mechanisms of glycosyltransferases and glycosidases	8
Figure 1.6: Major types of glycans	9
Figure 1.7: Biosynthesis and classes of <i>N</i> -glycans	10
Figure 1.8: β -elimination mechanism of <i>O</i> -glycans	12
Figure 1.9: <i>O</i> -GalNAc <i>O</i> -glycans	13
Figure 1.10: <i>O</i> -Mannose <i>O</i> -glycans	13
Figure 1.11: <i>O</i> -Fucose <i>O</i> -glycans	14
Figure 1.12: <i>O</i> -Xylose <i>O</i> -glycans	14
Figure 1.13: Structure of <i>O</i> -GlcNAc modification	15
Figure 1.14: Glycolipids as a type of glycan	16
Figure 1.15: Glycosylphosphatidylinositol as a glycan	17
Figure 1.16: <i>C</i> -Mannose as a glycan	17
Figure 1.17: Glycan extension and capping of different glycan cores	19
Figure 1.18: Chemical structures of mono- and di-sulfated LacNAc	20
Figure 1.19: SNFG depiction of the glycocalyx featuring <i>N</i> -glycans, <i>O</i> -glycans, glycolipids GPI-anchored glycans <i>C</i> -mannose glycans, and <i>O</i> -GlcNAc modified glycans	21
Figure 1.20: Biosynthesis of <i>N</i> -acetylneuraminic acid (Neu5Ac)	22
Figure 1.21: Different types of sialic acid	24
Figure 1.22: Biological roles of sialic acid	25
Figure 1.23: The human Siglec family	27
Figure 1.24: Evolution of Siglec-sialic acid axis	28
Figure 1.25: Siglec binding can be either in <i>cis</i> or <i>trans</i> and depending on the relationship between the ligand and the Siglec	29

Figure 1.26: Non-covalent interactions drive the interaction between a Siglec and sialosides.....	31
Figure 1.27: Protein ligand binding and comparison of the binding strength between receptor ligand pairs in biological systems.	32
Figure 1.28: Biological roles of Siglecs.	34
Figure 1.29: Glycan-arrays for discovering Siglec ligands.....	36
Figure 1.30: Cell-Based arrays for discovering Siglec ligands.....	37
Figure 1.31: Nanoparticle-based arrays for discovering Siglec ligands	38
Figure 1.32: Mass spectrometry-based approaches for discovering Siglec ligands	39
Figure 1.33: Thesis statement visual abstract.	42
Figure 1.34: Visual abstract of Chapter 2	42
Figure 1.35: Visual abstract for Chapter 3	43
Figure 1.36: Visual abstract for Chapter 4	44
Figure 2.1: Siglec-Fc chimera concept.....	47
Figure 2.2: Key contributions to the development of Siglec-Fc chimera over time	48
Figure 2.3: Flow cytometry as an approach to discover Siglec ligands	51
Figure 2.4: Transfection of Flp-In CHO cells to stably express Siglec-Fcs.....	53
Figure 2.5: Workflow for the expression and purification of Siglec-Fcs from stably transfect CHO Flp-In cells.....	54
Figure 2.6: Systemic profiling of cells treated with glycan modifying enzymes to more specifically describe Siglec ligands.....	56
Figure 2.7: Expression of an exogenous gene using lentiviral transduction	57
Figure 2.8: Depiction of gene knock-out by CRISPR/Cas9	58
Figure 2.9: Siglec-Fcs can be complexed with either Strep-Tactin or an anti-hlgG antibody	60
Figure 2.10: CD22-Fc staining of U937 cells treated with neuraminidase A.....	61
Figure 2.11: Siglec-Fc binding to U937 cells overexpressing CMAH	62
Figure 2.12: Binding of Siglec-Fcs to CHST1 and CHST2 expressing cells.....	63
Figure 2.13: Synergistic effect of use of knocking out glycan biosynthetic enzymes in cells which have been virally transduced towards enhanced Siglec binding towards describing Siglec ligands	64

Figure 2.14: Isoforms of CD33	64
Figure 2.15: Binding of CD33 and Siglec-8 Fc to U937 cells overexpressing CHST1 with various critical glycosylation enzymes knocked-out	66
Figure 2.16: Binding of Siglec-7 precomplexed with Strep-Tactin and anti-hIgG to primary human peripheral blood mononuclear cells	67
Figure 2.17: Gating Strategy to isolate primary immune cells from human spleens.....	68
Figure 2.18: Expression of Siglecs by human splenocytes.....	69
Figure 2.19: Siglec-Fc binding to human splenocytes	70
Figure 2.20: Siglec-Fc enzyme-linked immunosorbent assay used to probe Siglec ligands	71
Figure 2.21: Siglec-Fc bead assay used to measure the binding between Siglecs and liposome bearing Siglec ligands.....	73
Figure 2.22: Depiction of mass-spectrometry based assay for measuring the dissociation constant (K_d) between a Siglec and a ligand	74
Figure 2.23: Binding of different generations of hSiglec-15-Fc to CHST1 expressing U937 cells.....	77
Figure 2.24: Validation of CMAS KO CHO Flp-In cells	78
Figure 2.25: Comparing yield and function of Siglec-7-WT-Fc produced from wildtype and CMAS KO CHO cells.....	78
Figure 2.26: Summary of Siglec-Fc yields from wildtype and all C4 CMAS KO CHO cells	79
Figure 2.27: Binding of Siglec-Fc produced from wildtype and CMAS KO CHO cells to U937 cells overexpressing CHST1/2. U937 refers to CHST1/2 expressing U937 cells	80
Figure 2.28: Masking potential of Siglecs on primary human immune cells.	84
Figure 2.29: Masking of Siglec-Fcs by Siglec-Fcs.	88
Figure 3.1: Chemical structure of ganglioside GM1a	99
Figure 3.2: Biosynthetic pathway of gangliosides.	100
Figure 3.3: Comparison of human and murine Siglec families.	106
Figure 3.4: Schematic of cell-based assay used to probe Siglec-ganglioside interactions.....	108
Figure 3.5: Liposomes formulated with a high affinity Siglec-1 ligand appended to PEG45–DSPE engages wildtype hSiglec-1 expressing CHO cells.	109

Figure 3.6: Binding of GM1a liposomes formulated with varying amounts of PEG45-DSPE to hSiglec-1 expressing CHO cells.	110
Figure 3.7: Binding of GM1a liposomes in the presence of micellular GM1a to hSiglec-1 expressing CHO cells.....	111
Figure 3.8: Binding of liposomes formulated with increasing ganglioside content to CHO cells expressing hSiglec-1.....	112
Figure 3.9: Liquid Glycan Array (LIGA) profile of glycan density on phage against hSiglec-1 CHO cells.....	113
Figure 3.10: Binding of liposomes formulated with asialo-GM1 (GA1) and GM1a to CHO cells expressing hSiglec-1.....	114
Figure 3.11: Effect of ganglioside acyl chains on ganglioside liposome binding to CHO cell expressing hSiglec-1.....	115
Figure 3.12: Binding of GM1a liposomes formulated with various amounts of cholesterol and lipids with different acyl chains to CHO cells expressing hSiglec-1.....	116
Figure 3.13: Binding of liposomes formulated with gangliosides presented from an optimized liposome formulation compared to high affinity Siglec-1-ligand presented from <i>Doxil</i> liposome formulation to hSiglec-1 expressing CHO cells.....	117
Figure 3.14: Siglec expression validation, of CHO cell lines engineered to express each human Siglec and each corresponding arginine mutant.....	119
Figure 3.15: Ganglioside liposome binding to CHO cells expressing Siglec-4 in the absence of serum and after treatment with neuraminidase.....	120
Figure 3.16: Heatmap summarizing the binding interactions of each human Siglec.....	121
Figure 3.17: Ganglioside content titration to Siglec-6 and Siglec-7 expressing CHO cells.....	121
Figure 3.18: Heatmap summarizing the binding interactions of each murine Siglec and ganglioside bearing liposomes.....	123
Figure 3.19: Mouse Siglec-ganglioside binding outside a lipid bilayer using an ELISA.....	124
Figure 3.20: Optimization of liposome cholesterol content and bulk lipid structure for screening murine Siglec-1-ganglioside interactions.	125

Figure 3.21: Ganglioside content titration (1-10 mol%) against murine Siglecs that were found to bind gangliosides	126
Figure 3.22: Quantifying the interactions between m/hSiglec-1 and the oligosaccharide of GM1a and GM1b by a native mass spectrometry binding assay	128
Figure 3.23: Comparative binding of internal and externally linked sialic acids to m/hSig-1 in the bead assay and ELISA.....	129
Figure 3.24: Venn diagram comparison of murine and human Siglec ganglioside binding.	133
Figure 4.1: Model of Siglec-6 broken down into domains with important amino acids highlighted	145
Figure 4.2: Assessing the ability of Siglec-6 to bind gangliosides outside of a bilayer.	149
Figure 4.3: Siglec-6-Fc binding to cell lines in a cell-based glycan array..	150
Figure 4.4: Development of Liposome Over Lectin Assay (LOLA).	151
Figure 4.5: Investigating the ability of Siglec-6 to bind gangliosides in other liposome based assays... .	152
Figure 4.6: Schematic of a glycolipid broken down into three components.	153
Figure 4.7: Binding of neoglycolipids present from an amide dialkyl-glycerol scaffold	153
Figure 4.8: Binding of neoglycolipids presenting the GM3 oligosaccharide from triazole linked lipid scaffolds.....	154
Figure 4.9: Binding of neoglycolipids featuring a triazole linked di-O-hexadecyl glycerol.....	155
Figure 4.10: Optimizing the mol% of 5 in neoglycolipid bearing liposomes for engaging Siglec-6	156
Figure 4.11: Binding competition between GM1a bearing liposomes and liposomes bearing neoglycolipid 5 against WT Siglec-6 in the bead assay	157
Figure 4.12: Comparison of the binding of Siglec-6 to GM3 and nGL 5 outside of a lipid bilayer.....	158
Figure 4.13: Binding of Siglec-6 to liposomes bearing asialo-glycolipids	159
Figure 4.14: Expression and glycolipid binding to Siglec-6/8 full length chimeras in the cell assay	161
Figure 4.15: Binding of Siglec-6/8 chimeras to liposomes bearing neoglycolipid 5 in the LOLA	162
Figure 4.16: Amino acid sequence alignment of residues surrounding the critical arginine residue in the human Siglec family.....	164
Figure 4.17: Binding of liposomes formulated with nGL 5 to W127A Siglec-6 in the cell assay.....	164

Figure 4.18: Summary of liposome binding to mutants of Siglec-6.....	165
Figure 4.19: Blocking of neoglycolipid bearing liposome binding to Siglec-6 expressing CHO cells using anti-Siglec-6 antibody.	166
Figure 4.20: Gating scheme for defining Memory B-cells (CD19 ⁺ , CD22 ⁺ , CD27 ⁺ , IgD ⁻ , CD38 ⁻) and naïve B-cells (CD19 ⁺ , CD22 ⁺ , CD27 ⁻ , IgD ⁺) from human blood.....	167
Figure 4.21: Neoglycolipid liposome binding to human naïve and memory B cells.	167
Figure 4.22: Expression of Siglec-6 by primary mast cells isolated from human spleens.	168
Figure 4.23: Binding of liposomes bearing neoglycolipid 5 to LAD2 cells.....	169
Figure 4.24: Binding of liposome formulated 5 mol% 5 to human syncytiotrophoblasts.....	170
Figure 4.25: pHrodo is a pH sensitive fluorescent indicator that can be used to measure liposome internalization.....	171
Figure 4.26: Internalization of neoglycolipid liposomes by Siglec-6 expressing Daudi cells.....	172
Figure 4.27: Workflow for isolating extracellular vesicles from human blood.	173
Figure 4.28: Characterization of extracellular vesicles isolated from human blood.....	174
Figure 4.29: Binding of EVs to wildtype and mutant Siglec-6 expressing CHO cells.....	174
Figure 4.30: Binding competition between EVs and liposomes bearing neoglycolipid 5 against WT Siglec-6 in the bead assay.....	175
Figure 4.31: Binding of EVs treated with neuraminidase A and S to Siglec-6 in the bead assay.....	176
Figure 4.32: Binding of EVs isolated from wildtype and β 1-4GalNT1 ^{-/-} N2a cells to WT Siglec-6 in the bead assay.....	176
Figure 4.33: Binding of EVs to LAD2 cell pretreated with Siglec-6 blocking antibody.....	177
Figure 4.34: Internalization of EVs by Daudi cells virally transduce with Siglec-6.....	178
Figure 4.35: Model of Siglec-6 Binding glycolipids presented from a bilayer.....	183
Figure 5.1: Siglec-ganglioside interactions are likely most relevant in Siglec extracellular vesicle interactions.....	198
Figure 5.2: Summary of thesis.....	199

Figure A2.1: Siglec expression by B cells (CD19 ⁺) isolated from human spleen.....	227
Figure A2.2: Siglec expression by CD8 ⁺ T cells (CD3 ⁺ , CD4 ⁺) isolated from human spleen.....	227
Figure A2.3: Siglec expression by CD8 ⁺ T cells (CD3 ⁺ , CD8 ⁺) isolated from human spleen.....	228
Figure A2.4: Siglec expression by monocytes (CD14 ⁺) isolated from human spleen.....	228
Figure A2.5: Siglec expression by neutrophils (CD15 ⁺ , CD16 ⁺) isolated from human spleen	229
Figure A2.6: Siglec expression by mature NK cells (CD56 ⁺) isolated from human spleen.....	229
Figure A2.7: Siglec-Fc binding to B cells (CD19 ⁺) isolated from human spleen	230
Figure A2.8: Siglec-Fc binding to CD8 ⁺ T cells (CD3 ⁺ , CD4 ⁺) isolated from human spleen	231
Figure A2.9: Siglec-Fc binding to CD8 ⁺ T cells (CD3 ⁺ , CD8 ⁺) isolated from human spleen	232
Figure A2.10: Siglec-Fc binding to monocytes (CD14 ⁺) isolated from human spleen.....	233
Figure A2.11: Siglec-Fc binding to neutrophils (CD15 ⁺ , CD16 ⁺) isolated from human spleen..	234
Figure A2.12: Siglec-Fc binding to mature NK cells (CD56 ⁺) isolated from human spleen.	235
Figure A3.1: Interrogation of the human Siglec family against nine commercially available gangliosides using our optimized liposome formulation.....	236
Figure A3.2: Binding of ganglioside bearing liposomes to Siglec expressing CHO cells after treatment with neuraminidase A and S.....	237
Figure A3.3: Binding of ganglioside liposomes to the murine Siglec family in the bead assay	238
Figure A3.4: Binding of the murine Siglecs to gangliosides in an ELISA.....	239
Figure A3.5: GM1a, GM2 and GD1a ganglioside content titration against human Siglec-1 in the bead assay	240
Figure A3.6: GM1a, GM2 and GD1a ganglioside content titration against murine Siglec-1 in the bead assay	240
Figure A3.7: GM1a, GM2 and GD1a ganglioside content titration against murine Siglec-15 in the bead assay.....	241
Figure A3.8: GM1a, GM2 and GD1a ganglioside content titration against mSiglec-F in the bead assay	241

Figure A3.9: GM1a, GM2 and GD1a ganglioside content titration against mSiglec-F in the bead assay.
..... 242

Figure A4.1: Complex gangliosides in EVs are ligands for Siglec-6. a, Abbreviated ganglioside biosynthesis highlighting the role of β 1-4GalNT1 in ganglioside biosynthesis.....245

Table of Tables

Table 1.1: Current state of Siglec ligands.	40
Table 2.1: Dissociation constants derived between Siglec and glycans using Siglec fragment approach.....	75
Table 2.2: Publications that have used the Macauley Lab Siglec-Fc Chimeras.....	81
Table 3.1: Effects of defects in ganglioside biosynthesis on health.....	102
Table 3.2: Affinities of ganglioside oligosaccharides towards human Siglec-1	113
Table 3.3: Sources of gangliosides used to interrogate the human Siglec family	118
Table 5.1: Publications that have used the Macauley Lab Siglec-Fc chimera and or optimized ganglioside liposome.....	192
Table A2. 1: Markers used to define of immune cells from human spleen samples.....	225
Table A2.2: Anti-Siglec antibodies used in this study.	226
Table A4.1: Key amino acids in Siglec-6.....	243
Table A4. 2 Siglec-6 Mutagenesis Primers	244

List of Abbreviations

AD	alzheimer's disease
AML	acute myeloid leukemia
βGalNT1	β -1,4- <i>N</i> -acetyl-galactosaminyltransferase 1
C1GalT	core 1 galactose transferase
CD	cluster of differentiation
CHO	chinese hamster ovary
CHST	carbohydrate sulfotransferase
CMAH	cytidine monophospho- <i>N</i> -acetylneuraminic acid hydroxylase
CMAS	<i>N</i> -acetylneuraminic acid cytidyltransferase
CMP	cytidine monophosphate
COIN	concentration-independent native mass spectrometry
COSMC	core 1 β 3-Gal β -T-specific molecular chaperone
DAPC	1,2-diarachidoyl- <i>sn</i> -glycero-3-phosphocholine
DLPC	1,2-dilauroyl- <i>sn</i> -glycero-3-phosphocholine
DMPC	1,2-dimyristoyl- <i>sn</i> -glycero-3-phosphocholine
DNA	deoxyribonucleic acid
DOPC	1,2-dioleoyl- <i>sn</i> -glycero-3-phosphocholine
DPPC	1,2-dipalmitoyl- <i>sn</i> -glycero-3-phosphocholine
DSPC	1,2-distearoyl- <i>sn</i> -glycero-3-phosphocholine
DSPE	1,2-distearoyl- <i>sn</i> -glycero-3-phosphorylethanolamine
EGFR	epidermal growth factor repeats
ELISA	enzyme-linked immunosorbent assay
EndoH	endoglycosidase H
ER	endoplasmic reticulum

EV	extracellular vesicles
Fc	fraction crystallizable (constant region)
Fuc	L-fucose
Gal	Galactose
GalNAc	<i>N</i> -acetylgalactosamine
GDP	guanosine diphosphate
Glc	glucose
GlcA	glucuronic acid
GlcNAc	<i>N</i> -acetylglucosamine
GMA	glycan microarrays
GNE	UDP- <i>N</i> -acetylglucosamine 2-epimerase/ <i>N</i> -acetylmannosamine kinase
GPI	glycosylphosphatidylinositol
HIV	human immunodeficiency virus
hSiglec	human Siglec
IC₅₀	half maximal inhibitory concentration
IgE	immunoglobulin E
IgG	immunoglobulin G
ITIM	immunoreceptor tyrosine-based inhibitory motif
K_d	dissociation constant
Kdn	ketodeoxynonulonic acid
KO	Knock-out
LacNAc	β -Gal ρ -(1 \rightarrow 4)- β -Glc ρ NAc-(1 \rightarrow 3)
LiGA	liquid glycan array
LOLA	liposome over lectin assay
MAG	myelin associated glycoprotein

Man	mannose
MGAT1	α -1,3-mannosyl-glycoprotein 2- β - <i>N</i> -acetylglucosaminyltransferase
mSiglec	murine Siglec
MYA	million years ago
Neu5Ac	<i>N</i> -acetylneuraminic acid
Neu5Gc	<i>N</i> -glycolylneuraminic
NeuA	neuraminidase A (<i>Arthrobacter ureafaciens</i>)
NeuS	neuraminidase S (<i>Streptococcus pneumoniae</i>)
nGL	neoglycolipid
OGT	O-linked <i>N</i> -acetylglucosaminyltransferase
PAPS	3'-Phosphoadenosine-5'-phosphosulfate
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PNGase F	peptide: <i>N</i> -glycosidase F
POFUT1	protein O-fucosyltransferase 1
POFUT2	protein O-fucosyltransferase 2
POMT1	protein O-mannosyl-transferase 1
POMT2	protein O-mannosyl-transferase 2
POPC	1-palmitoyl-2-oleoyl- <i>sn</i> -glycero-3-phosphocholine
PSPC	1-palmitoyl-2-stearoyl- <i>sn</i> -glycero-3-phosphocholine
RNA	ribonucleic acid
SAMP	self-associated molecular pattern
SARS-CoV-2	severe acute respiratory syndrome corona Virus 2
Siglec	sialic-acid-binding immunoglobulin-type lectins

SNP	single nucleotide polymorphism
SNFG	symbolic nomenclature for glycans
SOPC	1-stearoyl-2-oleoyl-sn-glycero-3-phosphocholine
TLC	thin layer chromatography
TSP-1	thrombospondin type-1 repeat
U.V./Vis	ultraviolet/visible light
UCGC	UDP-glucose ceramide glucosyltransferase
UCGC	UDP-glucose ceramide glucosyltransferase
UDP	uridine diphosphate
UT	untransduced
WT	wildtype
XT1	xylose transferase 1
XT2	xylose transferase 2
Xyl	D-xylose

Chapter 1: Glycobiology and the Siglecs

1.1: Introduction to Glycobiology

1.1.1: Glycans as biological polymers

The surface of all cells across all domains of life are coated in a dense array of carbohydrate-decorated macromolecules known as glycans¹. These cell surface glycans are collectively referred to as the glycocalyx of a cell. Often, the first point of contact between two cells occurs through glycans and glycan binding proteins called lectins². Glycans are a critical part of nearly all physiological (protein folding³, blood typing⁴, reproduction⁵, immunomodulation⁶, *etc.*) and pathophysiological functions (cancer⁷, neurodegenerative disorders⁸, viral⁹ and bacterial infections¹⁰, autoimmune disorders¹¹, *etc.*). In fact, many breakthroughs in biology, such as HIV vaccines¹² and organ transplant/xenotransplants¹³ have come through understanding the interactions between glycans and lectins. However, glycans are difficult to study because they are branched polymers made from subunits many of which are isomers, and their sequence cannot be predicted according to the central dogma of biology, because the biosynthesis of glycans is not template-encoded. Thus, new tools and approaches are required to develop a better understanding of the roles of glycans in health and disease.

1.1.2: Glycans as a family of biological macromolecules

Glycans are polymers made from subunits known as monosaccharides. The structure of glycans can be relatively simple (*e.g.* single monosaccharide unit) to very large/complex (*e.g.* hundreds of monosaccharides with multiple branches)¹⁴. Compared to other biomolecules such as DNA, RNA, and protein, relatively less is known about the roles of glycans in biology (**Figure 1.1**). This discrepancy in knowledge likely stems from the inherent challenges in studying glycans. Glycans are particularly challenging to study because glycan biosynthesis is governed by the expression of glycosyltransferases rather than a direct template like RNA and proteins^{15, 16}. Studying glycans is further complicated by the fact that glycans can be branched polymers, whereas nucleic acids and proteins are linear polymers. Moreover, monosaccharides in glycans

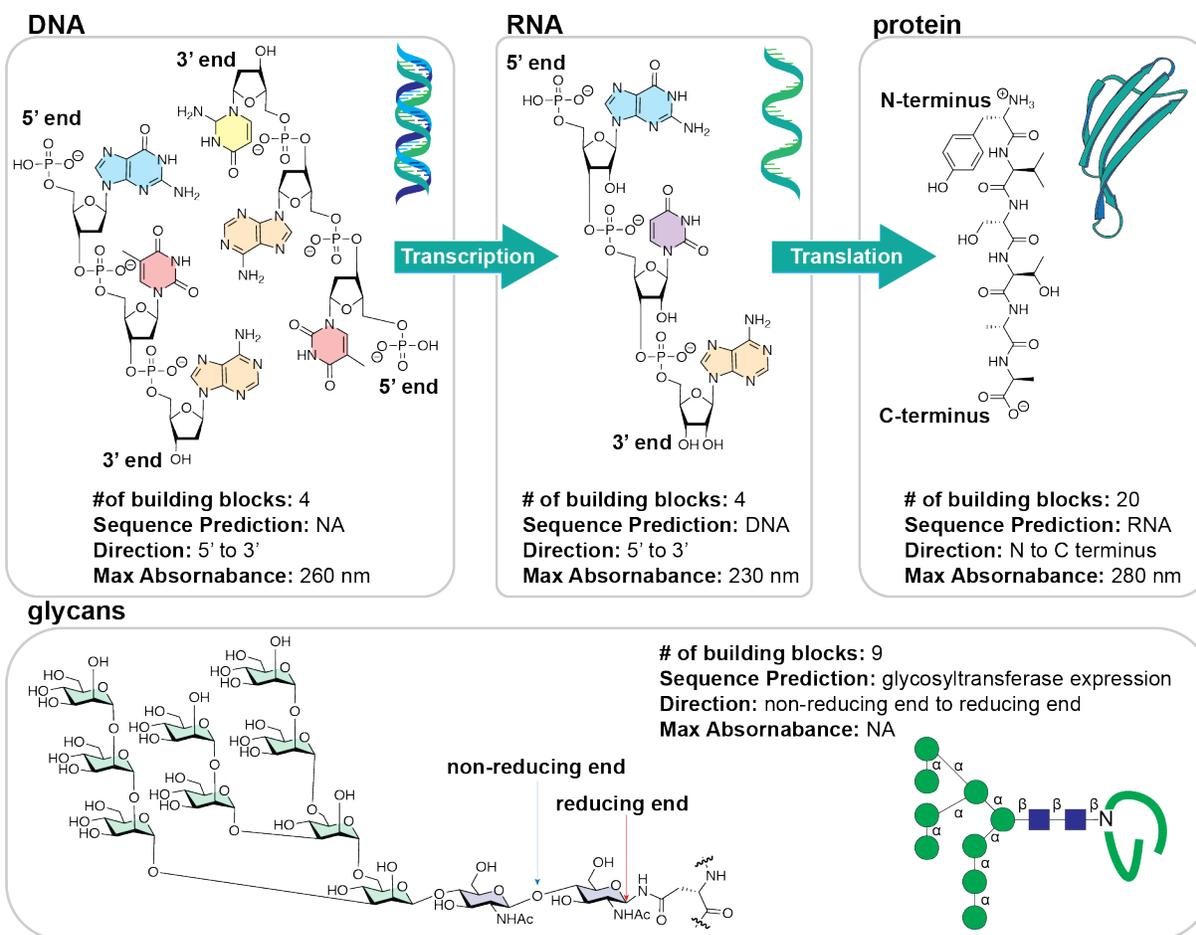


Figure 1.1: Comparison of biological macromolecule families. DNA, RNA, and proteins all fall along the central dogma of biology and can be predicted at each level based on the primary sequence of the macromolecule the proceeds it in the dogma (e.g. DNA can predict RNA can predict protein). Unlike DNA, RNA, and proteins, glycans can be branched structures and can have different stereochemical linkages between monosaccharides, and their structure cannot be predicted according to the central dogma of biology.

can be linked in different regio- and stereochemistries, unlike the other biological molecules that are generally linked in the same way between each subunit. Additionally, many monosaccharides and glycans are stereoisomers, making them difficult to study by mass spectrometry. Glycans also do not possess an innate chromophore, making it challenging to study glycans by ultraviolet/visible spectrophotometry-based techniques¹⁷. Moreover, when synthesizing glycans, the abundance of alcohol groups can make targeting a specific alcohol arduous, and each monosaccharide can adopt two anomers in solution, further complicating synthetic approaches (**Figure 1.2a**)¹⁸. Monosaccharides convert between these anomers through a process referred to as mutarotation (**Figure 1.2b, c**). Mutarotation is possible under physiological conditions due to

the presence of the oxygen in the cyclohexane/cyclopentane ring, which lowers the pK_a of the anomeric hydroxyl by five orders of magnitude (17 to 12)¹⁹. A pyranose (six-membered ring) or furanose (five-membered ring) forms when the hydroxyl on C5 or C4 attacks the aldehyde at C1 respectively. If the monosaccharide is in its more common pyranose form, it is denoted with a *p* (e.g. Glcp), and when it is in its furanose form is denoted with an *f* (e.g. Glcf)²⁰. Due to these reasons, monosaccharides are complex molecules to study.

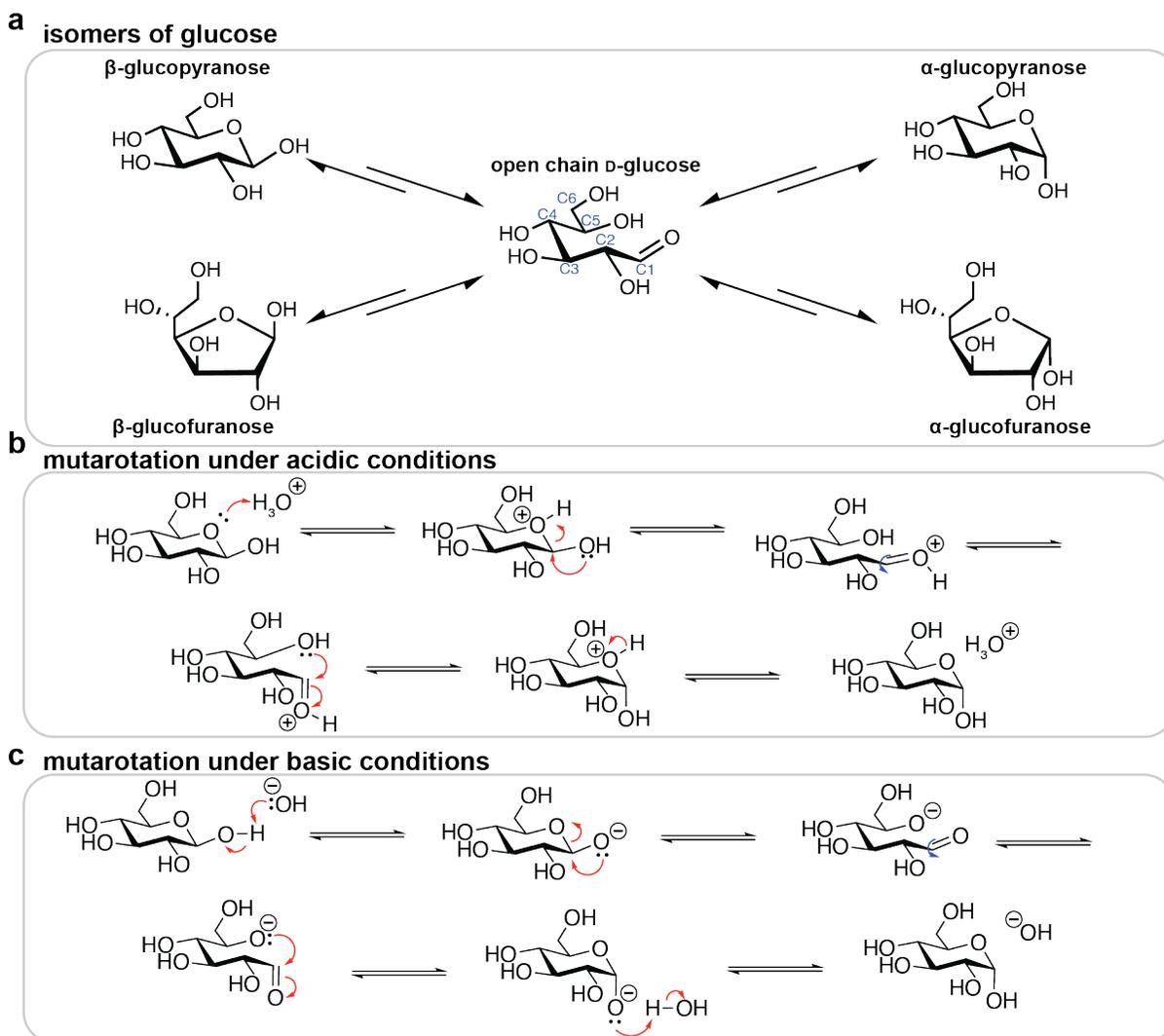


Figure 1.2: Proposed mechanism of interconversion of isomers of glucose. a, Different isomers of glucose with respect to pyranose/furanose and α/β - anomers. b, c, Interconversion of α/β -glucose under acidic and basic conditions respectively.

1.2: Types of Glycans

1.2.1: Building blocks and glycosyltransferases

In mammals, there are nine monosaccharide building blocks that contribute to the diversity and complexity of glycans (**Figure 1.3**)²¹. Monosaccharides can also be represented using a Fischer projection when in their linear form which is helpful when assigning D or L configuration of the monosaccharide. Nucleic acids and proteins are typically described as 5'-OH to 3'-OH and N-terminus to C-terminus from left to right respectively. Likewise, monosaccharides are described from non-reducing end to reducing end from left to right. The reducing end refers to the end of the sugar that features a ketone (*e.g.* sialic acid) or aldehyde (*e.g.* glucose) in its open chain form. As drawing the complete chemical structures of glycans can be rather cumbersome, a unified Symbolic Nomenclature for Glycans (SNFG) has been developed where each monosaccharide is represented by a different shape and colour²².

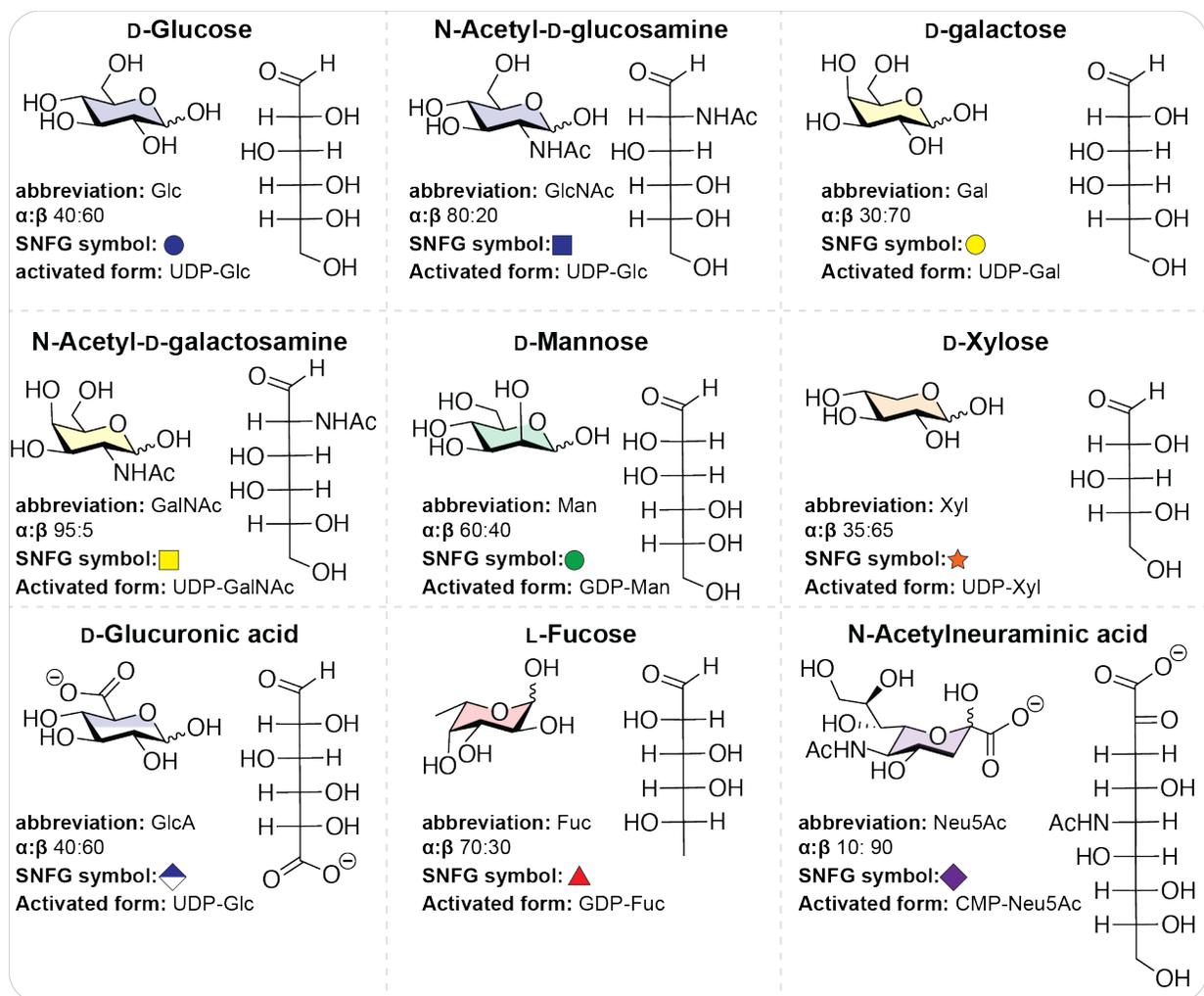


Figure 1.3: Chemical structures of the nine monosaccharide building blocks in mammals. Monosaccharides are represented in their pyranose form, Fischer projection and symbol according to the Symbolic Nature for Glycans (SNFG).

Monosaccharides are activated by the addition of a nucleotide diphosphate making a nucleophilic substitution reaction at the anomeric center energetically feasible. This nucleotide activated sugar is referred to as the *donor* (Figure 1.4). The donor acts as an electrophile and is attacked by a nucleophile, known as the *acceptor*, from another monosaccharide, a protein, or a lipid²³. Glycans are synthesized in a stepwise fashion by enzymes collectively referred to as glycosyltransferases that catalyze the assembly of sugar *donors* and *acceptors*, to thereby build the glycan. In addition to making the monosaccharide more electrophilic and serving as a better leaving group compared to water or hydroxide, the nucleotide diphosphate also contributes to the recognition of the sugar donor by the glycosyltransferase²⁴.

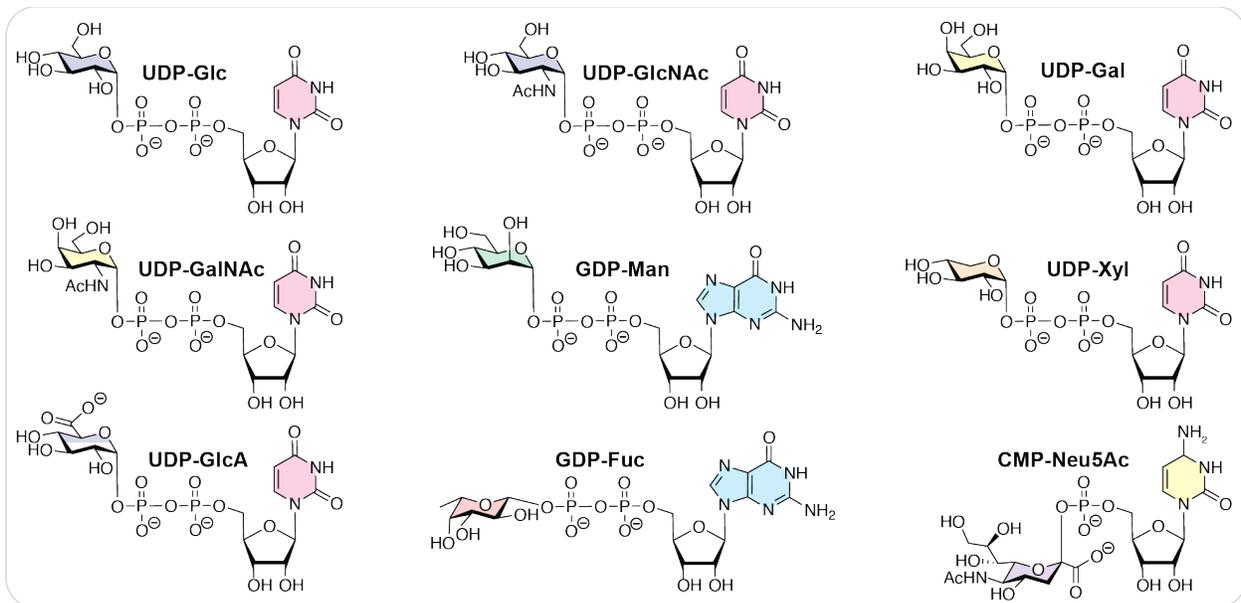


Figure 1.4: Chemical structures of the nine monosaccharide donors. Glycosyltransferases use activated forms of monosaccharides referred to as sugar *donors* to build glycans.

Glycosyltransferases are divided into many families and subtypes, which is determined by the substrates the enzyme acts on and linkage it forms¹⁶. These enzymes can be broken down into two general classes: retaining or inverting²⁵. This is based on the differences in the anomeric linkage in the product compared to the sugar *donor*²⁶. Inverting glycosyltransferase generally use an S_N2 -like mechanism (**Figure 1.5a**) and catalysis can be aided by a divalent cation such as magnesium²⁷, whereas a retaining glycosyltransferase generally use a substrate assisted or S_Ni mechanism (**Figure 1.5b**)^{25, 28}. Glycosyltransferases are typically expressed in the Golgi apparatus²⁹, but can also be expressed in the endoplasmic reticulum (ER)¹⁶ depending on the type of glycan/protein/lipid they act on. While building glycans is important for maintaining cellular health, it is just as important to degrade glycans. Enzymes that break down glycans are known as glycosyl hydrolases or glycosidases and they generally operate through a double S_N2 like mechanism (**Figure 1.5c**). Many of the cellular glycosidases reside in the lysosome to participate in recycling of monosaccharide building blocks.

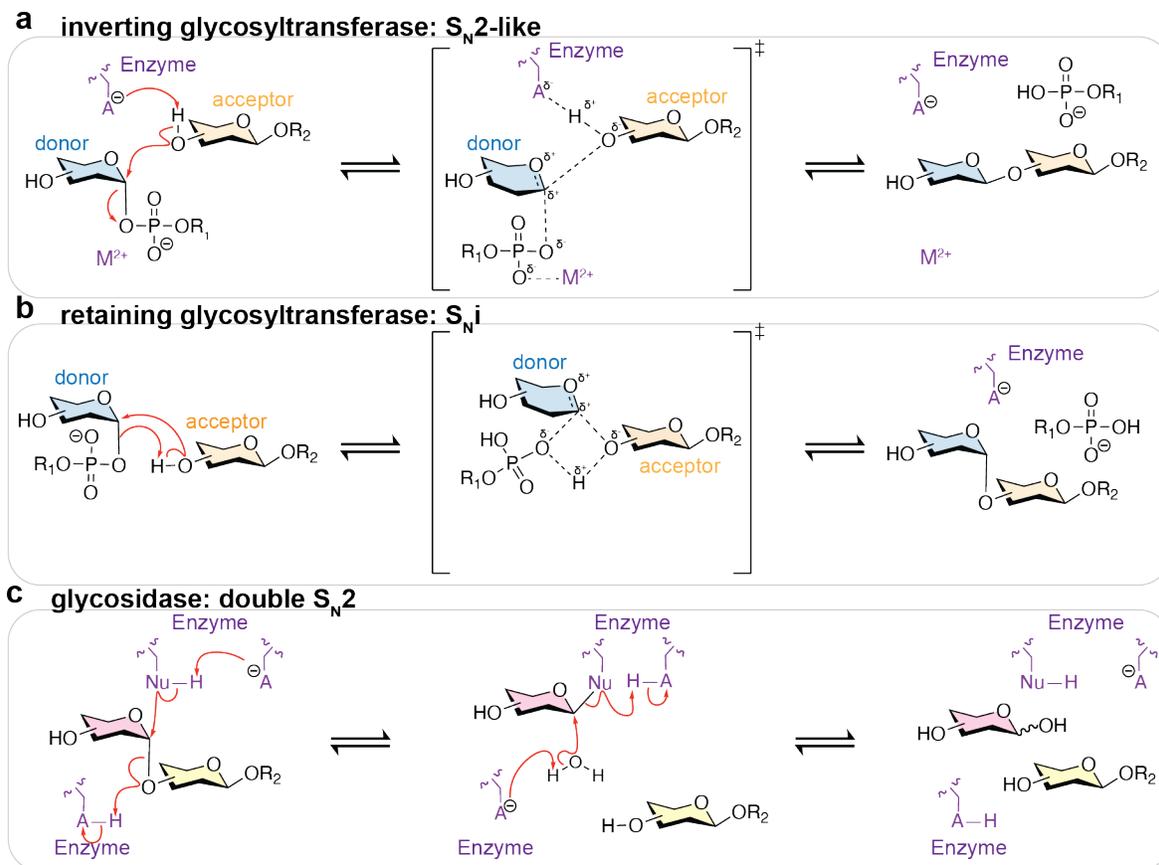


Figure 1.5: Proposed general mechanisms of glycosyltransferases and glycosidases. **a, b**, Proposed general mechanism of inverting S_N2-like and retaining S_Ni-like glycosyltransferase respectively. **c**, Proposed general mechanism of a glycosidase operating through a double S_N2-like mechanism. R₁ represents of the remainder of the nucleotide portion of the sugar-donor. R₂ represents the remainder of glycan, protein, or lipid.

In mammals, the three major types of glycans are *N*-glycans (**Figure 1.6a**), *O*-glycans (**Figure 1.6b**) and glycolipids (**Figure 1.6c**); however, these classes are further divided into many subtypes³⁰. Glycans are typically depicted using SNFG with the reducing end on the right-hand side. Oxford Notation can be used to further describe glycans represented using SNFG by using angles between the shapes to denote the type of linkage (e.g. 90° counter-clockwise from the reducing end represents a 2-8 linkage).

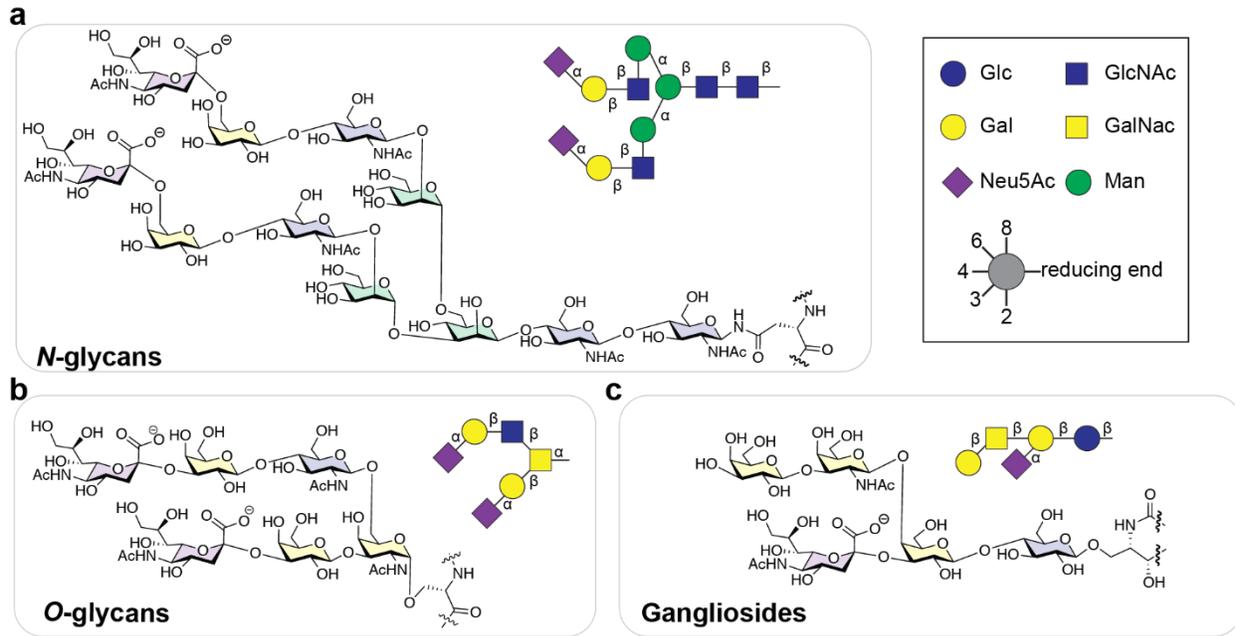


Figure 1.6: Major types of glycans. **a**, Representative structure of a complex *N*-glycan. **b**, Representative structure of a sialylated core-2 *O*-glycan. **c**, Representative structure of glycolipid ganglioside GM1a. Structures are represented as the chemical structure as well as SNFG with Oxford notation.

1.2.2: *N*-glycans

N-glycan biosynthesis begins with the synthesis of an *N*-glycan precursor known as Glc₃-Man₉-GlcNAc₂, which is appended to a dolicholdiphosphate. Glc₃-Man₉-GlcNAc₂ is synthesized one monosaccharide at a time by adding monosaccharides to dolicholdiphosphate, which is anchored to the membrane of the ER on the cytoplasmic face³⁰. Once the first seven monosaccharides are added, it is then flipped into the lumen of the ER where it is further elongated into the 14-mer oligosaccharide³¹. The glycan is then transferred to a protein while it is being translated by the ribosome (**Figure 1.7a**). As *N*-glycans are transferred to their protein substrate as the protein is being translated, *N*-glycans often increase the stability of the protein and in many cases, proteins cannot fold without the *N*-glycan³². Generally, *N*-glycosylation sites can be predicted by the sequon Asn-Xxx-Ser/Thr, where 'Xxx' is any amino acid other than proline³³. Unlike other glycans, *N*-glycans are linked through an amide bond (**Figure 1.7b**) rather than a glycosidic linkage; much like the peptide bond linking two amino acids³⁴. Once on the protein, the protein travels from the ER to the Golgi apparatus through the secretory pathway³¹.

While the protein is being shuttled through the Golgi apparatus, the *N*-glycan is further modified to yield three different types of *N*-glycans: high mannose, hybrid, or complex (**Figure 1.7c**)³⁵.

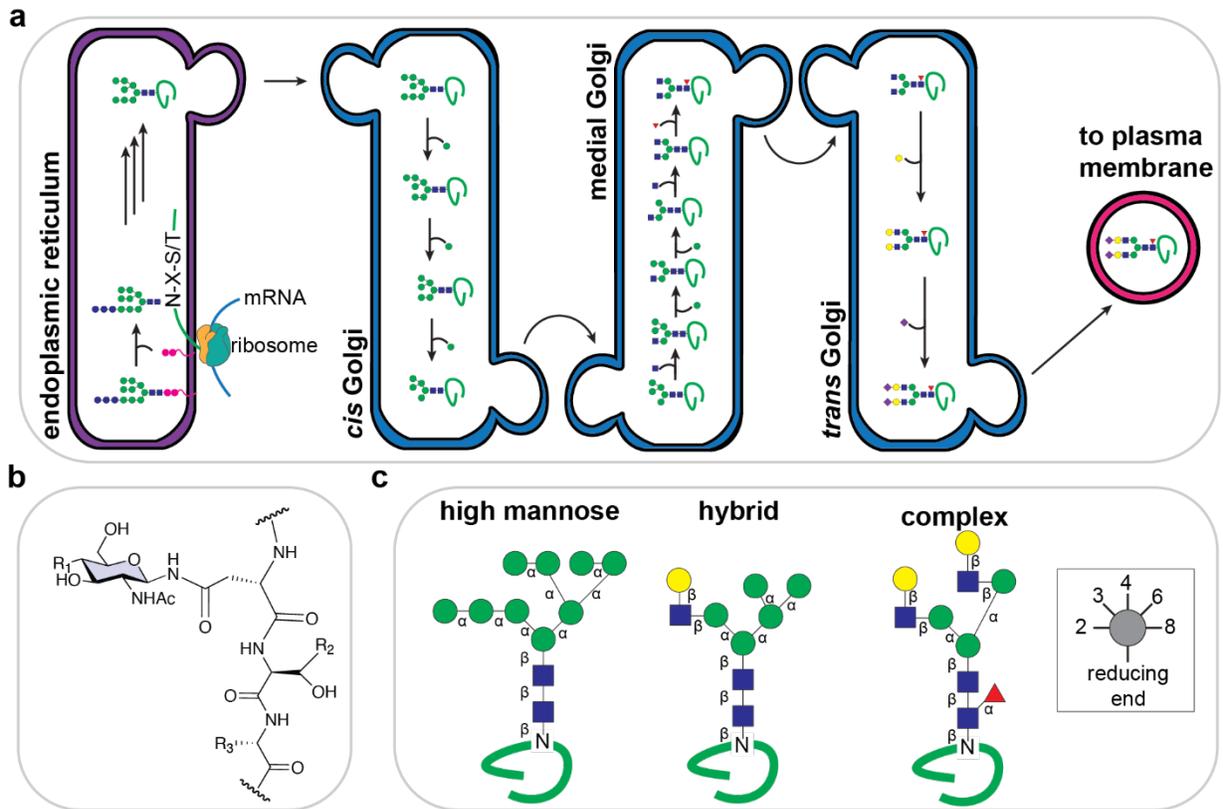


Figure 1.7: Biosynthesis and classes of *N*-glycans. **a**, Simplified schematic of *N*-glycan biosynthesis in mammals, adapted from *The Essentials of Glycobiology 4th Edition*. **b**, Chemical structure of an *N*-glycan attached to a protein. R₁ is the remainder of the glycan, R₂ is either H or CH₃, of a serine or threonine respectively, and R₃ is any R group of an amino acid other than proline. **c**, Different classes of *N*-glycans. For panels **a**, and **c**, the glycans are represented using SNFG representation.

N-glycans are a relatively well-studied type of glycan and many tools have been developed to study them. Common strategies used to modulate *N*-glycans are knocking-out enzymes critical for the biosynthesis of *N*-glycans³⁶, treatment of cells or proteins with enzymes that specifically act on *N*-glycans such as Endoglycosidase H (EndoH)³⁷, and Peptide: *N*-glycosidase F (PNGase F)³⁸, or chemical inhibitors of *N*-glycan biosynthesis such as kifunensine³⁹. The *mgat1* gene is often targeted in knock-out experiments, as it is required to mature high mannose *N*-glycans into complex *N*-glycans. Similarly, treatment of cells with kifunensine also results in immature high mannose *N*-glycans by inhibiting ER mannosidases that are involved in converting the *N*-glycan from a high mannose to complex-type. EndoH is a glycosidase which cleaves the glycosidic

linkage between the two GlcNAc residues on high mannose and hybrid *N*-glycans liberating the glycan from the protein while leaving behind the initial GalpNAc residue on the protein. While often referred to as a glycosidase, it is more appropriate to refer to PNGase F as an amidase as it hydrolyzes the amide linkage between the GlcpNAc on the reducing end of the glycan and the asparagine residue. Treatment of proteins with PNGase F leaves a characteristic *scar* as the asparagine residue is chemically converted to an aspartic acid residue.

1.2.3: O-glycans

O-glycans are another type of glycan frequently found on proteins; however, they differ from *N*-glycans in several ways⁴⁰. For instance, O-glycans are linked through a serine or threonine residue in a glycosidic linkage and there is no reliable sequon to predict O-glycosylation sites⁴⁰. Moreover, there is no precursor like there is for *N*-glycans and the glycans are built up one monosaccharide at a time. There are many types of O-glycans including: O-GalNAc, O-Mannose, O-Fucose, O-Xylose, and O-GlcNAc^{41, 42}. All of these factors make O-glycans very structurally diverse.

1.2.3.1: O-GalNAc glycans

Mucin-type O-GalNAc glycans are the most common forms of O-glycans and the initial GalNAc is added to a protein by a family of twenty polypeptidyl GalNAc transferases (ppGalNAcTs)⁴³. When identifying the roles of O-GalNAc glycans, knocking-out *cosmc*, a chaperon required for functional core 1 galactose transferase (C1GalT) the enzyme responsible for elongating the *Tn* antigen $\alpha(\text{GalpNAc}\rightarrow\text{Ser/Thr})$ to the *T* antigen $\beta\text{-Gal-(1}\rightarrow\text{3)-}\alpha\text{-GalpNAc-(Ser/Thr)}$ also known as core 1 O-GalNAc. This results in truncated O-GalNAc glycans⁴³. As an alternative to genetically interrupting the biosynthesis of O-GalNAc glycans, α -benzyl-GalNAc can be used as a pharmacological inhibitor of O-GalNAc glycans, which acts via a decoy acceptor⁴⁴. Unfortunately, there are no enzymatic tools equivalent to PNGaseF or EndoH to study O-glycans. Instead of removing O-glycans enzymatically, O-glycans can be liberated from their protein through β -elimination. However, these conditions are not ideal, as under basic conditions, open

chain monosaccharides can undergo further eliminations/rearrangements as well as truncated glycans due to the presence of an acidic proton (**Figure 1.8**)⁴⁵.

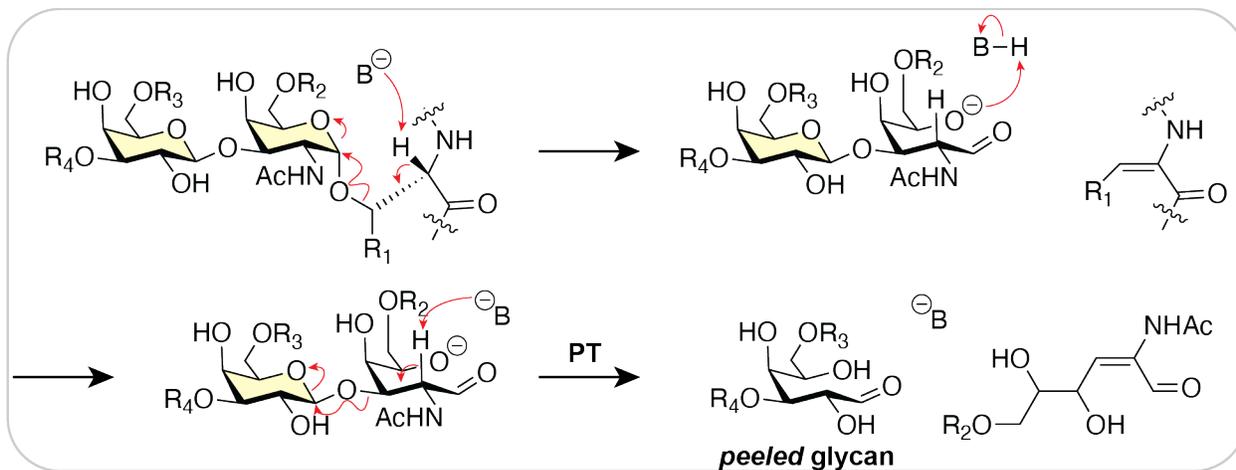


Figure 1.8: β-elimination mechanism of O-glycans. R₁ represents a hydrogen (serine) or methyl group (threonine). R₂, R₃, and R₄ represent possible extension points of the glycan or hydrogen if the glycan is not extended. PT, proton transfer.

Following the addition of the initial GalpNAc, many different *O*-glycan cores can be assembled, with cores 1-4 being the most common (**Figure 1.9**)⁴⁶. Alternatively, the glycan may not be modified (yielding the *Tn* antigen) or the GalpNAc can be sialylated (forming the *sialyl Tn* antigen). The *Tn* and *sialyl Tn* antigens are uncommon and are typically elongated in healthy humans. However, these truncated glycans can be a molecular marker for cancer^{47, 48}. While *O*-GalNAc glycans are found on many proteins, they are particularly abundant on a family of proteins known as mucins⁴⁹. In fact, a defining characteristic of a mucin is having a stretch of amino acids that is heavily glycosylated with *O*-GalNAc glycans (up to 80% by weight)⁴⁰. Mucins can either be soluble and released to the extracellular matrix⁵⁰ or they can be membrane-bound extending out from the surface of cells⁵¹. Cell surface mucins are characteristic of epithelial cells and contributes to the formation of a protective gel around the cells especially in the gut and lungs. These mucins are often considered the first line of defense between the outside world and tissues such as the lungs and gastro-intestinal track⁵².

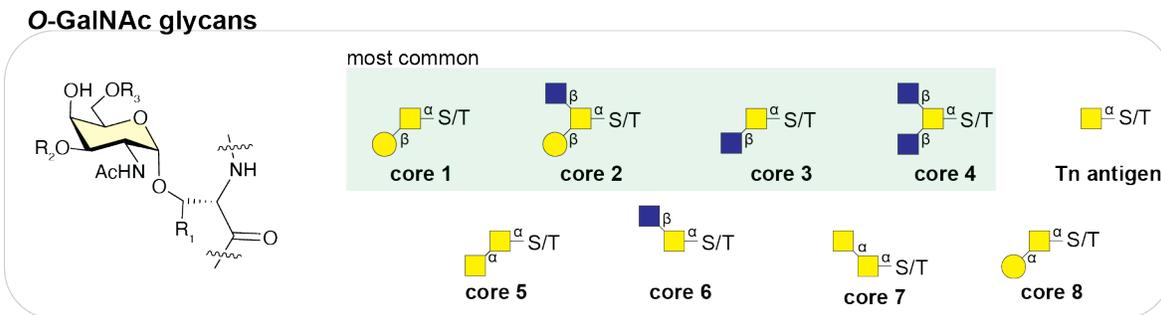


Figure 1.9: O-GalNAc O-glycans. Different types of O-GalNAc cores. Cores 1-4 (highlighted in green) are the most common. R₁ is either a H (serine) or CH₃ (threonine). R₂ and R₃ are possible glycan extension points.

1.2.3.2: O-Mannose glycans

The O-mannose modification (**Figure 1.10**) is less abundant (30% of O-glycans⁵³) and has fewer cores compared to O-GalNAc glycans, however it is still an important modification. Disruptions in O-mannosylation leads to neuronal disorders and muscular dystrophies⁵⁴. There are two enzymes Protein O-mannosyl-transferase 1 and 2 (POMT1 and POMT2), which are both required for transferring the initial mannose to the protein in the ER⁵⁴. There are three O-mannose core structures, which can be further modified as the protein travels through the Golgi apparatus⁵³.

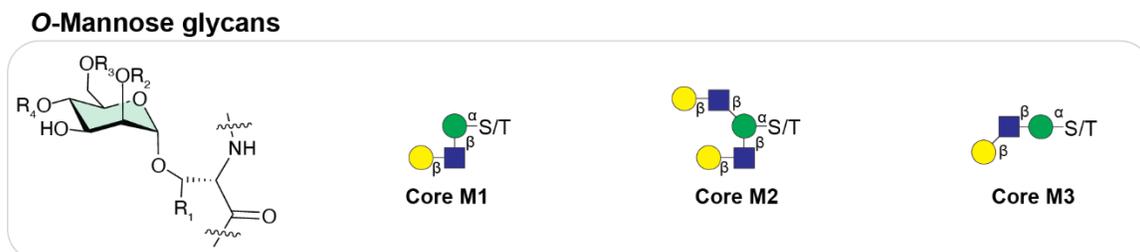


Figure 1.10: O-Mannose O-glycans. R₁ is either a H (serine) or CH₃ (threonine). R₂, R₃, and R₄ are possible glycan extension points.

1.2.3.3: O-Fucose glycans

A serine or threonine residue can also be modified with a fucose residue (**Figure 1.11**). Fucose is added to proteins by Protein O-fucosyltransferase 1 (POFUT1 or FUT12), which initiates the synthesis of *epidermal growth factor-like repeats* (EGFR) glycan motif and acts on the sequon of Cys-Xxx-Xxx-Xxx-Ser/Thr-Cys where Xxx is any amino acid⁵⁵. Protein O-fucosyltransferase 2 (POFUT2 or FUT13) also transfers a fucose to a protein through a serine or threonine residue however it proceeds the biosynthesis of *thrombosin type-1 repeats* glycan motif and recognizes the sequon Cys-Xxx-Xxx-Ser/Thr-Cys⁵⁵. Both POFUT1 and 2 act on their protein

substrates in the ER⁵⁵. Compared to other types of glycans, the biological functions of O-fucose are relatively unknown.

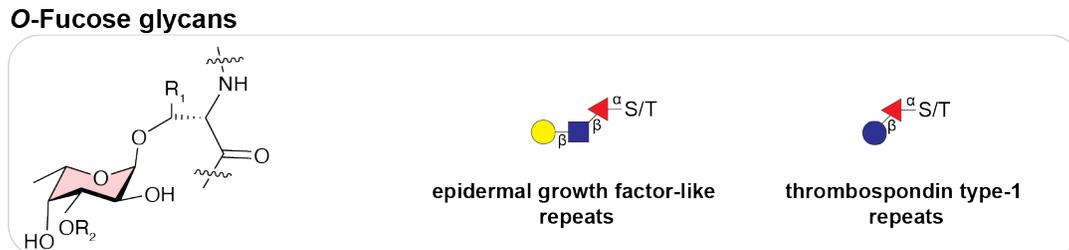


Figure 1.11: O-Fucose O-glycans. R₁ is either a H (serine) or CH₃ (threonine). R₂ is a possible glycan extension point.

1.2.3.4: O-Xylose glycans

O-Xylose is another type of O-glycosylation, which is the starting point for many proteoglycans which are proteins that are decorated with long linear glycan chains often with repeating subunits made of two monosaccharides (**Figure 1.12**)⁵⁶. The addition of xylose by xylose transferase 1 or 2 (XT1/2) to a protein begins the biosynthesis of heparan sulfate and chondroitin sulfate⁵⁶; both of these start from the same motif and are important components of the extracellular matrix⁴².

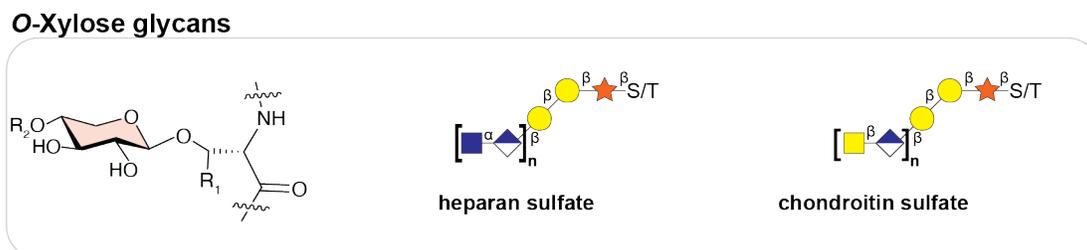


Figure 1.12: O-Xylose O-glycans. R₁ is either a H (serine) or CH₃ (threonine). R₂ is a possible glycan extension point.

1.2.3.5: O-GlcNAc glycans

O-GlcNAc is different than the other types of O-glycosylation because it is not elongated like the other types and only exists as a single monosaccharide linked to a protein (**Figure 1.13**)⁵⁷. Moreover, O-GlcNAc sites are found on intracellular proteins rather than cell surface proteins. Additionally, there is only one enzyme, O-GlcNAc transferase (OGT), that catalyzes the addition of GlcNAc to a protein⁵⁷. Another important distinction between O-GlcNAc modification and other glycans is that it is a reversible modification. The reversible nature of this modification allows it to

participate in cell signaling pathways and, in many cases, it competes with phosphorylation sites on proteins⁵⁸.

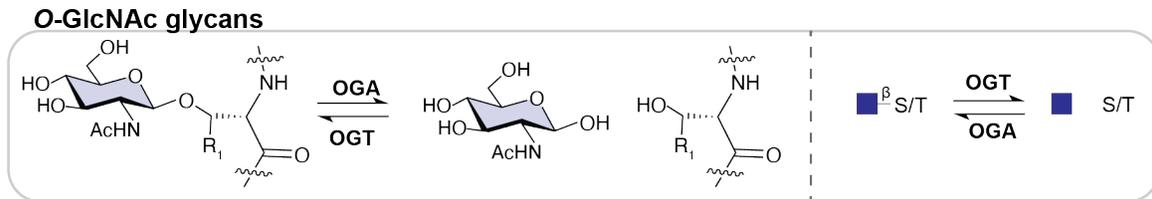


Figure 1.13: Structure of O-GlcNAc modification. R₁ is either a H (serine) or CH₃ (threonine). OGA, O-GlcNAcase; OGT, O-GlcNAc transferase.

1.2.4: Glycolipids

The last major class of glycans are the glycolipids, more specifically glycosphingolipids, that can be further divided into *series* according to the sequence of monosaccharides extending from the ceramide⁵⁹. Regardless of the series, nearly all glycolipids are derived from glucosylceramide, which is synthesized on the cytoplasmic face of the ER starting from serine and a coenzyme-A modified fatty acid (**Figure 1.14a**)⁶⁰. As glucosylceramide is the precursor to nearly all glycolipids, UDP-glucose ceramide glucosyltransferase (UGCG) is often a genetic or pharmacological target for the disruption of glycolipid biosynthesis⁶¹. Ceramide travels to the Golgi apparatus where it is converted into glucosylceramide and can be further elaborated with other monosaccharides (**Figure 1.14b**). Different glycolipid series have tissue specific expression. For example, the *ganglio*-series or gangliosides are the most abundant form of glycolipid in adult mammals, with high prevalence in the central nervous system⁶². *Lacto*-series glycolipids can be found on secretory organs and the *neolacto*-series and *globo*-series can be found on white and red blood cells, respectively⁵⁹. The *archo* and *mollu* series are generally found in invertebrates and are not relevant to mammalian biology⁵⁹.

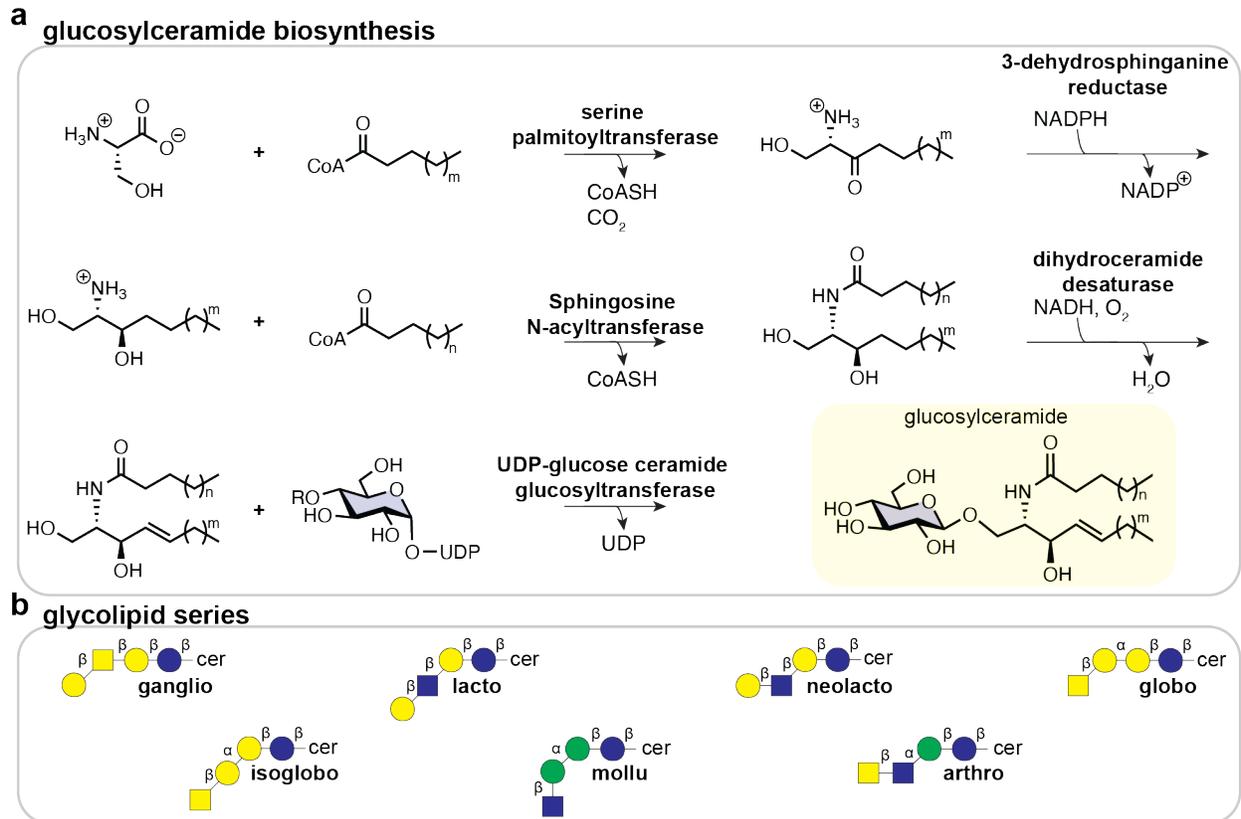


Figure 1.14: Glycolipids as a type of glycan. a, Biosynthesis of glucosylceramide the precursor to glycolipids. m (16-22), and n (18) methylene units. The alkyl chains may also be unsaturated. **b,** SNFG representation of different glycolipid series.

1.2.5: GPI-anchored proteins

Glycosylphosphatidylinositol (GPI) anchored glycans are another class of glycans but are less abundant compared to *N*-glycans, *O*-glycans, and glycolipids⁶³. The characteristics of a GPI-anchor are a glucosamine linked to a phospholipid through an inositol. This glycan seed can be further elaborated with other monosaccharides which anchor a protein to the extracellular side of the cell membrane (**Figure 1.15**)⁶⁴. The biosynthesis of the GPI anchor begins with the lipid portion in the ER, after which the protein is linked to a lipid through the glycans⁶⁴. More monosaccharides can then be added to the GPI anchor as it travels through the Golgi apparatus⁶⁴. An example of a GPI-anchored protein is CD24, which is important in cell signalling⁶⁵.

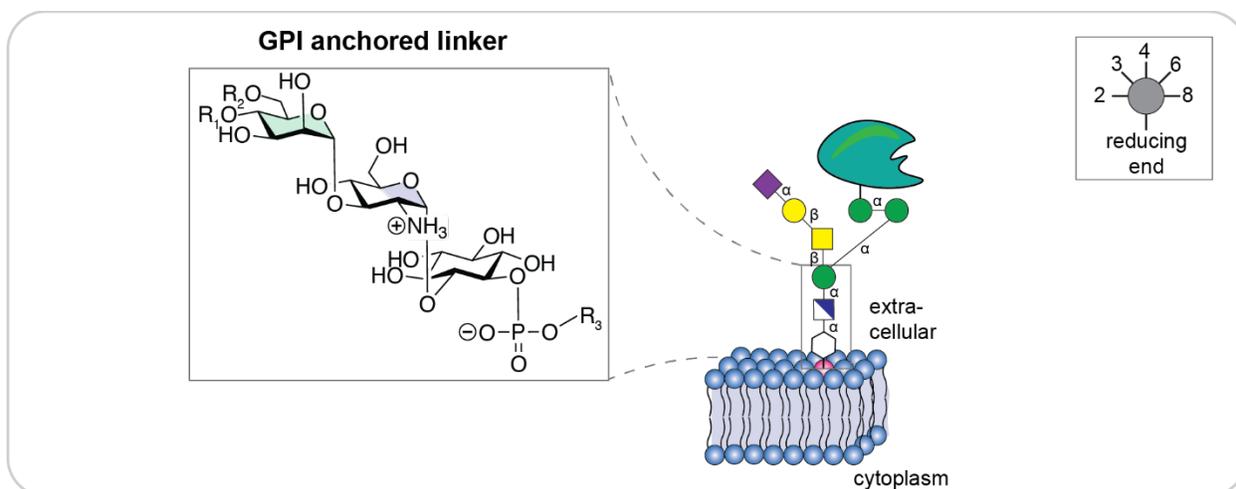


Figure 1.15: Glycosylphosphatidylinositol as a glycan. Chemical structure and SNFG representation of a GPI anchored protein. R_1 and R_2 are possible extension points of the glycan and R_3 is the remainder of the lipid.

1.2.6: C-mannose

Another infrequent but important type of glycan is C-mannosylation. Like N-glycans, C-mannosylation sites can be predicted by the sequon Trp-Xxx-Xxx-Trp/Cys where the mannose is added to the first tryptophan residue (**Figure 1.16**)⁶⁶. C-mannosylation is unique from other glycans as the monosaccharide is attached to the protein through a carbon-carbon bond catalyzed by C-mannosyltransferase, which acts in the ER. C-mannosylation is proposed to be important for protein folding and protein secretion⁶⁶. An examples of C-mannosylated proteins include type-1 cytokine receptors⁶⁶.

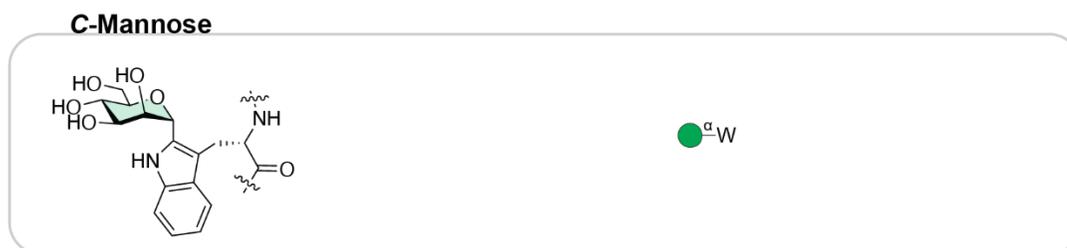


Figure 1.16: C-Mannose as a glycan. Chemical structure and SNFG representation of C-linked mannose to a tryptophan on a protein.

1.2.7: Extensions and capping of glycans

The diversity of glycans begins with the different types and subtypes of glycans (N-glycans, O-glycans, etc.), which is further increased by glycosyltransferases that build onto existing cores⁶⁷. Cores are said to be capped by specific glyco-patterns that terminate the growth

of the glycan chain⁶⁸. Common glycan caps are sialic acid and fucose. In some cases, sialic acid can be directly transferred onto a Gal in an $\alpha(2\rightarrow3)$ or $\alpha(2\rightarrow6)$ manner or $\alpha(2\rightarrow6)$ to a GalNAc residue, ending the glycan^{69, 70}. Alternatively, the glycan can be built up further before it is capped⁷¹. Many of the cores (**Figure 1.17a**) terminate with a GlcNAc residue(s), which can serve as an extension point to further diversify the glycan. One of the ways a terminal GlcNAc residue can be elongated is by either a $\beta(1\rightarrow3)$ or $\beta(1\rightarrow4)$ Gal residue, which are referred to as type-1 or type 2 LacNAc respectfully³⁰. Continuing from type-1 LacNAc, a fucose residue can be added $\alpha(1\rightarrow4)$ to the GlcNAc giving a biantennary glycan. The Gal can then be capped with an $\alpha(2\rightarrow3)$ sialic acid residue yielding the *Lewis A* antigen. *Lewis* antigens are commonly found on blood cells and many bodily fluids⁷². Alternatively, the Gal, from LacNAc, can be modified with an $\alpha(1\rightarrow2)$ fucose residue giving rise to the blood group *H* antigen. The *H* antigen can then be modified with either a $\alpha(1\rightarrow3)$ GalNAc or Gal resulting in blood group antigens *A* and *B* respectively⁴. Another possibility is for the LacNAc to be $\alpha(1\rightarrow3)$ fucosylated on the GlcNAc. This can then be followed by $\alpha(2\rightarrow3)$ sialylation of the Gal, completing the sialyl *Lewis X* antigen or the Gal can be capped with $\alpha(1\rightarrow2)$ fucose completing the *Lewis Y* antigen (**Figure 1.17b**)⁷².

The LacNAc motif can serve as a starting point for many other glycan epitopes. Starting with the simpler glycans, LacNAc can be modified with a $\alpha(1\rightarrow3)$ Gal yielding the *α -Gal* epitope, which is common in many mammals but not humans⁷³, and is the glycan epitope which can cause an allergy to red meat in response to infection by ticks⁷⁴. LacNAc can also be capped with either $\alpha(2\rightarrow3)$ or $\alpha(2\rightarrow6)$ sialic acid yielding $\alpha(2\rightarrow3/6)$ *Sialyl-LacNAc*. While this can be the end of the glycan epitope, more sialic acid residues can be added in an $\alpha(2\rightarrow8)$ linkage to the previously $\alpha(2\rightarrow3/6)$ sialic acid. This can happen just once⁷⁵, or it can happen many times yielding poly sialic acid (poly-Sia) with as many as 400 sialic acid residues⁷⁶. However, poly-Sia is not very common, with only six proteins being identified as poly-sialylated in humans⁷⁶. Alternatively, GalNAc can be added to in a $\beta(1\rightarrow4)$ fashion to $\alpha(2\rightarrow3)$ Sialyl LacNAc yielding the *Cad/Sda* epitope, which has implications in xenotransplants as it is not a common epitope in humans but is common in

other mammals such as pigs⁷⁷. Another possible outcome for LacNAc is to be further modified with more LacNAc⁷⁸. The poly LacNAc can then be capped with many motifs including $\alpha(2\rightarrow3)$ or $\alpha(2\rightarrow6)$ sialic acid (**Figure 1.17c**).

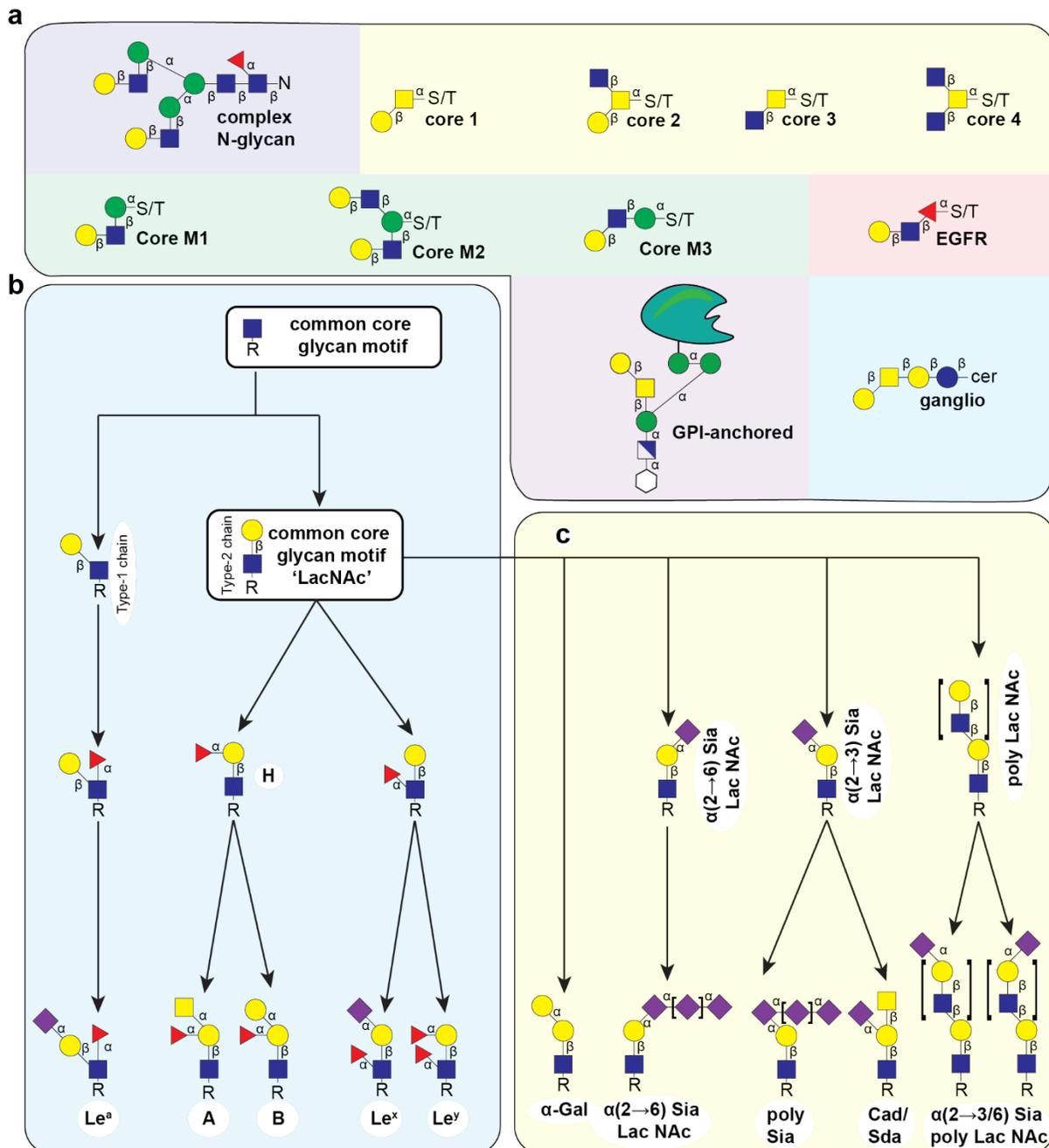


Figure 1.17: Glycan extension and capping of different glycan cores. a, Summary of glycan cores. **b**, elongation of terminal GlcNAc residue. **c**, Elongation of terminal LacNAc β -Galp-(1 \rightarrow 4)- β -GlcPNAc-(1 \rightarrow 3).

The glycan structure that is produced is governed by which glycosyltransferases and the specificity of the glycosyltransferases expressed in the Golgi apparatus²⁵. This means that the

same protein produced in two different cell types may have different types of glycosylation depending on the glycosylation machinery produced by that cell type. Moreover, glycans are very heterogeneous as not every glycosylation site may be acted on in every copy of a given protein or lipid, producing either truncated glycans or if the initial monosaccharide is missed, no glycan at all²⁵. This can complicate glycomic analysis as the same protein can be found with different glycosylation states or *glycoforms* even when produced from the same cell. Glycan structures are further diversified by post glycosylation modifications. Examples of this include Gal and GlcNAc sulfation (**Figure 1.18**)^{79, 80}. This takes place in the Golgi apparatus and is accomplished using a family of enzymes referred to as carbohydrate sulfotransferases (CHST) and an activated form of sulfate called 3'-phosphoadenosine-5'-phosphosulfate (PAPS)⁸¹. The addition of sulfation to glycans affords many properties such as changing interactions with lectins^{79, 80} and how glycans are metabolized⁸².

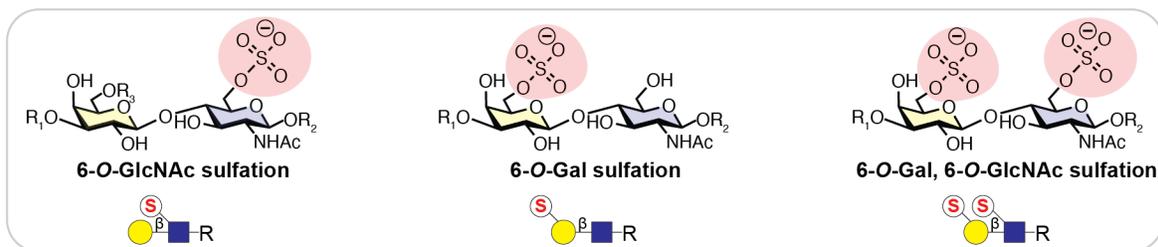


Figure 1.18: Chemical structures of mono- and di-sulfated LacNAc. R₁, R₂, and R₃ are possible branches of the glycans.

Each of these modifications and combinations of these modifications are possible, demonstrating the complexities and variety of structures that populate the glycocalyx of any given cell. However, this is not an exhaustive list of glycans, and this section serves only to illustrate the vast structural complexity of the glycocalyx (**Figure 1.19**).

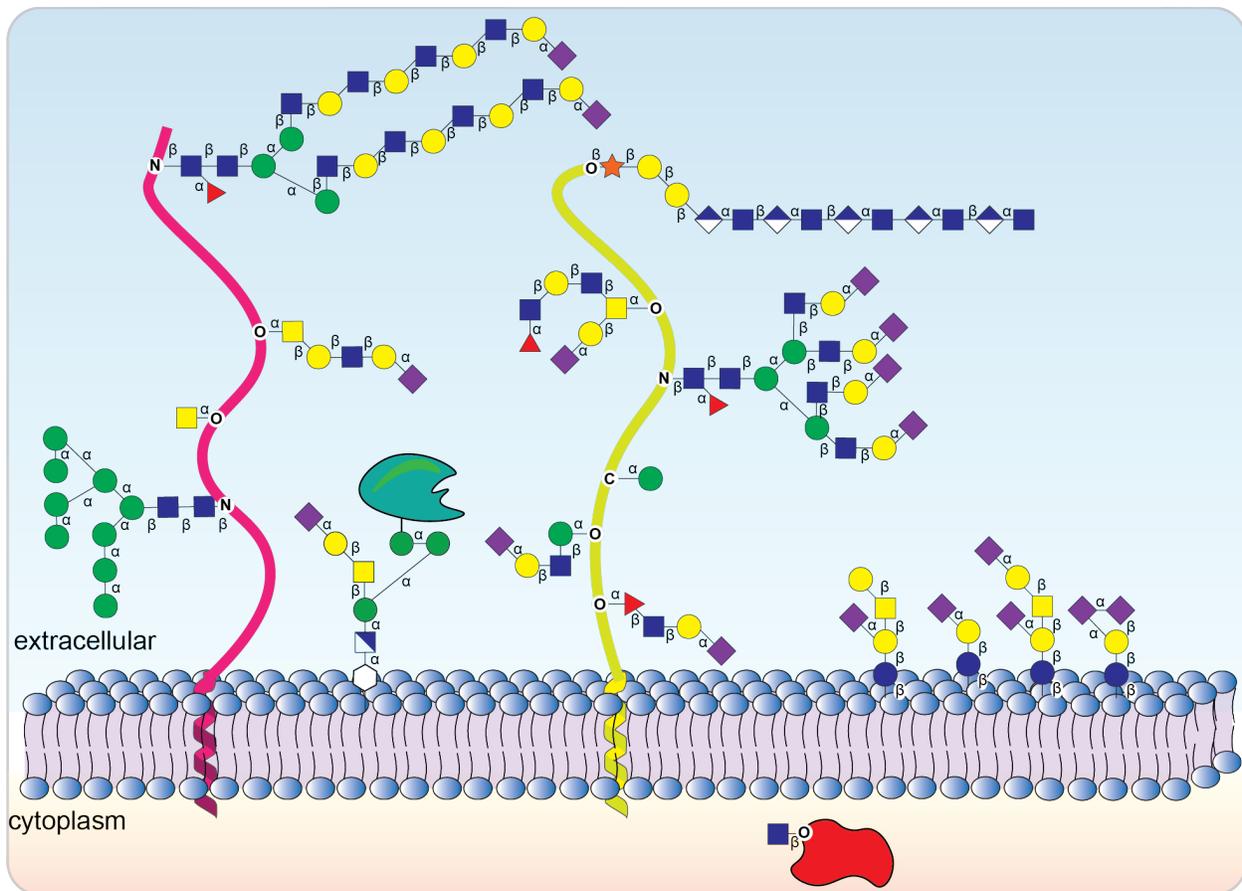


Figure 1.19: SNFG depiction of the glycocalyx featuring *N*-glycans, *O*-glycans, glycolipids GPI-anchored glycans C-mannose glycans, and *O*-GlcNAc modified glycans.

1.2.8: Sialic acid as a cap for glycans

One of the most common and biologically important caps for glycans in mammals is sialic acid^{41, 75, 83}. Sialic acid is a nine-carbon alpha-keto acid monosaccharide and its biosynthesis begins in the cytoplasm with the sugar donor UDP-GlcNAc which is converted into *N*-acetylmannosamine 6-phosphate (Man β NAc-6-P) by the bifunctional enzyme GNE (**Figure 1.20a, b**)⁸⁴. A carbon-carbon bond is then formed between Man β NAc-6-P and phosphoenolpyruvate by Neu5NAc-9-P-synthase to *N*-acetylneuraminic acid (Neu5Ac) 9-phosphate. The product is then dephosphorylated by a phosphatase yielding Neu5Ac that is subsequently transported into the nucleus for it to be charged into its donor form CMP-Neu5Ac by the enzyme *N*-acetylneuraminic acid cytidyltransferase (CMAS). CMP-Neu5Ac is then transported out of the nucleus and into the Golgi apparatus where it is then transferred to a glycan. CMP-Neu5Ac biosynthesis is tightly

regulated through feedback inhibiting GNE, the first enzyme in the biosynthesis of CMP-Neu5Ac⁸⁴.

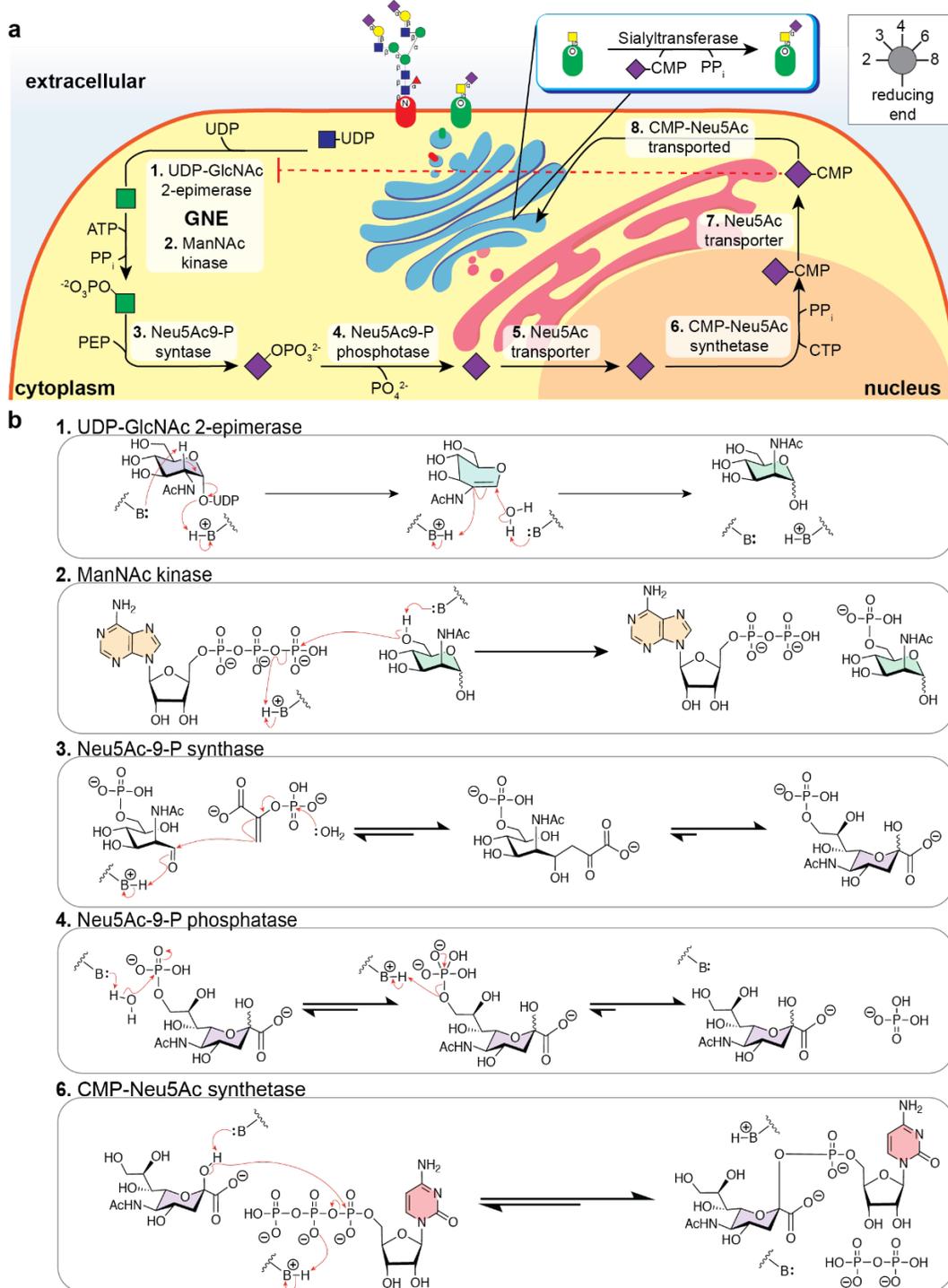
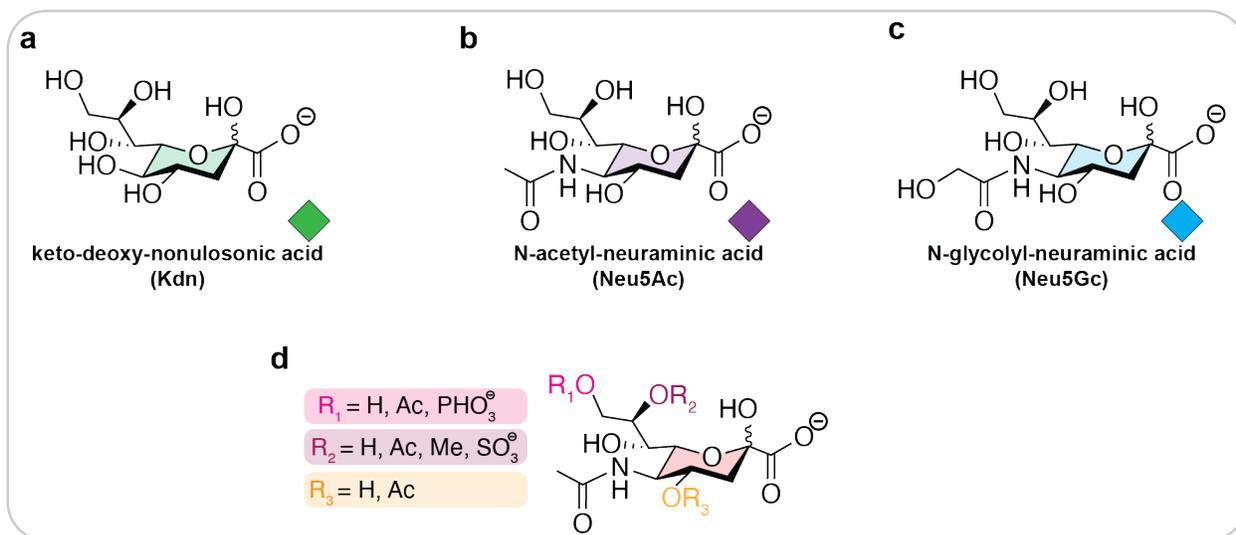
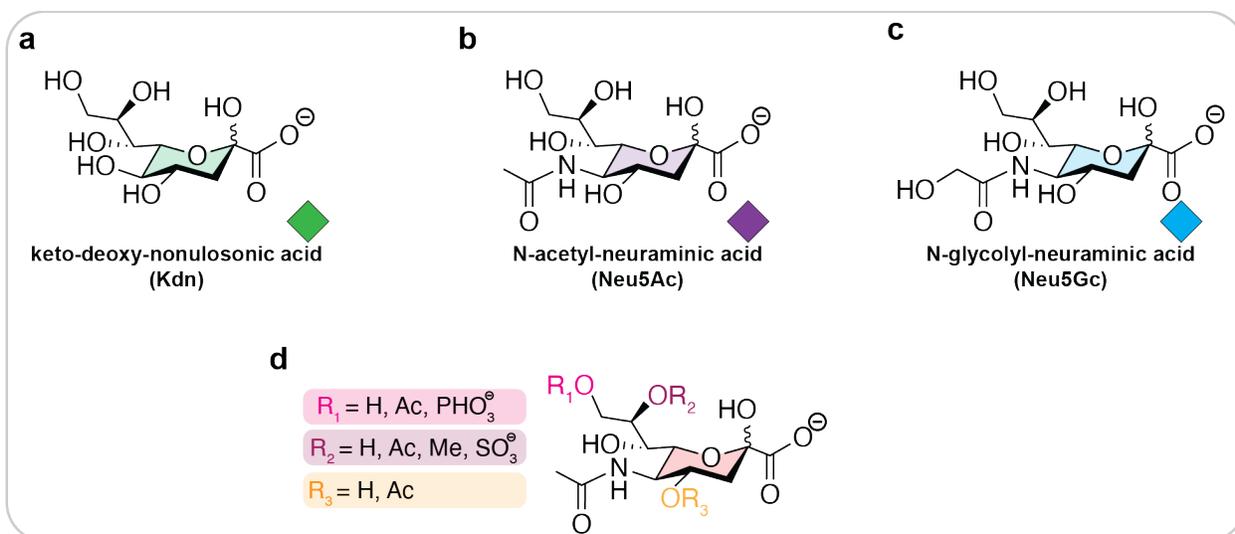


Figure 1.20: Biosynthesis of *N*-acetylneuraminic acid (Neu5Ac). a, b, biological and chemical depiction the biosynthesis of Neu5Ac respectively.

Sialic acid is generally specific to vertebrates, however there are many species-specific, tissue specific and cell to cell chemical variations⁸⁵. For instance, keto-deoxy-nonulonic acid (Kdn) is commonly found in the tissues of fish, whereas *N*-glycolyl-neuraminic acid (Neu5Gc) and *N*-acetyl-neuraminic acid (Neu5Ac) are the predominant forms of sialic acid in mammals(**Figure**

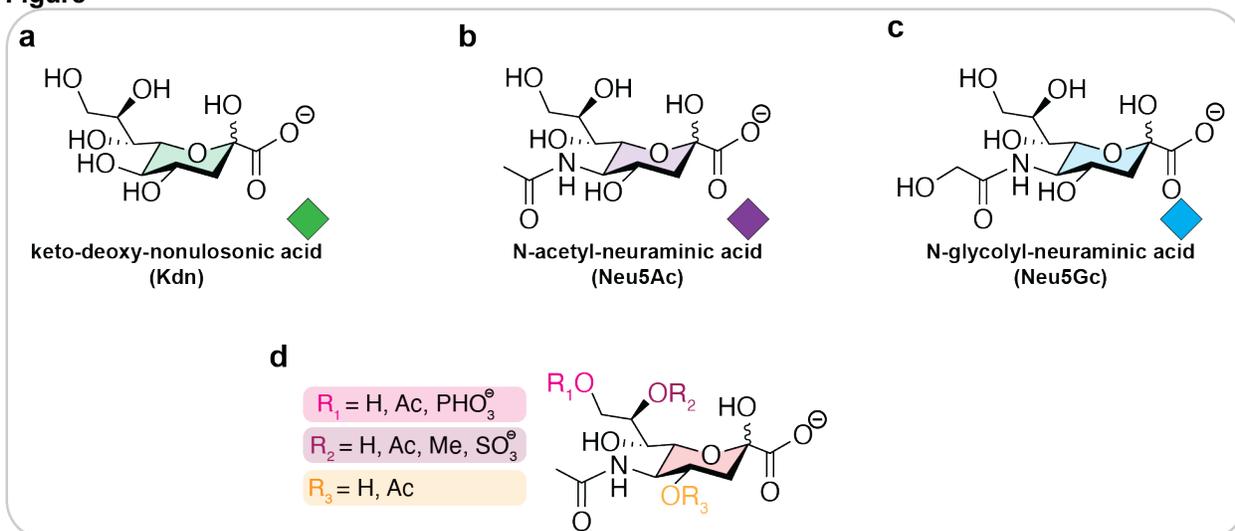


1.21a-c)^{86,87}. In many mammals, both Neu5Ac and Neu5Gc are present in a cell type/tissue dependant manner. However, Neu5Ac is the only type of sialic acid biosynthetically produced in humans, as humans have lost a functional gene for the enzyme (cytidine monophospho-*N*-acetylneuraminic acid hydroxylase-CMAH) that converts CMP-Neu5Ac into CMP-Neu5Gc^{85, 88}. Each form of sialic acid can be further modified at carbon 4, 8, and 9 with acetyl groups and with phosphorylation and sulfation being possible at carbon 8 and 9, respectively (**Figure**



1.21d)⁸⁹. However, the modification of sialic acid with phosphorylation and sulfation have not been robustly investigated.

Figure



1.21: Different types of sialic acid. **a**, **b**, and **c** Chemical structures and SNFG symbols of Kdn, Neu5Ac, and Neu5Gc respectively. **d**, Possible chemical modifications of sialic acids.

1.2.9: Sialic acid as a self-associated molecular pattern

As described above, sialic acid is generally specific to vertebrates and is not commonly (some species of bacteria produce sialic acid) found across the other kingdoms of life such as bacteria, fungi, and protozoa, which is important as organisms from these kingdoms are often pathogenic to mammals and other vertebrates⁹⁰. This means that sialic acid is a unique marker

that can be used to differentiate animal cells from other types of cells. To this end, sialic acid can be considered a self-associated molecular pattern (SAMP), which can help immune cells differentiate *self*-cells from pathogens preventing autoimmunity while destroying pathogens⁹¹. However, this system of *self*-differentiation has placed a selective pressure on pathogens and some bacteria have evolved to make their own sialic acid or scavenge sialic acid from their host to hide from the immune system⁹². Sialic acid is also important in viral infection as viral proteins are synthesised in their host's cells and can be glycosylated and sialylated by the same machinery that glycosylates the host proteins⁹³. Moreover, enveloped viruses take pieces of the host cells' bilayer when they bud off from the cells taking many of the host cells' glycans with it^{9,94}. Following the same motive, many cancerous cells increase the amount of sialic acid on their surface to avoid being destroyed by immune cells (**Figure 1.22**)^{95,96}. While some of the properties sialic acid imparts to the glycocalyx may be nuanced, one of the most obvious properties sialic acid affords glycans is the ability to bind to lectins and modulate immune cell function⁹⁷.

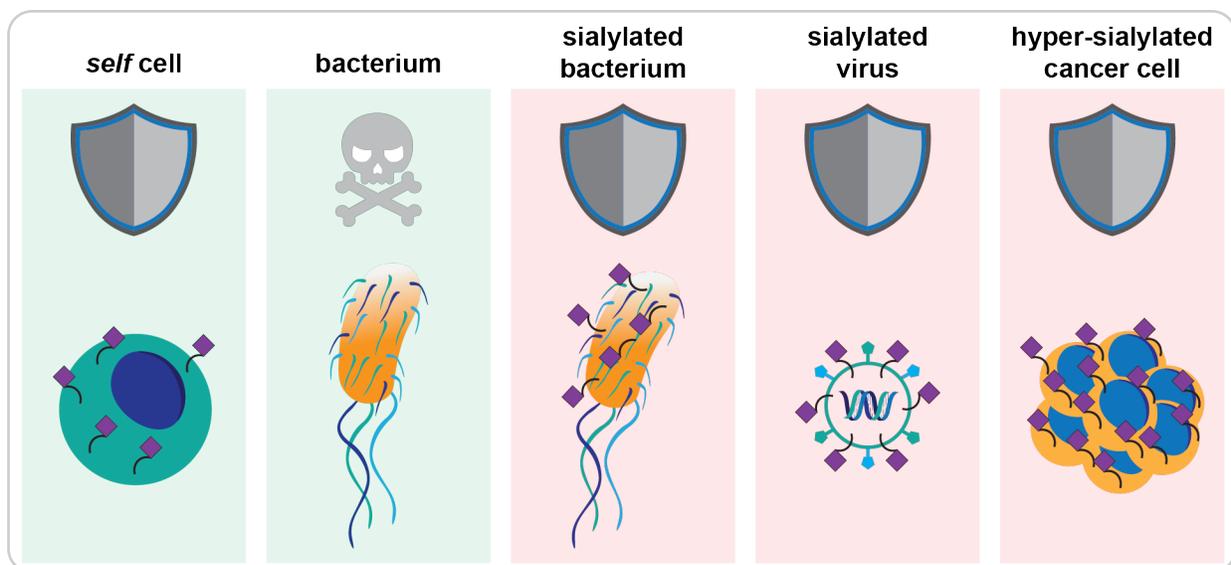


Figure 1.22: Biological roles of sialic acid. Sialic acid is a self-associated molecular pattern which can help the immune system differentiate *self*-cells from non-*self*-cells (highlighted in green). However, some pathogens and cancers take advantage of this system to avoid or exploit the immune system (highlighted in red).

1.3: The Siglec family of cell surface receptors

The sialic-acid-binding immunoglobulin-like lectins or Siglecs are a family of cell surface receptors that are typically found on the immune cells of vertebrates. Siglecs are defined as I-type lectins due to their homology with the immunoglobulin superfamily and their ability to bind sialylated glycans⁹⁸. There are several structural components that make up a Siglec (**Figure 1.23a**). First, there is the extracellular *N*-terminal glycan-binding domain known as a V (variable)-set domain, which is responsible for engaging with the glycan⁶. Within the V-set domain of a Siglec, there is a common positively charged arginine which forms a salt bridge with the negatively charged carboxylate of sialic acid (**Figure 1.23b**). All Siglecs also have at least one IgG-like (i.e. C (constant) 2) domain that serves as a spacer to place the V-set domain away from the plasma membrane of the cell. The number of C2 domains in a Siglec can range from sixteen (Siglec-1) to one (Siglec-15)⁹⁸. Another common structural feature between Siglecs is a disulfide bond between the V-set domain and the first C2 domain which, provides more structure to the V-set domain. Siglecs are anchored into the plasma membrane by a single pass transmembrane segment. Some Siglecs such as Siglec-15 have a positively charged residue in their transmembrane domain which requires an adaptor proteins such as DAP12 to exist in a membrane⁹⁹. The cytoplasmic tail of the Siglec is presumably unstructured and, with the exception of Siglec-1, contain signalling motifs¹⁰⁰. The majority of Siglecs contain motifs that antagonize immune cell signaling, although some Siglecs, such as Siglec-15, are activatory in nature owing to their pairing with DAP12 and the absence of an inhibitory motif¹⁰¹. The inhibitory signaling motifs are described as immunoreceptor tyrosine-based inhibitory (ITIM) or ITIM-like¹⁰⁰. The number of these motifs can also vary between Siglecs. These structural features unite the Siglecs as a family of receptors, however, finer aspects of their structure and function can vary across species¹⁰⁰.

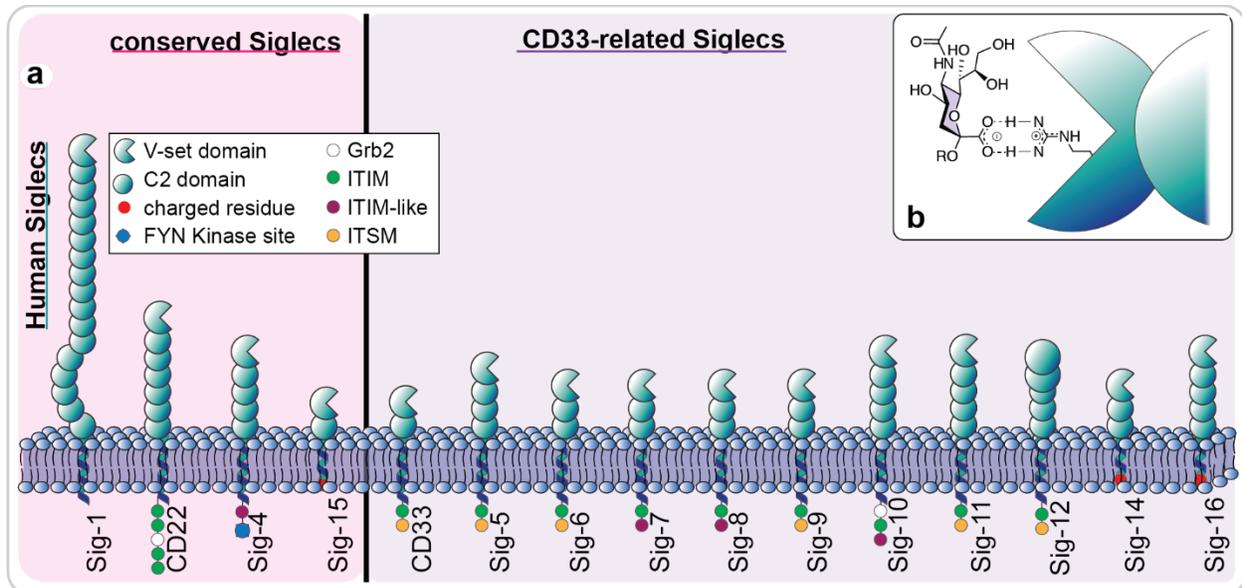


Figure 1.23: The human Siglec family. **a**, Representation of the human Siglec family. **b**, Depiction of the critical salt bridge that forms between an arginine residue in the V-set domain of the Siglec and the carboxylate of the sialoside.

1.3.1: Siglecs across species

Siglecs first evolved approximately 400 million years ago as all jawed vertebrates have Siglecs in some capacity (**Figure 1.24**)¹⁰². Siglecs can be divided into two groups: the conserved Siglecs and the CD33-related Siglecs¹⁰³. As their name implies, the conserved Siglecs are conserved across vertebrates with respect to primary sequence similarity and cell type expression¹⁰⁴. The CD33-related Siglecs are said to be rapidly evolving and tend to be unique to each species. However, some of the members of the CD33-related sub-family of Siglecs can be likened between species as orthologs or paralogs while others are unique to that species. Orthologs result from new genes evolving from a common ancestral gene and paralogs result from a gene duplication event¹⁰⁵. In humans, there are fifteen Siglecs and they are referred to by a number. The conserved Siglecs are Siglecs -1, -2, -4, and -15 while the CD33 related Siglecs are Siglec-3 through Siglec-12, Siglec-14, and Siglec-16. Siglec-12 is often left out when discussing the Siglec family as its glycan binding domain has evolved such that it cannot bind glycans¹⁰⁶. In addition to the human Siglec-sialic acid-axis being different from other mammals, due to the absence of Neu5Gc in their glycans, humans also lack Siglec-13 while it is present in

other mammals such as chimpanzees¹⁰⁷. As Siglecs are conserved across vertebrates many organisms such as zebra fish¹⁰⁸, mice¹⁰⁹ and great apes¹¹⁰ can serve as model organisms to study the roles of Siglecs in human physiology. However, there are structural differences between Siglecs from different species which may have biological consequences, and outside of a few examples such as CD22¹¹¹, how Siglecs differ functionally between species has been largely unexplored.

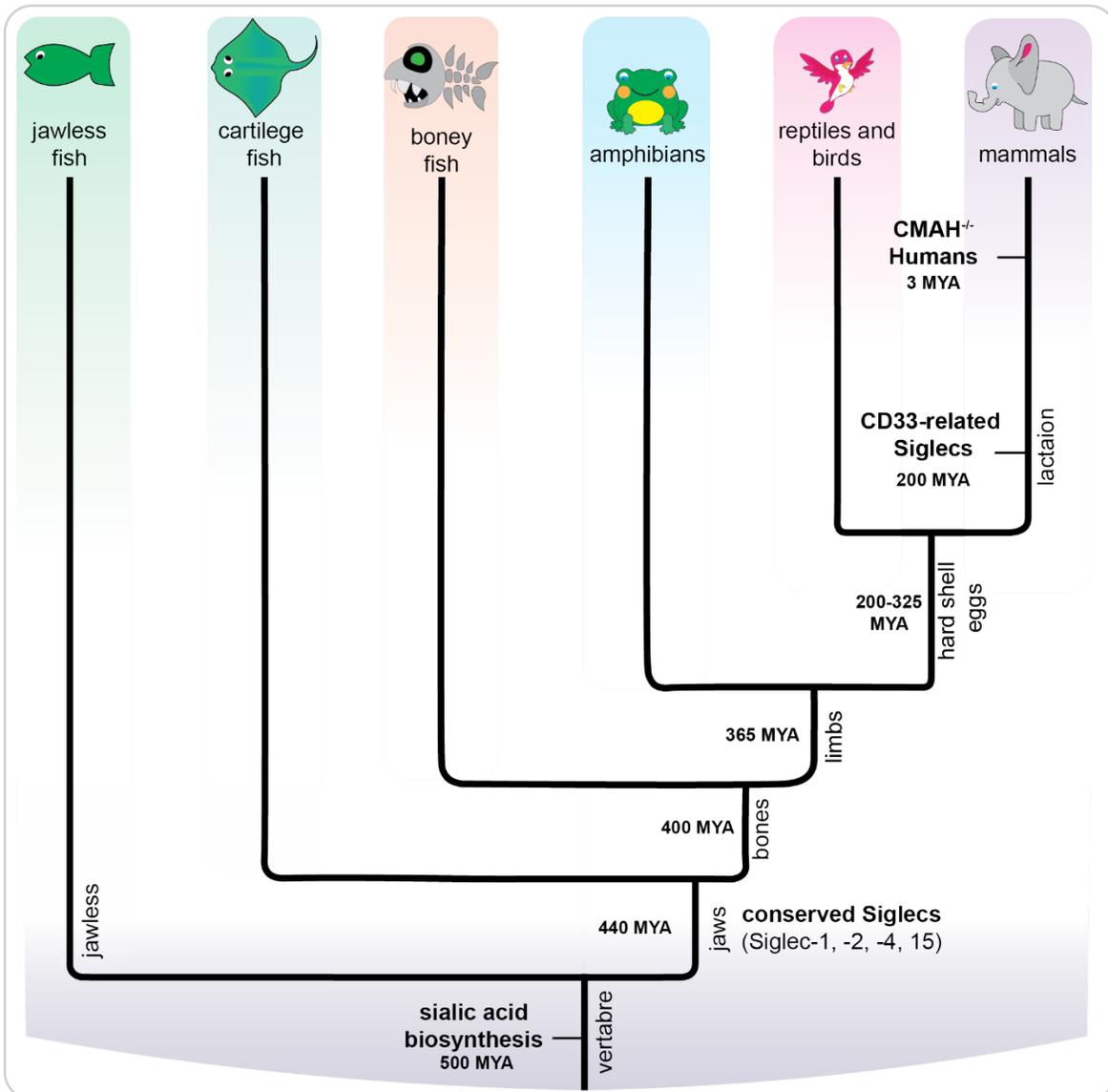


Figure 1.24: Evolution of Siglec-sialic acid axis. The relevant highlights of the evolution of Siglecs. The first appearance of sialic acid biosynthesis approximately 500 million years ago (MYA) shortly followed by the evolution of the conserved Siglecs (440 MYA). The CD33 related Siglecs evolved approximately 200 MYA and CMAH (cytidine monophospho-*N*-acetylneuraminic acid hydroxylase) was lost in humans approximately 3 MYA.

1.4: Siglec ligand binding

1.4.1 *cis* vs *trans* binding

Siglec-glycan binding can be divided into two types. One type is referred to as *cis* binding, which refers to when the Siglec and the glycan are on the same cell (**Figure 1.25a**). The other type of binding is known as *trans* binding, where the Siglec ligand is on another cell, particle, or soluble protein (**Figure 1.25b**)¹¹². When a Siglec is bound to a *cis* ligand, the Siglec is said to be *masked* as it is limited in its ability to bind in *trans* (**Figure 1.25c**)¹⁰⁰. However, if the *cis* ligands are removed, the Siglec is said to be *unmasked* (**Figure 1.25d**), making *trans* binding more favoured. This balance between *cis* and *trans* binding is relevant when studying the physiological roles of Siglecs.

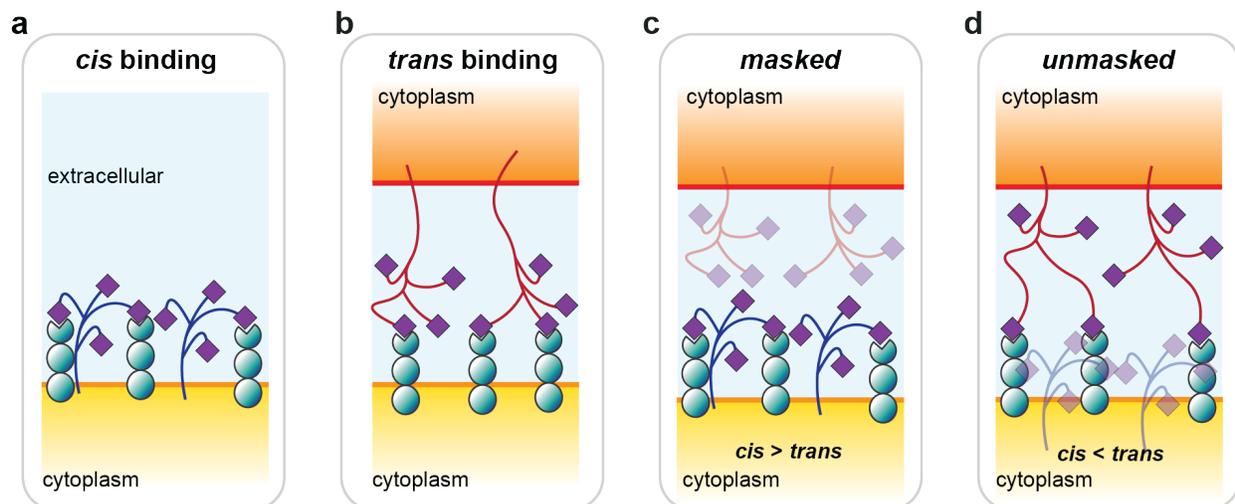


Figure 1.25: Siglec binding can be either in *cis* or *trans* and depending on the relationship between the ligand and the Siglec. a, b, depiction of *cis* and *trans* Siglec binding where the Siglec and the ligand are on the same cell/particle or different cell/particle respectively. c, d depiction of Siglecs being masked and binding *cis* ligands over *trans* ligands or unmasked and binding *trans* over *cis* ligands respectively.

1.4.2 Siglec-ligand interactions are driven by non-covalent interactions

While each Siglec has a different preference for different glycan motifs⁶, the binding between a Siglec and its ligand is a reversible non-covalent interaction that is driven by the sum of many weaker interactions¹¹³. Gibbs free energy is often used as a measure of the strength of a given bond or non-covalent interaction. For reference, the Gibbs free energy required to break a carbon-carbon bond is approximately 90 kJ/mol (**Figure 1.26**)¹¹⁴. The binding of a Siglec to its ligand is largely driven by the salt bridge formed between the carboxylate anion of the sialic acid

and the guanidinium cation of the arginine residue from the Siglec. Ionic interactions, such as this have a Gibbs free energy of approximately 5 kJ/mol¹¹⁵. Other non-covalent interactions that are important for Siglec binding are hydrogen bonds which typically have a Gibbs free energy of 1-2 kJ/mol¹¹⁶. In some cases, such as with Siglec-2 (CD22), a significant decrease in binding to sialosides with 9-O-acetylation was observed, which is likely due to the loss of the hydrogen bond between the 9 hydroxyl on the sialic acid and Siglec¹¹⁷. Dipole-dipole interactions likely also contribute to the binding between a Siglec and a sialoside¹¹⁸. Another key contributor to Siglec binding is the hydrophobic interaction. Binding between any two molecules is inherently entropically disfavoured, however the properties of an aqueous environment can be leveraged to make binding favourable due to the hydrophobic effect and Van der Waals or dispersion forces¹¹⁹. When a hydrophobic molecule or a hydrophobic patch of a protein is placed in an aqueous environment, it causes ordering of the water molecules around the hydrophobic area. Siglecs tend to have hydrophobic patches in their binding site that are occupied by ordered water molecules¹¹⁹. When the sialoside binds to the Siglec, the hydrophobic areas of the glycan can line up with these hydrophobic patches on the Siglec releasing the water to the solvent, where it is then disordered making the binding entropically favourable. Moreover, hydrophobic groups can form favourable interactions through London dispersion forces. Through the sum of all these relatively weak interactions, Siglec-sialoside binding is possible, however it is relatively weak³⁷ when compared to other biological interactions.

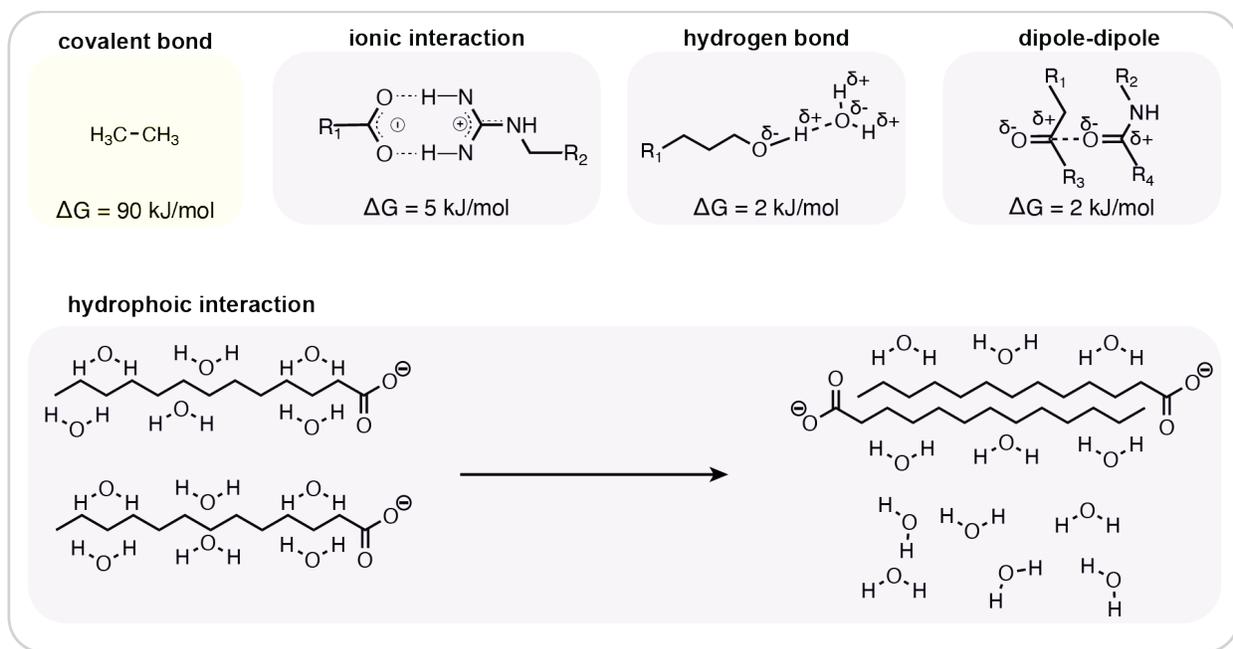


Figure 1.26: Non-covalent interactions drive the interaction between a Siglec and sialosides. Comparison of the Gibbs free energy between different kinds of non-covalent interactions as well as the amount of energy required to break a carbon-carbon bond for comparison.

1.4.3: Relative strength of Siglec-glycan interactions

Protein-ligand interactions can be simplified to an equilibrium illustrated in **Figure 1.27a**. Typically, protein-ligand interactions are quantified by a dissociation constant or K_d (**Figure 1.27b**). The K_d is often used as a measure of binding strength of a given interaction as it is the concentration at which half the binding sites on the protein are occupied by the ligand on the protein. Typical K_d ranges for a Siglec and a sialoside are 10^{-3} - 10^{-5} M (10 μM – 1 mM) depending on the Siglec and the sialoside³⁷. When comparing this with other non-covalent biological interactions such as insulin/insulin receptor (10^{-9} M)¹²⁰, antibody/antigen (10^{-9} - 10^{-12} M)¹²¹, or streptavidin/biotin (10^{-14} - 10^{-15} M)¹²², Siglec-sialoside interactions are relatively weak (**Figure 1.27c**). However, Siglecs overcome their relatively weak affinity for their ligands through a phenomenon known as avidity (**Figure 1.27d**)¹²³. Avidity can be simplified as many weak interactions adding up to a strong interaction or “a whole is greater than the sum of its parts” and interactions that leverage avidity are often likened to Velcro®. However, the effect of the number of binding pairs on avidity is not linear and modeling these interactions is non-trivial¹²⁴. Avidity is the multivalent binding strength between two cells, a cell and a particle, or two particles.

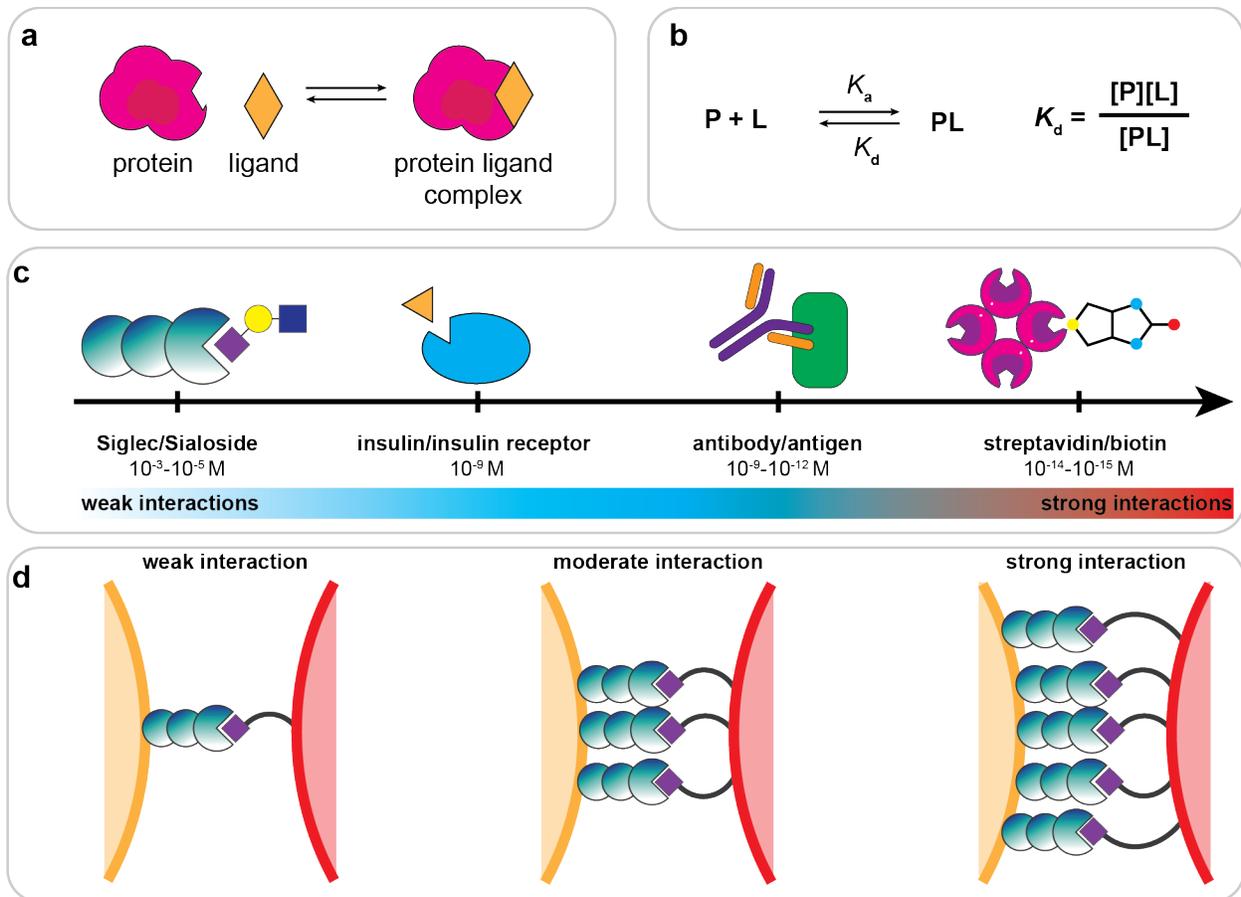


Figure 1.27: Protein ligand binding and comparison of the binding strength between receptor ligand pairs in biological systems. **a**, Simple depiction of protein ligand interaction. **b**, Formula for calculating the dissociation constant K_d for protein ligand interactions. **c**, comparison of the binding strength as measured by K_d between various protein ligand pairs. **d**, depiction of avidity between Siglecs and sialosides.

1.5: Siglecs in biological systems

1.5.1: Siglec functions

Siglecs, which are generally expressed by immune cells, have many functions. Most of these functions are controlled by the binding of a Siglec to its ligand(s)^{100, 125}. One of the functions of Siglecs is cell-adhesion. Indeed, Siglec-1 was first known as *Sialoadhesin* as it was found to help macrophages adhere to other cells through sialic acid containing glycans (**Figure 1.28a**)¹⁰⁰. Another function of Siglecs is to endocytose particles, such as extracellular vesicles and viruses. Many Siglecs including Siglec-1⁹⁴, -2¹²⁶, -3¹²⁷, -6¹²⁸, and -9¹²⁹ have all been known to endocytose (internalize) particles (**Figure 1.28b**). While internalization of particles by Siglecs is believed to be largely driven by the intracellular signaling motifs, interestingly Siglec-1 does not possess any signaling motifs, yet it is still capable of internalizing particles. This suggests that the mechanism of internalization may not be conserved across Siglecs. In addition to internalization, Siglecs are also known to regulate cell signaling through glycan binding, this can be in *trans* where the Siglecs on one cell affect the signalling pathway on another cell (**Figure 1.28c**) or in *cis* where the Siglec affects the signalling on the same cell as the Siglec (**Figure 1.28d**). Examples of Siglecs modulating cell signalling in *trans* are rare, but one example is Siglec-7 which has been shown to decrease T cell activation however the mechanism(s) of this are not well understood¹³⁰. Additionally, Siglec-9 has been proposed to induce apoptosis in cells by binding in *trans*¹³¹. Siglec *cis*-signaling has been more thoroughly studied and many Siglecs including Siglec-2, -3, -7, and -9 have been shown to inhibit cellular activation through Fc receptors^{132, 133}. Siglecs are not like other types of cell surface receptors such as G-protein coupled receptors¹³⁴ that undergo conformational change when they bind their ligands. Instead, Siglecs need to be in close spatial proximity to their counter receptor in order to dampen a signalling cascade¹³². Once in proximity to the activatory receptor, the ITIM and or ITIM-like motifs are phosphorylated by Src kinases. Following phosphorylation, the Siglec then recruits phosphatases such as SHP-1 or -2 which then act on the activatory receptor dampening the activatory signal¹³². All of these functions make Siglecs

multifunctional cell surface receptors that immune cells can use to battle pathogens and maintain immune homeostasis. Because the many functions of Siglecs are initiated by Siglec binding to ligand(s), understanding Siglec ligands is paramount for understanding Siglec function.

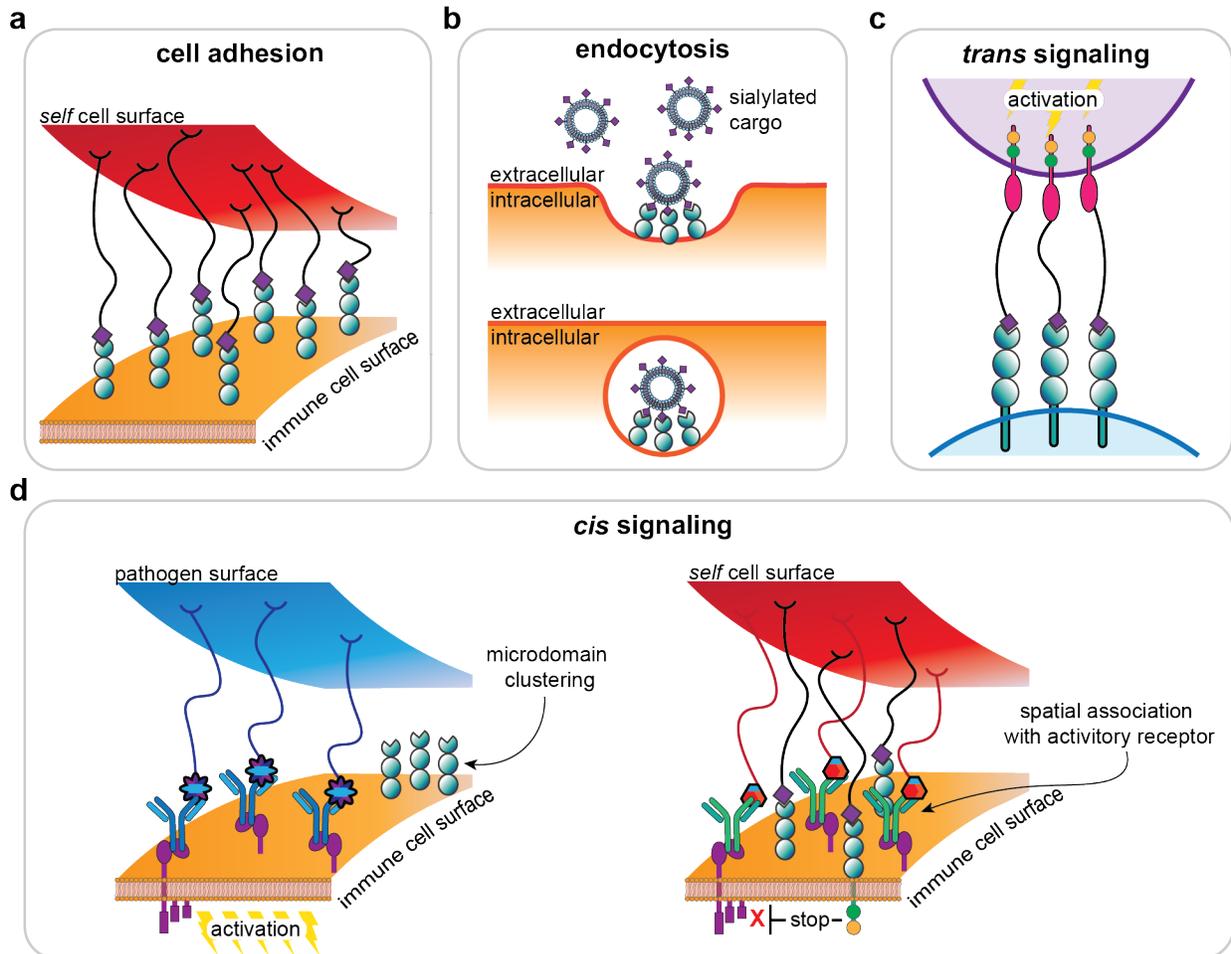


Figure 1.28: Biological roles of Siglecs. **a**, Siglecs contribute to cell adhesion and immune synapse formation through binding ligands in *trans*. **b**, Siglecs mediate endocytosis (internalization) of extracellular particles such as viruses or extracellular vesicles. **c**, Siglecs can affect cell signaling in *trans* by clustering ligands on other cells into microdomains. **d**, Siglecs can affect cell signaling in *cis* by being brought into close spatial proximity to an activatory receptor.

1.6: Current state of Siglec ligands

1.6.1 Approaches to describe Siglec Ligands

To understand the biological roles of a Siglec, a complete description of their ligands is required. Physiological and pathophysiological circumstances when these ligands change have the potential to drive biological effects through altered Siglec function. To this end, many studies have investigated which type(s) of sialoside a Siglec can bind. This has been accomplished with a variety of approaches including glycan microarrays, ELISAs, cell-based glycan arrays, probe/nanoparticle-based glycan arrays, lectin microarrays, and mass spectrometry-based approaches to name a few^{37, 79, 128, 135, 136}; and each approach has its own advantages and disadvantages.

1.6.2 Plate-based glycan arrays

Surface-based glycan arrays are a high throughput approach ideal for screening large libraries of glycans against many Siglecs. Surface-based glycan arrays can be divided further into glycan microarrays (GMA)^{136, 137, 138, 139, 140} and enzyme-linked immunosorbent assays (ELISAs)^{128, 141, 142, 143}. In both a GMA and ELISA, glycans are displayed from a solid surface such as a glass slide or microplate well respectively (**Figure 1.29**). A general difference between the two approaches is that in a GMA the glycans are chemically linked to the surface whereas in an ELISA the glycans are non-specifically adsorbed to the surface^{144, 145}. The glycans are then probed with a soluble version of a Siglec and is detected with a fluorescently labeled secondary antibody in a GMA or an antibody conjugated to a reporting enzyme in an ELISA. An additional difference between a GMA and ELISA is GMAs typically probe a larger library of compounds compared to ELISAs. A disadvantage of surface based glycan arrays is that the glycans are presented in an unnatural way which can affect the Siglec's ability to recognize the glycan. Moreover, the glycans in a surface-based glycan array are typically synthesised and, therefore, the pool of glycans is limited to what is synthetically accessible.

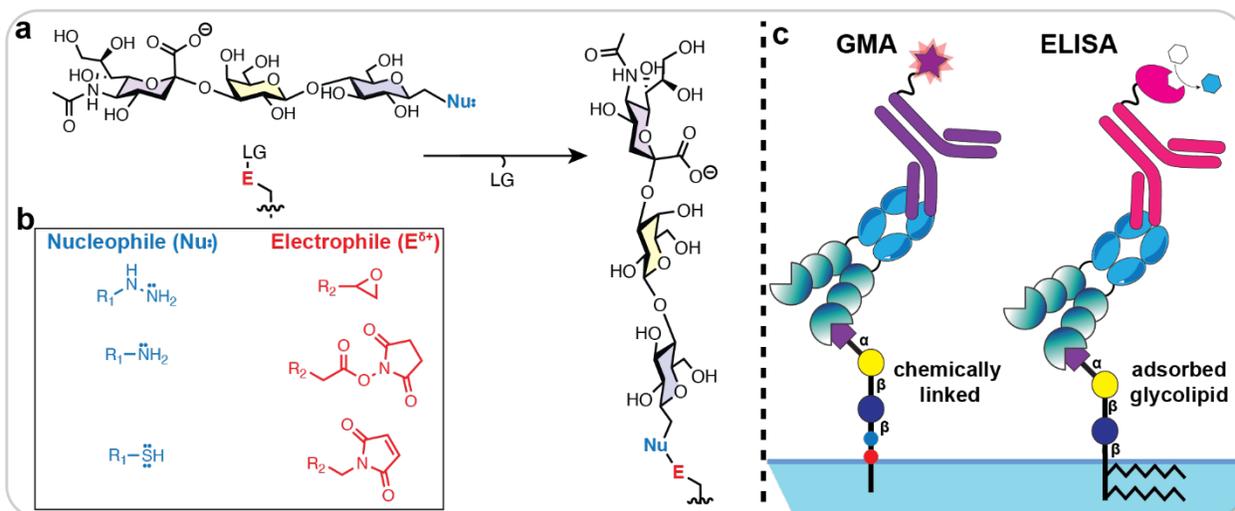


Figure 1.29: Glycan-arrays for discovering Siglec ligands. **a**, depiction of general synthetic scheme used to link a glycan to a surface. It is common for the glycan to be modified at the reducing end with a nucleophilic functional group whereas there is an electrophilic functional group on the surface. **b**, Examples of commonly used nucleophile and electrophiles for linking glycans to a surface. **c**, Schematic of how Siglec-glycan binding is reported. In a GMA (glycan microarray) the binding between a Siglec and glycan is typically reported using a fluorescently labeled secondary. In an ELISA (enzyme-linked immunosorbent assays), the binding between the Siglec and glycan is quantified colourimetrically by a secondary antibody that is conjugated to an enzyme such as horseradish peroxidase. Nu, nucleophile; LG, leaving group.

1.6.3 Cell-based glycan arrays

Another common approach is a cell-based glycan array where the glycans on cells are probed directly using a soluble version of a Siglecs (**Figure 1.30**)^{37, 79, 80}. Cell-based glycan arrays are ideal for discovering if Siglec ligands are present on a given cell. Using genuine/primary cells and tissues ensures the glycans are presented in a biologically relevant context; however, it can be challenging to determine specific Siglec-glycan interactions. Cell-based glycan arrays can be complimented by genetically engineering cells that express different carbohydrate modifying enzymes such as carbohydrate sulfotransferases⁷⁹, CMAH or specific sialyltransferases. The effect of a glycosyltransferase on Siglec ligands can also be probed by knocking-out different glycosylation pathway enzymes⁸⁰. Moreover, glycans on a cell can also be enzymatically modified with neuraminidases or can be built up further using exo-enzymatic modifications^{146, 147, 148}. Lastly, pharmacological inhibitors can also be used to modify the glycans of a cell¹⁴⁹.

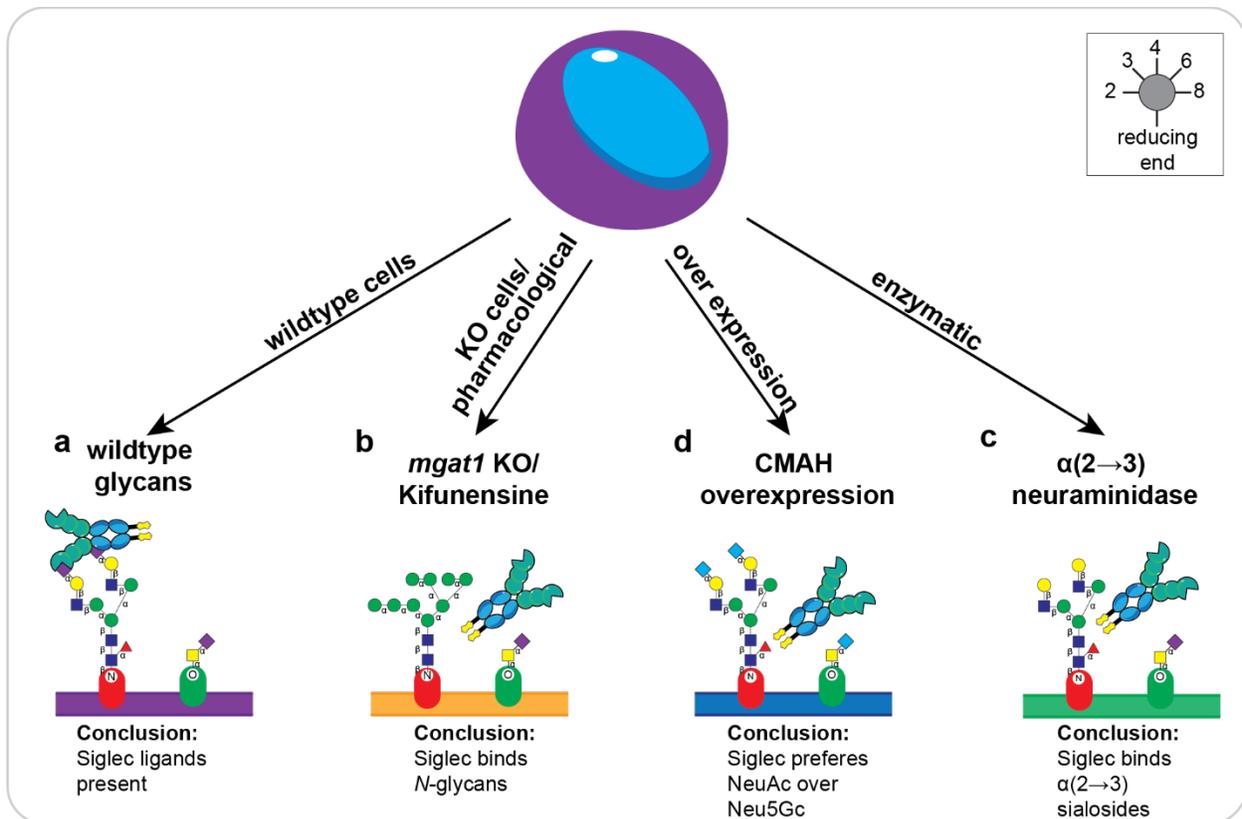


Figure 1.30: Cell-Based arrays for discovering Siglec ligands. Cell-based glycan arrays are ideal for probing Siglec ligands in a natural context on cell lines or primary cells (a). Moreover, cell glycans can be modified to have a defect in a glycosylation pathway either genetically, pharmacologically or by using small interfering RNAs (b). Alternatively, or overexpressing a glycan modifying enzyme can be used to explore the effect of an enzyme on Siglec binding. Another popular strategy is to treat cells with glycan modifying enzymes such as neuraminidases, mucinases, etc. to explore the effect of the enzyme on Siglec binding to the cells.

1.6.4 Nanoparticle-based glycan arrays

Nanoparticles decorated with glycans are another common approach for discovering Siglec ligands (Figure 1.31). There are many different types of nanoparticles, but common nanoparticles used for probing Siglec ligands include liposomes^{37, 128, 150, 151, 152} and bacteriophages^{128, 135, 153}. Alternatively, a glycan of interest can be covalently linked to a protein such as bovine serum albumin, but this is less common¹⁵⁴. Advantages of using glyco-nanoparticles are that the glycans are defined and homogenous and that the glycan itself is not the reporter. Additionally, for glycolipids presented from liposomes, the glycan is presented in a more biologically relevant context compared to glycan arrays or ELISAs, and that the glycan density can be easily modulated. Historically, Siglec-glycolipid interactions have been studied

outside the context of a lipid bilayer such as in a glycan microarray/ELISA^{137, 141}. Alternately, glycans have been formulated into lipids by conjugating the glycan through long polymers to lipids^{127, 155}; again, separating the glycan from the bilayer and presenting the ligand in an unnatural way. Moreover, both liposomes and phages are generally not immunogenic or toxic making them ideal for profiling Siglec-ligands *in vivo*^{156, 157}. Generally, glycan bearing liposomes are fluorescently labeled and then added to Siglec-expressing cells and the interaction between the particle and the cell can be quantified by flow cytometry. In the case of glycan decorated phages, the phages are labeled with a DNA barcode which enables the binding between the glyco-phage and a Siglec-expressing cell to be quantified through next generation sequencing. The downside of using glycan labeled nanoparticles is the work to prepare libraries of glycan bearing nanoparticles with a high degree of quality control, making it a relatively low throughput approach.

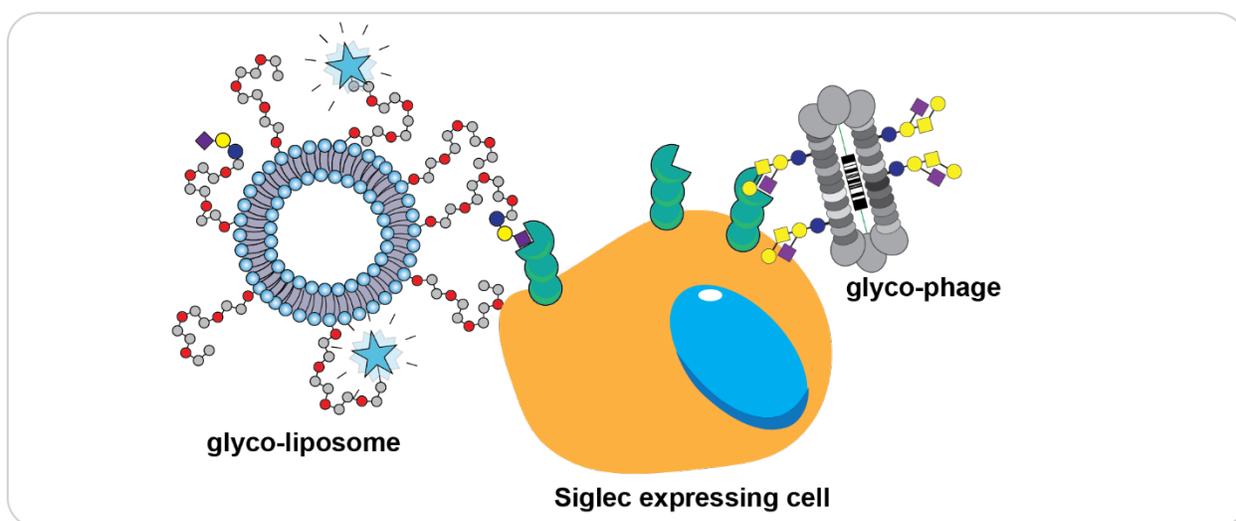


Figure 1.31: Nanoparticle-based arrays for discovering Siglec ligands. Nanoparticles displaying glycans such as liposomes or phages can be used to probe binding to Siglec expressing cells. Typically binding of liposomes and phages to Siglec expressing cells is reported using fluorescence and DNA sequencing respectively.

1.6.5 Mass spectrometry-based glycan arrays

Lastly Siglec-ligand interactions can be discovered through mass spectrometry-based approaches (**Figure 1.32**)^{37, 39, 79, 128}. One of the biggest advantages of mass-spectrometry-based approaches is that affinity can be directly measured in a truly label free manner. Recently, Concentration-Independent (COIN)- native mass spectrometry was developed allowing for the

determination of disassociation constants between lectins and their glycans independent of the glycan concentration. This strategy can be coupled with catch and release native mass spectrometry to identify and absolutely quantify the binding between a lectin and a glycan¹⁵⁸. The disadvantage of this approach is that regiochemistry of glycans cannot be assigned and glycans that share a molecular weight cannot be distinguished from each other. Moreover, as Siglecs are glycoproteins themselves, the heterogeneity of the glycoforms can make mass spectra difficult to interpret because there are many signals for a single Siglec.

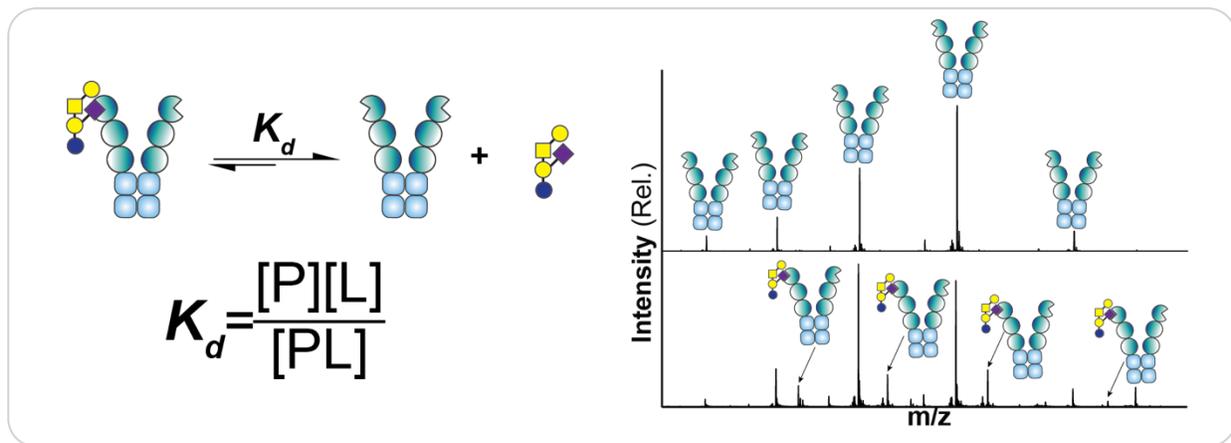


Figure 1.32: Mass spectrometry-based approaches for discovering Siglec ligands.

While there are many approaches to discover and dissect Siglec ligands, the approaches described above are the most common. Using these and approaches similar to these, there has been great success at describing the ligands of many Siglecs, however, the ligands of some Siglecs are better described than others. For some members of the human Siglec family, such as CD22, their ligands are well described (e.g. CD22 generally binds $\alpha(2\rightarrow6)$ sialic acid presented from an *N*-glycan which is installed by ST6Gal1). However, for Siglecs such as Siglec-6, Siglec-10, and Siglec-15 for example, a complete description of their ligands is lacking. Information regarding Siglecs and their ligands can be found in **Table 1.1**. Without an understanding of a Siglec's ligands, it is difficult to understand their role in a biological system. Due to the complexity of the glycocalyx and that one cell type can express multiple Siglecs, tools/approaches that enable the systematic description of Siglec ligands are greatly needed.

Table 1.1: Current state of Siglec ligands. N.D., not determine; Ac, *N*-acetylneuraminic acid; *N*-glycolylneuraminic acid; PDB, protein data bank.

Siglec	sialic acid linkage preference	key sialyltransferases(s)	high-affinity ligand developed	crystal structure (PDB ID)	preferred glycan	Neu5Ac/Neu5Gc ¹⁵⁹	enhanced by sulfation
hSiglec-1	$\alpha(2\rightarrow3)^{128, 152}$, $\alpha(2\rightarrow6)^{37}$, $\alpha(2\rightarrow8)^{128, 152}$	N.D.	Yes ¹⁶⁰	No	gangliosides ¹²⁸	Ac > Gc ¹⁶¹	No ⁷⁹
hCD22	$\alpha(2\rightarrow6)^{37, 162}$	ST6Gal1 ¹⁶³	Yes ¹⁶⁴	5VKJ	<i>N</i> -glycans ¹⁶⁵	Ac = Gc ^{161, 166}	6-O-GlcNAc ⁷⁹
hCD33	$\alpha(2\rightarrow3)^{37}$, $\alpha(2\rightarrow6)^{37}$	ST3Gal4 ⁸⁰	Yes ¹⁶⁷	No	O-glycans [§] , <i>N</i> -glycans [§]	Gc > Ac ^{137, 161, 166}	6-O-Gal ^{79, 80} > 6-O-GlcNAc ⁷⁹
hSiglec-4	$\alpha(2\rightarrow3)^{128}$, $\alpha(2\rightarrow6)^{80}$, $\alpha(2\rightarrow8)^{128}$	ST3Gal2 ⁸⁰	Yes ¹⁶⁸	No	gangliosides ¹²⁸	N.D.	N.D.
Siglec-5/14	$\alpha(2\rightarrow3)^{128, 141}$, $\alpha(2\rightarrow8)^{128, 141}$	N.D.	No	No	O-glycans, gangliosides ^{128, 141}	Gc > Ac ^{137, 166}	6-O-Gal ⁷⁹
Siglec-6	$\alpha(2\rightarrow3)^{128}$, $\alpha(2\rightarrow6)^{169}$, $\alpha(2\rightarrow8)^{128}$	N.D.	No	No	gangliosides ¹²⁸	Gc > Ac ¹⁶⁶	N.D.
Siglec-7	$\alpha(2\rightarrow3)^{128, 141}$, $\alpha(2\rightarrow6)^{170, 171}$, $\alpha(2\rightarrow8)^{128, 141}$	ST3Gal1 ^{80, 172/2⁸⁰} ST6GalINAc2 ^{80/3^{80/4⁸⁰}} , ¹⁷¹	Yes ¹³²	2HRL, 1O7V, 1NKO	O-glycans, gangliosides ^{128, 141}	Ac = Gc ¹⁷³	6-O-Gal ^{79, 80} > 6-O-GlcNAc ^{79, 80}
Siglec-8	$\alpha(2\rightarrow3)^{79, 95, 140, 174}$	ST3Gal4 [§]	Yes ¹⁷⁵	7QU6, 2N7A	O-glycans, <i>N</i> -glycans	Gc > Ac	6-O-Gal ^{79, 80, 174} 6-O-GlcNAc ⁷⁹
Siglec-9	$\alpha(2\rightarrow3)^{128, 140}$, $\alpha(2\rightarrow6)^{132}$	ST3Gal1 ^{172/4^{80/6⁸⁰}} ST3Gal4 ⁸⁰	Yes ¹⁷⁶	No	gangliosides ^{128, 141}	Ac > Gc ^{137, 173}	6-O-GlcNAc ⁷⁹
Siglec-10	$\alpha(2\rightarrow3)^{128, 177}$	N.D.	Yes ¹⁷⁶	No	<i>N</i> -glycans ¹⁷⁸ gangliosides ^{128, 177}	Gc > Ac ¹⁵⁹	N.D.
Siglec-11/16	poly $\alpha(2\rightarrow8)^{179}$	N.D.	No	No	poly sialylated glycans	Gc > Ac ^{180, 181}	N.D.
Siglec-15	$\alpha(2\rightarrow3)^{79, 80}$	ST6GalINAc1 ⁸⁰ ST3Gal4 ⁸⁰	Yes ¹⁸²	7ZOZ	O-glycans ⁸⁰ ,	Ac > Gc ¹⁵⁹	6-O-Gal ^{79, 80}

1.7: Thesis statement and objectives

The development of any therapeutic begins by understanding the underlying molecular mechanisms of the pathology and Siglecs are front and center in many pathologies. Viruses including HIV, Ebola, and SARS-CoV-2^{9, 183, 184, 185, 186}, use glycans to exploit the endocytic nature of Siglecs to infect host cells. Cancerous cells are known to hyper-sialylate which may help them avoid destruction by the immune system through exploiting Siglecs' ability to act as an immunological break⁹⁵. Dysregulation of the ability of Siglecs to act as an immunological break is thought to exacerbate autoimmune diseases such as Guillain Barré Syndrome among others¹⁷⁷. Siglecs also play a role in the progression of neurodegenerative diseases such as Alzheimer's Disease, Parkinson's Disease and Huntington's Disease¹⁸⁷. Moreover, Siglecs are thought to play a role in conception and healthy pregnancies^{5, 188, 189}. Understanding the physiological or pathophysiological role of a Siglec begins with describing the ligands of a Siglec as all Siglec functions are initiated by the binding of a glycan. While great progress has been made in describing Siglec ligands, many gaps remain (**Table 1.1**). There are many factors that complicate Siglec-ligand discovery, two of the largest obstacles are diversity of the glycans in the glycocalyx and that cells that express Siglecs almost always express multiple Siglecs; consequently, deconvoluting Siglec-ligand interactions is incredibly difficult. If approaches which enable the ligands of Siglecs to be systematically described are developed, then the physiological and pathophysiological roles of Siglecs can be better understood. Insights from these approaches can then be applied to skew physiological mechanisms around the Siglec-sialic axis for therapeutic benefit (**Figure 1.33**). This thesis aims to develop new approaches and improve previously developed approaches (**section 1.6**) to systematically study Siglec-glycan interactions and apply discoveries made through these approaches to better our understanding of the physiological and pathophysiological roles of Siglecs towards developing Siglec/sialoside based therapeutics.

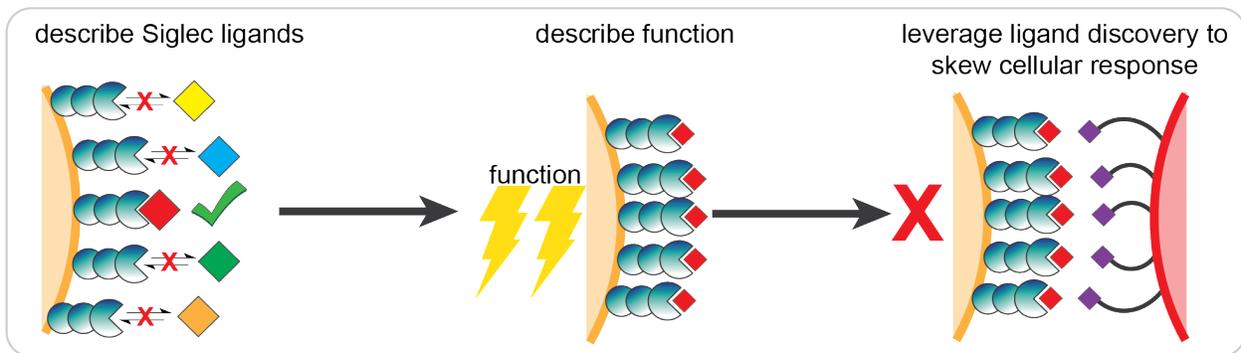


Figure 1.33: Thesis statement visual abstract. Siglecs can be modulated for therapeutic benefits if their ligands are understood.

1.7.1: Chapter 2: The versatility of Siglec-Fcs

In Chapter 2, the versatility of a novel Siglec-Fc chimera platform was demonstrated by its use in many distinct approaches for studying Siglec ligands which were used in numerous studies to describe Siglec ligands. Advancements in production and storage of Siglec-Fcs was optimized and used to discover Siglec ligands on primary cells and to dissect the ligands of various Siglecs including CD33 and Siglec-8 (**Figure 1.Error! Bookmark not defined.**).

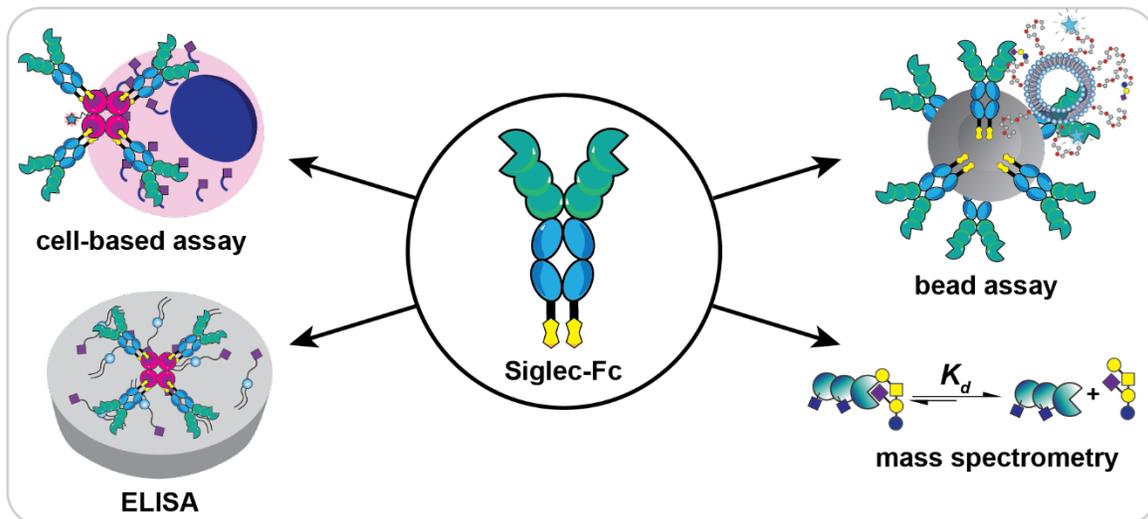


Figure 1.Error! Bookmark not defined.: Visual abstract of Chapter 2. Siglec-Fcs chimeras are versatile tools which can be used in a variety of approaches to describe Siglec ligands.

1.7.2: Chapter 3: Liposomes as tools to study Siglec ligands

In Chapter 3, a novel liposomal nanoparticle formulation which, was optimized to study Siglec-ganglioside interactions was developed. Using this optimized liposome formulation, the entire human and mouse Siglec family was interrogated against a panel of nine gangliosides. Many novel Siglec-ganglioside interactions were discovered most notably the ability of Siglec-6

to bind gangliosides independent of its conserved arginine residue. Moreover, the ganglioside binding profiles between human and mouse Siglecs were compared and the appropriateness of mouse models to study human Siglec-ganglioside interactions was evaluated (**Figure 1.Error!** Bookmark not defined.).

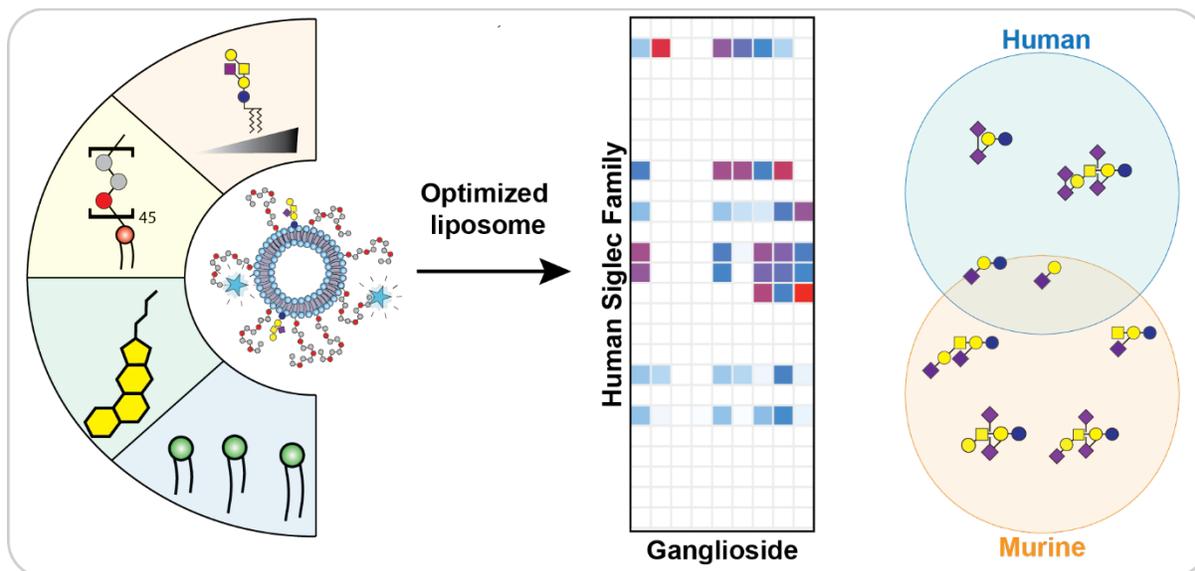


Figure 1.Error! Bookmark not defined.: **Visual abstract for Chapter 3.** Optimization of a liposomal formulation to dissect Siglec-ganglioside interactions.

1.7.3: Chapter 4: Exploring the ligands of Siglec-6

In Chapter 4, the ability of Siglec-6 to bind gangliosides was investigated using a panel of synthetic glycolipids and various genetic approaches discovering that a solvent exposed tryptophan residue is critical for ganglioside recognition by Siglec-6 and that ganglioside recognition by Siglec-6 is mostly driven by bilayer presentation. Leveraging the knowledge of glycolipid ligands of Siglec-6, targeting of physiological Siglec-6 was possible on primary human memory B cells, mast cells, and placental syncytiotrophoblasts. Moreover, a possible novel biological function of Siglec-6 was discovered in its ability to mediate the uptake of extracellular vesicles (**Figure 1.34**).

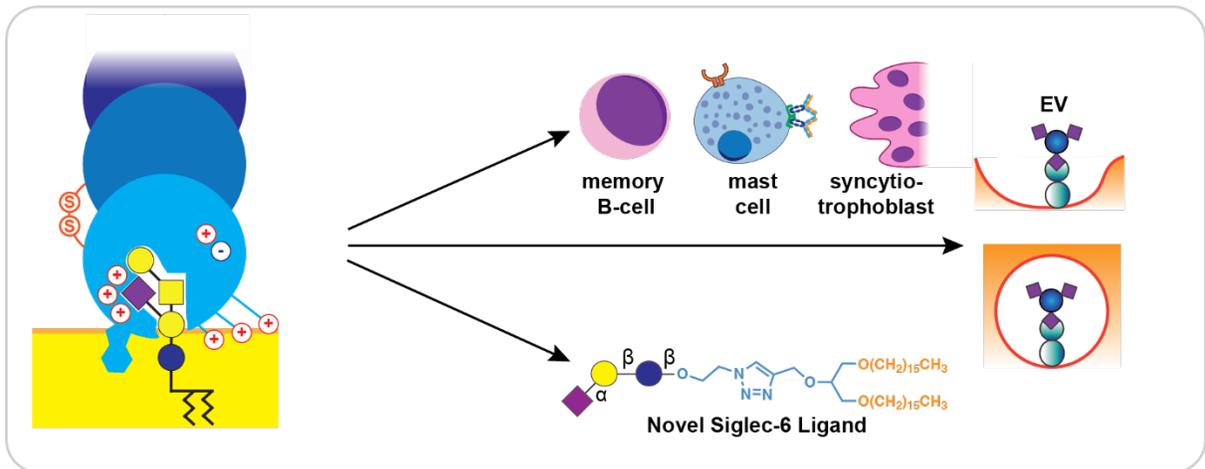


Figure 1.34: Visual abstract for Chapter 4. Liposomes and neoglycolipids were used together to understand the ligands and probe the biological roles of Siglec-6.

Chapter 2: Siglec-Fcs as Versatile Tools to Study Siglec

Ligands

2.1: Acknowledgements

Portions of this chapter were published as: Schmidt, E.N.; Jung, J.; Macauley, M.S.; Flow cytometry-based detection of Siglec ligands. *Carbohydrate-Protein Interactions: Methods and Protocols*, Springer US (2023).

Thanks to Dr. Jaesoo Jung for helping to develop the Siglec-Fc producing CHO cells. I would also like to thank Dr. Jaesoo Jung and Christopher D. St. Laurent for developing lentivirally transduced and CRISPR/Cas9 engineered cells. I would like to thank Zeinab Jame-Chenarboo for her assistance in performing the experiments on primary human splenocytes. I would like to thank Prof. Lori J. West for providing human splenocytes.

I would also like to acknowledge all those who donated tissue. Much of this work would be impossible without them.

2.2: Introduction

2.2.1: Importance of describing Siglec ligands

The cell surface is decorated in a dense array of complex carbohydrates which play roles in biology and are particularly important in cell-to-cell communication^{125, 190, 191}. Siglecs are an immunoglobulin superfamily of type-I membrane proteins found in all mammals with fifteen distinct Siglecs in humans⁶. Binding of Siglecs to their sialic acid-containing glycan ligands modulates immune responses by regulating the spatial proximity of Siglecs to other immune activity receptors¹²⁵. Under the appropriate physiological context, recruitment of Siglecs to activity receptors can modulate immune cell signaling pathways and the immunomodulatory properties of Siglecs can be exploited by pathogens and cancers. Indeed, both bacteria¹⁹² and viruses¹⁹³ can exploit Siglecs in order to skew the host immune system, and cancerous cells coat themselves with elevated levels of sialosides to exploit Siglecs^{95, 194, 195, 196}.

Despite their significance in health and disease, there is an incomplete description of their ligands. Siglec–ligand interactions can be broken down into two categories. When the ligand and the Siglec are on different cells or particles, it is referred to as a *trans* interaction whereas if the Siglec and the ligand are on the same cell or particle, it is referred to as a *cis* interaction¹⁹⁷. Siglecs recognize sialylated glycoconjugates (sialosides) through a conserved arginine in their V-set domain and there are three common sialic acid motifs: $\alpha(2\rightarrow3)$, $\alpha(2\rightarrow6)$ linked sialic acid linked to a galactose or *N*-acetylgalactosamine residue, and $\alpha(2\rightarrow8)$ linked sialic acid to another sialic acid residue. As Siglecs have relatively low affinities for their ligands (K_d in the low mM range), approaches that leverage avidity are needed to study them¹⁰⁰. For instance, Siglec-Fcs are often complexed with secondaries such as anti-IgG antibodies or Strep-Tactin to increase avidity.

2.2.2: Siglec-Fc chimeric proteins

Siglecs are naturally membrane bound proteins which makes studying them *in vitro* challenging as studying membrane proteins in solution will often cause them to misfold and/or precipitate. The most common scaffold to use as the Siglec probe is a soluble recombinant version

of a Siglec in which 2-3 extracellular domains of the Siglec are fused to the constant region of an IgG antibody, described as a Siglec-Fc chimera, creating a probe with two binding sites (**Figure 2.1**). This construct makes the Siglec soluble as well as increases avidity (dimer). Moreover, the Siglec-Fc can be complexed with anti-IgG to form a tetramer further increasing avidity. Generally, Siglec-Fcs are produced recombinantly in mammalian cells such as Chinese hamster ovary (CHO) cells because Siglecs are glycoproteins themselves and *E. coli* lack the biosynthetic equipment to install the glycans which are required for proper Siglec folding.

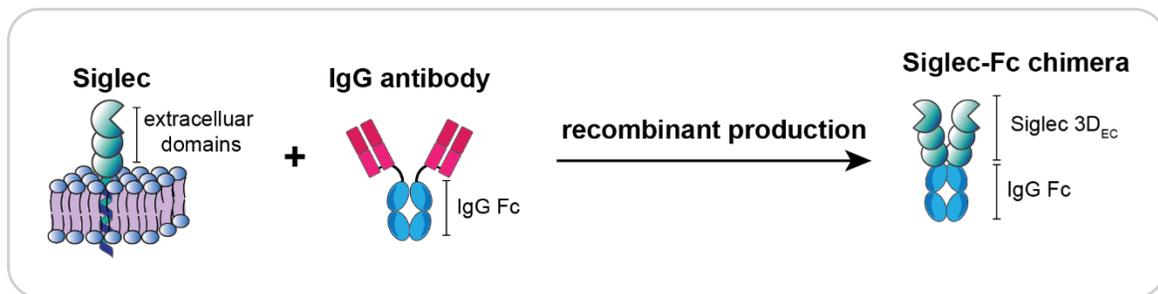


Figure 2.1: Siglec-Fc chimera concept. The extracellular domains of a Siglec, typically the first two or three extracellular domains including the V-set domain are combined with the constant domain of an IgG antibody and recombinantly produced as a Siglec-Fc chimera.

Siglec-Fc chimeras have been successfully deployed as tools for probing Siglec ligands for over three decades. Their use began in the early 1990's for immunoprecipitations^{198, 199}, which helped identify CD22 as being specific for $\alpha(2\rightarrow6)$ sialosides^{117, 199, 200}. The use of Siglec-Fc chimeras as tools to study Siglec ligands by flow cytometry accelerated in the early 2000's^{201, 202, 203, 204, 205}. Many recent studies continue to make fundamental insights into Siglec specificity using flow cytometry^{37, 79, 80, 206, 207, 208}.

2.2.3: Challenges with the previous versions of the Siglec-Fc

Siglec-Fcs became popular in the early 1990's with Stamenkovic *et al.* being one of the first teams to use a Siglec-Fc and Siglec-Fcs have been popular tools to describe Siglec ligands ever since¹⁹⁸. Despite their utility, traditional Siglec-Fc proteins have several drawbacks that may not be fully appreciated. The first drawback is that the Fc has the potential to engage Fc γ receptors. A second drawback of traditional Siglec-Fc chimeras is the need for an anti-human IgG secondary antibody for detection; while pre-complexing with a secondary can greatly enhance

sensitivity, it can potentially further exacerbate binding to Fc γ receptors and recognize human IgG coated to human cells and tissues²⁰⁹. A third underappreciated drawback, but known well to researchers in the field, is the instability of commercially supplied Siglec-Fc stored at 4 °C. To address the first two concerns, other scaffolds have emerged. In 2012, Gieseke *et al.* used a different approach where biotinylated monomeric Siglec fragments were complexed to streptavidin, effectively generating a tetrameric presentation while avoiding display of the Fc²¹⁰. Another alternative to the Siglec-Fc which, was developed by Gonzalez-Gil *et al.* was to fuse the extracellular domains of a Siglec with a cartilage oligomeric matrix proteins (COMP) to form a pentameric complex. These probes have had success in discovering Siglec ligands on glycan microarrays¹³⁹. The most recent update to the Siglec-Fc chimera was in 2020 when our laboratory developed a new version of the Siglec-Fc chimera with subtle changes that improves Siglec-Fcs versatility (Figure 2.2)³⁷.

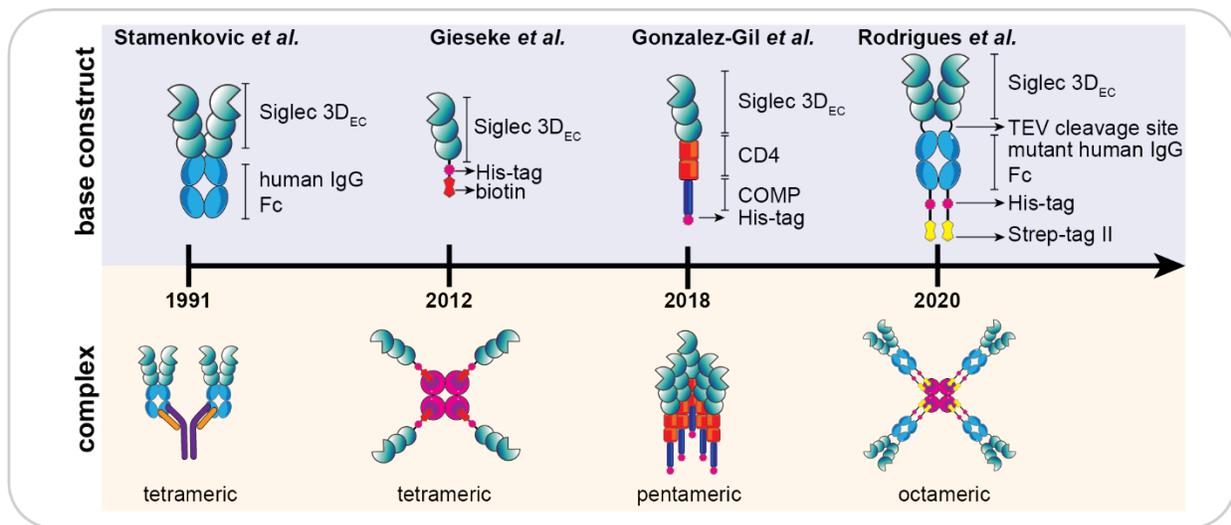


Figure 2.2: Key contributions to the development of Siglec-Fc chimera over time.

To circumvent some of the aforementioned limitations of traditional Siglec-Fc chimeras, our laboratory recently developed a new library of Siglec-Fc chimeras expressed from CHO cells, with modifications that make them versatile tools for use in flow cytometry as well as other approaches (immunohistochemistry, glycan microarrays, ELISA, *etc.*)³⁷. First, mutations were introduced into the Fc region to abolish their undesired association to the Fc receptor but still

allowing for anti-human IgG to be used as a secondary²¹¹. Second, the addition of a C-terminal His₆-tag allows for facile purification of the Siglec-Fc under mild conditions compared to Protein G purification that requires strong acidic elution, potentially causing the Siglec-Fc to precipitate²¹². Our constructs also contain a Strep Tag II²¹³, which can be used in parallel with the His₆-tag for two dimensional purification. Moreover, the Strep Tag II serves as a genetically encoded handle for building up a highly multivalent complex with Strep-Tactin (similar to the probe developed by Gieseke *et al.*), yielding a probe with up to eight glycan binding sites. To ensure the most appropriate control is used, like an isotype control in standard flow cytometry, we additionally created a version of each Siglec-Fc wherein the essential arginine, which forms a critical ionic interaction with sialosides, is mutated to an alanine. This new generation of Siglec-Fc constructs are reliable tools for studying Siglec ligands. The use of Strep-Tactin as a reporter has the additional advantage of avoiding off target binding of the secondary to IgG coated human cells and tissues. Additionally, our lab demonstrated that this version of the Siglec-Fc can be lyophilized which improves the shelf-life of these reagents. This new iteration of the Siglec-Fc construct was used to further describe Siglec ligands in a variety of approaches towards a better understanding of Siglec ligands to learn more about their physiological and pathophysiological roles.

2.2.4: Flow cytometry

Flow cytometry is a powerful analytical approach to study cells. As its name suggests, flow cytometers leverage fluid dynamics *-flow-* to analyze many cells *-cyto-* at a time. Flow cytometry is very similar to immunofluorescence microscopy as both approaches typically use fluorescently labeled antibodies to stain cells however flow cytometry allows for the screening of thousands of cells per second resulting in robust relative quantification of a fluorescent signal. For reference, a typical flow cytometry histogram in this work represents at least 10,000 cells or *events*. In a traditional flow cytometer, a bulk solution of cells is pushed through a nozzle where it is met by a relatively quickly moving sheath fluid which results in the cells falling into a single file line due to hydrodynamic focusing. The cells then pass one at a time through a channel that has many lasers

and detectors that are used to measure the light scattering and fluorescence emission intensity of each cell (**Figure 2.3a**). Once the cells have been analyzed, they are transferred to a waste reservoir.

Following processing by the flow cytometer, the data needs to be analyzed. There are two common plots used for the analysis of flow data. The first is a two-dimensional plot where a different fluorescence intensity is plotted on the x- and y-axis. This ideally produces a plot like **Figure 2.3b** which can be divided into quadrants. Each cell is then given a description based on which quadrant it lands in (e.g. **Q1**, x^-,y^- ; **Q2**, x^-,y^+ ; **Q3**, x^+,y^+ ; **Q4**, x^+,y^-). A *gate* can then be used to isolate a certain population of cells for further analysis by a similar approach. Another common way to present flow cytometry data is by using a histogram to show the fluorescence intensity of a given parameter (e.g. Siglec-Fc binding to a given population of cells) on the x-axis and the number of cells on the y-axis. Histograms from different populations or staining conditions can then be overlaid on each other to compare the fluorescence intensities (**Figure 2.3c**). Fluorescence intensity is often quantified by taking the median or mean fluorescence intensity (MFI or mFI) of a population. While similar in many respects, an analytical flow cytometer differs from fluorescence-assisted cell sorter (FACS) with respect to the fate of the analyzed cells. When cells are processed with a FACS, the cells are then captured in a new flask to be used for further experiments whereas cells are transferred to a waste reservoir post analysis in an analytical flow cytometer.

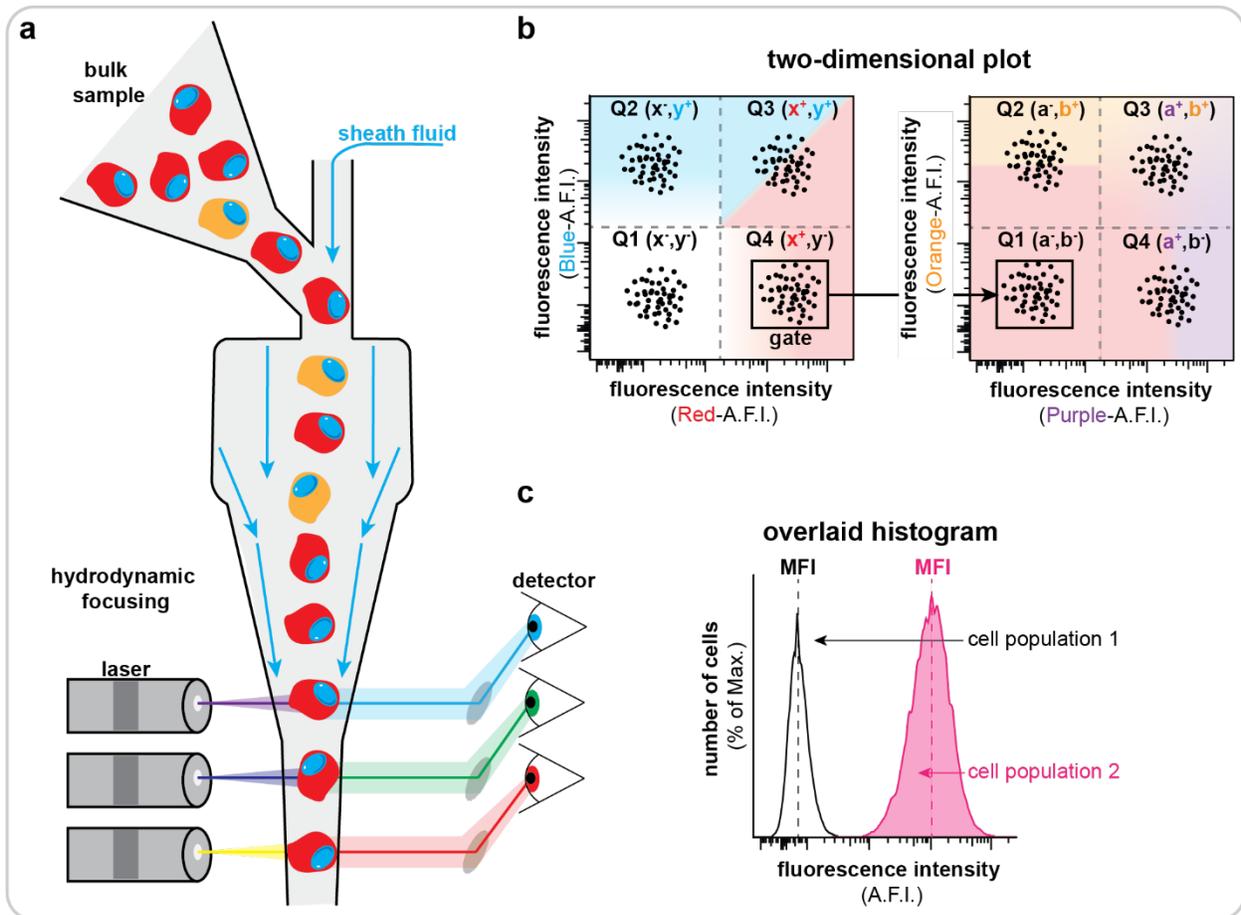


Figure 2.3: Flow cytometry as an approach to discover Siglec ligands. **a**, Schematic of an analytical flow cytometer. Cells from a suspension are focused using quickly moving sheath fluid which causes the cells to form a single file line due to hydrodynamic focusing. Once the cells are aligned, the cells pass through a chamber with many lasers and detectors. The amount of light scattered or emitted by fluorescently labeled cells is then measured. **b**, Representative two dimensional flow plots used to isolate specific cells from a heterogenous population. Cells can be classified as a double negative Q1 ($x-,y-$), single positive Q1 ($x+,y-$) or Q4 ($x-,y+$), or a double positive Q3 ($x+,y+$). Using this strategy, any of these populations of cells can then be isolated with a gate and examined by another two parameters. This can be repeated as necessary. **c**, Flow cytometry data can also be represented as a histogram which is used to visualize the variation in a population of cells with respect to a single parameter. Often the histograms from two different populations are overlaid in order to make a comparison in a given parameter between two populations. Flow cytometry is a relatively quantitative method and is excellent at measuring if there is a difference between two samples however it is not trivial to relate an arbitrary fluorescence intensity (AFI) to objective output such as number of copies of proteins on a cell or number of ligands etc.

2.2.5: Siglec-Fc protein production

Siglec-Fcs are generally produced in mammalian cells. Bacteria such as *Escherichia coli* lack the biosynthetic equipment to properly glycosylate the Siglec-Fcs which often causes them to misfold and Siglec-Fcs produced from bacteria tend to have very low yields if any soluble protein is produced at all. However smaller fragments of Siglecs have been produced from *E. coli* (PDB: 7QU6).

2.2.6: Siglec-Fc expression

We chose to use Chinese Hamster (*Cricetulus griseus*) Ovary (CHO) cells, as they are known to produce large quantities of recombinant protein and have previously been engineered to have the *Flp-In System* allowing them to be easily stably transfected²¹⁴. A stable transfection is different from a transient transfection as when cells are stably transfected, the recombinant DNA becomes part of the cells chromosomal DNA and, therefore, all the daughter cells will also have the recombinant gene; whereas, in a transient transfection, the recombinant DNA gets diluted with each generation however the cells do not need to be engineered with a system such as the Flp-In system²¹⁵. Flp-In cells have been engineered to have a Flp-In recombinase site in one of their chromosomes. However, when the Flp-In recombinase sites are introduced, it is added to a random location in the genome but once successfully introduced into the chromosome, all the progeny of that cell will have the recombinase site in the same location. The addition of recombinant DNA to a cell using the Flp-In system requires a helper plasmid, containing the DNA recombinase, in addition to a plasmid containing the recombinant DNA. To stably transfect Flp-In CHO cells, lipofectamine is used as a transfection reagent to shuttle the plasmid containing the Siglec-Fc DNA (pCDNA5) and the helper plasmid (pOG44) into the cells (**Figure 2.4a**)²¹⁶. Once inside the cell, the Flp recombinase is translated from pOG44 then acts on the Flp sites in pCDNA5 and the chromosome of the CHO cell and the desired recombinant DNA as well as a hygromycin resistance gene are added to the chromosome (**Figure 2.4b**). The cells are then

cultured in the presence of hygromycin B which selects for cells that successfully underwent transfection (**Figure 2.4c**).

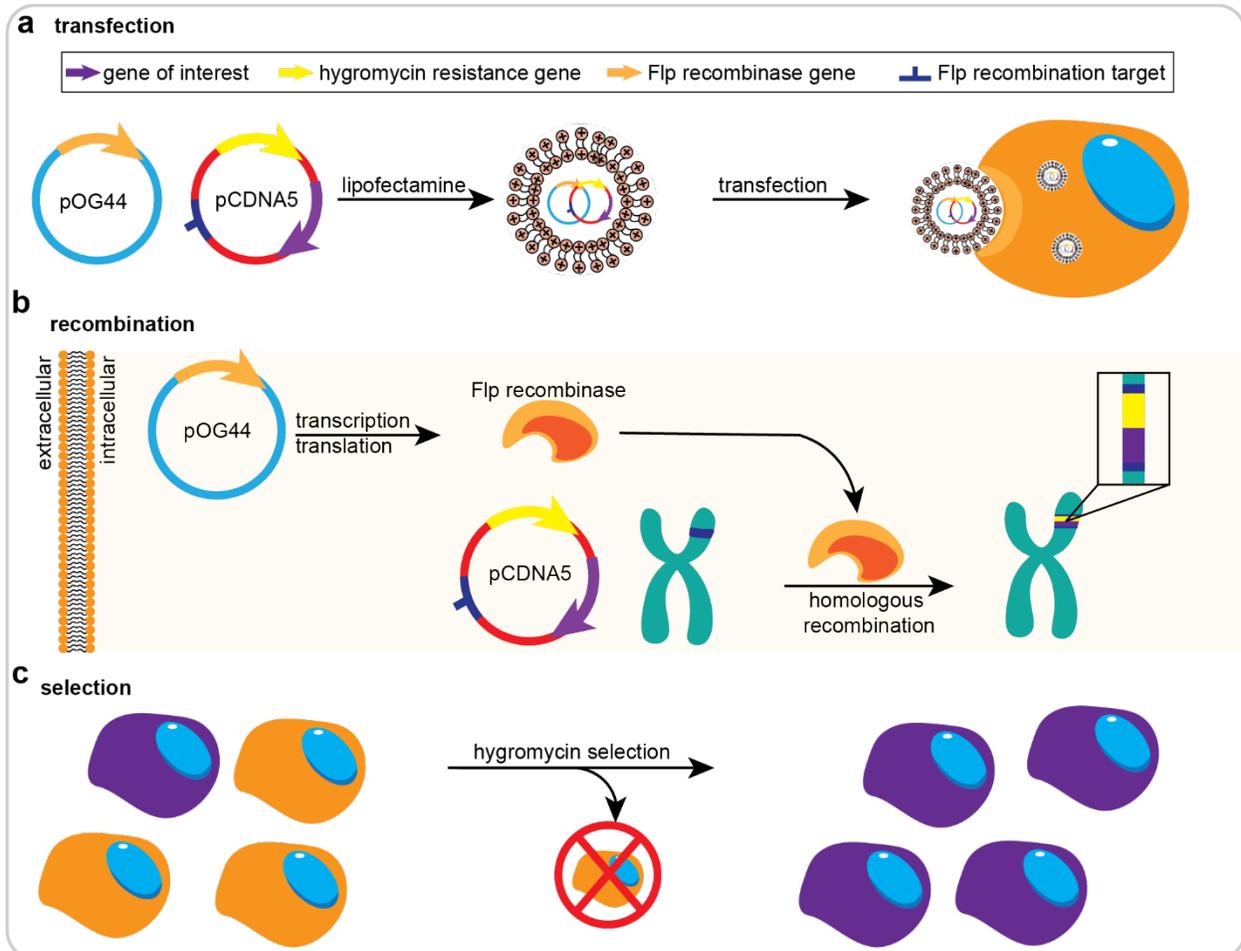


Figure 2.4: Transfection of Flp-In CHO cells to stably express Siglec-Fcs. **a**, After the Siglec-Fc gene is cloned into pCDNA5, pCDNA5 and pOG44 are co-transfected into CHO Flp-In cells using lipofectamine. **b**, pOG44 contains the gene for the Flp recombinase which gets transcribed and translated to produce the Flp recombinase which catalyses the addition of the gene of interest and a hygromycin resistance gene into the chromosome of the CHO Flp-In cell resulting in a stable transfection. **c**, The cells then undergo selection as cells that have successfully undergone recombination with pCDNA5 will survive in the presence of hygromycin B.

2.2.7: Siglec-Fc purification

Once the Siglec-Fc DNA has been successfully added to the CHO cells, the cells will start producing Siglec-Fcs. The Siglec-Fcs will be released to the media as the Siglec-Fcs still contain the signal peptide at the start of the Siglec which directs the Siglec-Fc to the cell surface. However, as the transmembrane domain is lacking from the Siglec, it is not anchored to the vesicle membrane and when the vesicle from the Golgi apparatus fuses with the cellular membrane the

Siglec-Fc is freed from the cell²¹⁷. The Siglec-Fc can then easily be separated from the cells by removing the media from the cells. After the cultured media is isolated from the cells, the Siglec-Fcs can be purified by a series of affinity chromatography techniques²¹⁸. Siglec-Fcs are not very stable in solution for long periods of time (days-weeks), so to improve the shelf-life of the Siglec-Fcs, they are lyophilized which allows them to be stored under ambient conditions and are stable for longer (years) periods of time (**Figure 2.5**).

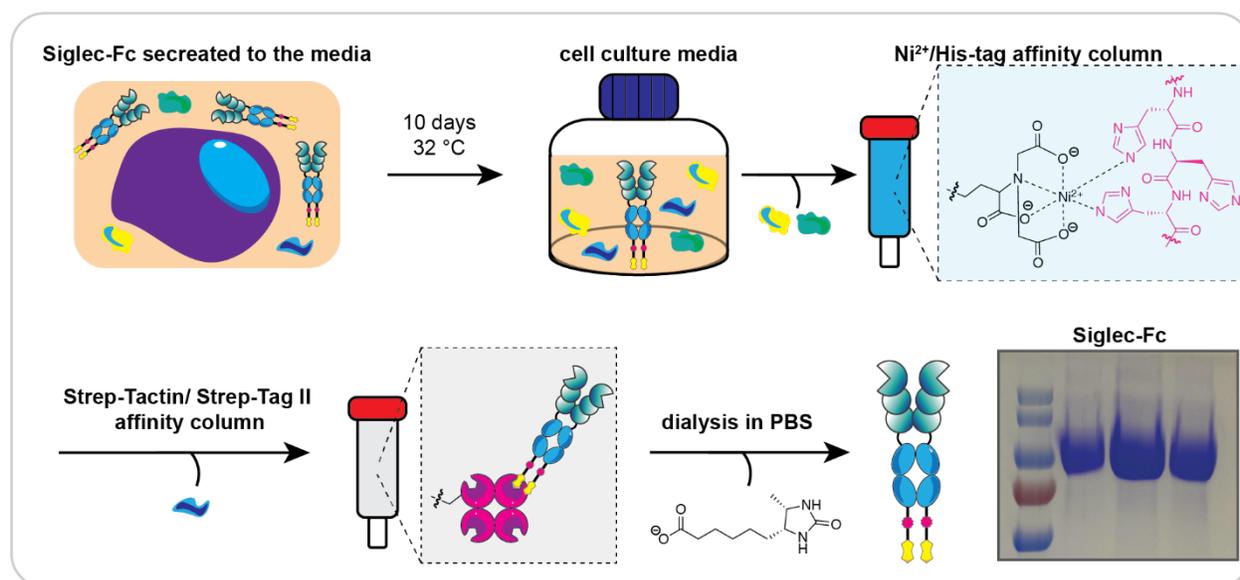


Figure 2.5: Workflow for the expression and purification of Siglec-Fcs from stably transfected CHO Flp-In cells. Siglec-Fc expressing CHO cells are incubated in growth media at 32 °C for ten days post confluence. The media is then isolated from the cells and purification begins using a Ni²⁺ affinity column. Following this, the Siglec-Fc is further purified using a Strep Tactin column. The pure Siglec-Fc is then dialyzed against phosphate buffer saline (PBS), concentrated, and lyophilized for storage.

2.2.8: Approaches for modulating cellular glycans

One of the strengths of a cell-based glycan array is that the glycans are presented in a very biologically relevant way which, makes this method a great approach to describe Siglec ligands. However, due to the heterogeneity of the glycoalyx, it can be very difficult to identify which glyco-epitopes the Siglec is engaging with. To better understand which glycans are the ligands for a Siglec, the glycans of the cell can be modified through many different means. In this section, common approaches for modulating cell glycans will be discussed.

2.2.8.1: Enzymatic digest of cellular glycans

A common approach for modulating Siglec ligands on cells is to treat the cells with a glycan modifying enzyme such as a neuraminidase (also known as a sialidase)^{37, 219, 220}. Neuraminidases/sialidases are glycosidases which cleave the linkage between the sialic acid and the underlying galactose or GalNAc and used by many different organisms but are perhaps most famously used by viruses such as Influenza²²¹. There are many different neuraminidases which have different specificities with respect to the type of sialoside they cleave. For instance, neuraminidase A (NeuA) from *Arthrobacter ureafaciens* cleaves all sialic acid linkages whereas neuraminidase S (NeuS) from *Streptococcus pneumoniae* only cleaves $\alpha(2\rightarrow3)$ linked sialosides³⁷. Using these enzymes in parallel can have synergistic effects when describing Siglec ligands. First, Siglec ligands must be present on the cells of interest under baseline conditions (**Figure 2.6a**). The cells can then be treated with NeuA, and if there is a decrease in Siglec-Fc binding, it is evidence that Siglec-Fc binding is sialic acid dependent (**Figure 2.6b**). Following this, cells can then be treated with NeuS, and if there is a decrease in Siglec-Fc binding, it suggest that $\alpha(2\rightarrow3)$ linked sialosides contribute to Siglec binding and if binding is not decrease (**Figure 2.6c**), it is indirect evidence that $\alpha(2\rightarrow6)$ or $\alpha(2\rightarrow8)$ sialosides are ligands for the Siglec. Neuraminidases are just one example of glycan modifying enzymes which can be used to describe Siglec ligands on cells in this way. Other enzymes such as mucinases can also be used in a similar manner²²⁰.

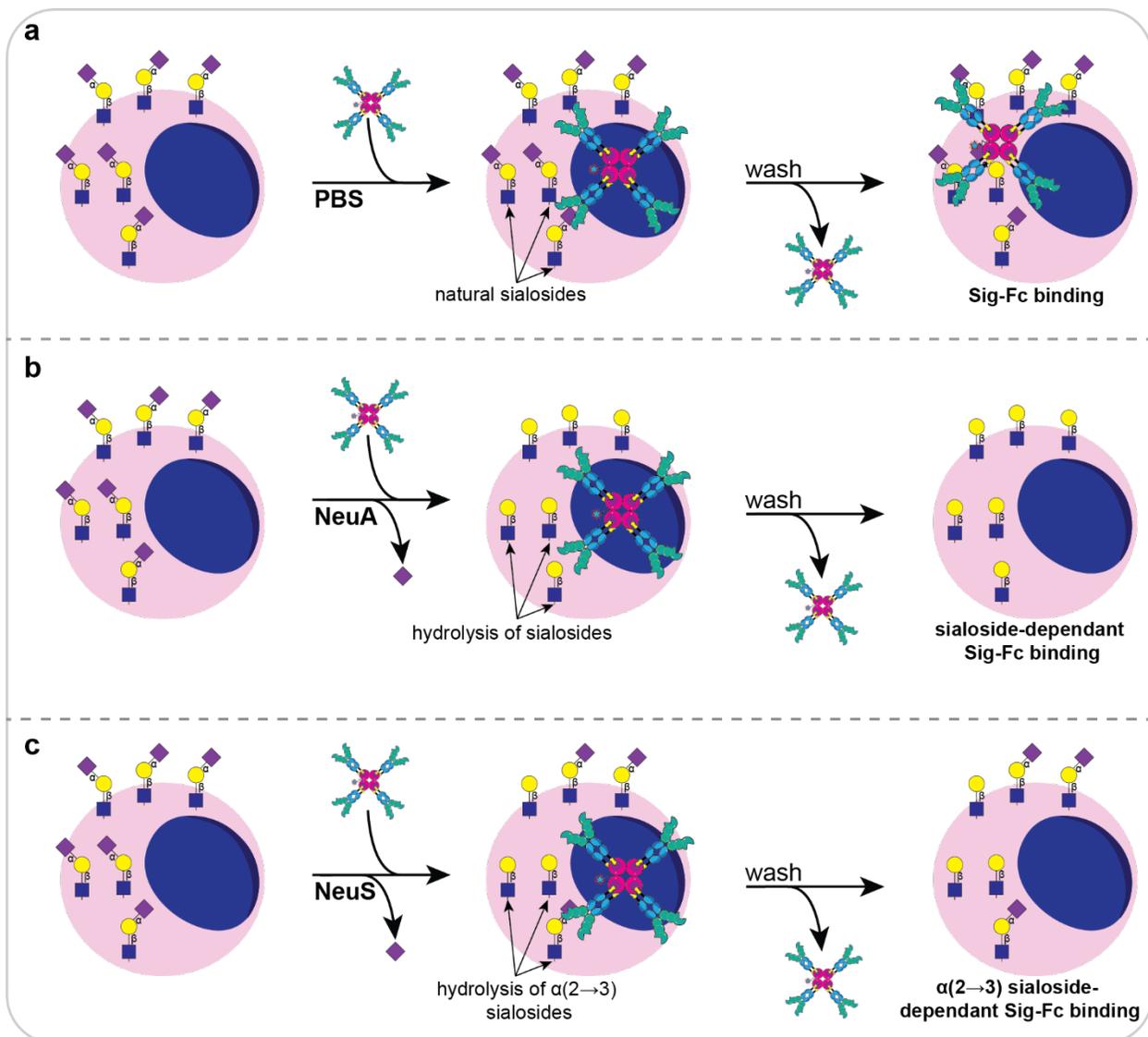


Figure 2.6: Systemic profiling of cells treated with glycan modifying enzymes to more specifically describe Siglec ligands. a, b, c, Staining of cells in a cell-based glycan array with soluble Siglec-Fc without enzymatic treatment and treatment with neuraminidase A and neuraminidase S respectively.

2.2.8.2: Introduction of exogenous glycan modifying enzymes

There are many different strategies for introducing an exogenous gene into a cell. For instance, the Flp-In system is a simple and facile approach to add a gene of interest into the chromosome of a cell. However, this approach requires the cells to be previously engineered to have the Flp-In recombinase site, which can be a cumbersome process limiting the cell types that can be engineered with the Flp-In system. Lentiviral transduction overcomes this problem by using a non-replicating virus to stably transfect any given cell. Lentiviral transduction begins by triple transfecting human embryonic kidney 293T (HEK293T) cells with three different plasmids

(RP172-transfer plasmid, gene of interest, zeocin resistance, fluorescent protein (mAmetrine) to con; RP18-packaging vector; RP19-envelope vector-**Figure 2.7a**). Following the transfection, the HEK293T cells produce and release non-replicating virus into the media. The media is then collected, and the virus is isolated by a series of centrifugations. Once the virus is isolated, it is then added to the cell line of interest. Cells which successfully underwent viral transduction will be resistant to zeocin and will express the fluorescent protein mAmetrine. Cells that survive zeocin selection are then checked for mAmetrine signal as well as for the protein of interest (**Figure 2.7b**). Once it is confirmed that the gene of interest was successfully introduced to the cell line of interest, the cells are ready for further experiments. Using this strategy, various glycan modifying enzymes can be exogenously introduced to a cell of interest and its effect on Siglec binding can be probed.

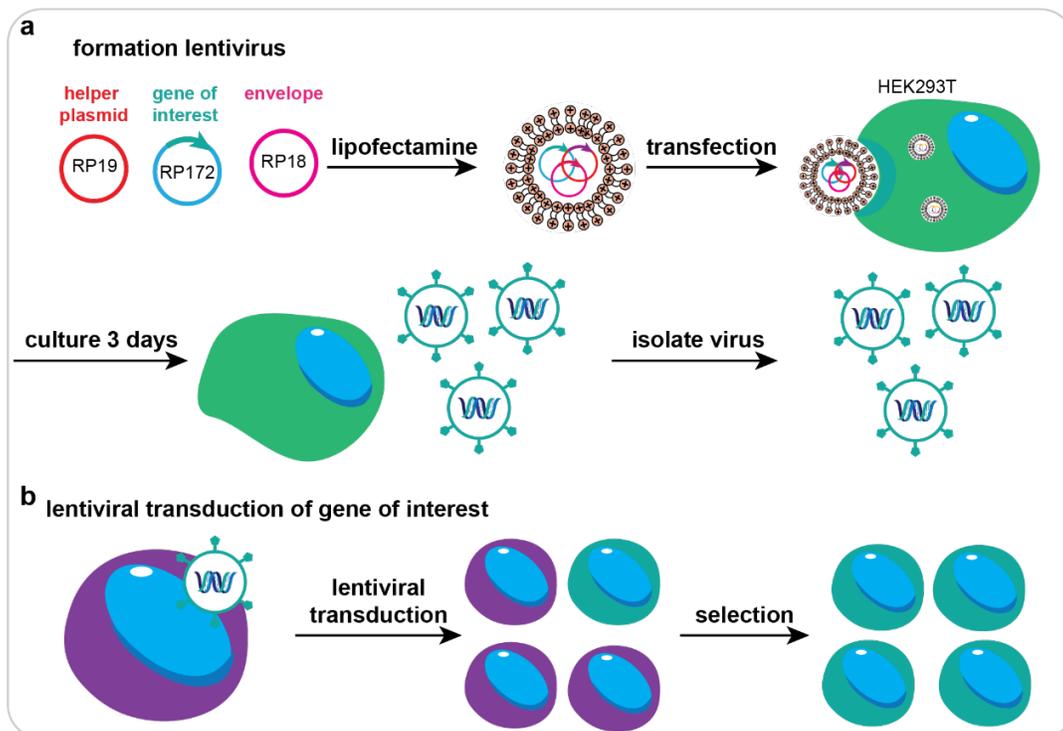


Figure 2.7: Expression of an exogenous gene using lentiviral transduction. a, Formation of lentivirus used for transduction. Three plasmids are transfected into HEK293T cells: 1, RP172 which is the vector for the gene of interest as well as mAmetrine- a marker for successful viral infection; RP19 which encodes the vesicular stomatitis virus G glycoprotein (VSVG) envelope protein; 3, RP19 which encodes for the other proteins that are needed to form a virion. These plasmids are transfected into HEK293T cells which produce viral particles that are then isolated and will be used to transduce the gene of interest into a cell line of interest. **b,** viral transduction schematic where viruses produced by HEK293T cells are used to introduce an exogenous gene to a cell line of interest.

2.2.8.3: Knock-out of glyco-genes

While introducing a new gene into a cell line to observe the effect on Siglec-Fc binding is a very fruitful approach to describe Siglec ligands, the opposite strategy of removing a gene to see the effect on Siglec-Fc binding can be just as informative. Typically, a gene is knocked-out using CRISPR/Cas9 system where a *trans*-activating RNA or trRNA and a CRISPR RNA or crRNA are complexed with the Cas9 enzyme (**Figure 2.8**). The trRNA improves the endonuclease activity of the Cas9 complex and the crRNA directs the complex to a specific location in the genome²²². The complex is transfected into the cell line of interest. The efficiency of gene editing through CRISPR is largely dependent on a cell's transfection efficiency (*e.g.* if the cell does not take up the CRISPR complex it will not undergo gene editing). To improve efficiency, the trRNA is modified with a fluorescent reporter, which enables cells that successfully underwent transfection to be selected for by fluorescence assisted cell sorting (FACS). After the cells have been sorted, they can then be screened for the absence of the gene of interest by traditional flow cytometry. Clones that display the phenotype of interest are then validated genotypically by sequencing the relevant portion of the genome.

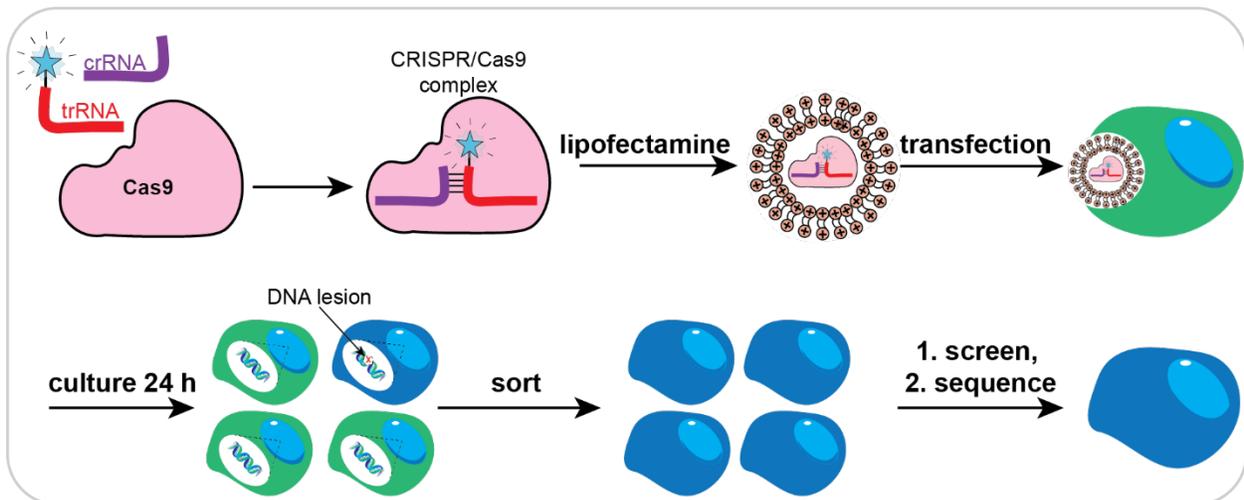


Figure 2.8: Depiction of gene knock-out by CRISPR/Cas9.

In this chapter, the versatility of the newly developed Siglec-Fc will be demonstrated by applying in a cell-based glycan array, ELISA, and bead assay. The emphasis will be studying Siglec ligands using the Siglec-Fcs in a cell-based assay on genetically/chemically modified cell lines and primary human cells. Improvements with respect to Siglec-Fc sensitivity and expression yields through the knock-out of CMAS will also be demonstrated.

2.3: Results

2.3.1: Cell-based glycan array

As Siglec-ligand interactions are relatively weak ($K_d=10^{-3-5}$ M) Siglec-Fcs are often complexed with a secondary to leverage avidity and increase sensitivity. The Siglec-Fc developed by Rodrigues *et al.*³⁷ can be complexed with either Strep-Tactin or an anti-human IgG antibody (**Figure 2.9a, b**). Strep-Tactin can offer some advantages because it can avoid interactions with Fcγ receptors expressed on cells and will not cross-react with IgG coating human primary cells and tissues.

2.3.1.1: Comparison of Strep-Tactin and anti-hIgG as a secondary

A cell-based glycan array was used to demonstrate that this novel Siglec-Fc-construct is compatible with both anti-hIgG and Strep-Tactin as a secondary. hSiglec-1 was precomplexed with fluorescently labeled Strep-Tactin, robust binding of the complex was observed to wildtype (WT) U937 cells (a monocytic human immortalized cell line) compared to just Strep Tactin alone or the corresponding hSiglec-1 arginine mutant (**Figure 2.9c**). Likewise, when Siglec-7-Fc was precomplexed with a polyclonal anti-hIgG, significant binding of the complex was observed to wildtype U937 compared to anti-hIgG alone or the corresponding arginine mutant of Siglec-7 (**Figure 2.9d**). These results demonstrate that this new version of the Siglec-Fc can be complexed with either Strep-Tactin or anti-hIgG to reliably describe Siglec ligands in a cell-based glycan array.

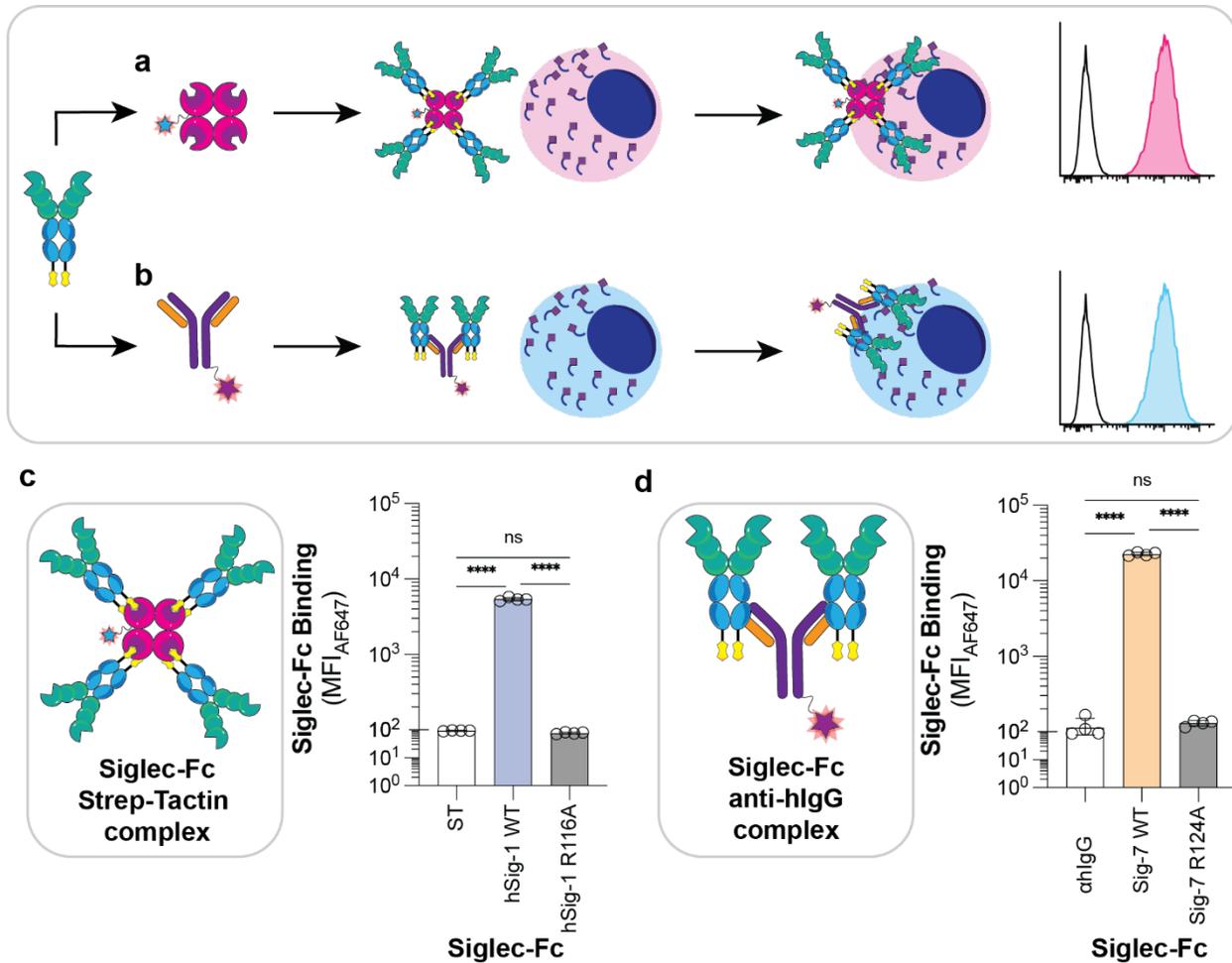


Figure 2.9: Siglec-Fcs can be complexed with either Strep-Tactin or an anti-hlgG antibody. **a**, **b**, complex of Siglec-Fc with Strep Tactin or an anti-hlgG antibody respectively. **c**, Staining of human hSiglec-1-Fc complexed with Strep Tactin to wildtype U937 cells. **d**, Staining of Siglec-7-Fc complexed with Strep-Tactin to wildtype U937 cells. Data is represented as the mean of four technical replicates and error bars represent one standard deviation from the mean. A one-way ANOVA test was used for statistical analysis. Not Significant (NS), $P > 0.05$; **** $P < 0.0001$.

2.3.1.2: Enzymatic treatment of cell glycans

Another approach for studying Siglec ligands using a cell-based glycan array is to treat the cells with a glycan modifying enzymes such as a sialidase (**Section 2.2.8.1**). NeuA cleaves sialosides regardless of the underlying regioselectivity (**Figure 2.10a**). Treatment of cells with neuraminidases is commonly used to demonstrate that Siglec-Fc binding is sialic acid dependant. Here it was used to determine if binding of our novel Siglec-Fc to cells was sialic acid dependent. As expected, when wildtype U937 cells were treated with NeuA, a significant reduction in hCD22 binding was observed (**Figure 2.10b**). While hCD22-Fc binding was significantly reduced

(approximately 10-fold) to the cells treated with NeuA, it is important to note that there was still greater binding with the wildtype CD22-Fc compared to the arginine mutant control. This suggests that not all CD22 ligands were removed during the treatment with NeuA and it could be important to consider in future experiments that enzymatic treatment of cells may not be 100% effective.

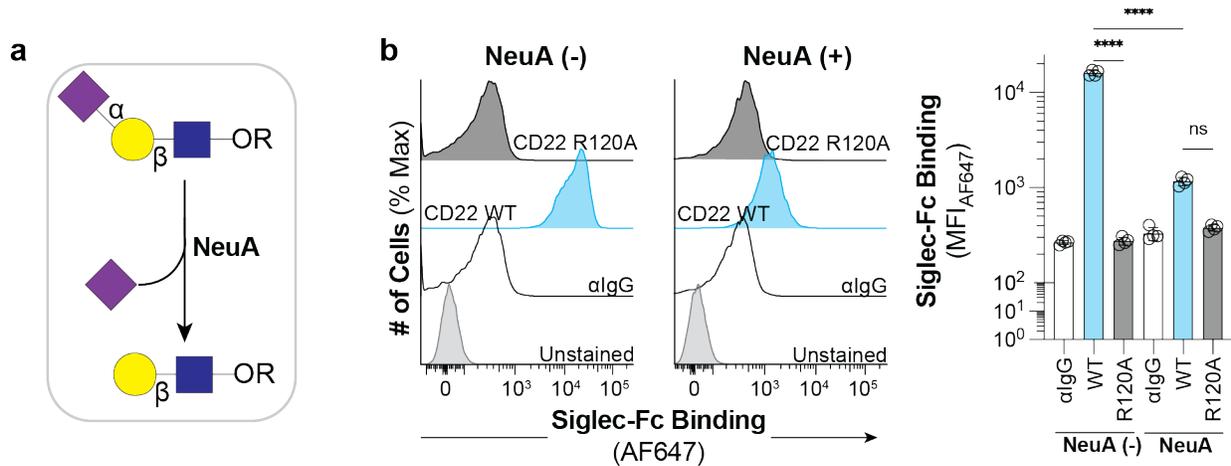


Figure 2.10: CD22-Fc staining of U937 cells treated with neuraminidase A. a, Neuraminidase A (NeuA) is a general sialic acid hydrolase that can cleave $\alpha(2\rightarrow3)$ and $\alpha(2\rightarrow6)$ linked sialic acids. b, CD22-Fc staining of U937 cells treated with and without NeuA. For panel b, data is presented with a representative flow cytometry histogram as well as a summary chart. Data in the summary chart is represented as the mean of four technical replicates and error bars represent one standard deviation from the mean. A one-way ANOVA test was used for statistical analysis. Not Significant (NS), $P > 0.05$; **** $P < 0.0001$.

2.3.1.3: Overexpression of glycan modifying enzymes

Using cells that have been engineered to overexpress glycan modifying enzymes can also be an efficient way to learn how glycan modifications effect Siglec-ligand interactions. This strategy was used to probe the effect of the presence of Neu5Gc in the glycocalyx of cells on the binding of human Siglecs. U937 cells virally transduced with CMAH (Cytidine monophospho-N-acetylneuraminic acid hydroxylase), which catalyzes the conversion of CMP-Neu5Ac to CMP-Neu5Gc and thereby artificially introducing Neu5Gc glycans into the glycocalyx of the human U937 cells. When binding of Siglec-Fcs to these cells was investigated, it was found that hSiglec-1, Siglec-9, and mSiglec-15 bound stronger in the absence of CMAH overexpression whereas CD33, Siglec-5, Siglec-8, Siglec-10, and Siglec-11 showed improved binding when CMAH was over expressed. The binding of the remaining Siglecs, (CD22 and Siglec-7) were unaffected by the overexpression of CMAH (Figure 2.11).

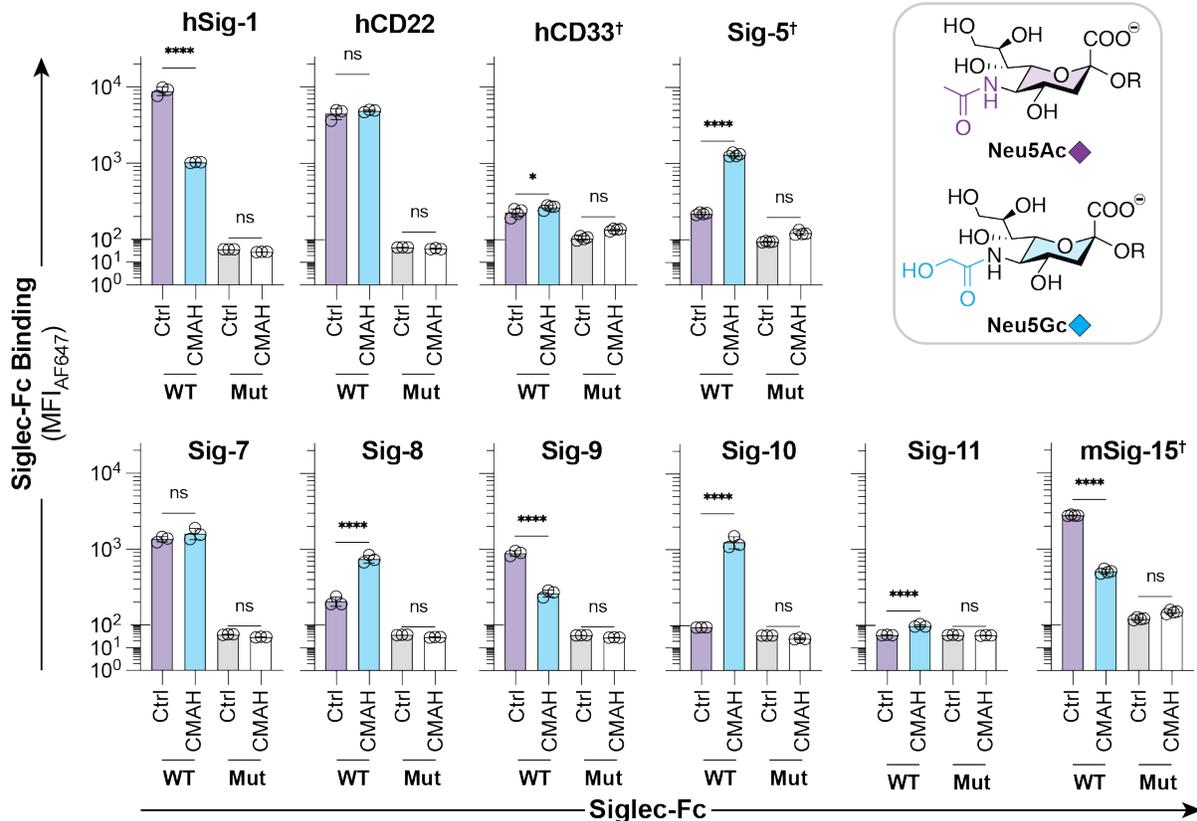


Figure 2.11: Siglec-Fc binding to U937 cells overexpressing CMAH. Data is presented as the mean of three technical replicates and error bars represent one standard deviation from the mean. † Represents that anti-hlgG was used as a secondary instead of Strep-Tactin. A one-way ANOVA test was used for statistical analysis. Not Significant (NS), $P > 0.05$; $*0.05 > P \geq 0.01$; **** $P < 0.0001$.

Another glycan modifying enzyme family that has recently been demonstrated to be very important in Siglec-ligand binding is the carbohydrate sulfotransferase family (CHST). CHST enzymes catalyze the addition of a sulfate group to either a galactose or GlcNAc at the sixth position. The importance of carbohydrate sulfation on Siglec binding was largely described using cell-based glycan arrays where cell lines were engineered to overexpress various members of the CHST family. Using this strategy Jung *et al.* probed the entire human and mouse Siglec family against a panel of cells overexpressing CHST1, 2, 4, 8, and 9⁷⁹. In this work, the authors found that carbohydrate sulfation increased binding to CD22, CD33, Siglec-5, -7, -8, -9, -14, and 15. As carbohydrate sulfation can enhance Siglec-ligand binding, using cells that overexpress these enzymes can be useful when elucidating the specific glycan ligands of a Siglec. The increase in

binding of mSiglec-15 and Siglec-9 to CHST1 and CHST2 U937 cells, respectively, compared to wildtype U937 cells is demonstrated in **Figure 2.12**.

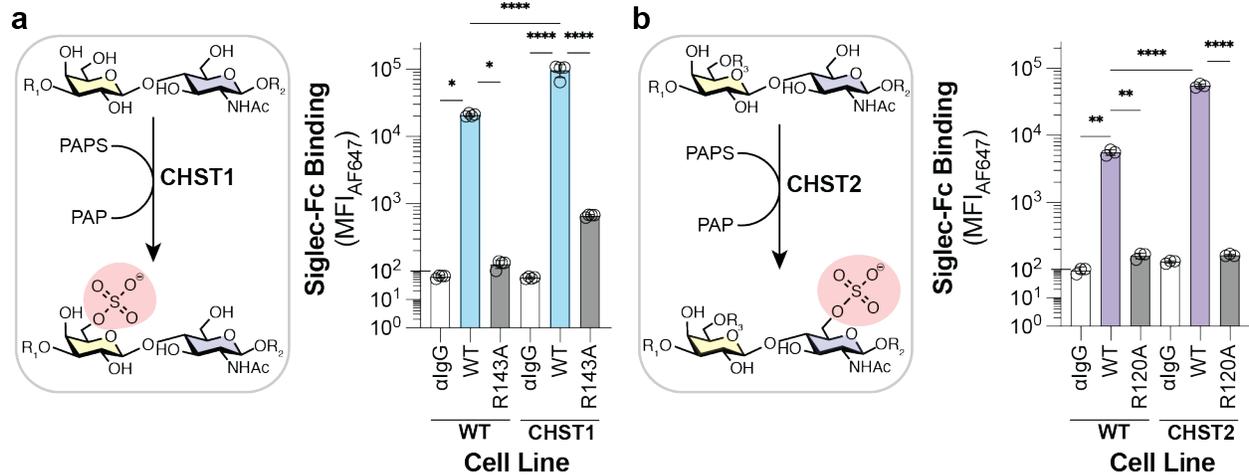


Figure 2.12: Binding of Siglec-Fcs to CHST1 and CHST2 expressing cells. **a**, Binding of mSiglec-15 to U937 CHST1 (catalyses the addition of a sulfate group to the 6th position on galactose) overexpressing cells. **b**, Binding of Siglec-9-Fc to U937 CHST2 (catalyses the addition of a sulfate group to the 6th position on GlcNAc) compared to wildtype U937 cells. R1, R2, and R3 are possible branches of the glycans. Data is presented as the mean of three technical replicates and error bars represent one standard deviation from the mean. A one-way ANOVA test was used for statistical analysis. Not Significant (NS), $P > 0.05$; **** $P < 0.0001$.

2.3.4: Applying engineered cells to describe Siglec ligands

Describing how Siglec binding is affected by the underlying chemistry (*e.g.* glycosidic linkage preference, enhancement by sulfation, *etc.*) of a cells' glycans is a good first step towards a better understanding of Siglec ligands. However, this alone is not particularly informative on which glyco-epitope(s) are the optimal ligand(s) for a Siglec as all sialic acid linkages and sulfation motifs can be found in most classes of glycans. To further describe which types of glycans are ligands for Siglecs, major glycan pathway enzymes (MGAT1, complex *N*-glycans; COSMC, elongated core-1 *O*-GalNAc *O*-glycans; *etc.*) can be knocked-out in cells. These knock-out cells can then be profiled with Siglec-Fcs to observe the effect of the knock-out on Siglec-Fc binding. However, in order to see the effect of the knock-out on Siglec binding to the cells, there needs to be robust binding to the cells before the gene is knocked-out. For some Siglecs like hSiglec-1, CD22, and Siglec-7 this is usually not an issue as these Siglecs bind to many different cell types with great avidity. Other Siglecs such as hCD33, Siglec-5, and Siglec-9 bind weakly to most cell

types in the absence of sulfation. In order to improve the binding of these Siglecs, the cells can be virally transduced with a gene which improves the binding of the Siglec to the cells (such as CHST1 for hCD33 and Siglec-5 or CHST2 for Siglec-9) before the gene of interest is knocked-out.

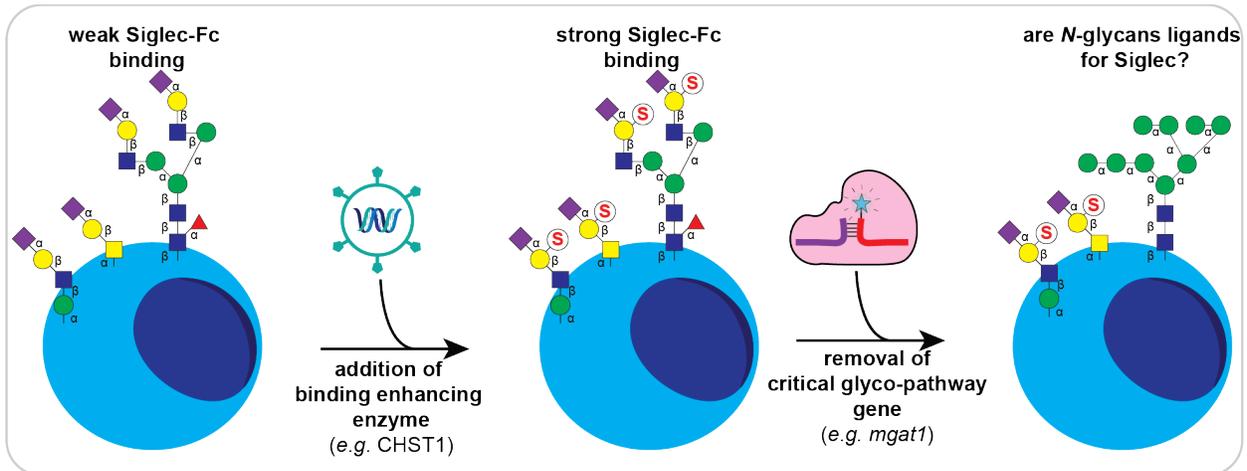


Figure 2.13: Synergistic effect of use of knocking-out glycan biosynthetic enzymes in cells which have been virally transduced towards enhanced Siglec binding towards describing Siglec ligands.

2.3.4.1: Describing the ligands of hCD33

Identifying the ligands of CD33 is of great interest due to its role in Alzheimer's disease (AD). In the human population there is a single nucleotide polymorphism (SNP), which is found to be homozygous in approximately 10% of the population²²³. This SNP results in the portion of the mRNA which codes for the V-set of CD33 to be spliced out of the transcript during mRNA maturation, leading to a protein product which lacks the V-set domain (**Figure 2.14**)²²⁴. Interestingly, individuals who are homozygous for this SNP have a decreased susceptibility to AD however it is unclear why this is the case²²⁴. In this context, full length CD33 is referred to as CD33M (for major) and the truncated protein is referred to as CD33m (for minor).

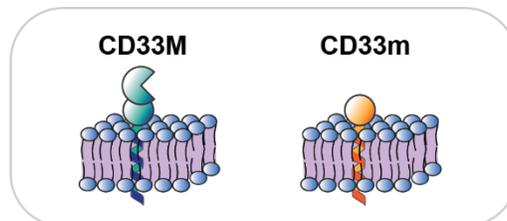


Figure 2.14: Isoforms of CD33. CD33M or the major isoform the protein product of the more common variant of CD33. CD33m is the minor isoform.

CD33 is expressed by microglia, the resident macrophages of the brain²²⁵. It has been proposed that this SNP reduces AD susceptibility through a gain-of-function mechanism, as microglia that express CD33m show decreased plaque burden compared to CD33 knock-out microglia²²⁴. The inability of CD33m to interact with its ligands may also contribute to AD pathology, motivating a better understanding of the ligands of CD33. However, to evaluate the role of CD33 ligands in AD, the ligand(s) of CD33 need to be described and this has been the focus of many studies^{37, 79, 80, 226}. To elucidate the ligands of CD33, a combination of cells overexpressing CHST1, and various glycosyltransferase knock-outs were used. Siglec-8 was used in parallel with CD33 as CD33 and Siglec-8 have similar sialoside ligand preferences²²⁶. Four different cell lines were used for this investigation. First, U937 cells expressing CHST1, the second was the same cells with MGAT1 knocked-out so that no complex or hybrid *N*-glycans can be produced by the cells¹⁵. The third cell line had COSMC knocked-out on top of MGAT1 and expressing CHST1. This prevents the extension of *O*-GalNAc *O*-glycans⁴³. The fourth cell line had POMGNT1 knocked-out in addition to MGAT1 and COSMC preventing elongation of *O*-mannose *O*-glycans (**Figure 2.15a**)⁵⁴. When the binding of the CD33-Fc was measured against these cell lines, it was found that binding decreased with each successive knock-out, yet binding was not reduced to the same level of the arginine mutant on the triple knock-out, suggesting that there are still CD33 ligands present. Siglec-8 showed a similar pattern as CD33, however binding increased to the double knockout compared to the wildtype CHST1 expressing cells. As there was still significant binding to the cells with MGAT1, COSMC, and POMGNT1 knocked-out, we posited that CD33/Siglec-8 may be binding to another type of sialic acid-containing glycan such as glycolipids, *O*-Fucose, or glyco-RNA.

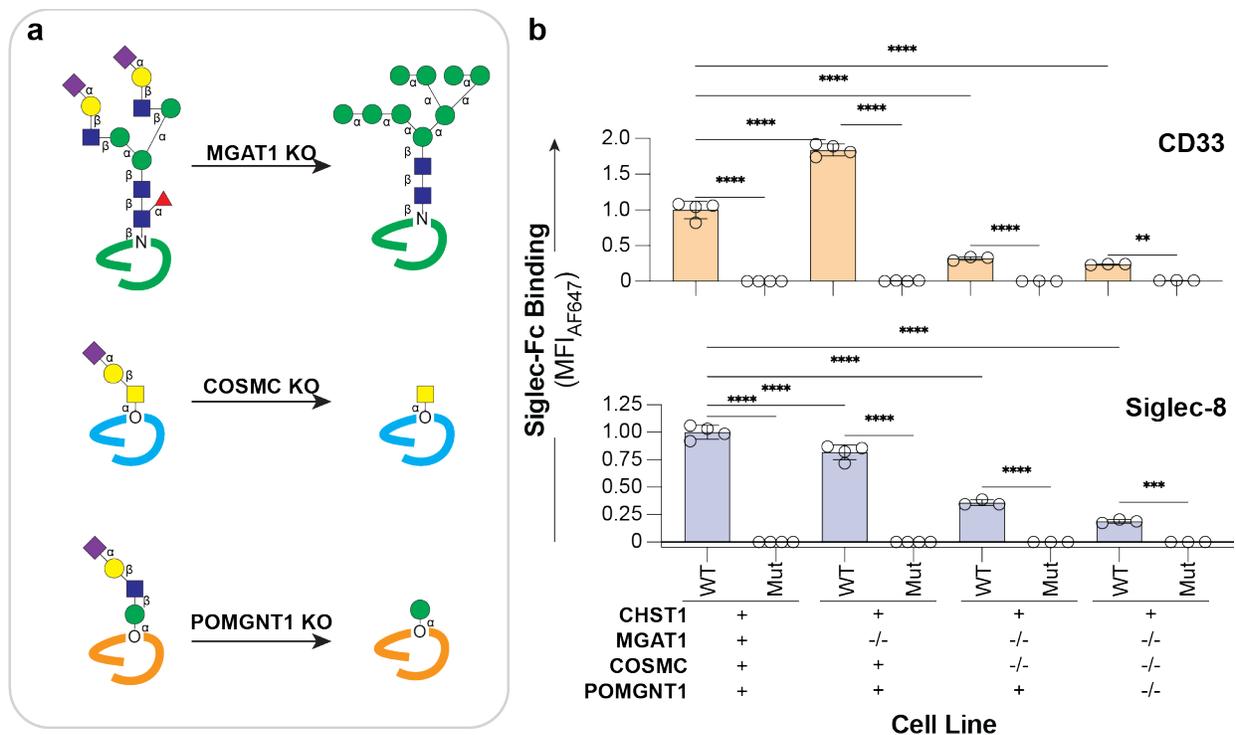


Figure 2.15: Binding of CD33 and Siglec-8 Fc to U937 cells overexpressing CHST1 with various critical glycosylation enzymes knocked-out. **a**, depiction of the biosynthetic consequence of knocking-out MGAT1, COSMC, and POMGNT1. **b**, Binding of CD33 and Siglec-8-Fc to glyco-knock-out U937 cells. Data is presented as the mean of four technical replicates and error bars represent one standard deviation from the mean. A one-way ANOVA test was used for statistical analysis. ** 0.01 > P ≥ 0.001; ***0.001 > P ≥ 0.0001; ****P < 0.0001.

2.3.5 Probing Siglec ligands on primary cells

Cell lines are an excellent starting point to study Siglec ligands. As cell lines are immortalized, they can be easily grown in large numbers and serve as an ideal platform to optimize tools to describe Siglec ligands such as the Siglec-Fc chimeras. However, it is important to recognize that cell lines are not a replacement for genuine biological samples and ultimately, the goal is to describe Siglec ligands in a genuine biological context such as on primary human cells and tissues. In this section, the Siglec-Fc probes which were validated and optimized above were used to probe Siglec ligands on primary human cells.

2.3.5.1 Strep-Tactin is a superior secondary when profiling Siglec ligands on primary cells

After working mainly with immortalized cell lines, we were eager to apply what we learned about Siglecs to genuine human cells. As human cells and tissues are coated with IgG isotype antibodies, anti-hIgG secondary can cause a dramatic increase in background signal or glycan

independent Siglec binding (**Figure 2.16**). This makes Strep-Tactin a better choice of secondary when probing Siglec ligands on primary human cells.

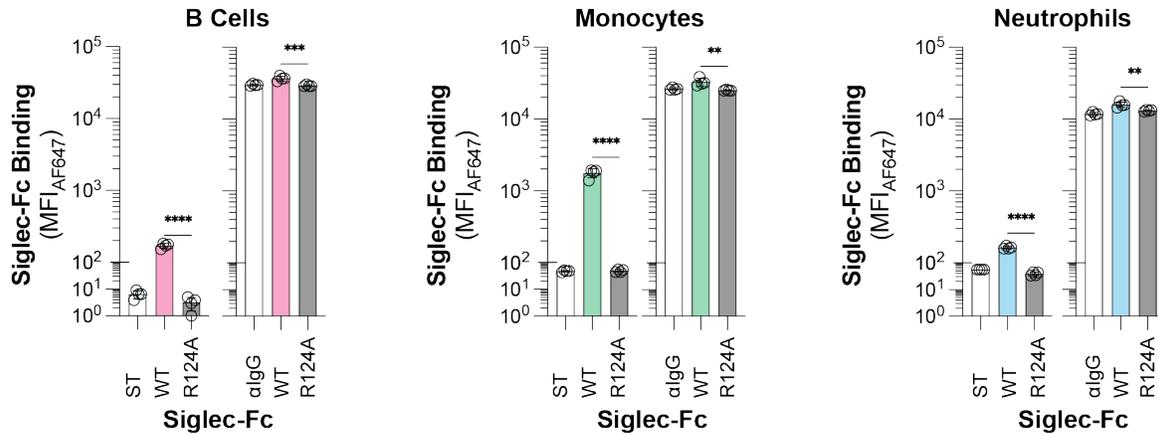


Figure 2.16: Binding of Siglec-7 precomplexed with Strep-Tactin and anti-hlgG to primary human peripheral blood mononuclear cells. Data is represented as the mean of four technical replicates and error bars represent one standard deviation from the mean. A one-way ANOVA test was used for statistical analysis. Not Significant (NS), $P > 0.05$; **** $P < 0.0001$.

2.3.5.2 Siglec expression on primary cells

Accordingly, we started by measuring Siglec expression on primary human cells. To this end, immune cells were isolated from the spleens of ten donors (5 males, 5 females) and probed for both Siglec ligands and Siglec expression. From these donors, six types of immune cells were isolated by traditional flow cytometry: neutrophils, B cells, CD4⁺ and CD8⁺ T cells, monocytes, and mature natural killer cells (**Figure 2.17**, **Figure A2.1-Figure A2.6**).

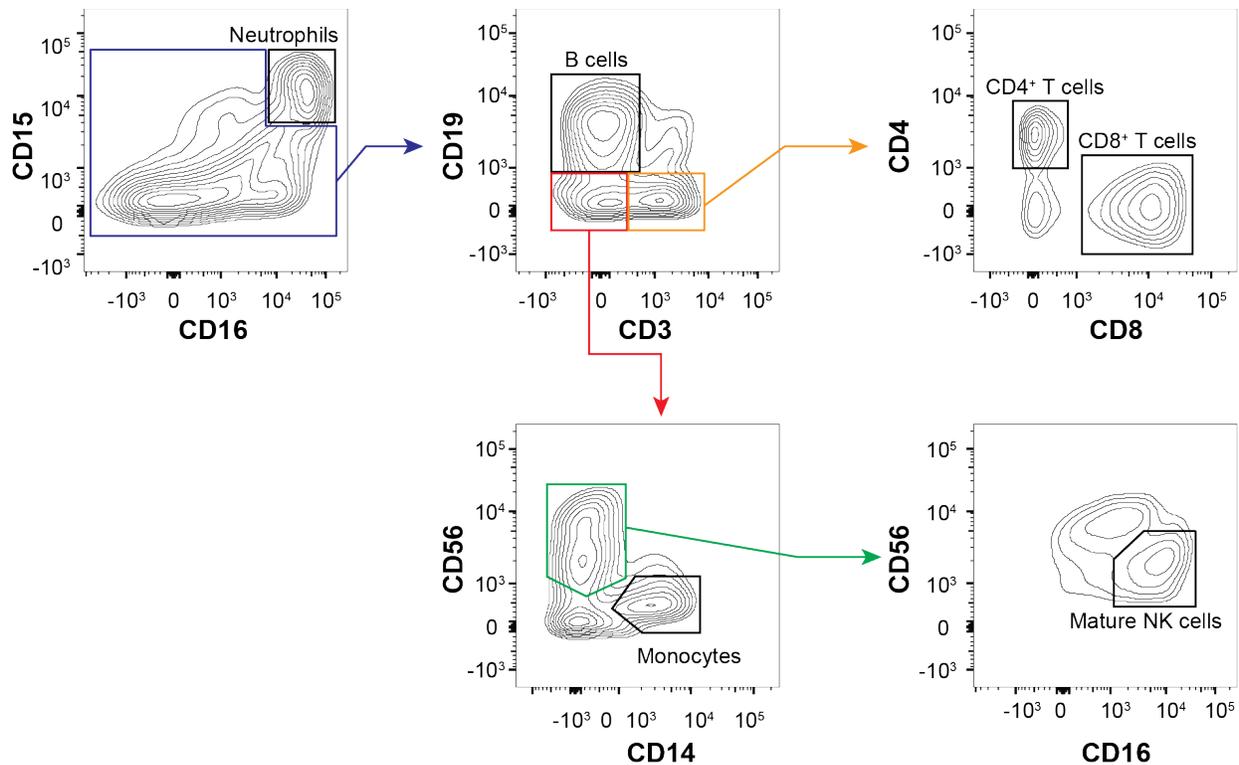


Figure 2.17: Gating Strategy to isolate primary immune cells from human spleens. Using multidimensional flow cytometry neutrophils ($CD15^+$, $CD16^+$), B cells ($CD19^+$), $CD4^+$ and $CD8^+$ T cells, monocytes ($CD14^+$) and mature natural killer (NK) cells ($CD56^+$, $CD16^+$) were isolated.

Starting with Siglec expression, it was found that Siglec-1, Siglec-6, Siglec-8, Siglec-11, and Siglec-15 were not expressed on any of the immune cells tested however, not all immune cell types were examined. CD22 was found to be unique to B cells and expressed relatively strongly^{6, 111, 125, 227}. CD33 was expressed by neutrophils and monocytes with monocytes expressing CD33 20-fold higher than neutrophils. Siglec-5 was found to be expressed by neutrophils, B cells, and monocytes with similar levels of expression between neutrophils and monocytes, which was approximately 20-fold higher than B cells. Siglec-7 and Siglec-9 had a similar expression profile being expressed by neutrophils, mature NK cells, and monocytes with Siglec-7 levels being the at similar levels on monocytes and mature NK cells and about 20-fold lower on neutrophils. Siglec-10 was expressed by B cells and monocytes but was expressed 10-fold more by monocytes.

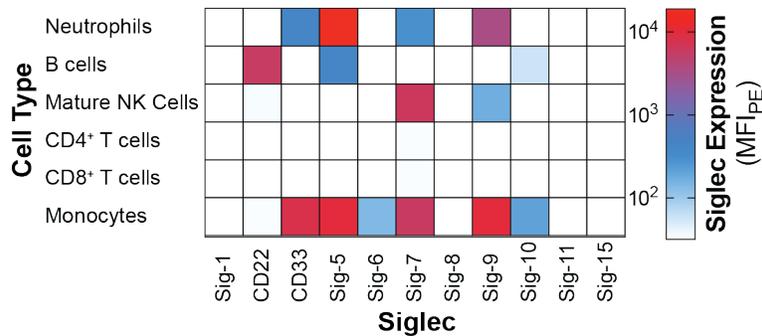


Figure 2.18: Expression of Siglecs by human splenocytes. Results are represented as the mean anti-Siglec antibody binding from ten biological replicates.

2.3.5.3 Siglec ligand expression on primary cells

To tie together the glycan binding preferences to Siglec ligands on cells with the Siglec expression, we probed human splenocytes from the same ten donors with Siglec-Fcs by flow cytometry. Typically, Siglec-Fc binding to cells requires pre-complexation with anti-hlgG but using anti-hlgG with primary cells leads to high background signal as demonstrated above (**Figure 2.16**, **Figure A2.7-Figure A2.12**). To overcome this limitation, the Siglec-Fcs were pre-complexed with Strep-Tactin instead of anti-hlgG³⁷. hSiglec-1, CD22, and Siglec-7 ligands were found ubiquitously across all the immune cell types tested. The binding of the CD22-Fc to B cells was the strongest of all the Siglec-Fcs tested, followed by Siglec-1-Fc to monocytes. The Siglec-7-Fc bound mature NK cells relatively strongly. The remaining Siglec-Fcs bound relatively weakly compared to hSiglec-1, hCD22 and Siglec-7. Ligands of CD33 were found on neutrophils, mature NK cells, CD8⁺ T cells and monocytes with similarly intense binding to neutrophils and monocytes and weaker binding to mature NK cells and CD8⁺ T cells. Siglec-5 ligands were found on neutrophils, and monocytes with the strongest binding to neutrophils. Interestingly Siglec-5-Fc bound neutrophils independent of its canonical arginine (Arg119). Siglec-6, Siglec-8, Siglec-9, Siglec-10, mSiglec-15 were all found to bind monocytes albeit rather weakly; however, Siglec-10 could also bind neutrophils. Siglec-11 ligands were not found on any of the cell types tested.

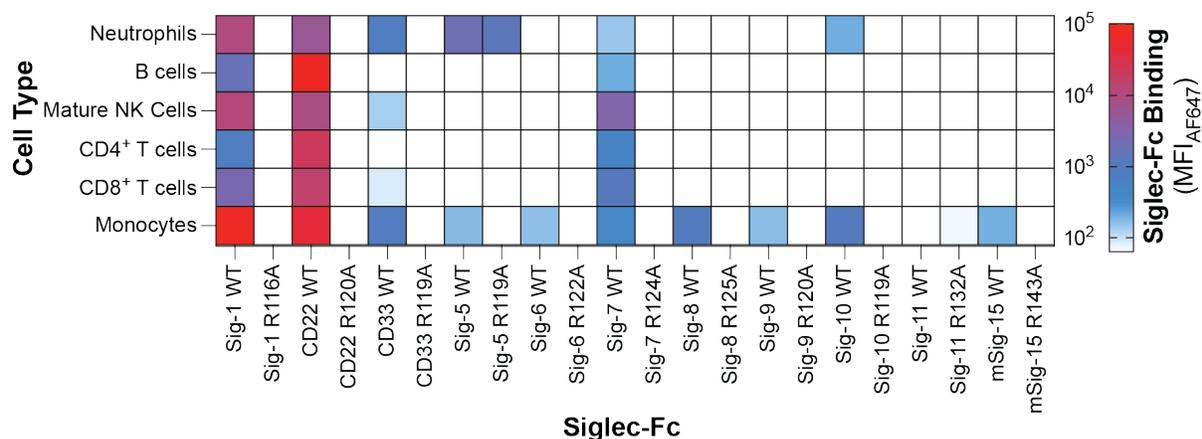


Figure 2.19: Siglec-Fc binding to human splenocytes. Results are represented as the mean Siglec-Fc-Strep-Tactin complex binding of each Siglec to ten biological replicates.

2.3.6: Other Siglec-Fc-based approaches

While Siglec-Fc chimeras are very useful for describing Siglec ligands in cell-based glycan arrays, our group's version of the Siglec-Fc chimera is a versatile tool that can be applied in many different approaches. Previously, Siglec-Fcs have been used in many other experimental approaches. In this section, it is demonstrated that our version of the Siglec-Fc can also be used in approaches beyond the cell-based glycan array.

2.3.6.1: Siglec-Fc ELISA

Siglec-Fcs have been used in ELISAs since the early 2000's and are still used to describe Siglec ligands presently (2024)^{128, 141, 142, 143, 228}. The elegance of this approach is that it is relatively simple to perform and no specialized equipment is needed. Traditionally, an anti-hIgG conjugated to an enzyme such as horseradish peroxidase (HRP) is used as a secondary¹⁴¹; however, we wanted to confirm that our new Siglec-Fc chimera also worked in an ELISA and to determine that the Strep-Tag II could be leveraged to form a complex with Strep-Tactin-HRP. To address this, the glycans (in this case synthetic glycolipids, referred to as neoglycolipids (**Figure 2.20a, b**) are non-specifically adsorbed to the surface of a microplate and the Siglec-Fc is precomplexed to Strep-Tactin-HRP. The Siglec-Fc-Strep-Tactin-HRP complex is added to the well of the microplate containing the glycolipid (**Figure 2.20c**). Unbound Siglec-Fc is washed away and the binding of the Siglec-Fc to the glycan is measured by adding an HRP substrate (3,3',5,5'-tetramethylbenzidine), which becomes coloured when oxidised by HRP. The reaction is quenched

by adding an acidic solution. Using this approach, binding between hSiglec-1¹⁶⁰ and hCD33¹⁶⁷ and their respective high affinity ligands was measured (**Figure 2.20d**). High affinity ligands were used as they are easy to synthesize compared to genuine glycoproteins and glycolipids and their hydrophobic tail facilitate absorption to the surface of the microplate. While these results themselves are not particularly novel, they demonstrate that our new Siglec-Fc construct works in an ELISA and having a handle on this assay was important in the following chapters.

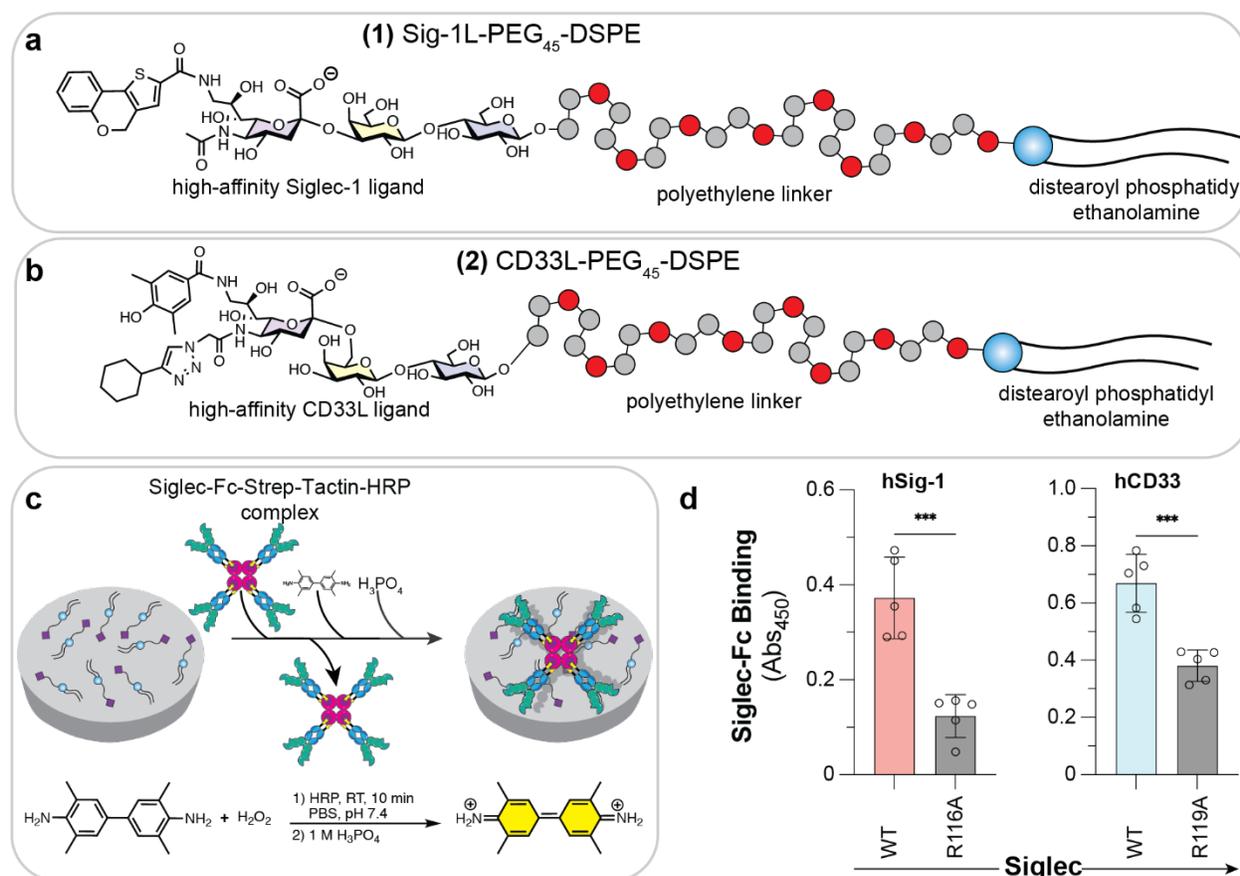


Figure 2.20: Siglec-Fc enzyme-linked immunosorbent assay used to probe Siglec ligands. **a, b**, chemical structures of high affinity Siglec-1 and hCD33 ligands appended to polyethylene glycol (PEG)₄₅ distearoyl phosphatidylethanolamine respectively. **c**, Depiction of ELISA where neoglycolipids are adsorbed to a well of a microplate followed by the addition of the Siglec-Fc-Strep-Tactin-Horseradish peroxidase (HRP). Unbound complex is then washed away and Siglec binding is quantified upon the addition of TMB producing a blue coloured compound. The reaction is then quenched with an acidic solution causing the substrate to turn yellow. The absorbance is then measured at 450 nm. **d**, ELISA results of hSiglec-1-Fc and hCD33 binding to their respective ligands in an ELISA. A two tailed Student's t-test was used for statistical analysis. ***0.001 > P ≥ 0.0001.

2.3.6.2: Siglec-Fc bead assay

While cell-based glycan arrays are excellent at addressing if Siglec ligands are present on a specific cell type, cell-based glycan arrays do not provide detailed information on the specific

glycan recognized by a Siglec. However, the Siglec-Fc platform can still be used to answer this question when applied in a different, more targeted, experimental approach. Previous investigations have developed assays where polyacrylamide glycan probes are complexed to microbeads. Following the complexation of the beads with the ligand polymers, a Siglec-Fc-complexed with fluorescently labelled antibody is added to the beads and the strength of the interaction is quantified by flow cytometry¹²³. Taking inspiration from this approach, we leverage the Strep-Tag II and streptavidin coated microbeads to form Siglec-decorated microbeads. As for the ligands, instead of using polyacrylamide glycan probes we used glycan bearing liposomes. The Siglec-Fc-bead assay is an ideal approach for addressing if a certain glycan is a ligand for a Siglec. For example, in a bead assay, a specific ligand of interest is chemically appended to a lipid, allowing it to be formulated into a fluorescently labeled liposome (**Figure 2.21a, b**). The Siglec-Fc is then precomplexed with streptavidin coated beads (approximately 1 μM in diameter) and the Siglec-coated beads are then incubated with the ligand bearing liposomes and binding is quantified by flow cytometry. This approach was validated using a previously designed Siglec-1 ligand¹⁶⁰, an FDA approved liposome formulation²²⁹, and hSiglec-1-Fc coated microbeads (**Figure 2.21c**).

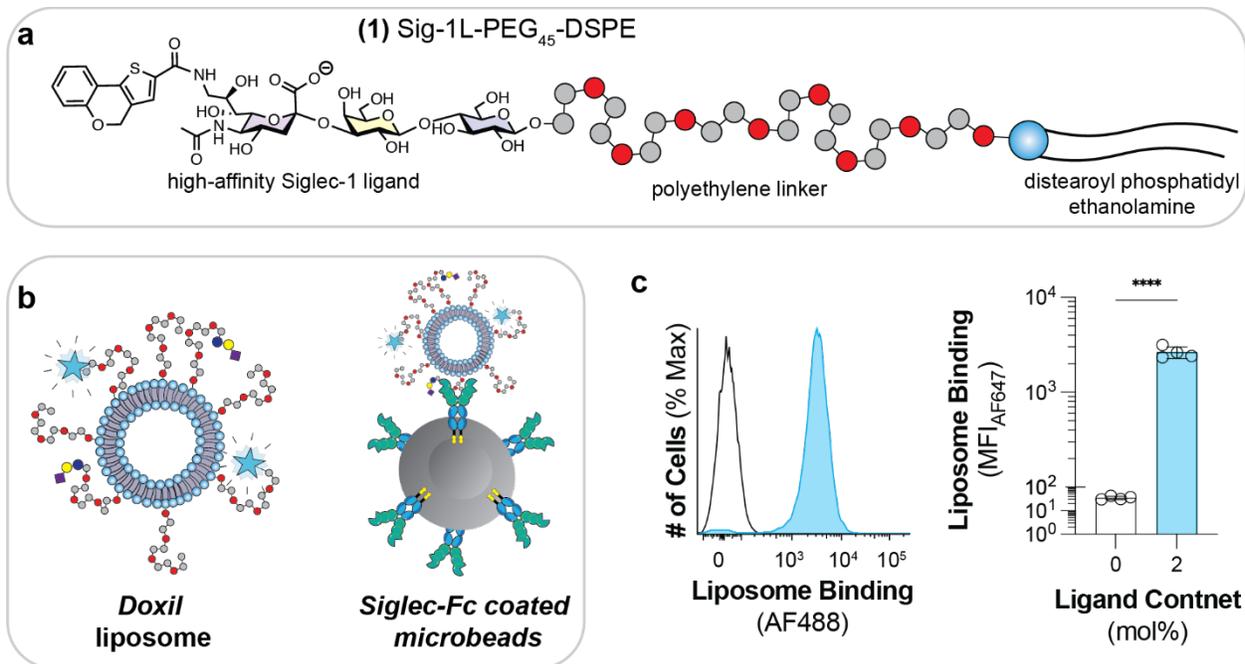


Figure 2.21: Siglec-Fc bead assay used to measure the binding between Siglecs and liposome bearing Siglec ligands. a, Structure of Siglec-1 high affinity ligand appended to PEG₄₅-DSPE. b, Depiction of liposomes and Siglec-Fc bead assay. c, Binding of liposomes formulated 2 mol% high affinity Siglec-1 ligand to beads coated with wildtype hSiglec-1. Data is represented as the mean of four technical replicates and error bars represent one standard deviation from the mean. A two tailed Student's t-test was used for statistical analysis. **** = $P < 0.0001$.

2.3.6.3: Siglec-Fc mass-spectrometry-based assay

Another application of the Siglec-Fc chimera is for it to be used in mass-spectrometry based approaches. Unlike the other approaches, binding between a Siglec-Fc and glycans in these types of assays does not require a reporter or to be precomplex with a secondary. These types of assays are excellent for measuring the affinity of a Siglec for its ligand whereas the other approaches described above (cell-based glycan array, bead assay, ELISA) measure avidity. In theory, this approach is quite simple as the relative amounts of free protein, free ligand, and protein ligand complex can be measured and used to calculate the dissociation constant (K_d). For this to work practically, a uniform protein spectrum is required. To achieve this, the Siglec-Fc is produced in Lec1 CHO cells so that all the *N*-glycans on the Siglec-Fc are high mannose *N*-glycans. The Siglec-Fc is then purified as described above. Once purified, the Siglec-Fc is treated with tobacco etch virus (TEV) protease to liberate the Siglec domains from the Fc generating a monomeric soluble Siglec fragment. The Siglec fragments are then treated with EndoH to remove

the high mannose *N*-glycans (**Figure 2.22a**)³⁷. In 2023, this strategy has been improved with the development of concentration-independent (COIN) native mass spectrometry developed by Bui D. *et al.* in 2023³⁹ where the Siglec-Fc can be used instead of the Siglec fragment.

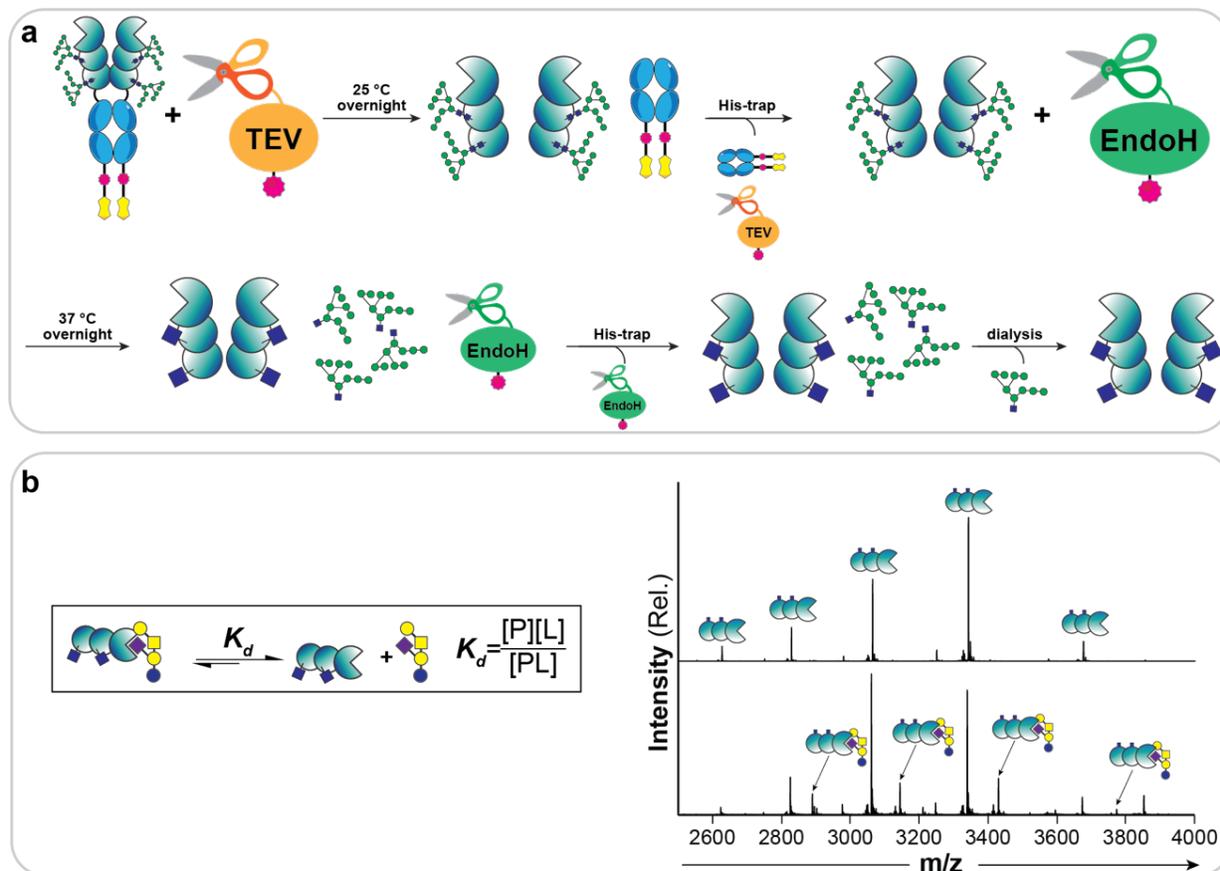


Figure 2.22: Depiction of mass-spectrometry based assay for measuring the dissociation constant (K_d) between a Siglec and a ligand. a, Depiction of the preparation of the monomeric Siglec from the Siglec-Fc. **b,** Representative mass spectra of Siglec fragment and Siglec-fragment-oligosaccharide complex.

After the Siglec fragment is prepared, it is then mixed with increasing concentrations of a potential ligand and the amount of protein-ligand complex is determined (**Figure 2.22b**) and the K_d is derived from this information. This strategy was used in many studies to determine the affinity of Siglecs towards various oligosaccharides (**Table 2.1**).

Table 2.1: Dissociation constants derived between Siglec and glycans using Siglec fragment approach. *Original study, †work is under review, OS, oligosaccharide.

Study	Siglec	Glycan	K_d (mM)
Rodrigues E., <i>et al.</i> 2020 ^{37*}	hSiglec-1	$\alpha(2\rightarrow6)$ Siayllactose	1.0 ± 0.2
		$\alpha(2\rightarrow3)$ Siayllactose	0.5 ± 0.1
		$\alpha(2\rightarrow6)$ SiayllacNAc	1.6 ± 0.3
		$\alpha(2\rightarrow3)$ SiayllacNAc	0.6 ± 0.1
	hCD22	$\alpha(2\rightarrow6)$ Siayllactose	0.075 ± 0.004
		$\alpha(2\rightarrow3)$ Siayllactose	Not detected
		$\alpha(2\rightarrow6)$ SiayllacNAc	0.060 ± 0.005
		$\alpha(2\rightarrow3)$ SiayllacNAc	Not detected
	hCD33	$\alpha(2\rightarrow6)$ Siayllactose	2.6 ± 0.8
		$\alpha(2\rightarrow3)$ Siayllactose	2.0 ± 0.3
		$\alpha(2\rightarrow6)$ SiayllacNAc	2.7 ± 0.4
		$\alpha(2\rightarrow3)$ SiayllacNAc	2.7 ± 0.1
Jung J., <i>et al.</i> 2021 ⁷⁹	hCD33	$\alpha(2\rightarrow3)$ SiayllacNAc	>20
		$\alpha(2\rightarrow3)$ Siayl 6-O-Sulfo-Gal lacNAc	2.5 ± 0.08
		$\alpha(2\rightarrow3)$ Siayl 6-O-Sulfo-GlcNAc lacNAc	8.5 ± 0.5
		$\alpha(2\rightarrow3)$ Siayl 6-O-Sulfo-Gal, 6-O-Sulfo-GlcNAc lacNAc	0.70 ± 0.1
Schmidt E.N., <i>et al.</i> 2023 ¹²⁸	hSiglec-1	GM1a OS	1.2 ± 0.1
		GM2 OS	0.9 ± 0.1
		GM3 OS	0.5 ± 0.1
		GD1a OS	1.3 ± 0.2
Schmidt E.N., <i>et al.</i> 2024 ²²⁸	hSiglec-1	GM1a OS	1.5 ± 0.1
		GM1b OS	0.89 ± 0.03
	mSiglec-1	GM1a OS	2.0 ± 0.1
		GM1b OS	1.2 ± 0.06

2.3.7: Improving the Siglec-Fc

Siglec-Fcs have been tools to study Siglec ligands for over thirty years, however, in 2020, Rodrigues *et al.* increased the versatility of the Siglec-Fc with the addition of the Strep-Tag II, TEV sequence, and mutant Fc region³⁷. The development of our panel of soluble Siglecs was started before I began my doctoral work. After performing many experiments with our Siglec-Fcs as described above, it was found that there was opportunity to improve these constructs with respect to the storage, expression and activity.

2.3.7.1: Improving Siglec-Fc activity through amino acid selection

To generate a Siglec-Fc, a section of the Siglec needs to be selected to be recombinantly joined to the Fc portion of the antibody. Deciding how much of the Siglec sequence will be included in the Siglec-Fc is very important with respect to its function and stability. Siglec-15 is conserved between mice and humans and the Siglecs share 82% primary sequence identity yet when they were compared head-to-head in a cell-based glycan array, the mSiglec-15-Fc significantly out-performed its human counterpart (**Figure 2.23a**). This was surprising as h/mSiglec-15 are very similar proteins and were expected to bind to cells with similar avidity. Consulting a sequence alignment between our hSiglec-15-Fc, mSiglec-15-Fc, and a commercial version of hSiglec-15 (from *R&D Systems-Figure 2.23b*), we noticed that our version of the hSiglec-15-Fc was two amino acids shorter. Indeed, when the extra two amino acids were added, hSiglec-15-Fc binding was rescued. This result serves as a cautionary tale for designing chimeric proteins, and if a chimeric protein such as a Siglec-Fc is not performing as expected, it may be due to its recombinant structure which causes the protein to misfold. Indeed, during the purification of hSiglec-15 generation 1, precipitate was observed during the dialysis; however, this was not observed with hSiglec-15 generation 2. This suggest that the reduced functionality of hSiglec-15 was likely due to an unstable recombinant structure which was stabilized by the addition of the two amino acids.

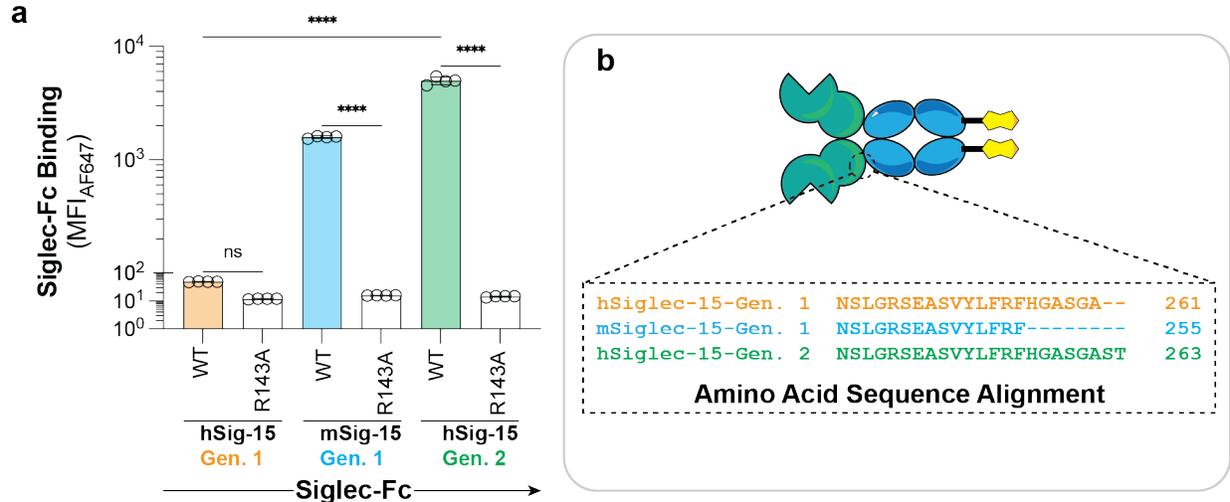


Figure 2.23: Binding of different generations of hSiglec-15-Fc to CHST1 expressing U937 cells. a, Binding of generation 1 and generation 2 hSiglec-15 as well as mSiglec-15-Fc to CHST1 overexpressing cells. **b,** Sequence alignment of the hSiglec-15 generation 1 and generation 2 as well as mSiglec-15-Fc. Data is represented as the mean of four technical replicates and error bars represent one standard deviation from the mean. A one-way ANOVA test was used for statistical analysis. Not Significant (NS), $P > 0.05$; **** $P < 0.0001$.

2.3.7.2: Improving Siglec-Fc expression and activity

CHO cells are famous for being able to produce large amounts of protein recombinantly in the order of 10-1000 mg/L which was one of the reason CHO cells were chosen to produce the Siglec-Fcs²³⁰. However, for the majority of our Siglec-Fc chimeras, the yields are closer to 0.25-1 mg/L. Some gains in Siglec-Fc yield were realized when cells were cultured at 32 °C instead of 37 °C, which is in line with previous findings²³¹. Interestingly we found that the arginine mutant typically expresses at double the yield of the wildtype Siglec-Fc. We hypothesized that this may be due to the ability of the wildtype Siglec-Fc to engage with sialosides on the CHO cells reducing the yield and posited that if this were the case, Siglec-Fcs produced from CHO cells lacking sialic acid containing glycans may improve the yield of the wildtype Siglecs. Moreover, Siglec-Fcs produced from cells which lack sialic acid containing glycans may have improved activity due to them being unmasked which has been proposed previously²³².

To test this hypothesis, CMAS was knocked-out of CHO cells using CRISPR/Cas9. Clones which had CMAS knocked-out, would not have any cell surface sialic acid and thus could be

identify by a lack of binding by Siglec-Fcs in a cell-based glycan array. Using Siglec-1 as a screen, three CMAS KO clones were identified (**Figure 2.24**).

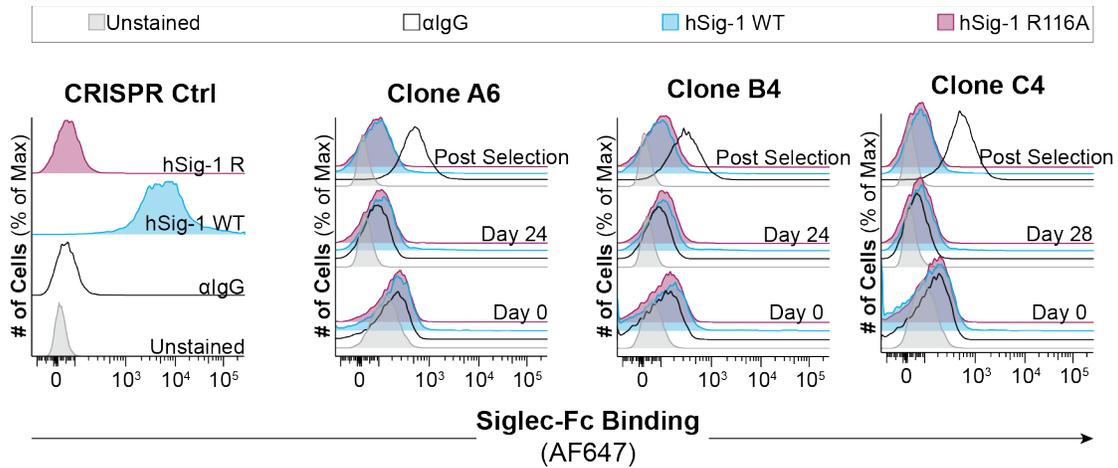


Figure 2.24: Validation of CMAS KO CHO Flp-In cells. CMAS KO CHO cells were checked for the presence of cell-surface sialic acid using hSiglec-1-Fc as hSiglec-1 has robust binding to WT CHO cells and should not bind to CMAS KO CHO cells.

Following the successful generation of CMAS KO CHO cells, the three clones were transfected with Siglec-7-WT-Fc and Siglec-7-WT-Fc was isolated from these cells. It was found that all three of the CMAS KO cell lines had increased Siglec-Fc expression compared to wildtype CHO cells (**Figure 2.25a**). Moreover Siglec-Fcs produced from CMAS KO cells showed as good or better performance when used in a cell-based glycan array against wildtype U937 cells (**Figure 2.25b**) or in an ELISA against ganglioside GD3 (**Figure 2.25c**)

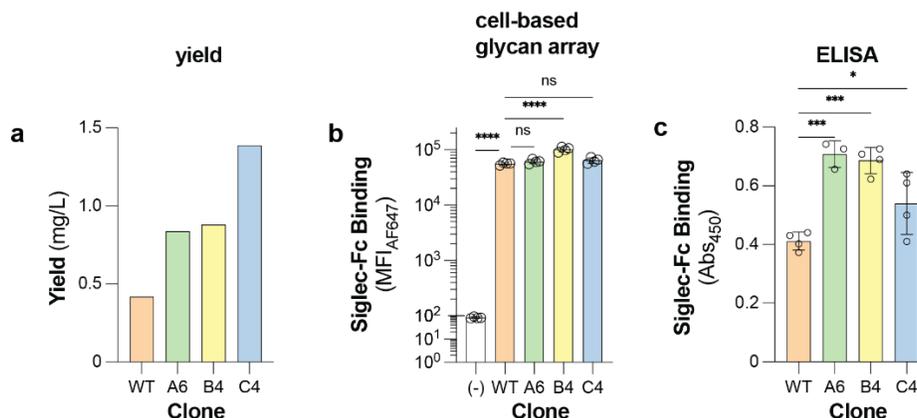


Figure 2.25: Comparing yield and function of Siglec-7-WT-Fc produced from wildtype and CMAS KO CHO cells. **a**, yield of Siglec-7-WT-Fc from wildtype CHO cells and three CMAS KO CHO cell clones (A6, B4 and C4, one preparation from each clone). **b**, **c**, Functional comparison of Siglec-7-WT-Fc expressed from wildtype CHO cells and CMAS KO clones in a cell-based glycan array using WT U937 cells and in an ELISA using ganglioside GD3 respectively. For panel a, the yield of Siglec-7-WT-Fc from WT CHO cells were compared to the yield from three CMAS KO CHO cells. For panel b, and c data is represented as the

mean of at least three technical replicates and error bars represent one standard deviation from the mean. A one-way ANOVA test was used for statistical analysis. Not Significant (NS), $P > 0.5$; $0.05 > P \geq 0.01$; $0.001 > P \geq 0.0001$; $P < 0.0001$.

While these initial results with Siglec-7 were encouraging, it needed to be determined if this was a general improvement to the Siglec-Fc platform or if it was unique to Siglec-7. To answer this question, more members of the Siglec family were produced as Fc chimeras from CMAS KO CHO cells. It was found that Siglec-Fcs produced from CMAS KO cells were expressed on average 4-fold higher than from wildtype CHO cells, yet there was no significant increase in the yield for the arginine mutants produced from the CMAS KO cells compared to the wildtype CHO cells (Figure 2.26). While the yields are still an order of magnitude lower than other proteins expressed from CHO cells, expressing Siglec-Fcs from CHO CMAS KO cells does improve the yield significantly.

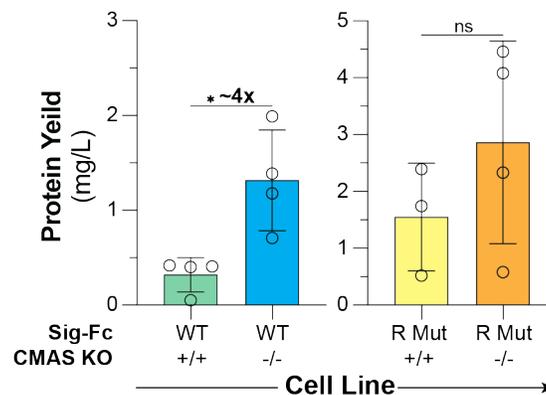


Figure 2.26: Summary of Siglec-Fc yields from wildtype and all C4 CMAS KO CHO cells. Siglec-Fcs used in the comparison for the wildtype Siglec-Fc were Siglec-5 WT, Siglec-9 WT, Siglec-7 WT, and Siglec-7 WT (initial test) from wildtype CHO cells and hCD33 WT, Siglec-5 WT, Siglec-8 WT, and Siglec-7-WT (initial test) from CMAS KO CHO cells. Siglec-Fcs used in the comparison for the arginine mutant Siglec-Fc expression comparison were Siglec-5 R119A, Siglec-7 R124A, and Siglec-9 R120A from CHO WT cells and hCD33 R119A, Siglec-5 R119, Siglec-8 R125A, and Siglec-9 R120A for CMAS KO CHO cells.

When the activity of Siglec-Fcs from CMAS KO CHO cells was compared to Siglec-Fcs produce from wildtype CHO cells in a cell-base glycan array, for each Siglec, there was an improvement with Siglec-9 showing the greatest improvement in binding at nearly 100-fold compared to the Siglec-9-Fc produced from wildtype CHO cells. CD33 and Siglec-8 also showed improvement with -Fc binding increasing nearly 3-fold compared to the Siglec-Fc produced from wildtype CHO cells. (Figure 2.27). Together these results demonstrate that Siglec-Fcs produced

from CMAS KO CHO cells are produced at higher yields and have modestly but significantly improved function which could be due to unmasking of the Siglec-Fc due to the Siglec Fc being produced in CMAS KO cells.

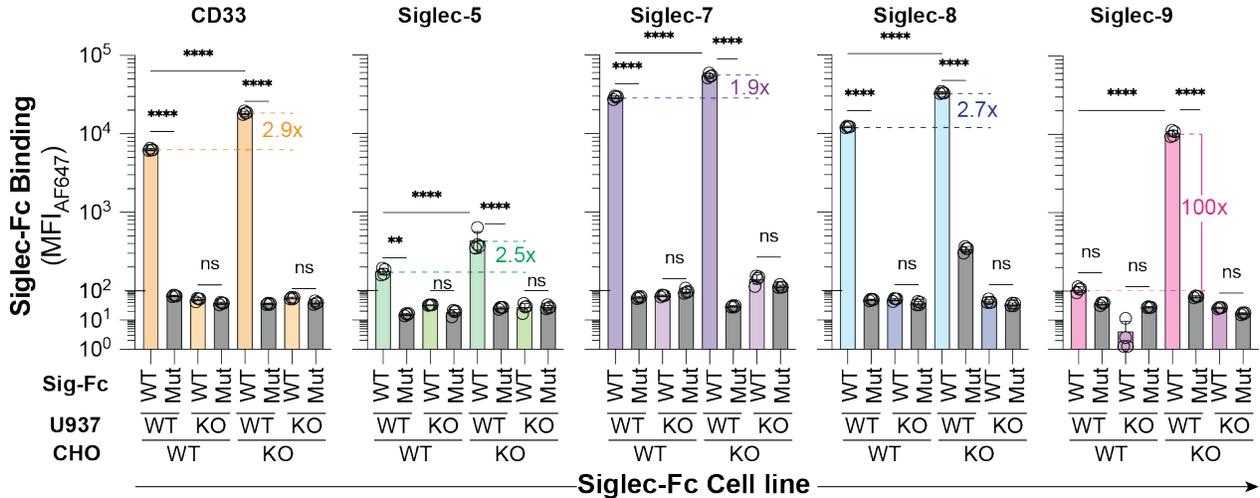


Figure 2.27: Binding of Siglec-Fc produced from wildtype and CMAS KO CHO cells to U937 cells overexpressing CHST1/2. U937 refers to CHST1/2 expressing U937 cells. KO U937 cells refers to cell expressing CHST1 but have CMAS KO, making the cell devoid of sialylated glycans.

2.4: Discussion

2.4.1: Siglec-Fcs as tool to describe Siglec ligands

Siglecs are important receptors for maintaining immune homeostasis and are implicated in many pathologies such as cancers, neurodegenerative diseases, and infections^{6, 97, 101, 187}. Due to their relevance in many diseases, Siglecs are a promising target for immunomodulating therapeutics. However, before Siglecs can be modulated, their ligands need to be determined. To this, Siglec-Fc chimeras are versatile tools which can be applied in many approaches to describe Siglec ligands (**Table 2.2**).

Table 2.2: Publications that have used the Macauley Lab Siglec-Fc chimeras.*Original work, †work is under review.

Publication	Application	Year
Rodrigues E., <i>et al.</i> ^{*37}	cell-based glycan array immunofluorescence mass spectrometry	2020
Jung J., <i>et al.</i> ⁷⁹	cell-based glycan array mass spectrometry	2021
Schmidt E.N., <i>et al.</i> ¹²⁸	ELISA bead assay mass spectrometry	2023
Bui D., <i>et al.</i> ³⁹	mass spectrometry (COIN)	2023
Garnham R. <i>et al.</i> ¹⁷²	cell-based glycan array	2024
Hodgson K. <i>et al.</i> ²³³	immunofluorescence	2024
Schmidt E.N., <i>et al.</i> ²²⁸	ELISA-bead assay mass spectrometry	2024
Lima G.M., <i>et al.</i> [†]	cell-based glycan array	2024

2.4.2: Siglec expression by primary splenocytes cells

With a few minor exceptions, Siglecs are generally expressed by immune cells. Using traditional flow cytometry analysis, the Siglec expression profiles of six immune cell types were described. The results from this investigation largely agree with the literature, however we were unable to observe expression of some Siglecs such as Siglec-1, Siglec-6, Siglec-11, and Siglec-15. However, this is most likely due to the types of immune cells that were examined in this study. Macrophages are reported to express Siglec-1, Siglec-11, and Siglec-15 and dendritic cells are

known to express Siglec-1 and Siglec-15, as both of these cell types were excluded from this study for practical reasons (relatively low abundance in spleen, availability and robustness of fluorophores available, availability of robust antibodies, *etc.*) so it is not surprising that no expression of these Siglecs was observed⁶. Siglec-6 is expressed by memory B cells and mast cells (excluded from this investigation). Memory B cells make up a small percent of the total B cell population (less than 5%)¹²⁸; moreover, Siglec-6 is expressed at relatively low levels on memory B cells. Without the markers to specifically identify memory B cells and the low expression of Siglec-6, the Siglec-6⁺ B cells likely get drowned out and thus Siglec-6 was not observed on B cells^{6, 128}. Using markers to identify memory B cells, Siglec-6⁺ B cells were identified in Chapter 4. Siglec-8 is expressed by eosinophils, basophils, and mast cells, which were not investigated in this study for the same reasons described above⁶. Siglec-11 is also expressed by microglia; however, microglia are generally found in nervous tissue and would not be expected in the spleen. Overall, these results agree with what was previously reported⁶.

2.4.3: Siglec ligands expressed by primary splenocytes

Using the Siglec-Fc probes, Siglec ligands were investigated on six types of immune cells isolated from human spleens and some interesting trends emerged. Siglec-1, CD22, and Siglec-7 ligands were found on all of the immune types tested. CD33 bound the majority of the immune cell types (Neutrophils, Monocytes, mature NK cells, and CD8⁺, T cells) tested. Siglec-5 and Siglec-10 bound weakly to B cells and monocytes and the remaining Siglecs bound minimally to monocytes. As the general ligand preferences with respect to regiospecificity for some Siglec family members is known, based on which Siglecs bind to which immune cell the types of sialosides on each immune cells can be proposed. For instance, CD22 is a strict $\alpha(2 \rightarrow 6)$ binder and its binding to all cell types is evidence of $\alpha(2 \rightarrow 6)$ linked sialosides being present on all cell types²⁰¹. On the other hand, Siglec-1 is generally regarded as a $\alpha(2 \rightarrow 3)$ binder and the binding the Siglec-1 to all cell types suggests that $\alpha(2 \rightarrow 3)$ linked sialosides is likely present on all immune cells¹⁶⁰. Other Siglecs such as CD33, Siglec-7, Siglec-5 can recognize most or all sialic acid motifs

making it difficult to speculate which types of sialosides/glycans may be present³⁷. Initially, it was surprising to observe that R119A Siglec-5-Fc bound to neutrophils. However, after consulting the literature, it was found that Siglec-5 was previously reported to bind Heat Shock Protein-70 (HSP70) independent of its conserved arginine residue (Arg119), however as a cytoplasmic protein, it is not obvious why this would be the case²³⁴. Currently it is unclear if Siglec-5 is binding HSP70 on neutrophils, however this could be the subject of future investigations.

2.4.4: Relationship between *cis* and *trans* Siglec ligands

It is also worth noting that more often than not, cells which express a Siglec tend to also express the ligands for the same Siglec. For instance, B cells express CD22, Siglec-5, and Siglec-10 and ligands of all these Sigeles were found on B cells. The presence of both the Siglec and Siglec ligands on the same cells suggest that the Siglec is likely masked, and this has been demonstrated with CD22 on B cells^{162, 235}. With this in mind, the expression of a Siglec can be plotted vs the abundance of Siglec ligands to predict which Sigeles are most available to engage in *trans* binding (**Figure 2.28**). Using this analysis, CD22, Siglec-5, and Siglec-7 are likely masked on B cells, neutrophils, and mature NK cells, respectively, From the other perspective, Siglec-9 is expressed by neutrophils, mature NK cells and monocytes, but weak to no binding of the Siglec-9-Fc was observed on these cells. This suggest that Siglec-9 may be more available to bind in *trans*. The relationship between *cis* and *trans* binding may be incredibly important for understanding Siglec biology and their roles in various pathophysiologies. For instance, Siglec-9 is known to antagonize Fcγ receptor-based cellular activation¹³² and if it is readily available to bind in *trans*, it would be an excellent target for pathogens and cancers to target towards pacifying neutrophils, NK cells and monocytes. Indeed, Siglec-9 has been proposed as a target for cancer cells to pass immune checkpoints^{236, 237}. However, it should be recognized^{236, 237} that this investigation probed Siglec expression and Siglec ligand expression on healthy cells. How the balance between *cis* and *trans* ligands changes in different disease states could be very interesting. Indeed, it has been proposed that cancerous cells increase the amount of Siglec ligands on their

surface to protect themselves from dangerous immune cells⁹⁵. Could a similar glycan remodeling be applicable in other disease conditions such as bacterial/viral infections or neurodegenerative diseases? If so what are the biological consequence(s) of this glycan remodeling? These questions could be the subject of future investigations now that a baseline has been described with respect to Siglec and Siglec-ligand expression.

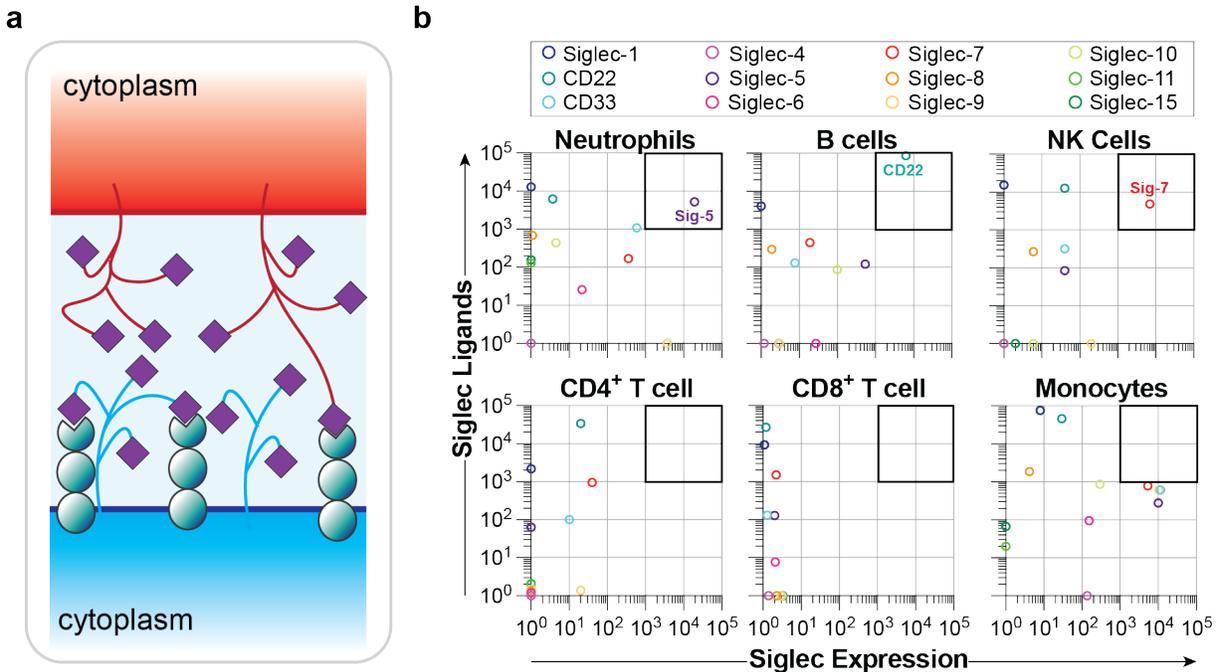


Figure 2.28: Masking potential of Siglecs on primary human immune cells. **a**, depiction of the balance between *cis* and *trans* Siglec-ligand binding. **b**, Masking potential of Siglec expressed immune cells under healthy physiological conditions.

2.4.5: Siglec-Fcs combine with engineered cell lines for describing Siglec-ligands

Cell-based glycan arrays are an excellent approach for identifying which cell types bear Siglec ligands. However, cell-based glycan arrays are less useful when trying to identify which specific glyco-epitope(s) are the ligands for a specific Siglec. To better describe Siglec ligands using cell-based glycan arrays, cell lines can be engineered to express glycan modifying enzymes such as enzymes from the CHST family or CMAH. Alternatively, key biosynthetic enzymes such as MGAT1, COMSC, *etc.* can be knocked-out to narrow down the class of glycan a Siglec prefers, Moreover, a combination of overexpressing glycan modifying enzymes while simultaneously knocking-out glycan modifying enzymes can have synergistic effects for identifying Siglec ligands.

In the case of CD33, which has a rather weak affinity for non-sulfated glycans, CHST1 was overexpressed to increase the sensitivity of the CD33-Fc. Following this, MGAT1, COSMC, and POMGNT1 were systematically knocked-out on top of each other. CD33 showed an increase in binding to cells when MGAT1 was knocked-out. This may be due to their being an excess of other glycans such as *O*-glycans in the absence of *N*-glycans. Many of the glycan-elongating biosynthetic equipment is shared between different types of glycans. Accordingly in the absence of *N*-glycans, more *O*-glycans could be extended or be sialylated to produce CD33 ligands. Alternatively, the difference was only 1.7-fold which could be accounted for by clonal variation which arose during the knocking-out of MGAT1. However, more MGAT1^{-/-} clones would be needed to rule this out. CD33 showed a significant decrease in binding when both MGAT1 and COSMC were knocked-out suggesting that *O*-GalNAc glycans are the major ligands for CD33 which is in line with previous observations²³⁸. However, CD33 binding to the double knock-out was still present. It was posited that *O*-mannose *O*-glycans were responsible for the remaining binding to the double knock-out cells. To address this, POMGNT1 was knocked-out on top of MGAT1 and COSMC. Binding was reduced compared to the double knock-out yet there was still significant CD33 binding to the cells compared to the arginine mutant. Taken together, these results suggest that CD33 ligands are primarily $\alpha(2\rightarrow3)$ linked 6-*O*-Sulfo-Gal *O*-GalNAc sialosides; however, as the triple knock-out was not able to completely abolish binding, it suggests that other types of glycans, such as *O*-fucose *O*-glycans, GPI-anchors, gangliosides or glyco-RNA²³⁹, could be potential ligands for CD33. Enzymes such as RNAses²³⁹ and phospholipases²⁴⁰ could be used to treat these triple knock-out cells to resolve if glyco-RNA or GPI-anchors are facilitating CD33 binding. Moreover, treatment of the triple knock-out cells with GENZ could help resolve if CD33 is binding gangliosides⁶¹. Further resolving which types of glycans can facilitate binding to CD33 could be an excellent extension of this work. A better understanding of CD33 ligands will serve to further describe the roles of CD33 in AD.

2.4.6: Improvements to the Siglec-Fc chimeras

Siglec-Fc chimeras were first used in the early 1990's. Since then, there have been a few iterations that each served to improve the construct. In 2012, Gieseke *et al.* biotinylated a soluble Siglec monomer so that it could be complexed with streptavidin enabling the formation of a tetrameric probe. However, the chemoenzymatic-addition of biotin requires additional steps and reagents to increasing the effort in the process. Moreover, chemical modifications are not always 100 percent effective, and the degree of modification could vary between preparations. Ideally, if biotin or a biotin mimic could be encoded genetically these obstacles could be overcome. In 2018, Gonzalez-Gil *et al.* fused the extracellular domains of a Siglec with cartilage oligomeric matrix proteins forming a pentameric probe. The advantage of this probe was that it did not require a secondary to increase the multivalency as the COMP domains spontaneously form pentamers¹³⁹. However the lack of biological/biochemical handles (e.g. IgG domain or Strep-Tag) limits the approaches the construct can be used in and perhaps this is one of the reasons that neither probe has gained popularity in the scientific community. Regardless, both of these excellent probes removed the IgG-Fc portion of the Siglec-Fc to minimize non-sialoside specific interactions while maintaining multivalency.

Building on all these works, Rodrigues *et al.* designed and produce a new version of the Siglec-Fc which combines the best of each of these constructs. The mutations in the Fc portion of the Siglec-Fc decrease the interactions between the Siglec-Fc and Fcγ receptors, improving their utility for probing Siglec-ligands on primary cells and tissues while still enabling anti-hIgG to be used as a secondary. The Siglec-Fc was further improved by applying the work of Schmidt and Skerra²¹³ as a Strep-Tag II was added to the Siglec-Fc construct, enabling the Siglec-Fcs to be purified under more gentle conditions compared to protein A or G columns, and served as a genetically encoded handle which can be grasped with Strep-Tactin. Lastly, the addition of the TEV sequence allows for the Siglec-Fc to be easily modified to a monomeric Siglec fragment for

studying monomeric interactions. All of these changes to the initial Siglec-Fc construct improve its functionality and its versatility.

Further improvements were made with respect to expression, storage and activity. Producing Siglec-Fcs from CHO CMAS KO cells not only improved the yield by 4-fold on average but also improved the activity of the Siglec-Fcs by 2-3-fold. Moreover, storing freeze dried Siglec-Fcs extended the shelf-life from weeks to months or even potentially years under relatively benign storage conditions. The improved yield of the Siglec-Fcs produced from CMAS KO cells may be due to the lack of sialosides on the CHO cells. This would prevent the Siglec-Fcs from binding to the surface CHO cells, releasing the Siglec-Fcs to the media allowing them to be easily isolated. This hypothesis is supported by the fact that the arginine mutant version of the Siglec-Fcs almost always expresses 2-4-fold higher than the wildtype Siglec-Fc in wildtype CHO cells and that there was not a significant difference in the yields between the arginine mutants expressed by CHO WT and CMAS KO cells.

The improvement in Siglec-Fc binding may be due to the lack of *cis* binding between the Siglec-Fcs. Siglecs themselves are glycoproteins, and some Siglecs have as many as six glycosylation sites on the first three extracellular domains alone²²⁴. Moreover, the Fc region can also be glycosylated²⁴¹. When the Siglec-Fcs are forced into close proximity by the secondary (either Strep-Tactin or anti-hIgG) the Siglecs may engage in *cis* binding preventing binding in *trans* (**Figure 2.29a**). Alternatively or additionally, the formation of complexes may serve as multivalent sialoside bearing complexes which could also compete with the ligands of interest such as the sialoside on a cell in a cell-based glycan array. Both of these factors could explain why Siglec-Fc binding is improved when Siglec-Fcs are produced from CMAS KO CHO cells.

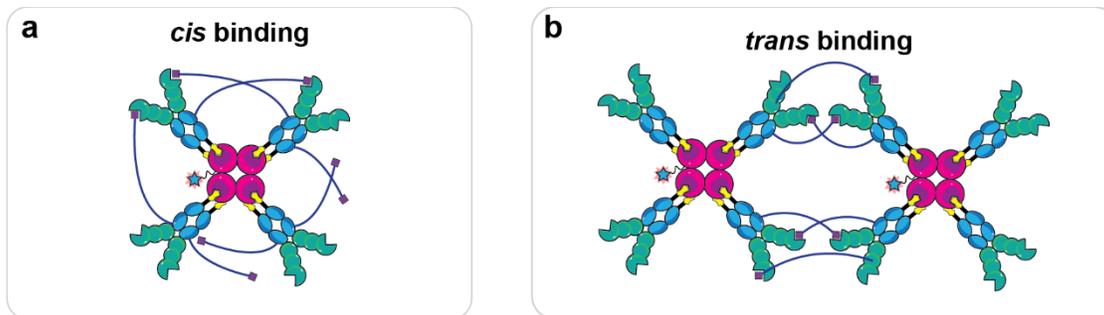


Figure 2.29: Masking of Siglec-Fcs by Siglec-Fcs.

2.5: Conclusion

Siglecs play important roles in many pathophysiological conditions. Understanding how these immunomodulatory receptors are exploited by pathogens and how they can be modulated therapeutically, begins with understanding Siglec ligands. Siglec-Fc chimeras are versatile tools which can be used in a variety of approaches to describe Siglec ligands. This chapter demonstrates the utility of these constructs and highlights works that have applied Siglec-Fcs to describe Siglec ligands. Moreover, improvements were made with respect to expression/yield and activity. Siglec-Fcs have, are and will continue to be one of the most versatile and sensitive tools to describe Siglec ligands.

2.6: Methods

Human Samples. All experiments involving human blood samples and placental sample collection were approved by the human research ethics board (HREB) biomedical panel at the University of Alberta.

Cell Culture and Growth Medium. CHO Flp-In cells (ATTC) were cultured in DMEM/F12 Media (Gibco) supplemented with 5% (V/V) fetal bovine serum (Gibco), penicillin (100 U/mL) and streptomycin (100 µg/mL). U937 cells were cultured in RPMI 10% (V/V) fetal bovine serum, penicillin (100 U/mL) and streptomycin (100 µg/mL). Cells were grown at 37 °C and 5% CO₂ in tissue culture flasks (VWR).

Cloning of Siglec Constructs. The genes for Siglec-1-11 and 15 were synthesized by *GeneArt* (Thermo Fisher) and designed with a 5' *NheI* and 3' *AgeI* site immediately before the start codon and stop codon respectively. When appropriate, silent mutations were introduced to remove

internal *NheI* and *AgeI* cut sites. Each gene was cut out from the initial vector via a double restriction enzyme digest with *NheI* and *AgeI* (Thermo Fisher). Digestion was confirmed by running a 1 % agarose gel after which the digested gene was excised from the gel and a gel purification was performed using a GeneJET Gel Extraction Kit (Thermo Fisher) to isolate the double digested gene. The digested gene was ligated into *NheI* and *AgeI* digested pCDNA5 vector and then transformed into chemically competent *Escherichia coli* DH5 α . Colonies were then picked and transferred to liquid culture (lysogeny broth plus 50 μ g/mL ampicillin) and left to grow overnight in a shaking incubator at 37 °C. The plasmid was then isolated from the bacterial culture using GeneJET Plasmid Miniprep Kit (Thermo Fisher) and the successful incorporation of the Siglec gene into pCDNA5 was validated by restriction enzyme digestion and Sanger sequencing.

Site Directed Mutagenesis. Mutations were introduced into each Siglec via gene overlap extension PCR mutagenesis (primers were ordered through IDT). Mutagenesis primers can be found in Supplementary Tables 2 and 3 for the critical arginine mutants of the entire Siglec family and the additional Siglec-6 mutants respectively.

Stable Transfection. CHO Flp-In cells were initially cultured as described above. The transfection began by seeding 400,000 cells in a 6-well cell culture dish. The next day, 0.2 μ g of the desired Siglec DNA in pCDNA, 2 μ g of pOG44 plasmid and 7 μ g of Lipofectamine Plus reagent (Thermo Fisher) were added to 250 μ L of Opti-MEM (Gibco) and the mixture was incubated at room temperature for 15 min. Next, 8 μ L of Lipofectamine LTX reagent (Thermo Fisher) was added to the mixture; and the mixture was left at room temperature for 30 min. During the incubation, the seeded cells were gently washed with Opti-MEM. The transfection mixture was then added to the seeded cells and the cells were left in the growing conditions described above overnight. The next day the media was aspirated from the well and replaced with CHO growth media (described above). The cells were then selected over two weeks by gradually increasing the amount of Hygromycin B from 0.5 mg/mL to 1 mg/mL, replacing the media every other day.

Liposome Production. Stock lipids (DSPC, Cholesterol, PEG₄₅-DSPE) solution were prepared by solvating the lipids in chloroform. An appropriated volume of each lipid solution was transferred to a glass test tube to reach the desired mol% of the component. The chloroform was then removed under N₂ gas to yield a lipid thin film. Once dry, 100 μ L of DMSO was added to the lipid containing test tube and then the functional lipids (ganglioside, AF647-PEG₄₅-DSPE, Siglec ligand-PEG₄₅-DSPE, pHrodo-PEG₄₅-DSPE, NGL) was added from their respect DMSO stock solution. The lipid mixture was then stored at -80 °C until completely frozen at which point excess DMSO was removed via lyophilization. Dry lipids were then stored at -80 °C until they were extruded. Lipid suppliers can be found in Supplementary Table 4.

Liposome Extrusion. Dry lipids were then hydrated in 1 mL of PBS. The lipids were then sonicated in a bath sonicator in a cycle of 1 min on, 5 min off for a total of 5 cycles. Liposomes were then extruded using an Avanti mini extruder first using an 800 nm filter and then a 100 nm filter yielding liposomes that were 130 \pm 35 nm (Supplementary Table 5) measured using a Malvern Panalytical Zetasizer Nano S. Liposomes were stored at 4 °C.

Cell-based Glycan Array. 200,000 cells/well were plated into a 96-well U-bottom microplate and centrifuged at 300 x g for 5 min. The cells were then resuspended in 30 μ L of 100 nM (α hlgG) or 250 nM (Strep-Tactin) Siglec-Fc-Secondary complex and incubated on ice for 30 min. Unbound Siglec-Fc complex was then removed via centrifuged at 300 x g for 5 min followed an addition wash with FACS solution. The cells were resuspended in FACS buffer solution and were then ready for analysis by flow cytometry.

Flow cytometry. Flow cytometry measurements were collected on a 5-laser Fortessa X-20 (BD Bioscience). All the resulting data was analyzed using FlowJo (10.5.3) software (BD Biosciences).

Siglec-Fc Production. Siglec-Fc constructs designed by Rodrigues *et al.*³⁷ were used in this work. Siglec-Fc constructs were expressed using wildtype CHO cells in cell culture media as described above with 1 % HEPES. Cells incubated as described above for one week post confluency. The supernatant was the harvested and stored at 4 °C.

Siglec-Fc Purification. The purification was heavily inspired by Rodrigues *et al.*³⁷ The purification of the Siglec-Fcs from supernatant began with Ni²⁺ affinity chromatography using an AKTA FPLC equipped with a HisTrap Excel column (GE). The column was equilibrated with fifteen column volumes (CVs) of equilibrium buffer (500 mM NaCl, 20 mM NaPO₄H₂, pH 7.4). The Siglec-Fc containing supernatant was then loaded in its entirety onto the column. The column was then washed with fifteen CVs of wash buffer (500 mM NaCl, 30 mM imidazole, 20 mM NaPO₄H₂, pH 7.4). The Siglec-Fc was then eluted in twenty 1 mL fractions with twenty CVs of elution buffer (500 mM NaCl, 500 mM imidazole, 20 mM NaPO₄H₂, pH 7.4). Fractions containing Siglec-Fc were compiled and then prepared for the next stage of purification by diluting the fractions 1:4 with buffer W (100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0). A Strep-Tactin column (IBA life sciences) was washed with fifteen CVs of buffer W. The Siglec-Fc-buffer W solution was then loaded in its entirety onto the column. The bound Siglec-Fc was then washed with fifteen CVs of buffer W. Siglec-Fc was eluted from the column with fifteen column volumes of buffer E (100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 10 mM *d*-Desthiobiotin, pH 8.0). Fractions containing Siglec-Fc were pooled and dialyzed against PBS to remove any *d*-Desthiobiotin. After dialysis, Siglec-Fcs were concentrated using an ultra-centrifugal device (30 kDa MWCO) and aliquoted into 5 µg aliquots and frozen. The Siglec-Fcs were then lyophilized overnight, and the lyophilized Siglec-Fcs were stored at -20 °C.

Improving Siglec-Fc storage. Many who work with Siglec-Fcs acknowledge that Siglec-Fcs are not very stable in solution and tend to lose activity with time (in as little as a week) even when stored at 4 °C. To overcome this, Siglec-Fcs are typically aliquoted and stored at -80 °C. While this is an effective solution, it can be an expensive way to store the Siglec-Fcs. Moreover, if Siglec-Fcs are stored at -80 °C, when the Siglec-Fcs are shipped they need to be on dry ice to prevent the Siglec-Fcs from thawing in transit. This makes sharing these constructs impractical, difficult, and expensive. As an alternative, we aliquot and lyophilize the Siglec-Fcs. Once freeze dried, the

Siglec-Fcs can be stored at more mild conditions such as -20 °C and can even be shipped at room temperature while retaining their activity when hydrated.

Siglec preparation for direct binding assay. Monomeric Siglec fragment was prepared as described previously^{37, 79, 128} but in short, Siglec-Fc were produced as described above from CHO Lec1 cells and purified as described above. The purified was then treated with a 10-fold molar excess of TEV protease overnight at room temperature. TEV and the Fc portion were then removed using a HisTrap Excel column where the monomeric Siglec fragment was in the flow through. The Siglec fragment was then treated with EndoH at 37 °C overnight. The EndoH was then removed the same way as TEV using a HisTrap Excel column. The Siglec fragment was dialyzed against 200 mM NH₄OAc. Following the buffer exchange the Siglec fragment was then concentrated by ultra-centrifugal device until the concentration was 1 mg/ml and stored at 4 °C.

Direct ESI-MS Binding Assay (Performed by the Klassen Group). The affinities of oligosaccharides oligosaccharides (L) for Siglec fragments were measured by the direct ESI-MS assay.¹ A reference protein (P_{ref}) was added to the ESI solutions in order to correct mass spectra for any nonspecific binding that occurred during the ESI process.² The dissociation constant (K_d) was calculated from the total abundance (Ab) ratio (R) of the ligand bound (PL) to free Siglec-1 fragment (P) ions (equation 4) measured by ESI-MS for solutions of known initial concentration of Siglec-1 fragment (P₀) and ligand (L₀), equation 5:

$$R = \frac{Ab(PL)}{Ab(P)} = \frac{[PL]}{[P]} \quad (1.1)$$

$$K_d = \frac{[P][L]}{[PL]} = \frac{[L]_0}{R} - \frac{[P]_0}{R + 1} \quad (1.2)$$

The reported K_d values correspond to average values measured at 3.6 μM of Siglec-1 fragment and 20, 40, 80, and 140 μM of each oligosaccharide tested. Direct ESI-MS measurements were performed with nanoflow ESI in positive ion mode (voltage ~1 kV) on a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific). Automatic gain control (AGC) target, the maximum inject

time, capillary temperature and S-lens RF level were set to 1×10^6 , 100 ms, 150 °C and 100 °C, respectively. The resolution was 17500 at m/z 200. Data acquisition and processing were carried out using Xcalibur (Thermo Fisher, version 4.1).

Unmasking/Neuraminidase treatment. Cells were harvested and resuspended in PBS (pH 7.4) media with either Neuraminidase A or S (0.3 mg/mL). The cells were then placed in a 37 °C shaking incubator for 1 h. The cells were then washed with complete media and then the liposome binding assay was performed as described above.

then lyophilized overnight, and the lyophilized Siglec-Fcs were stored at -20 °C.

ELISA. Our ELISA approach was carried out similar to previous work by Rapoport²⁴² and Yamakawa¹⁴³. Ganglioside ethanol solutions (10 μ M) were prepared and then transferred in 50 μ L increments to a 96 well microplate. The ethanol was removed by drying the plates overnight at room temperature. The plates were then washed with PBS, dried, and then blocked with 5 % (m/V) BSA PBS for 1 h. The plates were then washed with PBS and then 2 μ g/mL Siglec-Fc precomplexed to Strep-Tactin (2 Siglec-Fc :1 Strep-Tactin monomer) horse radish peroxidase was added to the microplate. The complex was incubated with the plate at room temperature for 2 h. The unbound complexes were then removed by washing in PBS and the plate was developed with Sera care TMB solution. The amount of binding was then quantified by using the background (microplate well with no ligand) subtracted absorbance at 450 nm using a Molecular Devices SpectraMAX ® iD5.

Bead Assay. Pierce™ Streptavidin Magnetic Beads (Thermo Scientific) were blocked with 2 % (m/V) BSA on ice for 1 h. Siglec-Fc in PBS was added then added to the bead solution such that the final concentration was 25 μ g/mL. The Siglec-Fc was complexed to the beads for 1 h on ice. Excess Siglec-Fc was removed by washing the beads with 2 % BSA solution and 50 μ M liposome or 50 μ g/mL (protein concentration) EV solution was then added to beads. The beads were incubated with the liposomes/EVs for 30 min at 37 °C. The beads were then washed with 2 %

BSA solution and flow cytometry was used to measure the binding between the beads and the liposomes.

Isolation of White Blood Cells from Human Spleen. Spleen tissue was cut into pieces approximately 1 cm³ and placed into a petri dish containing RPMI supplemented with FBS (10 % V/V) and penicillin (100 U/mL) and streptomycin (100 µg/mL) that was chilled to 4 °C. White blood cells were then separated from the rest of tissue using a *Miltenyi gentleMACS Dissociator*. The tissue homogenate was then passed through a tea strainer and the filtrate was centrifuged at 400 x g for 10 min. The supernatant was removed, and the cell pellet was resuspended in 4 °C red blood cell lysis solution (StemCell Technologies) and the cells were incubated for 10 min. The white blood cells in the lysis solution were then passed through a 75 µm cell strainer and the filtrate was then diluted 5-fold with supplemented RPMI and centrifuged for at 400 x g for 10 min. The pellet was washed three more times with supplemented RPMI. Resuspend the cells in DMSO/FBS (1:9 V/V) and freeze cells using a Mr. Frosty™ at -80 °C for two days. The cells were then moved into liquid nitrogen and stored until needed.

Lentivirus Production. Various lentiviruses were produced as previously reported by Bhattacharjee, A. *et al.*²⁴³. Briefly, 1,000,000 HEK293T cells were plated in a 6-well dish containing 1.5 mL of DMEM growth medium (Gibco) containing 10 % (V/V) fetal bovine serum (FBS; Gibco), 100 U/mL penicillin (Gibco) and 100 µg/mL streptomycin (Gibco). 24 h later, a mixture of 625 ng RP18, 625 ng RP19, 1250 ng hSiglec-6 vector, 7.5 µL TransIT®-LT1 Reagent (Mirus Bio), and Opti-MEM media (Gibco) was added to the HEK293T cells. Cells were incubated with this transfection mixture at 37 °C, 5 % CO₂ for 72 h. Following transfection, the cell supernatant was harvested and concentrated using Lenti-X concentrator reagent (Takara Bio) following the manufacturer instructions.

Viral Transduction. 150,000 cells were plated in a 24-well plate in 250 µL of growth media. A range of 10X concentrated lentivirus (1, 2, 5, 10 µL) was added to the corresponding wells and incubated for approximately 8 h at 37 °C, 5 % CO₂ in a tissue culture incubator. After incubation,

the media in each well was topped up to 750 μ L. Three days post transduction, 200 μ L of cells were plated in a 96-well U-bottom plate and centrifuged at 300 x g for 5 min. The cell pellets were resuspended in 150 μ L of flow cytometry buffer (HBSS containing 1 % (V/V) FBS, 500 μ M EDTA) and the titer of each virus was determined by measuring the mAmetrine+ cells in each well by flow cytometry. The mAmetrine+ cells, ranging from 0.5 to 5 %, were re-plated in 6-well plates. The mAmetrine+ virally transduced cells were selected for using 300 μ g/mL zeocin until the mAmetrine+ population was \geq 95%.

Data Collection Software. Flow cytometry data was collected with BD FACSDiva™ software Version 8.0.1 and analyzed with FlowJo LLC. Version 10.5.3.. Xcalibur (Thermo Fisher Scientific, version 4.4) was used for mass spectrometry data acquisitions. ELISA was collected using Molecular Devices Soft Max Pro 7.0.3.

Statistical Analyses. For datasets comparing only two conditions, a Student's *t*-test was used. When datasets had three or more conditions a one-way ANOVA was used. All statistical analysis was carried out using GraphPad Prism version 8.4.

Chapter 3: Glycolipid bearing liposomes as tools for Discovering Siglec Ligands

3.1: Acknowledgements

Portions of this chapter were published as:

Schmidt, E.N.; *et al.* Siglec-6 mediates the uptake of extracellular vesicles through a noncanonical glycolipid binding pocket, *Nat. Commun.* (2023).

Schmidt, E.N.; *et al.* Dissecting the abilities of murine Siglecs to interact with gangliosides, *J. Biol. Chem.*, (2024).

I would like to thank Jaesoo Jung for assistance in developing the CHO cell lines used in this chapter, Xue Yan Guo for assistance with performing liposome binding assays and ELISAs, Dr. Mirat Sojitra under the supervision of Prof. Ratmir Derda for performing LIGA experiments, Dr. Ling Han, Dr. Elena Kitova, Dr. Duong T. Bui under the supervision of Prof. John S. Klassen Group for performing mass-spectrometry-based analysis of gangliosides incorporation and K_d analysis between Siglecs and ganglioside oligosaccharides.

3.2: Introduction

3.2.1: Glycolipids as Siglec ligands

Siglecs are a family of immuno-modulatory cell surface receptors that recognize sialylated glycan ligands¹²⁵. A leading hypothesis for the role of Siglecs in maintaining immune homeostasis is that binding to sialylated glycans represents a form of *self*-recognition that operates as an immune checkpoint⁶. Sialic acid residues decorate proteins and lipids, both of which can act as Siglec ligands³⁵. In all cases, sialoside recognition by Siglecs is critically dependent on a conserved canonical arginine residue in their N-terminal V-set domain that forms an ionic interaction between the cationic guanidinium of the essential arginine in Siglecs and the anionic carboxylate of the sialic acid²⁴⁴. In addition to their physiological roles in human health, Siglecs also play key roles in pathophysiological conditions, as they can be exploited by viruses^{9, 245, 246}, bacteria²⁴⁷, and cancers⁶ for immune evasion. Despite the growing understanding for the roles of Siglecs, there remains an incomplete description of their glycan ligands.

Due to the relatively weak affinity of Siglecs for their glycan ligands, Siglecs are often studied outside of their natural context using approaches that leverage multivalency. This is particularly true for glycolipids as there are challenges associated with studying the binding of Siglecs to glycolipids in a lipid bilayer²⁴⁸. Accordingly, the majority of Siglec–glycolipid interactions have been established through plate-based approaches in which a soluble, recombinant Siglec is used to probe glycolipids or neoglycolipids adsorbed on a hydrophobic surface^{140, 242, 249, 250, 251, 252, 253} or via glycan microarrays where the oligosaccharide portion of the glycolipid is covalently-linked to a surface¹³⁷. Using these approaches, Siglec-1, -4, -5, -7, -9, -10, and -15 have been reported to bind the oligosaccharide portion of gangliosides^{242, 249, 254, 255}, the major class of sialylated glycolipids in mammals. However, only some Siglec–ganglioside interactions have been validated in the context of a lipid bilayer, and even fewer have been studied in a biological membrane. Beyond the challenges associated with studying the two species in a lipid bilayer, deconvoluting Siglec–ganglioside interactions in a biological setting are further complicated by

the cell-type specific combination of gangliosides⁷⁵ and expression of multiple Siglecs on immune cells⁶.

Despite the challenges associated with studying Siglec-ganglioside interactions, several biological roles have been credited to them. For example, Siglec-1 (CD169/Sialoadhesin; Sn) on macrophages/dendritic cells mediates the uptake of viruses and extracellular vesicles (EVs) through binding gangliosides in their membrane^{9, 185, 245, 256, 257}. Siglec-4 (Myelin Associated Glycoprotein; MAG) on oligodendrocytes binds gangliosides on neurons to regulate neurite growth²⁵⁸. Moreover, Siglec-7 on natural killer cells recognizes gangliosides on cancer cells or EVs from cancer cells to prevent immune cell activation^{257, 259}. The fact that gangliosides are abundant on all mammalian cells makes understanding Siglec–ganglioside interactions important due to their potential of serving broader immunomodulatory roles.

3.2.2 Chemical characteristics of gangliosides

Gangliosides are amphipathic molecules with a water soluble glycan that protrudes from the cell surface and a lipid soluble ceramide portion that anchors it into the outer leaflet of the plasma membrane. Gangliosides are defined by a core carbohydrate backbone consisting of β -Galp-(1→3)- β -GalpNAc-(1→4)- β -Galp-(1→4)- β -GlcP linked to ceramide (**Figure 3.1**)²⁶⁰. Sialic acid can be linked to each of the monosaccharide units, except for the glucose⁷⁵. Glycolipids have a systematic nomenclature, and in the case of gangliosides, they are given a ‘G’ for belonging to the *ganglio* series of glycolipids, then a Greek prefix (M-mono, D-di, T-tri, *etc.*) to denote the number of sialic acid residues in the oligosaccharide, and lastly they are given an number based on the relative order a ganglioside would run on a thin layer chromatography (TLC) plate compared to other members of the *ganglio* series with smaller gangliosides (such as GM3) traveling further up the TLC plate than larger members (GM1a). While this is practical and historically accurate, in more simple terms, gangliosides with all four backbone monosaccharides

are denoted with a 1, with three are denoted with a 2, with two are denoted with a 3 and with just one of the monosaccharides it is denoted with a 4.

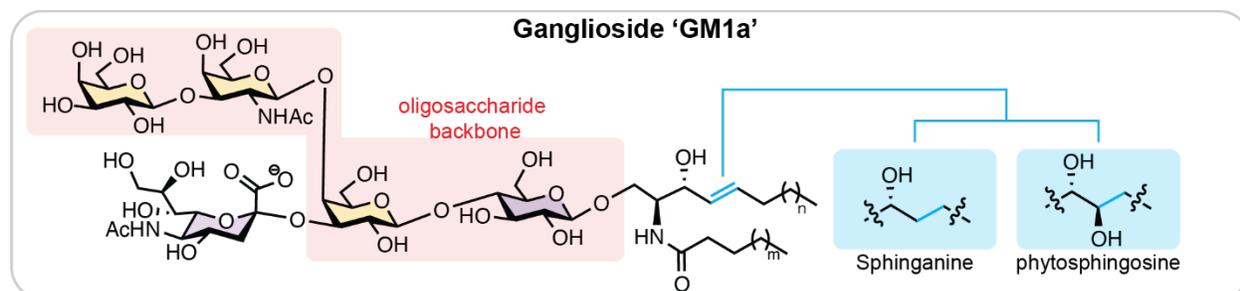


Figure 3.1: Chemical structure of ganglioside GM1a. The ganglioside carbohydrate motif (β -Galp-(1 \rightarrow 3)- β -GalpNAc-(1 \rightarrow 4)- β -Galp-(1 \rightarrow 4)- β -Glc) is highlighted in red. Possible modification of the hydrocarbon tails are highlighted in blue.

While there is diversity with respect to the oligosaccharide portion of the ganglioside, there can also be chemical diversity with respect to the ceramide portion. Most commonly, there is an unsaturated carbon chain. The alkene can be reduced to an alkane which is referred to as sphinganine. Alternatively, an alcohol can be added across the alkene generating a diol hydrocarbon chain, known as phytosphingosine. Moreover, the lengths of the hydrophobic chains can vary with respect to length and degree of unsaturation but are often between 18 and twenty carbons in length. As gangliosides are challenging to synthesize, gangliosides used for research purposes are typically isolated from animal products such as porcine, bovine brain, or bovine milk. The structure of the acyl chain varies between the different tissues and gangliosides isolated from the brain tend to be homogenous with respect to their acyl chains whereas gangliosides isolated from milk tend to be heterogenous with respect to their acyl chains.

3.2.3: Biosynthesis of gangliosides

The biosynthesis of gangliosides begins with the synthesis of ceramide from serine and activated fatty acids (**Figure 3.2**). Ceramide can be extended with either a galactose or glucose residue; if a galactose is added, it can be extended with an α (2 \rightarrow 3) sialic acid residue, yielding ganglioside GM4 or the galactose can be modified with a sulfate group yielding 3-O-sulfo-galactoceramide or sulfatide. More commonly, a glucose is added which can be further elongated yielding the rest of the ganglio-series glycosphingolipids/gangliosides.

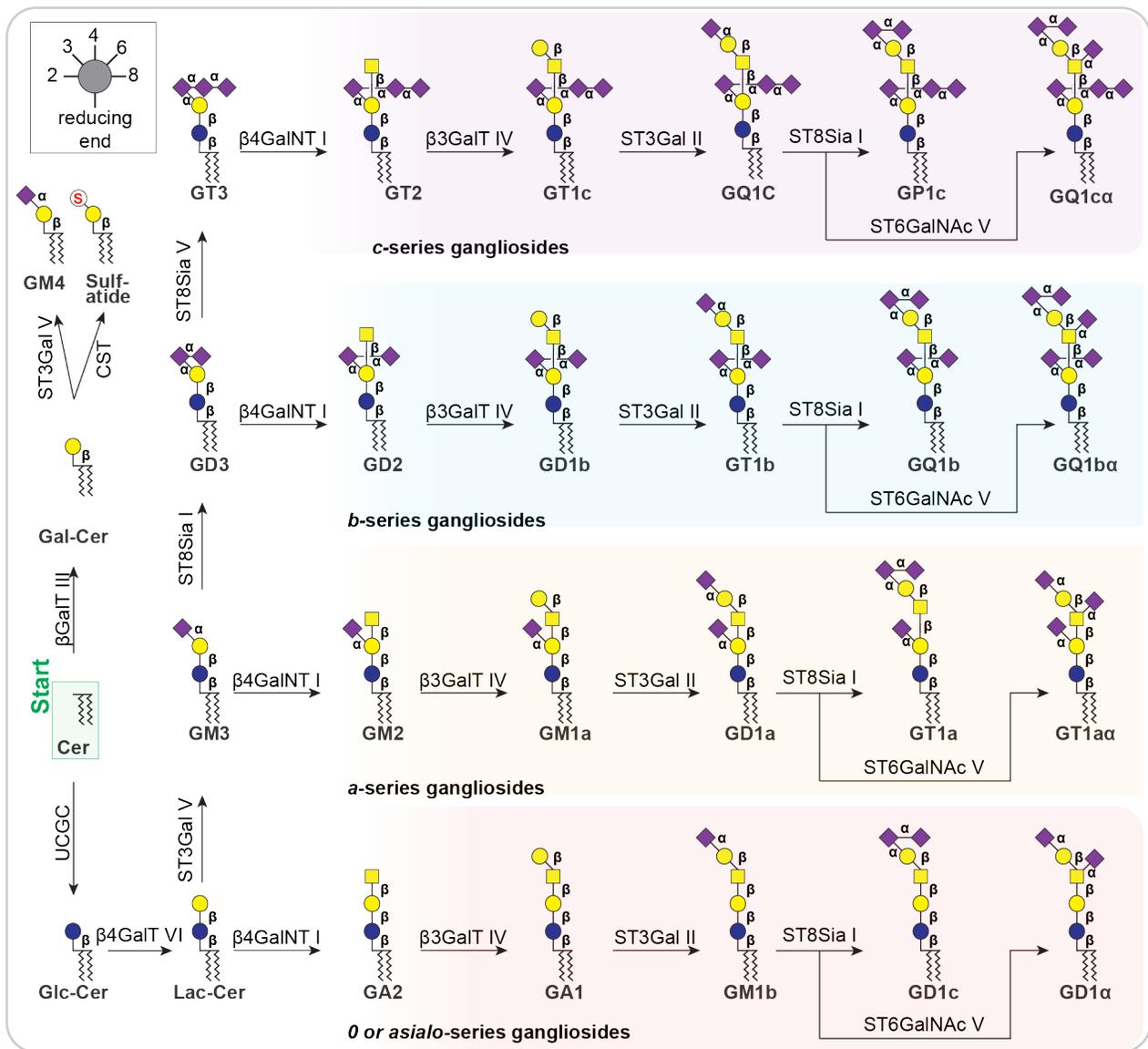


Figure 3.2: Biosynthetic pathway of gangliosides.

Key enzymes in the biosynthesis of gangliosides are UDP-glucose ceramide glucosyltransferase (UCGC) that catalyses the addition of the initial glucose residue to ceramide, $\beta 4\text{GalT I}$ that elongates lacto-ceramide to GA2 which is the precursor for most of the gangliosides, and ST3GalV that also acts on lacto-ceramide to produce GM3, another precursor for more complex gangliosides. Gangliosides can be divided into series based on the number of sialic acid residues added to the internal galactose. Gangliosides with no sialic acid on the internal galactose are said to be part of the *0 series* or *asialo series*. Gangliosides with one, two, or three sialic acids linked to the internal galactose are referred to as the *a*, *b*, and *c* series, respectively. All three

sialic acid motifs ($\alpha(2\rightarrow3)$, $\alpha(2\rightarrow6)$, and $\alpha(2\rightarrow8)$) are found in gangliosides, however $\alpha(2\rightarrow6)$ gangliosides are not common. Gangliosides are found ubiquitously across all cell types and tissues but are particularly abundant in the nervous system, where approximately 75% of sialic acid is presented from a ganglioside⁶². Ganglioside composition can vary between tissues; for instance, GD1b (27%), GD1a (26%), GT1b (24%) and GM1a (17%) make up the majority of gangliosides in the brain in a cell type dependant manner whereas in tissues such as the liver, GM3 is the most abundant ganglioside^{62, 250}. In addition to cell membranes, gangliosides can also be found in natural nanoparticles known as extracellular vesicles (EVs)^{128, 257, 261, 262}. The importance of gangliosides in human health is demonstrated by individuals who have genetic defects in the ganglioside biosynthetic pathways. Individuals who have defects in these genes unfortunately have a relatively low quality of life (

Table 3.1). Ganglioside biosynthesis is conserved between mammals, which makes animal models excellent tools to study human ganglioside biosynthesis and the importance of gangliosides in human health²⁵⁰.

Table 3.1: Effects of defects in ganglioside biosynthesis on health.

Biosynthetic Enzyme	Biosynthetic Result	Effects on Individual
- GM3 Synthase deficiency (ST3GalV ²⁶³)	-Only 0 series gangliosides	-Seizure susceptibility -Blindness -Developmental stagnation -Deafness -Protective effect from insulin resistance* -Ineffective CD4 ⁺ TCR dependent activation
GM2 synthase deficiency; (β4GalNTI ²⁶³)	-GM3 is the only a series ganglioside	-Spastic paraplegia -Deafness -Blindness -Seizure susceptibility
Tay Sachs Disease; (hexosaminidase-A ²⁶⁴)	-GM2 cannot be broken down and accumulates	-Muscle weakness -Deafness -Blindness -Seizure susceptibility

3.2.4: Roles of Siglec-ganglioside interactions in disease

In addition to being crucial in healthy development, gangliosides are also relevant in many host-pathogen interactions. For instance, envelope viruses such as HIV^{94, 183, 185, 245}, Ebola⁹, SARS-CoV-2²⁶⁵ among others take components of the host cells bilayer, including gangliosides, when they bud off from the cell. It has been proposed that the acquisition of gangliosides in this manner increases the virulence of viruses through exploiting Siglec-1. Bacteria, such as *Campylobacter jejuni*, use ganglioside like oligosaccharides to avoid detection by the hosts immune system²⁶⁶, however, in some cases the host can generate antibodies to these oligosaccharides that cross-react with endogenous gangliosides, resulting in an autoimmune condition known as Guillain-Barré syndrome²⁶⁷. Interestingly, it has been suggested that the development of Guillain-Barré syndrome may be due to a mutation in Siglec-10, which impairs binding to sialosides such as gangliosides¹⁷⁷. Moreover, it has been proposed that cancerous cells release ganglioside containing EVs to enhance their own survival, possibly by exploiting Siglec-7^{261, 262}. Siglec-ganglioside interactions are clearly important in many aspects of health and disease, yet the ability of most Siglecs to engage gangliosides has only been described to a limited degree. A more complete description of which gangliosides are ligands for which Siglecs will lead to a better understanding of Siglec-gangliosides in health and disease and possible novel therapeutic strategies.

3.2.5: Murine models to study human physiology and pathophysiology

Mice are among the most commonly used organisms for studying human physiological and pathophysiological processes²⁶⁸. However, challenges arise in using mice as a model organism when proteins involved in processes and pathways differ substantially between mice and humans. While some differences with respect to sialoside binding have been described between human and murine Siglecs, if ganglioside binding is conserved between humans and murine Siglecs has not been determined. Siglecs are cell surface receptors expressed by immune cells whose functions are regulated by their sialoglycan ligands^{224, 269, 270}. These functions include

regulation of immune cell signaling, internalizing extracellular cargo, and cell adhesion¹⁰⁰. Given that the immunomodulatory roles of Siglecs are directly tied to their ability to bind their ligands, it motivates a better understanding of their ligands, particularly as it relates to functional analogs between mice and humans. Importantly, in humans there are fifteen Siglecs, while in mice there are only nine Siglecs (**Figure 3.3**)⁶. Siglec-1, -2, -4, and -15 are well conserved between mice and humans with respect to their expression pattern and primary sequence similarity, but the rest are significantly divergent and classified as the CD33-related Siglecs⁶. Among the CD33-related sub-family, there are orthologs, which have resulted from new genes evolving from a common ancestral gene, and paralogs that result from a gene duplication event²⁷¹. For example, Siglec-7 and -9 are orthologs of Siglec-E, Siglec-8 is a paralog of Siglec-F, Siglec-10 is the ortholog of Siglec-G, and the rest are specific to their respective species²⁷².

In this chapter, a liposome formulation was optimized towards Siglec-ganglioside binding. Using this optimized liposome formulation, the human and mouse Siglec families were interrogated against a panel of nine commercially available gangliosides and similarities between human and murine Siglecs with respect to ganglioside binding are discussed.

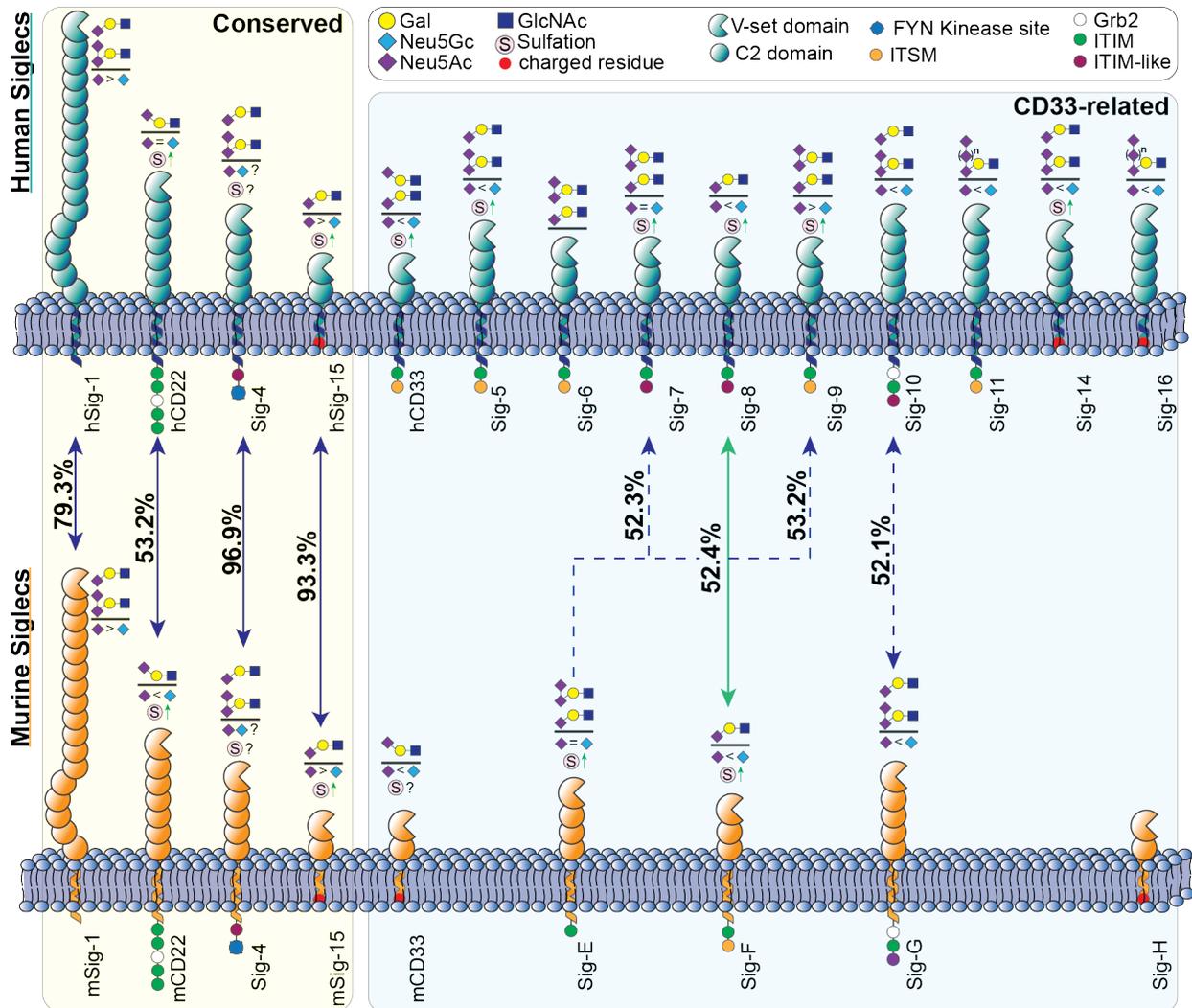


Figure 3.3: Comparison of human and murine Siglec families. The number to the left of the arrow denotes the sequence identity between the amino acid sequence of the V-set domain of the two Siglecs. A solid blue arrow denotes that the Siglecs are conserved between mice and humans. A dotted blue arrow denotes that the Siglecs are orthologs and a solid green arrow denotes that the Siglecs are paralogs.

3.3: Results

3.3.1: Cell-based Siglec-ganglioside binding assay

One obstacle in studying Siglec-ganglioside interactions is that immune cells often express more than one Siglec and many different gangliosides, making it challenging to deconvolute which Siglec binds to which ganglioside. To overcome this, we developed a cell line that only expresses one Siglec. This was achieved by stably transfecting a full length Siglec into a cell line that naturally lacks Siglecs. CHO cells were chosen because they naturally lack Siglecs and have previously been engineered with the FlpIn system making them easy to stably transfect²¹⁵. As for the ganglioside, liposomes were used to present the ganglioside from a more biologically relevant context. Moreover, known liposome formulations can easily be adapted to include gangliosides as well as fluorescent lipids allowing for detection via flow cytometry. This cell-based system places both the Siglec and the ganglioside in a more biologically relevant context compared to traditional approaches that use a Siglec-Fc and gangliosides adsorbed to a surface such as an ELISA. In our cell-based assay, the ganglioside bearing liposomes were added to Siglec-expressing cells and the binding of the liposomes to the cells was quantified by flow cytometry (**Figure 3.4a,b**).

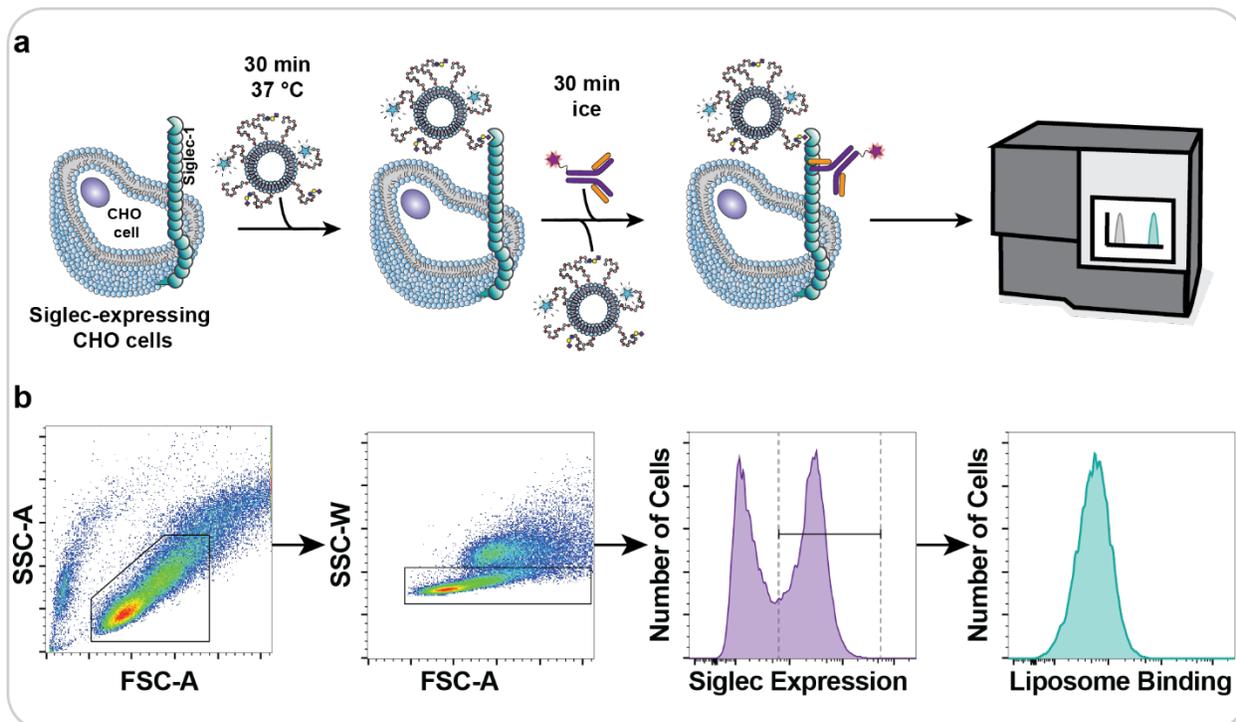


Figure 3.4: Schematic of cell-based assay used to probe Siglec-ganglioside interactions. **a**, Workflow of cell-based assay where fluorescently labeled ganglioside bearing liposomes are incubated with Chinese hamster ovary (CHO) cells expressing full length Siglecs measured by flow cytometry. **b**, Flow cytometry gating strategy for cell-based assay.

3.3.2 Optimization of liposome formulation for Siglec binding

Many different liposome formulations exist; however, how the liposome formulation influences the ability of the Siglec to engage a ganglioside was unknown. To address this, we decided to use a known liposome formulation as a starting point and further optimize the formulation parameters: polyethylene glycol (PEG) content, ganglioside content, cholesterol content, and bulk lipid structure for Siglec binding. We chose the FDA-approved *Doxil* liposomal formulation as the starting point for optimizing our glycolipid-containing liposomes. The initial formulation consisted of 57 mol% 1,2-distearoyl-*sn*-glycero-3-phosphocholine (DSPC), 38 mol% cholesterol, and 5 mol% polyethylene glycol 1,2-distearoyl-*sn*-glycero-3-phosphatidylethanolamine (PEG₄₅-DSPE MW 2000)²²⁹. This formulation was chosen as it has been reliably used to target Siglecs with high affinity glycan ligands linked to PEGylated lipids^{155, 273}. To detect liposome binding to cells by flow cytometry in a cell assay, 0.1 mol% AF647-PEG₄₅-DSPE was included in the formulation. As human Siglec-1 (hSiglec-1) is an established

ganglioside binder^{249, 274, 275}, it was used to optimize the liposomal formulation. Liposomes displaying a previously developed high affinity Siglec-1 ligand¹⁶⁰ linked to PEG₄₅-DSPE (**Figure 3.5a**) showed robust binding to wildtype hSiglec-1 and no binding when the canonical essential arginine (Arg116) was mutated to alanine in the cell assay (**Figure 3.5b**).

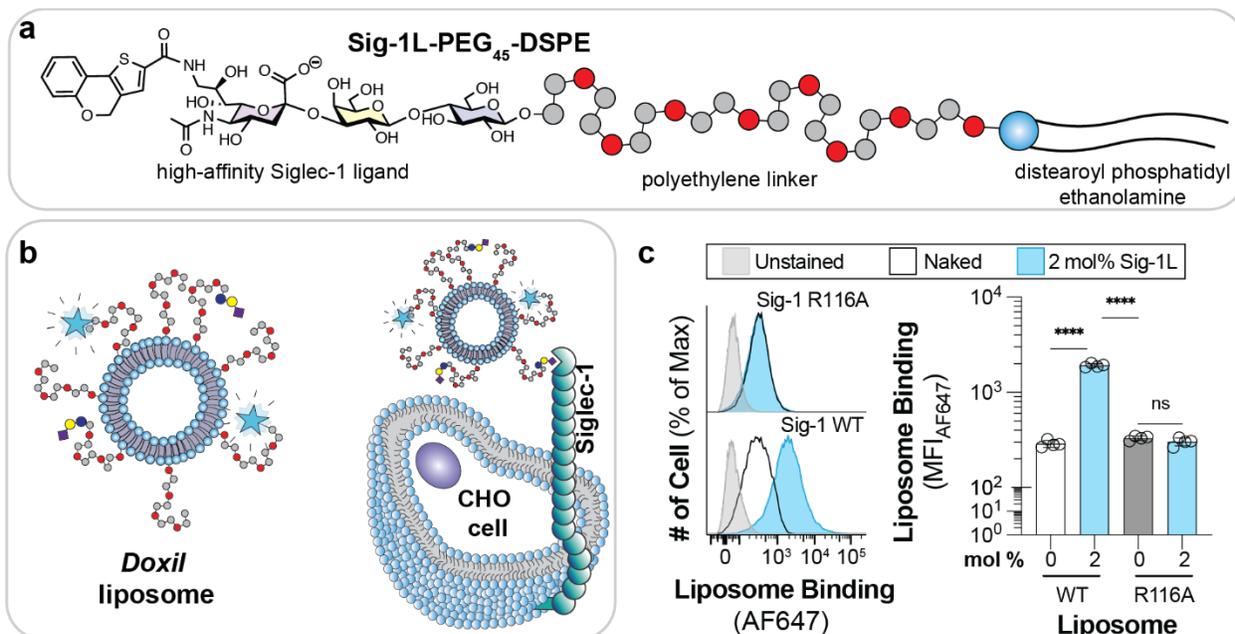


Figure 3.5: Liposomes formulated with a high affinity Siglec-1 ligand appended to PEG₄₅-DSPE engages wildtype hSiglec-1 expressing CHO cells. **a**, Chemical structure of high affinity Siglec-1-ligand appended to PEG₄₅-DSPE. **b**, representation of Doxil liposome formulation and depiction of CHO cell overexpressing hSiglec-1 binding to Siglec-1-ligand liposomes. **c**, Binding of liposomes formulated with 2 mol% Siglec-1-ligand to CHO cells expressing wildtype and R116A hSiglec-1. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. For panel d, a one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.5$; **** = $P < 0.0001$.

When the high affinity Siglec-1L-PEG₄₅-DSPE was replaced in this formulation with GM1a (3 mol%), a known ligand for hSiglec-1²⁷⁶, little to no binding of these ganglioside liposomes was observed to wildtype hSiglec-1 CHO cells (**Figure 3.6a**). To test if the hydrodynamic shell created by the PEG₄₅-DSPE²⁷⁷ prevents binding, the amount of PEG in the liposome formulation was systematically decreased, revealing ganglioside liposome binding (**Figure 3.6b**). No further increase in ganglioside liposome binding was observed when PEG₄₅-DSPE content was reduced below 0.5 mol%; however, non-specific liposome binding increased. Thus, 0.5 mol% PEG₄₅-DSPE was used in favour of the initial 5 mol% in liposomal formulations moving forward.

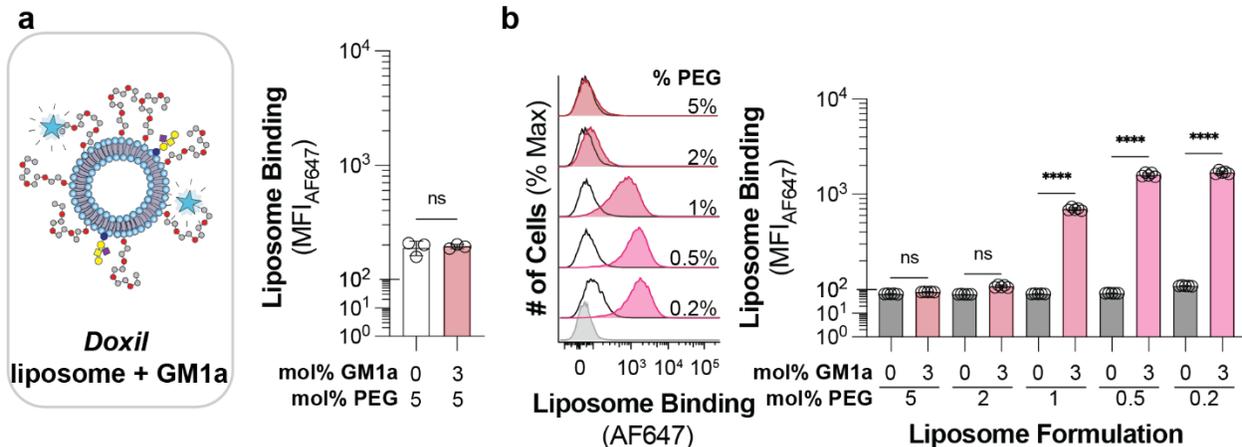


Figure 3.6: Binding of GM1a liposomes formulated with varying amounts of PEG45-DSPE to hSiglec-1 expressing CHO cells. **a**, Binding of Doxil liposomes formulated with 3 mol% GM1a to hSiglec-1 expressing CHO cells. **b**, Binding of GM1a liposomes to CHO expressing hSiglec-1 with varying amounts of PEG45-DSPE. For panel **b**, data is presented with representative flow cytometry histograms and a summary chart. Data was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from three (a) and five (b) technical replicates. A one-way ANOVA was used for statistical analysis. NS, $P > 0.05$; **** = $P < 0.0001$.

Now that ganglioside dependant liposome binding had been observed, we wanted to know if micellar GM1a could interfere with ganglioside liposome binding. To address this, a competitive assay was used where free/micellar GM1a (critical micellar concentration 20 nM²⁷⁸) was added to the cell mixture alongside ganglioside liposomes bearing GM1a. A decrease in liposome binding would be expected as micellar ganglioside concentration increased if hSiglec-1 could bind the GM1a micelles (**Figure 3.7a**). No decrease in liposome binding was observed in the presence of GM1a micelles, suggesting that hSiglec-1 does not bind micellar GM1a (**Figure 3.7b**). It is noteworthy that liposome binding was not affected by the presence of micellar gangliosides even though the concentration was over 100-fold higher (ganglioside concentration in liposome 0.75 μ M, micellar concentration 100 μ M) binding demonstrating that liposomal presentation is important for ganglioside binding by hSiglec-1.

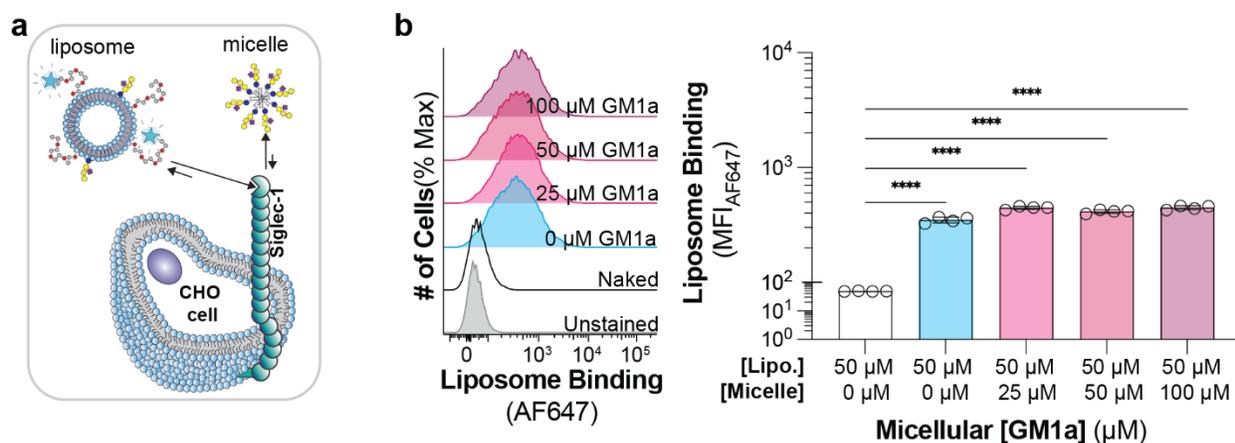


Figure 3.7: Binding of GM1a liposomes in the presence of micellular GM1a to hSiglec-1 expressing CHO cells. **a**, schematic describing competitive assay between GM1a micelles and GM1a liposomes. **b**, binding of GM1a liposomes in the presence of GM1a to hSiglec-1 expressing CHO cells. Black, naked liposome 0 μM micellular GM1a; blue, 3% GM1a liposome, 0 μM micellular GM1a; pink, 3% GM1a liposomes with increasing amounts of micellular GM1a. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median MFI from four technical replicates. A one-way ANOVA was used to compare the binding between GM1a liposomes in the presence of GM1a micelles to naked liposomes. **** = $P < 0.0001$.

The next parameter to optimize was the ganglioside content of the liposomes. The relative amount of ganglioside GM1a, GM2, GM3, and GD1a was titrated in the liposomes against hSiglec-1 expressing CHO cell (**Figure 3.8**). A significant decrease in ganglioside liposome binding to hSiglec-1 CHO cells was observed when the ganglioside content exceeded a threshold of 1–3 mol% for GM1a, GM2, and GD1a. In contrast, GM3 showed a different trend, with a progressive increase in ganglioside liposome binding up to 20 mol% GM3. To rule out the possibility of different incorporation efficiencies into the liposomes for each ganglioside, we quantified ganglioside incorporation into the liposomes using a mass spectrometry-based approach²⁷⁸ which uses deuterated standards and found similarly high rates of incorporation for GM1 (96%²⁷⁸), GM2 (98%), and GM3 (96%); suggesting that gangliosides are incorporated at similar levels.

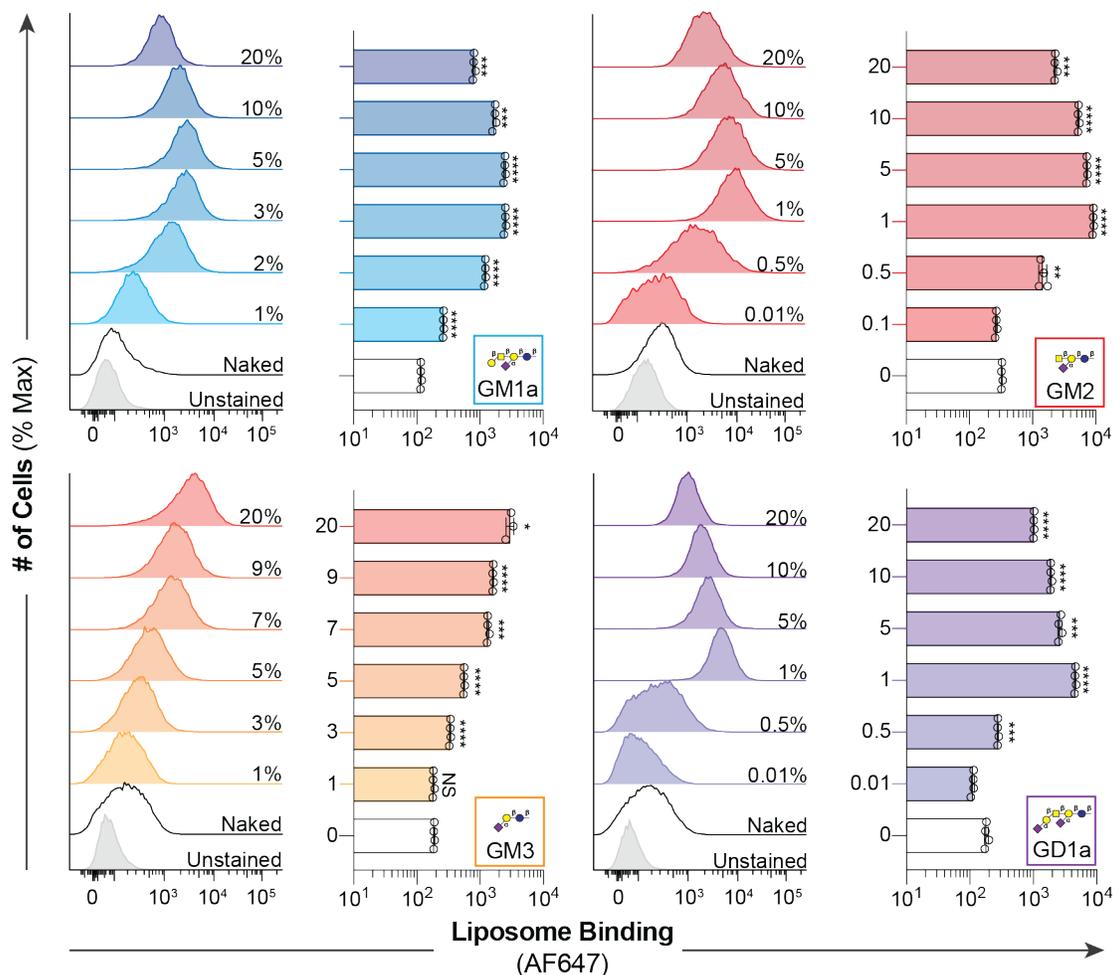


Figure 3.8: Binding of liposomes formulated with increasing ganglioside content to CHO cells expressing hSiglec-1. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from at least three technical replicates. A one-way ANOVA was used to compare if liposomes formulated with increasing amounts of ganglioside were significantly higher than a naked liposome. Not Significant (NS); $P > 0.05$; * = $0.05 > P \geq 0.01$; ** = $0.01 > P \geq 0.001$; *** = $0.001 > P \geq 0.0001$; **** = $P < 0.0001$.

Another possible explanation for the weaker binding of GM3 liposomes, compared to other ganglioside liposomes, is that the intrinsic affinity of hSiglec-1 for the GM3 trisaccharide is weaker than the oligosaccharide portion of the other gangliosides. To test this, we used a mass spectrometry-based Siglec binding assay³⁷ to determine the dissociation constant (K_d) of hSiglec-1 to the oligosaccharide moieties of GM1a, GM2, GM3, and GD1a (**Table 3.2**). In contrast with the cell assay, the GM3 trisaccharide exhibited the highest affinity towards hSiglec-1, with a K_d value of 0.5 ± 0.1 mM. Therefore, factors beyond the intrinsic affinity of hSiglec-1 to the

oligosaccharide moiety of the ganglioside influences avidity of hSiglec-1 for glycolipids presented in a lipid bilayer.

Table 3.2: Affinities of ganglioside oligosaccharides towards human Siglec-1.

Ganglioside Oligosaccharide	Dissociation Constant (K_d)
GM1a	1.2 ± 0.1 mM
GM2	0.9 ± 0.1 mM
GM3	0.5 ± 0.1 mM
GD1a	1.3 ± 0.2 mM

3.3.3: Ganglioside crowding impedes Siglec-ganglioside interactions

Another interesting observation in the ganglioside titration was that binding did not follow the expected sigmoidal curve where binding approaches saturation as the concentration of the ligand increases. Instead, a unimodal binding curve was observed with liposome binding peaking between 1-3 mol%. We hypothesized that the unimodal density-dependent binding of hSiglec-1 to GM1, GM2, and GD1a ganglioside liposomes is related to steric crowding²⁷⁹. To test this hypothesis, we used a liquid glycan array (LiGA)¹³⁵, wherein the oligosaccharide moieties of gangliosides were conjugated to a bacteriophage at different densities using a minimum of four independently barcoded preparations of the phage at each density, enabling binding to be read out by next generation sequencing of hSiglec-1 CHO cells incubated with LiGA (**Figure 3.9**). For GM2 and GD1a, binding was maximal at a density of approximately 26 nm²/ligand. This optimal density was in a similar range as the optimal density of 18 nm²/ligand on liposomes.

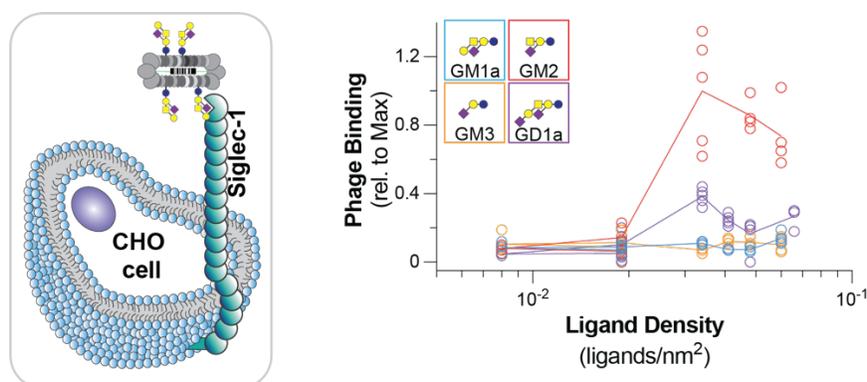


Figure 3.9: Liquid Glycan Array (LiGA) profile of glycan density on phage against hSiglec-1 CHO cells. Oligosaccharides of GM1am GM2, GM3 and GD1a were covalently linked to phages at different ligand densities and phage binding to hSiglec-1 expressing CHO cell was quantified by DNA sequencing.

To test this hypothesis further, we investigated molecular crowding using liposomes formulated with asialo GM1a (GA1). GA1 alone did not mediate binding to hSiglec-1, yet excess GA1 (20 mol%) with GM1 (3 mol%) in the ganglioside liposome formulation impaired binding to hSiglec-1 compared to GM1a alone (**Figure 3.10**). In fact, liposomes with excess GA1 showed equivalent binding to that of high density (20 mol%) GM1a ganglioside liposomes. As excess ganglioside densities are detrimental for hSiglec-1 binding, 3 mol% gangliosides were used in ganglioside liposome formulations moving forward. When comparing the results from the ganglioside titrations compared to the absolute affinities of the oligosaccharides, these results suggest that the binding of a Siglec and a ganglioside in solution is not representative of the same interaction in a bilayer.

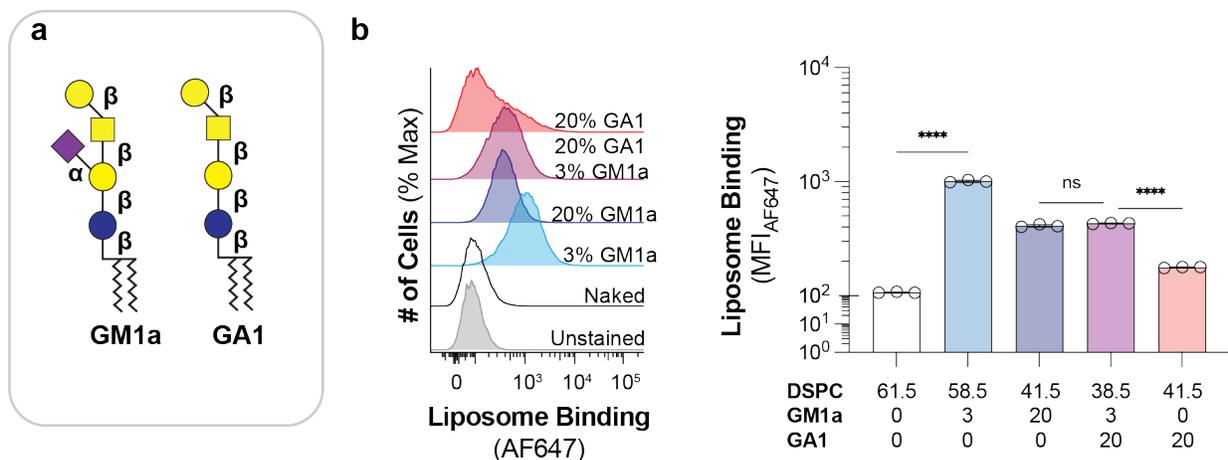


Figure 3.10: Binding of liposomes formulated with asialo-GM1 (GA1) and GM1a to CHO cells expressing hSiglec-1. **a**, Symbolic nomenclature for glycan structures of GM1a and GA1. **b**, Binding of GM1a liposomes formulated with and without GA1 to hSiglec-1 expressing CHO cells. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from three technical replicates. A one-way ANOVA was used for statistical analysis. Not Significant (NS); $P > 0.05$; **** = $P < 0.0001$.

As stated above, the acyl chains of gangliosides can vary between species and tissues²⁶⁰, so we assessed how the acyl chain structures affected the binding of GM1a ganglioside liposomes to hSiglec-1 CHO cells using GM1a from four sources: bovine, ovine, porcine, and synthetic. As expected, GM1a isolated from the different sources did have different acyl chains (**Figure 3.11a**). When ganglioside liposomes bearing GM1a isolated from each of the different

sources were tested for binding, no significant differences in binding were observed to hSiglec-1 expressing CHO cells (**Figure 3.11b**).

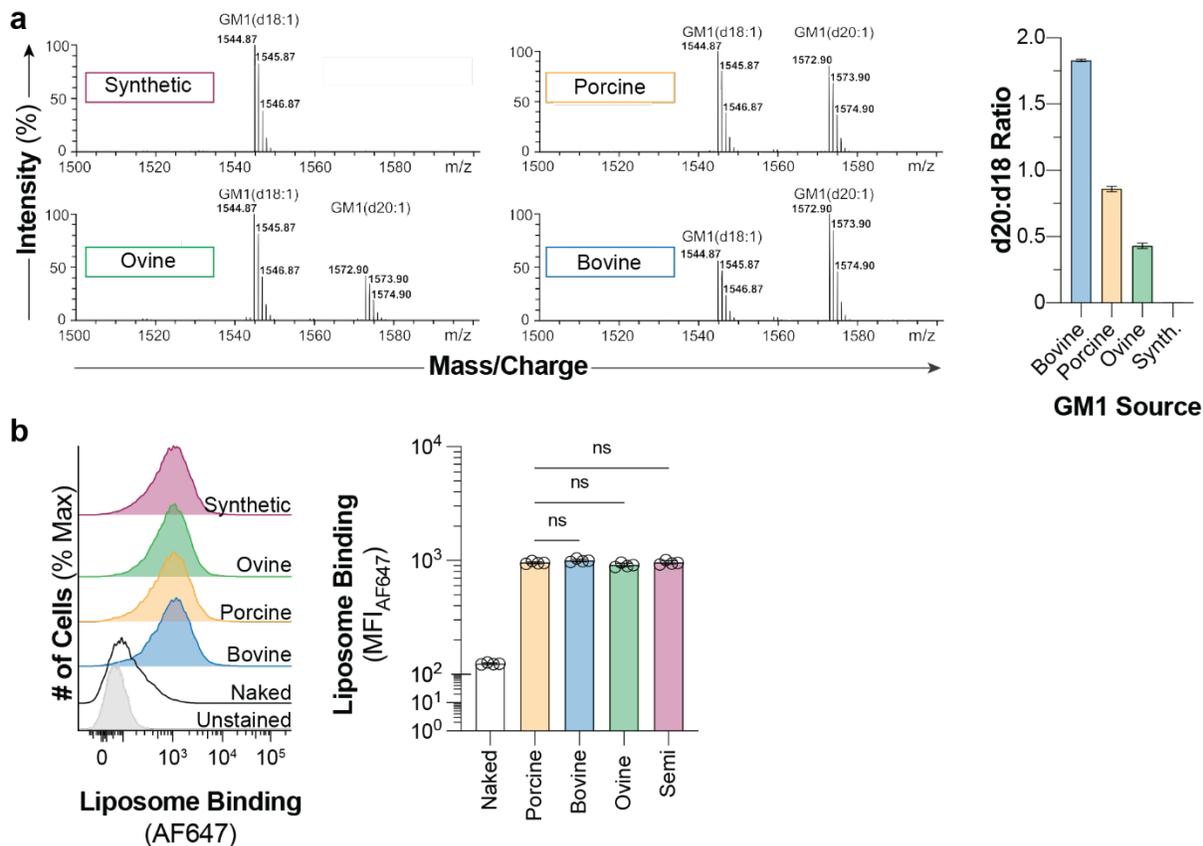


Figure 3.11: Effect of ganglioside acyl chains on ganglioside liposome binding to CHO cell expressing hSiglec-1. **a**, Mass spectra of ganglioside GM1a isolated from different sources. Results are quantified as the mean of six technical replicates \pm one standard deviation. **b**, Binding of liposomes formulated with GM1a isolated from different sources to CHO cells expressing hSiglec-1. Flow cytometry data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. For panel b, a one-way ANOVA was used for statistical analysis to compare liposomes formulated with GM1a from porcine brain (initial formulation parameter) to liposomes with GM1a from different sources. Not Significant (NS).

The last two parameters to be investigated were cholesterol content and acyl chain structure (**Figure 3.12a**). When the cholesterol content was titrated from 38-8 mol% no trend between cholesterol content and ganglioside liposome binding to hSiglec-1 CHO cells was observed (**Figure 3.12b**). As cholesterol content had limited to no effect on ganglioside liposome binding, 38 mol% was used moving forward as a cholesterol content of approximately 40 mol% is similar to what is found in mammalian cells and important for maintaining liposome integrity²⁸⁰. Conversely, the structure of the bulk lipid in the liposome formulation did impact ganglioside

liposome binding to hSiglec-1 CHO cells, with 1-palmitoyl-2-stearoyl-sn-glycero-3-phosphocholine (PSPC) showing the highest binding (**Figure 3.12c and d**). Accordingly, PSPC was used as the bulk lipid for the screening of the Siglec family.

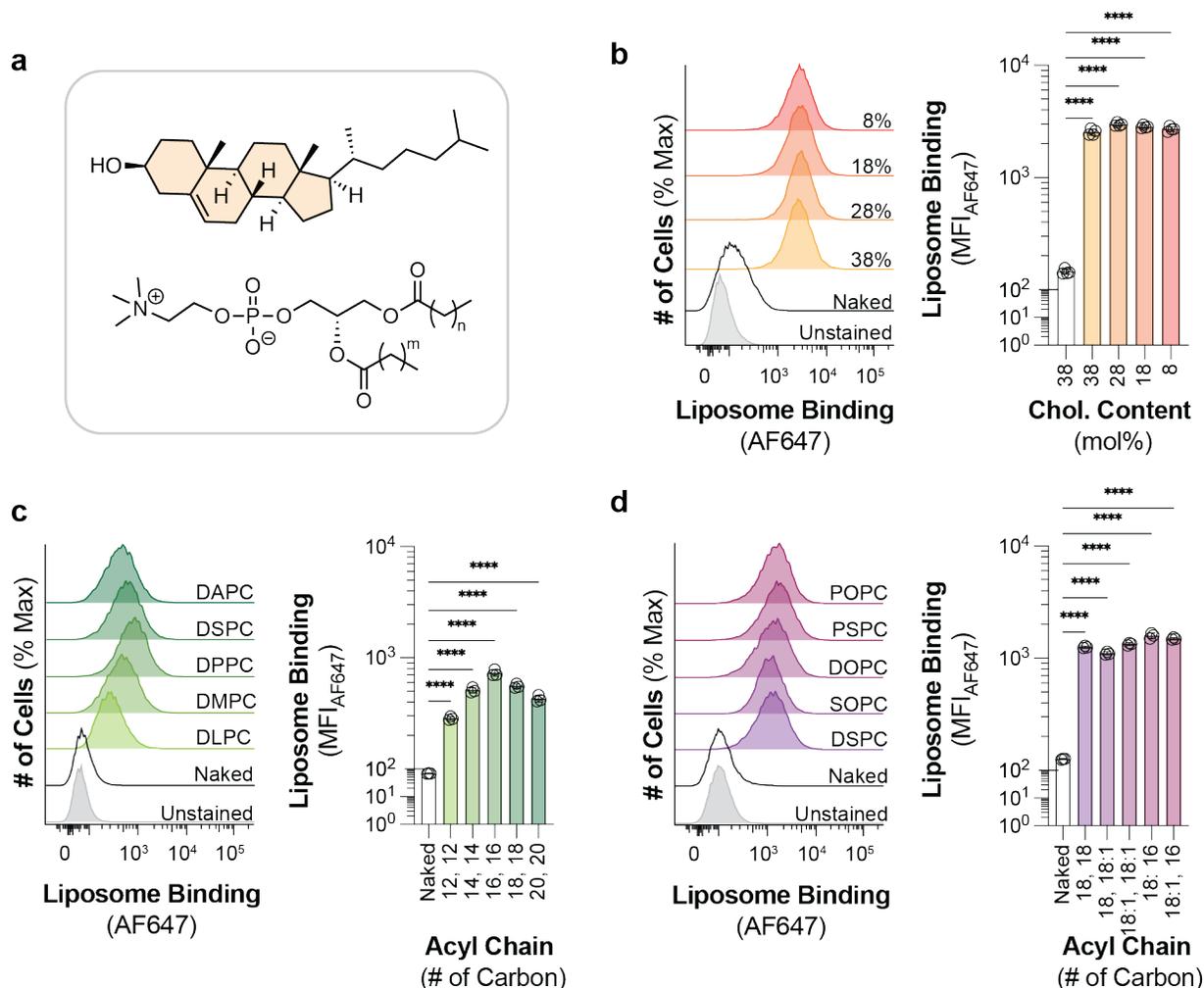


Figure 3.12: Binding of GM1a liposomes formulated with various amounts of cholesterol and lipids with different acyl chains to CHO cells expressing hSiglec-1. **a**, chemical structure of cholesterol and phosphatidyl choline lipids. **b**, Binding of liposome formulated with 3 mol% GM1a and varying amounts (38-8 mol%) of cholesterol to hSiglec-1 expressing CHO cells. **c**, Binding of 3 mol% GM1a liposomes formulated with bulk lipids with varying acyl chain lengths to hSiglec-1 expressing CHO cells. **d**, Binding of 3 mol% GM1a liposomes formulated with bulk lipids with asymmetric and unsaturated acyl chains to hSiglec-1 expressing CHO cells. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. A one-way ANOVA was used for statistical analysis. **** = $P < 0.0001$.

After completing the liposome formulation optimization, we compared the ability of our newly optimized liposome formulation and the starting *Doxil* liposome formulation displaying the high-affinity Siglec-1 ligand to bind to hSiglec-1 expressing CHO cells (**Figure 3.13**). Interestingly,

liposomes bearing GM2, bound to the hSiglec-1 expressing CHO cells with almost the same avidity as the high-affinity Siglec-1 bearing liposomes despite the affinity of Siglec-1 towards the oligosaccharide of GM2 being nearly 1000-fold lower than the synthetic ligand; further demonstrating that ganglioside presentation from a liposome can enhance avidity.

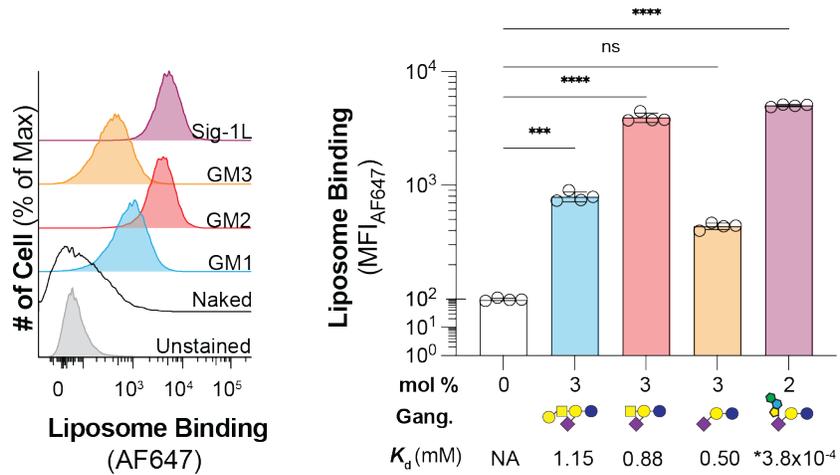


Figure 3.13: Binding of liposomes formulated with gangliosides presented from an optimized liposome formulation compared to high affinity Siglec-1-ligand presented from *Doxil* liposome formulation to hSiglec-1 expressing CHO cells. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. **** = $P < 0.0001$.

3.3.4 Human Siglec-ganglioside interactions

After completing the liposomal formulation optimizations, we were ready to move on to interrogating the rest of the members of the human Siglec family. As synthesizing genuine gangliosides is very challenging, we were limited to a panel of nine commercially available gangliosides which were isolated from animal sources.

Table 3.3: Sources of gangliosides used to interrogate the human Siglec family

Ganglioside	Source	Supplier
GM1	Porcine Brain	TRB Chemedica Inc.
GM2	Bovine, Semi-synthetic	Matreya
GM3	Bovine Milk	Sigma
GM4	Chicken Egg	Matreya
GD1a	Bovine, Natural	Matreya
GD1b	Bovine, Natural	Matreya
GD3	Bovine Buttermilk	Matreya
GT1b	Bovine, Natural	Matreya
GQ1b	Porcine, Natural	Matreya

3.3.5 Development and validation of Siglec expressing CHO cells

As stated above, CHO cells were chosen as a cell line because they lack Siglecs and are easy to stably transfect²⁸¹. A novel panel of 24 CHO cells were developed where each cell line expresses a full-length, membrane-bound wildtype human Siglec or their corresponding arginine mutant, wherein the canonical arginine residue critical for sialic acid recognition is mutated³⁷ (**Figure 3.14**). The extracellular domains of Siglec-14 and -16 are nearly identical to Siglec-5 and -11, respectively; therefore, Siglec-14 and -16 were not included in our panel. Siglec-15 requires the adapter protein DAP12 for cell surface expression due to a lysine in its transmembrane segment⁹⁹, but this DAP12-dependency was eliminated through a K274L mutation.

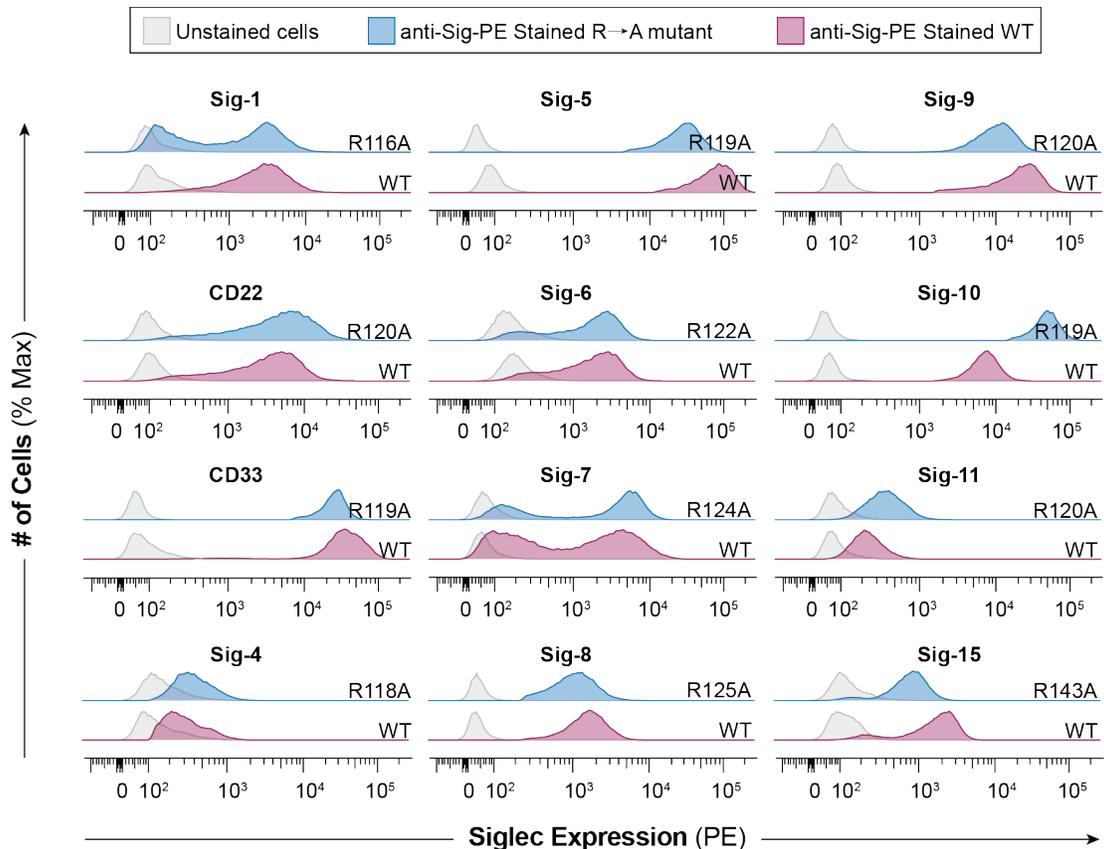


Figure 3.14: Siglec expression validation, of CHO cell lines engineered to express each human Siglec and each corresponding arginine mutant.

3.3.6: Interrogating human Siglecs with ganglioside liposomes

Using the optimized liposomal formulation and our newly developed panel of CHO cells, the human Siglec family was tested against nine gangliosides in ganglioside liposomes using untransfected (UT) CHO cells as a background. Binding of Siglec-1, -5, -6, -9, and -10 to multiple gangliosides was observed (**Figure A3.1**). Many novel interactions were observed, including: Siglec-5 with GM1 and GD1b, Siglec-6 with numerous gangliosides, Siglec-9 with GM2, and Siglec-10 with GM1 and GD3. Binding of ganglioside liposomes to Siglec-1, -5, -9, and -10 was abrogated when their canonical essential arginine was mutated. Unexpectedly, mutation of the canonical essential arginine in Siglec-6 (Arg122) did not abrogate ganglioside liposome binding. Although Siglec-6 has not been reported to bind gangliosides, Arg122 was reported as being critical for recognition of sialylated ligands on cells²⁰⁵. These findings suggest that Siglec-6 binding to gangliosides is not dependant on the conserved arginine residue. It was surprising that initially,

Siglec-4 and -7 showed minimal ganglioside liposome binding, given previous reports showing that GD1a and GT1b were ligands for Siglec-4, while GQ1b and GD3 were ligands for Siglec-7^{143, 282}. We considered that these Siglecs may be masked by *cis* ligands on the CHO cells, preventing interactions with ganglioside bearing liposomes. After treatment of CHO cells expressing human Siglec-4 with neuraminidase S (Neu S), a neuraminidase which specifically cleaves $\alpha(2\rightarrow3)$ linked sialosides between the sialic acid and the galactose, binding of GD1a liposome to cells was observed; demonstrating that *cis* ligands can prevent Siglec-liposome binding (**Figure 3.15**).

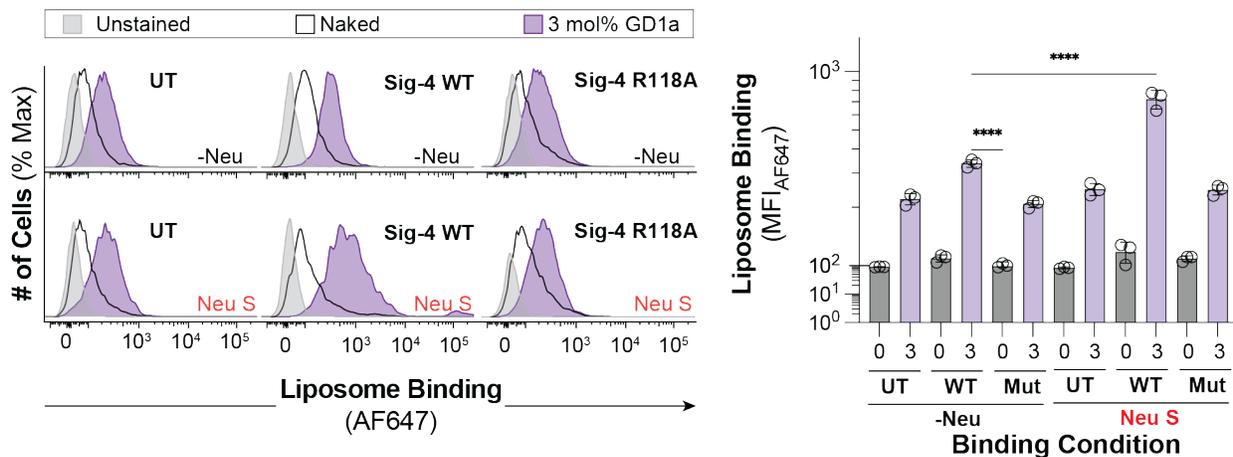


Figure 3.15: Ganglioside liposome binding to CHO cells expressing Siglec-4 in the absence of serum and after treatment with neuraminidase. 3 mol% GD1a ganglioside liposomes binding to hSiglec-4 expressing CHO cells when measured after treatment with neuraminidase S. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from three technical replicates. A one-way ANOVA was used for statistical analysis. ****= $P < 0.0001$.

Therefore, several of the binding experiments were repeated after treatment with Neu A (neuraminidase that cleaves all sialic acid linkages) and Neu S (neuraminidase that cleaves only $\alpha(2\rightarrow3)$ sialosides). In addition to improving ganglioside binding to Siglec-expressing cells, unmasking conditions also modestly enhanced glycolipid liposome binding to Siglec-5 and -9 (**Figure A3.2**). The binding between ganglioside bearing liposomes and the human Siglec family are summarized in **Figure 3.16**.

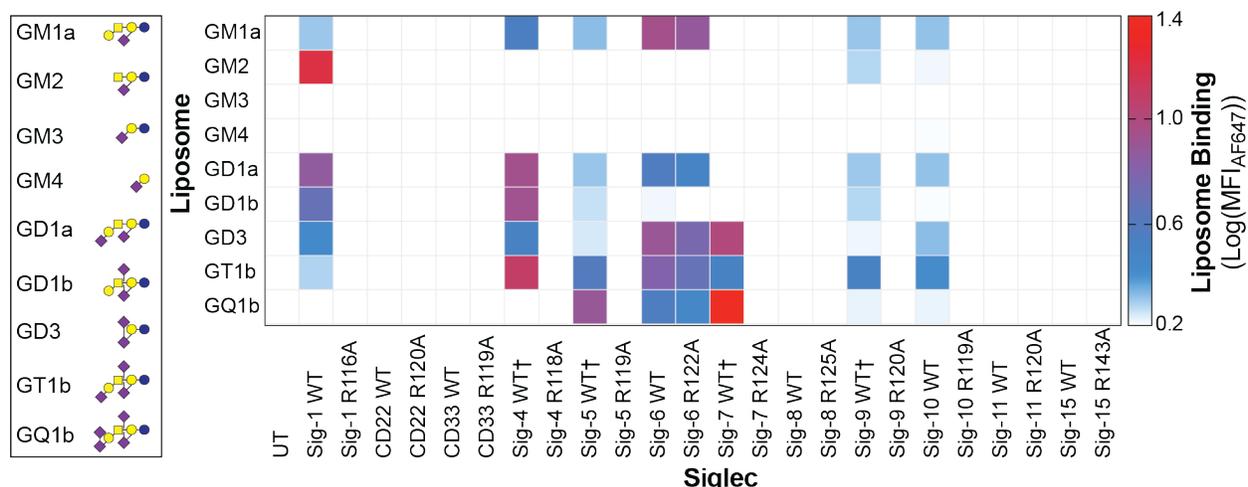


Figure 3.16: Heatmap summarizing the binding interactions of each human Siglec. Color is representative of the $\log_{10}(\text{MFI}_{\text{AF647}})$ of each ganglioside liposome subtracted from the $\log_{10}(\text{MFI}_{\text{AF647}})$ of the same ganglioside against untransfected (UT) CHO cells. † Denotes binding was measured after treatment of cells with neuraminidase.

3.3.7: Unimodal binding profiles may be a general Siglec phenomenon

As ligand content was important for ganglioside binding to hSiglec-1 expressing cells, we also titrated the amount of GM1a and GD3 in ganglioside liposomes and tested binding against Siglec-6 and -7, respectively (**Figure 3.17**). Similar to the binding of Siglec-1 to GM1a, GM2 and GD1a, binding of ganglioside liposomes to Siglec-6 and -7 decreased above an optimal ganglioside content.

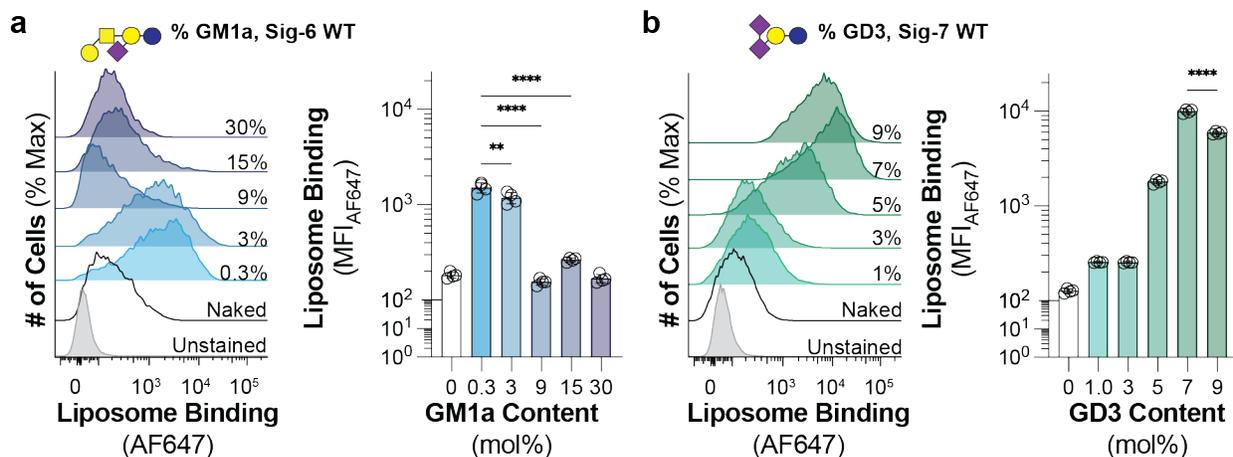


Figure 3.17: Ganglioside content titration to Siglec-6 and Siglec-7 expressing CHO cells. a, b, binding of GM1a and GD3 liposomes formulated with varying ganglioside content to Siglec-6 and Siglec-7 expressing CHO cells respectively. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from at least three technical replicates. A one-way ANOVA was for statistical analysis. Not Significant (NS); $P > 0.05$; **** = $P < 0.0001$.

3.3.8 Mouse Siglec-ganglioside interactions

After optimizing a liposome formulation for Siglec engagement and screening the entire human Siglec family against a panel of gangliosides, we were able to move on to screening the murine family against the same panel of gangliosides. This would allow for the comparison of the ganglioside binding profiles between murine and human Siglecs; a key step in assessing the appropriateness of using mouse models to study human Siglec-ganglioside physiology. To avoid the problem of ligands being missed due to masking caused by *cis* ligands, we decided to use the bead assay to measure the murine Siglec-ganglioside interactions instead of the cell assay.

3.3.8.1 Profiling mouse Siglec-ganglioside interactions within the context of a bilayer

To study the ability of mSiglecs to recognize glycolipids, we profiled the binding of Siglecs to a panel of nine gangliosides using our previously optimized liposome formulation, which consists of 3 mol% ganglioside, 0.5 mol% PEG₄₅-DSPE, 58.5 mol% PSPC, and 38 mol% cholesterol¹²⁸. The binding between the mSiglecs and gangliosides was studied using a bead assay wherein a Siglec-Fc bearing the Strep-tag II was immobilized on streptavidin microbeads, incubated with fluorescent liposomes containing gangliosides, and quantified by flow cytometry. Using this approach, it was found that mSiglec-1, Siglec-E, Siglec-F, and mSiglec-15 all bound multiple gangliosides, whereas no binding was observed between the other Siglecs and the ganglioside liposomes (**Figure 3.18**). mSiglec-1, Siglec-E, and Siglec-F bound to nearly all ganglioside liposomes, albeit with different preferences. Surprisingly, none of these Siglecs engaged liposomes bearing GM1a. For mSiglec-1, GD1a bearing liposomes had the strongest binding, whereas liposomes formulated with GT1b bound the strongest to Siglec-E, and Siglec-F had the strongest binding to liposomes formulated with GM2. mSiglec-15 interacted preferentially with liposomes formulated with gangliosides, which possessed an $\alpha(2\rightarrow3)$, Sia on the terminal Gal, such as GM4, GT1b, and GD1a while also binding GQ1b.

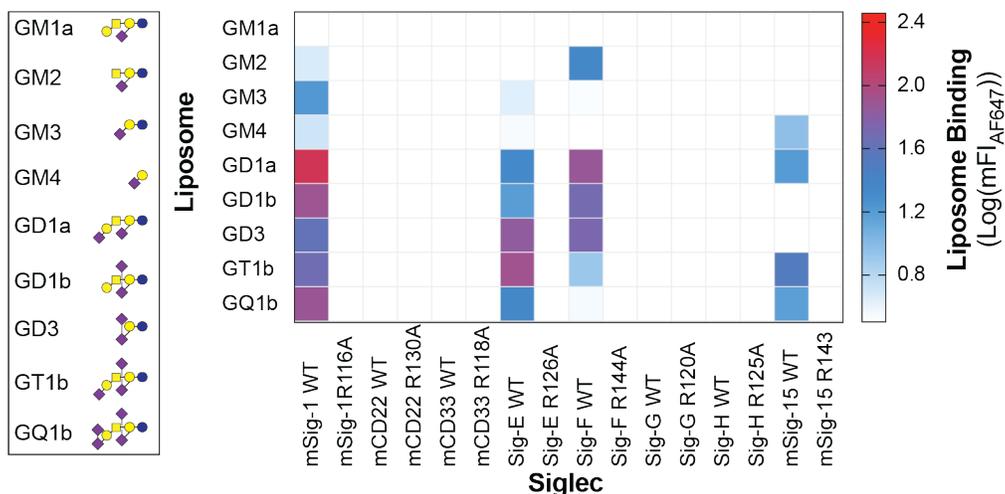


Figure 3.18: Heatmap summarizing the binding interactions of each murine Siglec and ganglioside bearing liposomes. Color is representative of the $\log_{10}(\text{mFI}_{\text{AF647}})$ of each ganglioside liposome subtracted from the $\log_{10}(\text{mFI}_{\text{AF647}})$ of the same ganglioside liposome against the corresponding arginine mutant.

3.3.8.2 Profiling mouse Siglec-ganglioside interactions outside the context of a bilayer

It was unexpected that no interactions between any mSiglec and GM1a was observed. As we learned that oligosaccharide presentation is important for ganglioside recognition by a Siglec, we posited that a different oligosaccharide presentation may be able to rescue binding between a murine Siglec and GM1a. To this, we assessed mSiglec-ganglioside binding through a traditional plate-based ELISA wherein the gangliosides are adsorbed to a microplate. The Siglec-Fc was precomplexed with Strep-Tactin Horseradish peroxidase, incubated with the immobilized gangliosides in a well, washed, and colorimetrically developed. Similar to the bead assay, mSiglec-1, Siglec-E, Siglec-F, and mSiglec-15 bound multiple gangliosides. Nevertheless, not every interaction from the bead assay was observed in the ELISA. For example, the mSiglec-1-complex engaged with GM3, GD1a, GT1b, and GQ1b in the ELISA whereas mSiglec-1-beads interacted with all gangliosides other than GM1a to some extent in the bead assay. mSiglec-15-complex did not bind GQ1b in the ELISA, but mSiglec-15-beads did bind GQ1b in the bead assay. Conversely, some Siglec-ganglioside interactions were observed in the ELISA that were not observed in the bead assay. For instance, the Siglec-E-complex engaged GM1a in the ELISA but not the bead assay. Overall, the results from the bead assay and the ELISA largely agree, demonstrating that mSiglec-1, Siglec-E, Siglec-F, and mSiglec-15 are proficient at interacting with

gangliosides, but there are subtle differences in the ganglioside binding profile between approaches and not all interactions are observed when the glycolipid is presented from a lipid bilayer.

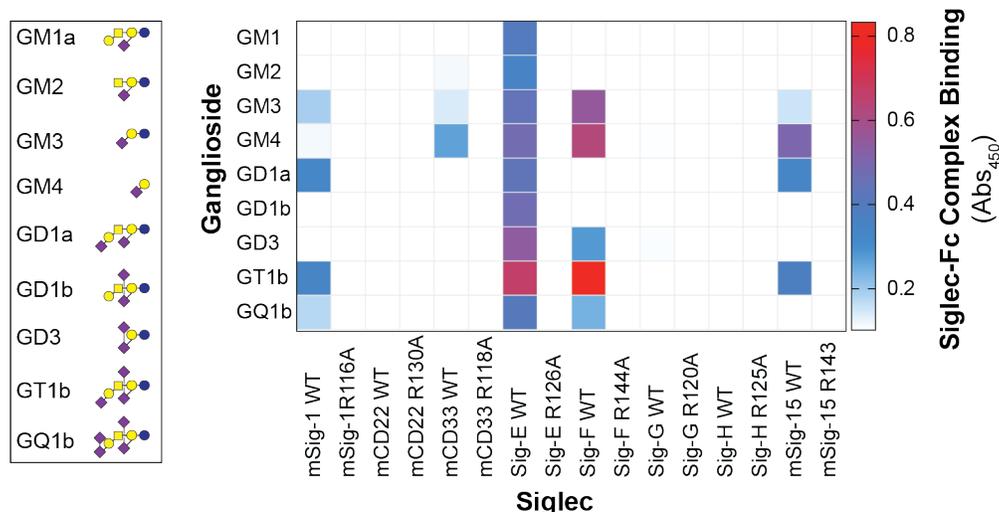


Figure 3.19: Mouse Siglec-ganglioside binding outside a lipid bilayer using an ELISA. Heatmap summarizing mSiglec-Fc binding to adsorbed gangliosides. Color is representative of the mean binding of the WT mSiglec-Fc complex to the adsorbed gangliosides liposome subtracted from the corresponding mutant mSiglec-Fc complex binding to the same ganglioside at least four technical replicates.

3.3.8.3: Re-optimization of liposome formulation does not reveal GM1a-mSiglec-1 interaction

With the exception of Siglec-E in the ELISA, it was unexpected that no interactions between any mSiglec and GM1a were observed. The structural similarity between the oligosaccharide of GM1a and other gangliosides that did bind murine Siglecs motivated an investigation into why no murine Siglec interacted with GM1a bearing liposomes. We posited that since our previous formulation was optimized against hSiglec-1¹²⁸, it may not be optimal for murine Siglecs. Accordingly, we titrated the cholesterol content (38, 20, and 10 mol%) and the length of the acyl chain used for the bulk lipid in the liposome (12 carbon-DLPC, 16 carbon-DPPC, 20 carbon DAPC) in 3 mol% GM1a and GD1a liposomes against mSiglec-1 (**Figure 3.20**).

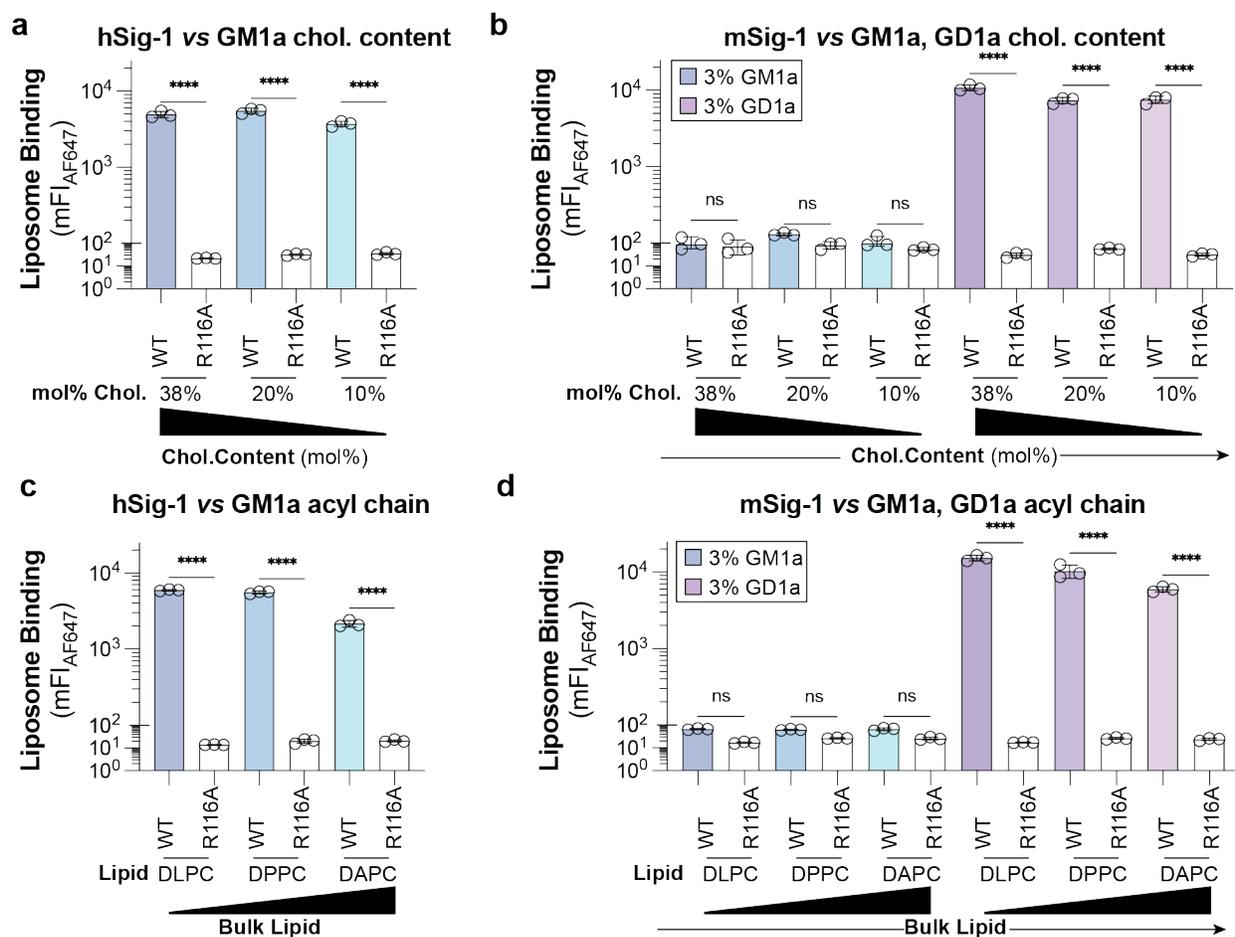


Figure 3.20: Optimization of liposome cholesterol content and bulk lipid structure for screening murine Siglec-1-ganglioside interactions. a, b, Cholesterol content titration of GM1a and GD1a liposomes against h/mSiglec-1 respectively. c, d, acyl chain length titration of GM1a and GD1a liposomes against h/mSiglec-1 respectively. Data is represented as the mean of at least three technical replicates and error bars are representative of one standard deviation from the mean. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; **** $P < 0.0001$.

While both cholesterol content and acyl chain length had modest effects on Siglec engagement of ganglioside liposomes, no significant interaction between mSiglec-1 and GM1a bearing liposomes was observed against any of the formulations tested. During the optimization of the liposome formulation with the human Siglecs, ganglioside content (ligand density) was found to strongly influence Siglec-liposome interactions. To assess if the ganglioside content was responsible for the lack of GM1a binding to mSiglecs, we titrated the amount of three gangliosides (GM1a, GM2, and GD1a) from 1-10 mol% against the four mSiglecs that were found to engage with gangliosides using hSig-1 as a reference (**Figure 3.21**).

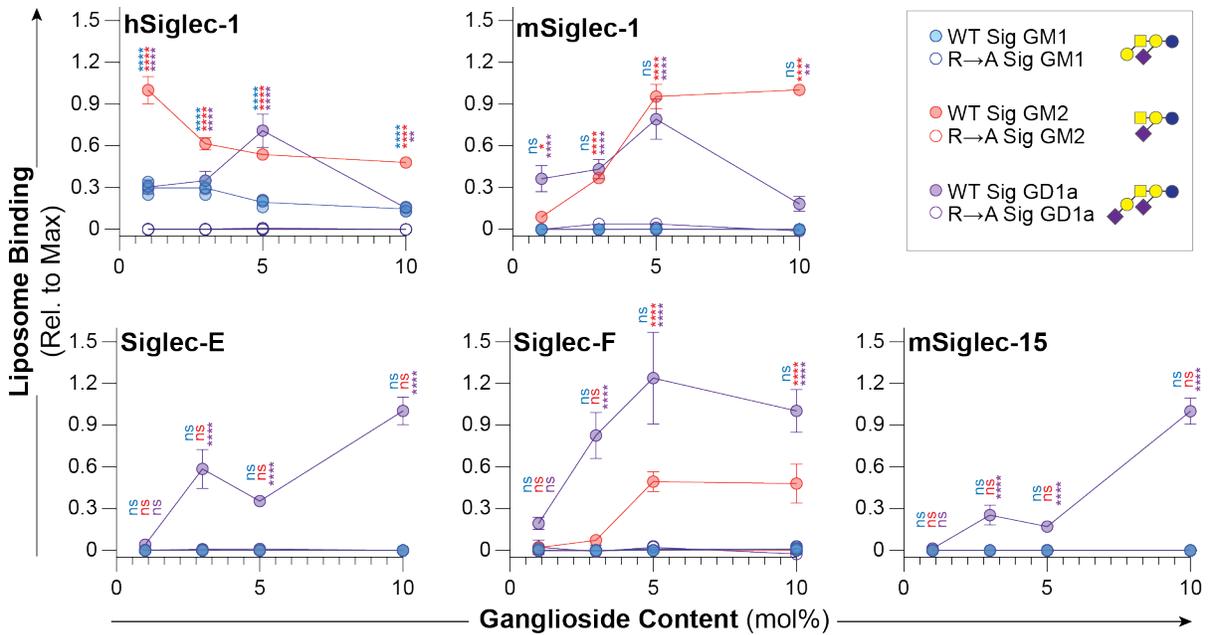


Figure 3.21: Ganglioside content titration (1-10 mol%) against murine Siglecs that were found to bind gangliosides. Data is represented as the mean of at least three technical replicates and error bars are representative of one standard deviation from the mean. a one-way ANOVA was used to compare the binding between the wildtype Siglec and the corresponding arginine mutant at each mol% of ganglioside. Not Significant (NS), $P > 0.05$; $*0.05 > P \geq 0.01$; $** 0.01 > P \geq 0.001$; $****P < 0.0001$.

In line with the initial results in the cell assay, hSiglec-1 bound GM1a, GM2, and GD1a, with GM2 being the best ligand¹²⁸. For the mSiglecs, GM1a did not bind to mSiglec-1, Siglec-E, Siglec-F, or mSiglec-15 at any mol% tested, whereas GD1a was found to be a ligand for all the Siglecs tested. Like our initial screen of the mSiglecs, GM2 was found to be a ligand for mSiglec-1 and Siglec-F and not for Siglec-E or mSiglec-15. However, there was a difference in the ligand density effects between mSiglec-1 and hSiglec-1 with GM2. Interestingly, the binding profile between Siglec-E and Siglec-F with liposomes bearing increasing amounts of GD1a continue up to 10 mol% rather than decreasing like other Siglecs such as hSiglec-1, mSiglec-1 and Siglec-F. Overall, these results suggest that the lack of engagement of GM1a by murine Siglecs is not due to an unoptimized liposomal formulation.

3.3.8.4: Murine Siglec-1 prefers terminally linked $\alpha(2 \rightarrow 3)$ linked gangliosides

After ruling out that the liposome formulation was the reason for the lack of observed binding between mSiglec-1 and GM1a, we consider that it may be the structure of GM1a that mSiglec-1 has difficulty engaging with. As mSiglec-1 engages liposomes bearing GD1a and GM2 but not

GM1a and considering where the sialic acid(s) are presented from the oligosaccharide structures of GM1a, GM2, and GD1a, we hypothesized that the sialic acid presented from the internal Gal residue is difficult for mSiglec-1 to access. It follows that because GD1a contains Sia at both the internal and external Gal residues, and was recognized by mSiglec-1, it suggests that recognition of GD1a by mSiglec-1 is through the terminal Gal residue. To understand the contribution to mSiglec-1 binding for each sialic acid residue on GD1a, we tested the ability of mSiglec-1 to bind GM1b, a linear structural isomer of GM1a where the sialic acid is linked to the terminal Gal instead of the internal Gal. The ability of h/mSiglec-1 to bind the oligosaccharides of GM1a and GM1b was tested using a quantitative native mass spectrometry-based assay developed previously to study Siglec-ligand interactions^{37, 79, 128}. The concentration of ganglioside oligosaccharide was then increased (100-500 μ M for GM1a, 20-100 μ M GM1b) and the change in the amount of Siglec-oligosaccharide complex was determined (**Figure 3.22a and b**). The interaction between hSiglec-1 and GM1b was found to be the strongest ($K_d = 0.89$ mM), followed by mSiglec-1 with GM1b ($K_d = 1.2$ mM), hSiglec-1 with GM1a ($K_d = 1.5$ mM), and the weakest interaction was between mSiglec-1 and GM1a ($K_d = 2.0$ mM). These results suggest that both mSiglec-1 and hSiglec-1 prefer terminal $\alpha(2\rightarrow3)$ sialic acids, but that hSiglec-1 is better able to accommodate the internal $\alpha(2\rightarrow3)$ linked sialic acid in GM1a compared to mSiglec-1.

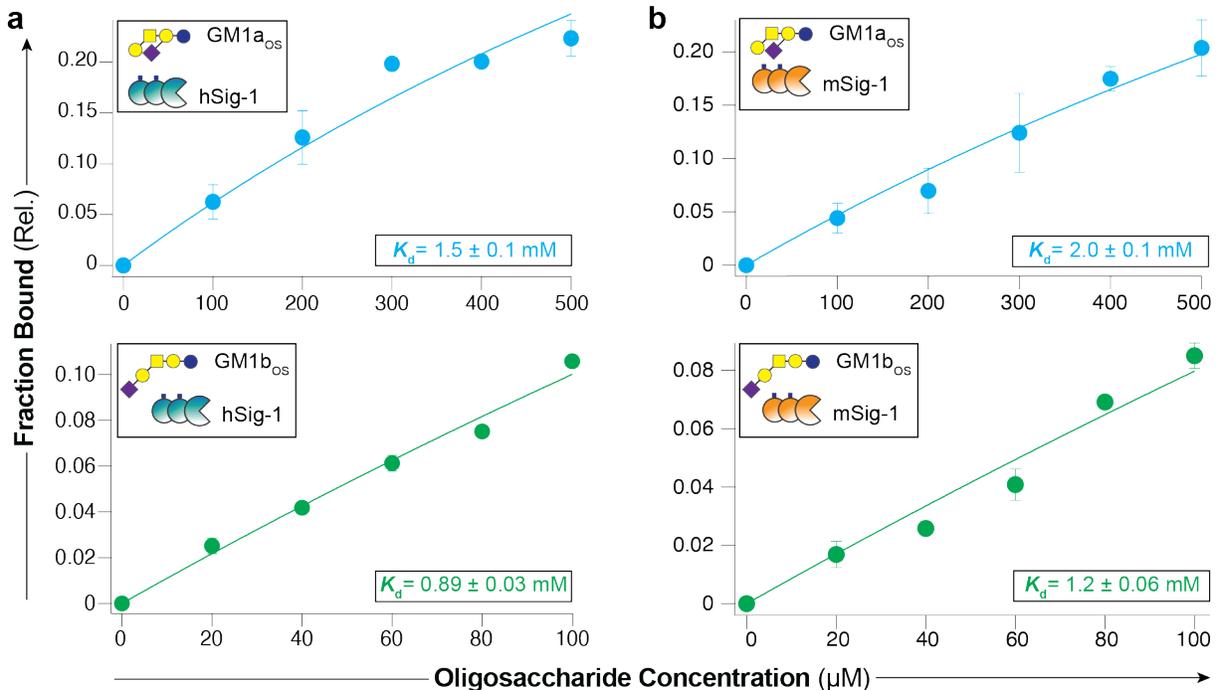


Figure 3.22: Quantifying the interactions between m/hSiglec-1 and the oligosaccharide of GM1a and GM1b by a native mass spectrometry binding assay. a, b, Summary of ganglioside oligosaccharides titrations binding between m/hSiglec-1 fragment to the oligosaccharide of GM1a (100-500 μM) and GM1b (20-100 μM)

The results from the mass spectrometry assay, above, were followed up using the ELISA and bead assay. In the ELISA, hSiglec-1 showed the best binding to GD1a, followed by GM1b, and then GM1a. Recognition of GD1a and GM1b by mSiglec-1 was similar, but no significant binding to GM1a was observed (**Figure 3.23a**). In the bead assay, GD1a was found to be a superior ligand compared to GM1b for both murine and human Siglec-1 and GM1a was found to be a ligand for hSiglec-1 but not mSiglec-1 (**Figure 3.23b**). These results further support that mSiglec-1 poorly recognizes GM1a, particularly in the context of a lipid bilayer due to the sialic acid residue being linked to the internal Gal residue (**Figure 3.23c**).

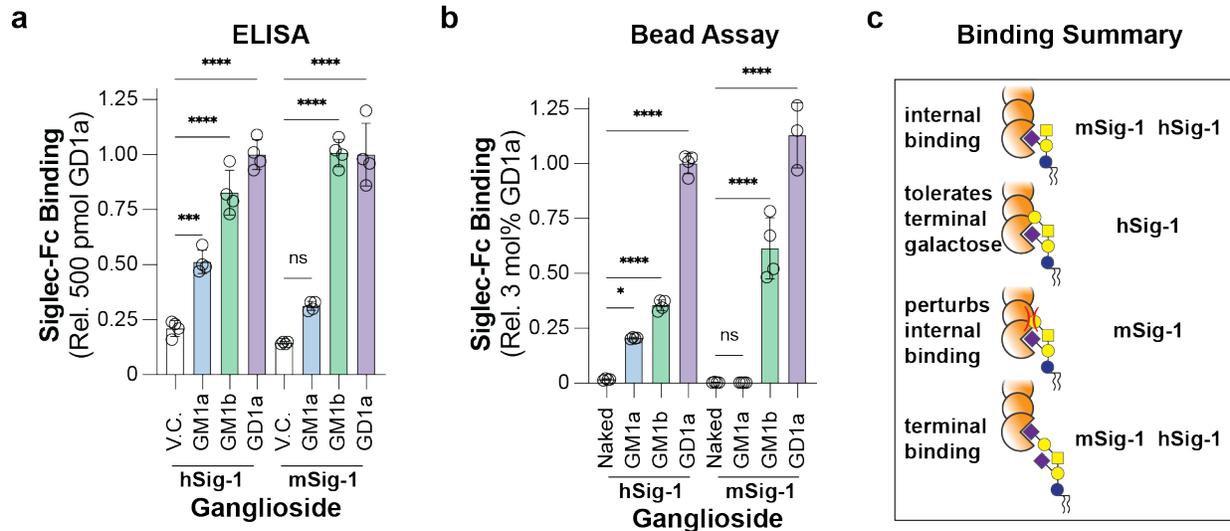


Figure 3.23: Comparative binding of internal and externally linked sialic acids to m/hSig-1 in the bead assay and ELISA. a, b, hSig-1, and mSig-1 binding to GM1a, GM1b and GD1a in the ELISA and bead assay respectively. c, proposed model for human mSig-1 binding of gangliosides. Data is represented as the mean of at least three technical replicates and error bars are representative of one standard deviation from the mean. For a and b, a one-way ANOVA was used to compare the binding of each ganglioside to either a vehicle control (V.C.) or a naked liposome respectfully. Not Significant (NS); $P > 0.05$; *** = $0.001 > P \geq 0.0001$; **** = $P < 0.0001$.

3.4: Discussion

3.4.1: Siglec-ganglioside interactions

Sialic acid-containing glycoproteins and glycolipids play many biological roles through serving as ligands for sialic acid-dependent lectins, such as the Selectins and Siglecs. Gangliosides are highly abundant on mammalian cell membranes but studying the interactions between proteins and glycolipids in a biological membrane has many challenges. In addition to the complex mixture of glycolipids in a cell membrane, membrane dynamics and composition can influence the conformation of the carbohydrate headgroup of gangliosides^{62, 278, 283}. Here, we reduced this complexity by using liposomes displaying glycolipids to study their ability to engage immunomodulatory Siglecs. By developing an understanding of how the composition of the lipid bilayer influences the ability of Siglecs to engage gangliosides, we developed an optimal liposome formulation leading to the discovery of new Siglec–ganglioside interactions.

3.4.2: Liposome formulation effects Siglec-ganglioside binding

In the process of optimizing the liposomal formulation, a striking observation was the unimodal nature of the binding between hSiglec-1 and ganglioside liposomes formulated with GM1a, GM2, and GD1a, which was not observed for GM3. Previous work examining GM3 content in liposomes and its impact on recognition by hSiglec-1 found that binding increased up to 5 mol%, but higher ganglioside content was not reported¹⁵⁰. This phenomenon was not unique to hSiglec-1, as Siglec-6, -7, and mSiglec-1 also showed similar binding patterns with GM1a, GD3, and GD1a respectively. Interestingly, this trend was not observed with Siglec-E mSiglec-15 with GD1a. Perhaps difference in the effect of ganglioside content on the ability of the Siglec to bind to the ganglioside bearing liposomes sheds light on its biological role. For example, the natural ligands for mSiglec-15 may be higher density compared to Siglec-1. However, this is purely speculation and if ganglioside content can reach these levels in cells or EVs has not been reported.

With respect to the Siglecs that were less efficient at binding gangliosides above an optimal amount, several lines of evidence in our studies, suggests that this behavior is due to steric crowding at high ganglioside content that negatively impacts binding. Steric crowding may arise from ganglioside–ganglioside interactions, which are dependent on the structure of the oligosaccharide²⁸⁴, as a form of phase separation²⁸⁵. The divergent behavior of GM3 may be related to observations made in molecular modeling studies, which showed that the relatively small, linear GM3 trisaccharide is buried in the bilayer, making it difficult for hSiglec-1 to access²⁸⁶. Except for GM3 and GM4, gangliosides presented from a lipid bilayer demonstrated remarkable avidity for Siglecs. For example, the soluble equivalent of the high affinity Siglec-1 ligand was previously found to have an IC_{50} value of 0.38 μ M with Siglec-1¹⁶⁰, which is over 1000-fold stronger than the affinity we report (0.9 ± 0.1 mM) for GM2 to Siglec-1. Yet, when presented in a liposome at a similar ligand content, both the high affinity Siglec-1 ligand and GM2 have a similar ability to engage hSiglec-1. This discrepancy may be related to the high affinity Siglec-1 ligand being presented at the end of a long PEG linker, leading to entropic costs in organizing multiple copies of the ligand for multivalent engagement of hSiglec-1²⁸⁷.

Our observation that ganglioside recognition on liposomes is shielded by PEG could be analogous to the glycocalyx of cells. Indeed, there are few examples of *trans* interactions between Siglecs on one cell and gangliosides on another²⁸⁸. These results point to other relevant biological locations where gangliosides may serve as Siglec ligands, such as on EVs and viruses. Gangliosides are found ubiquitously across all cell types and tissues and contribute to the composition of EVs²⁸⁹. hSiglec-1²⁵⁶ and -7^{259, 262} have both been shown to engage EVs in a manner that is predicted to be through glycolipids.

3.4.3: Siglec-ganglioside binding partners

Using the cell-based assay, many of the established human Siglec-ganglioside interactions were reproduced, specifically with hSiglec-1, -4, -5, -7, -9, and -10^{242, 249, 254}. In addition, we found novel interactions including: hSiglec-4 with GM1a, GD1b, and GD3; Siglec-5

with GM1a; Siglec-9 recognizes GM1a; and Siglec-10 recognizes GM1a, GM2, and GD3. However, it is important to consider how membrane dynamics influence avidity. There are discrepancies in the literature and in this study with respect to Siglec–ganglioside interactions that is likely due to the experimental format used. For example, Siglec-6 was found to bind many gangliosides in the cell assay but not as many in the ELISA. Moreover, not every interaction was observed when in both the bead assay and the ELISA with the murine Siglecs. Moreover, binding of Siglec-6 to glycolipids had never been observed, although Siglec-6-ganglioside interactions have only been investigated to a limited degree with the only report being the use of an ELISA to measure the binding of Siglec-6 towards GD3²⁰⁵. In summary, interpretation of these results with respect to biological significance requires careful consideration of the experimental platform.

Comparing the results of the interrogation of the human and murine Siglec, some of the mSiglecs bound more similarly to the analogous hSiglec than others when gangliosides are presented from a liposome (**Figure 3.24**). For the conserved Siglecs, hSiglec-1 and mSiglec-1 bound all the same gangliosides with the exception of mSiglec-1 being unable to bind GM1a. The inability of mSiglec-1 to bind GM1a was previously reported using a cell adhesion assay²⁴⁹. In this assay, which, is similar to an ELISA, gangliosides were adsorbed to a microplate and then COS cells expressing full-length membrane bound Siglecs were added to the plate. Unbound cells would then be washed away and the level of binding between the Siglec and the ganglioside was quantified by the activity of lactate dehydrogenase post cell lysis which, was proportional to the strength of the interaction between the Siglec and the ganglioside. Both m/hSiglec-1 are generally considered to recognize $\alpha(2\rightarrow3)$ sialic acid⁶; however, the ability of both m/hSiglec-1 to bind $\alpha(2\rightarrow8)$ linked sialosides may be unappreciated, as both m/hSiglec-1 were found to bind $\alpha(2\rightarrow8)$ Sia gangliosides such as GD3, GD1b, and GQ1b¹²⁸. However, it is difficult to fully rule out that a portion of the $\alpha(2\rightarrow8)$ linked sialoside was hydrolyzed during the course of liposome preparation and led to these results. Consulting crystal structures of mSiglec-1, PDB 1OD7²⁹⁰ and 1QFO²⁹¹, the latter of which is co-crystalized with sialyllactose (GM3 oligosaccharide), a co-

crystal structure of m/hSig-1 with a $\alpha(2\rightarrow8)$ linked sialoside would be very insightful with respect to how Siglec-1 engages/accommodates the $\alpha(2\rightarrow8)$ sialic acid. In the interrogation of the human Siglec family study, no binding between hSiglec-15 and any of the gangliosides was observed, yet mSiglec-15 did show ganglioside binding.

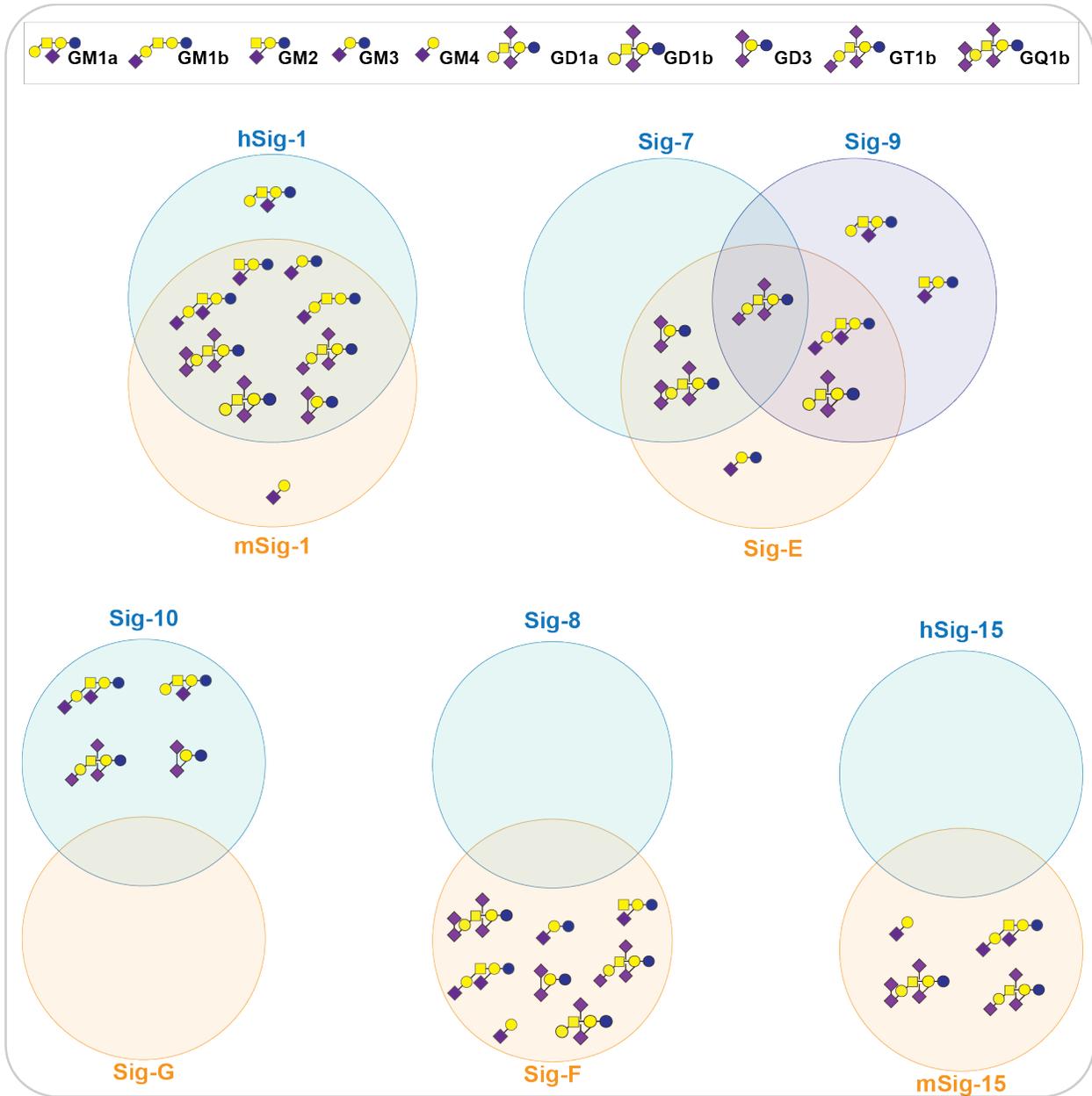


Figure 3.24: Venn diagram comparison of murine and human Siglec ganglioside binding. Interactions reported are a summary of interactions identified in this thesis.

This is surprising given that m/hSiglec-15 share strong sequence homology (V-set identity 94%). However, as it was found in chapter 2, our first generation hSiglec-15-Fc was unstable and

binding was likely not observed due to inactive protein rather than differences in ligand preferences between the two Siglecs. Moreover, Siglec-15 required a K274L mutation to be expressed in as a membrane bound Siglec and while cell surface expression of hSiglec-15 was achieved, it may not have been at levels sufficient to observe ganglioside liposome binding. Alternatively, Siglec-15 may be masked on CHO cell which caused *trans* binding by ganglioside liposomes to be missed. Regardless of the reason, this work could be extended by repeating the hSiglec-15 ganglioside binding using the second generation hSiglec-15-Fc. Regardless, mSiglec-15 appeared to have a strong preference for a terminal $\alpha(2\rightarrow3)$ linked Neu5Ac as demonstrated by its ability to bind GM4, GD1a and GT1b while not binding internally $\alpha(2\rightarrow3)$ linked gangliosides such as GM1a and GM2. This is in line with previous investigations that also observed Siglec-15 binding to $\alpha(2\rightarrow3)$ linked sialic acids^{79, 80}.

For the CD33-related Siglecs, while hCD33 and mCD33 share a name, the two Siglecs are structurally and functionally different²⁴³. With respect to hCD33, ganglioside binding has only been observed using approaches where the ganglioside is outside a bilayer²⁹². In agreement with this, we observed only minimal binding between mCD33 and GM3 in the ELISA but not in the bead assay. For orthologs, Siglec-7, Siglec-9, and Siglec-E, all bound most gangliosides in this study and in other investigations^{128, 140}. Previously, Siglec-10 was found to bind many gangliosides^{128, 177} whereas in this study, Siglec-G was unable to bind any gangliosides regardless of presentation. With respect to the paralogs Siglec-F and Siglec-8, Siglec-8 has not been reported to bind gangliosides when presented from a liposome¹²⁸ but can bind gangliosides outside a bilayer¹⁴⁰. In this study, Siglec-F could bind all the gangliosides that Siglec-8 was reported to bind to in an ELISA but could also bind GM4, GD1a, and GD1b when presented from a liposome. Both the bead assay and the ELISA are robust assays for measuring Siglec-ganglioside interactions. However, it is important to consider that these assays are artificial and due to the avidity leveraged in both platforms, the observed interactions, and effects of ganglioside content on Siglec binding need to be validated in a biological context, which could be the focus of future studies. To this,

using cells engineered to produce specific gangliosides or exogenous loading of gangliosides to cells may be fruitful approaches to validate Siglec-ganglioside interactions in a more biologically relevant context.

3.4.4: Mouse models are ideal for studying human Siglec-ganglioside interactions

When comparing the ganglioside binding profiles between hSiglecs and mSiglecs it is important to consider that all the gangliosides used in this study contained Neu5Ac as the form of sialic acid, and not Neu5Gc. Mice produce Neu5Gc in addition to Neu5Ac as their sialic acid and Neu5Gc-containing gangliosides could be preferentially recognized by some mSiglecs over their Neu5Ac-containing ganglioside counterparts^{159, 293}. While it would have been ideal to test Neu5Gc-containing gangliosides, unfortunately these are not readily available, and their synthesis is non-trivial. How Neu5Gc on ganglioside affects Siglec binding is an excellent future direction. Nevertheless, the use of Neu5Ac gangliosides enabled the direct comparison of the results from this work to our previous study where we analyzed binding of human Siglecs to gangliosides in similar types of assays.

3.4.5: GM1a is a poor ligand for mSiglec-1

GM1a is an important molecule in mammalian physiology and pathophysiology as it is one of the most common gangliosides^{62, 294}. In Huntington's Disease murine models, delivering GM1a can restore healthy motor functions²⁹⁵. With GM1a involved in many physiological and pathophysiological processes, it was surprising that none of the murine Siglecs bound GM1a in a bilayer. Even re-optimizing the liposome formulation with respect to cholesterol content, acyl chain length, and ganglioside content did not reveal any binding between mSiglec-1 and GM1a. However, these optimizations are not exhaustive and perhaps in the context of a *bonafide* biological membrane, mSiglec-1 may be able to interact with GM1a as binding between mSiglec-1 and GM1a could be observed in the mass-spectrometry assay. However, the interaction between mSiglec-1 and GM1a was the weakest interaction measured. While mSiglec-1 does have a weaker affinity to GM1a compared to hSiglec-1, the difference is not large enough to

explain the observed difference in the bead assay and the ELISA and is more likely due to other factors such as the oligosaccharide presentation.

In our focused ELISA, hSiglec-1 bound GD1a with greater avidity than GM1b whereas mSiglec-1 bound the two gangliosides equally. This may be because hSig-1 can bind to the internal and terminal Sia residues whereas mSiglec-1 can only bind the terminal Sia residue. In the GM1-focused bead assay, both hSiglec-1 and mSiglec-1 bound GD1a with greater avidity than GM1b. This may be due to GM1b having a different optimal ganglioside content than GD1a, resulting in an unfavourable oligosaccharide conformation and decreased engagement of the Siglec. Alternatively, the branched nature of GD1a may position the terminal Sia towards the solvent allowing for engagement by Siglec-1, whereas the linear GM1b oligosaccharide may adopt a more buried or 'laid down' conformation proposed for other linear gangliosides like GM3²⁸⁶. In the future, techniques such as STD-NMR or co-crystal structures can be used to directly assess which sialic acid residues are recognized by Siglecs. Overall, it appears that Siglec-ganglioside binding profiles are reasonably conserved between human Siglecs and their equivalent murine Siglecs.

3.5: Conclusion

The foundation of our understanding of the physiological and pathophysiological roles of Siglecs are rooted in studies that discovered and described Siglec ligands. Through screening the entire human and murine Siglec family against a panel of gangliosides presented in a lipid bilayer within an optimized liposome revealed many previously undiscovered Siglec–glycolipid interactions, most notably between Siglec-6 and several gangliosides. Additionally, while many Siglecs are conserved or may serve the same function at the organism or cellular level in humans and mice, subtle differences in structure may have profound biological effects that need to be considered. The results from this chapter suggest that murine models are generally appropriate to study human Siglec-ganglioside interactions at a global level but are less appropriate when studying specific interactions such as Siglec-1 and GM1a.

3.6: Methods

Cell culture. CHO cells were cultured as described in Chapter 2.

Stable Transfection. Full length Siglecs were transfected as described in Chapter 2.

Siglec-Fc expression and purification. Siglec-Fcs were expressed and purified as described in Chapter 2.

Liposome Extrusion. Liposomes were produced as described in chapter 2.

Siglec preparation for direct binding assay (Performed by the Klassen Group). Siglec fragments were produced as described in chapter 2.

Direct ESI-MS Binding Assay. Direct ESI-MS Binding Assay was performed as described in chapter 2.

Cell Assay. 200,000 cells/well were plated into a 96-well U-bottom microplate and centrifuged at 300 x g for 5 min. The cells were then resuspended in 50 μ L of 50 μ M liposome/50 μ g/mL (protein concentration) of EV solution in DMEM F12 5 % FBS and incubated at 37 °C for 30 min. Excess liposomes were then removed via centrifuged at 300 x g for 5 min followed by washing the cells with FACS solution. The cells were then resuspended in (1:250 V/V) anti-Siglec-flow buffer (1 % V/V FBS, 500 μ M, EDTA Hank's Balanced Salt Solution pH 7.4), solution and left to rest on ice for 30 min. The cells were then centrifuged again at 300 x g for 5 min and the cells were resuspended in FACS buffer solution and were then ready for analysis by flow cytometry. Information regarding antibody clone, catalog number supplier, label isotype and dilution can be found in chapter 2.

Flow cytometry. Flow cytometry measurements were collected as described in chapter 2.

Ganglioside Content Quantification (Performed by the Klassen Group). The average ganglioside content in a given liposome sample was measured by electrospray ionization mass spectrometry (ESI-MS) using an internal standard (IS)²⁹⁶. Briefly, the liposome sample was first disassembled in methanol solution of 0.15% formic acid and a known amount of IS added. The solution was then analyzed by ESI-MS. The total ion abundance (*Ab*) ratio of the ganglioside

(GSL) to the IS measured by ESI-MS is related to corresponding solution concentration ratio by **equation 3.1**:

$$R = \frac{Ab(GSL)}{Ab(IS)} = \frac{[GSL]}{[IS]} \quad (3.1)$$

and thus the ganglioside concentration can be obtained from linear regression. The corrected ganglioside percentage (GSL %_{corr}) and the ganglioside incorporation in the liposome can be found from **equation 3.2** and **3.3**, respectively, where [GSL]_{nominal} is the nominal concentration of the ganglioside used for liposome preparation:

$$GSL \%_{corr} = \frac{[GSL]}{[Total Lipid]} \quad (3.2)$$

$$Incorporation = \frac{[GSL]}{[GSL]_{nominal}} \quad (3.3)$$

The corresponding deuterium labelled gangliosides, N- ω -CD3-octadecanoyl monosialoganglioside GM2 (GM2-d3), and N- ω -CD3-octadecanoyl monosialoganglioside GM3 (GM3-d3), purchased from Matreya LLC (State College, PA), were used as IS for GM2 and GM3 content quantification, respectively. All these measurements were carried out in negative ion mode using a Q Exactive Hybrid Quadrupole-Orbitrap mass spectrometer coupled with the Nanospray Flex Ion Source (Thermo Fisher Scientific, Bremen, Germany). Tips pulled from a borosilicate capillary (1.0 mm o.d., 0.78 mm i.d.) by a micropipette puller (P-1000, Sutter Instruments, Novato, CA) were used to perform the nanoflow ESI. The sample solution was loaded into the nanoflow ESI tip and a voltage of -0.7 kV was applied to a platinum wire inserted into the tip and in contact with the sample solution. For the Orbitrap mass spectrometer, the key parameters were: capillary temperature 160 °C, maximum inject time 100 ms, microscans 2, and resolution 140,000. All other parameters were set at default values. Data acquisition and processing were performed using Xcalibur (Thermo Fisher Scientific, version 4.4).

Direct ESI-MS Binding Assay (*Performed by the Klassen Group*). The affinities of ganglioside (GM1, GM2, GM3 and GD1a) oligosaccharides (L) for Siglec-1 fragment were measured by the direct ESI-MS assay.¹ A reference protein (P_{ref}) was added to the ESI solutions in order to correct mass spectra for any nonspecific binding that occurred during the ESI process.² The dissociation constant (K_d) was calculated from the total abundance (Ab) ratio (R) of the ligand bound (PL) to free Siglec-1 fragment (P) ions (**equation 3.4**) measured by ESI-MS for solutions of known initial concentration of Siglec-1 fragment (P_0) and ligand (L_0), **equation 3.5**:

$$R = \frac{Ab(PL)}{Ab(P)} = \frac{[PL]}{[P]} \quad (3.4)$$

$$K_d = \frac{[P][L]}{[PL]} = \frac{[L]_0}{R} - \frac{[P]_0}{R + 1} \quad (3.5)$$

The reported K_d values correspond to average values measured at 3.6 μ M of Siglec-1 fragment and 20, 40, 80, and 140 μ M of each ganglioside oligosaccharide tested. Direct ESI-MS measurements were performed with nanoflow ESI in positive ion mode (voltage \sim 1 kV) on a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific). Automatic gain control (AGC) target, the maximum inject time, capillary temperature and S-lens RF level were set to 1×10^6 , 100 ms, 150 $^{\circ}$ C and 100 $^{\circ}$ C, respectively. The resolution was 17500 at m/z 200. Data acquisition and processing were carried out using Xcalibur (Thermo Fisher, version 4.1).

Modification of Phage Clones with Glycans (*Performed by the Derda Group*). Solution of five SDB phage clone (10^{12} - 10^{13} PFU/mL in PBS) was combined in equal amounts (50 μ L) to create a multiplexed silent barcode (MSDB). Twenty-four such MSDB were created and combined DBCO-NHS (50 mM in DMF) to afford final concentration ranging from 0.25 mM to 2 mM. The reaction was incubated for 45 min at room temperature. The conjugates were purified by ZebaTM Spin Desalting column (7 kDa MWCO, 0.5 mL, Thermo Fisher) and pVIII modification rate was confirmed by MAL/DI using a previously reported protocol. Typically, 1 mM DBCO-NHS yields 25% of pVIII modification after 45 min incubation. Six MSDB was combined and solution of azido-

glycans (10 mM stock in Nuclease Free H₂O) were added to the solution to afford a 2 mM concentration of glycan-azide and the solutions were further incubated overnight at 4 °C. The glycan conjugation was confirmed using MALDI-TOF. The conjugates were purified by Zeba column and supplemented with glycerol and stored as 50% glycerol stock at -20 °C. LiGA mixture was prepared by combining these solutions.

Binding of LiGA to CHO Cells expressing Siglec-1 (*Performed by the Derda Group*). Confluent CHO cells expressing Siglec-1 or UT CHO cells were detached using PBS plus 5 mM EDTA and washed with PBS (2x5 mL). Suspension of at 2 x10⁶ cells in HEPES-1, % BSA (20 mM HEPES, 150 mM NaCl, 2 mM CaCl₂, pH 7.4, 1% BSA) in a round bottom 3 mL tube (Corning, #352054) was combined with LiGA (10⁸ PFU) and incubated for 1 h at 4 °C . The cells were then washed HEPES-0.1 % BSA in (2x3 mL) and HEPES buffer (1x1 mL). The washed cell pellet was resuspended in 30 µL nuclease free H₂O. The solution was incubated at 90 °C for 15 min, centrifuged at 21,000 x g for 10 min, and 25 µL of the supernatant was used for PCR and sequencing.

Liposome Ligand Density Calculation. The total number of lipids in a liposome (N_L) were calculated using **equation 3.6**. Liposomes were assumed to be a sphere 100 nm in diameter (d). The thickness of a DSPC bilayer (h) was assumed to be 5 nm. The area of a phosphatidyl choline head group (b) was assumed to be 0.71 nm².

$$N_L = \frac{4\pi \left(\frac{d}{2}\right)^2 + 4\pi \left[\frac{d}{2} - h\right]^2}{b} \quad (3.6)$$

The ligand density (LD) was then calculated using **equation 3.7** where χ is the appropriate mol% of ganglioside in the liposome formulation assumed that 50 % is on the outer leaflet of the bilayer, N_L was calculated using equation 3.6, N_A is Avogadro's number and d is the diameter of the liposome.

$$LD = \frac{\left(\frac{\chi}{2}\right)(N_L)(N_A)}{4\pi \left(\frac{d}{2}\right)^2} \quad (3.7)$$

Unmasking Cell Assay. Cells were treated with neuraminidase as described in Chapter 2.

Bead Assay. The bead assay was performed as described in Chapter 2.

ELISA. The ELISA was performed in the same way as described in Chapter 2.

Chapter 4: Exploring the ligands of and functions of Siglec-6

4.1: Acknowledgements

Portions of this chapter were published as: Schmidt, E.N.; *et al.* Siglec-6 mediates the uptake of extracellular vesicles through a noncanonical glycolipid binding pocket, *Nat. Commun.* (2023).

I would like to thank Xue Yan Guo for assistance with performing liposome binding assays and site directed mutagenesis, Dr. Maju Joe and Fahima Mozaneh for synthesizing and characterizing neoglycolipids; Prof. Elisa Fadda for their expertise in molecular modeling; Dr. Kelli McCord for virally transducing Siglec-6 expressing B cell lines and performing imaging flow cytometry; Prof. Lori J. West for providing human spleen samples; Professor Marianne Kulka for providing LAD2 cells, Prof. Meghan R. Riddell, Jasmine Nguyen, and Dr. Dimitra Lamprinaki for performing binding experiments with syncytiotrophoblasts; Dr. Dimitra Lamprinaki under the supervision of Prof. Lara K. Mahal for isolation and characterization of EVs; Amanda R. Krysler, Christopher R. Cromwell, under supervision of Prof. Basil P. Hubbard for the generation of β GalNT1^{-/-} N2a cells and John Monyor under the supervision of Prof. Simonetta Sipione for production and isolation of EVs from these cells.

I would also like to acknowledge all those who donated tissues and blood. Much of this work would be impossible without them.

4.2: Introduction

4.2.0: Overview of Siglec-6

Siglec-6 - also known as CD327 or OB-BP1 - is a member of the CD33-related Siglecs. Siglec-6 was first discovered in the late 1990's being expressed in the human placenta²⁹⁷. While Siglec expression is generally restricted to immune cells, Siglec-6 breaks away from this dogma; in addition to being expressed on immune cells²⁹⁸, Siglec-6 is also expressed on placental syncytiotrophoblasts. Interestingly, Siglec-6 expression in the placenta is unique to humans²⁰⁵. In addition to being expressed by syncytiotrophoblast, Siglec-6 is also expressed by memory B cells^{128, 299}, and mast cells^{128, 298}. Compared to other Siglecs, there are relatively few investigations into Siglec-6 and its ligands and, consequently, its biological roles are poorly understood. After discovering that Siglec-6 can bind gangliosides, we aimed to leverage this discovery to probe deeper into the biological roles of Siglec-6. Using our optimized liposome platform, we discovered that bilayer presentation is critical for sialoside recognition by Siglec-6, which is mediated by a solvent exposed tryptophan residue. Moreover, this tryptophan residue contributes more to the binding of liposomes than the conserved arginine residue. Additionally, we found that Siglec-6 preferentially binds $\alpha(2 \rightarrow 3)$ and $\alpha(2 \rightarrow 8)$ linked sialosides over $\alpha(2 \rightarrow 6)$ linked sialosides.

4.2.1: Siglec-6 structure

Like all Siglecs, Siglec-6 has the typical structural components: a glycan binding or V-set domain, two C2 domains, a transmembrane domain and a cytoplasmic tail featuring an immunoreceptor tyrosine-based inhibitory motif (ITIM) and an ITIM-like motif. Unfortunately, there is no crystal structure to consult when describing the structure of Siglec-6. However, recent advances in bioinformatics have enabled the development of AlphaFold; which, is a software that combines machine learning with published crystal structures to predict the structure of proteins^{300, 301}. While AlphaFold is impressive and reasonably accurate, it does have its limitations and is not a replacement for crystal structures but can be helpful when lacking a crystal structure³⁰². Using the AlphaFold model of Siglec-6, the different structural components can be highlighted (**Figure**

4.1). Using this model, the critical structural components of the Siglec can be observed. First there is the V-set or glycan binding domain. There are several important amino acids in this domain. These residues include Cys46 (one half of the interdomain disulfide bridge connecting the V-set to the first C2 domain) and an ionic pair of Arg92 and Asp115, which forms a salt bridge critical for proper protein folding. In this domain, there is also the conserved arginine residue (Arg122) and Trp127 which is important for binding glycolipids presented from a lipid bilayer. Key residues in the first C2 domain of Siglec-6 are Asn149, 163, and 223 all of which are predicted to be *N*-glycosylation sites, as well as the other half of the cysteine bridge with the V-set domain (Cys172). There is also an intradomain disulfide bond between Cys166 and 215. In the last C2 domain, there is another interdomain disulfide bridge between Cys274 and 319. After a short, disordered spacer, there is the transmembrane domain that anchors the protein into the membrane. The last part of the Siglec is the cytoplasmic tail which is structurally disordered and contains the ITIM and ITIM like motifs which have key tyrosine residues at the 426 and 445 positions respectively.

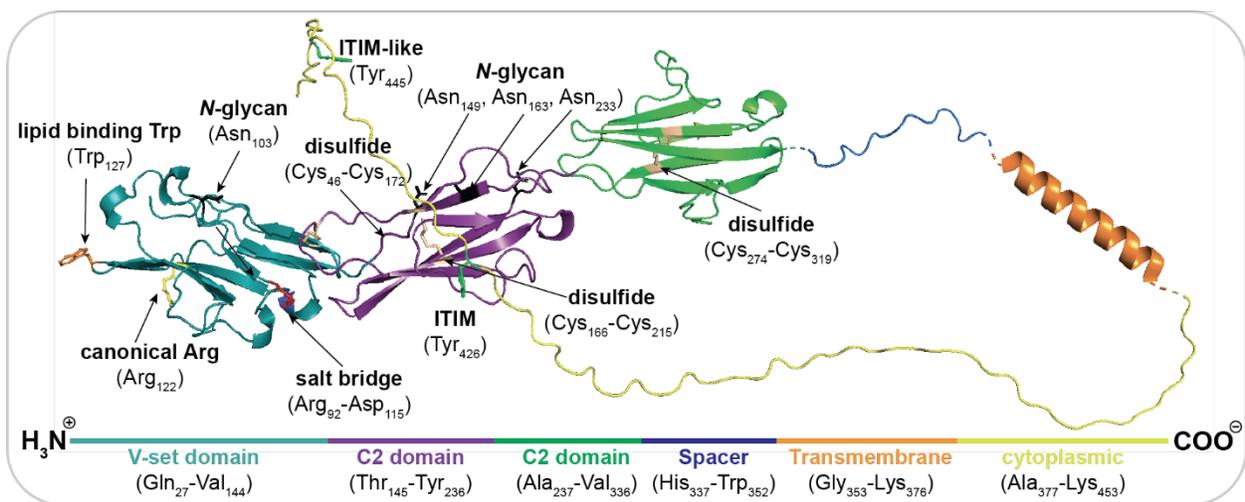


Figure 4.1: Model of Siglec-6 broken down into domains with important amino acids highlighted.

4.2.2: Siglec-6 expression

Siglec-6 can be found on memory B cells, mast cells, and syncytiotrophoblasts. Memory B cells are a subset of B cells and a component of the adaptive immune system which serve to help protect the organism from repeated infection from the same pathogen through the production of antibodies³⁰³. General markers used to identify B cells are CD19 and CD22. Memory B cells

can be isolated from other types of B cells using the marker set CD27⁺IgD⁻CD38⁻³⁰⁴. During the initial infection or vaccination, B cells that are specific for the pathogen will be selected for in the germinal center of a lymph node or other secondary lymphoid organs through a process known as somatic hypermutation¹³³. Once a B cell makes it through the selection process, it will either become a memory B cell or an antibody-secreting plasma cell. After the infection has cleared, memory B cells will remain in circulation in waiting for decades³⁰³. When a memory B cell encounters the pathogen it is specific for, it will begin to proliferate and differentiate into antibody producing plasma cells, neutralizing the pathogen. Exactly when in the process Siglec-6 expression begins and why Siglec-6 only appears on memory B cells, and not naïve B cells, remains unclear, however, the restricted nature of this expression makes it an excellent target for diseases such as acute myeloid leukemia (AML)³⁰⁵.

Siglec-6 is also expressed by mast cells, an immune cell which participates in both the innate and adaptive immune responses and are typically found in tissues and mucosal membranes (respiratory track, skin, gastro-intestinal tract, *etc.*)³⁰⁶. Occupying these physiological locations makes mast cells part of the first line of defense against many pathogens and allergens. Mast cells are most known for their roles in inflammation and allergic reactions. When mast cells are activated by the cross-linking of FcεRI-IgE complexes due to the presence of an antigen, mast cells release many signaling molecules (cytokines, thromboxanes, chemokines, *etc.*) that induce an anaphylactic response/inflammation³⁰⁷. Mast cells are also believed to be relevant in many autoimmune diseases/conditions, such as insulin independent diabetes mellitus, multiple sclerosis, and rheumatoid arthritis to name a few³⁰⁸. Due to their roles in many diseases and conditions, being able to modulate mast cells is of great therapeutic interest³⁰⁹.

Lastly, Siglec-6 can also be expressed on placental syncytiotrophoblasts. These multinucleated cells form the outer most layer of the placenta and are in direct contact with the maternal blood. Interestingly, while other primates express Siglec-6 on their B cells, Siglec-6 expression by syncytiotrophoblasts has only been identified in human placentas^{205, 310}. While the

niche filled by Siglec-6 expressed by syncytiotrophoblasts remains unclear, it is known that Siglec-6 expression increases in pregnancies and is affected by preeclampsia. Preeclampsia is a complication during pregnancy that is characterized by abnormally high levels of protein in the urine (proteinuria) and high blood pressure (hypertension)³¹⁰. One in twenty pregnancies are affected by preeclampsia and it is one of the leading causes of maternal/fetal deaths³¹¹.

In this chapter, the ability of Siglec-6 to engage glycolipids independent of its conserved arginine was investigated. This was accomplished through a combination of genetic means (chimera constructs and mutagenesis) as well as chemical means (a panel of synthetic glycolipids). This investigation led to the discovery of a relatively high affinity Siglec-6 ligand which, when formulated into a liposome could engage Siglec-6 on a variety of primary human cells. Liposomes formulated with this ligand were also used to demonstrate that Siglec-6 could mediate the internalization of these nanoparticles. Lastly, a possible biological role for Siglec-6-ganglioside interactions was proposed when it was found that Siglec-6 could also mediate the internalization of extracellular vesicles.

4.3: Results

4.3.1: Siglec-6 ligand presentation

The robust binding between CHO cells expressing Siglec-6 and ganglioside-bearing liposomes observed in chapter 3 came as a surprise as Siglec-6 was previously reported to bind $\alpha(2\rightarrow6)$ linked sialosides such as the *sialyl Tn antigen* and none of the gangliosides used in our study featured an $\alpha(2\rightarrow6)$ linked sialic acid motif¹⁶⁶. Moreover, GD3 was found not to be a ligand for Siglec-6 using an ELISA in a previous study¹⁶⁶ but it was found to be a ligand for Siglec-6 in our cell-based assay. One explanation for this difference is that the oligosaccharide presentation afforded by the liposome is important for Siglec-6 recognition of glycolipids. To assess how the environment of the glycolipid affects the binding between Siglec-6 and glycolipids, we probed the binding between Siglec-6 and glycolipids in a variety of different experimental approaches.

4.3.1.0: Siglec-6-ganglioside binding outside the context of a bilayer

We started by probing the ability of Siglec-6 to bind gangliosides outside a lipid bilayer using an ELISA approach. As a control, we also evaluated the ability of hSiglec-1 and -7 to bind gangliosides in an ELISA together with Siglec-6, as hSiglec-1 and -7 are both established ganglioside binders and have been shown to bind gangliosides outside of a bilayer^{249, 251}. Using an ELISA, where gangliosides were adsorbed to a microplate and probed with our soluble dimeric Siglec-Fc proteins precomplexed with tetrameric Strep-Tactin-Horseradish Peroxidase (HRP)³⁷, hSiglec-1 and -7 showed similar ganglioside binding profiles in the ELISA and the cell assay. On the other hand, Siglec-6 did not recognize the gangliosides that it interacted with in the cell assay (GM1a, GD1a, GD3, GT1b, and GQ1b) (**Figure 4.2**). Interestingly, there was a small amount of binding of Siglec-6 to GM3 and GM4 in the ELISA and, notably, the same binding profile was observed between WT Siglec-6 and R122A Siglec-6. The inability of Siglec-6 to recognize gangliosides when presented outside the confines of a bilayer supports that bilayer presentation of Siglec-6 ligands is important for recognition.

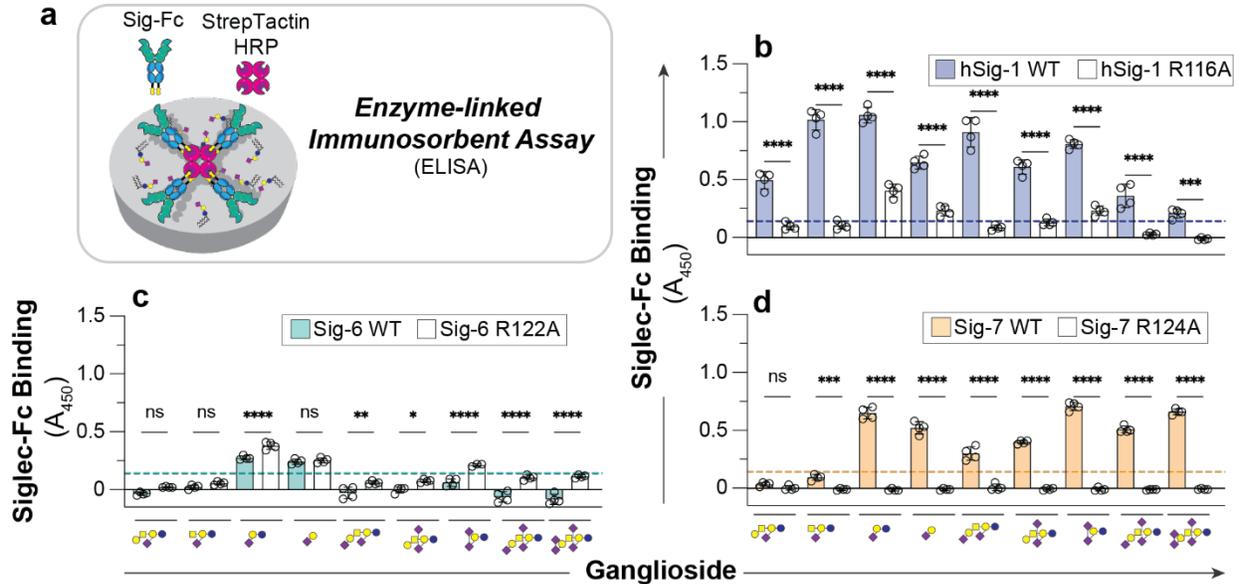


Figure 4.2: Assessing the ability of Siglec-6 to bind gangliosides outside of a bilayer. **a**, Schematic of the enzyme-linked immunosorbent assay (ELISA) used to investigate Siglec–ganglioside interactions outside the context of a lipid bilayer. **b**, **c**, and **d**, ELISA results of hSiglec-1, -6, and -7 binding to nine gangliosides respectively. All results are represented as the mean \pm one standard deviation of at least four technical replicates. The dotted line represents two standard deviations above the blank well. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; $0.05 > P \geq 0.01$; ** $0.01 > P \geq 0.001$; *** $0.001 > P \geq 0.0001$; **** $P < 0.0001$

4.3.1.1: Siglec-6-Fc does not bind in cell-based glycan array

Siglec-6 binding to glycolipids was dependent on how its ligands were presented, as it did not robustly bind to gangliosides in the ELISA but did when displayed from a lipid bilayer in the cell assay. In Chapter 2, a cell-based glycan array was used to probe Siglec ligands on primary cells and no Siglec-6 ligands were found on any of the immune cell types that were screened (B cells, NK cells, $CD4^+$ and $CD8^+$ T cells, monocytes, and mature neutrophils). As Siglec ligands are often found on cells that express that Siglec (e.g. $CD22$ ligands are found on B cells), we hypothesized that cell lines derived from biologically relevant cells may carry Siglec-6 ligands. Accordingly, we screened BeWo and Daudi cell lines against the Siglec-6-Fc in a cell-based glycan array. These cell lines were chosen because Siglec-6 is expressed by syncytiotrophoblast and memory B cells and BeWo and Daudi cells are trophoblast and B cell derived cell lines, respectively. As well, BeWo cells were previously reported to have Siglec-6 ligands²⁰⁵. However, in our hands, Siglec-6 did not bind either cell line. It is unclear if this is because these cells do not

express the right types of glycans or if glycolipids are too buried in the glycocalyx to be recognized by soluble Siglec-6-Fc.

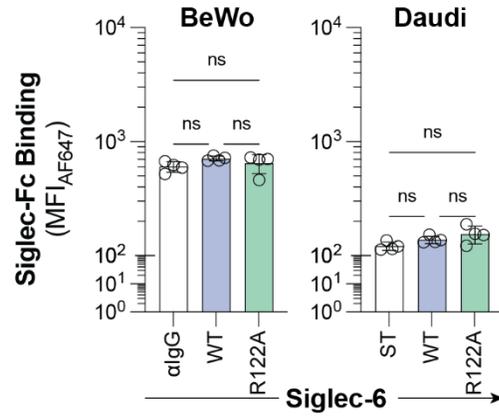
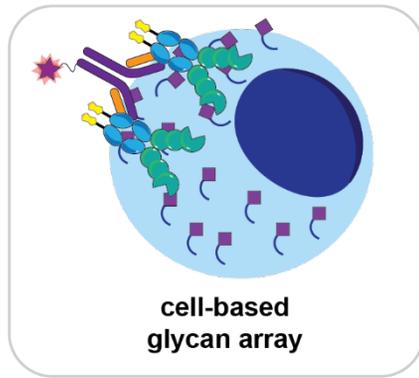


Figure 4.3: Siglec-6-Fc binding to cell lines in a cell-based glycan array. The cell lines tested were BeWo cells (a trophoblast cell line and Daudi cell (a B cell line). Data is represented as the mean \pm one standard deviation of four technical replicates. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$.

4.3.1.2: Bilayer presentation rescues Siglec-6 ganglioside binding

To further investigate how ganglioside presentation influences Siglec binding, we used two other liposome-based assays. The first we refer to as the Liposome Over Lectin Assay (LOLA), wherein a Siglec-Fc is adsorbed to a microplate, followed by probing with the same fluorescently labeled ganglioside liposomes used in the cell assay (**Figure 4.4a**). Using hSiglec-1 and GM1a bearing liposomes, the optimal amount of Siglec adsorbed in the well was determined to be 1 μg per well (**Figure 4.4b**). The results of the cell-based assay were repeated with respect hSiglec-1 binding to ganglioside bearing liposomes and, as expected, hSiglec-1 bound GM1a bearing liposomes in an Arg116 dependant manner (**Figure 4.4c**) and bound GM1a, GM2, and GM3 bearing liposomes with the same relative avidity (GM2>GM1a>GM3) as the cell-based assay (**Figure 4.4d**).

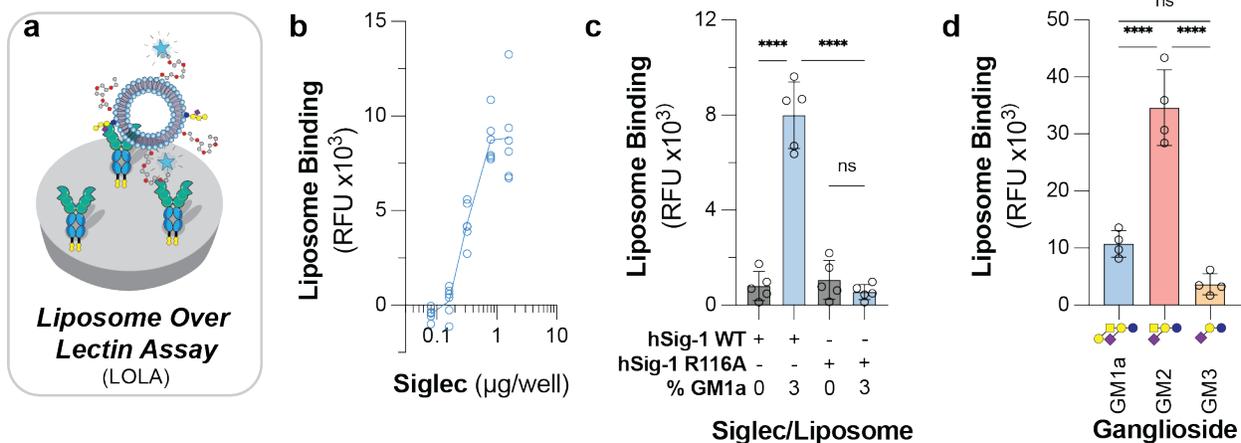


Figure 4.4: Development of Liposome Over Lectin Assay (LOLA). **a**, Depiction of the LOLA where the Siglec-Fc is adsorbed non-specifically to a microplate and ganglioside bearing liposomes are added to the wells and binding is determined via fluorescence intensity by a spectrophotometer. **b**, Binding of GM1a liposomes to hSiglec-1 adsorbed in different amounts to the microplate wells. **c**, Binding of naked liposomes and liposomes formulated with 3 mol% GM1a to WT and Arg116A hSiglec-1. **d**, Binding of GM1a, GM2, and GM3 bearing liposomes to hSiglec-1 WT in the LOLA. All results are represented as the mean \pm one standard deviation of at least four technical replicates. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; $0.05 > P \geq 0.01$; $****P < 0.0001$.

With this assay developed, it could be used to evaluate if Siglec-6 could bind gangliosides when presented from a bilayer in a different experimental approach. Again, hSiglec-1 and Siglec-7 were used in parallel with Siglec-6. All three Siglecs showed similar binding in the LOLA

4.3.2: Probing Siglec-6 with synthetic glycolipids

A glycolipid can be broken down into three components, the oligosaccharide, the linker, and the acyl chains (Figure 4.6); all of which impact oligosaccharide presentation. To probe the specificity of Siglec-6 with respect to the glycan beyond which glycolipids are commercially available, we prepared a panel of synthetic glycolipids or neoglycolipids (nGLs) and systematically varied the oligosaccharide, headgroup/linker, and acyl chains and assessed the effect of each component on Siglec-6 binding.

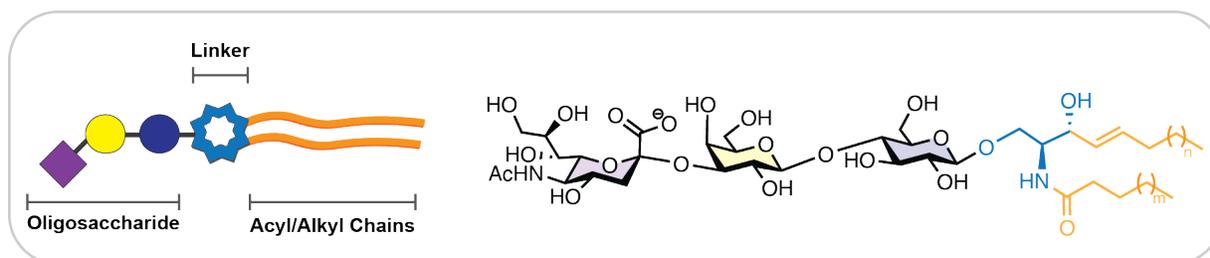


Figure 4.6: Schematic of a glycolipid broken down into three components. A glycolipid can be broken down into three components: the oligosaccharide, the linker, and the acyl/alkyl chains of the lipid.

4.3.2.1: Siglec-6 binding is dependent on oligosaccharide presentation

Starting with the glycan moiety, nGLs **3** and **4** were prepared, which feature the oligosaccharide moiety of GM1a and GM3, respectively, linked to 1,3-di-O-hexadecyl glycerol via an amide linker (Figure 4.7a). Neoglycolipid liposomes formulated with 3 mol% **3** bound minimally to Siglec-6 CHO cells, whereas liposomes containing 3 mol% **4** showed a five-fold increase in binding compared to native GM3 ganglioside liposomes (Figure 4.7b).

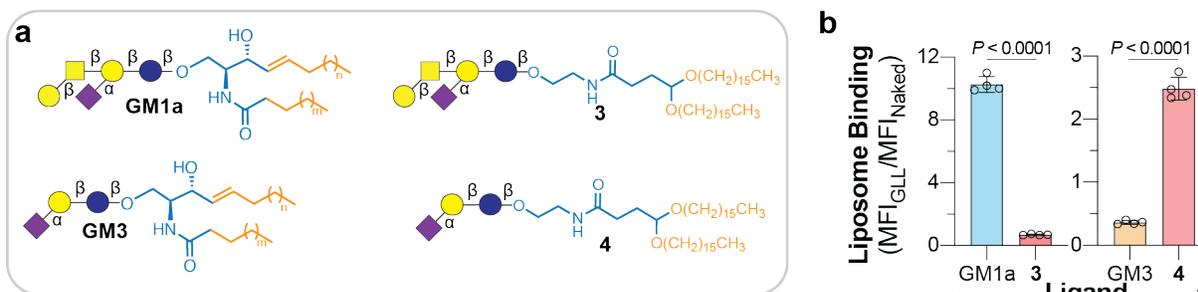


Figure 4.7: Binding of neoglycolipids present from an amide dialkyl-glycerol scaffold. a, Structures of nGLs **3** and **4** presenting the oligosaccharide of GM1a and GM3, respectively, through an amide-linkage to a 1,3-di-O-hexadecyl glycerol scaffold. b Binding of liposomes formulated with 3 mol% **3** and **4** to WT Siglec-6 CHO cells in the cell assay. Liposome binding was quantified as the binding of the liposome bearing the glycolipid over the binding of a naked liposome. Data is represented as the mean \pm one standard deviation of at least four technical replicates. A one-way ANOVA was used for statistical analysis.

As the oligosaccharide of GM3 (α -Neup5Ac-(2 \rightarrow 3)- β -lactose) was preferred in this artificial presentation, the chemical diversity of our nGL panel was increased by linking the GM3 oligosaccharide to three different lipid groups by coupling the β -azidoethyl glycoside of α -Neup5Ac-(2 \rightarrow 3)- β -lactose to three different lipids through a triazole linkage to form nGLs **5**, **6**, and **7** (**Figure 4.8a**). Liposomes formulated with 3 mol% **5** displayed relatively high binding to Siglec-6, whereas **6** and **7** showed lower binding (**Figure 4.8b**).

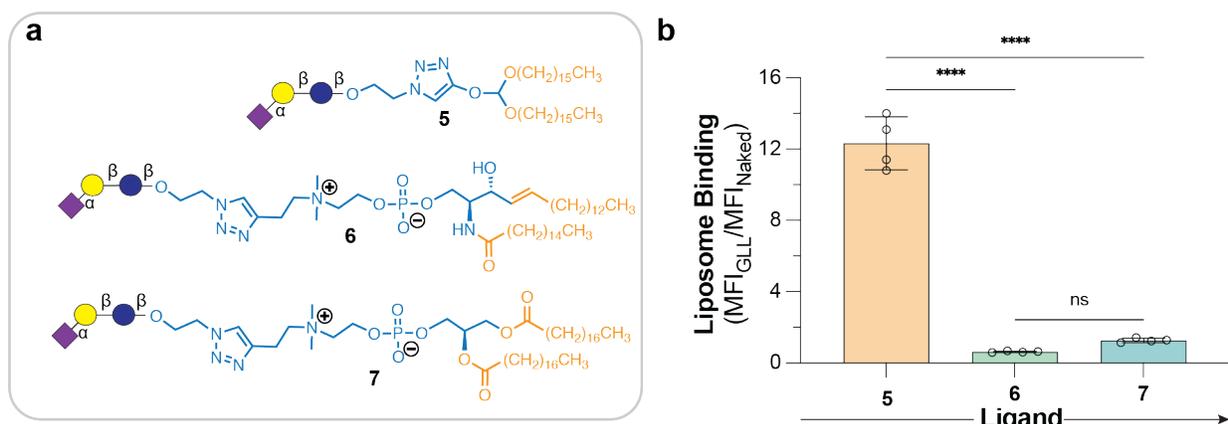


Figure 4.8: Binding of neoglycolipids presenting the GM3 oligosaccharide from triazole linked lipid scaffolds. **a**, Structures of nGLs **5**, **6**, and **7** presenting the oligosaccharide of GM3 triazole-linked to 1,3-di-O-hexadecyl glycerol, phosphatidyl sphingomyelin, and distearoylphosphatidylcholine scaffold, respectively. **b**, Binding of liposomes formulated with nGLs **5**, **6**, and **7** to WT Siglec-6 CHO cells in the cell assay relative to naked liposomes. Data is representative of the mean \pm one standard deviation of four technical replicates. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; **** $P < 0.0001$.

Using the triazole-linked di-O-hexadecyl glycerol nGL scaffold, we further explored the glycan specificity of Siglec-6 with respect to regiospecificity by synthesizing three nGLs with triazole-linked α (2 \rightarrow 6)-sialyl-lactose (**8**), α (2 \rightarrow 3)-sialyl-LacNAc (**9**), and α (2 \rightarrow 6)-sialyl-LacNAc (**10**) (**Figure 4.9a**). Compared to liposomes with 3 mol% **5**, both WT and R122A Siglec-6 bound minimally to the conjugates containing an α (2 \rightarrow 6)-linked sialoside (**Figure 4.9b**). Binding was also modestly lower when the lactose moiety of **5** was replaced with LacNAc (**9**).

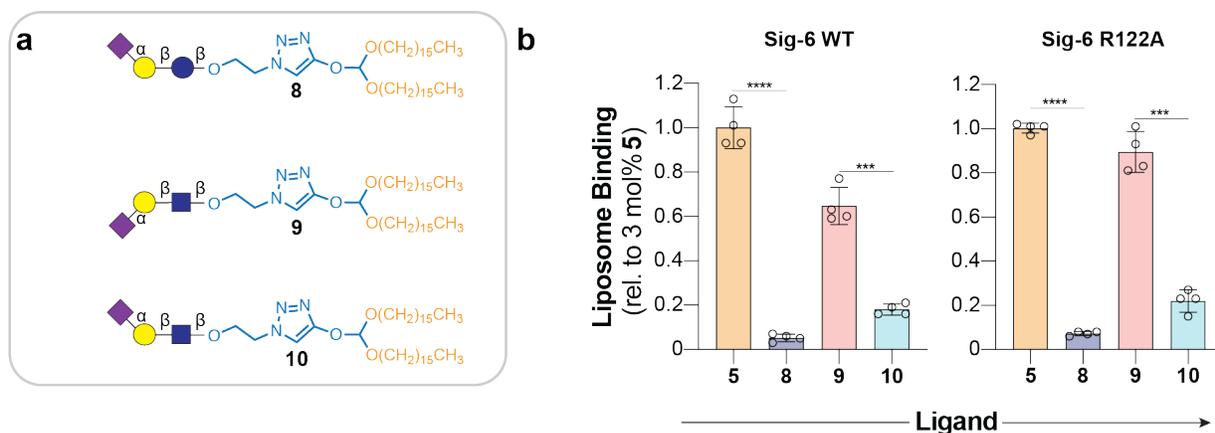


Figure 4.9: Binding of neoglycolipids featuring a triazole linked di-O-hexadecyl glycerol. a, Structures of nGLs 8, 9, and 10 presenting an $\alpha(2 \rightarrow 3)$ or $\alpha(2 \rightarrow 6)$ -linked sialoside on an underlying lactose or LacNAc core, triazole-linked to 1,3-di-O-hexadecyl glycerol scaffold. **b,** Binding of liposomes formulated with 5 and 8–10 to WT and R122A Siglec-6 CHO cells in the cell assay relative to liposomes formulated with 3 mol% nGL 5. Data is representative of the mean \pm one standard deviation of four technical replicates. A one-way ANOVA was used for statistical analysis. *** $0.001 > P \geq 0.0001$; **** $P < 0.0001$.

4.3.2.2: Optimizing neoglycolipid liposomes

As nGL 5 was to be used for targeting Siglec-6 in subsequent experiments, the content of 5 in liposomes was titrated and the optimal amount of nGL 5 was determined in three different binding experiments. The first was using Siglec-6 expressing CHO cells, the second was using Siglec-6-Fc complexed beads, and the last was using Raji cells (an immortalized cell line derived from human B cells) virally transduced with Siglec-6 (**Figure 4.10a-c**). In all cases, binding of liposomes formulated with nGL 5 plateaued at 5 mol% and accordingly, 5 mol% 5 was used in liposomes moving forward. However, as nGL 5 content exceeded 5 mol%, Siglec-6-independent binding could be observed in the Raji cells, suggesting that at higher ligand content, the liposomes can engage other Siglecs/lectins expressed by the Raji cells, demonstrating that nGL-5 is not specific for Siglec-6.

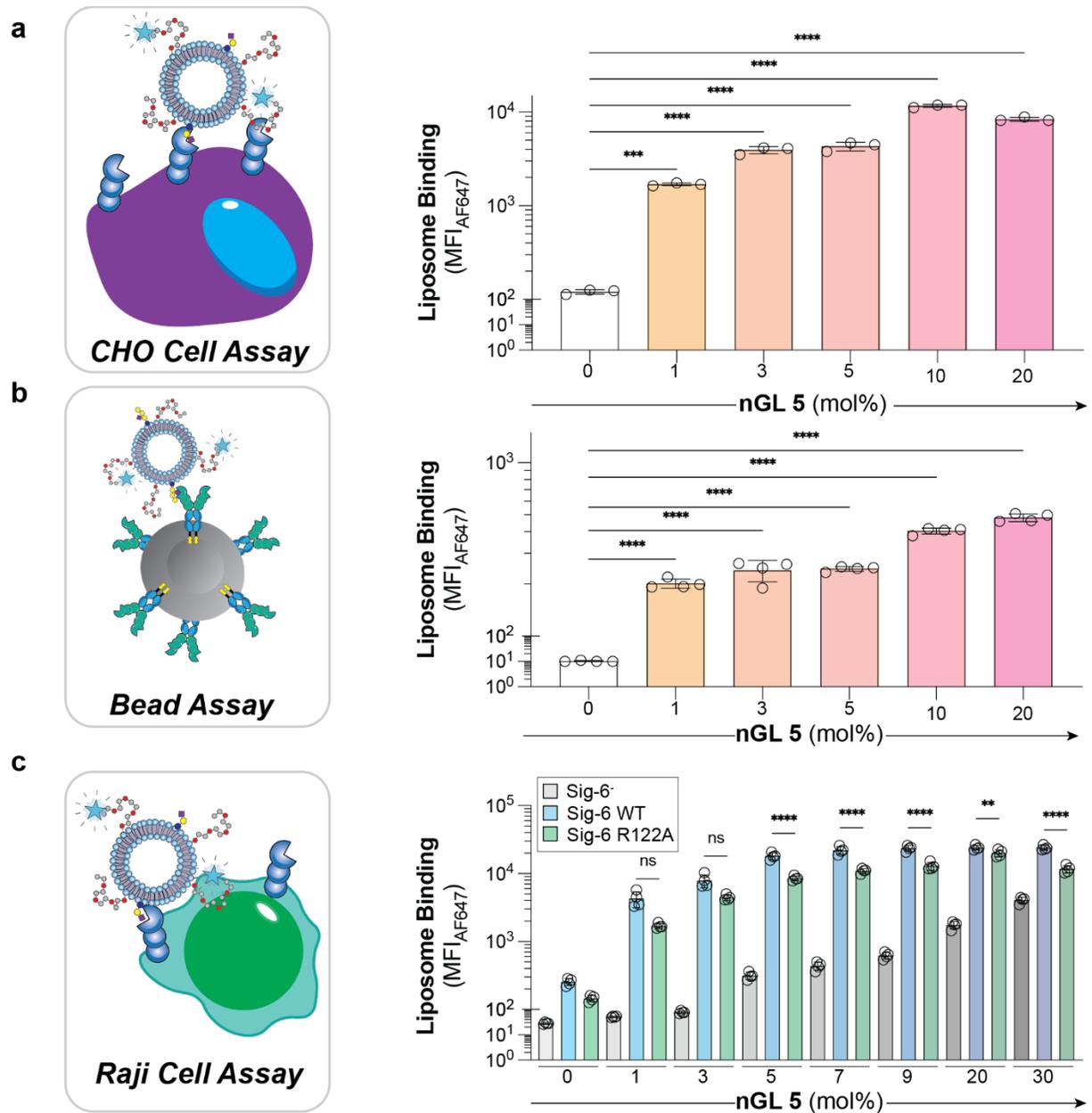


Figure 4.10: Optimizing the mol% of 5 in neoglycolipid bearing liposomes for engaging Siglec-6. **a, b, c,** Binding of liposomes formulated with increasing amounts of neoglycolipid 5 against CHO cells expressing WT Siglec-6, Siglec-6-Fc- complexed beads and Raji cells virally transduce with Siglec-6 respectively. Data is quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; ** $0.01 > P \geq 0.001$; *** $0.001 > P \geq 0.0001$; **** $P < 0.0001$

To further our understanding of how 5 neoglycolipid liposomes engage Siglec-6, we performed a competition assay between GM1a bearing liposomes and liposomes bearing nGL 5. The binding of GM1a ganglioside liposomes decreased as the concentration of 5 mol% 5

neoglycolipid liposomes increased, suggesting that the two ligands compete for the same binding site (**Figure 4.11a and b**).

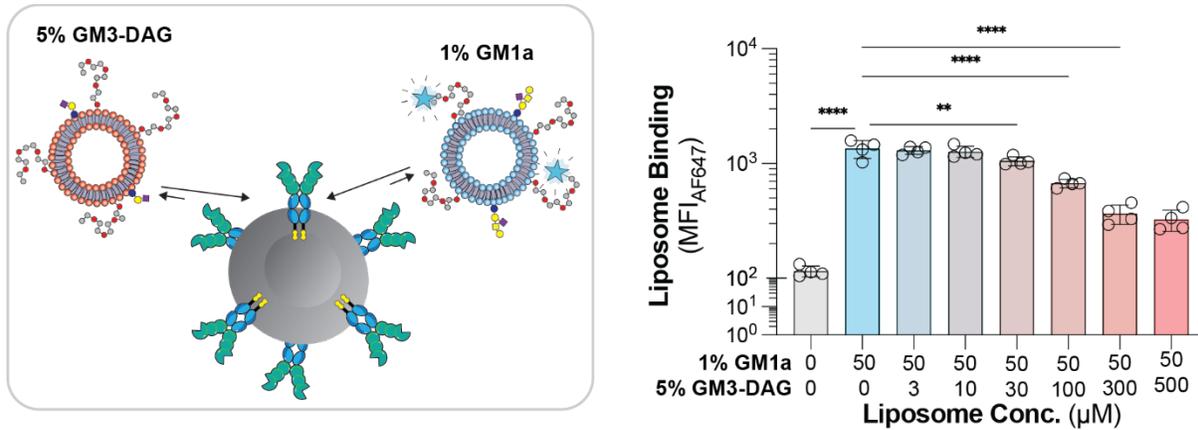


Figure 4.11: Binding competition between GM1a bearing liposomes and liposomes bearing neoglycolipid 5 against WT Siglec-6 in the bead assay. Binding of 1 mol% GM1a liposomes to WT Siglec-6 in the presence of an increasing concentration of 5 mol% neoglycolipid 5 bearing liposomes. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four three technical replicates. A one-way ANOVA was used for statistical analysis. ** = 0.01 > P \geq 0.001; ****P < 0.0001.

The strong ability of **5** to engage Siglec-6 was unexpected. Moreover, the improved binding of Siglec-6 to liposomes bearing neoglycolipid **5** compared to natural GM3 was also surprising as both ligands have the same oligosaccharide. After observing such stark differences in binding between the two glycolipids, which share an oligosaccharide, we posited that this difference was due to liposomal presentation. To address this, we used the ELISA to measure the binding between GM3 and neoglycolipid **5** to Siglec-6. Indeed, the improved binding to nGL **5** over GM3 was unique to a liposomal display, as **5** and GM3 showed only a 1.5-fold difference in binding to Siglec-6 by ELISA compared to the 34-fold difference in the cell assay (**Figure 4.12**).

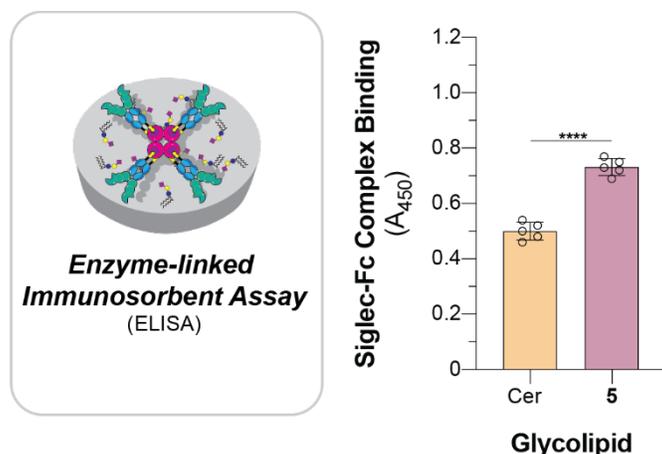


Figure 4.12: Comparison of the binding of Siglec-6 to GM3 and nGL 5 outside of a lipid bilayer. Data is represented as the mean \pm one standard deviation of five technical replicants of the background corrected $A_{450\text{nm}}$. The background was measured using an ethanol vehicle control. A two tailed Student's t-test was used for statistical analysis. **** = $P < 0.0001$.

Siglec-6 binding to glycolipids was further complicated as Siglec-6 could bind glycolipids independent of its canonical essential arginine residue. We considered that if the Arg122 was expendable, then perhaps Siglec-6 could bind asialo glycolipids. This was addressed by formulating glycolipid liposomes with a variety of asialo glycolipids and it was found that none of the asialo glycolipids were able to mediate binding to Siglec-6 expressing CHO cells (**Figure 4.13a, b**). Taken together, the results from this section suggest that Siglec-6 binds $\alpha(2\rightarrow3)$ linked sialylated glycolipids when presented from a lipid bilayer independent of its canonical arginine residue.

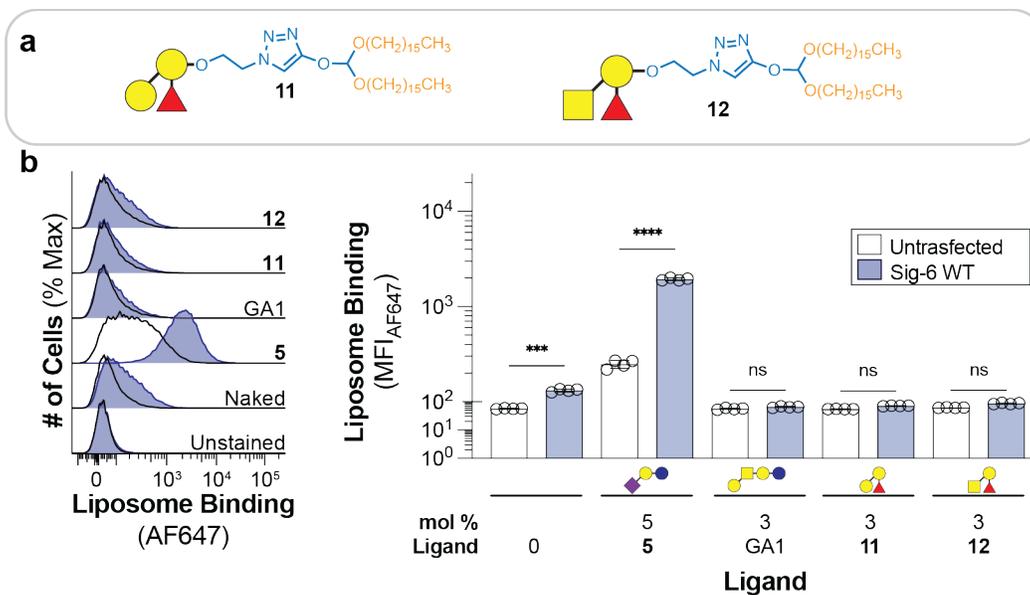


Figure 4.13: Binding of Siglec-6 to liposomes bearing asialo-glycolipids. **a**, Chemical structures of nGLs 11 and 12. **b**, Binding of liposomes bearing glycolipids 5, GA1, 11, and 12 to untransfected and CHO cells expressing WT Siglec-6. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; **** $P < 0.0001$

4.3.3: Identifying which domain(s) of Siglec-6 are important for glycolipid recognition

4.3.3.1: Identifying which domain(s) of Siglec-6 are important for glycolipid recognition

To narrow down which domain(s) of Siglec-6 were important for glycolipid recognition, we created a series of chimeric Siglecs that consisted of different combinations of the extracellular domains of Siglec-6 and Siglec-8. Siglec-8 was chosen because it has 46% sequence identity with Siglec-6 but does not bind glycolipid liposomes. Surprisingly, none of the three individual Siglec-6 domains supported binding to liposomes formulated with nGL **5** in the cell assay; however, when the first two domains of Siglec-6 were used, binding of **5** neoglycolipid liposomes was greater than that of untransfected cells (**Figure 4.14a and b**). A similar binding pattern was observed between the Siglec-6/8 chimeras and GD1a bearing liposomes suggesting that liposomes formulated with nGL **5** and genuine gangliosides bind to Siglec-6 in the same manner.

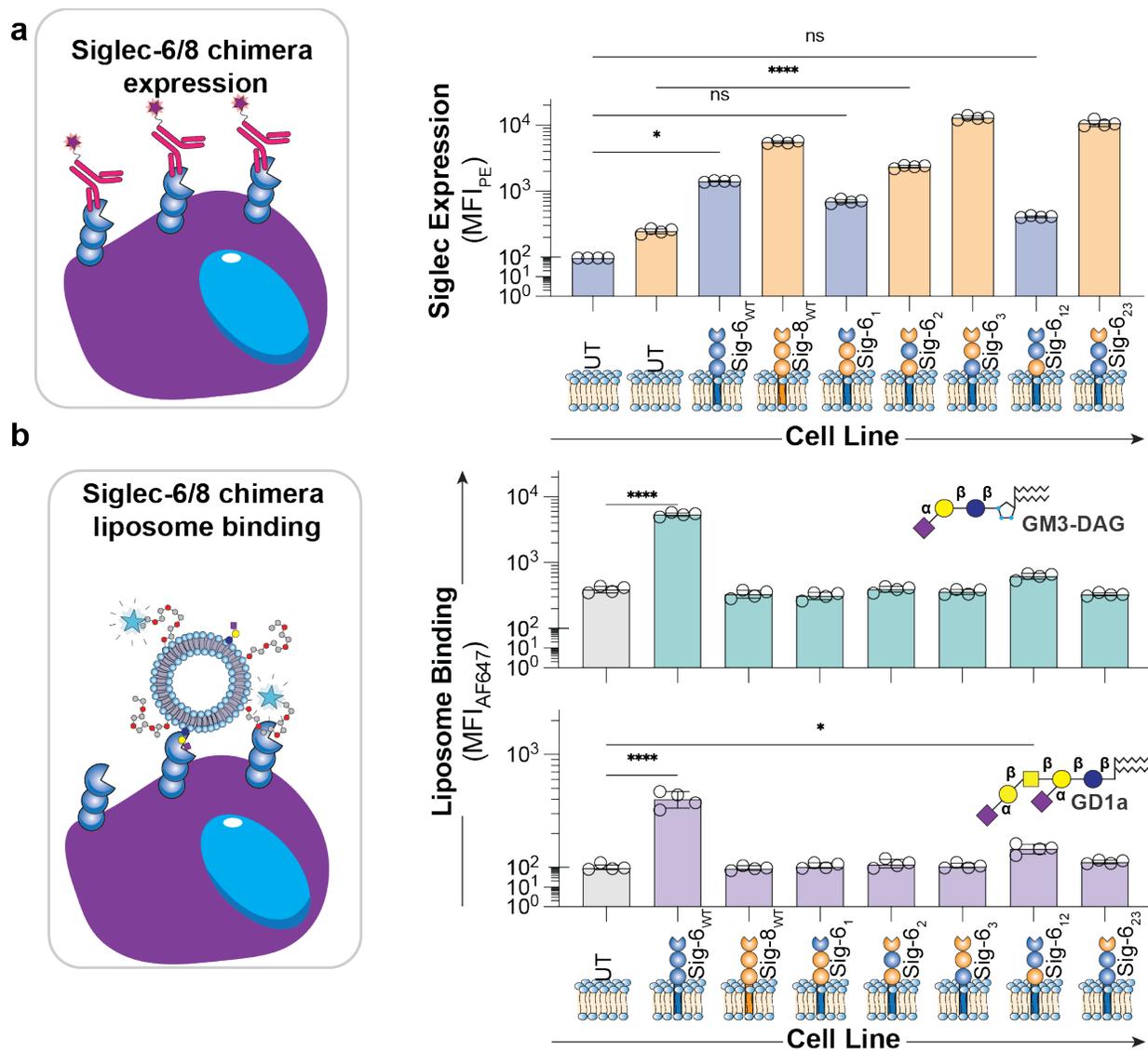


Figure 4.14: Expression and glycolipid binding to Siglec-6/8 full length chimeras in the cell assay. **a**, Expression of Siglec-6/8 chimeras compared to wildtype untransfected CHO cells (UT-grey), WT Siglec-6 (blue), and WT Siglec-8 (orange). **b**, Binding of liposome formulated with neoglycolipid 5 to CHO cells expressing each Siglec-6/8 chimera. Data is represented as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; $0.05 > P \geq 0.01$; **** $P < 0.0001$

Reduced binding of the chimeric construct consisting of the first two domains of Siglec-6, relative to WT Siglec-6, was likely due to a 76% reduction in expression of this construct. Indeed, when expressed as a soluble Fc conjugate and used in the LOLA, this construct displayed comparable levels of binding to WT Siglec-6 (Figure 4.15). These results suggest that both the V-set and the first C2 domain are needed for binding.

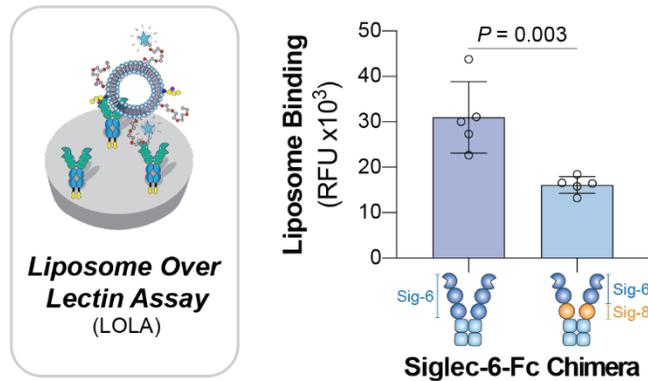


Figure 4.15: Binding of Siglec-6/8 chimeras to liposomes bearing neoglycolipid 5 in the LOLA. Data is represented as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from five technical replicates. A two tailed Student's t-test was used for statistical analysis.

4.3.3.2: Mutagenesis analysis of Siglec-6

The experiments with the Siglec-6/8 chimeras suggested that the interface between the V-set domain and the first C2 domain were important for glycolipid binding. Due to the results from the chimeric Siglec-6/8 constructs, this region was targeted for mutagenesis to determine which residues were important for glycolipid recognition. When performing mutagenesis studies, it is important to consider how the mutation could impact the structure of the protein. If a mutation disrupts the binding between a Siglec and a glycolipid, this could be for two reasons. First, the residue directly interacts with the ligand, which is likely the case when the canonical arginine residue is mutated in Siglecs such as Siglec-1, -2 *etc.* Second, the residue may be important for the structure of the Siglec. The effect of a mutation on protein folding can be approximated by comparing the expression of the mutant protein to that of the wildtype protein; if the expression is much lower than the wildtype, that residue is likely important for the proteins structure. With this in mind, mutagenesis studies were performed on Siglec-6 using the cell assay and liposomes bearing nGL 5. As a control, binding of neoglycolipid liposomes to wildtype Siglec-6 was measured in parallel with each mutant and cells were gated on the same level of Siglec-6 expression to ensure that differences in binding were not due to differences in Siglec-6 expression.

In the AlphaFold model of Siglec-6, two cysteine residues, Cys46 and Cys172, form an interdomain disulfide bridge, which creates the interdomain interface (**Figure 4.1**). Despite even

higher levels of expression compared to WT Siglec-6, the C46A and C172A mutants of Siglec-6 did not recognize **5** nGLs. The guanidinium functional group within an arginine residue commonly supports sialic acid binding, even for lectins outside of the Siglec family³¹². Accordingly, seven additional arginine residues in the V-set domain and at the interface of the V-set/C2 domain were mutated. Several mutants (R109A and R114A) displayed a modest reduction in binding to **5** neoglycolipid liposomes. A more pronounced reduction in binding was observed for the R92A mutant, however it was expressed at near background levels. Contributions from amino acids surrounding Arg92 were also investigated and F93A, L95A, and G175M mutants showed significantly decreased binding to **5** neoglycolipid liposomes. Notably, Gly175 is present within a loop in the underlying C2 domain that is predicted to protrude into the interdomain cleft. These mutants further support the interface between the V-set and the first C2-domain being important for glycolipid recognition.

4.3.3.3: Bilayer binding residues in Siglec-6

After thorough mutagenesis analysis of Siglec-6, no single mutation was able to abolish **5** nGL liposome recognition to background levels without disrupting Siglec-6 expression. However, one observation made during these studies was that cells expressing Siglec-6 routinely displayed significant binding to naked liposomes, which is an observation that has not been seen for other Siglecs. As liposome binding persisted despite these mutations, we consulted a sequence alignment of the Siglecs to see if there was anything unique about Siglec-6. Interestingly, there is a tryptophan residue (Trp127), which is expected to be solvent exposed in a loop in the V-set domain of Siglec-6 (**Figure 4.16a and b**).

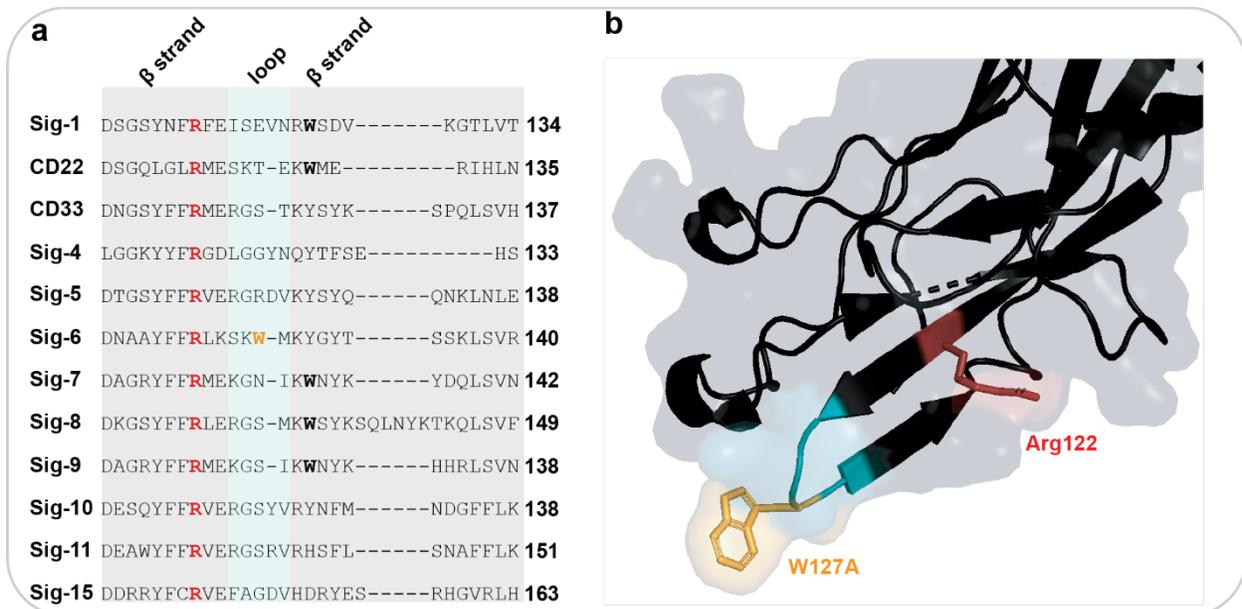


Figure 4.16: Amino acid sequence alignment of residues surrounding the critical arginine residue in the human Siglec family. **a**, sequence alignment of the human Siglecs surround the conserved arginine residue essential for binding sialosides. **b**, AlphaFold model of the V-set domain of Siglec-6 with solvent exposed tryptophan and canonical arginine residue highlighted.

Because tryptophan residues are known to facilitate lipid binding³¹³, we hypothesized that this tryptophan residue may insert itself into the liposome and facilitate binding. Accordingly, an alanine was substituted for Trp127 in Siglec-6. This mutation had no effect on Siglec-6 expression, suggesting that this residue is not structurally important (**Figure 4.17a**).

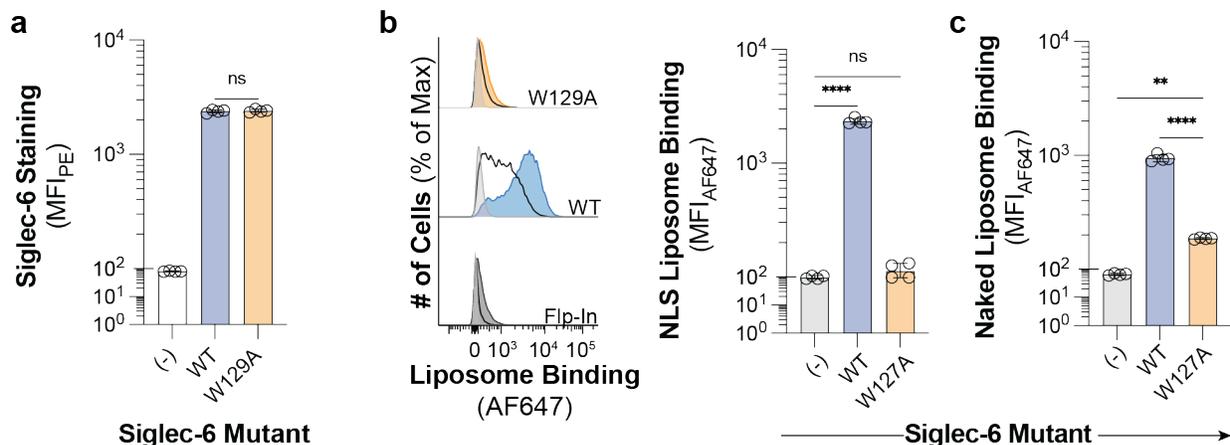


Figure 4.17: Binding of liposomes formulated with nGL 5 to W127A Siglec-6 in the cell assay. **a**, expression of W127A Siglec-6 compared to wildtype Siglec-6 by CHO cells. **b**, binding of naked liposomes and 5 mol% 5 nGL bearing liposomes to W127A Siglec-6. Siglec expression and liposome binding was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; ** = $0.01 > P \geq 0.001$; **** $P < 0.0001$.

No binding of 5 nGL bearing liposomes was observed to the tryptophan mutant compared to untransfected cells (**Figure 4.17b**). Moreover, there was a significant reduction in binding between naked liposomes and the W127A mutant compared to WT Siglec-6(**Figure 4.17c**). The binding between all mutants and Siglec-6 are summarized in **Figure 4.18**.

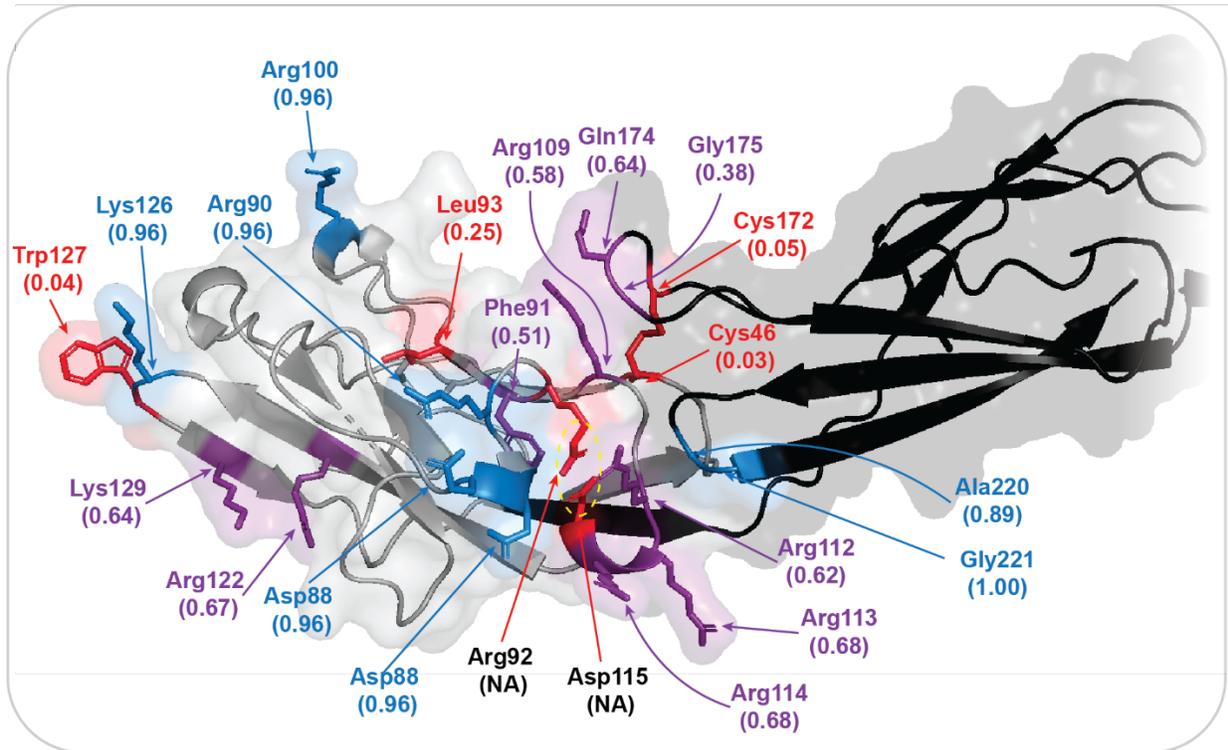


Figure 4.18: Summary of liposome binding to mutants of Siglec-6. Mutated residues are highlighted and shown as 'sticks'. Binding is reported as relative to wildtype Siglec-6, blue, 1.00-0.75; purple, 0.75-0.25; red 0.25-0, NA, no Siglec-6 expression detected.

4.3.4: Targeting physiological Siglec-6⁺ cells and tissues *ex vivo*

The strong binding of **5** neoglycolipid liposomes to Siglec-6 prompted an investigation into their use for targeting Siglec-6 on physiologically relevant cells. Siglec-6 has an unusual expression pattern and is found on mast cells, memory B-cells, placental syncytiotrophoblasts, and has no mouse ortholog¹²⁵. As primary cells lack a genetic control for Siglec-6 and we found that there can be Siglec-6 independent binding to certain cells, a tool to measure if liposome binding was specific to Siglec-6 was needed. One strategy to prevent binding between a Siglec and a ligand is to use a blocking antibody. In this approach, the cells are first treated with an antibody for the Siglec. After the Siglec is treated with the antibody, the liposome is added. If the Siglec is blocked by the antibody, a decrease in liposome binding is expected. Indeed, pre-treatment of Siglec-6 expressing CHO cells with an anti-Siglec-6 antibody could be used to block **5** neoglycolipid liposome binding (**Figure 4.19**).

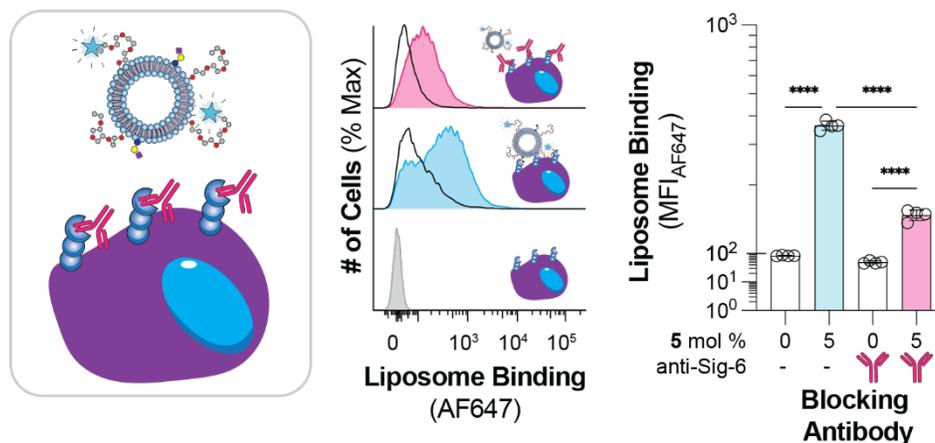


Figure 4.19: Blocking of neoglycolipid bearing liposome binding to Siglec-6 expressing CHO cells using anti-Siglec-6 antibody. CHO cells expressing Siglec-6 were pre-incubated with anti-Siglec-6 antibody prior to incubation with liposomes bearing neoglycolipid **5**. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. A one-way ANOVA was used to compare **5** liposome binding before and after treatment with the anti-Siglec-6 antibody to Siglec-6 expressing CHO cells. **** = $P < 0.0001$.

4.3.4.0: Targeting Siglec-6 on memory B cells

After confirming that the Siglec-6 antibody could block liposome binding, we began assessing the ability of **5** nGL bearing liposomes to engage Siglec-6 in physiologically relevant contexts. First, we tested the binding of liposomes formulated with 5 mol% **5** against memory B

cells from peripheral human blood. Memory B cells were defined as CD19⁺, CD22⁺, CD27⁺, IgD⁻, and CD38⁻. We used naïve B cells (CD19⁺, CD22⁺, CD27⁻, IgD⁺) as a negative control as Siglec-6 expression is higher on memory B cells compared to naïve B cells (**Figure 4.20**).

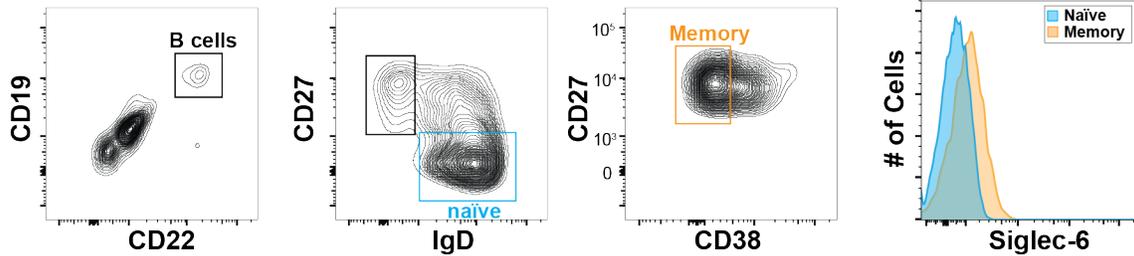


Figure 4.20: Gating scheme for defining Memory B-cells (CD19⁺, CD22⁺, CD27⁺, IgD⁻, CD38⁻) and naïve B-cells (CD19⁺, CD22⁺, CD27⁻, IgD⁺) from human blood.

B cells were isolated from whole blood from four healthy donors. We observed significant binding of the liposomes formulated with nGL **5** to memory B cells compared to the naïve B cells (**Figure 4.21a**). The binding of the neoglycolipid liposome to memory B cells was reduced to the same level as naïve B cells when cells were pretreated with the anti-Siglec-6 antibody (**Figure 4.21b**). These results suggest that liposomes formulated with nGL **5** can target genuine memory B cells through Siglec-6.

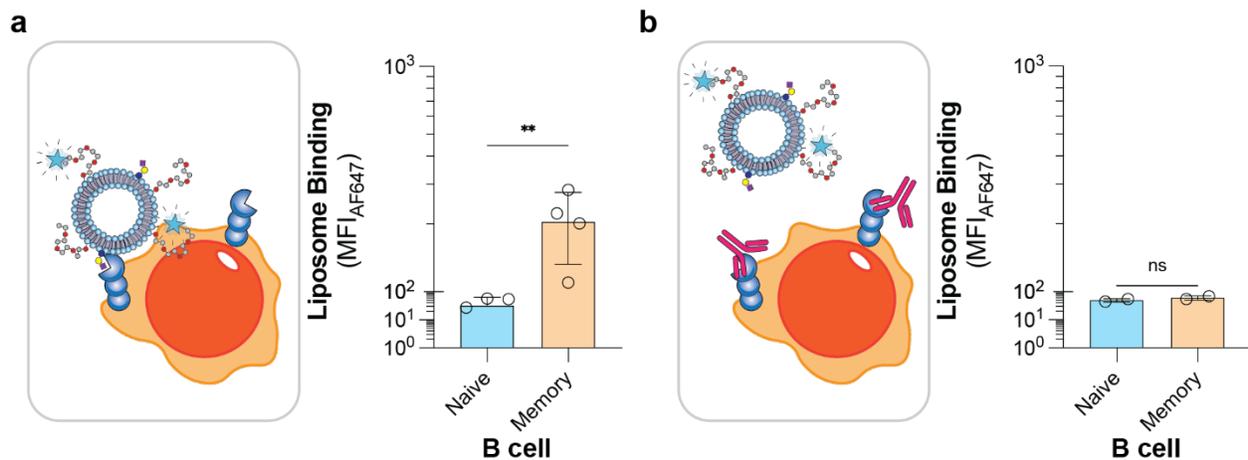


Figure 4.21: Neoglycolipid liposome binding to human naïve and memory B cells. a, b, B cells were isolated from the blood of healthy donors with and without and with pre-treatment of Siglec-6 blocking antibody, respectively. Each datum is representative of a biological replicate. Not Significant (NS), $P > 0.05$; ** $0.01 > P \geq 0.001$.

4.3.4.1: Targeting Siglec-6 on mast cells

After the success of targeting memory B cells with nGL **5** bearing liposomes, we investigated if Siglec-6 on mast cells could also be targeted with these liposomes. Mast cells were defined as FcεR1⁺ and c-KIT/CD117⁺. Mast cells were successfully identified via flow cytometry from human spleen samples and showed robust Siglec-6 expression (**Figure 4.22a, b**), however mast cells were in very low abundance (approximately 0.01% of total white blood cells) and it was felt that it was impractical to continue working with genuine human mast cells isolated from spleens.

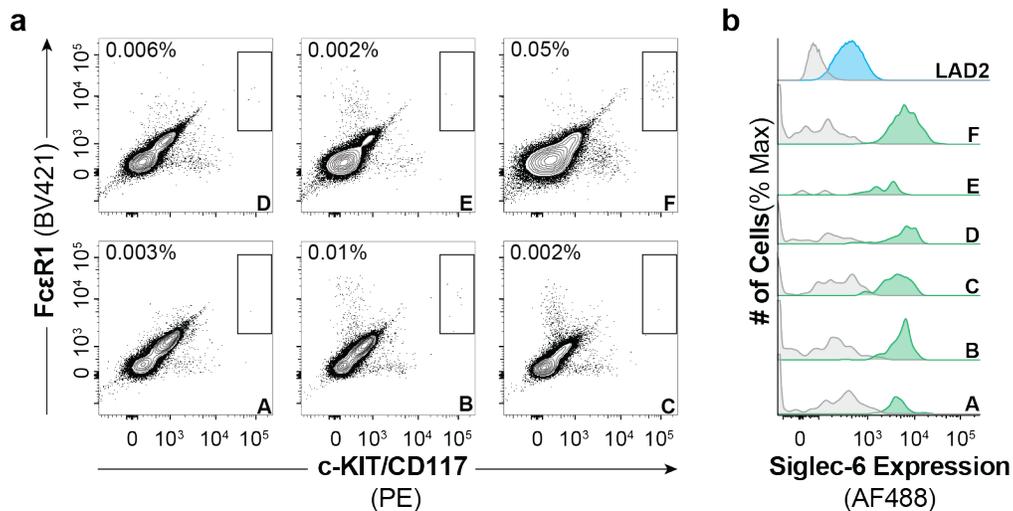


Figure 4.22: Expression of Siglec-6 by primary mast cells isolated from human spleens. **a**, Mast cells from six different healthy donors (represented by capital letters) identified as FcεR1⁺ and c-KIT/CD117⁺. Percentages of mast cells in total white blood cells is in upper left corner of each plot. **b**, Representative flow cytometry of Siglec-6 expression on LAD2 cells and primary mast cells stained with anti-Siglec-6-AF488.

As an alternative, LAD2 (immortalized human mast-cell derived) cells were used. LAD2 were an appropriate cell model to use as they are a mast cell line³¹⁴ and express Siglec-6 at comparable levels to genuine human mast cells. Just like with memory B cells, we found that **5** neoglycolipid liposomes bound robustly to LAD2 cells and that this binding was reduce to levels equivalent to naked liposomes when the LAD2 cells were pre-incubated with a Siglec-6 antibody (**Figure 4.23**).

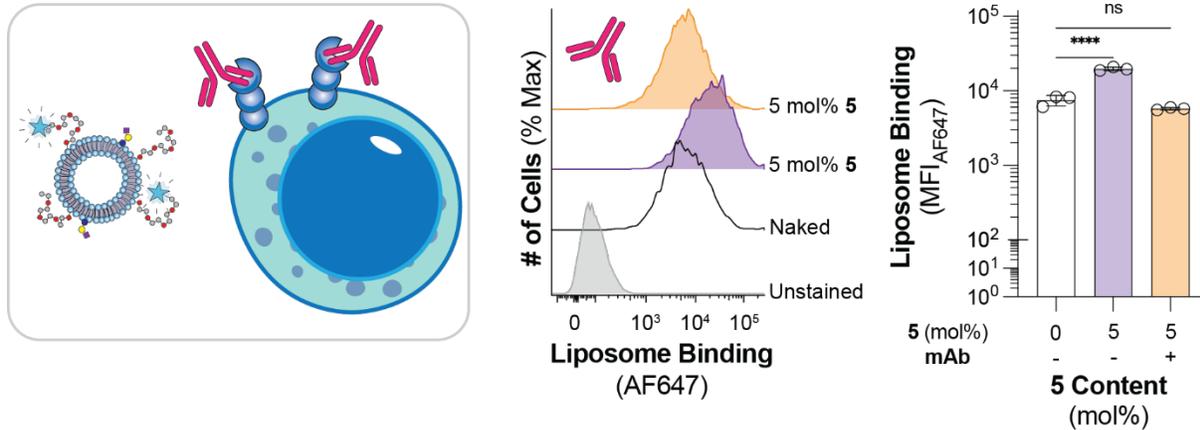


Figure 4.23: Binding of liposomes bearing neoglycolipid 5 to LAD2 cells. LAD2 cells were pre-incubated with anti-Siglec-6 antibody prior to incubation with liposomes bearing neoglycolipid 5. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from three technical replicates. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; **** = $P < 0.0001$.

4.3.4.2: Targeting Siglec-6 on placental syncytiotrophoblast

Lastly, we also investigated whether **5** neoglycolipid liposomes could engage Siglec-6 expressed on syncytiotrophoblasts. Working with live human placental explant tissue cultures, syncytiotrophoblasts could be identified by the distinct pattern of phalloidin staining, we showed that **5** neoglycolipid liposomes are strongly associated with the syncytiotrophoblasts compared to liposomes without the nGL (**Figure 4.24a**). In four independent biological replicates, **5** neoglycolipid liposomes showed significantly more puncta on these cells, suggesting that neoglycolipid liposomes bind to Siglec-6 on the syncytiotrophoblasts in a glycolipid dependant manner (**Figure 4.24b**). Consistent with this, **5** neoglycolipid liposomes showed a strong colocalization with Siglec-6 compared to naked liposomes (**Figure 4.24c**). Moreover, pre-treatment with an anti-Siglec-6 antibody blocked these interactions (**Figure 4.24d**). These results demonstrate that Siglec-6 can be targeted on primary human cells and tissues using synthetic epitopes.

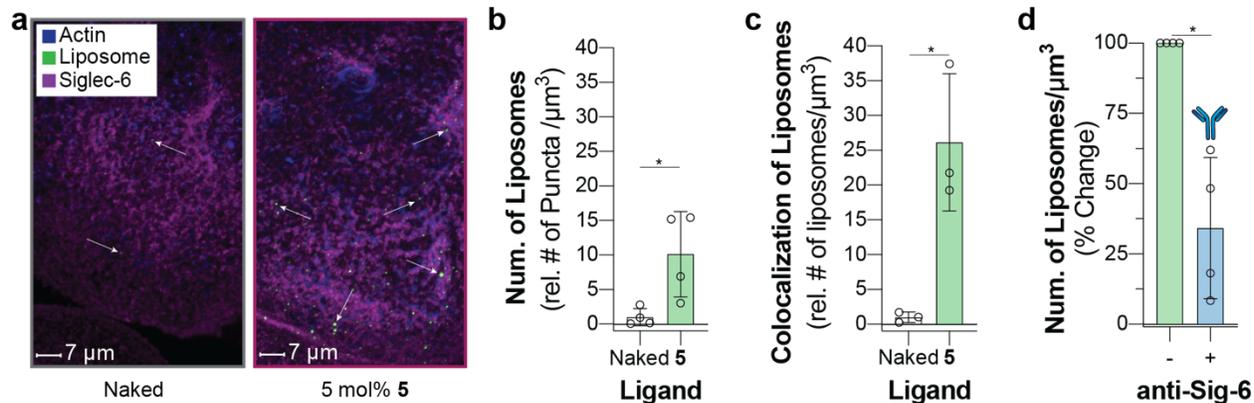


Figure 4.24: Binding of liposome formulated 5 mol% 5 to human syncytiotrophoblasts. **a**, representative confocal microscopy images of human placental explants stained with anti-Siglec-6 and liposomes formulated with 5 mol% 5. **b**, **c**, Quantification of the total number of naked liposomes and 5 neoglycolipid liposomes binding to human placental explants per μm^3 and the number of 5 neoglycolipid liposomes colocalized with Siglec-6 per μm^3 respectively. **d**, Binding of liposomes bearing neoglycolipid 5 to human syncytiotrophoblasts after treatment with an anti-Siglec-antibody. Each datum represents a biological replicate. A two-tailed Student's t-test was used for statistical analysis. $*0.05 > P \geq 0.01$

4.3.5: Siglec-6 internalizes neoglycolipid liposomes

As Siglecs are generally endocytic receptors¹⁰¹, we investigated if Siglec-6 can internalize cargo such as liposomes. Using Daudi cells, a human B-cell line, transduced to over express Siglec-6, we assessed internalization of 5 neoglycolipid liposomes using two different approaches. The first was using liposomes formulated with a pHrodo-labeled lipid. pHrodo is a derivative of rhodamine which features a *push-pull electron system* which is when an electron donating group is connected to an electron withdrawing group through a conjugated/aromatic (**Figure 4.25a**)³¹⁵. Under acidic conditions, the push-pull system is disrupted due to protonation of the electron rich nitrogen resulting in increased red fluorescence (λ_{ex} 560 nm, λ_{em} 587 nm-**Figure 4.25b**). During endocytosis, the endosome acidifies placing the particle in an acidic environment (pH 4.5-6.5)³¹⁶. Therefore, a pH sensitive fluorescent reported such as pHrodo can be used to report on cellular internalization (**Figure 4.25c**)¹²⁷.

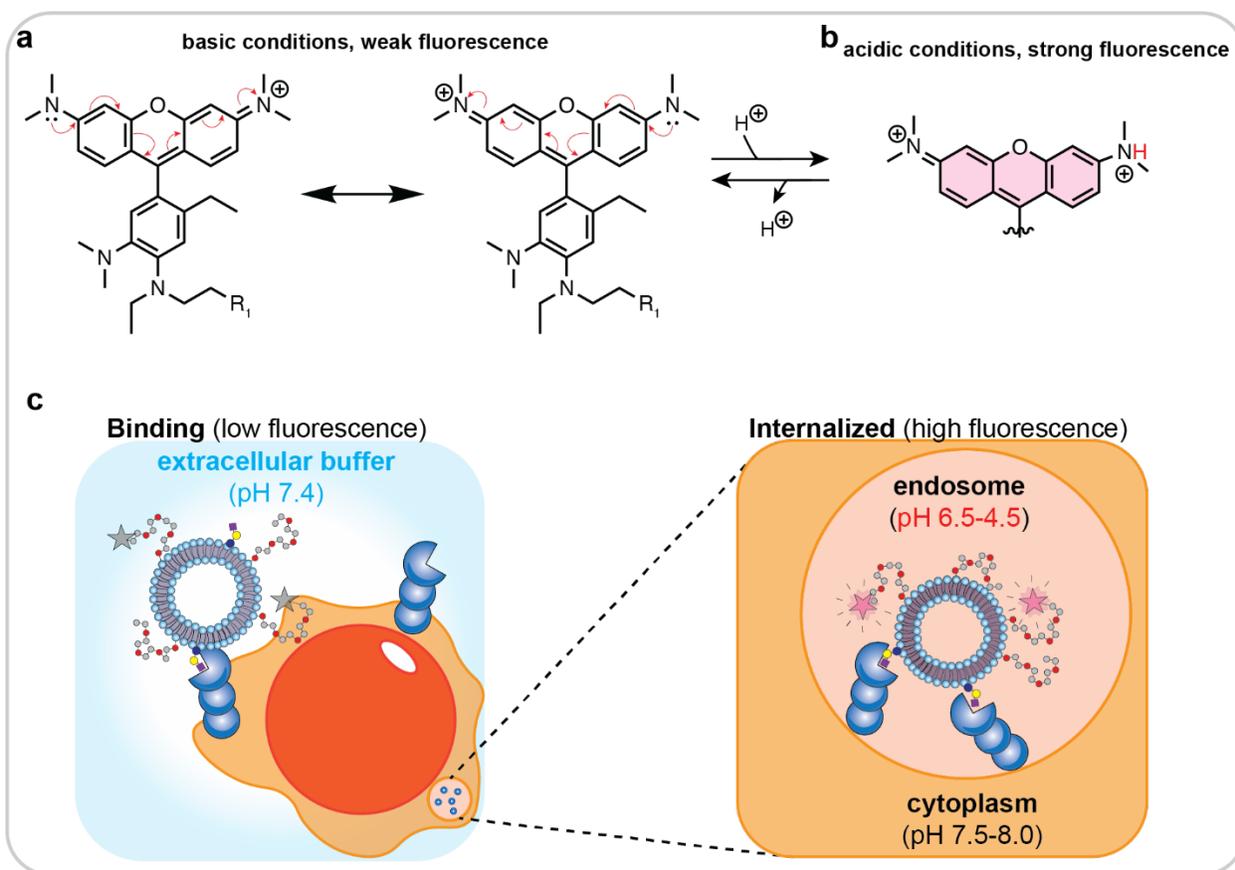


Figure 4.25: pHrodo is a pH sensitive fluorescent indicator that can be used to measure liposome internalization. **a**, chemical structure of pHrodo and depiction of the push-pull electron system. R₁ represents a common location where chemical handles such as amines, NHS esters, etc. are covalently linked. **b**, Protonation of the electron rich nitrogen under acidic conditions causing an increase in fluorescence (λ_{ex} 560 nm, λ_{em} 587 nm). **c**, Depiction of liposome binding and internalization where pHrodo becomes increasingly fluorescent due to the drop in pH from the buffer to the endosome.

Time-dependent increase in pHrodo signal for **5** neoglycolipid liposomes was observed independent of Arg122 and only at 37 °C, which is indicative of cellular internalization (**Figure 4.26a**). Imaging flow cytometry more directly revealed that Siglec-6 Daudi cells internalize **5** neoglycolipid liposomes at 37 °C, but not 4 °C, further supporting that Daudi cells internalize liposomes in a Siglec-6 dependant manner (**Figure 4.26b**).

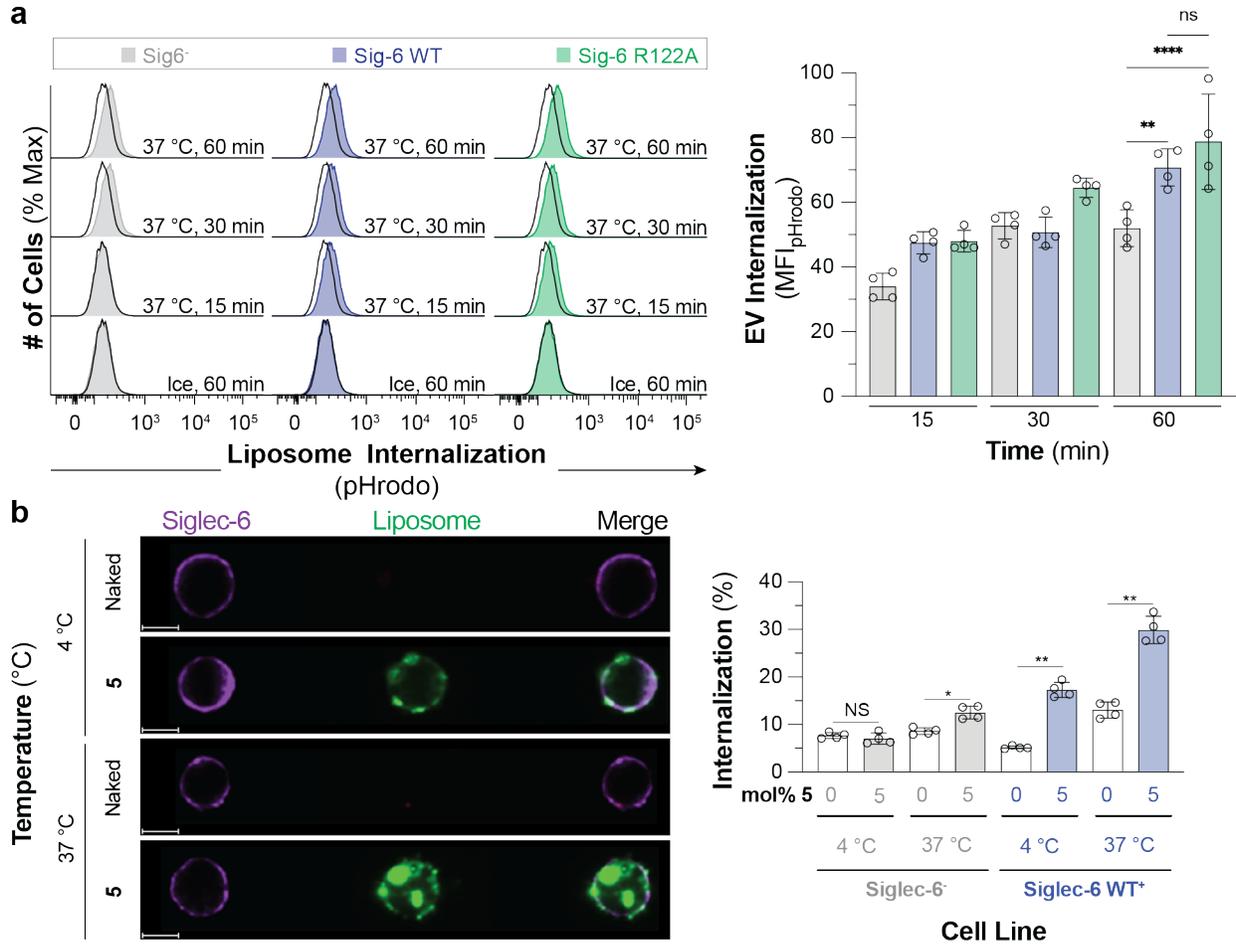


Figure 4.26: Internalization of neoglycolipid liposomes by Siglec-6 expressing Daudi cells. a, Fluorescence of pHrodo labeled liposomes bearing neoglycolipid 5 incubated with Daudi cells virally transduced with empty vector and WT Siglec-6 over 60 min at 4 °C and 37 °C. **b,** Imaging flow cytometry fluorescence of Daudi cells transduced with Siglec-6 (purple) incubated with AF647-labeled liposomes (green) for 60 min at 4 °C or 37 °C. anti-Siglec-6-AF488 antibody staining shows the cell surface expression of Siglec-6. Scale bars represent 7 μ m. For panels a and b, data was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates and as the mean \pm one standard deviation of the percent internalization respectively. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; $0.05 > P \geq 0.01$; $0.01 > P \geq 0.001$; $****P < 0.0001$.

4.3.6: Siglec-6 binding to extracellular vesicles

The observation of Siglec-6 binding gangliosides suggested that Siglec-6-ganglioside interactions may be biologically significant. Gangliosides are commonly found in the outer leaflet of plasma membranes of cells; however, we were unable to detect Siglec-6-glycan interactions in the cell-based assay so probing gangliosides on the surfaces of cells (**Figure 4.3**). We considered other locations where gangliosides reside. Extracellular vesicles (EVs) are cell membrane-derived nanoparticles ranging from 50 nm to 200 nm in diameter^{262, 317}. When EVs are released from the cell, they are comprised of components of the membrane of that cell, including gangliosides. Indeed, several Siglecs have been shown to interact with ligands on EVs^{256, 262}. With this in mind, we investigated if Siglec-6 could bind EVs.

4.3.6.0: Siglec-6 binds extracellular vesicles

Human blood was used as a source of EVs and EVs were isolated from human blood through a series of centrifugal separations (**Figure 4.27**). Once the EVs were isolated, they were fluorescently labeled using NHS chemistry and characterized by transmission electron microscopy (**Figure 4.28a**). Fluorescently labeled EVs showed robust binding to Siglec-6 CHO cells in the cell assay (**Figure 4.28b, c**).

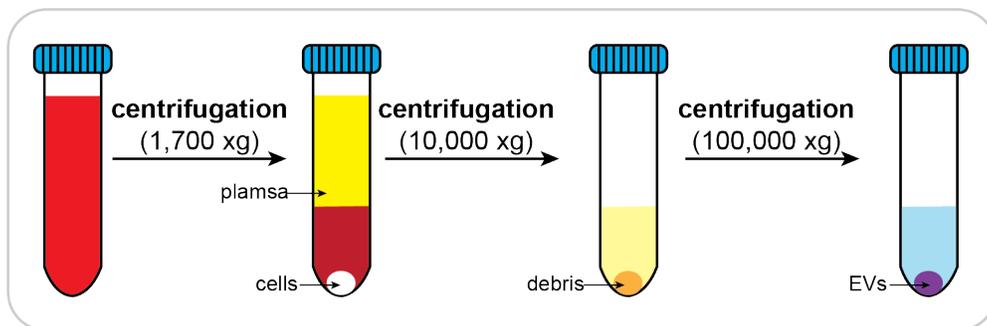


Figure 4.27: Workflow for isolating extracellular vesicles from human blood.

However, there are many receptors and lectins on CHO cells and many glycans and proteins on EVs. To determine if Siglec-6 mediated the binding of EVs to CHO cells the Siglec-6 blocking antibody was used. Indeed, EV binding to Siglec-6 expressing CHO cells was reduced upon pre-treatment with the anti-Siglec-6 antibody; suggesting that EV binding to CHO cells is at least in part due to Siglec-6.

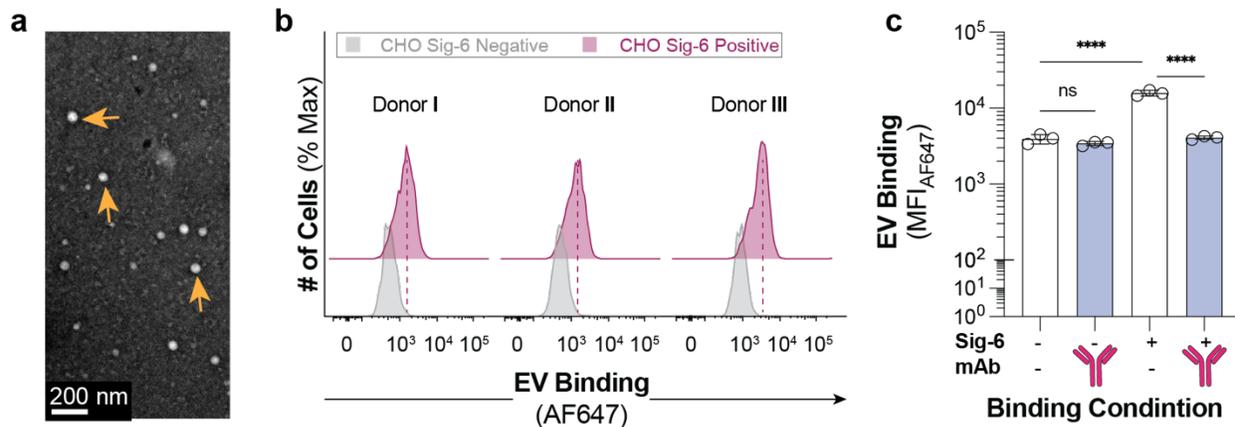


Figure 4.28: Characterization of extracellular vesicles isolated from human blood. **a**, Reversative transmission electron microscopy image of EVs isolated from peripheral human blood. **b**, Binding of EVs from three different donors labeled with AF647 to CHO cells expressing WT Siglec-6. Data is presented as flow cytometry histograms. **c**, Binding of EVs to CHO cell expressing WT Siglec-6 pretreated with the anti-Siglec-6 antibody. Data was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from three technical replicates. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$; **** $P < 0.0001$.

4.3.6.1: Siglec-6 binds extracellular independent of canonical arginine residue

After establishing that Siglec-6 can bind to EVs, comparisons between the binding of ganglioside liposomes and EVs were carried out. Just like with glycolipid-bearing liposomes, Siglec-6 engage EVs independent of Arg122, while the disulfide bridge between the V-set domain and the first C2 domain was important for EV binding as demonstrated by the reduction in EV binding in both the C46A and C172A Siglec-6 mutants (**Figure 4.29**).

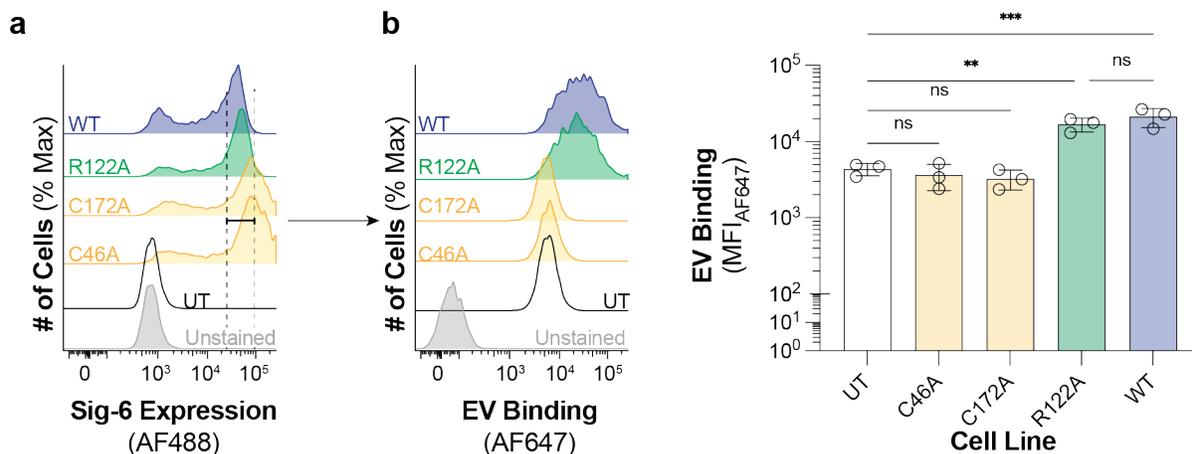


Figure 4.29: Binding of EVs to wildtype and mutant Siglec-6 expressing CHO cells. **a**, Expression of Siglec-6 mutants by CHO cells. **b**, Binding of EVs isolated from human blood to CHO cells expressing Siglec-6 mutants. For panel **b**, Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from three technical replicates. A one-way ANOVA was used to compare 5 liposome binding before and after treatment with the anti-Siglec-6 antibody to Siglec-6 expressing CHO cells. Not Significant (NS), $P > 0.05$; ** $0.01 > P \geq 0.001$; *** $0.001 > P \geq 0.0001$.

To further the understanding of how Siglec-6 binds EVs, we measured the binding of EVs in the presence of increasing amounts of 5% **5** neoglycolipid liposomes. This was done to assess if the EVs and liposomes compete for the same binding site on Siglec-6. EV binding decreased as liposome concentration increased, suggesting that the neoglycolipid liposomes and EVs compete for the same binding pocket (**Figure 4.30**). These results suggest that **5** and ligands on EVs compete for the same, or overlapping, binding site on Siglec-6.

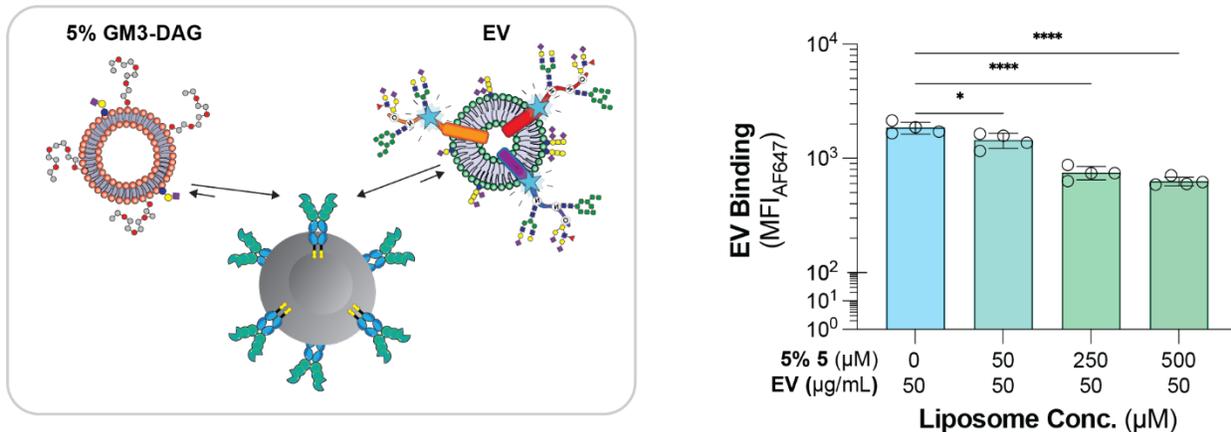


Figure 4.30: Binding competition between EVs and liposomes bearing neoglycolipid 5 against WT Siglec-6 in the bead assay. Binding of EVs to WT Siglec-6 in the presence of an increasing amount of liposomes bearing 5 mol% neoglycolipid 5. Data was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. A one-way ANOVA was used for statistical analysis $*0.05 > P \geq 0.01$; $****P < 0.0001$.

4.3.6.2: Siglec-6 binds extracellular through $\alpha(2 \rightarrow 3)$ gangliosides

After investigating how Siglec-6 engages EVs from the perspective of Siglec-6, we investigated how the properties of the EV effect the ability to engage Siglec-6. First, EVs were treated with a broadly acting neuraminidase A (Neu A) or an $\alpha(2 \rightarrow 3)$ -specific neuraminidase S (Neu S) to determine if Siglec-6 binds $\alpha(2 \rightarrow 3)$ linked sialosides as it did in the neoglycolipid profiling (**Figure 4.31**). Indeed, significantly decreased EV binding was observed after treatment with either neuraminidase. These results suggest that EV binding is sialic acid dependent and that $\alpha(2 \rightarrow 3)$ -linked sialosides on EVs mediate binding to Siglec-6.

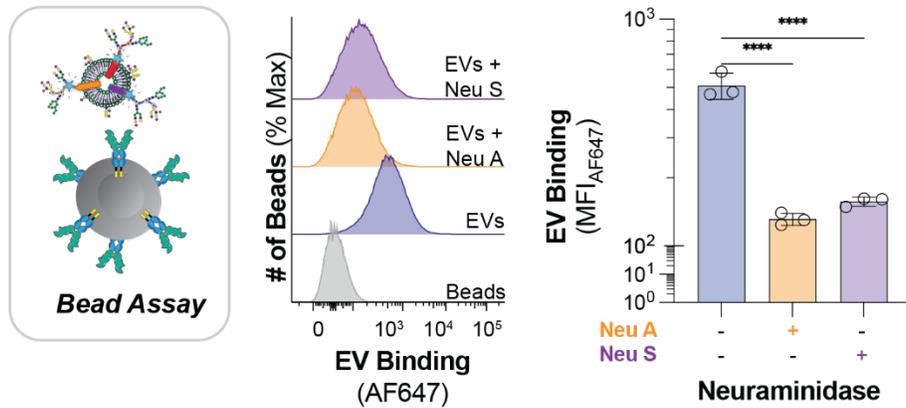


Figure 4.31: Binding of EVs treated with neuraminidase A and S to Siglec-6 in the bead assay. EVs were treated with neuraminidase S ($\alpha(2 \rightarrow 3)$ specific), neuraminidase A (cleaves all sialic acid linkages), or BSA prior to incubation with Streptavidin microbeads containing immobilized WT Siglec-6-Fc. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from three technical replicates. A one-way ANOVA was for statistical analysis. **** $P < 0.0001$.

To examine whether gangliosides in EVs mediate binding to Siglec-6, we prepared EVs from WT and $\beta 1-4\text{GalNT1}^{-/-}$ N2a cells. If $\beta 1-4\text{galnt1}$ is knocked-out, then gangliosides more complex than GM3, such as GM1a and GD1a that are ligands for Siglec-6, cannot be synthesized (**Figure 4.32a**). If these more complex gangliosides contribute to the binding of EVs, it would be expected to observe less binding to EVs from the knock-out cells compared to EVs isolated from the wildtype cells. A significant reduction in binding of $\beta 1-4\text{GalNT1}^{-/-}$ -derived EVs compared to WT EVs was observed (**Figure 4.32b**), suggesting that complex glycolipids in EVs support binding to Siglec-6.

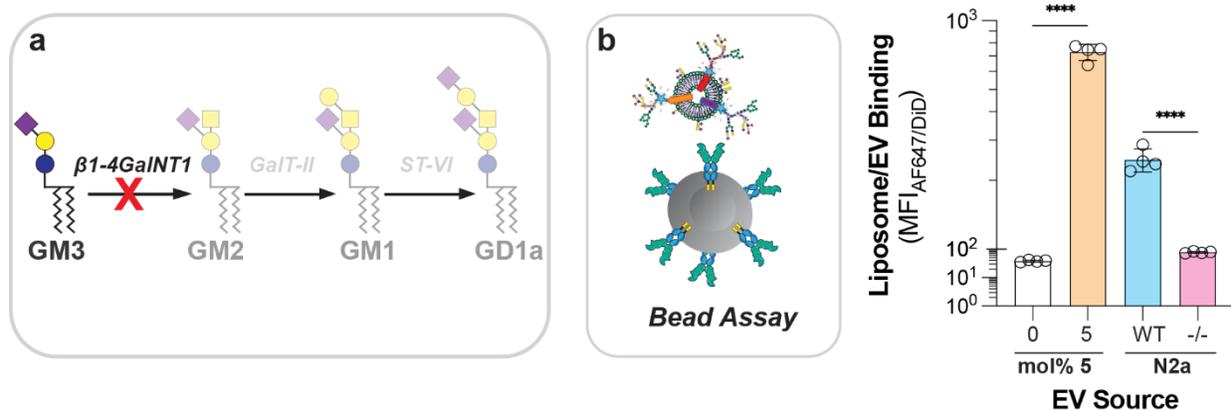


Figure 4.32: Binding of EVs isolated from wildtype and $\beta 1-4\text{GalNT1}^{-/-}$ N2a cells to WT Siglec-6 in the bead assay. **a**, Abbreviated ganglioside biosynthetic pathway highlighting the role of $\beta 1-4\text{GalNT1}$ in the biosynthesis of complex gangliosides. **b**, Binding of EVs isolated from WT and $\beta 1-4\text{GalNT1}^{-/-}$ N2a cells in the bead assay. Data was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from four technical replicates. A one-way ANOVA was for statistical analysis. **** $P < 0.0001$.

Lastly, we were examined whether Siglec-6 dependent binding to EVs was possible under more physiologically relevant conditions. To address this, binding of EVs isolated from two different donors to LAD2 cells was measured in the presence and absence of the Siglec-6 ligand blocking antibody. A modest, but significant, reduction in EV binding with EVs from two different donors was observed to LAD2 cells blocked with an anti-Siglec-6 antibody (**Figure 4.33**). These results demonstrate that physiologically relevant expression levels of Siglec-6 support EV binding.

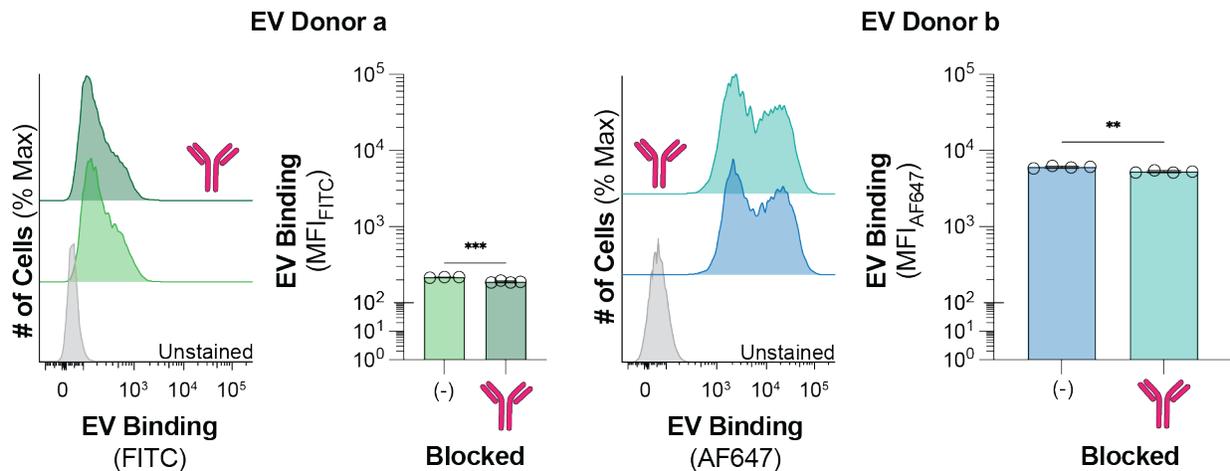


Figure 4.33: Binding of EVs to LAD2 cell pretreated with Siglec-6 blocking antibody. LAD2 cells were pre-incubated with anti-Siglec-6 antibody prior to incubation with EVs followed by analysis by flow cytometry. Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from three technical replicates. A two tailed Student's t-test was used for statistical analysis. ** $0.01 > P \geq 0.001$; *** $0.001 > P \geq 0.0001$.

4.3.5.3: Siglec-6 internalizes Extracellular Vesicles

As we previously demonstrated that Siglec-6 could facilitate the uptake of neoglycolipid liposomes, we posited that Siglec-6 may also be able to internalize EVs. Using pHrodo label EVs, the ability of Daudi cells expressing Siglec-6 to internalize EVs was assessed. It was found that pHrodo signal increased in a time and temperature dependant manner by flow cytometry (**Figure 4.34a**). Similar to liposome internalization, EV internalization was not dependant on Arg122 but was dependent of Siglec-6. Moreover, imaging flow cytometry also showed Siglec-6 dependant uptake of EVs (**Figure 4.34b and c**). Therefore, Siglec-6 engages EVs and mediates their internalization.

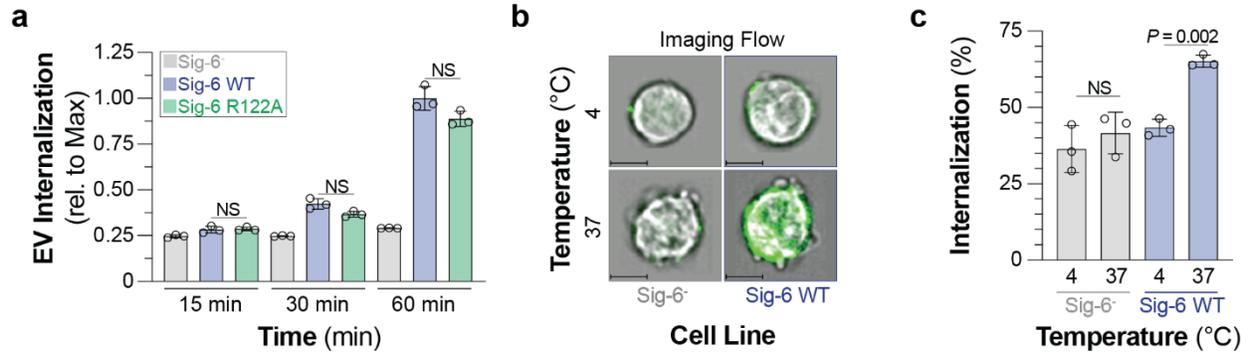


Figure 4.34: Internalization of EVs by Daudi cells virally transduce with Siglec-6. **a**, Daudi cell stably transduced with Siglec-6 or empty lentiviral vector were incubated with pHrodo labelled EVs for different amounts of time and analyzed by flow cytometry. Data was quantified as the mean \pm one standard deviation of the media fluorescent intensity (MFI) from three technical replicates. **b**, Representative imaging flow cytometry images of empty vector and WT Siglec-6 virally transduced Daudi cells incubated with AF488 labeled EVs at 4 or 37 °C with the EV fluorescence overlaid over the brightfield image. Scale bars represent 7 μ m. **c** Quantification of internalization of EVs at 4 or 37 °C by Daudi cells transduced with WT Siglec-6 and an empty vector. Data is represented by the mean \pm one standard deviation of at least three technical replicants. A one-way ANOVA was used for statistical analysis. Not Significant (NS), $P > 0.05$.

4.4: Discussion

Relatively little is known about Siglec-6 with respect to its ligands and its biological functions when compared to other Siglecs such as CD22, CD33, and Siglec-7. A large step towards describing the ligands of Siglec-6 was taken in the previous chapter when it was revealed that Siglec-6 bound gangliosides when presented from an optimized liposome formulation. While the ability of Siglec-6 to bind gangliosides was unexpected, it was even more surprising that Siglec-6 bound these gangliosides independent of its conserved arginine residue Arg122. This was exciting and warranted an investigation into how Siglec-6 recognized its ganglioside ligands.

4.4.0: Bilayer presentation is important for glycolipid recognition by Siglec-6

Binding of Siglec-6 towards gangliosides has not been studied in depth, however one study measured the binding of Siglec-6 towards GD3 using an ELISA but found that Siglec-6 did not bind GD3²⁰⁵. As GD3 was found to be a ligand for Siglec-6 in the cell assay, we posited that this contradiction may be due to the experimental approach. Accordingly, we interrogated Siglec-6 against our panel of gangliosides in an ELISA and found that Siglec-6 did not bind to many of the gangliosides that it did in the cell assay (GM1a, GD1a, GD3, *etc.*) but could bind to others that Siglec-6 did not bind in the cell assay (GM3 and GM4). Interestingly, hSiglec-1 and -7 also showed improved binding to gangliosides with smaller oligosaccharides (GM3 and GM4) in the ELISA compared to the cell assay. This is in line with a modeling study that proposed that GM3 adopts a more buried or laid down conformation in a bilayer which may make the oligosaccharide less assessable to the Siglec²⁸⁶. However, when the binding is measure outside a bilayer, the oligosaccharide is accessible to the Siglec resulting in binding being observed.

The lack of binding of Siglec-6-Fc to primary immune cells in Chapter 2 may support that glycolipid presentation/accessibility is important for binding gangliosides. Moreover, no binding between the Siglec-6-Fc and any of the cell lines tested was observed. Indeed, very few examples of a Siglec binding to a ganglioside in *trans*. Perhaps this is due to the glycocalyx of the cell shielding the gangliosides from the Siglec-6-Fc, similar to how increasing PEG₄₅-DSPE content

disrupted binding between hSiglec-1 and GM1a in the cell assay in Chapter 3. Based on these results, we hypothesize that bilayer presentation/accessibility is important for glycolipid recognition by Siglec-6 and in line with this, Siglec-6-ganglioside binding was observed in other approaches such as the bead assay and the LOLA where the gangliosides were presented from a bilayer, supporting this hypothesis.

4.4.1: Siglec-6 recognizes $\alpha(2\rightarrow3)$ linked glycolipids

Intrigued by how the presentation of the glycolipid affected the ability of Siglec-6 to bind gangliosides we sought to explore this further using a panel of neoglycolipids. It was found that the chemical structure of the neoglycolipid dramatically effected the ability of Siglec-6 to recognize a glycolipid when presented from a bilayer. The best ligand found was the oligosaccharide of GM3 linked through a triazole to a di-O-hexadecyl glycerol (compound **5**). Interestingly, the improved binding of **5** compared to GM3 was unique to bilayer presentation as **5** show comparable binding to GM3 in an ELISA but was 34-fold higher in the cell-based assay. Using this di-O-hexadecyl glycerol scaffold, it was also found that Siglec-6 does not engage with $\alpha(2\rightarrow6)$ linked sialosides which is in contrast with previous reports which found that the *sialyl Tn* antigen (α -Neup5Ac-(2 \rightarrow 6)- α -GalpNAc-Ser/Thr) is a ligand for Siglec-6. However, $\alpha(2\rightarrow6)$ sialyl lactose and $\alpha(2\rightarrow6)$ sialyl LacNAc was used rather than the *sialyl Tn* antigen, so it is possible that Siglec-6 does not recognize $\alpha(2\rightarrow6)$ linked sialic acid from these scaffolds. While **5** was a relatively potent Siglec-6 ligand, it also had off target effects as observed in the experiments with Raji cells. This ligand could be improved in the future by chemically modifying it to be more selective for Siglec-6.

4.4.2: Trp127 is critical for Siglec-6 binding to glycolipid bearing liposomes

Another unique trait of Siglec-6 was that binding of ganglioside bearing liposomes was not dependant on the conserved arginine residue Arg122, albeit some binding was lost with this mutant compared to wildtype Siglec-6 (approximately 30%). To gain a better understanding of how Siglec-6 binds its ligands independent of this arginine residue, a series of chimeras and

mutants were made to determine which residues are important for ganglioside recognition by Siglec-6. The Siglec-6/8 chimeras revealed that both the V-set and the first C2-set domain were critical for glycolipid recognition. Accordingly, the interface between these domains was targeted for mutagenesis studies. First, both cysteine residues required to form the interdomain disulfide bridge (Cys46-Cys172) were mutated. Unsurprisingly both cysteine mutants show no liposome binding compared untransfected cells. After determining that this interdomain area is structurally important for Siglec-6-liposome binding, we probed residues in this area to determine if they impacted Siglec-6 binding to glycolipid liposomes. We found that when residues in a loop in the first C2-set domain (residues 173-175) were mutated, significant reductions in binding were observed. Moreover, when residues buried in the V-set domain were mutated (Leu93, Phe91) reductions in binding were also observed. While mutagenesis of these residues effected glycolipid liposome binding, it also effected Siglec-6 expression, suggesting that these mutations change the structure of Siglec-6 such that it is less able to bind its ligands rather than directly interacting with the ligand.

Following the determination of the interface between the V-set and first C2-set domain as being critical for Siglec-6 binding to glycolipids, all arginine residues in this area of the Siglec were mutated but none of these mutants lost binding to the same level as other Siglec arginine mutants; although, Arg112-114 showed modestly diminished binding (67% of wildtype). However, as these residues project into the solvent, it is unlikely that these residues directly interact with the sialoside; perhaps the form ionic interactions with the phosphate head groups in the liposomes. The one exception was Arg92; however, when it was mutated to an alanine little to no Siglec-6 was observed on the surface of the cells. Some expression was achieved with the conserved R92K mutant however no binding to neoglycolipid liposomes was observed. After consulting the AlphaFold models, it was found that this arginine residue is likely involved in a salt bridge with Asp115 and indeed, when Asp115 was mutated to and alanine or an asparagine, no cell surface Siglec-6 expression was observed further supporting that this salt bridge is important for Siglec-

6 folding. After finding that no arginine residue supported binding on its own, we considered that an arginine residue was not facilitating the binding to the liposomes.

After consulting a sequence alignment of the Siglecs, it was found that there was a solvent exposed tryptophan in a loop near the conserved arginine residue. Solvent exposed tryptophan residues are not common as they are very hydrophobic, however they are known to insert into bilayers and contribute to binding interactions³¹⁸. As a tryptophan in this position was unique to Siglec-6 and Siglec-6 tended to bind liposomes lacking a ligand more than any other Siglec, we posited that lipid binding through this tryptophan was important for liposome binding. Indeed, when the tryptophan was mutated to an alanine, binding was reduced to levels consistent with untransfected CHO cells. Moreover, naked liposomes bound to W127A CHO cells to the same degree as untransfected CHO cells. These results suggest that liposome binding by Siglec-6 is driven by Trp127 and why Siglec-6 binding to glycolipids is weaker outside the context of a bilayer. Lastly, when a lysine (Lys129) proximal to the conserved arginine residue was mutated to an alanine, binding was diminished to similar levels as when the conserved arginine residue was mutated. Moreover, when the double mutant (R122A, K129A) version of Siglec-6 was tested, binding was reduced even further; however, it was still not reduced to the same level as naked liposomes. There is one other lysine residue in this area (Lys124) that may also contribute to ganglioside recognition, which deserves further investigation.

With all this data in mind, a proposed model of Siglec-6 binding to glycolipids when presented from a bilayer is described in **Figure 4.35**. The first step is for Siglec-6 to bind to the bilayer through the Trp127 and arginine 112, 113, and 114. Following this engagement of the bilayer, the conserved arginine residue forms a salt bridge with the carboxylate of the sialoside. However, in the absence of the conserved arginine residue, Lys129, and potentially Lys124, may potentially compensate for this missing residue and facilitate interactions with sialic acid-containing glycolipids.

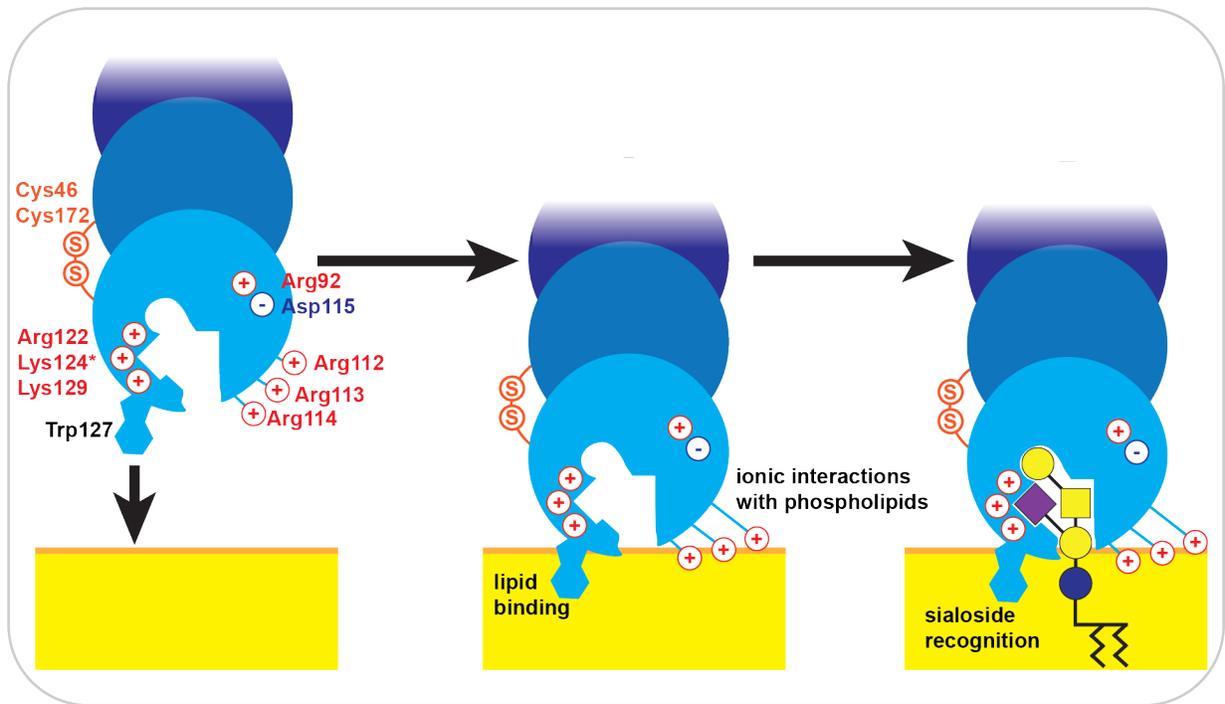


Figure 4.35: Model of Siglec-6 Binding glycolipids presented from a bilayer.

4.4.3: Targeting Siglec-6 under physiological conditions

Using our optimized liposome formulation and neoglycolipid **5**, memory B cells, mast cells, and placental syncytiotrophoblast were all targeted in a Siglec-6 dependant manner. The ability of these liposomes to target Siglec-6 under physiological conditions opens up a novel therapeutic avenue to treat many diseases such as acute myeloid leukemia, allergies, and preeclampsia. Liposomes bearing neoglycolipid **5** could be formulated to encapsulate therapeutics such as anti-cancer drugs (AML), allergy medication (mast cells) or anti-pyrototic drugs such as monomethyl fumarate to treat preeclampsia. However, before the therapeutic potential of this liposomes can be realized, neoglycolipid **5** needs to be chemically derivatized to make binding specific to Siglec-6 to reduce off target effects.

4.4.4: Biological roles of Siglec-6

These studies demonstrate that Siglec-6 is proficient at binding and internalizing EVs independent of its canonical essential arginine. We identified the interface between the V-set domain and the first C2 domain as being critical for glycolipid recognition of EVs and that EVs are internalized in a Siglec-6 dependant manner. While more work is needed to probe the biological roles of Siglec-6 on memory B cells and mast cells, the ability of Siglec-6 to recognize EVs is particularly interesting in the context of syncytiotrophoblasts of the placenta, which are bathed in maternal blood containing an abundance of maternal EVs³¹⁹. Fetal-maternal immunological tolerance relies on communication between the fetus and mother³²⁰, which Siglec-6 has the potential to participate in through recognition of maternal EVs. Similarly, the ability of Siglec-6 to be engaged by ligands on EVs may have important implications for helping to maintain immunological tolerance.

4.5: Conclusion

Using our optimized liposome formulation, it was found that recognition of ganglioside by Siglec-6 is driven by interactions between Trp127 and the liposomal bilayer rather than its conserved arginine residue. Probing the ability of Siglec-6 to recognize a panel of synthetic nGLs yielded neoglycolipid GM3-DAG, which has greater avidity for Siglec-6 compared to natural gangliosides, enabling targeting of liposomes to Siglec-6-expressing cells and tissues, which opens future drug delivery applications. Siglec-6 was also found to engage and internalizes EVs, demonstrating its utility as a versatile immunomodulatory receptor with the potential of participating in immunological tolerance at numerous cellular locations.

4.5: Methods

Cell Assay. The cell assay was performed as described in chapter 3.

Human Samples. All experiments involving human blood samples and placental sample collection were approved by the human research ethics board (HREB) biomedical panel at the University of Alberta.

ELISA. The ELISA was performed as described in chapter 2.

Liposome Over Lectin Assay (LOLA). Siglec-Fc in PBS was adsorbed to a 96-well flat-bottom fluorescent microplate by adding 1 µg/well and incubating the plate at 4 °C overnight. The plate was then washed with PBS followed by blocking with 5 % (m/V) BSA in PBS for 1 h at room temperature. The plate was again washed with PBS and 100 µM liposome in PBS was added to each well. The plate was then incubated at 37 °C for 30 min followed by washing with PBS. PBS was then added to the plate and the fluorescence intensity (λ_{ex} 640 nm, λ_{em} 680 nm) of each well was measured using a Molecular Devices SpectraMAX® iD5.

Generation of Siglec-6/8 Chimeric Constructs. The genes for the constructs containing each domain of Siglec-6 to the two others from Siglec-8 were synthesized by *GeneArt* and designed with a 5' *NheI* and 3' *AgeI* site immediately before the start codon and stop codon respectively. When appropriate, silent mutations were introduced to remove internal *NheI* and *AgeI* cut sites. The additional two constructs containing the two domains of Siglec-6 next to each other were generated by gene overlap extension mutagenesis using the primers in Supplementary Table 7 and using the constructs above as a template.

Isolation of White Blood Cells from Human Spleen. White blood cells were isolated as described in chapter 2.

Liposome Uptake Assay in the Placenta Explants (Performed by the Riddell & Mahal Groups). Was performed as described previously by Shaha *et al.*³²¹. 6.5-7.5 weeks gestation placentas were obtained from elective pregnancy terminations after informed patient consent in accordance with methods approved by the University of Alberta Human Ethics Research Board. Whole

placenta was rinsed in cold PBS and the placenta was cut into uniform 2 mm³ pieces and incubated overnight at 37 °C in Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 10 % (V/V) fetal calf serum. Following overnight incubation, placental explants were serum-starved in serum-free IMDM with 0.5 % (m/V) bovine serum albumin and 25 mM HEPES buffering agent for 1 h then incubated in 50 μM liposome media solution (5 % **5**, 0 % liposome control) for 2 h and washed with cold PBS before fixation in 4 % paraformaldehyde (PFA). For Siglec-6 blocking assays, following overnight incubation, placental explants were first blocked with human IgG (1:50 dilution; Thermo Fisher) for 30 min at 37 °C, and then, explants were incubated either with Siglec-6 antibody (1:50 dilution) or serum-free media for another 30 min at 37 °C. Following blocking, placental explants were incubated with 50 mM **5** liposomes with or without the addition of Siglec-6 antibody (1:100 dilution) for 2 h and washed with cold PBS before overnight fixation in 2 % PFA.

Lentivirus Production. Viral production was performed as described in chapter 2.

Viral Transduction. Viral transduction was performed as described in chapter 2.

Imaging Flow Cytometry and Quantification (Performed by Kelli A. McCord, Macauley Group).

Approximately 2x10⁵ Daudi cells/well were plated into a 96-well U-bottom microplate and centrifuged at 300 x g for 5 min. The cell pellets were placed on ice, and 50 μL of 100 μM of liposome or 1:20 μL of EVs in RPMI growth medium (Gibco) containing 10 % (V/V) fetal bovine serum (FBS; Gibco), 100 U/mL penicillin (Gibco), and 100 μg/mL streptomycin (Gibco) was added to their corresponding wells. Plates were incubated for 1 h at either 4 °C or 37 °C. After incubation, plates were centrifuged at 300 x g, 5 min, 4 °C then incubated with fluorescently conjugated anti-Siglec-6 antibody (Alexa Fluor 488; 1:250 dilution; R&D Systems) for 25 min at 4 °C. The plates were centrifuged once more at 300 x g, 4 °C, 5 min and the resulting cell pellet was resuspended in 40 μL of flow buffer (1 % FBS, 500 μM EDTA Hank's Balanced Salt Solution pH 7.4). 5000 events were collected for each sample using the ImageStream®X Mk II Flow cytometer (excitation

lasers 488 nm and 642 nm, 60X magnification). Data analysis was performed using IDEAS software, version 6.2.

Placenta Explant Immunofluorescence Imaging (Performed by the Riddell Group). Following fixation, tissue was washed and blocked with 5 % normal donkey serum and 0.3 % Triton x100 in PBS and incubated with fluorescently conjugated Siglec-6 antibody (Alexa Fluor 594; 1:200 dilution; R&D Systems) overnight at 4 °C. Samples were then washed and incubated with donkey anti-mouse-AF594 (Alexa Fluor® 594; 5 µg/mL; Invitrogen) and fluorescently conjugated phalloidin (iFluor™ 405; 1:400; AAT Bioquest) for 2 h at room temperature and protected from light. Following incubation, explants were washed with 1X PBS-Tween-20 and PBS and mounted on slides with imaging spacers. 1 µm z-stack slices were taken with a Zeiss LSM700 confocal microscope with a Zeiss Plan Apochromat 63x lens (NA 1.4). Quantification of images was done using Volocity Acquisition Software (Quorum Technologies). For co-localization assays, puncta were defined as objects > 0.1 µm and < 1.5 µm. Total number of puncta per µm³ and number of puncta colocalized with Siglec-6 per µm³ were normalized to liposomal control. For blocking assays, puncta were defined as objects > 0.1 µm and < 1.0 µm. One sample t and Wilcoxon test was run to test the significance of change of normalized total number of puncta per µm³ in the blocking condition compared to non-blocking control. Microscopy Images were processed with ImageJ

Human Extracellular vesicle (EV) Isolation and Labeling (Performed by the Mahal Group). Human peripheral blood, taken under an approved institutional ethics protocol, was centrifuged at 1,700 x g for 5 min at 4 °C. The upper phase (plasma) was collected and centrifuged at 10,000 x g for 30 min at 4 °C. Following 0.22 mm filtration (Millipore), the supernatant was diluted 10 times using PBS and ultra-centrifuged at 110,000 x g for 2 h at 4 °C using Type 70 Ti rotor (Beckman Coulter). The pellet containing the EVs were resuspended in PBS, aliquoted, and stored in -80 °C until further use. For EV fluorescent labeling, EVs were incubated with 1.6 % (V/V) of NHS-Alexa (Alexa Fluor™ 647 or 488 NHS Ester (Succinimidyl Ester); 10 mg/mL in

DMSO stock, Thermo Fisher) or pHrodo (pHrodo™ Red, (Succinimidyl Ester); 10 mg/mL in DMSO stock; Thermo Fisher) overnight at 4 °C or 1 h at room temperature, respectively. Excess dye was removed by spin-filtration, using Ultra-0.5 Centrifugal Filter Units (Millipore). EVs were recovered in PBS, aliquoted, and stored in -80 °C. For EV de-sialylation, 10 µg of EVs were incubated with 5 µL of neuraminidase A or S for 1 h at 37 °C in PBS.

Culturing of LAD2 Cells (*Performed by the Kulka Group*). The LAD2 cell line (a gift from Arnold Kirshenbaum) was cultured in StemPro-34 SFM media (Life Technologies) supplemented with 2 mM L-glutamine, 100 U/ml penicillin, 50 mg/ml streptomycin, and 100 ng/ml recombinant human SCF (PeproTech, Rocky Hill, NJ). Cells were maintained at 1×10^5 cells/ml at 37°C and 5% CO₂ and periodically tested for expression of c-KIT and FcεRI by flow cytometry.

Generation of N2a β1-4GalNT1^{-/-} Cells (*Performed by the Hubbard Group*). The N2a β4GalNT1^{-/-} cell line was generated using CRISPR/Cas9. Briefly, guide RNAs were designed to target the β1-4galnt1 gene and mixed with equimolar quantities of Cas9 to construct ribonucleoprotein complexes which were then transfected into N2a cells using Lipofectamine CRISPRMAX (Thermo Fisher, USA) according to manufacturer's instructions. After 24 h of incubation, cells were sorted for the presence of the ATTO 550 fluorescent marker on the tracrRNA, using FACS Aria™ III cell sorter (BD Biosciences, USA), at the Faculty of Medicine and Dentistry Flow Cytometry Facility, University of Alberta, Canada. To generate clonal cell lines, positive cells were sorted one cell per well into a 96-well plate and further expanded. Confirmation of gene knock-out was obtained by immunoblotting and PCR.

Culturing Neuro2a mouse neuroblastoma cells (N2a) and N2a β1-4GalNT1^{-/-} cells (*Performed by the Sipione Group*). Cells were cultured in DMEM (Cytvia Life Sciences): Opti-MEM I (1:1) supplemented with 10% heat inactivated fetal bovine serum, 2 mM L-glutamine and 0.11 g/L sodium pyruvate and maintained in a 5% CO₂ atmosphere at 37 °C.

EV Isolation from N2a Cell Lines (*Performed by the Sipione Group*). N2a and N2a β1-4galnt1^{-/-} cells were incubated in DMEM supplemented with 2 mM L-glutamine and 0.11 g/L sodium

pyruvate in the absence of serum for 48 h to prevent uptake of gangliosides from the fetal bovine serum. After 48 h, cells were labelled with the lipophilic dye DiD, as previously described³²². Briefly, 5 μ L of DiD (Thermo Fisher) was added to each mL of cell suspension containing 1×10^6 cells in DMEM and incubated for 20 min at 37°C, followed by centrifugation at 400 x g for 5 min at room temperature. The stained cell pellet was further subjected to three rounds of centrifugation in medium containing serum to remove unbound dye. One 500 cm² dish per cell type was seeded with 14×10^6 cells each in phenol red-free DMEM:Opti-MEM I (1:1) supplemented with 1X N-2 supplement, 2 mM L-glutamine and 0.11 g/L sodium pyruvate filtered through a 0.1 μ m polyethersulfone filter for 24 h.

Cell debris and apoptotic bodies were removed from the conditioned medium by centrifugation at 2,000 x g for 10 min at 4 °C in an Eppendorf® Centrifuge 5810 R, using an A-4-62 swinging bucket rotor. The EVs remaining in the cleared conditioned medium were isolated by sequential ultrafiltration and size-exclusion chromatography. Briefly, Amicon® Ultra-15 Centrifugal Filters (100,000 MWCO) were used to concentrate the cleared conditioned medium. To minimize EV loss, the filter membranes were first blocked by centrifugation with 5% Tween-80³²³ in DPBS at 2,600 x g for 10 min at 4 °C, followed by three centrifugations in DPBS for 5 min each. Once blocked, the membrane was kept in DPBS until use to prevent drying. The supernatant containing EVs was concentrated by centrifugation at 2,600 x g at 4 °C until the concentrate volume reached 500 μ L.

The concentrate was applied to a qEVoriginal Gen 2 Size Exclusion Column and fractions were collected using the Automatic Fraction Collector V1 (iZon Science®). The buffer volume was set to 2.9 mL and thirteen 0.5 mL fractions were eluted with DPBS. The presence of EVs in the fractions was determined by measuring DiD fluorescence ($\lambda_{Ex}/\lambda_{Em} = 644/674$ nm) in each fraction using a SpectraMax® i3x multi-mode microplate reader (Molecular Devices, USA). The fractions enriched with EVs (fractions 1-4) were pooled and concentrated using Amicon® Ultra-4mL Filters

(10,000 MWCO) that were blocked with Tween-80 as described above. Purified and concentrated EVs were then stored at -70°C until use.

Transmission Electron Microscopy (*Performed by the Mahal Group*). EVs suspended in PBS or Millipore water, respectively, was placed onto a 300 mesh formvar/copper coated grid (Ted Pella) and left for 3 min for liposomes and 5 min for EVs. Then, the excess liquid was removed, and a 10 μ l drop of 4% uranyl acetate solution was placed onto the grid and left for 5 min and 1 min for liposomes and EVs, respectively. The excess liquid was removed, and the grids were left to dry completely. The grids were analyzed by a Morgagni 268 transmission electron microscope at 80 kV with a Gatan Orius CCD camera.

Chemical Synthesis

General. All reagents were purchased from commercial sources and were used without further purification. THF used in reactions was purified by successive passage through columns of alumina and copper under nitrogen. All reactions were monitored by TLC on silica gel 60-F254 (0.25 mm). Visualization of the reaction components was achieved using UV fluorescence (254 nm) and/or by charring with acidified *p*-anisaldehyde solution in ethanol. Organic solvents were evaporated under reduced pressure below 40 °C, and the products were purified by flash column chromatography on silica gel (230–400 mesh), reverse-phase flash column chromatography (C18) or size exclusion column chromatography (Sephadex-LH-20). HPLC grade CH₃OH was used in the reactions as well as all column purifications. ¹H NMR spectra were recorded at 700 MHz, 600 MHz or 500 MHz, and chemical shifts were referenced to either TMS (0.0, CDCl₃) or CD₃OD (3.30, CD₃OD) or HOD (4.78, D₂O). ¹H data were reported as though they were first order. ¹³C NMR spectra was recorded at 125 MHz and ¹³C chemical shifts were referenced to external acetone (31.07, D₂O). Electrospray mass spectra (HRMS-ESI) were recorded on samples suspended in mixtures of THF with CH₃OH and added NaCl.

Chapter 5: Conclusion and Future Directions

From fish to flowers, fungi to flamingos, in all shapes, forms, sizes, and colours, life is truly amazing. One of the many aspects that makes biology so interesting is that all of the diversity appreciated at the macroscopic level is built from the same molecular building blocks. Some building blocks and their respective polymers like amino acids (proteins) and nucleotides (nucleic acids), are relatively well understood at the molecular level due to their linear nature and their predictable template driven synthesis. On the other hand, monosaccharides (glycans) are comparatively not as well understood as their biosynthesis is not template driven and are branch polymers. Thanks to advancements in the chemical tools such as the development and advancement of the Siglec-Fc^{37, 198}, lectin microarrays^{324, 325}, glycan microarrays^{136, 137}, cell-based glycan arrays^{37, 79, 80}, the glycan/lectin language is starting to be decoded and more defined biological roles for glycans are emerging. This is particularly true with respect to the Siglec family of immunomodulatory lectins. In this thesis, two major tools (Siglec-Fc and glycolipid bearing liposomes) with applications in several assays were developed to help discover and dissect Siglec ligands towards a better understanding of the pathological and pathophysiological roles of Siglecs and glycans (Table 5.1).

Table 5.1: Publications that have used the Macauley Lab Siglec-Fc chimera and or optimized ganglioside liposome.*Original work, †work is under review.

Publication	Tool	Year
Rodrigues E. <i>et al.</i> ^{*37}	Siglec-Fc*	2020
Jung J. <i>et al.</i> ⁷⁹	Siglec-Fc	2021
Galleguillos D. <i>et al.</i> ³²⁶	Ganglioside Liposome	2022
Han L. <i>et al.</i> ²⁷⁸	Ganglioside Liposome	2022
Schmidt E. <i>et al.</i> ^{*128}	Siglec-Fc	2023
	Ganglioside Liposome*	
Bui D.T. <i>et al.</i> ³⁹	Siglec-Fc	2023
Schmidt E.N. <i>et al.</i> ²²⁸	Siglec-Fc	2024
	Ganglioside Liposome	
Garnham R. <i>et al.</i> ¹⁷²	Siglec-Fc	2024
Hodgson K. <i>et al.</i> ²³³	Siglec-Fc	2024
Lima G.M. <i>et al.</i> [†]	Siglec-Fc	2024

In Chapter 2, the Siglec-Fc chimera scaffold which was first developed in the 1990's was improved with respect to versatility, sensitivity, and stability. All of these modifications increase the Siglec-Fc's utility for describing Siglec ligands. The versatility of this advanced version of the Siglec-Fc was demonstrated by its use in various biochemical approaches including cell-based glycan arrays, bead assays, ELISAs, and mass spectrometry-based assays. Using the Siglec-Fc construct in a cell-based glycan array Siglec ligands on primary human immune cells were probed in parallel with Siglec expression. In this screen, it was found that Siglec-1, CD22, and Siglec-7 ligands were found on all immune cell types tested, CD33 ligands were observed on a subset of immune cells and lastly, the ligands of some Siglecs such as Siglec-6, Siglec-11, and Siglec-15 were not found on any of the cells tested. While not always true, it can be generally stated that Siglec ligands can be found on a cell that expresses that Siglec. This implies that many Siglecs are masked at least to some degree under physiological conditions. This could be for many reasons but two will be discussed. One function of Siglec-masking by *cis* ligands could be to directly modulate Siglecs with respect to *cis* and *trans* signalling pathways^{130, 132}. Another possible reason for this basal masking of Siglecs is to prevent Siglecs from binding in *trans* unless forced into an immunological synapse as a mechanism to make Siglecs more difficult to exploit by pathogens. An exception to this was Siglec-9, which was found to be expressed by neutrophils, mature NK cells, and monocytes, yet Siglec-9-Fc only bound weakly to monocytes, suggesting that Siglec-9 is more available to bind in *trans* on neutrophils and mature NK cells. While the results of probing Siglec ligands on primary cells are interesting, a future direction could be to perform the same analysis under different pathophysiological conditions (tumor microenvironment, viral/bacterial infection *etc.*) to see how Siglec expression and Siglec ligand expression changes in response to immunological stress.

Additionally, a cell-based glycan array was used to investigate which types of glycans are the ligands for CD33. Using engineered cells, it was found that the major ligands for CD33 are α -Neup5Ac-(2→3)-(6-O-Sulfo)-Galp-O-GalpNAc O-glycans. However, while the majority of binding

could be attributed to these O-glycans, CD33-Fc binding was still detectable in the absence of complex N-glycans, O-GalNAc glycans and complete O-mannose O-glycans, suggesting that CD33 ligands are found on another class of glycan in addition to O-glycans. Other possible classes of glycans which could be CD33 ligands are gangliosides, GPI-anchors, O-fucose O-glycans, and glyco-RNA. It is noteworthy that results in Chapter 3 did not provide any evidence that CD33 can bind any gangliosides. The caveat to this is that ligand presentation can strongly affect the ability for a Siglec to engage a ligand. Indeed, the glyco-motif α -Neup5Ac-(2→3)-Gal β -O-GalpNAc is not unique to O-glycans and can be found in gangliosides such as GM1b, GD1a, and GT1b and GPI anchors. However, CD33 binding to glycolipids presented from a bilayer has yet to be demonstrated. While it may be biosynthetically possible, sulfated glycolipids have not been well described and the ability for these glycans to be sulfated will be critical for recognition by CD33. Future investigations could probe the possibility of GPI-anchors, O-fucose O-glycans and glyco-RNA to be CD33 ligands using engineered cells treated with various glycan modifying enzymes. Lastly, it was also found that Siglec-Fcs produced from CHO cells with CMAS knocked-out had better sensitivity and were produced at higher yields compared to cells produced from their wildtype counter parts. While the exact reasons as to why this is the case remains unclear, producing Siglec-Fcs from CMAS KO CHO cells and applying them in the assays developed in this chapter will enable Siglec ligands to be better described in the future.

In Chapter 3, a liposome formulation was optimized for engagement of Siglecs through glycolipids. The formulation was optimized with respect to PEG, ganglioside, and cholesterol content, as well as bulk lipid structure. One interesting finding from these optimizations was that high PEG (greater than 2 mol%) content in liposomes impeded Siglec-ganglioside interactions. Another interesting phenomenon that was observed during the optimization of the liposome formulation was that Siglec-ganglioside binding peaked at an optimal ganglioside density. While this decrease in binding was observed at relatively high ganglioside density, if this finding has any genuine biological implications or is just an artifact of the liposome platform has yet to be explored.

This could be relevant in the budding of enveloped viruses or release of EVs. It is possible that viruses which bud from cells with the optimal ganglioside content would be better at exploiting Siglecs and therefore increase their virulence. Moreover, EVs released with just the right amount of ganglioside may target Siglecs much more effectively. However, this is merely speculation. While exploring this hypothesis will be difficult, an excellent starting point would be to isolate viruses (or pseudoviruses) and EVs produced from different cells and quantify the ganglioside content within each particle. The binding of these particles to Siglecs could then be quantified using the approaches developed in Chapter 2 and 3, and the relationship between ganglioside content of these particles and Siglec binding could then be described. This could be the subject of future studies.

While there were many interesting observations made during the optimization of the liposome formulation, applying the formulation to more broadly describe Siglec-ganglioside interactions also yielded many novel findings. Using the optimized formulation, the entire human and mouse Siglec family were interrogated against a panel of nine gangliosides. Many novel interactions were observed, but of most interest was Siglec-6 with many gangliosides. What was most interesting about Siglec-6 binding to glycolipids was that it was able to bind gangliosides independent of its canonical arginine residue. This exciting finding was followed up in Chapter 4. It was also found that despite being a ligand for many of the human Siglecs, GM1a was not found to be a ligand for any of the murine Siglecs. A series of subsequent experiments which applied the approaches developed in Chapter 2, revealed that this was likely due to the sialic acid being presented from the internal galactose rather than the terminal galactose. This observation has implications in using murine models to explore human biology and could be particularly relevant when studying modes of viral infection as many viruses use gangliosides such as GM1a to exploit Siglecs. This preference for GM1a over GM1b by murine Siglecs may suggest that GM1a is more prevalent in humans than mice, however this has not been well described but an investigation into the differential distribution of gangliosides between mice and humans could be the focus of

future investigations. Other than the specific interaction between GM1a and murine Siglecs, Siglec-ganglioside interactions seem to be largely conserved between mice and humans, suggesting that they are an appropriate model to learn more about the biological roles of Siglec-ganglioside interactions in humans.

In Chapter 4, how Siglec-6 binds to glycolipids independent of its canonical arginine residue was investigated using a panel of neoglycolipids. Using these neoglycolipids, it was found that Siglec-6 preferentially binds $\alpha(2\rightarrow3)$ linked sialosides over $\alpha(2\rightarrow6)$ linked sialosides and that binding is significantly improved when the glycolipid is presented from a bilayer. Mutagenesis analysis of Siglec-6 revealed that glycolipid binding is driven by a solvent exposed tryptophan residue and when this residue is mutated binding is lost. One of the neoglycolipids developed in this chapter bound to Siglec-6 with great avidity when presented from an optimized liposome. Using liposomes formulated with this neoglycolipid, Siglec-6 was targeted to primary human tissues which express Siglec-6, demonstrating that liposomes bearing this ligand could be used to direct liposomes to Siglec-6 expressing tissues. This has implications in treating many diseases/conditions in which Siglec-6 is relevant. For instance, anti-cancer drugs could be encapsulated and directed using neoglycolipid bearing liposomes to cancerous cells such as in the case of AML. However, before this formulation can be used for drug delivery, neoglycolipid **5** needs to be chemically modified to be specific for Siglec-6 to avoid off target effects. Modifying sialic acid with hydrophobic groups which exploit hydrophobic pockets in the Siglec binding site has proven fruitful in the past to develop higher affinity and selective Siglec ligands. Neoglycolipid **5** serves as a great starting point to develop a high affinity and selective Siglec-6 ligand using the assays developed in chapter 2. In this chapter, it was also found using a combination of chemoenzymatic and genetic approaches that Siglec-6 binds and internalizes EVs in an $\alpha(2\rightarrow3)$ linked sialic acid dependant manner and that EV binding to Siglec-6 is at least in part mediated by gangliosides. The ability of Siglec-6 to bind and internalize EVs is very interesting as compared to most other Siglecs, Siglec-6 seems to have a very restricted ligand binding preference. Using

glycolipid liposomes and EVs to further explore the immunological niche of Siglec-6 would be a great subject for a future investigation. A biological role for Siglec-6 could be to facilitate the transfer of maternal EVs across the placenta to the fetus. However, this is just a hypothesis and needs to be investigated. Unfortunately, Siglec-6 is not found in mice and there is also no ortholog or paralog making murine models a less than ideal choice to evaluate this hypothesis²⁷². Additionally, there are structural differences between mouse and human placentas further making murine models impractical to address this hypothesis³²⁷. Siglec-6 transgenic guinea pigs may be the most appropriate model to further probe Siglec-6-ganglioside interactions at the fetal maternal interface³²⁸.

In the literature, there are very few studies which have demonstrated the ability for Siglecs to bind gangliosides in genuine cell membranes; however, a couple instances of Siglec-ganglioside binding in this context have been reported. These include MAG (Siglec-4) which inhibits nerve regeneration through engaging with gangliosides, and Siglec-7 binding to GD3 in immune-synapses^{259, 329}. In line with the sparse examples of Siglecs binding gangliosides in cell membranes, no binding to any primary cells or cell lines was observed with Siglec-6-Fc (a ganglioside binder) in a cell-based glycan array. Another finding in this thesis was that high PEG content in liposomes impede ganglioside binding. Perhaps the PEG and the glycocalyx act in a similar capacity and shield the ganglioside from the Siglec, which could be a reason why gangliosides in cell membranes are not often found to be Siglec ligands (**Figure 5.1a**). With this in mind, gangliosides may serve as Siglec ligands in other biological contexts such as EVs. Indeed, many Siglecs (*e.g.* Siglec-1²⁵⁶, Siglec-6¹²⁸, Siglec-7^{257, 261}, and Siglec-9³³⁰) are known to bind EVs and all of these Siglecs are known to bind gangliosides. Perhaps it is in this context where Siglec-ganglioside interactions are most biologically relevant (**Figure 5.1b**). However, if this is the case and what biological purpose this serves still needs to be described which could be the focus of future investigations.

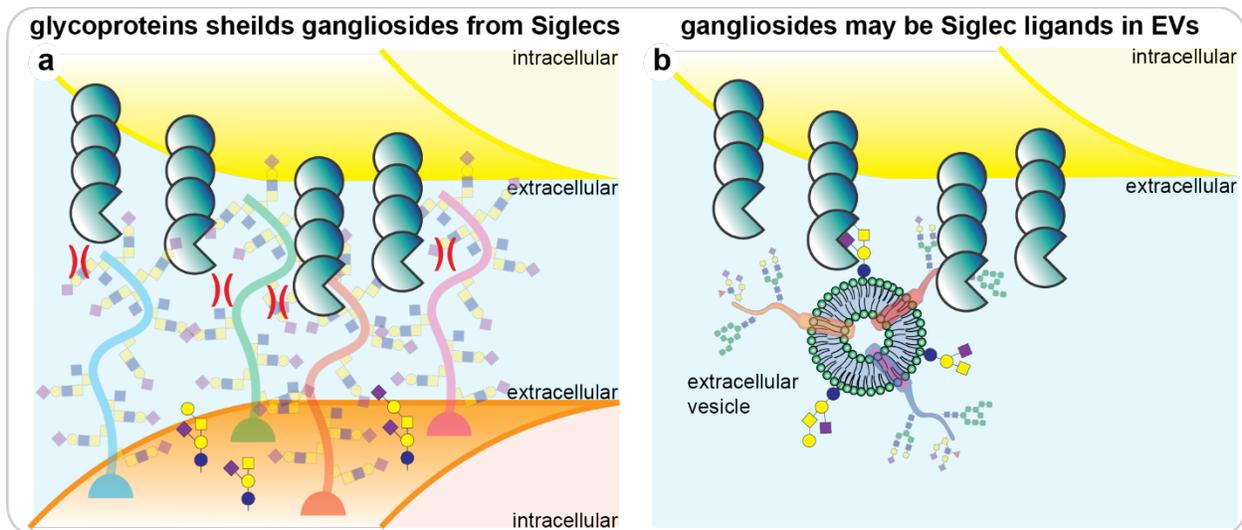


Figure 5.1: Siglec-ganglioside interactions are likely most relevant in Siglec extracellular vesicle interactions.

To best learn about the biological roles of Siglecs, their ligands need to be described and tools that make the systematic dissection of Siglec ligands are needed. In this thesis, two tools (Siglec-Fc and glycolipid bearing liposome) for describing Siglec ligands were refined and applied to study Siglec ligands leading to many novel discoveries. The most exciting discovery was that Siglec-6 can engage glycolipids independent of its conserved arginine residue when the glycolipid is presented from a lipid bilayer. Under physiological conditions, Siglec-6-ganglioside interactions may be important for the endocytosis of EVs by Siglec-6 expressing cells. This thesis highlights how the physiology of Siglecs can be discovered by first describing their ligands. In the future these optimized tools can be used to form a more complete description of Siglec ligands and lead to the development of therapeutics that operate around the Siglec-sialic acid axis (**Figure 5.2**).

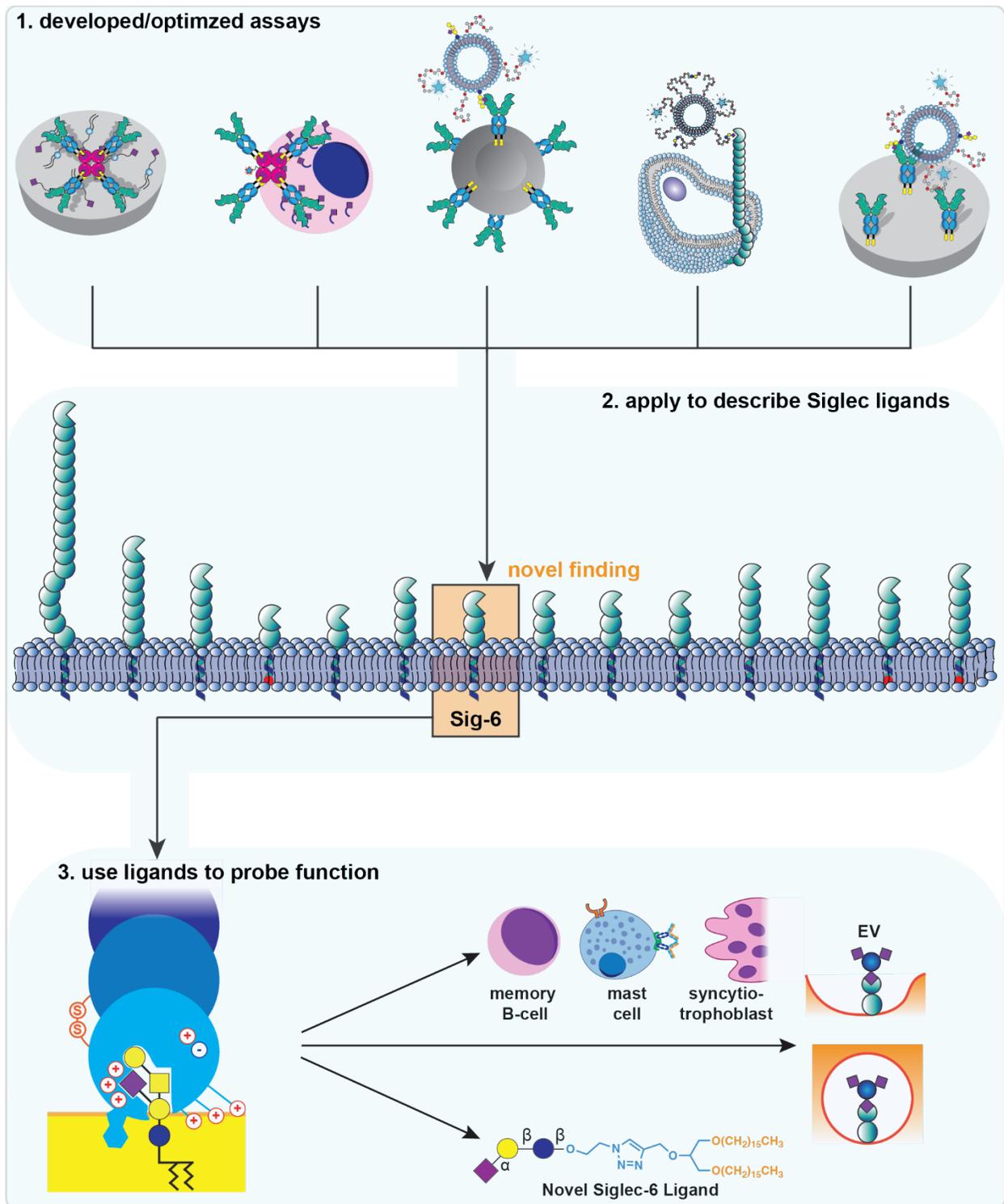


Figure 5.2: Summary of thesis. The development of approaches to study Siglec ligands lead to novel ligand discoveries and then these ligands were used to probe Siglec function.

References

1. Varki A. Biological roles of glycans. *Glycobiology* **27**, 3-49 (2017).
2. Gagneux P, Hennet T, Varki A. Biological Functions of Glycans. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn. Cold Spring Harbor (2022).
3. Breitling J, Aebi M. N-linked protein glycosylation in the endoplasmic reticulum. *Cold Spring Harb Perspect Biol* **5**, a013359 (2013).
4. Jajosky RP, *et al.* ABO blood group antigens and differential glycan expression: Perspective on the evolution of common human enzyme deficiencies. *iScience* **26**, 105798 (2023).
5. Teclé E, Reynoso HS, Wang R, Gagneux P. The female reproductive tract contains multiple innate sialic acid-binding immunoglobulin-like lectins (Siglecs) that facilitate sperm survival. *J Biol Chem* **294**, 11910-11919 (2019).
6. Duan S, Paulson JC. Siglecs as Immune Cell Checkpoints in Disease. *Annu Rev Immunol* **38**, 365-395 (2020).
7. Dube DH, Bertozzi CR. Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nat Rev Drug Discov* **4**, 477-488 (2005).
8. Haukedal H, Freude KK. Implications of Glycosylation in Alzheimer's Disease. *Front Neurosci* **14**, 625348 (2020).
9. Perez-Zsolt D, *et al.* Anti-Siglec-1 antibodies block Ebola viral uptake and decrease cytoplasmic viral entry. *Nat Microbiol* **4**, 1558-1570 (2019).
10. Martinez-Duncker I, Chiodo F, Mora-Montes HM, Vasta GR. Editorial: The Role of Glycans in Infectious Disease. *Front Microbiol* **13**, 921436 (2022).
11. Dejean G, *et al.* Synergy between Cell Surface Glycosidases and Glycan-Binding Proteins Dictates the Utilization of Specific Beta(1,3)-Glucans by Human Gut Bacteroides. *mBio* **11**, (2020).
12. Horiya S, MacPherson IS, Krauss IJ. Recent strategies targeting HIV glycans in vaccine design. *Nat Chem Biol* **10**, 990-999 (2014).
13. Eisenson DL, Hisadome Y, Yamada K. Progress in Xenotransplantation: Immunologic Barriers, Advances in Gene Editing, and Successful Tolerance Induction Strategies in Pig-To-Primate Transplantation. *Front Immunol* **13**, 899657 (2022).
14. Jin C, *et al.* Structural diversity of human gastric mucin glycans. *Mol Cell Proteomics* **16**, 743-758 (2017).

15. Bieberich E. Synthesis, Processing, and Function of N-glycans in N-glycoproteins. *Adv Neurobiol* **9**, 47-70 (2014).
16. Kellokumpu S, Hassinen A, Glumoff T. Glycosyltransferase complexes in eukaryotes: long-known, prevalent but still unrecognized. *Cell Mol Life Sci* **73**, 305-325 (2016).
17. Cummings RD, Pierce JM. The challenge and promise of glycomics. *Chem Biol* **21**, 1-15 (2014).
18. Seeberger PH, Overkleeft HS. Chemical Synthesis of Glycans and Glycoconjugates. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn. Cold Spring Harbor Laboratory Press (2022).
19. Downey AM, Hocek M. Strategies toward protecting group-free glycosylation through selective activation of the anomeric center. *Beilstein J Org Chem* **13**, 1239-1279 (2017).
20. Seeberger PH. Monosaccharide Diversity. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn (2022).
21. Werz DB, Ranzinger R, Herget S, Adibekian A, von der Lieth CW, Seeberger PH. Exploring the structural diversity of mammalian carbohydrates ("glycospace") by statistical databank analysis. *ACS Chem Biol* **2**, 685-691 (2007).
22. Varki A, *et al.* Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **25**, 1323-1324 (2015).
23. Mikkola S. Nucleotide Sugars in Chemistry and Biology. *Molecules* **25**, (2020).
24. Kapitonov D, Yu RK. Conserved domains of glycosyltransferases. *Glycobiology* **9**, 961-978 (1999).
25. Rini JM, Moremen KW, Davis BG, Esko JD. Glycosyltransferases and Glycan-Processing Enzymes. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn (2022).
26. Breton C, Snajdrova L, Jeanneau C, Koca J, Imberty A. Structures and mechanisms of glycosyltransferases. *Glycobiology* **16**, 29R-37R (2006).
27. Lairson LL, Henrissat B, Davies GJ, Withers SG. Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* **77**, 521-555 (2008).
28. Schuman B, Evans SV, Fyles TM. Geometric attributes of retaining glycosyltransferase enzymes favor an orthogonal mechanism. *PLoS One* **8**, e71077 (2013).

29. Joud M, Moller M, Olsson ML. Identification of human glycosyltransferase genes expressed in erythroid cells predicts potential carbohydrate blood group loci. *Sci Rep* **8**, 6040 (2018).
30. Stanley P, Wuhrer M, Lauc G, Stowell SR, Cummings RD. Structures Common to Different Glycans. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn. Cold Spring Harbor Laboratory Press (2022).
31. Stanley P, Moremen KW, Lewis NE, Taniguchi N, Aebi M. N-Glycans. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn (2022).
32. Zhou Q, Qiu H. The Mechanistic Impact of N-Glycosylation on Stability, Pharmacokinetics, and Immunogenicity of Therapeutic Proteins. *J Pharm Sci* **108**, 1366-1377 (2019).
33. Rao RS, Buus OT, Wollenweber B. Evolutionary Pattern of N-Glycosylation Sequon Numbers in Eukaryotic ABC Protein Superfamilies. *Bioinform Biol Insights* **4**, 9-17 (2010).
34. Shirakawa A, Manabe Y, Fukase K. Recent Advances in the Chemical Biology of N-Glycans. *Molecules* **26**, (2021).
35. Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. *Nat Rev Nephrol* **15**, 346-366 (2019).
36. Byrne G, *et al.* CRISPR/Cas9 gene editing for the creation of an MGAT1-deficient CHO cell line to control HIV-1 vaccine glycosylation. *PLoS Biol* **16**, e2005817 (2018).
37. Rodrigues E, *et al.* A versatile soluble siglec scaffold for sensitive and quantitative detection of glycan ligands. *Nat Commun* **11**, 5091 (2020).
38. Maley F, Trimble RB, Tarentino AL, Plummer TH, Jr. Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Anal Biochem* **180**, 195-204 (1989).
39. Bui DT, *et al.* Absolute Affinities from Quantitative Shotgun Glycomics Using Concentration-Independent (COIN) Native Mass Spectrometry. *ACS Cent Sci* **9**, 1374-1387 (2023).
40. Brockhausen I, Wandall HH, Hagen KGT, Stanley P. O-GalNAc Glycans. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn (2022).
41. Wilkinson H, Saldova R. Current Methods for the Characterization of O-Glycans. *J Proteome Res* **19**, 3890-3905 (2020).

42. Briggs DC, Hohenester E. Structural Basis for the Initiation of Glycosaminoglycan Biosynthesis by Human Xylosyltransferase 1. *Structure* **26**, 801-809 e803 (2018).
43. Wang Y, *et al.* Cosmc is an essential chaperone for correct protein O-glycosylation. *Proc Natl Acad Sci U S A* **107**, 9228-9233 (2010).
44. Huet G, *et al.* GalNAc-alpha-O-benzyl inhibits NeuAcalpha2-3 glycosylation and blocks the intracellular transport of apical glycoproteins and mucus in differentiated HT-29 cells. *J Cell Biol* **141**, 1311-1322 (1998).
45. Kameyama A, Thet Tin WW, Toyoda M, Sakaguchi M. A practical method of liberating O-linked glycans from glycoproteins using hydroxylamine and an organic superbase. *Biochem Biophys Res Commun* **513**, 186-192 (2019).
46. Wang S, *et al.* Chemoenzymatic modular assembly of O-GalNAc glycans for functional glycomics. *Nat Commun* **12**, 3573 (2021).
47. Ju T, Otto VI, Cummings RD. The Tn antigen-structural simplicity and biological complexity. *Angew Chem Int Ed Engl* **50**, 1770-1791 (2011).
48. Munkley J. The Role of Sialyl-Tn in Cancer. *Int J Mol Sci* **17**, 275 (2016).
49. Bergstrom KS, Xia L. Mucin-type O-glycans and their roles in intestinal homeostasis. *Glycobiology* **23**, 1026-1037 (2013).
50. Mantelli F, Argueso P. Functions of ocular surface mucins in health and disease. *Curr Opin Allergy Clin Immunol* **8**, 477-483 (2008).
51. Lillehoj EP, Kato K, Lu W, Kim KC. Cellular and molecular biology of airway mucins. *Int Rev Cell Mol Biol* **303**, 139-202 (2013).
52. Kang Y, Park H, Choe BH, Kang B. The Role and Function of Mucins and Its Relationship to Inflammatory Bowel Disease. *Front Med (Lausanne)* **9**, 848344 (2022).
53. Praissman JL, Wells L. Mammalian O-mannosylation pathway: glycan structures, enzymes, and protein substrates. *Biochemistry* **53**, 3066-3078 (2014).
54. Manya H, *et al.* Demonstration of mammalian protein O-mannosyltransferase activity: coexpression of POMT1 and POMT2 required for enzymatic activity. *Proc Natl Acad Sci U S A* **101**, 500-505 (2004).
55. Schneider M, Al-Shareffi E, Haltiwanger RS. Biological functions of fucose in mammals. *Glycobiology* **27**, 601-618 (2017).
56. Merry CLR, Lindahl U, Couchman J, Esko JD. Proteoglycans and Sulfated Glycosaminoglycans. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn (2022).

57. Zachara NE, Akimoto Y, Boyce M, Hart GW. The O-GlcNAc Modification. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn (2022).
58. Yang X, Qian K. Protein O-GlcNAcylation: emerging mechanisms and functions. *Nat Rev Mol Cell Biol* **18**, 452-465 (2017).
59. Schnaar RL, Sandhoff R, Tiemeyer M, Kinoshita T. Glycosphingolipids. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn (2022).
60. Kolter T. Ganglioside biochemistry. *ISRN Biochem* **2012**, 506160 (2012).
61. Nguyen L, *et al.* Sialic acid-containing glycolipids mediate binding and viral entry of SARS-CoV-2. *Nat Chem Biol* **18**, 81-90 (2022).
62. Sipione S, Monyror J, Galleguillos D, Steinberg N, Kadam V. Gangliosides in the Brain: Physiology, Pathophysiology and Therapeutic Applications. *Front Neurosci* **14**, 572965 (2020).
63. Dolezal S, Hester S, Kirby PS, Nairn A, Pierce M, Abbott KL. Elevated levels of glycosylphosphatidylinositol (GPI) anchored proteins in plasma from human cancers detected by *C. septicum* alpha toxin. *Cancer Biomark* **14**, 55-62 (2014).
64. Komath SS, Fujita M, Hart GW, Ferguson MAJ, Kinoshita T. Glycosylphosphatidylinositol Anchors. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn (2022).
65. Chantziou A, *et al.* Glycosylation Modulates Plasma Membrane Trafficking of CD24 in Breast Cancer Cells. *Int J Mol Sci* **22**, (2021).
66. Minakata S, *et al.* Protein C-Mannosylation and C-Mannosyl Tryptophan in Chemical Biology and Medicine. *Molecules* **26**, (2021).
67. Gagneux P, Panin V, Hennet T, Aebi M, Varki A. Evolution of Glycan Diversity. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn (2022).
68. Sharma S, Shekhar S, Sharma B, Jain P. Decoding glycans: deciphering the sugary secrets to be coherent on the implication. *RSC Adv* **10**, 34099-34113 (2020).
69. Moody AM, *et al.* Sialic acid capping of CD8beta core 1-O-glycans controls thymocyte-major histocompatibility complex class I interaction. *J Biol Chem* **278**, 7240-7246 (2003).
70. Hobbs SJ, Nolz JC. Regulation of T Cell Trafficking by Enzymatic Synthesis of O-Glycans. *Front Immunol* **8**, 600 (2017).

71. Brazil JC, Parkos CA. Finding the sweet spot: glycosylation mediated regulation of intestinal inflammation. *Mucosal Immunol* **15**, 211-222 (2022).
72. Regan FAM. Blood Cell Antigens and Antibodies. In: *Dacie and Lewis Practical Haematology* (2017).
73. Huai G, Qi P, Yang H, Wang Y. Characteristics of alpha-Gal epitope, anti-Gal antibody, alpha1,3 galactosyltransferase and its clinical exploitation (Review). *Int J Mol Med* **37**, 11-20 (2016).
74. Commins SP, Platts-Mills TA. Delayed anaphylaxis to red meat in patients with IgE specific for galactose alpha-1,3-galactose (alpha-gal). *Curr Allergy Asthma Rep* **13**, 72-77 (2013).
75. Yu RK, Tsai YT, Ariga T, Yanagisawa M. Structures, biosynthesis, and functions of gangliosides--an overview. *J Oleo Sci* **60**, 537-544 (2011).
76. Villanueva-Cabello TM, Gutierrez-Valenzuela LD, Salinas-Marin R, Lopez-Guerrero DV, Martinez-Duncker I. Polysialic Acid in the Immune System. *Front Immunol* **12**, 823637 (2021).
77. Zhao C, Cooper DKC, Dai Y, Hara H, Cai Z, Mou L. The Sda and Cad glycan antigens and their glycosyltransferase, beta1,4GalNAcT-II, in xenotransplantation. *Xenotransplantation* **25**, e12386 (2018).
78. Kadirvelraj R, Yang JY, Kim HW, Sanders JH, Moremen KW, Wood ZA. Comparison of human poly-N-acetyl-lactosamine synthase structure with GT-A fold glycosyltransferases supports a modular assembly of catalytic subsites. *J Biol Chem* **296**, 100110 (2021).
79. Jung J, *et al.* Carbohydrate Sulfation As a Mechanism for Fine-Tuning Siglec Ligands. *ACS Chem Biol* **16**, 2673-2689 (2021).
80. Bull C, *et al.* Probing the binding specificities of human Siglecs by cell-based glycan arrays. *Proc Natl Acad Sci U S A* **118**, (2021).
81. Hemmerich S, Rosen SD. Carbohydrate sulfotransferases in lymphocyte homing. *Glycobiology* **10**, 849-856 (2000).
82. Luis AS, *et al.* Sulfated glycan recognition by carbohydrate sulfatases of the human gut microbiota. *Nat Chem Biol* **18**, 841-849 (2022).
83. Bhide GP, Colley KJ. Sialylation of N-glycans: mechanism, cellular compartmentalization and function. *Histochem Cell Biol* **147**, 149-174 (2017).
84. Li Y, Chen X. Sialic acid metabolism and sialyltransferases: natural functions and applications. *Appl Microbiol Biotechnol* **94**, 887-905 (2012).

85. Varki A. Colloquium paper: uniquely human evolution of sialic acid genetics and biology. *Proc Natl Acad Sci U S A* **107 Suppl 2**, 8939-8946 (2010).
86. Shiozaki K, Uezono K, Hirai G, Honda A, Minoda M, Wakata R. Identification of novel fish sialidase genes responsible for KDN-cleaving activity. *Glycoconj J* **37**, 745-753 (2020).
87. Nemanichvili N, *et al.* Wild and domestic animals variably display Neu5Ac and Neu5Gc sialic acids. *Glycobiology* **32**, 791-802 (2022).
88. Paul A, Bachar Abramovitch S, Padler-Karavani V. Specific Detection of Neu5Gc in Animal Tissues by Immunohistochemistry. *Methods Mol Biol* **2110**, 59-72 (2020).
89. Park SS. Post-Glycosylation Modification of Sialic Acid and Its Role in Virus Pathogenesis. *Vaccines (Basel)* **7**, (2019).
90. Varki A. Since there are PAMPs and DAMPs, there must be SAMPs? Glycan "self-associated molecular patterns" dampen innate immunity, but pathogens can mimic them. *Glycobiology* **21**, 1121-1124 (2011).
91. Laubli H, Varki A. Sialic acid-binding immunoglobulin-like lectins (Siglecs) detect self-associated molecular patterns to regulate immune responses. *Cell Mol Life Sci* **77**, 593-605 (2020).
92. Jennings MP, Day CJ, Atack JM. How bacteria utilize sialic acid during interactions with the host: snip, snatch, dispatch, match and attach. *Microbiology (Reading)* **168**, (2022).
93. Feng T, *et al.* Glycosylation of viral proteins: Implication in virus-host interaction and virulence. *Virulence* **13**, 670-683 (2022).
94. Izquierdo-Useros N, *et al.* Siglec-1 is a novel dendritic cell receptor that mediates HIV-1 trans-infection through recognition of viral membrane gangliosides. *PLoS Biol* **10**, e1001448 (2012).
95. Rodrigues E, Macauley MS. Hypersialylation in Cancer: Modulation of Inflammation and Therapeutic Opportunities. *Cancers (Basel)* **10**, (2018).
96. Dobie C, Skropeta D. Insights into the role of sialylation in cancer progression and metastasis. *Br J Cancer* **124**, 76-90 (2021).
97. Murugesan G, Weigle B, Crocker PR. Siglec and anti-Siglec therapies. *Curr Opin Chem Biol* **62**, 34-42 (2021).
98. Angata T, von Gunten S, Schnaar RL, Varki A. I-Type Lectins. In: *Essentials of Glycobiology* (eds Varki A, *et al.*). 4th edn (2022).

99. Angata T. Siglec-15: a potential regulator of osteoporosis, cancer, and infectious diseases. *J Biomed Sci* **27**, 10 (2020).
100. Crocker PR, Paulson JC, Varki A. Siglecs and their roles in the immune system. *Nat Rev Immunol* **7**, 255-266 (2007).
101. Pillai S, Netravali IA, Cariappa A, Mattoo H. Siglecs and immune regulation. *Annu Rev Immunol* **30**, 357-392 (2012).
102. Bornhöff K, Goldammer T, Rebl A, Galuska SP. Siglecs: A journey through the evolution of sialic acid-binding immunoglobulin-type lectins. *Dev Comp Immunol* **86**, 219-231 (2018).
103. Cao H, Crocker PR. Evolution of CD33-related siglecs: regulating host immune functions and escaping pathogen exploitation? *Immunology* **132**, 18-26 (2011).
104. Varki A, Angata T. Siglecs--the major subfamily of I-type lectins. *Glycobiology* **16**, 1R-27R (2006).
105. Jensen RA. Orthologs and paralogs - we need to get it right. *Genome Biol* **2**, INTERACTIONS1002 (2001).
106. Mitra N, *et al.* SIGLEC12, a human-specific segregating (pseudo)gene, encodes a signaling molecule expressed in prostate carcinomas. *J Biol Chem* **286**, 23003-23011 (2011).
107. Wang X, *et al.* Specific inactivation of two immunomodulatory SIGLEC genes during human evolution. *Proceedings of the National Academy of Sciences* **109**, 9935-9940 (2012).
108. Lehmann F, Gathje H, Kelm S, Dietz F. Evolution of sialic acid-binding proteins: molecular cloning and expression of fish siglec-4. *Glycobiology* **14**, 959-968 (2004).
109. Schmassmann P, *et al.* Targeting the Siglec-sialic acid axis promotes antitumor immune responses in preclinical models of glioblastoma. *Sci Transl Med* **15**, eadf5302 (2023).
110. Nguyen DH, Hurtado-Ziola N, Gagneux P, Varki A. Loss of Siglec expression on T lymphocytes during human evolution. *Proc Natl Acad Sci U S A* **103**, 7765-7770 (2006).
111. Macauley MS, *et al.* Unmasking of CD22 Co-receptor on Germinal Center B-cells Occurs by Alternative Mechanisms in Mouse and Man. *J Biol Chem* **290**, 30066-30077 (2015).
112. Jiang HS, Zhuang SC, Lam CH, Chang LY, Angata T. Recent Progress in the Methodologies to Identify Physiological Ligands of Siglecs. *Front Immunol* **12**, 813082 (2021).

113. Salahudeen MS, Nishtala PS. An overview of pharmacodynamic modelling, ligand-binding approach and its application in clinical practice. *Saudi Pharm J* **25**, 165-175 (2017).
114. Lange NA, Speight J. *Lange's Handbook of Chemistry, 70th Anniversary Edition*. McGraw-Hill Education (2005).
115. Ge M, Xia XY, Pan XM. Salt bridges in the hyperthermophilic protein Ssh10b are resilient to temperature increases. *J Biol Chem* **283**, 31690-31696 (2008).
116. Sedov IA, Solomonov BN. Determining the Gibbs Energies of Hydrogen-Bonding Interactions of Proton-Accepting Solutes in Aqueous Solutions from Thermodynamic Data at 298 K with Regard to the Hydrophobic Effect. *Journal of Chemical & Engineering Data* **56**, 1438-1442 (2011).
117. Sjoberg ER, Powell LD, Klein A, Varki A. Natural ligands of the B cell adhesion molecule CD22 beta can be masked by 9-O-acetylation of sialic acids. *J Cell Biol* **126**, 549-562 (1994).
118. Du X, *et al.* Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int J Mol Sci* **17**, (2016).
119. Sun Q. The Hydrophobic Effects: Our Current Understanding. *Molecules* **27**, (2022).
120. Yunn NO, Kim J, Ryu SH, Cho Y. A stepwise activation model for the insulin receptor. *Exp Mol Med* **55**, 2147-2161 (2023).
121. Koide A, Tereshko V, Uysal S, Margalef K, Kossiakoff AA, Koide S. Exploring the capacity of minimalist protein interfaces: interface energetics and affinity maturation to picomolar KD of a single-domain antibody with a flat paratope. *J Mol Biol* **373**, 941-953 (2007).
122. Luong JHT, Vashist SK. Chemistry of Biotin-Streptavidin and the Growing Concern of an Emerging Biotin Interference in Clinical Immunoassays. *ACS Omega* **5**, 10-18 (2020).
123. O'Reilly MK, Paulson JC. Multivalent ligands for siglecs. *Methods Enzymol* **478**, 343-363 (2010).
124. Erlendsson S, Teilum K. Binding Revisited-Avidity in Cellular Function and Signaling. *Front Mol Biosci* **7**, 615565 (2020).
125. Macauley MS, Crocker PR, Paulson JC. Siglec-mediated regulation of immune cell function in disease. *Nat Rev Immunol* **14**, 653-666 (2014).

126. O'Reilly MK, Tian H, Paulson JC. CD22 is a recycling receptor that can shuttle cargo between the cell surface and endosomal compartments of B cells. *J Immunol* **186**, 1554-1563 (2011).
127. Bhattacharjee A, *et al.* Increasing phagocytosis of microglia by targeting CD33 with liposomes displaying glycan ligands. *J Control Release* **338**, 680-693 (2021).
128. Schmidt EN, *et al.* Siglec-6 mediates the uptake of extracellular vesicles through a noncanonical glycolipid binding pocket. *Nat Commun* **14**, 2327 (2023).
129. Biedermann B, Gil D, Bowen DT, Crocker PR. Analysis of the CD33-related siglec family reveals that Siglec-9 is an endocytic receptor expressed on subsets of acute myeloid leukemia cells and absent from normal hematopoietic progenitors. *Leuk Res* **31**, 211-220 (2007).
130. Stewart N, *et al.* The glycoimmune checkpoint receptor Siglec-7 interacts with T-cell ligands and regulates T-cell activation. *J Biol Chem* **300**, 105579 (2024).
131. von Gunten S, *et al.* Siglec-9 transduces apoptotic and nonapoptotic death signals into neutrophils depending on the proinflammatory cytokine environment. *Blood* **106**, 1423-1431 (2005).
132. McCord KA, *et al.* Dissecting the Ability of Siglecs To Antagonize Fcγ Receptors. *ACS Cent Sci* **10**, 315-330 (2024).
133. Enterina JR, Jung J, Macauley MS. Coordinated roles for glycans in regulating the inhibitory function of CD22 on B cells. *Biomed J* **42**, 218-232 (2019).
134. Lu S, *et al.* Activation pathway of a G protein-coupled receptor uncovers conformational intermediates as targets for allosteric drug design. *Nat Commun* **12**, 4721 (2021).
135. Sojitra M, *et al.* Genetically encoded multivalent liquid glycan array displayed on M13 bacteriophage. *Nat Chem Biol* **17**, 806-816 (2021).
136. Padler-Karavani V, *et al.* Cross-comparison of protein recognition of sialic acid diversity on two novel sialoglycan microarrays. *J Biol Chem* **287**, 22593-22608 (2012).
137. Padler-Karavani V, *et al.* Rapid evolution of binding specificities and expression patterns of inhibitory CD33-related Siglecs in primates. *FASEB J* **28**, 1280-1293 (2014).
138. Campanero-Rhodes MA, *et al.* Carbohydrate microarrays reveal sulphation as a modulator of siglec binding. *Biochem Biophys Res Commun* **344**, 1141-1146 (2006).
139. Gonzalez-Gil A, *et al.* Sialylated keratan sulfate proteoglycans are Siglec-8 ligands in human airways. *Glycobiology* **28**, 786-801 (2018).

140. Yu H, *et al.* Siglec-8 and Siglec-9 binding specificities and endogenous airway ligand distributions and properties. *Glycobiology* **27**, 657-668 (2017).
141. Rapoport E, Mikhalyov I, Zhang J, Crocker P, Bovin N. Ganglioside binding pattern of CD33-related siglecs. *Bioorg Med Chem Lett* **13**, 675-678 (2003).
142. Yoshimura A, Hatanaka R, Tanaka H, Kitajima K, Sato C. The conserved arginine residue in all siglecs is essential for Siglec-7 binding to sialic acid. *Biochem Biophys Res Commun* **534**, 1069-1075 (2021).
143. Yamakawa N, *et al.* Discovery of a new sialic acid binding region that regulates Siglec-7. *Sci Rep* **10**, 8647 (2020).
144. Hyun JY, Pai J, Shin I. The Glycan Microarray Story from Construction to Applications. *Acc Chem Res* **50**, 1069-1078 (2017).
145. Pochechueva T, *et al.* Comparison of printed glycan array, suspension array and ELISA in the detection of human anti-glycan antibodies. *Glycoconjugate Journal* **28**, 507-517 (2011).
146. Babulic JL, *et al.* One-Step Selective Labeling of Native Cell Surface Sialoglycans by Exogenous alpha2,8-Sialylation. *ACS Chem Biol* **18**, 2418-2429 (2023).
147. De Leon Gonzalez FV, Boddington ME, Kofsky JM, Prindl MI, Capicciotti CJ. Glyco-Engineering Cell Surfaces by Exo-Enzymatic Installation of GlcNAz and LacNAz Motifs. *ACS Chem Biol* **19**, 629-640 (2024).
148. Sheikh MO, *et al.* Cell surface glycan engineering reveals that matriglycan alone can recapitulate dystroglycan binding and function. *Nat Commun* **13**, 3617 (2022).
149. Nguyen L, *et al.* Sialic acid-containing glycolipids mediate binding and viral entry of SARS-CoV-2. *Nat Chem Biol* **18**, 81-90 (2022).
150. Nijen Twilhaar MK, *et al.* Optimization of Liposomes for Antigen Targeting to Splenic CD169(+) Macrophages. *Pharmaceutics* **12**, (2020).
151. Affandi AJ, *et al.* CD169 Defines Activated CD14(+) Monocytes With Enhanced CD8(+) T Cell Activation Capacity. *Front Immunol* **12**, 697840 (2021).
152. Affandi AJ, *et al.* Selective tumor antigen vaccine delivery to human CD169(+) antigen-presenting cells using ganglioside-liposomes. *Proc Natl Acad Sci U S A* **117**, 27528-27539 (2020).
153. Lin CL, *et al.* Chemoenzymatic synthesis of genetically-encoded multivalent liquid N-glycan arrays. *Nat Commun* **14**, 5237 (2023).

154. Kelm S, *et al.* Sialoadhesin, myelin-associated glycoprotein and CD22 define a new family of sialic acid-dependent adhesion molecules of the immunoglobulin superfamily. *Curr Biol* **4**, 965-972 (1994).
155. Duan S, *et al.* Nanoparticles Displaying Allergen and Siglec-8 Ligands Suppress IgE-FcepsilonRI-Mediated Anaphylaxis and Desensitize Mast Cells to Subsequent Antigen Challenge. *J Immunol* **206**, 2290-2300 (2021).
156. Champagne-Jorgensen K, Luong T, Darby T, Roach DR. Immunogenicity of bacteriophages. *Trends Microbiol* **31**, 1058-1071 (2023).
157. Inglut CT, *et al.* Immunological and Toxicological Considerations for the Design of Liposomes. *Nanomaterials (Basel)* **10**, (2020).
158. El-Hawiet A, Shoemaker GK, Daneshfar R, Kitova EN, Klassen JS. Applications of a catch and release electrospray ionization mass spectrometry assay for carbohydrate library screening. *Anal Chem* **84**, 50-58 (2012).
159. Angata T. Possible Influences of Endogenous and Exogenous Ligands on the Evolution of Human Siglecs. *Front Immunol* **9**, 2885 (2018).
160. Nycholat CM, Rademacher C, Kawasaki N, Paulson JC. In silico-aided design of a glycan ligand of sialoadhesin for in vivo targeting of macrophages. *J Am Chem Soc* **134**, 15696-15699 (2012).
161. Brinkman-Van der Linden EC, Sjoberg ER, Juneja LR, Crocker PR, Varki N, Varki A. Loss of N-glycolylneuraminic acid in human evolution. Implications for sialic acid recognition by siglecs. *J Biol Chem* **275**, 8633-8640 (2000).
162. Razi N, Varki A. Masking and unmasking of the sialic acid-binding lectin activity of CD22 (Siglec-2) on B lymphocytes. *Proc Natl Acad Sci U S A* **95**, 7469-7474 (1998).
163. Collins BE, Smith BA, Bengtson P, Paulson JC. Ablation of CD22 in ligand-deficient mice restores B cell receptor signaling. *Nat Immunol* **7**, 199-206 (2006).
164. Collins BE, *et al.* High-affinity ligand probes of CD22 overcome the threshold set by cis ligands to allow for binding, endocytosis, and killing of B cells. *J Immunol* **177**, 2994-3003 (2006).
165. Wasim L, *et al.* N-Linked Glycosylation Regulates CD22 Organization and Function. *Front Immunol* **10**, 699 (2019).
166. Brinkman-Van der Linden EC, Varki A. New aspects of siglec binding specificities, including the significance of fucosylation and of the sialyl-Tn epitope. Sialic acid-binding immunoglobulin superfamily lectins. *J Biol Chem* **275**, 8625-8632 (2000).

167. Rillahan CD, *et al.* Disubstituted Sialic Acid Ligands Targeting Siglecs CD33 and CD22 Associated with Myeloid Leukaemias and B Cell Lymphomas. *Chem Sci* **5**, 2398-2406 (2014).
168. Shelke SV, *et al.* A fragment-based in situ combinatorial approach to identify high-affinity ligands for unknown binding sites. *Angew Chem Int Ed Engl* **49**, 5721-5725 (2010).
169. Patel N, *et al.* OB-BP1/Siglec-6. a leptin- and sialic acid-binding protein of the immunoglobulin superfamily. *J Biol Chem* **274**, 22729-22738 (1999).
170. Kawasaki Y, *et al.* Ganglioside DSGb5, preferred ligand for Siglec-7, inhibits NK cell cytotoxicity against renal cell carcinoma cells. *Glycobiology* **20**, 1373-1379 (2010).
171. Chang LY, *et al.* Molecular Basis and Role of Siglec-7 Ligand Expression on Chronic Lymphocytic Leukemia B Cells. *Front Immunol* **13**, 840388 (2022).
172. Garnham R, *et al.* ST3 beta-galactoside alpha-2,3-sialyltransferase 1 (ST3Gal1) synthesis of Siglec ligands mediates anti-tumour immunity in prostate cancer. *Commun Biol* **7**, 276 (2024).
173. Sonnenburg JL, Altheide TK, Varki A. A uniquely human consequence of domain-specific functional adaptation in a sialic acid-binding receptor. *Glycobiology* **14**, 339-346 (2004).
174. Kroezen BS, *et al.* A Potent Mimetic of the Siglec-8 Ligand 6'-Sulfo-Sialyl Lewis(x). *ChemMedChem* **15**, 1706-1719 (2020).
175. Nycholat CM, *et al.* A Sulfonamide Sialoside Analogue for Targeting Siglec-8 and -F on Immune Cells. *J Am Chem Soc* **141**, 14032-14037 (2019).
176. Rillahan CD, Schwartz E, McBride R, Fokin VV, Paulson JC. Click and pick: identification of sialoside analogues for siglec-based cell targeting. *Angew Chem Int Ed Engl* **51**, 11014-11018 (2012).
177. Alborzian Deh Sheikh A, *et al.* A Guillain-Barre syndrome-associated SIGLEC10 rare variant impairs its recognition of gangliosides. *J Autoimmun* **116**, 102571 (2021).
178. Wang S, Chen C, Guan M, Liu D, Wan X-F, Li L. Terminal Epitope-Dependent Branch Preference of Siglecs Toward N-Glycans. *Frontiers in Molecular Biosciences* **8**, (2021).
179. Angata T, Kerr SC, Greaves DR, Varki NM, Crocker PR, Varki A. Cloning and characterization of human Siglec-11. A recently evolved signaling molecule that can interact with SHP-1 and SHP-2 and is expressed by tissue macrophages, including brain microglia. *J Biol Chem* **277**, 24466-24474 (2002).
180. Hayakawa T, Angata T, Lewis AL, Mikkelsen TS, Varki NM, Varki A. A human-specific gene in microglia. *Science* **309**, 1693 (2005).

181. Hayakawa T, *et al.* Coevolution of Siglec-11 and Siglec-16 via gene conversion in primates. *BMC Evol Biol* **17**, 228 (2017).
182. Briard JG, Jiang H, Moremen KW, Macauley MS, Wu P. Cell-based glycan arrays for probing glycan-glycan binding protein interactions. *Nat Commun* **9**, 880 (2018).
183. Gutierrez-Martinez E, *et al.* Actin-regulated Siglec-1 nanoclustering influences HIV-1 capture and virus-containing compartment formation in dendritic cells. *Elife* **12**, (2023).
184. Izquierdo-Useros N, *et al.* Sialyllactose in viral membrane gangliosides is a novel molecular recognition pattern for mature dendritic cell capture of HIV-1. *PLoS Biol* **10**, e1001315 (2012).
185. Izquierdo-Useros N, Lorizate M, McLaren PJ, Telenti A, Krausslich HG, Martinez-Picado J. HIV-1 capture and transmission by dendritic cells: the role of viral glycolipids and the cellular receptor Siglec-1. *PLoS Pathog* **10**, e1004146 (2014).
186. Jalloh S, *et al.* CD169-mediated restrictive SARS-CoV-2 infection of macrophages induces pro-inflammatory responses. *PLoS Pathog* **18**, e1010479 (2022).
187. Siddiqui SS, *et al.* Siglecs in Brain Function and Neurological Disorders. *Cells* **8**, (2019).
188. Erickson JJ, *et al.* Pregnancy enables antibody protection against intracellular infection. *Nature* **606**, 769-775 (2022).
189. Awoyemi T, *et al.* Glycosylated Siglec-6 expression in syncytiotrophoblast-derived extracellular vesicles from preeclampsia placentas. *Biochem Biophys Res Commun* **533**, 838-844 (2020).
190. Gonzalez-Cuesta M, Ortiz Mellet C, Garcia Fernandez JM. Carbohydrate supramolecular chemistry: beyond the multivalent effect. *Chem Commun (Camb)* **56**, 5207-5222 (2020).
191. Brown GD, Willment JA, Whitehead L. C-type lectins in immunity and homeostasis. *Nat Rev Immunol* **18**, 374-389 (2018).
192. Chang YC, Nizet V. Siglecs at the Host-Pathogen Interface. *Adv Exp Med Biol* **1204**, 197-214 (2020).
193. Perez-Zsolt D, Martinez-Picado J, Izquierdo-Useros N. When Dendritic Cells Go Viral: The Role of Siglec-1 in Host Defense and Dissemination of Enveloped Viruses. *Viruses* **12**, (2019).
194. Stanczak MA, Laubli H. Siglec receptors as new immune checkpoints in cancer. *Mol Aspects Med*, 101112 (2022).

195. Smith BAH, Bertozzi CR. The clinical impact of glycobiology: targeting selectins, Siglecs and mammalian glycans. *Nat Rev Drug Discov* **20**, 217-243 (2021).
196. Adams OJ, Stanczak MA, von Gunten S, Laubli H. Targeting sialic acid-Siglec interactions to reverse immune suppression in cancer. *Glycobiology* **28**, 640-647 (2018).
197. Lubbers J, Rodriguez E, van Kooyk Y. Modulation of Immune Tolerance via Siglec-Sialic Acid Interactions. *Front Immunol* **9**, 2807 (2018).
198. Stamenkovic I, Sgroi D, Aruffo A, Sy MS, Anderson T. The B lymphocyte adhesion molecule CD22 interacts with leukocyte common antigen CD45RO on T cells and α 2-6 sialyltransferase, CD75, on B cells. *Cell* **66**, 1133-1144 (1991).
199. Sgroi D, Varki A, Braesch-Andersen S, Stamenkovic I. CD22, a B cell-specific immunoglobulin superfamily member, is a sialic acid-binding lectin. *Journal of Biological Chemistry* **268**, 7011-7018 (1993).
200. Powell LD, Sgroi D, Sjoberg ER, Stamenkovic I, Varki A. Natural ligands of the B cell adhesion molecule CD22 beta carry N-linked oligosaccharides with alpha-2,6-linked sialic acids that are required for recognition. *Journal of Biological Chemistry* **268**, 7019-7027 (1993).
201. Collins BE, *et al.* Constitutively unmasked CD22 on B cells of ST6Gal I knockout mice: novel sialoside probe for murine CD22. *Glycobiology* **12**, 563-571 (2002).
202. Kelm S, Gerlach J, Brossmer R, Danzer CP, Nitschke L. The ligand-binding domain of CD22 is needed for inhibition of the B cell receptor signal, as demonstrated by a novel human CD22-specific inhibitor compound. *J Exp Med* **195**, 1207-1213 (2002).
203. Jones C, Virji M, Crocker PR. Recognition of sialylated meningococcal lipopolysaccharide by siglecs expressed on myeloid cells leads to enhanced bacterial uptake. *Mol Microbiol* **49**, 1213-1225 (2003).
204. Avril T, North SJ, Haslam SM, Willison HJ, Crocker PR. Probing the cis interactions of the inhibitory receptor Siglec-7 with alpha2,8-disialylated ligands on natural killer cells and other leukocytes using glycan-specific antibodies and by analysis of alpha2,8-sialyltransferase gene expression. *J Leukoc Biol* **80**, 787-796 (2006).
205. Brinkman-Van der Linden EC, *et al.* Human-specific expression of Siglec-6 in the placenta. *Glycobiology* **17**, 922-931 (2007).
206. Santegoets KCM, *et al.* Expression profiling of immune inhibitory Siglecs and their ligands in patients with glioma. *Cancer Immunol Immunother* **68**, 937-949 (2019).
207. Gonzalez-Gil A, *et al.* Human brain sialoglycan ligand for CD33, a microglial inhibitory Siglec implicated in Alzheimer's disease. *J Biol Chem* **298**, 101960 (2022).

208. Wisnovsky S, *et al.* Genome-wide CRISPR screens reveal a specific ligand for the glycan-binding immune checkpoint receptor Siglec-7. *Proc Natl Acad Sci U S A* **118**, (2021).
209. Castro-Dopico T, Clatworthy MR. IgG and Fcγ Receptors in Intestinal Immunity and Inflammation. *Front Immunol* **10**, 805 (2019).
210. Gieseke F, *et al.* Siglec-7 tetramers characterize B-cell subpopulations and leukemic blasts. *Eur J Immunol* **42**, 2176-2186 (2012).
211. Vafa O, *et al.* An engineered Fc variant of an IgG eliminates all immune effector functions via structural perturbations. *Methods* **65**, 114-126 (2014).
212. Fishman JB, Berg EA. Protein A and Protein G Purification of Antibodies. *Cold Spring Harb Protoc* **2019**, (2019).
213. Schmidt TG, Skerra A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc* **2**, 1528-1535 (2007).
214. Wiberg FC, *et al.* Production of target-specific recombinant human polyclonal antibodies in mammalian cells. *Biotechnology and Bioengineering* **94**, 396-405 (2006).
215. Fus-Kujawa A, *et al.* An Overview of Methods and Tools for Transfection of Eukaryotic Cells in vitro. *Frontiers in Bioengineering and Biotechnology* **9**, (2021).
216. Hacker DL, Balasubramanian S. Recombinant protein production from stable mammalian cell lines and pools. *Current Opinion in Structural Biology* **38**, 129-136 (2016).
217. Barlowe CK, Miller EA. Secretory Protein Biogenesis and Traffic in the Early Secretory Pathway. *Genetics* **193**, 383-410 (2013).
218. Schmidt EN, Jung J, Macauley MS. Flow Cytometry-Based Detection of Siglec Ligands. *Methods Mol Biol* **2657**, 181-193 (2023).
219. Lenza MP, *et al.* Structural insights into Siglec-15 reveal glycosylation dependency for its interaction with T cells through integrin CD11b. *Nature Communications* **14**, (2023).
220. Park S, *et al.* Immunoengineering can overcome the glycocalyx armour of cancer cells. *Nature Materials* **23**, 429-438 (2024).
221. Creytens S, Pascha MN, Ballegeer M, Saelens X, de Haan CAM. Influenza Neuraminidase Characteristics and Potential as a Vaccine Target. *Frontiers in Immunology* **12**, (2021).
222. Karvelis T, Gasiunas G, Miksys A, Barrangou R, Horvath P, Siksnys V. crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biology* **10**, 841-851 (2013).

223. Teich K, *et al.* Cluster of differentiation 33 single nucleotide polymorphism rs12459419 is a predictive factor in patients with *nucleophosmin1*-mutated acute myeloid leukemia receiving gemtuzumab ozogamicin. *Haematologica* **106**, 2986-2989 (2021).
224. Eskandari-Sedighi G, Jung J, Macauley MS. CD33 isoforms in microglia and Alzheimer's disease: Friend and foe. *Mol Aspects Med* **90**, 101111 (2023).
225. Colonna M, Butovsky O. Microglia Function in the Central Nervous System During Health and Neurodegeneration. *Annual Review of Immunology* **35**, 441-468 (2017).
226. Gonzalez-Gil A, *et al.* Human brain sialoglycan ligand for CD33, a microglial inhibitory Siglec implicated in Alzheimer's disease. *Journal of Biological Chemistry* **298**, (2022).
227. Clark EA, Giltiay NV. CD22: A Regulator of Innate and Adaptive B Cell Responses and Autoimmunity. *Frontiers in Immunology* **9**, (2018).
228. Schmidt EN, Guo XY, Bui DT, Jung J, Klassen JS, Macauley MS. Dissecting the abilities of murine Siglecs to interact with gangliosides. *Journal of Biological Chemistry*, (2024).
229. Gabizon A, Shmeeda H, Barenholz Y. Pharmacokinetics of pegylated liposomal Doxorubicin: review of animal and human studies. *Clin Pharmacokinet* **42**, 419-436 (2003).
230. Xu W-J, Lin Y, Mi C-L, Pang J-Y, Wang T-Y. Progress in fed-batch culture for recombinant protein production in CHO cells. *Applied Microbiology and Biotechnology* **107**, 1063-1075 (2023).
231. Chi K, Xu H, Li H, Yang G, Zhou X, Gao X-D. Expression of a Siglec-Fc Protein and Its Characterization. *Biology* **12**, (2023).
232. Naito Y, *et al.* Germinal Center Marker GL7 Probes Activation-Dependent Repression of N-Glycolylneuraminic Acid, a Sialic Acid Species Involved in the Negative Modulation of B-Cell Activation. *Molecular and Cellular Biology* **27**, 3008-3022 (2023).
233. Hodgson K, *et al.* Sialic acid blockade inhibits the metastatic spread of prostate cancer to bone. *eBioMedicine* **104**, (2024).
234. Fong JJ, *et al.* Immunomodulatory activity of extracellular Hsp70 mediated via paired receptors Siglec-5 and Siglec-14. *The EMBO Journal* **34**, 2775-2788 (2015).
235. Collins BE, Blixt O, DeSieno AR, Bovin N, Marth JD, Paulson JC. Masking of CD22 by cis ligands does not prevent redistribution of CD22 to sites of cell contact. *Proceedings of the National Academy of Sciences* **101**, 6104-6109 (2004).

236. Wang Y, *et al.* Siglec-9+tumor-associated macrophages delineate an immunosuppressive subset with therapeutic vulnerability in patients with high-grade serous ovarian cancer. *Journal for ImmunoTherapy of Cancer* **11**, (2023).
237. Mei Y, *et al.* Siglec-9 acts as an immune-checkpoint molecule on macrophages in glioblastoma, restricting T-cell priming and immunotherapy response. *Nature Cancer* **4**, 1273-1291 (2023).
238. Sun L, *et al.* Installation of O-glycan sulfation capacities in human HEK293 cells for display of sulfated mucins. *Journal of Biological Chemistry* **298**, (2022).
239. Flynn RA, *et al.* Small RNAs are modified with N-glycans and displayed on the surface of living cells. *Cell* **184**, 3109-3124.e3122 (2021).
240. Lloyd KL, Kubes P. GPI-linked endothelial CD14 contributes to the detection of LPS. *American Journal of Physiology-Heart and Circulatory Physiology* **291**, H473-H481 (2006).
241. Vattepu R, Sneed SL, Anthony RM. Sialylation as an Important Regulator of Antibody Function. *Frontiers in Immunology* **13**, (2022).
242. Rapoport E, Mikhalyov I, Zhang J, Crocker P, Bovin N. Ganglioside Binding Pattern of CD33-related Siglecs. *Bioorganic & Medicinal Chemistry Letters* **13**, 675-678 (2003).
243. Bhattacharjee A, *et al.* Repression of phagocytosis by human CD33 is not conserved with mouse CD33. *Commun Biol* **2**, 450 (2019).
244. Movsisyan LD, Macauley MS. Structural advances of Siglecs: insight into synthetic glycan ligands for immunomodulation. *Org Biomol Chem* **18**, 5784-5797 (2020).
245. Puryear WB, Yu X, Ramirez NP, Reinhard BM, Gummuluru S. HIV-1 incorporation of host-cell-derived glycosphingolipid GM3 allows for capture by mature dendritic cells. *Proc Natl Acad Sci U S A* **109**, 7475-7480 (2012).
246. Perez-Zsolt D, *et al.* SARS-CoV-2 interaction with Siglec-1 mediates trans-infection by dendritic cells. *Cell Mol Immunol* **18**, 2676-2678 (2021).
247. Chang YC, Nizet V. The interplay between Siglecs and sialylated pathogens. *Glycobiology* **24**, 818-825 (2014).
248. Hallock KJ, Henzler Wildman K, Lee D-K, Ramamoorthy A. An Innovative Procedure Using a Sublimable Solid to Align Lipid Bilayers for Solid-State NMR Studies. *Biophysical Journal* **82**, 2499-2503 (2002).

249. Collins BE, *et al.* Binding specificities of the sialoadhesin family of I-type lectins. Sialic acid linkage and substructure requirements for binding of myelin-associated glycoprotein, Schwann cell myelin protein, and sialoadhesin. *J Biol Chem* **272**, 16889-16895 (1997).
250. Schnaar RL. Glycolipid-mediated cell-cell recognition in inflammation and nerve regeneration. *Arch Biochem Biophys* **426**, 163-172 (2004).
251. Yoshimura A, Hatanaka R, Tanaka H, Kitajima K, Sato C. The conserved arginine residue in all siglecs is essential for Siglec-7 binding to sialic acid. *Biochem Biophys Res Commun*, (2020).
252. Schnaar RL, *et al.* Myelin-associated Glycoprotein Binding to Gangliosides. Structural Specificity and Functional Implications. *Ann N Y Acad Sci* **845**, 92-105 (1998).
253. Liu Y, *et al.* Neoglycolipid probes prepared via oxime ligation for microarray analysis of oligosaccharide-protein interactions. *Chem Biol* **14**, 847-859 (2007).
254. Ito A, Handa K, Withers DA, Satoh M, Hakomori S-i. Binding specificity of siglec7 to disialogangliosides of renal cell carcinoma: possible role of disialogangliosides in tumor progression. *FEBS Letters* **504**, 82-86 (2001).
255. Murugesan G, *et al.* Siglec-15 recognition of sialoglycans on tumor cell lines can occur independently of sialyl Tn antigen expression. *Glycobiology* **31**, 44-54 (2021).
256. Saunderson SC, Dunn AC, Crocker PR, McLellan AD. CD169 mediates the capture of exosomes in spleen and lymph node. *Blood* **123**, 208-216 (2014).
257. Shenoy GN, *et al.* Sialic Acid-Dependent Inhibition of T Cells by Exosomal Ganglioside GD3 in Ovarian Tumor Microenvironments. *J Immunol* **201**, 3750-3758 (2018).
258. Pronker MF, *et al.* Structural basis of myelin-associated glycoprotein adhesion and signalling. *Nat Commun* **7**, 13584 (2016).
259. Nicoll G, Avril T, Lock K, Furukawa K, Bovin N, Crocker PR. Ganglioside GD3 Expression on Target Cells can Modulate NK Cell Cytotoxicity via Siglec-7-dependent and -Independent Mechanisms. *Eur J Immunol* **33**, 1642-1648 (2003).
260. Kolter T. Ganglioside Biochemistry *ISRN Biochem* **2012**, 506160 (2012).
261. Yesmin F, *et al.* Extracellular vesicles released from ganglioside GD2-expressing melanoma cells enhance the malignant properties of GD2-negative melanomas. *Sci Rep* **13**, 4987 (2023).
262. Otake AH, de Freitas Saito R, Duarte APM, Ramos AF, Chammas R. GD3 ganglioside-enriched extracellular vesicles stimulate melanocyte migration. *Biochim Biophys Acta Mol Cell Biol Lipids* **1864**, 422-432 (2019).

263. Sandhoff R, Schulze H, Sandhoff K. Ganglioside Metabolism in Health and Disease. *Prog Mol Biol Transl Sci* **156**, 1-62 (2018).
264. Picache JA, Zheng W, Chen CZ. Therapeutic Strategies For Tay-Sachs Disease. *Front Pharmacol* **13**, 906647 (2022).
265. Perez-Zsolt D, *et al.* SARS-CoV-2 interaction with Siglec-1 mediates trans-infection by dendritic cells. *Cell Mol Immunol* **18**, 2676-2678 (2021).
266. Ang CW, Noordzij PG, de Klerk MA, Endtz HP, van Doorn PA, Laman JD. Ganglioside mimicry of *Campylobacter jejuni* lipopolysaccharides determines antiganglioside specificity in rabbits. *Infect Immun* **70**, 5081-5085 (2002).
267. Latov N. *Campylobacter jejuni* Infection, Anti-Ganglioside Antibodies, and Neuropathy. *Microorganisms* **10**, (2022).
268. Perlman RL. Mouse models of human disease: An evolutionary perspective. *Evol Med Public Health* **2016**, 170-176 (2016).
269. Cheon DJ, Orsulic S. Mouse models of cancer. *Annu Rev Pathol* **6**, 95-119 (2011).
270. Smeekens JM, Kulis MD. Mouse Models of Food Allergy in the Pursuit of Novel Treatment Modalities. *Front Allergy* **2**, 810067 (2021).
271. von Gunten S, Bochner BS. Basic and clinical immunology of Siglecs. *Ann N Y Acad Sci* **1143**, 61-82 (2008).
272. McCord KA, Macauley MS. Transgenic mouse models to study the physiological and pathophysiological roles of human Siglecs. *Biochem Soc Trans* **50**, 935-950 (2022).
273. Kawasaki N, *et al.* Targeted delivery of mycobacterial antigens to human dendritic cells via Siglec-7 induces robust T cell activation. *J Immunol* **193**, 1560-1566 (2014).
274. Collins BE, Ito H, Sawada N, Ishida H, Kiso M, Schnaar RL. Enhanced binding of the neural siglecs, myelin-associated glycoprotein and Schwann cell myelin protein, to Chol-1 (alpha-series) gangliosides and novel sulfated Chol-1 analogs. *J Biol Chem* **274**, 37637-37643 (1999).
275. Ishida H, Kiso M. Synthetic Study on Neural Siglecs Ligands. Systematic Synthesis of .ALPHA.-Series Polysialogangliosides and Their Analogues. *Journal of Synthetic Organic Chemistry, Japan* **58**, 1108-1113 (2000).
276. Affandi AJ, *et al.* Selective tumor antigen vaccine delivery to human CD169(+) antigen-presenting cells using ganglioside-liposomes. *Proc Natl Acad Sci U S A*, (2020).

277. Li SD, Huang L. Stealth nanoparticles: high density but sheddable PEG is a key for tumor targeting. *J Control Release* **145**, 178-181 (2010).
278. Han L, *et al.* How Choice of Model Membrane Affects Protein-Glycosphingolipid Interactions: Insights from Native Mass Spectrometry. *Anal Chem* **94**, 16042-16049 (2022).
279. Schuster MC, Mortell KH, Hegeman AD, Kiessling LL. Neoglycopolymers produced by aqueous ring-opening metathesis polymerization: decreasing saccharide density increases activity. *Journal of Molecular Catalysis A: Chemical* **116**, 209-216 (1997).
280. Nsairat H, *et al.* Liposome bilayer stability: emphasis on cholesterol and its alternatives. *Journal of Liposome Research* **34**, 178-202 (2023).
281. Kim JY, Kim YG, Lee GM. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl Microbiol Biotechnol* **93**, 917-930 (2012).
282. Rinaldi S, *et al.* Analysis of lectin binding to glycolipid complexes using combinatorial glycoarrays. *Glycobiology* **19**, 789-796 (2009).
283. Patel DS, *et al.* Influence of Ganglioside GM1 Concentration on Lipid Clustering and Membrane Properties and Curvature. *Biophys J* **111**, 1987-1999 (2016).
284. Gu RX, Ingolfsson HI, de Vries AH, Marrink SJ, Tieleman DP. Ganglioside-Lipid and Ganglioside-Protein Interactions Revealed by Coarse-Grained and Atomistic Molecular Dynamics Simulations. *J Phys Chem B* **121**, 3262-3275 (2017).
285. Liu Y, Barnoud J, Marrink SJ. Gangliosides Destabilize Lipid Phase Separation in Multicomponent Membranes. *Biophys J* **117**, 1215-1223 (2019).
286. DeMarco ML, Woods RJ. Atomic-resolution conformational analysis of the GM3 ganglioside in a lipid bilayer and its implications for ganglioside-protein recognition at membrane surfaces. *Glycobiology* **19**, 344-355 (2009).
287. Chang CE, Chen W, Gilson MK. Ligand configurational entropy and protein binding. *Proc Natl Acad Sci U S A* **104**, 1534-1539 (2007).
288. Huang ML, Purcell SC, Verespy S, III, Wang Y, Godula K. Glycocalyx scaffolding with synthetic nanoscale glycomaterials. *Biomater Sci* **5**, 1537-1540 (2017).
289. He X, Guan F, Lei L. Structure and function of glycosphingolipids on small extracellular vesicles. *Glycoconj J* **39**, 197-205 (2022).
290. Zaccai NR, *et al.* Structure-guided design of sialic acid-based Siglec inhibitors and crystallographic analysis in complex with sialoadhesin. *Structure* **11**, 557-567 (2003).

291. May AP, Robinson RC, Vinson M, Crocker PR, Jones EY. Crystal structure of the N-terminal domain of sialoadhesin in complex with 3' sialyllactose at 1.85 Å resolution. *Mol Cell* **1**, 719-728 (1998).
292. Schnaar RL. Gangliosides as Siglec ligands. *Glycoconj J* **40**, 159-167 (2023).
293. Altman MO, Gagneux P. Absence of Neu5Gc and Presence of Anti-Neu5Gc Antibodies in Humans-An Evolutionary Perspective. *Front Immunol* **10**, 789 (2019).
294. Huang F, *et al.* Analysis and Comparison of Mouse and Human Brain Gangliosides via Two-Stage Matching of MS/MS Spectra. *ACS Omega* **7**, 6403-6411 (2022).
295. Di Pardo A, *et al.* Ganglioside GM1 induces phosphorylation of mutant huntingtin and restores normal motor behavior in Huntington disease mice. *Proc Natl Acad Sci U S A* **109**, 3528-3533 (2012).
296. Han L, Kitov PI, Li J, Kitova EN, Klassen JS. Probing Heteromultivalent Protein-Glycosphingolipid Interactions using Native Mass Spectrometry and Nanodiscs. *Anal Chem* **92**, 3923-3931 (2020).
297. Takei Y, Sasaki S, Fujiwara T, Takahashi E, Muto T, Nakamura Y. Molecular cloning of a novel gene similar to myeloid antigen CD33 and its specific expression in placenta. *Cytogenet Cell Genet* **78**, 295-300 (1997).
298. Smiljkovic D, *et al.* Expression and regulation of Siglec-6 (CD327) on human mast cells and basophils. *J Allergy Clin Immunol* **151**, 202-211 (2023).
299. Jellusova J, Nitschke L. Regulation of B cell functions by the sialic acid-binding receptors siglec-G and CD22. *Front Immunol* **2**, 96 (2011).
300. Jumper J, *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
301. Varadi M, *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439-D444 (2022).
302. Bertoline LMF, Lima AN, Krieger JE, Teixeira SK. Before and after AlphaFold2: An overview of protein structure prediction. *Front Bioinform* **3**, 1120370 (2023).
303. Inoue T, Kurosaki T. Memory B cells. *Nature Reviews Immunology* **24**, 5-17 (2023).
304. Sanz I, *et al.* Challenges and Opportunities for Consistent Classification of Human B Cell and Plasma Cell Populations. *Frontiers in Immunology* **10**, (2019).

305. Jetani H, *et al.* Siglec-6 is a novel target for CAR T-cell therapy in acute myeloid leukemia. *Blood* **138**, 1830-1842 (2021).
306. Krystel-Whittemore M, Dileepan KN, Wood JG. Mast Cell: A Multi-Functional Master Cell. *Frontiers in Immunology* **6**, (2016).
307. Katsoulis-Dimitriou K, Kotrba J, Voss M, Dudeck J, Dudeck A. Mast Cell Functions Linking Innate Sensing to Adaptive Immunity. *Cells* **9**, (2020).
308. Xu Y, Chen G. Mast Cell and Autoimmune Diseases. *Mediators of Inflammation* **2015**, 1-8 (2015).
309. O’Sullivan JA, Youngblood BA, Schleimer RP, Bochner BS. Siglecs as potential targets of therapy in human mast cell- and/or eosinophil-associated diseases. *Seminars in Immunology* **69**, (2023).
310. Rumer KK, Uyenishi J, Hoffman MC, Fisher BM, Winn VD. Siglec-6 Expression Is Increased in Placentas From Pregnancies Complicated by Preterm Preeclampsia. *Reproductive Sciences* **20**, 646-653 (2013).
311. Khan B, Allah Yar R, Khakwani Ak, Karim S, Arslan Ali H. Preeclampsia Incidence and Its Maternal and Neonatal Outcomes With Associated Risk Factors. *Cureus*, (2022).
312. Zeng FY, Gabius HJ. Sialic acid-binding proteins: characterization, biological function and application. *Z Naturforsch C J Biosci* **47**, 641-653 (1992).
313. Kraft CA, Garrido JL, Leiva-Vega L, Romero G. Quantitative analysis of protein-lipid interactions using tryptophan fluorescence. *Sci Signal* **2**, p14 (2009).
314. Kirshenbaum AS, Yin Y, Sundstrom JB, Bandara G, Metcalfe DD. Description and Characterization of a Novel Human Mast Cell Line for Scientific Study. *Int J Mol Sci* **20**, (2019).
315. Hopper TR, *et al.* Control of Donor–Acceptor Photophysics through Structural Modification of a “Twisting” Push–Pull Molecule. *Chemistry of Materials* **31**, 6860-6869 (2019).
316. Hu Y-B, Dammer EB, Ren R-J, Wang G. The endosomal-lysosomal system: from acidification and cargo sorting to neurodegeneration. *Translational Neurodegeneration* **4**, (2015).
317. Williams C, *et al.* Glycosylation of extracellular vesicles: current knowledge, tools and clinical perspectives. *J Extracell Vesicles* **7**, 1442985 (2018).

318. Huntley JJA, Fast W, Benkovic SJ, Wright PE, Dyson HJ. Role of a solvent-exposed tryptophan in the recognition and binding of antibiotic substrates for a metallo- β -lactamase. *Protein Science* **12**, 1368-1375 (2009).
319. Gude NM, Roberts CT, Kalionis B, King RG. Growth and function of the normal human placenta. *Thromb Res* **114**, 397-407 (2004).
320. Rizzuto G, Erlebacher A. Trophoblast antigens, fetal blood cell antigens, and the paradox of fetomaternal tolerance. *J Exp Med* **219**, (2022).
321. Shaha S, *et al.* Human placenta and trophoblasts simultaneously express three isoforms of atypical protein kinase-c. *Placenta* **119**, 39-43 (2022).
322. Viveiros A, *et al.* In-Cell Labeling Coupled to Direct Analysis of Extracellular Vesicles in the Conditioned Medium to Study Extracellular Vesicles Secretion with Minimum Sample Processing and Particle Loss. *Cells* **11**, 351 (2022).
323. Lee KJ, *et al.* Modulation of nonspecific binding in ultrafiltration protein binding studies. *Pharmaceutical research* **20**, 1015-1021 (2003).
324. Pilobello KT, Krishnamoorthy L, Slawek D, Mahal LK. Development of a lectin microarray for the rapid analysis of protein glycopatterns. *Chembiochem* **6**, 985-989 (2005).
325. Hu S, Wong DT. Lectin microarray. *Proteomics Clin Appl* **3**, 148-154 (2009).
326. Galleguillos D, *et al.* Anti-inflammatory role of GM1 and other gangliosides on microglia. *Journal of Neuroinflammation* **19**, (2022).
327. Hemberger M, Hanna CW, Dean W. Mechanisms of early placental development in mouse and humans. *Nature Reviews Genetics* **21**, 27-43 (2019).
328. Costa J, *et al.* The Role of the 3Rs for Understanding and Modeling the Human Placenta. *Journal of Clinical Medicine* **10**, (2021).
329. Vyas AA, *et al.* Gangliosides are functional nerve cell ligands for myelin-associated glycoprotein (MAG), an inhibitor of nerve regeneration. *Proceedings of the National Academy of Sciences* **99**, 8412-8417 (2002).
330. Dusoswa SA, *et al.* Glycan modification of glioblastoma-derived extracellular vesicles enhances receptor-mediated targeting of dendritic cells. *J Extracell Vesicles* **8**, 1648995 (2019).
331. Hosoguchi K, *et al.* An efficient approach to the discovery of potent inhibitors against glycosyltransferases. *J Med Chem* **53**, 5607-5619 (2010).

332. Dziadek S, Jacques S, Bundle DR. A novel linker methodology for the synthesis of tailored conjugate vaccines composed of complex carbohydrate antigens and specific TH-cell peptide epitopes. *Chemistry* **14**, 5908-5917 (2008).
333. Han L, *et al.* Neoglycolipids as Glycosphingolipid Surrogates for Protein Binding Studies Using Nanodiscs and Native Mass Spectrometry. *Anal Chem* **92**, 14189-14196 (2020).
334. Johansson SM, Arnberg N, Elofsson M, Wadell G, Kihlberg J. Multivalent HSA conjugates of 3'-sialyllactose are potent inhibitors of adenoviral cell attachment and infection. *Chembiochem* **6**, 358-364 (2005).
335. Blixt O, Collins BE, van den Nieuwenhof IM, Crocker PR, Paulson JC. Sialoside specificity of the siglec family assessed using novel multivalent probes: identification of potent inhibitors of myelin-associated glycoprotein. *J Biol Chem* **278**, 31007-31019 (2003).
336. Yudina ON, Sherman AA, Nifantiev NE. Synthesis of propyl and 2-aminoethyl glycosides of α -d-galactosyl-(1 \rightarrow 3')- β -lactoside. *Carbohydrate Research* **332**, 363-371 (2001).
337. Vasiliu D, *et al.* Large-scale chemoenzymatic synthesis of blood group and tumor-associated poly-N-acetyllactosamine antigens. *Carbohydr Res* **341**, 1447-1457 (2006).

Appendices

Appendix A2:

Table A2. 1: Markers used to define of immune cells from human spleen samples.

Antibody	Supplier	Cat. No.	Label	Clone	Isotype	Dilution (V/V)
anti-CD19	Biologend	982406	FITC	HIB19	Mouse IgG1, κ	1/250
anti-CD3	Biologend	317323	BV650	OKT3	Mouse IgG2a, κ	1/250
HLA-DR	Biologend	307617	APC-Cy7	L243	Mouse IgG2a, κ	1/250
anti-CD15	Biologend	323027	BV510	W6D3	Mouse IgG1, κ	1/250
anti-CD56	Biologend	362551	BV421	5.1H11	Mouse IgG1, κ	1/250
anti-CD4	Biologend	300557	BV711	RPA-T4	Mouse IgG1, κ	1/250
anti-CD8	Biologend	300913	PE-Cy7	HIT8a	Mouse IgG1, κ	1/250
anti-CD14	Biologend	301833	BV605	M5E2	Mouse IgG2a, κ	1/250
anti-CD16	Biologend	302045	BV786	3G8	Mouse IgG1, κ	1/250
PI			PE-Texas Red	NA	NA	1/1000

Table A2.2: Anti-Siglec antibodies used in this study.

Antibody	Supplier	Cat. No.	Label	Clone	Isotype	Dilution
anti-human CD169	Biolegend	346003	PE	7-239	Mouse IgG1, κ	1/250 (V/V)
anti-human CD22	Biolegend	302406	PE	HIB22	Mouse IgG1, κ	1/250 (V/V)
anti-human CD33	Biolegend	983904	PE	WM53	Mouse IgG1, κ	1/250 (V/V)
anti-MAG	Santa Cruz	sc-166849	PE	A-11	Mouse IgG2a, κ	1/250 (V/V)
anti-human Siglec-5	Biolegend	452003	PE	1A5	Mouse IgG1, κ	1/250 (V/V)
anti-Siglec-6	R&D Systems	FAB2859T	PE	767329	Mouse IgG2A	1/250 (V/V)
anti-human CD328	Biolegend	339203	PE	6-434	Mouse IgG1, κ	1/250 (V/V)
anti-human Siglec-8	Biolegend	347103	PE	7C9	Mouse IgG1, κ	1/250 (V/V)
anti-human Siglec-9	Biolegend	351503	PE	K8	Mouse IgG1, κ	1/250 (V/V)
anti-human Siglec-10	Biolegend	347603	PE	5G6	Mouse IgG1, κ	1/250 (V/V)
anti-Siglec-11	Biolegend	681702	Unlabeled	4C4	Mouse IgG2b	1/250 (V/V)
anti-Siglec-15	Non-commercial		Unlabeled	-	Mouse IgG1	1/250 (V/V)
anti-mouse IgG1	Biolegend	406607	PE	RMG1-1	Rat IgG	1/250 (V/V)
anti-mouse IgG2b	Biolegend	406707	PE	RMG2b-1	Rat IgG, κ	1/250 (V/V)
anti-mouse IgG	Thermofischer	A-21208	AF594	Polyclonal	Donkey, IgG	1/250 (V/V)

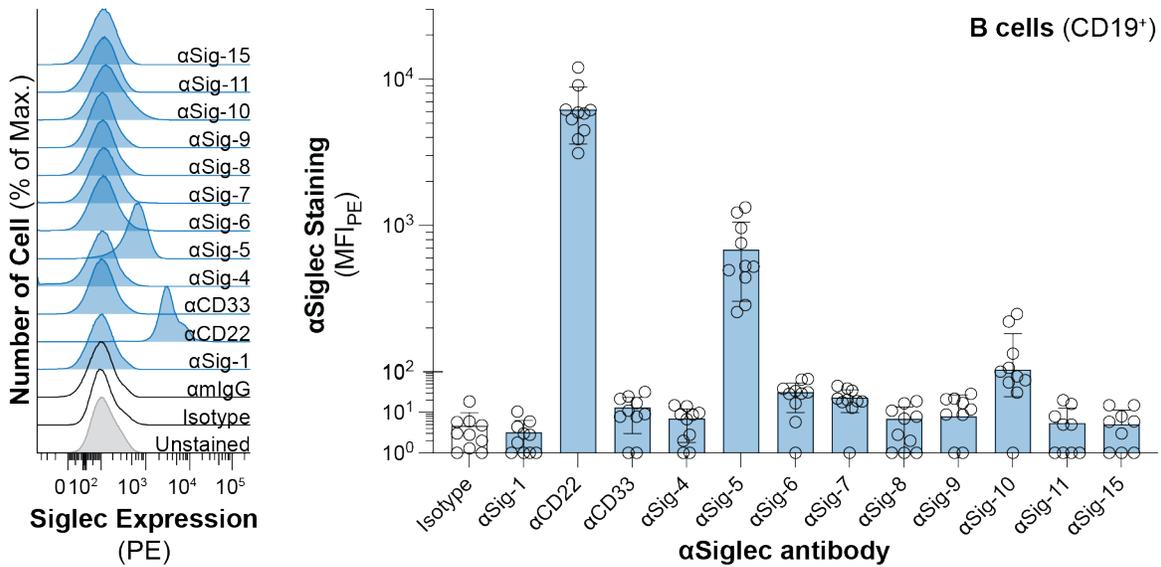


Figure A2.1: Siglec expression by B cells (CD19⁺) isolated from human spleen. Each datum represents a biological replicate.

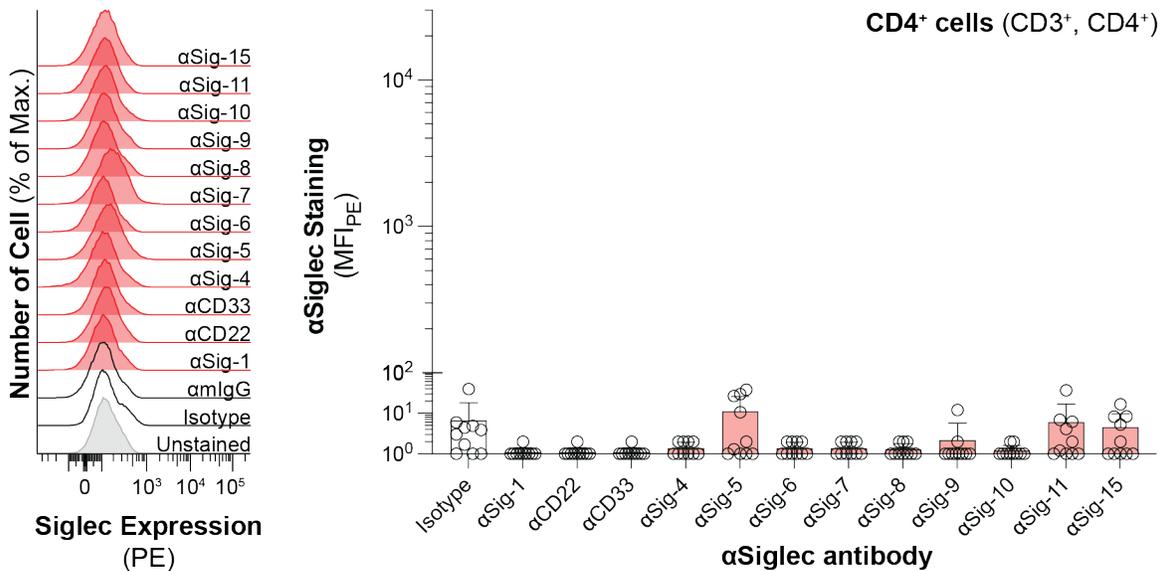


Figure A2.2: Siglec expression by CD8⁺ T cells (CD3⁺, CD4⁺) isolated from human spleen. Each datum represents a biological replicate.

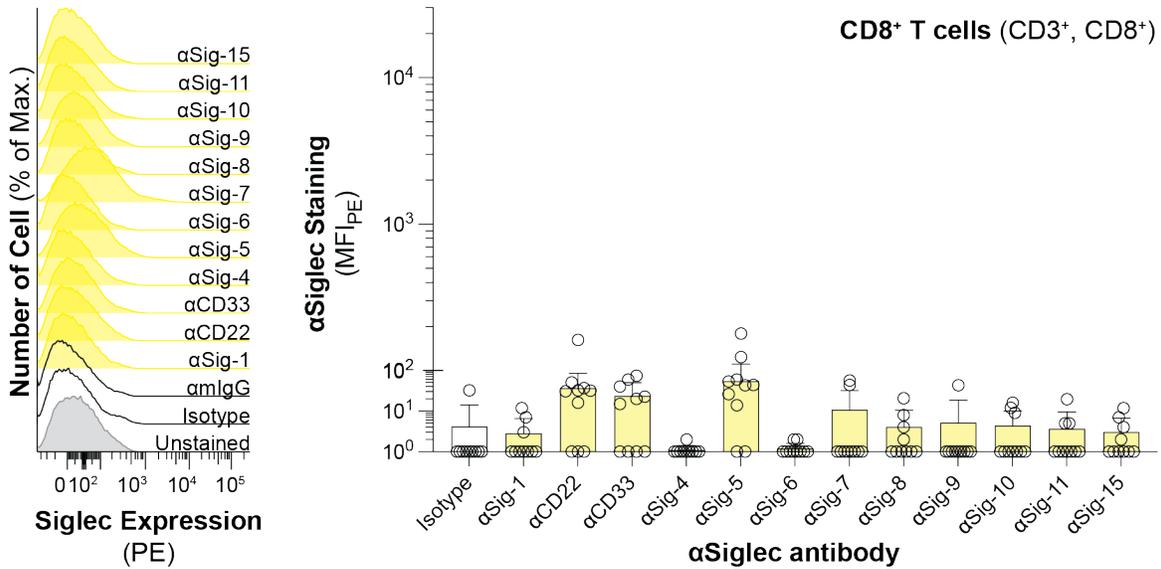


Figure A2.3: Siglec expression by CD8⁺ T cells (CD3⁺, CD8⁺) isolated from human spleen. Each datum represents a biological replicate.

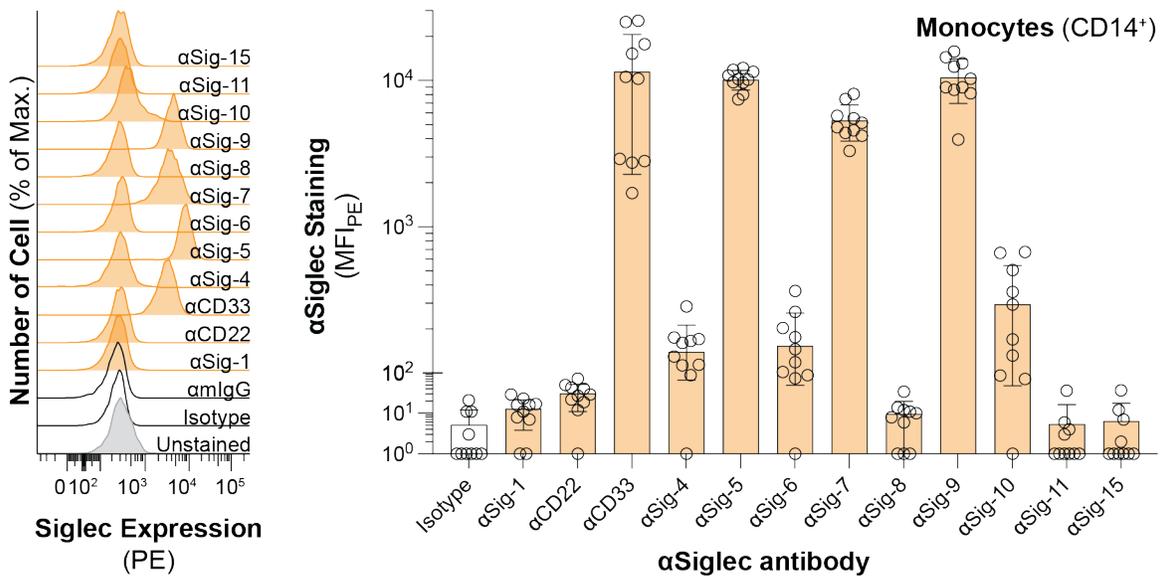


Figure A2.4: Siglec expression by monocytes (CD14⁺) isolated from human spleen. Each datum represents a biological replicate.

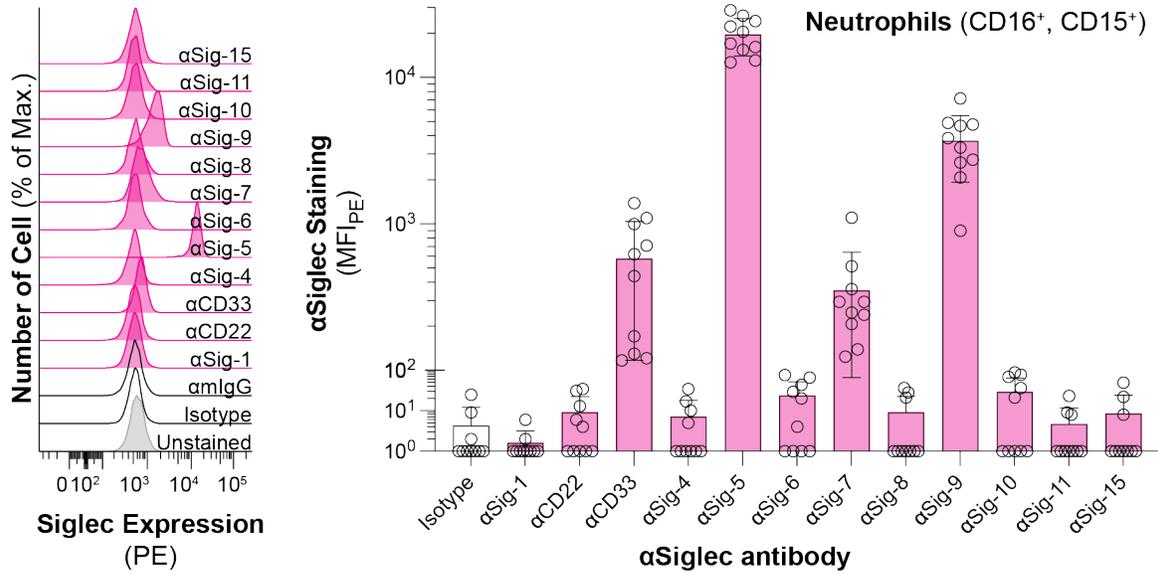


Figure A2.5: Siglec expression by neutrophils (CD15⁺, CD16⁺) isolated from human spleen. Each datum represents a biological replicate.

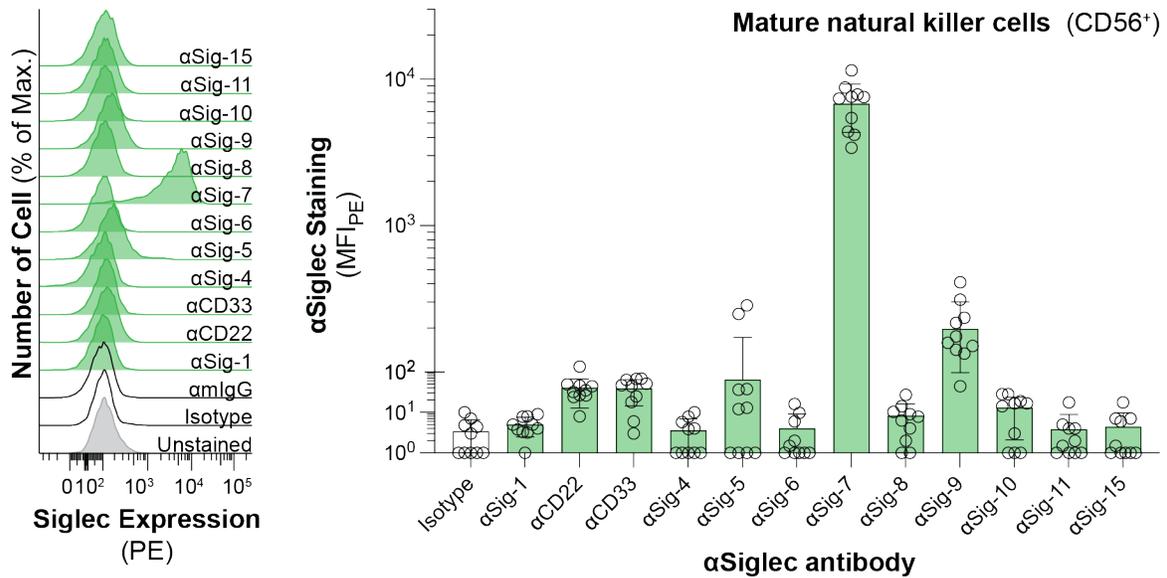


Figure A2.6: Siglec expression by mature NK cells (CD56⁺) isolated from human spleen. Each datum represents a biological replicate.

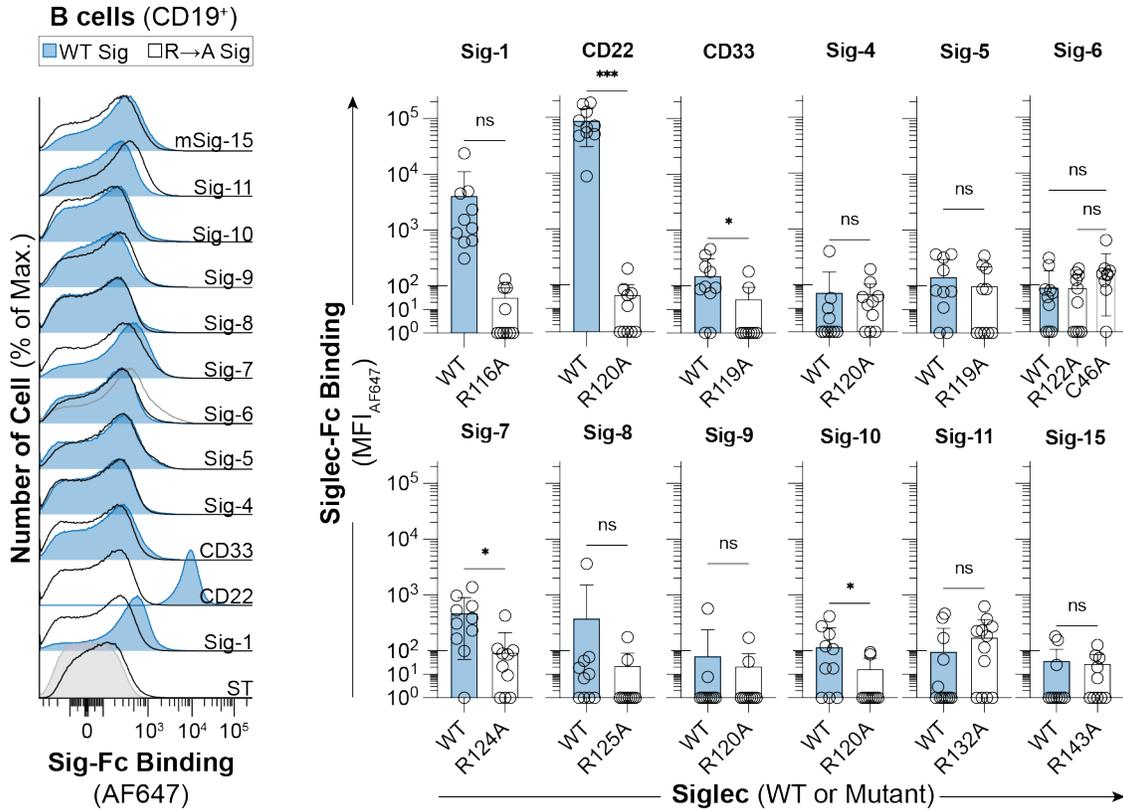


Figure A2.7: Siglec-Fc binding to B cells (CD19⁺) isolated from human spleen. Each datum represents a biological replicate.

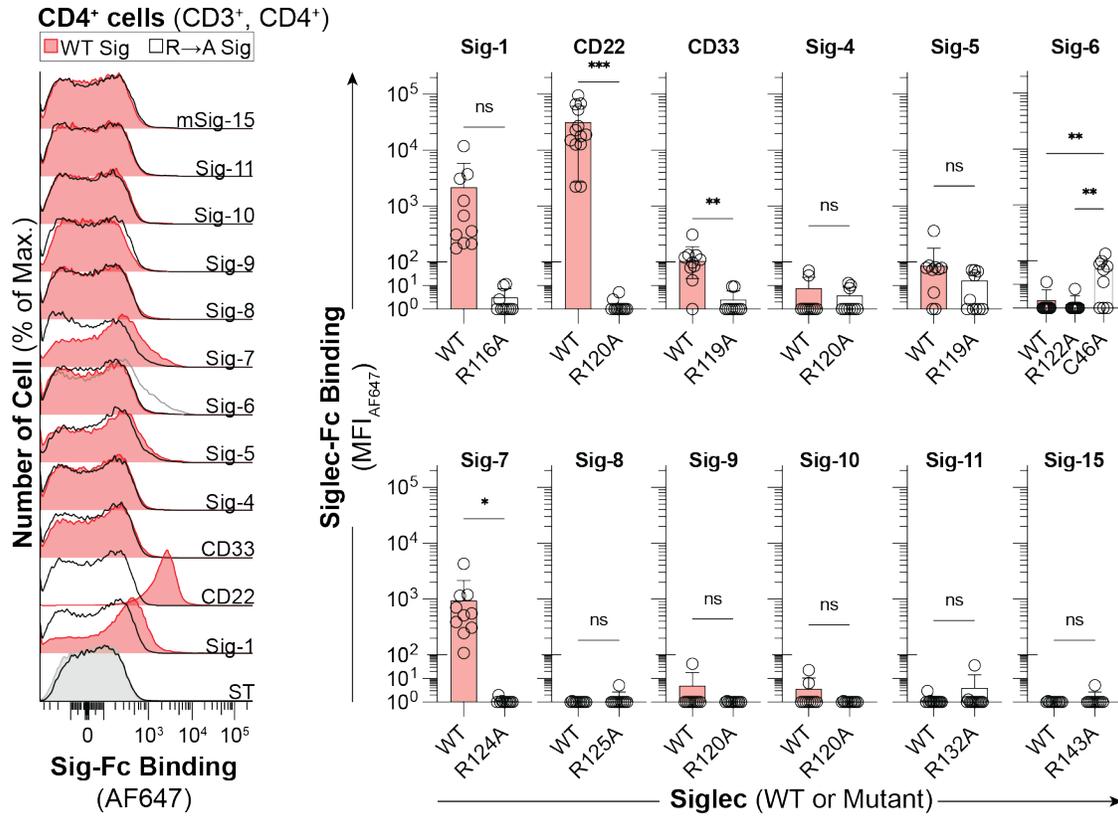


Figure A2.8: Siglec-Fc binding to CD8⁺ T cells (CD3⁺, CD4⁺) isolated from human spleen. Each datum represents a biological replicate.

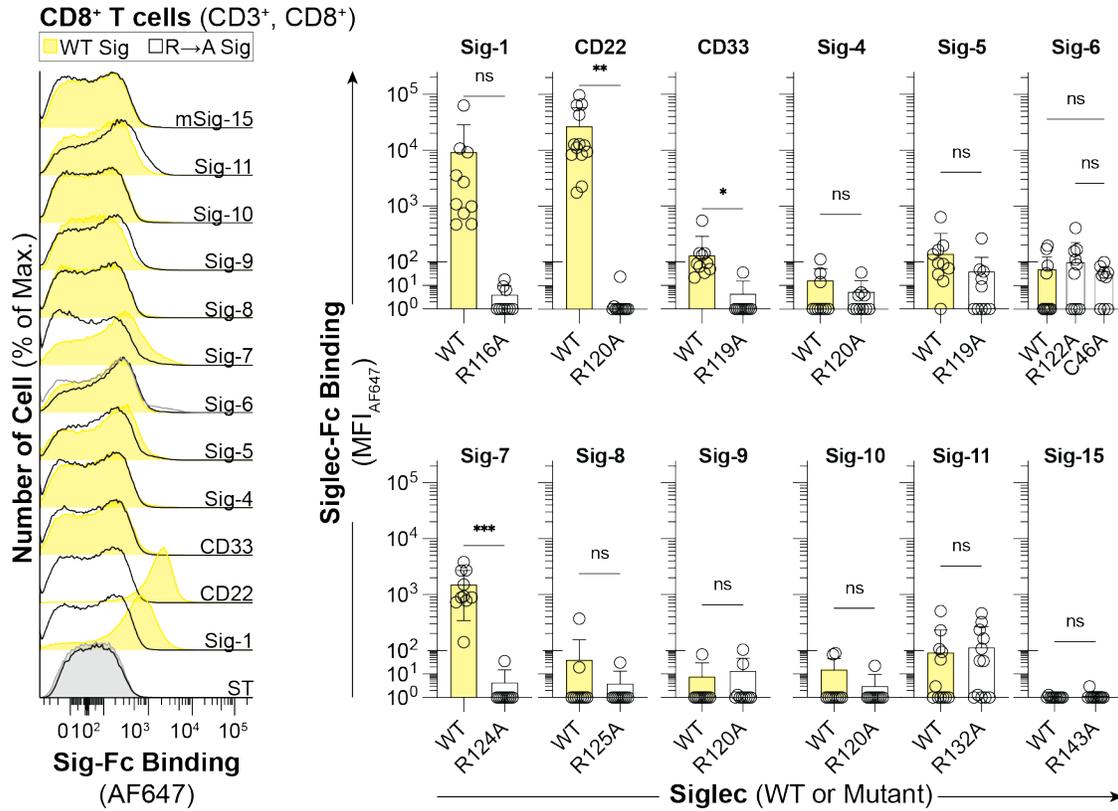


Figure A2.9: Siglec-Fc binding to CD8⁺ T cells (CD3⁺, CD8⁺) isolated from human spleen. Each datum represents a biological replicate.

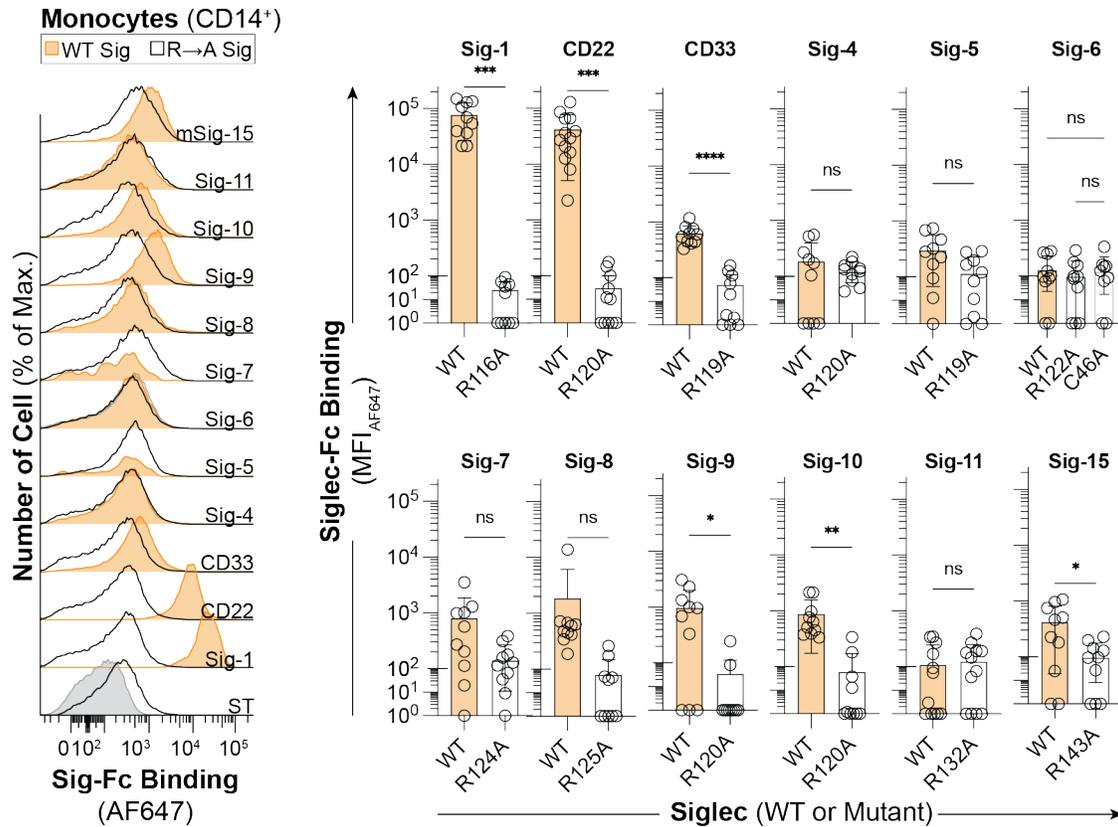


Figure A2.10: Siglec-Fc binding to monocytes (CD14⁺) isolated from human spleen. Each datum represents a biological replicate.

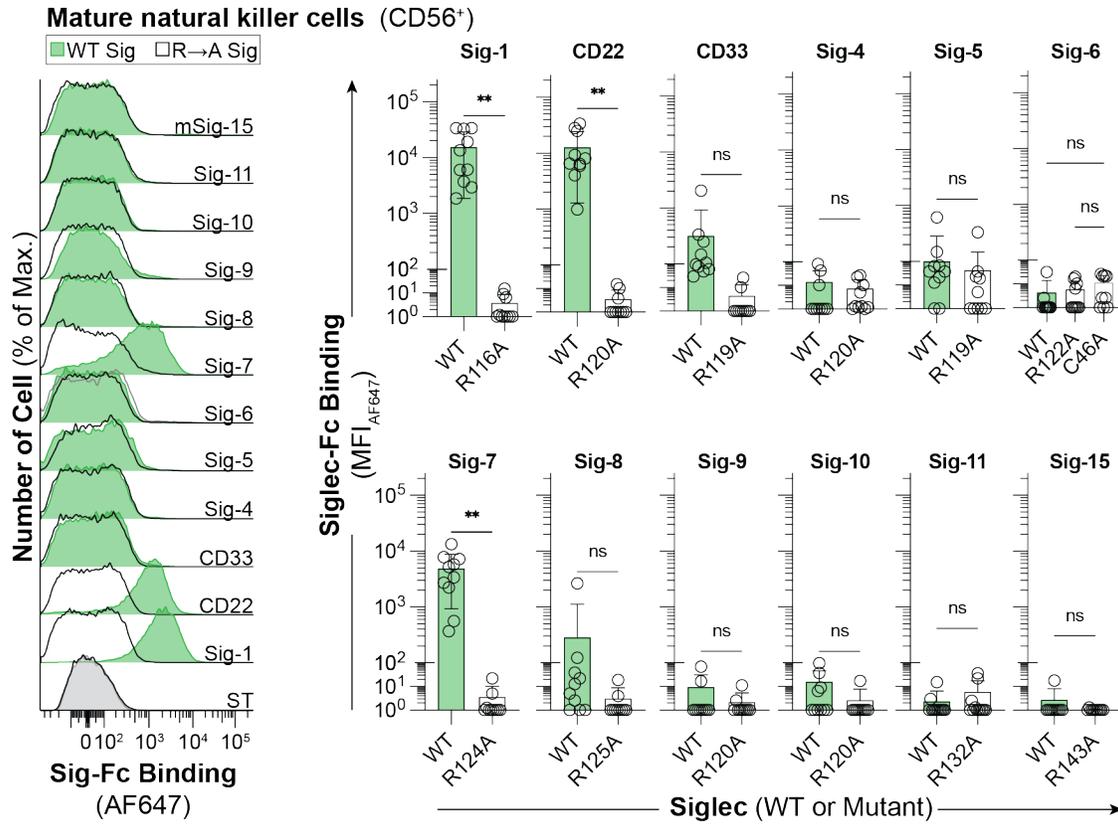


Figure A2.12: Siglec-Fc binding to mature NK cells (CD56⁺) isolated from human spleen. Each datum represents a biological replicate.

Appendix A3:

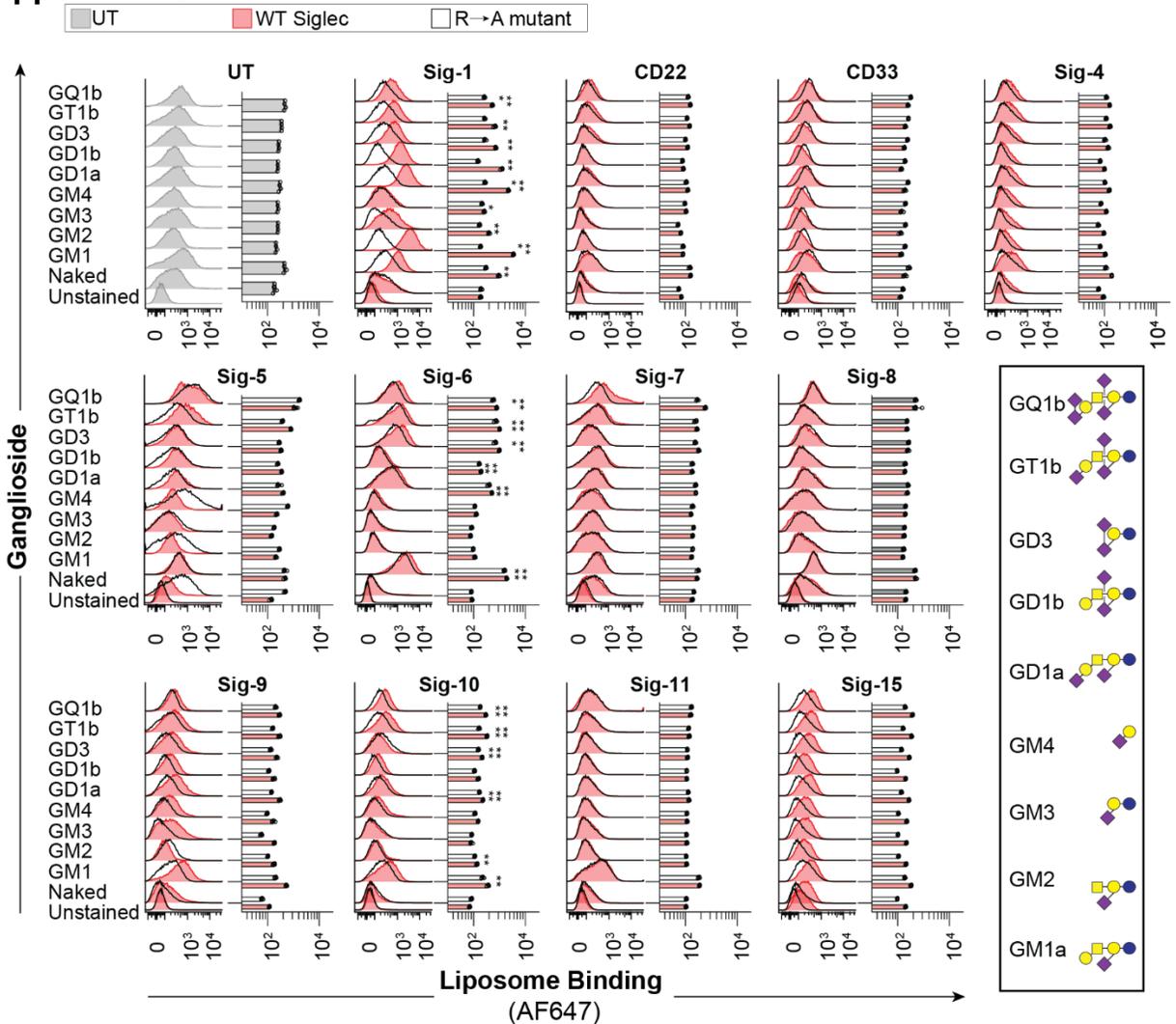


Figure A3.1: Interrogation of the human Siglec family against nine commercially available gangliosides using our optimized liposome formulation. UT-grey, WT (red), Mutant (black). Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from at least three technical replicates ($4 \geq n \geq 3$). A Brown-Forsythe and Welch one-way ANOVA was used to determine if the binding of liposome formulated with a ganglioside was significantly higher than a naked liposome to CHO cells expressing WT Siglec. Not Significant (NS); $P > 0.5$; * = $0.05 > P \geq 0.01$; ** = $0.01 > P \geq 0.001$; *** = $0.001 > P \geq 0.0001$; **** = $P < 0.0001$.

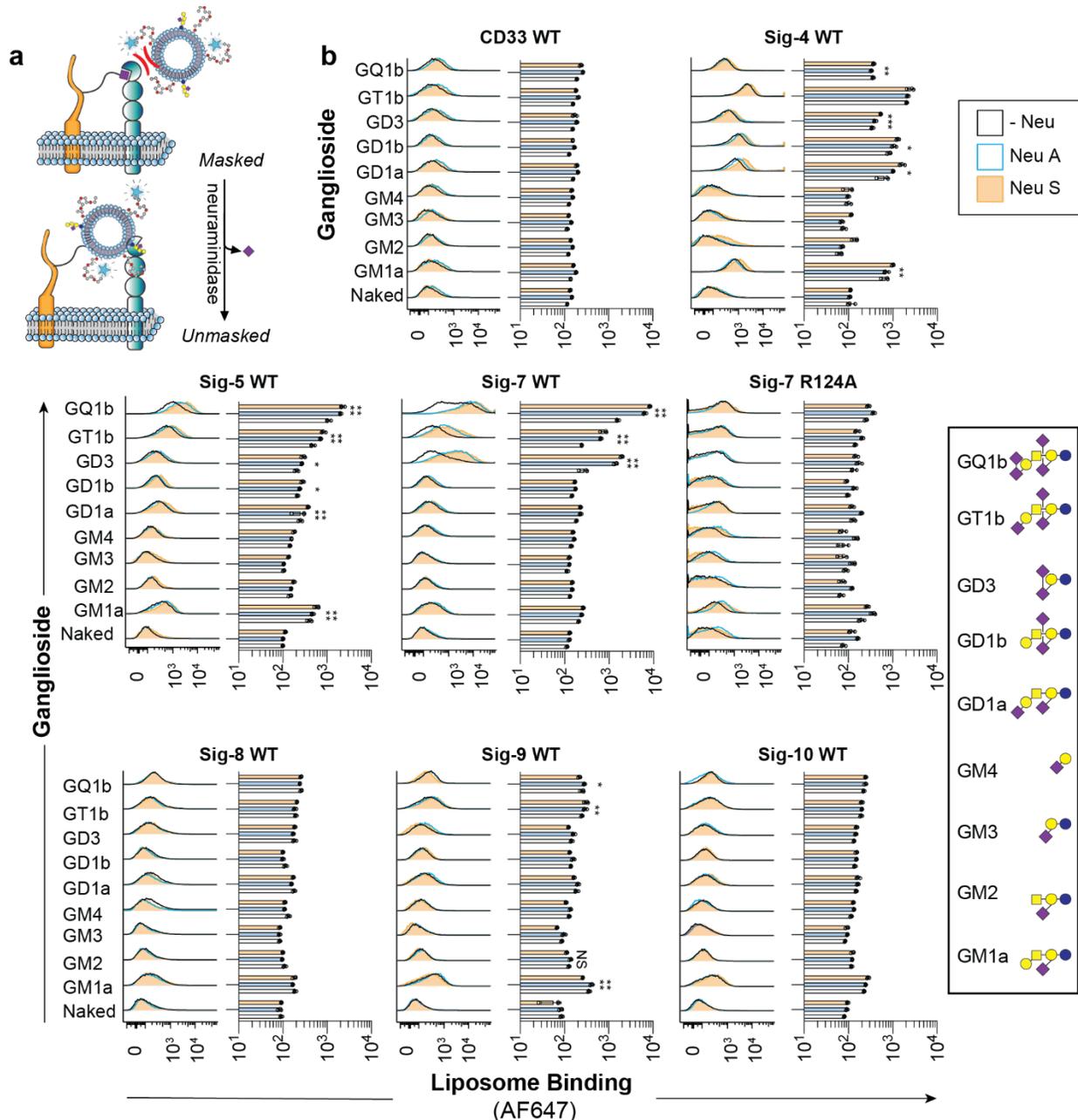


Figure A3.2: Binding of ganglioside bearing liposomes to Siglec expressing CHO cells after treatment with neuraminidase A and S. **a**, Schematic representation of the effect of neuraminidase treatment of cells on liposome binding. **b**, Binding of liposome formulated with the nine commercially available gangliosides in our optimized liposome formulation to CHO cells expressing select human Siglecs after treatment with neuraminidase A (blue) and neuraminidase S (orange). Data is presented with a representative flow cytometry histogram and was quantified as the mean \pm one standard deviation of the median fluorescent intensity (MFI) from at least three technical replicates ($4 \geq n \geq 3$). A Brown-Forsythe and Welch one-way ANOVA was used to determine if the binding of liposome formulated with a ganglioside was significantly higher than a naked liposome after treatment with neuraminidase S. Not Significant (NS); $P > 0.5$; * = $0.05 > P \geq 0.01$; ** = $0.01 > P \geq 0.001$; *** = $0.001 > P \geq 0.0001$; **** = $P < 0.0001$.

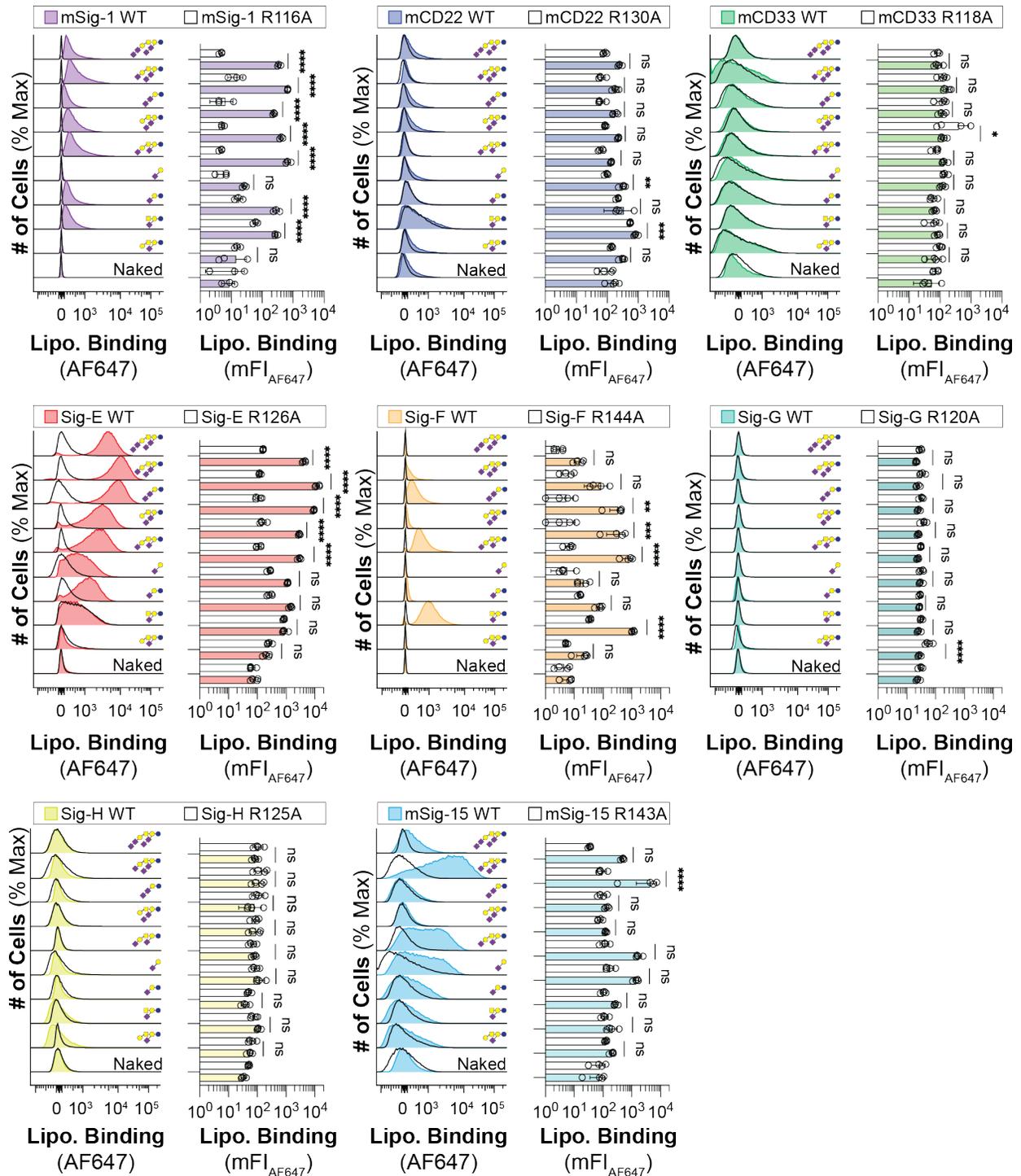


Figure A3.3: Binding of ganglioside liposomes to the murine Siglec family in the bead assay. Data is presented with a representative flow cytometry histogram as well as a summary chart. The data in the summary charts is presented as the mean of the mean fluorescence intensity from three technical replicates. A one-way ANOVA was for statistical analysis. Not Significant (NS); $P > 0.5$; $** = 0.01 > P \geq 0.001$; $*** = 0.001 > P \geq 0.0001$; $**** = P < 0.0001$.

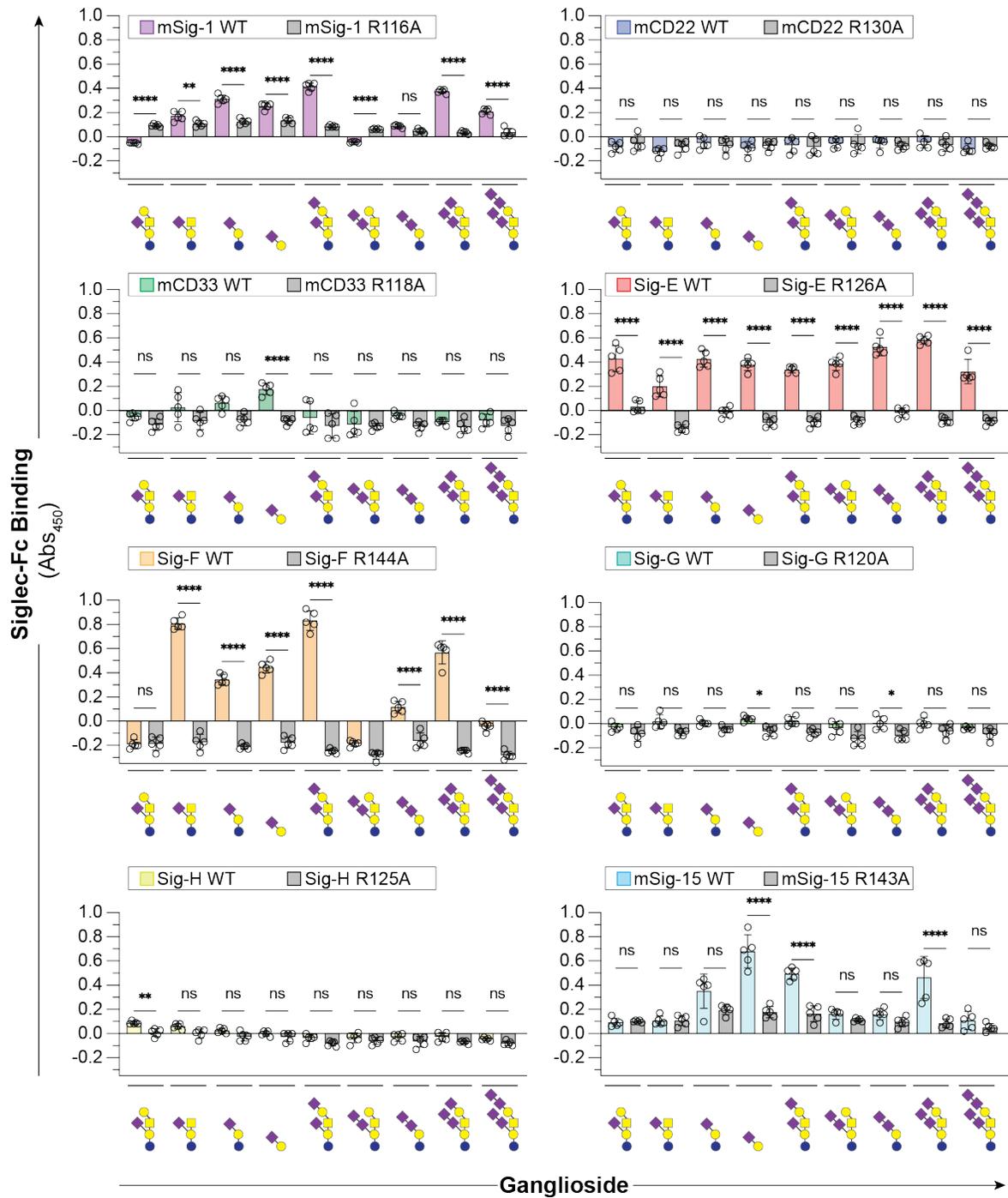


Figure A3.4: Binding of the murine Siglecs to gangliosides in an ELISA. The data presented as the mean of the absorbance at 450 nm from at least four technical replicates. A one-way ANOVA was for statistical analysis. Not Significant (NS); $P > 0.5$; $* = 0.05 > P \geq 0.01$; $** = 0.01 > P \geq 0.001$; $**** = P < 0.0001$.

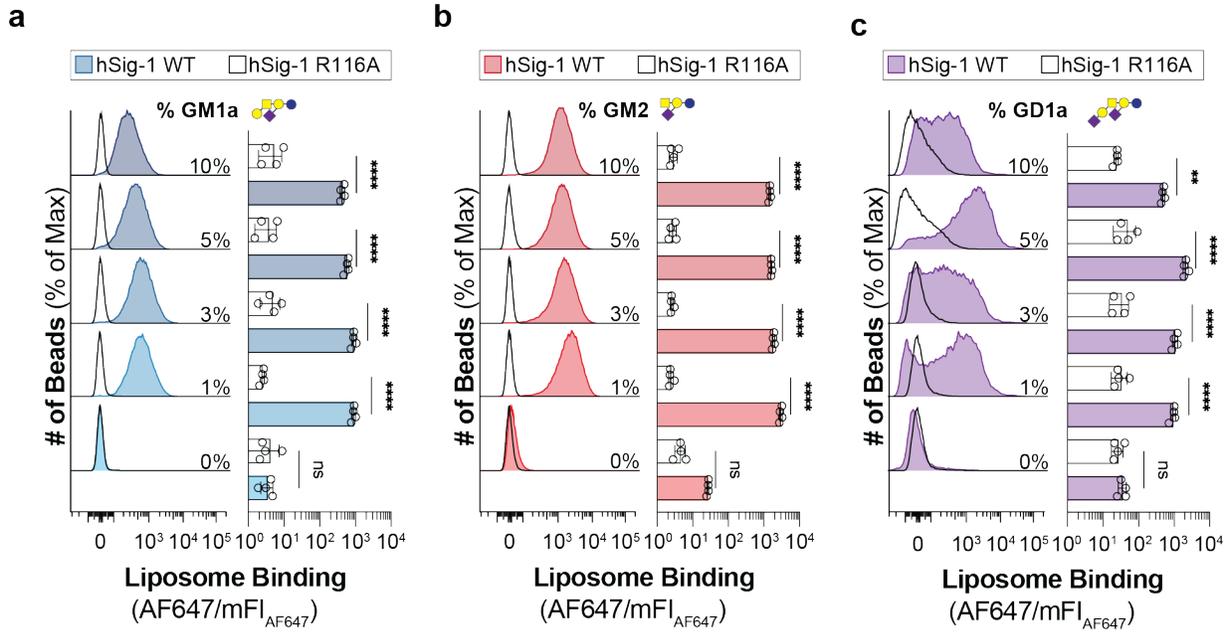


Figure A3.5: GM1a, GM2 and GD1a ganglioside content titration against human Siglec-1 in the bead assay. Data is presented with a representative flow cytometry histogram as well as a summary chart. The data in the summary charts is presented as the mean of the mean fluorescence intensity from three technical replicates. A one-way ANOVA was for statistical analysis. Not Significant (NS); $P > 0.5$; ** = $0.01 > P \geq 0.001$; *** = $0.001 > P \geq 0.0001$; **** = $P < 0.0001$.

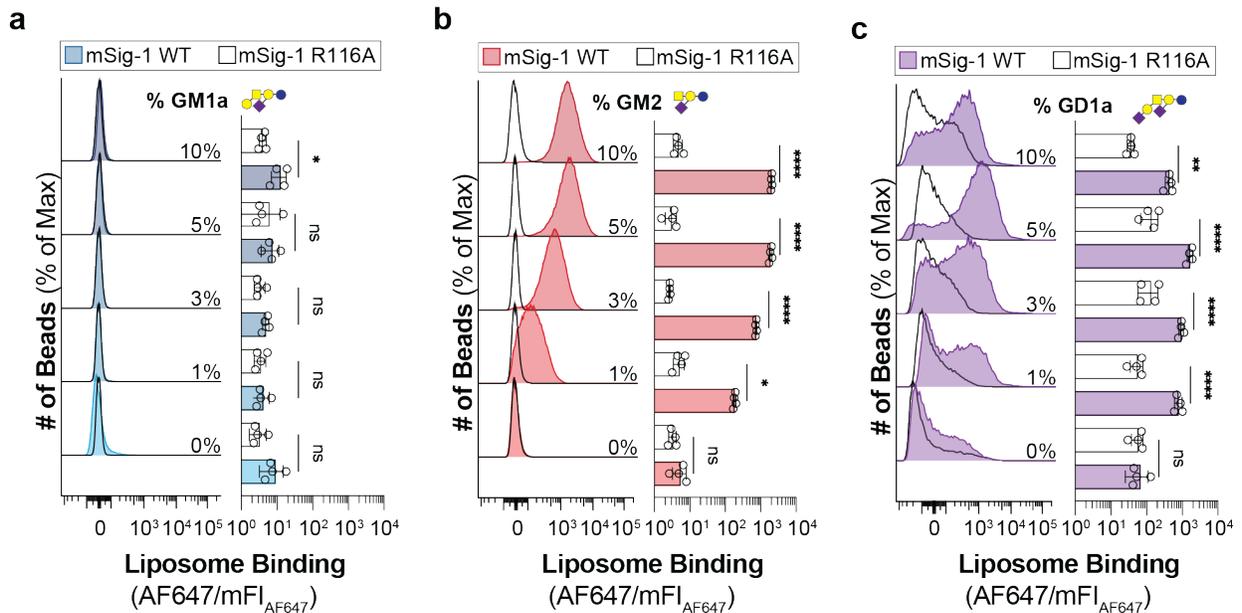


Figure A3.6: GM1a, GM2 and GD1a ganglioside content titration against murine Siglec-1 in the bead assay. Data is presented with a representative flow cytometry histogram as well as a summary chart. The data in the summary charts is presented as the mean of the mean fluorescence intensity from three technical replicates. A one-way ANOVA was for statistical analysis. Not Significant (NS); $P > 0.5$; ** = $0.01 > P \geq 0.001$; *** = $0.001 > P \geq 0.0001$; **** = $P < 0.0001$.

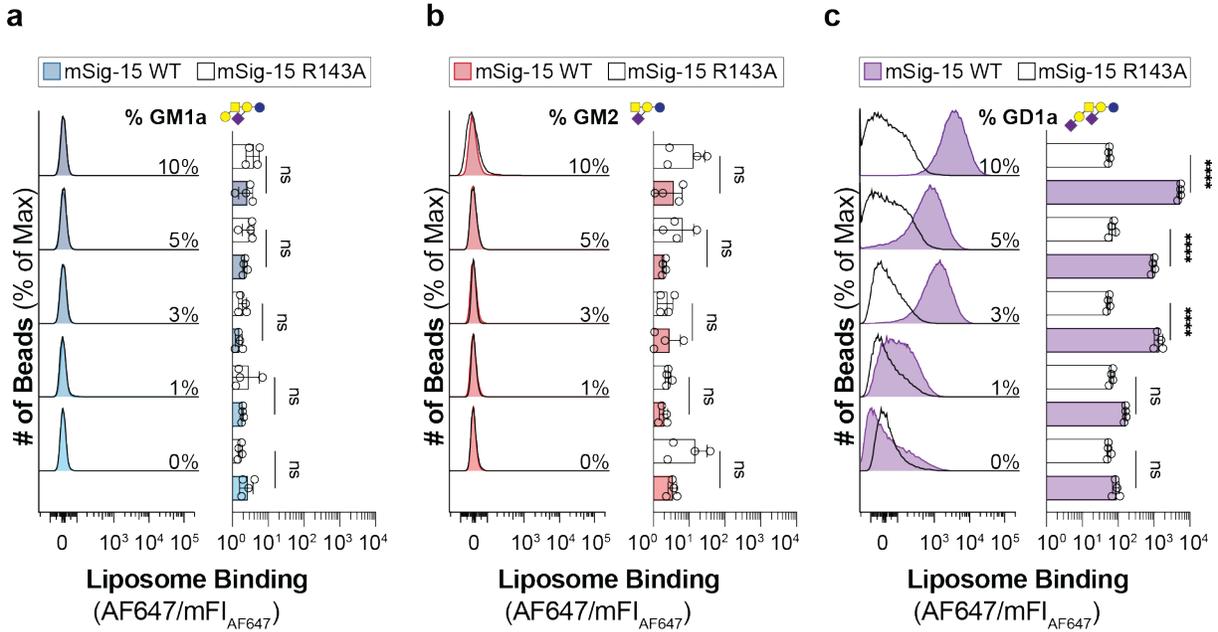


Figure A3.7: GM1a, GM2 and GD1a ganglioside content titration against murine Siglec-15 in the bead assay. Data is presented with a representative flow cytometry histogram as well as a summary chart. The data in the summary charts is presented as the mean of the mean fluorescence intensity from three technical replicates. A one-way ANOVA was for statistical analysis. Not Significant (NS); $P > 0.5$; $** = 0.01 > P \geq 0.001$; $*** = 0.001 > P \geq 0.0001$; $**** = P < 0.0001$.

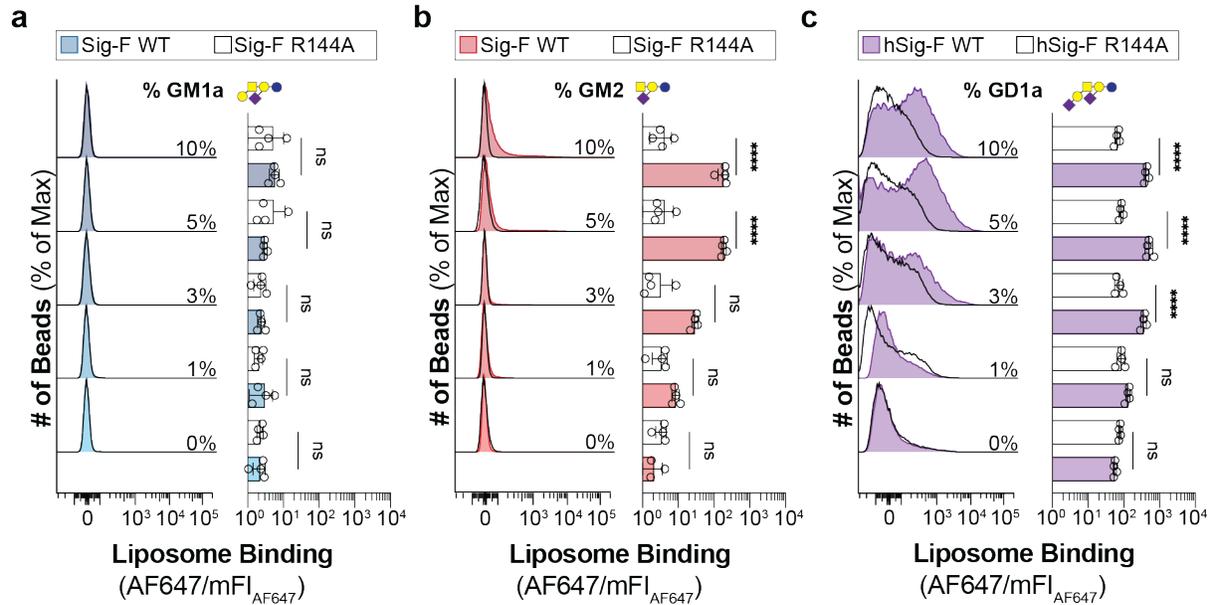


Figure A3.8: GM1a, GM2 and GD1a ganglioside content titration against mSiglec-F in the bead assay. Data is presented with a representative flow cytometry histogram as well as a summary chart. The data in the summary charts is presented as the mean of the mean fluorescence intensity from three technical replicates. A one-way ANOVA was for statistical analysis. Not Significant (NS); $P > 0.5$; $** = 0.01 > P \geq 0.001$; $*** = 0.001 > P \geq 0.0001$; $**** = P < 0.0001$.

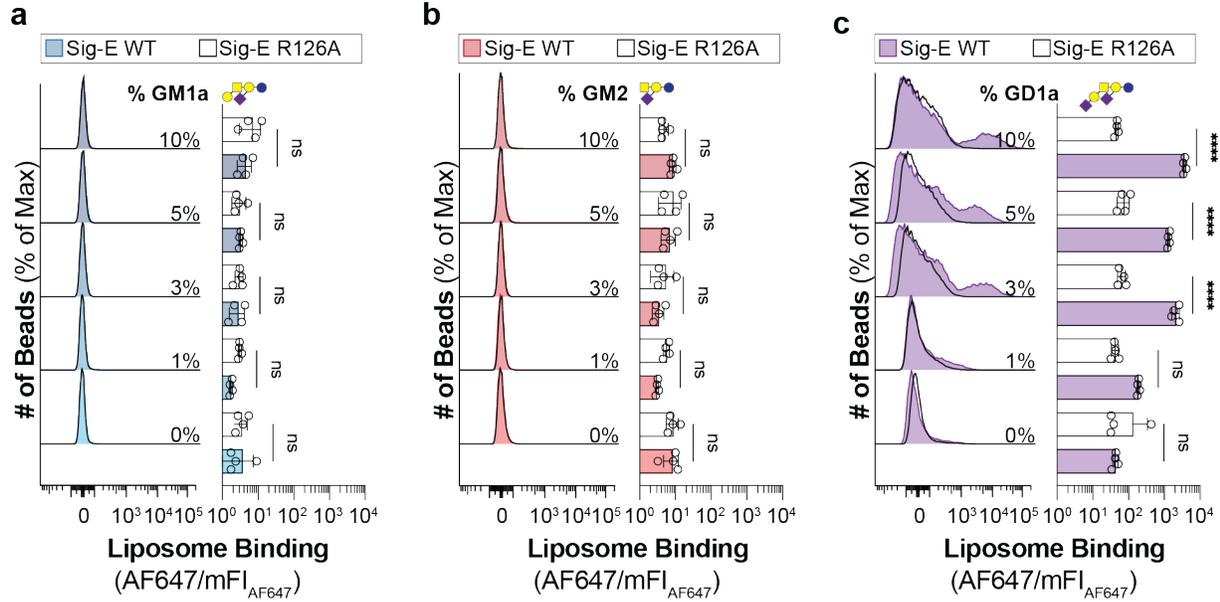


Figure A3.9: GM1a, GM2 and GD1a ganglioside content titration against mSiglec-F in the bead assay. Data is presented with a representative flow cytometry histogram as well as a summary chart. The data in the summary charts is presented as the mean of the mean fluorescence intensity from three technical replicates. A one-way ANOVA was for statistical analysis. Not Significant (NS); $P > 0.5$; $** = 0.01 > P \geq 0.001$; $*** = 0.001 > P \geq 0.0001$; $**** = P < 0.0001$.

Appendix A4:

Table A4.1: Key amino acids in Siglec-6.

Residue	Importance
Met ₁ -Ala ₂₆	signal peptide
Cys ₄₆ -Cys ₁₇₂	Interdomain disulfide
Cys ₅₁ -Cys ₁₀₄	Interdomain disulfide
Arg ₉₂	Salt bridge
Asn ₁₀₃	N-linked glycosylation
Asp ₁₁₅	Salt bridge
Arg ₁₂₂	Sialoside binding
Asn ₁₄₉	N-linked glycosylation
Asn ₁₆₃	N-linked glycosylation
Asn ₂₃₃	N-linked glycosylation
Cys ₂₇₄ -Cys ₃₁₉	Interdomain disulfide
Tyr ₄₂₆	Phosphorylation/cell signaling
Tyr ₄₄₅	Phosphorylation/cell signaling

Table A4.2 Siglec-6 Mutagenesis Primers

Primer Name	Sequence
Fwd Sig-6	AGCAGCGCTAGCATGCAGGGAGCCCAGGAAGCC
Rvs Sig-6	AGCAGCACCGGTTCACTTGTGTATCTTGATTTCC
Fwd C46A	CAGGAGGGTCTGGCCGTCTCGTACCCTG
Rvs C46A	CAGGGTACGAGGACGGCCAGACCCTCCTG
Fwd E87A	CGA AGA AGT GCA GGC GGA GAC CCG GG
Rvs E87A	CCC GGG TCT CCG CCT GCA CTT CTT CG
Fwd E88A	GAA GTG CAG GAG GCG ACC CGG GGC CG
Rvs E88A	CGG CCC CGG GTC GCC TCC TGC ACT TC
Fwd R90A	GGAGGAGACCGCGGGCCGATTCCA
Rvs R90A	GAATCGGCCCGCGGTCTCCTCCTG
Fwd R92A	CCCGGGGCGCTTTCCACCTCCTCTG
Rvs R92A	CCAGAGGAGGTGAAAGCGCCCCGGGTC
Fwd R92K	GACCCGGGGCAAATTCCACCTCCTC
Rvs R92K	GAGGAGGTGGAATTTGCCCGGGTC
Fwd F93A	CCC GGG GCC GAG CCC ACC TCC TC
Rvs F93A	GAG GAG GTG GGC TCG GCC CCG GG
Fwd L95A	CGG GGC CGA TTC CAC GCC CTC TGG GAT
Rvs L95A	ATC CCA GAG GGC GTG GAA TCG GCC CCG
Fwd R100A	CTCTGGGATCCCGCAAGGAAGAAGTCTGCTC
Rvs R100A	GAGCAGTTCTTCTTGGCGGATCCCAGAG
Fwd R101A	CTCTGGGATCCCAGAGCGAAGAAGTCTGCTC
Rvs R101A	GAGCAGTTCTTCTGCTCTGGGATCCCAGAG
Fwd R109A	CCTGAGCATCGCGGATGCCCGGAG
Rvs R109A	TCCGGGCATCGCGGATGCTCAGG
Fwd R112A	CAGAGATGCCGCCAGGAGGGACAATGC
Rvs R112A	GCATTGTCCCTCCTGGCGGCATCTCTG
Fwd R113A	CAGAGATGCCCGGGCGAGGGACAATGC
Rvs R113A	GCATTGTCCCTCGCCCGGGCATCTCTG
Fwd R114A	GAGATGCCCGGAGGGCGGACAATGCTG
Rvs R114A	CAGCATTGTCCGCCCTCCGGGCATCTC
Fwd R147A	CCCTGACCCACGCGCCCAACATCTCC
Rvs R147A	GGAGATGTTGGGCGCGTGGGTCAGGG
Fwd C172A	GCCCTGGGTCGCTGAGCAGGGGAC
Rvs C172A	GTCCCCTGCTCAGCGACCCAGGGC
Fwd E173A	CCC TGG GTC TGT GCG CAG GGG ACG
Fwd E173A	CGT CCC CTG CGC ACA GAC CCA GGG
Fwd Q174A	GGG TCT GTG AGG CGG GGA CGC CCC C
Rvs Q174A	GGG GGC GTC CCC GCC TCA CAG ACC C
Fwd G175M	CTG TGA GCA GAT GAC GCC CCC CAT CTT C
Rvs G175M	GAA GAT GGG GGG CGT CAT CTG CTC ACA G

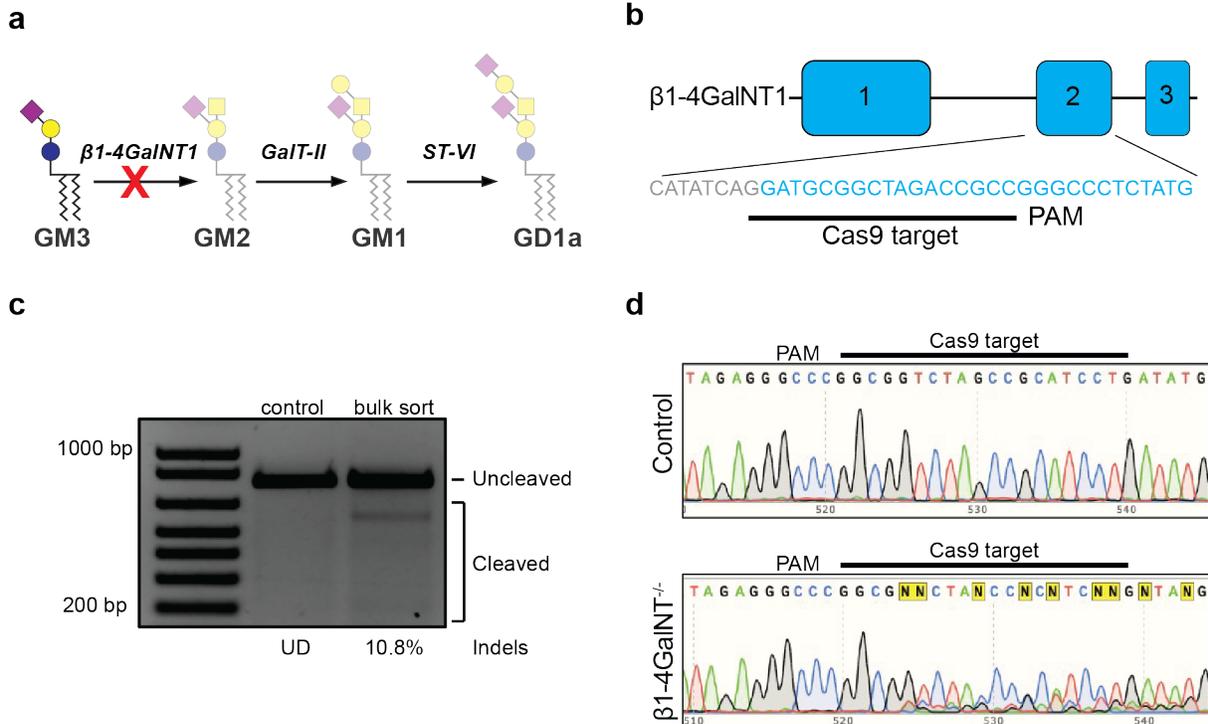
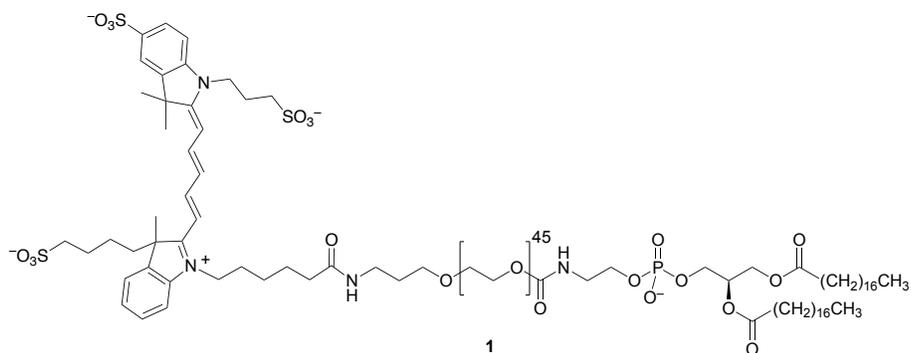


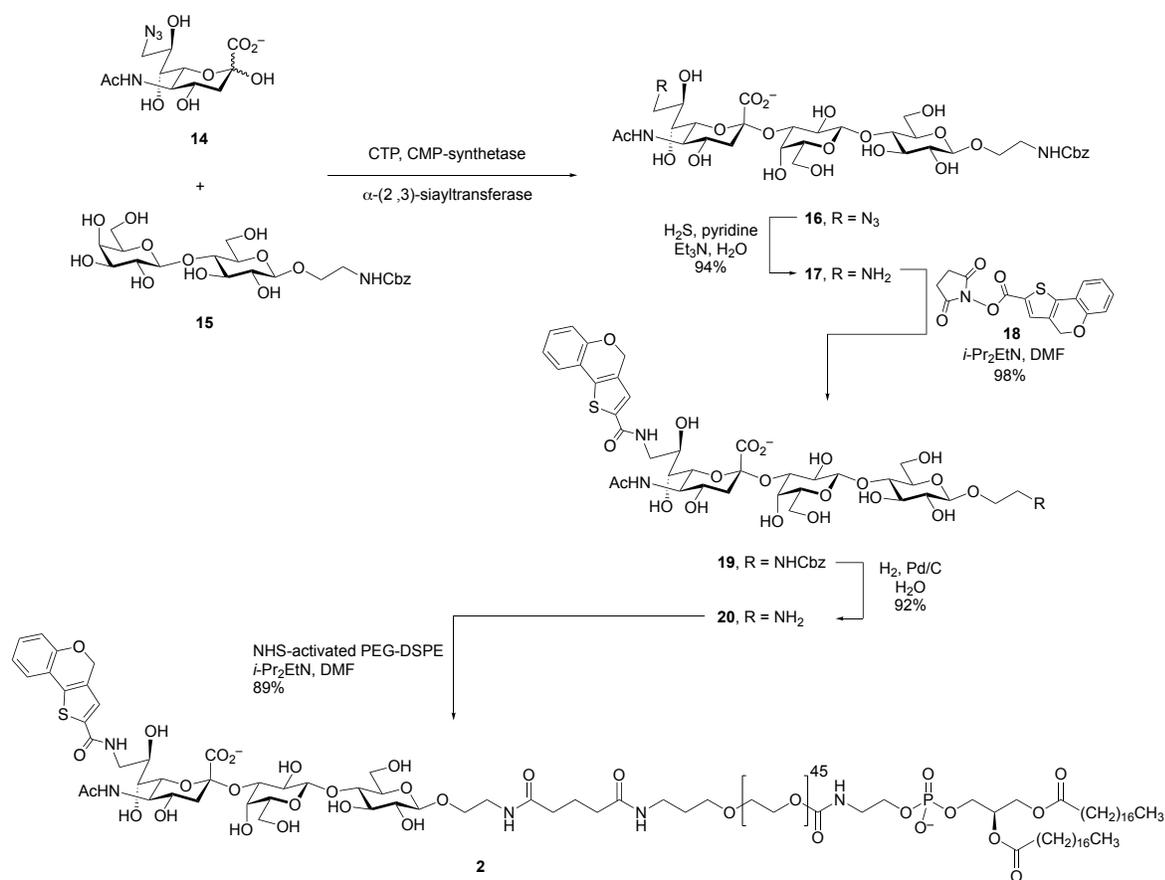
Figure A4.1: Complex gangliosides in EVs are ligands for Siglec-6. a, Abbreviated ganglioside biosynthesis highlighting the role of $\beta 1-4GalNT1$ in ganglioside biosynthesis. b, Schematic illustration of the Cas9 target site used to generate $\beta 1-4GalNT1^{-/-}$ cells. Intronic sequences are indicated by grey lettering, while blue lettering indicates exons. c, Gel showing relative cellular cleavage efficiencies of the untreated, parental cells against a population of cells FACS sorted for ATTO-550 fluorescence following Cas9 RNP transfection as determined by T7 endonuclease I digestion (n=1 technical replicate). d, Sanger sequencing trace of the $\beta 1-4galnt1$ target site for either parental control cells (above), or the monoclonal $\beta 1-4GALNT1^{-/-}$ cells (below).

Synthetic Schemes and Chemical Characterization

*Characterization and experimental procedures were described by Dr. Todd L Lowary, Dr. Maju Joe, Fahima Mozaneh, and Dr. Gour C. Daskhan.



AF647-PEG₄₅-DSPE 1 was synthesized as described by Bhattacharjee *et al.*¹²⁷



Siglec-1 ligand-PEG₄₅-DSPE 2:

A mixture of 9-azido sialic acid derivative **14**³³¹ (10 mg, 0.029 mmol), CTP (31 mg, 0.062 mmol), 12 mM MgCl₂ in Tris-HCl buffer (pH 8.6, 285 mL) was treated with CMP-synthetase (25 μ L) and the pH of the solution was adjusted to ~9.0 to 9.5 by adding 75 mL of aqueous 1M NaOH and the solution was incubated at 37 °C for 30 min. The progress of the reaction was monitored by TLC and when complete disaccharide **15**³³² (19.3 mg, 0.037 mmol), 0.5 M MgSO₄ and of Tris-HCl buffer (pH 8.6, 560 mL, final concentration 15 mM) and α -(2 \rightarrow 3)-sialyltransferases (0.15 mg/mL, PmSTI) were added. The reaction mixture was further incubated at 37 °C for 2 h. The excess enzymes were quenched by the addition of EtOH (500 μ L), the mixture was centrifuged, crude product was washed with EtOH (2 x 100 mL) and the supernatants were collected. The EtOH in the supernatant was removed under reduced pressure and the resulting aqueous solution was frozen at -80 °C and lyophilized to afford a white solid. The crude residue was subjected to P2-gel filtration chromatography using H₂O as eluent to afford **16** (22.8 mg, 90 %) as a white powder after lyophilization; ¹H NMR (600 MHz, D₂O) δ (ppm) 7.41–7.37 (m, 5H), 5.09 (s, 2H), 4.72 (d, 1H, *J* = 8.0 Hz), 4.42 (d, 1H, *J* = 8.0 Hz), 4.04 (ddd, 1H, *J* = 21.0, 12.0, 3.0 Hz), 3.94–3.90 (m, 4H), 3.81–3.76 (m, 3H), 3.70

(t, 1H, $J = 10.0$ Hz), 3.65–3.40 (m, 12H), 3.39 (dd, 1H, $J = 15.0, 7.8$ Hz), 2.98 (dd, 1H, $J = 13.2, 9.6$ Hz), 2.67 (dd, 1H, $J = 12.6, 4.8$ Hz), 1.99 (s, 3H, NHCOCH_3), 1.70 (t, $J = 12.6$ Hz, 1H). HRMS (ESI) calculated for m/z $[\text{M} - \text{H}]^+$ calcd for $\text{C}_{33}\text{H}_{49}\text{N}_5\text{O}_{20}$: 835.2971, found: 834.2889.

Trisaccharide **16** (10 mg, 0.011 mmol) was dissolved in a mixture of pyridine (1.8 mL), Et_3N (125 μL) and H_2O (125 μL) and the reaction mixture were cooled to 0 °C in an ice bath for approximately 15 min. H_2S gas was bubbled through the solution for 5 min, the reaction mixture turned an intense blue colour and the round bottom flask was capped and stirred overnight at room temperature. After completion of the reaction, the solvent was evaporated, the crude product was dissolved in methanol (10 mL) and the resulting solution was centrifuged. The supernatant was collected, concentrated and the crude mixture was purified by P2-gel filtration chromatography using H_2O as eluent to afford **17** (9.1 mg, 94 %) as a white powder after lyophilization. HRMS (ESI) calculated for m/z $[\text{M} - \text{H}]^+$ calcd for $\text{C}_{33}\text{H}_{51}\text{N}_3\text{O}_{20}$: 808.2993, found: 808.3005.

Amine **17** (5.0 mg, 6.18 mmol, 1 equiv.) and NHS-activated ester **18**¹⁶⁰ (2.5 mg, 7.72 μmol , 1.25 equiv.) were dissolved in anhydrous DMF (0.5 mL). *N,N*-Diisopropylethylamine (1.87 mL, 18.54 mmol, 3 equiv.) was added to the reaction mixture, and the solution was stirred overnight at room temperature. DMF was removed under reduced pressure and the residue was co-evaporated with toluene (2 x 5 mL). The crude product was dissolved in H_2O (5 mL) and hydrolyzed **18** was separated from the desired product **19** through successive extraction with EtOAc (5 mL, 5–6 times). TLC confirmed removal of hydrolyzed **18**. The aqueous layers were collected and lyophilized to afford **19** (5.52 mg, 98 %) as a white powder; ^1H NMR (500 MHz, D_2O) δ (ppm) 7.82 (t, 1H, $J = 7.5$ Hz), 7.76–7.72 (m, 2H), 7.68–7.66 (m, 2H), 7.42–7.40 (m, 5H), 5.10 (s, 2H), 4.49 ($J = 8.0$ Hz, 1H), 4.43 (d, 1H, $J = 8.0$ Hz), 4.08 (dd, 1H, $J = 10.0, 3.5$ Hz), 4.06–4.04 (m, 2H), 3.94–3.85 (m, 4H), 3.78–3.52 (m, 15H), 3.39–3.35 (m, 2H), 3.01 (dd, 1H, $J = 12.5, 9.5$ Hz), 2.73 (dd, 1H, $J = 12.5, 4.5$ Hz), 2.01 (s, 3H, NHCOCH_3), 1.78 (t, 1H, $J = 12.5$ Hz). HRMS (ESI) calculated for m/z $[\text{M} - \text{H}]^+$ calcd for $\text{C}_{45}\text{H}_{56}\text{N}_3\text{O}_{22}\text{S}$: 1023.3154, found: 1023.3145.

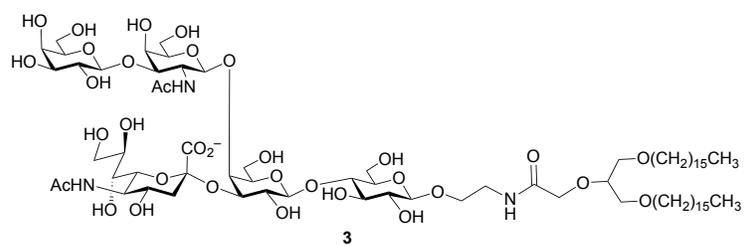
Trisaccharide **19** (5.0 mg, 4.88 mmol) was dissolved in H_2O (2 mL), followed by addition of Pd/C (10 mg, 5 mol%). Hydrogen gas was bubbled through the solution and the mixture was stirred overnight

under a hydrogen atmosphere. After completion of reaction, the catalyst was removed by filtration through a Celite pad. The solvent was evaporated, and the mixture was purified using a C18 column (gradient elution from H₂O to MeOH/H₂O (10 % to 50 %, V/V)) to produce **20** (4.0 mg, 92 %) as white solid after lyophilization of the fractions containing the desired product; ¹H NMR (600 MHz, D₂O) δ (ppm) 7.38–7.37 (m, 2H), 7.24 (t, 1H, *J* = 7.4 Hz), 7.01 (t, 1H, *J* = 7.4 Hz), 6.93 (d, 1H, *J* = 7.8 Hz), 5.19 (s, 2H), 4.34 (d, 2H, *J* = 7.8 Hz), 4.03 (dd, 1H, *J* = 9.6, 3.0 Hz), 3.96 (ddd, 1H, *J* = 11.4, 9.6, 3.0 Hz), 3.88 (d, 1H, *J* = 3.0 Hz), 3.83–3.80 (m, 4H), 3.70–3.64 (m, 8H), 3.56–3.46 (m, 8H), 3.44–3.32 (m, 2H), 3.16 (dd, 1H, *J* = 12.5, 9.5 Hz), 2.01 (t, 1H, *J* = 4.8 Hz), 2.73 (dd, 1H, *J* = 12.5, 4.8 Hz), 1.99 (s, 3H, NHCOCH₃), 1.75 (t, 1H, *J* = 12.5 Hz). ¹³C NMR (125 MHz, D₂O) δ (ppm) 176.0, 174.6, 164.8, 152.8, 137.8, 137.1, 133.3, 131.4, 127.5, 124.8, 123.8, 120.2, 117.6, 103.7, 102.88, 100.9, 79.1, 77.0, 76.3, 75.6, 75.0, 73.9, 73.6, 71.9, 70.9, 70.2, 69.2, 68.3, 67.6, 66.7, 63.5, 62.0, 60.8, 52.7, 49.9, 43.8, 40.5, 23.0. HRMS (ESI-MS) calculated for *m/z* [M + Na]⁺ calcd for C₃₇H₅₁N₃NaO₂₀S: 912.2679, found: 912.2681.

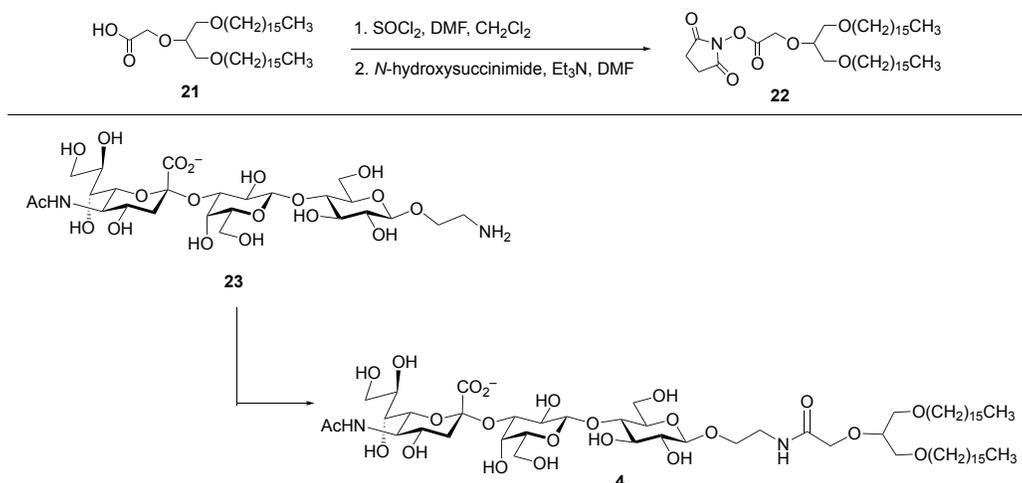
A mixture of **20** (1 mg, 0.91 mmol, 1.25 equiv.) and NHS-activated PEG-DSPE (2.5 mg, 0.83 mmol, 1 equiv.) were dissolved in anhydrous DMF (100–150 μL, ~10 mM) and placed in a 0.5 mL centrifuge tube at room temperature. The reaction mixture was degassed with N₂. A diluted solution of DIPEA (1.50–2.0 equiv.) in dry DMF was added carefully to adjust pH of the solution ~8.0 and the reaction mixture was stored at room temperature overnight. An aliquot of the reaction mixture was taken for TLC (CHCl₃–MeOH–H₂O = 75:23:2) analysis. The coupling was performed under anhydrous condition to avoid hydrolysis of the NHS-activated PEG-DSPE. The solvent was removed under *vacuum*, and the crude product was dissolved in water. The crude product was loaded to Sephadex G-100 gel filtration column using H₂O as an eluent to afford **20**-PEG-DSPE conjugate as a white powder after lyophilization of fractions having the desired product. Yield: (3.8 mg, 89 %), coupling efficiency 62 %.

¹H NMR (600 MHz, MeOD₄): δ 7.49 (d, *J* = 7.8 Hz, 1H), 7.38–7.33 (m, 1H), 7.19 (t, *J* = 7.8 Hz, 1H), 6.97 (t, *J* = 7.8 Hz, 1H), 7.99 (d, *J* = 7.8 Hz, 1H), 5.25 (s, 2H), 4.44 (d, *J* = 8.0 Hz, 1H), 4.42 (d, *J* = 8.0 Hz, 1H), 4.31 (d, *J* = 7.8 Hz, 1H), 4.19–3.98 (m, 2H), 4.97 (t, *J* = 5.4 Hz, 1H), 3.90–4.86 (m, 2H), 3.74–3.71 (m, 2H), 3.64 (broad s, 123H), 3.3–3.34–3.22 (m, 8H), 2.35–2.29 (m, 4H), 2.23–2.19 (m, 4H), 1.99 (s, 3H), 1.89–1.86 (m, 2H), 1.76–1.73 (m, 2H), 1.60–1.58 (m, 2H), 1.28 (broad s, 64H), 0.89 (s, 6H); The MALDI-TOF-MS

spectrum showed the average mass centered at 3.8 kDa and the expected average mass was 3.8 kDa. The coupling efficiency was determined through assigning underline signals of the aromatic protons signals at 7.49 ppm (d), 7.38–7.33 ppm (m), 6.97 (t) and 7.99 (d) of the phenyl moieties at C9 position of bifunctionally substituted Neu5Ac with terminal methyl groups at 0.89 ppm (s) of the DSPE lipid.

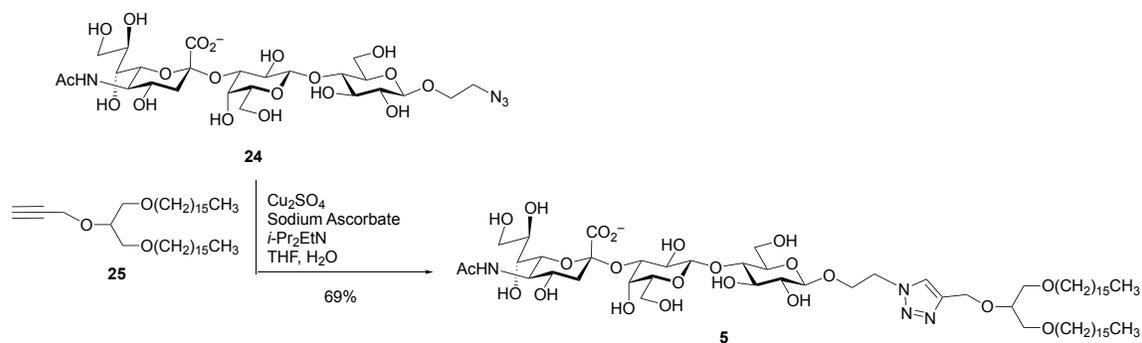


Neoglycolipid 3 was prepared as described previously by Han et al.³³³

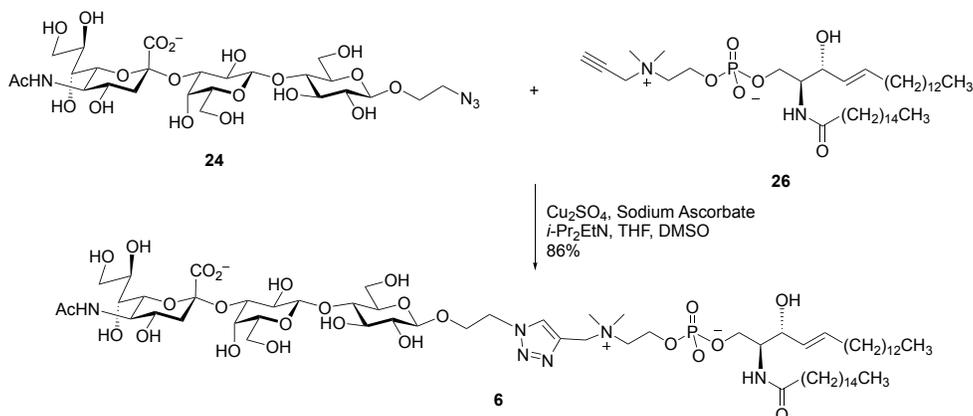


Neoglycolipid 4. To a stirred solution of acid **21** (8.0 mg, 13.4 μmol) in CH_2Cl_2 (3 mL) were added thionyl chloride (5.0 μL , 68.0 μmol) and dry DMF (20.0 μL) successively at room temperature. After heating at reflux overnight, the solution was cooled, the solvent was evaporated, and the residue was dried under high vacuum (2.5 h). The dried acyl chloride was dissolved in CH_2Cl_2 (3 mL), *N*-hydroxysuccinimide (6.0 mg, 52.0 μmol) and triethylamine (5.0 μL , 36.0 μmol) were added at 0 °C and the solution was heated at reflux overnight. After cooling to room temperature, the solution was diluted with CH_2Cl_2 (5 mL), washed with brine, dried over Na_2SO_4 , filtered, and the filtrate was concentrated to give the NHS-ester **22**, which was used directly for the next step without any further purification.

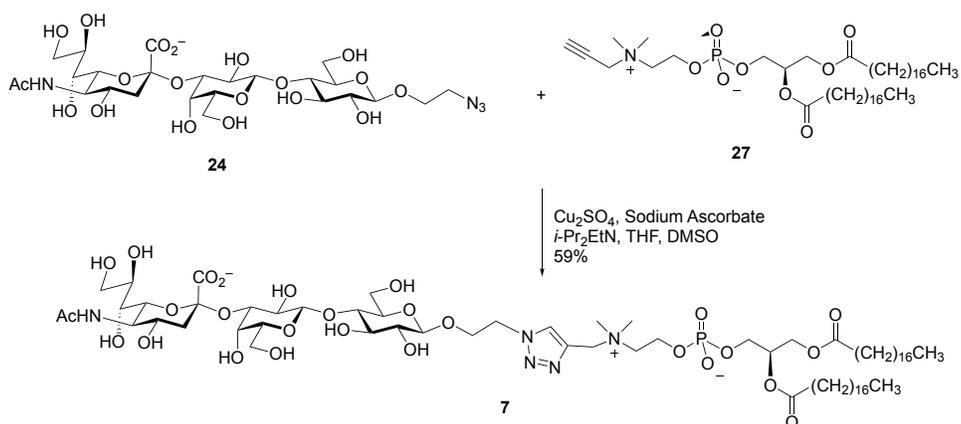
Trisaccharide amine **23**³³⁴ (2.0 mg, 3.0 μmol) was dissolved in *N,N*-dimethylacetamide (0.4 mL), and *N,N*-diisopropylethylamine (5 μL , 28.0 μmol) was added. This solution was then added to a glass vial containing **22** followed by THF (0.3 mL) and the mixture was stirred overnight at room temperature. The reaction mixture was directly loaded and purified by size exclusion column chromatography (Sephadex-LH20, CH_2Cl_2 – CH_3OH , 1:1) to afford **4** (0.9 mg, 25 %). R_f 0.3, EtOAc – CH_3OH – HOAc – H_2O (36:9:9:6); ^1H NMR (700 MHz, CD_3OD plus a few drops of CDCl_3 , δ_{H}) 4.42 (d, 1H, $J = 7.8$ Hz, H-1), 4.31 (d, 1H, $J = 7.8$ Hz, H-1), 4.15–4.10 (m, 2H), 4.05 (dd, 1H, $J = 3.1, 9.7$ Hz), 3.96–3.90 (m, 3H), 3.90–3.82 (m, 3H), 3.80–3.74 (m, 2H), 3.74–3.40 (m, 24H), 3.26 (dd, 1H, $J = 7.9, 7.9$ Hz), 3.21 (ddd, 2H, $J = 1.6, 3.2, 4.8$ Hz), 2.85 (dd, 1H, $J = 4.2, 12.5$ Hz), 2.79 (dd, 1H, $J = 7.5, 7.5$ Hz), 2.56 (dd, 1H, $J = 7.5, 7.5$ Hz), 2.00 (s, 3H, NHCOCH_3), 1.75–1.60 (m, 5H), 1.40–1.20 (m, 56H), 0.86 (dd, 6H, $J = 6.9$ Hz); HRMS (ESI) calcd. for $[\text{M} - \text{H}]^-$ $\text{C}_{62}\text{H}_{116}\text{N}_2\text{O}_{23}$ 1255.7896, found 1255.7909.



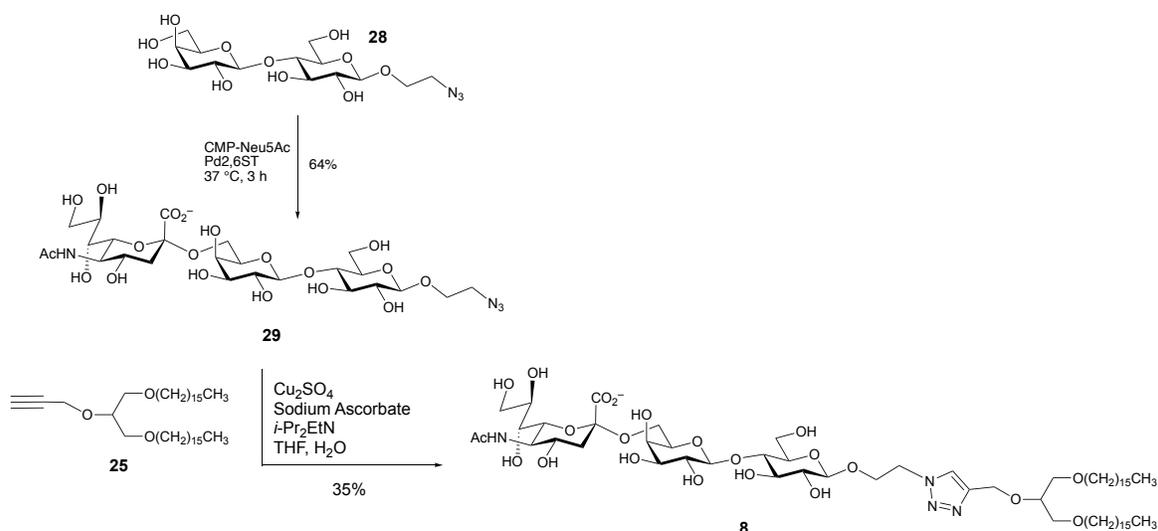
Neoglycolipid 5. To a stirred solution of trisaccharide azide **24**³⁵ (8.0 mg, 11.0 μmol) and alkyne **25**⁶⁷ (15.0 mg, 26.0 μmol) in a mixture of THF (3 mL) and water (3 mL) at room temperature were added *N,N*-diisopropylethylamine (6.0 μL , 34.0 μmol), copper (II) sulfate pentahydrate (53.0 mg, 212.0 μmol) and L-ascorbic acid sodium salt (81.0 mg, 408.0 μmol) successively. The reaction mixture was shielded from light (aluminum foil) and stirred overnight. The reaction mixture was then concentrated, and the residue was purified by size exclusion column chromatography (Sephadex-LH-20, CH_2Cl_2 – CH_3OH , 1:1) to afford **5** (10.0 mg, 69 %). R_f 0.3, EtOAc – CH_3OH – HOAc – H_2O (36:9:9:6); $^1\text{H NMR}$ (700 MHz, CD_3OD plus a few drops of CDCl_3 , δ_{H}) 8.03 (s, 1H), 4.65–4.55 (m, 2H), 4.39 (d, 1H, $J = 7.0$ Hz, H-1), 4.31 (d, 1H, $J = 7.8$ Hz, H-1), 4.24–4.20 (m, 2H), 3.95–3.40 (m, 28H), 3.22–3.20 (m, 1H), 2.90–2.80 (m, 1H), 2.10–2.00 (m, 2H), 1.99 (s, 3H, NHCOCH_3), 1.70–1.50 (m, 5H), 1.40–1.20 (m, 56H), 0.86 (dd, 6H, $J = 6.9$ Hz); HRMS (ESI) calcd. for $[\text{M} - \text{H}]^- \text{C}_{63}\text{H}_{116}\text{N}_4\text{O}_{22}$ 1279.8008, found 1279.8014.



Neoglycolipid 6. To a stirred solution of trisaccharide azide **24**³³⁵ (2.2 mg, 3.1 μmol) in water (0.5 mL), alkyne **26** (Avanti Polar Lipids, Inc., USA) (1.0 mg, 1.4 μmol) in a mixture of THF (0.3 mL) and DMSO (0.5 mL) was added at room temperature followed by *N,N*-diisopropylethylamine (3.0 μL , 17.0 μmol). Copper (II) sulfate pentahydrate (21.0 mg, 84.0 μmol) and L-ascorbic acid sodium salt (37.0 mg, 186.0 μmol) were then added successively. The reaction mixture was shielded from light (aluminum foil) and stirred overnight. After 24 h, copper (II) sulfate pentahydrate (14.0 mg, 56.0 μmol), and L-ascorbic acid sodium salt (30.0 mg, 151.0 μmol) each dissolved in 0.1 mL water were added again successively and stirring continued overnight. The reaction mixture was then diluted with CH₃OH–H₂O (1:1, 0.2 mL) stirred well and filtered through a cotton plug to remove most of the insoluble salts. The filtrate was directly loaded on to a C-18 column and purified by gradient elution (H₂O to CH₃OH–H₂O to neat CH₃OH) to afford **6** (1.7 mg, 86 %). *R*_f 0.1, EtOAc–CH₃OH–HOAc–H₂O (18:9:9:6); ¹H NMR (700 MHz, CD₃OD plus a few drops of CDCl₃, δ_{H}) 8.4 (s, 1H), 5.74–5.66 (m, 1H), 5.48–5.42 (m, 1H), 4.75 (s, 2H), 4.73–4.70 (m, 2H), 4.56 (s, 1H), 4.42 (d, 1H, *J* = 7.6 Hz, H-1), 4.39–4.34 (m, 2H), 4.32 (d, 1H, *J* = 7.8 Hz, H-1), 4.28–4.24 (m, 1H), 4.20–4.10 (m, 1H), 4.10–3.40 (m, 28H), 3.23–3.20 (m, 2H), 3.20 (s, 3H), 3.23–3.20 (m, 2H), 2.90–2.84 (m, 1H), 2.65 (s, 3H), 2.22–2.14 (m, 2H), 2.06–2.02 (m, 2H), 2.00 (s, 3H, NHCOCH₃), 1.76–1.50 (m, 5H), 1.40–1.20 (m, 52H), 0.90 (dd, 6H, *J* = 6.9 Hz); HRMS (ESI) calcd. for [M – H][–] C₆₆H₁₂₁N₆O₂₅P 1427.8046, found 1427.8048.



Neoglycolipid 7. To a stirred solution of trisaccharide azide **24**³³⁵ (2.2 mg, 3.1 μmol) in water (0.6 mL), alkyne **27** (Avanti Polar Lipids, Inc., USA) (1.0 mg, 1.2 μmol) in a mixture of THF (0.3 mL) and DMSO (0.5 mL) was added at room temperature followed by *N,N*-diisopropylethylamine (3.0 μL , 17.0 μmol). Copper (II) sulfate pentahydrate (23.0 mg, 92.0 μmol) and L-ascorbic acid sodium salt (54.0 mg, 272.0 μmol) were then added successively. The reaction mixture was shielded from light (aluminum foil) and stirred overnight. After 24 h, copper (II) sulfate pentahydrate (16.0 mg, 64.0 μmol), and L-ascorbic acid sodium salt (34.0 mg, 171.0 μmol) each dissolved in 0.05 mL water were added again successively along with another 0.05 mL each of THF and DMSO and stirring continued overnight. The reaction mixture was then diluted with $\text{CH}_3\text{OH-H}_2\text{O}$ (1:1, 0.2 mL) stirred well and filtered through a cotton plug to remove most of the insoluble salts. The filtrate was directly loaded onto a C-18 column and purified by gradient elution (H_2O to $\text{CH}_3\text{OH-H}_2\text{O}$ to neat CH_3OH) to afford **7** (1.1 mg, 59 %). R_f 0.21, $\text{EtOAc-CH}_3\text{OH-HOAc-H}_2\text{O}$ (18:9:9:6); $^1\text{H NMR}$ (700 MHz, CD_3OD plus a few drops of CDCl_3 , δ_{H}) 8.5 (s, 1H), 5.38–5.40 (m, 2H), 5.40–5.32 (m, 1H), 4.76 (s, 2H), 4.73–4.70 (m, 2H), 4.48–4.43 (m, 2H), 4.42 (d, 1H, $J = 7.9$ Hz, H-1), 4.39–4.34 (m, 2H), 4.32 (d, 1H, $J = 7.9$ Hz, H-1), 4.24–4.17 (m, 1H), 4.10–3.40 (m, 28H), 3.24–3.19 (m, 2H), 3.18 (s, 3H), 2.90–2.84 (m, 1H), 2.65 (s, 3H), 2.36–2.30 (m, 2H), 2.22–2.16 (m, 2H), 2.00 (s, 3H, NHCOCH_3), 1.66–1.50 (m, 6H), 1.40–1.20 (m, 60H), 0.90 (dd, 6H, $J = 6.9$ Hz); HRMS (ESI) calcd. for $[\text{M} - \text{H}]^-$ $\text{C}_{71}\text{H}_{130}\text{N}_5\text{O}_{27}\text{P}$ 1514.8618, found 1514.8631.

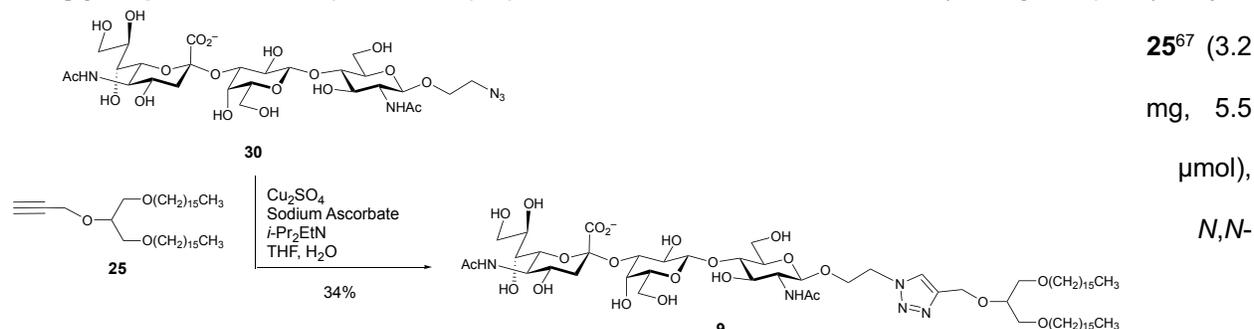


Neoglycolipid 8. Lactoside **28**³³⁶ (1.5 mg, 0.012 mmol), CMP-Sialic acid (3.4 mg, 0.016 mmol), and MgSO₄ (10 mM) were dissolved in Tris-HCl buffer (100 mM, 340 μ L, pH 8.8). Pd_{2,6}ST (0.15 mg/mL) recombinant shrimp alkaline phosphatase³⁰⁰ (1 μ L) were added to the mixture and the reaction was placed in a shaking incubator (37 °C, 3 h). The reaction was monitored using TLC in *i*-PrOH–NH₄OH–H₂O (5:2:1) and stopped by dilution with 4 volumes of cold 95 % ethanol. The precipitated protein was centrifuged (3700 rcf, 15 min) and the supernatant was carefully decanted into a round bottom flask and concentrated. The residue was resuspended in Milli-Q water and purified on a P2 gel filtration equilibrated in Milli-Q water giving compound **29** (1.1 mg, 64 %). ¹H NMR (700 MHz, D₂O) δ = 4.54 (d, 1H, *J* = 7.7 Hz), 4.42 (d, 1H, *J* = 7.7 Hz), 4.02–4.08 (m, 1H), 3.94–4.01 (m, 5H), 3.78–3.93 (m, 4H), 3.59–3.77 (m, 5H), 3.49–3.58 (m, 6H), 3.42–3.62 (m, 2H), 2.74 (dd, 1H, *J* = 7.7, 12.6 Hz), 2.03 (s, 3 H), 1.80 (app t, 1H, *J* = 11.9 Hz). HRMS (ESI) calcd for *m/z* [M – H][–] C₂₅H₄₂N₄O₁₉ 702.2443, found 701.2352.

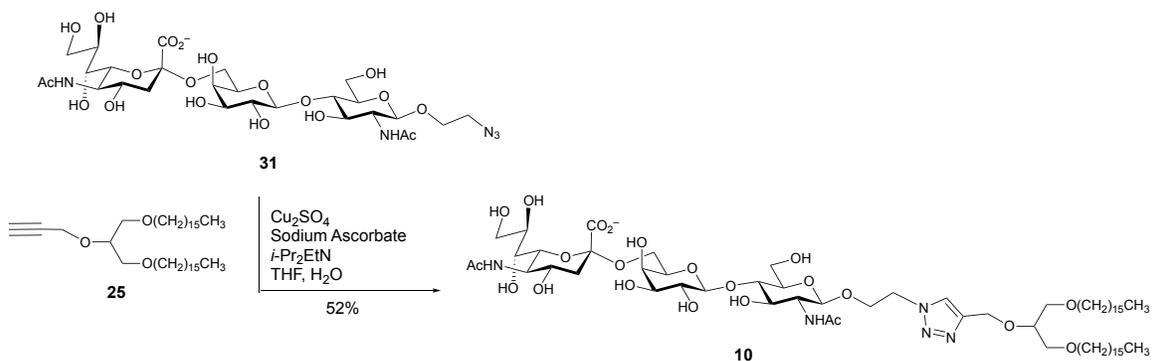
To a stirred solution of trisaccharide azide **29** (1.1 mg, 1.6 μ mol) in water (0.4 mL) was added a separately prepared solution of alkyne **25**⁶⁷ (3.1 mg, 5.3 μ mol) and *N,N*-diisopropylethylamine (1 μ L, 5.7 μ mol) in THF (0.4 mL). Additional THF (0.2 mL) was added followed by successive additions of copper (II) sulfate pentahydrate (13 mg, 52.0 μ mol; in 0.1 mL water) and L-ascorbic acid sodium salt (26 mg, 131.0 μ mol; in 0.1 mL water). The reaction mixture was shielded from light (aluminum foil) and stirred overnight. The reaction mixture was then transferred to another flask and concentrated under vacuum and a solution of CH₂Cl₂–CH₃OH (1:1, 10 mL) was added. The sides of the flask were scraped well to ensure complete

dissolution of the product and filtered through a cotton plug. The filtrate was concentrated, and the residue was purified by size exclusion column chromatography (Sephadex-LH-20, CH₂Cl₂-CH₃OH, 1:1) to afford **8** (0.7 mg, 35 %). *R_f* 0.23, EtOAc-CH₃OH-HOAc-H₂O (36:9:9:6); ¹H NMR (700 MHz, CD₃OD plus a few drops of CDCl₃, δ_H) 8.09 (s, 1H), 4.75 (s, 2H), 4.66-4.62 (m, 2H), 4.35 (d, 1H, *J* = 7.8 Hz, H-1), 4.31 (d, 1H, *J* = 7.7 Hz, H-1), 4.25-4.17 (m, 2H), 4.05-3.35 (m, 28H), 3.18 (dd, 2H, *J* = 0, 0 Hz), 2.90-2.80 (m, 1H), 2.10-2.05 (m, 2H), 2.0 (s, 3H, NHCOCH₃), 1.70-1.45 (m, 6H), 1.40-1.16 (m, 56H), 0.86 (dd, 6H, *J* = 6.9 Hz); HRMS (ESI) calcd. for [M - H]⁻ C₆₃H₁₁₆N₄O₂₂ 1279.8008, found 1279.8018.

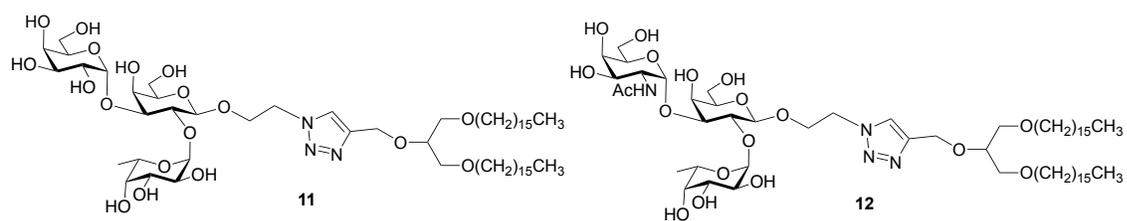
Neoglycolipid 9. This compound was prepared from trisaccharide azide **30**³³⁷ (1.0 mg, 1.3 μmol), alkyne



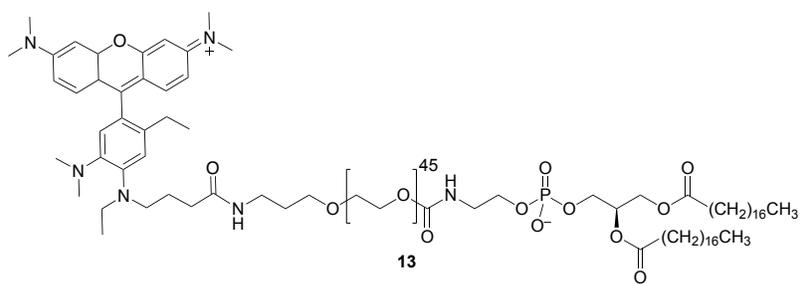
diisopropylethylamine (1 μL , 5.7 μmol), copper (II) sulfate pentahydrate (13 mg, 52.0 μmol ; in 0.1 mL water) and L-ascorbic acid sodium salt (26 mg, 131.0 μmol ; in 0.1 mL water) as described for the preparation of **8** to afford **9** (0.6 mg, 34 %). *R_f* 0.25, EtOAc–CH₃OH–HOAc–H₂O (36:9:9:6); ¹H NMR (700 MHz, CD₃OD plus a few drops of CDCl₃, δ_{H}) 7.91 (s, 1H), 4.75 (s, 2H), 4.44 (d, 1H, *J* = 7.8 Hz, H-1), 4.37 (d, 1H, *J* = 8.4 Hz, H-1), 4.26–4.20 (m, 1H), 4.04 (dd, 1H, *J* = 3.0, 9.7 Hz), 3.97–3.82 (m, 4H), 3.80–3.40 (m, 24H), 3.21 (ddd, 2H, *J* = 1.7, 3.3, 5.0 Hz), 2.85 (dd, 1H, *J* = 12.2, 4.7 Hz), 2.10–2.05 (m, 2H), 2.0 (s, 3H, NHCOCH₃), 1.91 (s, 3H, NHCOCH₃), 1.70–1.50 (m, 6H), 1.40–1.20 (m, 56H), 0.86 (dd, 6H, *J* = 6.9 Hz); HRMS (ESI) calcd. for [M – H][–] C₆₅H₁₁₉N₅O₂₂ 1320.8274, found 1320.8280.



Neoglycolipid 10. This compound was prepared from trisaccharide azide **31**³³⁷ (2.4 mg, 3.2 μ mol), alkyne **25**⁶⁷ (5.8 mg, 10.0 μ mol), *N,N*-diisopropylethylamine (3 μ L, 17.0 μ mol), copper (II) sulfate pentahydrate (21 mg, 84.0 μ mol; in 0.1 mL water) and L-ascorbic acid sodium salt (37 mg, 186.0 μ mol; in 0.1 mL water) as described for the preparation of **8** to afford **10** (2.2 mg, 52 %). R_f 0.18, EtOAc–CH₃OH–HOAc–H₂O (36:9:9:6); ¹H NMR (700 MHz, CD₃OD plus a few drops of CDCl₃, δ_H) 7.95 (s, 1H), 4.75 (s, 2H), 4.63–4.55 (m, 2H), 4.56 (d, 1H, J = 8.4 Hz, H-1), 4.32 (d, 1H, J = 7.7 Hz, H-1), 4.26–4.20 (m, 1H), 4.03 (dd, 1H, J = 9.6, 9.6 Hz), 3.94–3.78 (m, 4H), 3.80–3.40 (m, 24H), 3.21 (ddd, 2H, J = 1.7, 3.3, 5.0 Hz), 2.78 (dd, 1H, J = 4.5, 12.1 Hz), 1.99 (s, 3H, NHCOCH₃), 1.94 (s, 3H, NHCOCH₃), 1.70–1.50 (m, 5H), 1.40–1.20 (m, 56H), 0.89 (dd, 6H, J = 6.9 Hz); HRMS (ESI) calcd. for [M – H]⁻ C₆₅H₁₁₉N₅O₂₂ 1320.8274, found 1320.8271.



Neoglycolipids 11 and 12, which were prepared as described by Han et al.³³³



pHrodo-PEG₄₅-DSPE 13: 13 was prepared as described by Bhattacharjee et al.¹²⁷