

Multiplexed Liquid Glycan Array (LiGA) for Serological Assays

by

Revathi Reddy

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

Master of Science

Department of Chemistry

University of Alberta

© Revathi Reddy, 2021

Abstract

In this thesis, I describe a multiplexed liquid glycan array tool for serological assays. Human serum has a diverse assortment of carbohydrate binding antibodies that are generated as a response to glycans present on the cell surface of pathogens, tumor cells and vaccines. Glycan arrays are an established high-throughput tool that allows characterization of the interactions between glycans and glycan binding receptors (proteins, antibodies). One of the drawbacks of conventional “solid” printed array is the necessity for specialized array reader instruments that are not available in most clinical laboratories. Most questions previously answered by DNA-array and reader technologies are now answered by deep-sequencing of DNA. Building on this change, I demonstrate the utility of a “liquid” glycan array (LiGA) technology for specific detection of ABO blood group antibodies in human serum.

In the first chapter of this thesis, I briefly reviewed the different glycan array platforms developed for the detection of carbohydrate binding antibodies in human serum. I also collated the corresponding glycan array data from different research groups that would potentially benefit the glycobiology community.

In Chapter 2, I constructed a multiplexed LiGA (m-LiGA) which is an upgraded variant of previously published LiGA. I introduced multiple technical replicates within the same bulk assay to allow for quantification of reproducibility and robustness using Z' -score. I conjugated ABO blood group glycans to m-LiGA and performed quality control using MALDI-TOF, ELISA and PFU assays. ELISA was used to fine-tune the concentration of

targets - anti-A and anti-B antibodies as well as of the antigens - ABO LiGA library to be used for binding experiments. I used PFU assay as a proxy for deep sequencing during optimization of binding assays.

In Chapter 3, I checked the integrity and functional performance of ABO LiGA using PCR and NGS technologies. I demonstrate how multiple DNA barcodes encoding the same glycan structure and same glycan density makes it possible to calculate the Z' -score and characterize the confidence of LiGA-based assay.

In Chapter 4, I demonstrate the utility of ABO LiGA for specific detection of anti-A and anti-B IgG and IgM antibodies in human serum samples. I also compared the antibody binding response to ABO LiGA with other existing glycan array platforms like the traditional glass, Luminex and neoglycoprotein arrays.

In summary complex, I developed a LiGA-based tool that can be used for profiling anti-glycan antibodies in a mixture like the human serum. This tool was functionally validated using different quality control methods such as ELISA, PFU assays and NGS. However, in this project I specifically focused on profiling antibodies that bind to the 3 ABO blood group determinants. Construction of an ABO array with 18 different subtypes of glycans is needed to thoroughly profile the human serum for blood group antibodies. I believe that this project laid the groundwork for using LiGA to screen biologically relevant carbohydrate binding antibodies in diseased patient samples.

Preface

In **Chapter 1**, I briefly reviewed the different glycan array technologies used for profiling carbohydrate binding antibodies in human serum and is not published. I wrote the manuscript, collected and organized glycan array data from literature with advice from R. Derda. Portion of this chapter is from the published work of Sojitra, M.; Sarkar, S.; Maghera, J.; Rodrigues, E.; Carpenter, E. J.; Seth, S.; Ferrer Vinals, D.; Bennett, N. J.; **Reddy, R.**; Khalil, A.; Xue, X.; Bell, M. R.; Zheng, R. B.; Zhang, P.; Nycholat, C.; Bailey, J. J.; Ling, C.-C.; Lowary, T. L.; Paulson, J. C.; Macauley, M. S.; Derda, R., “Genetically encoded multivalent liquid glycan array displayed on M13 bacteriophage,” *Nat. Chem. Biol.* **2021**, 17 (7), 806-816. My contribution to this publication was performing LiGA screening on anti-A and anti-B antibodies. **Figure 1-4** is reused with permission from Purohit, S.; Li, T.; Guan, W.; Song, X.; Song, J.; Tian, Y.; Li, L.; Sharma, A.; Dun, B.; Mysona, D.; Ghamande, S.; Rungruang, B.; Cummings, R. D.; Wang, P. G.; She, J.-X., “Multiplex glycan bead array for high throughput and high content analyses of glycan binding proteins,” *Nat. Commun* **2018.**, 9 (1), 258. Copyright © Nature Publishing Group. **Figure 1-5** is reused with permission from Yan, M.; Zhu, Y.; Liu, X.; Lasanajak, Y.; Xiong, J.; Lu, J.; Lin, X.; Ashline, D.; Reinhold, V.; Smith, D. F.; Song, X., “Next-Generation Glycan Microarray Enabled by DNA-Coded Glycan Library and Next-Generation Sequencing Technology,” *Anal. Chem.* **2019**, 91 (14), 9221-9228. Copyright © American Chemical Society. **Figure 1-6** is reused from Kondengaden, S. M.; Zhang, J.; Zhang, H.; Parameswaran, A.; Kondengadan, S. M.; Pawar, S.; Puthengot, A.; Sunderraman, R.; Song, J.; Polizzi, S. J.; Wen, L.; Wang, P. G., “DNA

Encoded Glycan Libraries as a next-generation tool for the study of glycan-protein interactions,” bioRxiv **2020**, 2020.03.30.017012. **Figure 1-7** is reused with permission from Celik, E.; Fisher, A. C.; Guarino, C.; Mansell, T. J.; DeLisa, M. P., “A filamentous phage display system for N-linked glycoproteins,” *Protein Sci.* **2010**, 19 (10), 2006-13. Copyright © The Protein Society. **Figure 1-8** and **Figure 1-9** are reused with permission from Sojitra, M.; Sarkar, S.; Maghera, J.; Rodrigues, E.; Carpenter, E. J.; Seth, S.; Ferrer Vinals, D.; Bennett, N. J.; Reddy, R.; Khalil, A.; Xue, X.; Bell, M. R.; Zheng, R. B.; Zhang, P.; Nycholat, C.; Bailey, J. J.; Ling, C.-C.; Lowary, T. L.; Paulson, J. C.; Macauley, M. S.; Derda, R., “Genetically encoded multivalent liquid glycan array displayed on M13 bacteriophage,” *Nat. Chem. Biol.* **2021**, 17 (7), 806-816. Copyright © Nature Publishing Group.

Chapter 2 is based on non-published work initiated by R. Derda. I was responsible for the construction and characterization of mLiGA. I performed the optimization of screening conditions for anti-A and anti-B antibodies, functional validation of mLiGA on anti-A and anti-B antibodies using ELISA and PFU assay, organized and analyzed the data. N. Bennett constructed the SDB M13 phage library. N. Bennett, S. Sarkar, J. Maghera and M. Sojitra isolated SDB clones. I was responsible for amplifying the clones previously isolated in the lab. S. Sarkar monitored phage viability in CuAAC reaction and performed regioselectivity experiments on phage pVIII. ABO glycans were synthesized and characterized by T. Lowary’s group at the University of Alberta. Anti-A and anti-B antibodies produced by Immucor Inc. were donated by L. West group at the University of Alberta. **Figure 2-3** is adapted from the publication Sojitra, M.; Sarkar, S.; Maghera, J.; Rodrigues, E.; Carpenter, E. J.; Seth, S.; Ferrer Vinals, D.; Bennett, N. J.; Reddy, R.; Khalil,

A.; Xue, X.; Bell, M. R.; Zheng, R. B.; Zhang, P.; Nycholat, C.; Bailey, J. J.; Ling, C.-C.; Lowary, T. L.; Paulson, J. C.; Macauley, M. S.; Derda, R., “Genetically encoded multivalent liquid glycan array displayed on M13 bacteriophage,” *Nat. Chem. Biol.* 2021, 17 (7), 806-816.

Chapter 3 is based on non-published work initiated by R. Derda. I was responsible for performing all the binding experiments on anti-A and anti-B antibodies. Two-step semi-nested PCR method was developed with help from A. Atrazhev. LiGA-65 was constructed by M. Sojitra. I analyzed the sequencing data with R script written by E. Carpenter. I calculated the Z'-score for the binding assays.

Chapter 4 is based on non-published work initiated by R. Derda. Serum samples were obtained from the L. West group at the University of Alberta and the J. Gildersleeve group at NIH. I was responsible for designing screening assay for human serum, optimizing screening conditions and performing all the binding experiments on serum using mLiGA. Binding experiments on glass and Luminex bead glycan arrays were performed by the L. West group. Binding experiments on neoglycoprotein array was done by the J. Gildersleeve group. Sequencing results were analyzed using the R script written by E. Carpenter. I wrote the R script to generate heatmaps. I was responsible for comparing ABO LiGA data with glass, Luminex and neoglycoprotein arrays.

Acknowledgements

Pursuing this degree during a global pandemic came with its own unique challenges. There are many people I need to thank for making this possible.

First, I would like to thank my parents who have always supported me in all my endeavors. I want to thank them for being the best support system, albeit virtually for the past 2 years. I also want to thank my sister Shweta for her unconditional love and care for me.

I want to thank my advisor Prof. Ratmir Derda for allowing me to work on a research topic that interested me the most. Ratmir's unwavering passion for science and research inspires me to be a better scientist everyday. Even though it's work in progress, from strong documentation of experiments in the lab to succinct scientific writing, I am thankful to him for teaching me all aspects of good scientific practice. I also want to thank him for his financial support, mentorship, and guidance, especially during COVID-19 that helped keep my research on track.

I would like to thank Prof. Robert Campbell and Prof. Matthew Macauley for being my committee members and providing constructive advice throughout my program. I am thankful to Dr. Campbell for giving me the opportunity to develop my leadership and mentorship skills by recruiting me as the graduate student mentor for the UAlberta iGEM team. I also want to thank Prof. Warren Wakarchuk for graciously agreeing to be my arm's length examiner.

I am thankful to Anne Halpin, Jean Pearcey, Tess Ellis, Bruce Motyka from the West lab at the University of Alberta and Prof. Lori West for the generous advice on my project and donating serum samples. I am also thankful to Prof. Jeff Gildersleeve at the NIH for his constructive advice on using glycan arrays for serological assays and donating serum samples.

I would like to thank Derda Lab colleagues who turned into friends over the years. Most of all, I want to thank Arunika Ekanayake who made surviving graduate school easier. She has been a pillar of support for me in the lab (and otherwise) for the past two years. She is an excellent scientist and one of the humblest and kind persons I have ever met. Aside from science, my discussions with Arunika on inclusivity, mental health and women in science inspires me to educate myself on these topics and strive to create a positive work environment.

I am thankful to Mirat Sojitra for all the intellectual (and other) discussions, training me on phage techniques and teaching me everything about LiGA. Mirat introduced me to many new topics in Glycobiology (and other fields of science) which motivated me to be inquisitive and read a variety of scientific articles. I want to thank Jeffrey Wong for his assistance in the lab and helping with formatting of my thesis. I am thankful to Eric Carpenter for patiently answering all my questions regarding data analysis and helping me write R scripts. I also want to thank Kejia Yan and Ganesh Sable for being excellent lab mates.

Last but not the least, I want to thank Alison Chou, Ryan Snitynsky and Elizabeth Nanak from GlycoNet and members of the GlycoNet Trainee Association Exec Committee 2019-2021 for giving me the opportunities that contributed to my professional development. I also

want to thank my colleagues at the After iGEM Science Communication Steering Group (Nannan, Hassnain, Marissa) for our weekly discussions on communicating science to a broader audience and developing scicomm initiatives for the next generation of scientists. Pursuing interests outside the lab really helped me recharge and explore different non-academic career paths.

Table of Contents

Abstract	ii
Preface	iv
Acknowledgements	vii
List of figures	xiv
Abbreviations	xvi
Chapter 1: Glycan arrays as a tool to profile anti-glycan antibodies in the human serum	1
1.1. Glycans and the immune system	1
1.2. The ABCs of ABO antibodies in human serum	3
1.3. Historical review of the development of glycan arrays for the detection of anti-glycan antibodies in the human serum	6
1.3.1. Enzyme linked immunosorbent assay (ELISA)	8
1.3.2. Printed glycan array	8
1.3.3. PGA for antibody profiling of healthy serum.....	9
1.3.4. Infectious disease diagnosis	10
1.3.4.1. Host-microbe interactions and autoimmune disorders	10
1.3.4.2. Development of carbohydrate vaccines.....	12
1.3.4.3. The neoglycoprotein array.....	13
1.3.5. Bead based arrays	14
1.3.5.4. Luminex bead array	15
1.3.6. Next generation sequencing (NGS) based arrays	17
1.3.6.5. DNA-encoded glycan array	17
1.3.7. Phage-display based glycan microarray	20
1.3.8. Genetically encoded multivalent phage-display enabled liquid glycan array	22
1.4. Thesis overview	24
Chapter 2: Synthesis and validation of ABO LiGA array	26
2.1. Introduction	26
2.1.1. Robustness and reproducibility of the assay.....	27
2.1.2. Conjugation strategies	30
2.1.3. Glycan presentation	31

2.2. Quality control of glycophage conjugates and components of LiGA.....	32
2.2.1. Enzyme Linked Immunosorbent Assay (ELISA).....	32
2.2.2. Plaque forming Units (PFU) and PFU assays	33
2.3. Results and discussion	34
2.3.1. Construction and characterization of MSDB phage-ABO glycoconjugates ..	34
2.3.2. Validation of ABO LiGA using ELISA	39
2.3.3. Functional validation of MSDB phage-ABO glycoconjugates using PFU assay	41
2.4. Conclusion	42
2.5. Materials and methods	43
2.5.1. Construction of silent distal barcode (SDB) M13 phage library	43
2.5.2. SDB clone isolation and amplification.....	46
2.5.3. Monitoring phage viability in CuAAC reaction.....	47
2.5.4. Ligation of ABO glycans to M13 phage	48
2.5.5. Characterization of phage-ABO glycoconjugates by MALDI-TOF MS	48
2.5.6. Binding of phage-ABO glycoconjugates to anti-A and anti-B measured by .	49
2.5.7. Construction of Multiplex Silent Distal Barcodes (MSDBs)	50
2.5.8. Binding of ABO glycosylated clones to different dilutions of anti-A and anti-B antibodies measured by ELISA.....	50
2.5.9. Dose titration study of ABO glycosylated clones measured by ELISA.....	51
Chapter 3: Validation of MSDB ABO LiGA by deep sequencing	61
3.1. Introduction.....	61
3.1.1. Sequencing technologies	61
3.1.1.1. Sanger sequencing	61
3.1.1.2. Roche 454 sequencing	63
3.1.1.3. Ion Torrent Sequencing	64
3.1.1.4. Illumina deep sequencing	65
3.1.2. DNA encoding strategies.....	67
3.1.3. Errors and biases in DNA sequencing and NGS screening.....	68
3.1.4. PCR bias	70
3.2. Results and discussion	72

3.2.1.	Design and validation of a semi-nested PCR for amplification of MSDB library	72
3.2.2.	Comparison of sequencing performance of ABO MSDB LiGA amplified by one vs. two-step PCR	75
3.2.3.	Screening m-LiGA on anti-A and anti-B antibodies	77
3.2.4.	Screening a complex LiGA library on anti-A and anti-B antibodies	79
3.2.5.	Avidity based response of m-LiGA	81
3.3.	Conclusion	83
3.4.	Materials and methods	83
3.4.1.	Binding assay of mLiGA on 96 well plates	83
3.4.2.	Binding assay of mLiGA on Protein L beads	84
3.4.3.	One-step PCR Protocol and Illumina sequencing	84
3.4.4.	Two-step semi-nested PCR	85
3.4.5.	Processing of Illumina data	86
Chapter 4:	Testing of MSDB ABO LiGA on human serum samples	88
4.1.	Introduction	88
4.2.	Experimental design	90
4.3.	Results and discussion	91
4.3.1.	Detection of blood group antibodies in human serum tested by MSDB ABO LiGA	91
4.3.2.	Detection of blood group antibodies in human serum tested by ABO LiGA mixed with LiGA-65	92
4.3.3.	Comparison of MSDB ABO LiGA with glass, neoglycoprotein and Luminex bead array	95
4.4.	Conclusion	97
4.5.	Materials and methods	98
4.5.1.	Serum samples	98
4.5.2.	Microarray fabrication and binding assay	98
4.5.2.1.	ABO glass array	98
4.5.2.2.	Neoglycoprotein array	100
4.5.2.3.	ABO Luminex array	102
4.5.2.4.	MSDB ABO LiGA array	102

4.5.3. PCR protocol and Illumina sequencing	103
4.5.4. Generation of heatmaps	103
References:	105

List of figures

Figure 1-1. Illustration of PAMPs and SAMPs.....	2
Figure 1-2. Subtypes of ABO glycan.....	5
Figure 1-3. History of glycan arrays for serological profiling.....	6
Figure 1-4. Design and workflow of Printed Glycan Array.....	14
Figure 1-5. Design and workflow of Luminex bead array.....	16
Figure 1-6. Design and workflow of DNA-encoded glycan array.....	18
Figure 1-7. DNA-encoded glycan array on multivalent headpiece.....	19
Figure 1-8. Illustration of glycophage array.....	21
Figure 1-9. Illustration of Liquid Glycan Array (LiGA) and a typical LiGA workflow.....	22
Figure 1-10. Heatmap showing binding of lectins to LiGA.....	23
Figure 1-11. Screening of LiGA on anti-A and anti-B antibodies.....	24
Figure 2-1. Scheme of silent double barcode (SDB) and multiplex silent barcode (MSDB).....	29
Figure 2-2. Structures of ABO glycans before and after conjugation to phage.....	35
Figure 2-3. Toxicity of chemical glycosylation to phage.....	36
Figure 2-4. Synthesis and characterization of ABO LiGA.....	37
Figure 2-5. MALDI-TOF spectra of ABO LiGA conjugated to ABO glycans with varying densities.....	38
Figure 2-6. Functional validation of ABO LiGA using ELISA.....	40
Figure 2-7. Functional validation of ABO LiGA using PFU assay.....	42

Figure 3-1. Illustration of PCR Stochastic Bias.....	73
Figure 3-2. Scheme of qPCR and semi-nested PCR.....	74
Figure 3-3. Serial dilution qPCR of m-LiGA.....	75
Figure 3-4. Comparison between one-step and two-step PCR for low copy number phage clones.....	76
Figure 3-5. Comparison of Illumina deep sequencing results between one-step and two-step PCR.....	78
Figure 3-6. Deep sequencing analysis and Z'-score calculation for m-LiGA screening on plates and beads.....	80
Figure 3-7. Screening ABO m-LiGA with LiGA-65 on anti-A and anti-B antibodies.....	82
Figure 3-8. Density scan of ABO m-LiGA on anti-A and anti-B antibodies.....	92
Figure 4-1. Heatmap of serum IgM and IgG binding response to ABO m-LiGA.....	93
Figure 4-2. Heatmap of serum IgM signals to ABO m-LiGA mixed with LiGA-65.....	94
Figure 4-3. Comparison of IgM signals to ABO mLiGA with ABO m-LiGA+LiGA-65.....	95
Figure 4-4. Comparison of serum response to m-LiGA with glass, neoglycoprotein and Luminex glycan array.....	96

List of Abbreviations

PAMP	Pathogen Associated Molecular Pattern
DAMP	Danger Associated Molecular Pattern
SAMP	Self Associated Molecular Pattern
ELISA	Enzyme Linked Immunosorbent Assay
IgG	Immunoglobulin G
IgM	Immunoglobulin M
PGA	Printed Glycan Array
MGM	Printed Microbial Glycan Microarray
BSA	Bovine Serum Albumin
DNA	Deoxyribonucleic Acid
DEL	DNA Encoded Library
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
NGS	Next Generation Sequencing
<i>E. coli</i>	<i>Escherichia Coli</i>
LiGA	Liquid Glycan Array
m-LiGA	Multiplexed Liquid Glycan Array
SDB	Silent Double Barcode
MSDB	Multiplexed Silent Double Barcode
PFU	Plaque Forming Units

CuAAC	Copper Catalyzed Azide-Alkyne Cycloaddition
SPAAC	Strain Promoted Azide-Alkyne Cycloaddition
DBCO	Dibenzocyclooctyne
MALDI-TOF	Matrix Assisted Laser Desorption Ionization – Time of Flight
Tri-AN3	Trisaccharide Antigen A
Tri-BN3	Trisaccharide Antigen B
Di-N3	Disaccharide Antigen H
FC	Fold Change
MFI	Mean Fluorescence Intensity

Chapter 1: Glycan arrays as a tool to profile anti-glycan antibodies in the human serum

1.1. Glycans and the immune system

Along with nucleic acids, proteins, and lipids, glycans are one of the four fundamental classes of molecules that make up all living systems.¹ Every cell in the human body is decorated with a thick layer of glycans called the glycocalyx. Since the immune system is tasked with surveying the body for “danger”, the glycocalyx is routinely engaged when an immune cell contacts another cell or any component of its environment.² In such situations, glycans help modulate the behaviour of immune cells and play a major role in the discrimination of “self” and “non-self”. For the immune system to elicit a response against invading pathogens, it needs to identify “foreign” bodies as “non-self”. Several discoveries since the 1990’s have demonstrated that the human body uses “pathogen-associated molecular patterns” (PAMPs) and “danger associated molecular patterns” (DAMPs) to respond to foreign antigenic invasion. Many of these PAMPs and DAMPs are glycoconjugates on the bacterial cell surface and their immune receptors are lectins on the host cell surface or anti-glycan antibodies.^{3,4} For example, glycan structures on eosinophils are known to be modified during an event of inflammation such as bronchial asthma.⁵ ⁶ These structural changes on the glycocalyx lead to the formation of DAMPs which stimulates the production of anti-glycan antibodies.⁷ Along with PAMPs and DAMPs, “self-associated molecular patterns” (SAMPs) are also known to be equally important in preventing robust immune responses to non-pathogenic external stimuli.

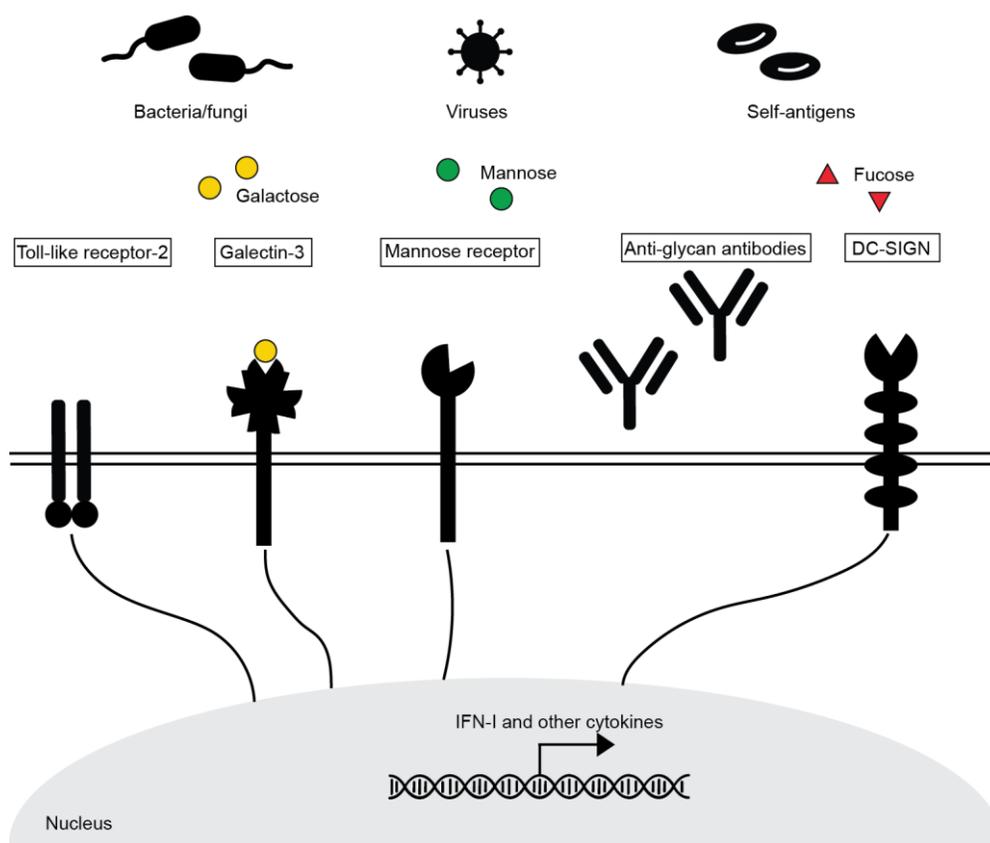


Figure 1-1: Illustration of Pathogen associated molecular patterns (PAMPs).and Self associated molecular patterns (SAMPs)

Glycans from pathogens and self-antigens act as PAMPs and SAMPs respectively and elicit an immune response in hosts by binding to glycan binding proteins and antibodies

Over the years, several pathogenic microbes have been discovered to mimic SAMPs to evade or dampen the host's immune response.^{8,9} The mechanisms discussed above generate a diverse pool and high levels of anti-glycan antibodies in human serum at different stages and events of life. Anti-glycan antibodies are essential for both fundamental research as well as therapeutic and diagnostic purposes. Since direct detection of glycans from complex biological samples is difficult, most studies rely on antibodies to detect and monitor expression of carbohydrate antigens using techniques such as immunohistochemical staining, Western blotting, and ELISA.¹⁰

More recently, anti-glycan antibodies are also used in clinical applications. For example, an anti-GD2 antibody Unituxin was granted FDA approval in March 2015 and ever since it is an important part of line-of-treatment of neuroblastoma.^{11,12} Several other antiglycan antibodies are in clinical trials for cancer therapy, such as antibodies to GD3, GM2, fucosyl-GM1, and Lewis Y.¹³⁻²¹ Examples of antiglycan antibodies used for diagnostic applications include antibodies that target Sialyl Lewis A (CA19.9) currently used in detection and monitoring of advanced pancreatic cancer.²²⁻²⁴

1.2. The ABCs of ABO antibodies in human serum

In humans, the most well-known examples of anti-glycan antibodies are those that bind A, B and O (H) antigens. When Karl Landsteiner and colleagues discovered the ABO blood group system in the 20th century, they were not aware of the underlying glycan basis. Their work allowed humans to be categorized in different groups based on serum substances that agglutinated antigens on red blood cells. We now know that the antigens are glycans and the serum constituents are anti-blood group antibodies. A, B and O (H) blood group antigens are oligosaccharides expressed as glycoproteins or glycolipids on the surface of cells and tissues. The monosaccharides are sequentially added by glycosyltransferases encoded by the ABO blood group locus to different precursor chains – types I-VI (Figure 1.1). At least four of these six subtypes are known to be present in humans.²⁵⁻²⁷ The variations in precursor structures that carry ABH antigens create unique antigen epitopes.^{26,28,29} West and co-workers recently demonstrated that subtype antigens are expressed differentially between erythrocytes and tissues or organs which is important to consider especially during an organ transplantation.³⁰

Early in life, an individual's immune system generates IgM class antibodies directed against the ABO oligosaccharide antigens that are absent from their red blood cells. The antibodies likely represent an immune response to oligosaccharide antigens synthesized by bacterial and fungal organisms in the environment, and whose structures are similar or identical to those of the A and B blood group molecules. For instance, type-O individuals do not make A and B antigens and therefore maintain relatively high titers of circulating IgM antibodies that react with A and B blood group molecules. Similarly, blood group B individuals maintain circulating IgM class anti-A antibodies, but they do not make antibodies against the blood group B antigen, which is, in these individuals, a "self" antigen. Sera taken from blood type A individuals usually contain high titers of IgM class anti-B antibodies, but not anti-A antibodies. Finally, individuals with the AB blood group do not make either anti-A or anti-B IgM antibodies. Anti-H antibodies are not made in most individuals because a substantial fraction of the H structures is not converted to A or B antigens, even in those with a functional A or B transferase allele.

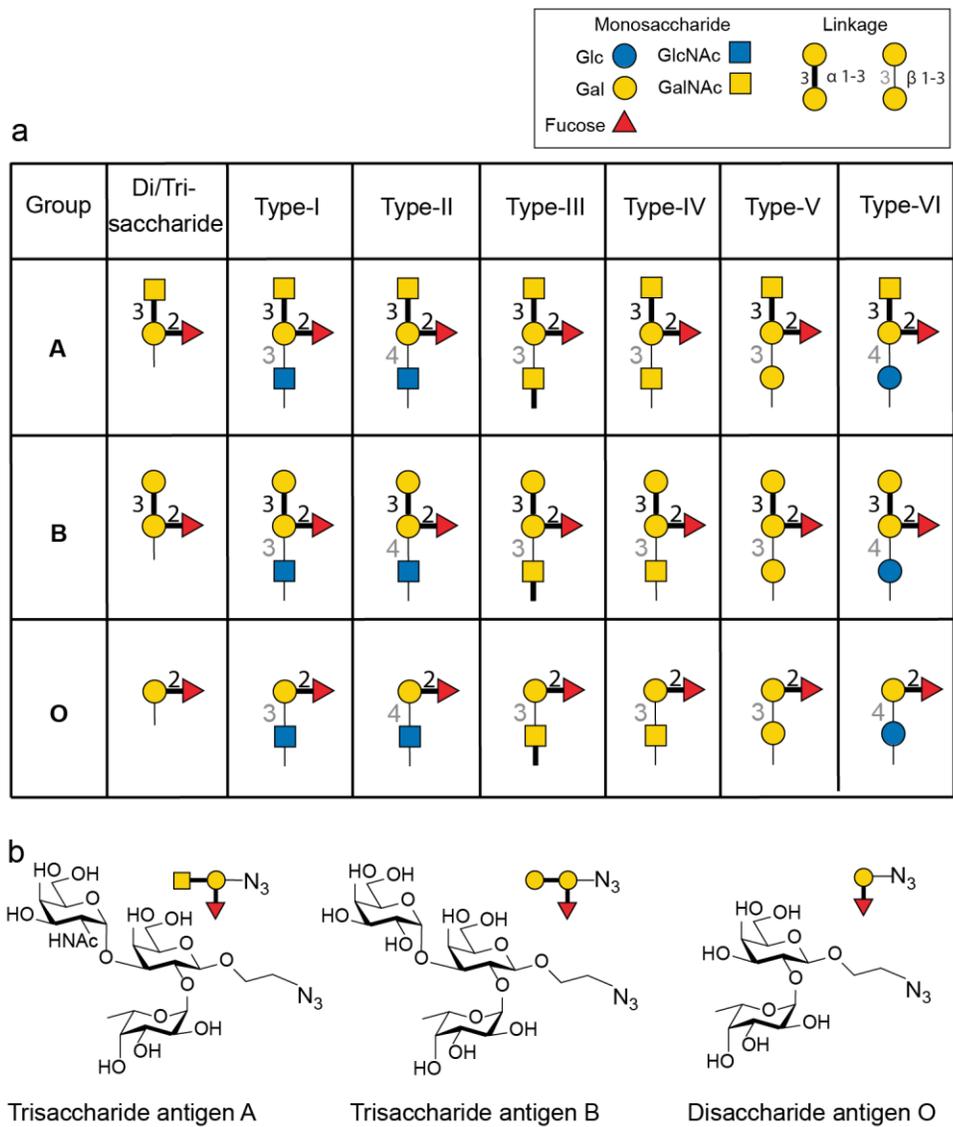


Figure 1-2: The ABO blood group glycans **a**, Six different subtypes of A, B and O (H) antigens classified based on the precursor chains **b**, Chemical structures of ABO blood group determinants

1.3. Historical review of the development of glycan arrays for the detection of anti-glycan antibodies in the human serum

Previously, glycan-lectin or glycan-antibody interactions were investigated by indirect methods such as hemagglutination inhibition and inhibition of precipitation assays. These assays were performed by using free carbohydrates as inhibitors of binding between glycan-presenting substances, such as red blood cells and glycan binding proteins such as anti-glycan antibodies. These assays led to the discovery of ABO blood group antigens by Morgan and Watkins, and Kabat and colleagues between the 1950's and 1970's.³¹

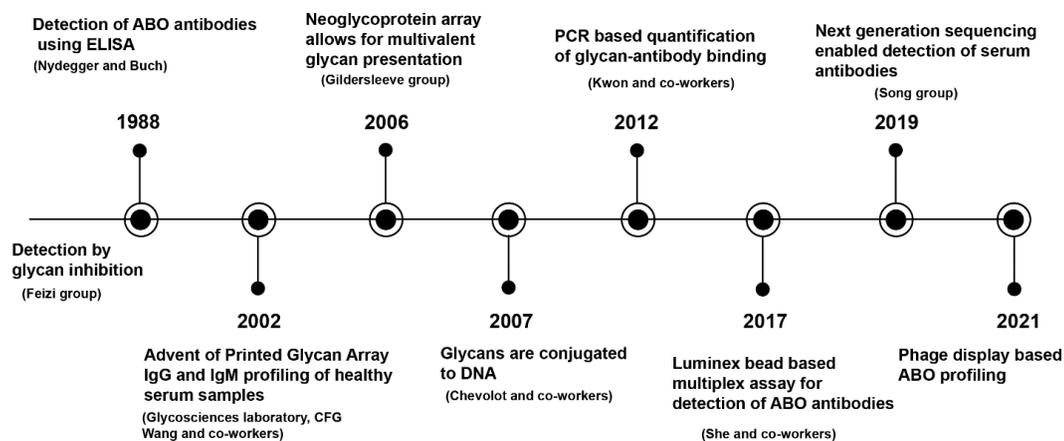


Figure 1-3: Scheme of the history of glycan arrays developed for serum antibody profiling

The method of measuring glycan binding by inhibition was later replaced by solid phase assays where anti-glycan antibodies were radio-labeled, and the assays were performed on thin layer chromatography (TLC). Some seminal discoveries that used this technique include identification of stage specific embryonic antigens (SSEA)^{32, 33}, the pancreatic cancer associated biomarker CA19-9³⁴ and GM-1 as cholera toxin ligand³⁵. In early 2000s, this method was replaced by

enzyme linked assays where secondary antibodies conjugated to an enzyme or fluorophore were used to detect anti-glycan antibodies. This assay format formed the basis of glycan microarrays where multiple glycans could be probed at the same time.

Glycan arrays can be made from individual, synthetically made components or from natural sources like cells, pathogens, tissues, milk and urine.³⁶ Glycan arrays made from natural sources complements arrays made from chemically prepared glycans, especially where the actual biological target of a glycan is unknown. Smith and Cummings pioneered a “shotgun microarray” approach where glycolipids were harvested from human erythrocytes and attached to glass surface for interrogation by GBPs and carbohydrate binding antibodies in human serum.³⁷ Since then, shotgun glycan arrays have been used to identify biologically relevant glycans such as brain glycosphingolipids³⁸, endogenous receptors for influenza viruses³⁹, O-glycans on *Schistosoma Mansoni* cell surface⁴⁰, glycans on ovarian cancer cell surface³⁷, and milk oligosaccharides⁴¹.

In this chapter, I review various types of glycan arrays that were developed using glycans that are “definitive synthetics” or “natural isolates” and are currently being used for profiling anti-glycan antibodies in the human serum. The goal of this review is to not only disseminate knowledge about glycan arrays used in serological profiling but also share the data collected by scientists to the glycobiology community. Data from all publications discussed in each section is uploaded to github (<https://github.com/derdalab/Serum-glycan-array-data>) in a .csv format. There are 4 sections with 9, 40, 8, 3, 1 glycan array datasets each from ELISA, printed glycan array, Luminex bead glycan array, DNA-encoded glycan array and phage-display library respectively (a total of 61 array datasets).

1.3.1. Enzyme linked immunosorbent assay (ELISA)

The ELISA was developed in 1971 as an alternative to radioimmunoassay and first used to determine the levels of IgG antibodies in rabbit serum.^{42,43,44} Half a century later, it is still the gold standard for serological testing and is currently one of the most commonly used assays for serum testing of SARS-CoV-2 antibodies.^{45,46,47} In 1988, Buchs and Nydegger at the University of Bern isolated antigen A and B from horse gastric mucosa and coated them on 96 well polystyrene plates, incubated with human serum and probed it with alkaline phosphatase conjugated to polyclonal IgG or IgM secondary antibodies.⁴⁸ This was the first time when ELISA was used for the detection of IgG and IgM anti-A and anti-B blood group antibodies from the human serum. The same group reported an extension to the ABO ELISA where synthetic ABO trisaccharides conjugated to BSA were immobilized on to the plate. Serum titers of IgG and IgM ABO antibodies was determined in healthy serum samples using monoclonal subtype specific secondary antibodies and compared to the classical hemagglutination assay.⁴⁹ Over the years, several research groups have used ELISA as a diagnostic test to detect anti-glycan antibodies specific to diseases⁵⁰ such as inflammatory bowel disease (IBD)^{51, 52}, multiple sclerosis (MS)⁵³,⁵⁴ Crohn's disease (CD)^{51, 55}. The glycans that bind to these antibodies are laminarin, chitin, laminaribioside, chitobioside in IBD (ELISA.csv, A1 and A2), galactocerebroside and glycans with terminal glucose in MS (ELISA.csv, A3 and A4), and laminaribioside and chitobioside in CD (ELISA.csv, A5).

1.3.2. Printed glycan array

Since their invention, glycan microarrays printed on a solid surface have played a key role in unveiling the binding specificities of known glycan binding proteins, anti-glycan antibodies as

well as identifying novel ones. The Glycosciences Laboratory at the Imperial College London developed the world's first robotically printed glycan array (PGA) in 2002. Today, along with the Consortium of Functional Glycomics, it is one of the most widely used printed microarrays by the scientific community. The high-throughput nature of the PGA was first put to test by Wang and co-workers in 2002.⁵⁶ 48 polysaccharides were printed onto a glass slide and used to profile antibodies in 20 healthy human serum samples. The authors observed anti-glycan antibodies binding to specific glycans with a sample volume as low as 1 μ L. Since then, PGAs have emerged as an indispensable tool for not only studying the anti-glycome in normal sera⁵⁷ but also disease prognosis, diagnosis, risk prediction and monitoring the immune response.^{58, 59}

1.3.3. PGA for antibody profiling of healthy serum

Bovin and co-workers at the Russian Academy of Sciences were among the first ones to delve deeper into the anti-glycan repertoire of healthy individuals.⁶⁰ 106 normal serum samples were screened with a PGA composed of 205 glycan structures including blood group antigens, N- and O- linked glycans, tumor associated antigens and lipopolysaccharides (Printed Glycan Array.csv, A1). The sera were observed to interact with at least 50 normal human glycan motifs. Although all the serum samples were obtained from healthy volunteers with no previous history of a disease, the authors observed some variations in their antibody profiles. The serological profile of the same samples obtained by PGA was observed to be different than that when the antibody samples were processed by affinity purification of pooled sera. More recently, the same group studied the serum antibody profiles of healthy infants in their third, sixth and twelfth month of life.⁶¹ A representative PGA of 487 glycans consisting of fragments of N- and O- linked glycans, glycolipids and bacterial lipopolysaccharide was used to study the innate immunity in

the first year of life and the impact of nutrition type on antibody development. The total amount and diversity of IgG observed at 3 months of age corresponded to the maternal repertoire and decreased after 12 months time. In contrast, IgM levels elevated after the first 6 months of life and corresponded closely with an adult human serum IgM profile.

1.3.4. Infectious disease diagnosis

A 2021 study by the Seeberger group emphasized the importance of applying glycan microarrays to diagnostics and studying antibody development at different stages of an infectious disease.⁶² Glycan specific IgA, IgG and IgM antibodies were profiled in the serum of 3 SARS-CoV-2 patients (1 moderate, 2 mild disease) using 3 assays – i.) Full proteome peptide arrays ii.) ELISA iii.) PGAs and the results were compared to each other. Glycans such as $\text{Man}_2\text{GlcNAc}_2$ that correspond to the N-glycan core on the SARS-CoV-2 spike protein showed significant binding to the antibodies. An increase in binding to $\alpha 1\text{-}2\text{Man}_3$ glycan was observed in patients with mild infection (Printed Glycan Array.csv, A35). Even though this was the first time PGA was used to detect glycan specific SARS-CoV-2 antibodies, it is important to note that this study involved only 3 subjects and a more extensive study is needed to fully understand the role of anti-glycan antibodies in COVID-19.

1.3.4.1. Host-microbe interactions and autoimmune disorders

Microarray based technologies have enabled researchers to study and characterize autoantibodies generated during autoimmune diseases.⁶³⁻⁶⁵ Overexpression of galactocerebroside, galactolaminaribioside and galactochitobioside and other glycan moieties in brain tissues belonging to the glycolipid class are known to trigger autoantibody generation and cause Multiple

Sclerosis (MS). Wang and co-workers developed an integrated microarray consisting of proteins, lipids and carbohydrates to screen autoantibodies in the serum of mice with an experimental autoimmune disorder and cerebrospinal fluid of MS patients.⁶⁶

Glycans expressed on the surface of bacteria are known to elicit an immune response in humans by presentation of foreign glycans to immune receptors present on the host cell surface. Sometimes during an infection, bacteria start to synthesize glycans similar to those on the surface of host cells. Presentation of such structures to host immune cells was reported to cause an autoimmune response.⁶⁷ Printed Microbial Glycan Microarrays (MGM) were produced to understand the interaction of immune system with microbial glycans. Arrays that contain sufficiently large number of glycans can also yield insights into the structural motifs driving these interactions. Several groups have used MGM for serological diagnosis of distinct microbial infections caused by pathogens belonging to the *Klebsiella* class⁵⁶(A36), *E.coli*^{56, 68}(A36), *Pneumococcus*⁵⁶(A36), *Meningococcus*⁵⁶(A36), *Burkholderia*^{69, 70}(A37, A38) *Mycobacterium*⁷¹(A39), *Saccharomyces*⁷¹(A39), *Salmonella*^{68, 72}, *Streptococcus*⁷³ class and *Bacillus anthracis*⁷⁴, *Francisella tularensis*⁷⁰(A38) and *Haemophilus influenzae*⁵⁶(A36) (Numbers in the bracket indicate the sheet name in the Printed Glycan Array.csv file where corresponding PGA data can be found).

Glycosphingolipids (GSLs) expressed on several *Borrelia* correspond to that of humans and induce an autoimmune response. Smith and Cummings used a shotgun microarray approach to construct an array of GSLs to determine the levels of antibodies in the sera of patients affected with Lyme disease.⁷⁵ GSLs harvested from *Borrelia* cells were fractionated, characterized by MALDI-TOF/TOF and immobilized onto glass slides to generate a small focused PGA and used

to screen the sera of patients with Lyme disease. A strong IgG response to a disialylated ganglioside GD1-b lactone was observed.

1.3.4.2. Development of carbohydrate vaccines

PGA were instrumental in development of vaccines by making it possible to understand the interaction between carbohydrate antigens present on tumor and viral cell surfaces and antibodies in human serum. Stage specific embryonic antigen 3 and 4 (SSEA 3/4) and Globo-H are among the most widely investigated tumor associated carbohydrate antigens. Globo-H antigen found on epithelial cells are observed to be over expressed in various types of cancer.^{76, 77} After more than 30 years of research conducted out of labs based in the Scripps Research Institute in USA, Academia Sinica in Taiwan and RIKEN Institute in Japan, the Wong group developed a Globo-H cancer vaccine candidate that is currently in Phase III clinical trials.⁷⁸ Initial studies used PGAs to identify elevated serum antibody titers against Globo-H in breast cancer patients.⁷⁷ Preliminary studies of two Globo-H vaccine candidates in mice revealed elevated antibody levels for a glycoconjugate with CD1 as ligand, cross reacting material 197 (CRM197) as carrier protein and C34 as an adjuvant as opposed to keyhole limpet haemocyanin (KLH) and QS-21 used in phase II clinical trials.^{59, 79} Based on these results obtained using a PGA, phase III global trials with this new vaccine candidate are being conducted by the Taiwanese biopharma company OBI Pharma, Inc.

One of the key goals of HIV vaccine development is induction of broadly neutralizing monoclonal antibodies (bnmAbs).⁸⁰ N-glycans of the high mannose type conceal the protein epitope of the virus and PGA experiments identified many bnmAbs targeting the HIV glycan shield in patients with an active infection. 2G12 was the first reported anti-glycan bnmAb that showed a nanomolar

affinity to glycan residues with $\alpha(1-2)$ -mannosides.⁸¹ Some potent bnmAbs of the PGT series bound to high mannose glycans whereas the others were characterized as binding to complex-type N glycans.⁸²

1.3.4.3. The neoglycoprotein array

Shortly after the discovery of printed glycan array, in 2006 the Gildersleeve group at the NIH pioneered a novel glycan array technology known as the neoglycoprotein array. In contrast to traditional arrays, in which glycans are attached directly to the surface, production of neoglycoprotein array is a sequence of two steps. 20-50 copies of glycans were first covalently attached to a carrier protein bovine serum albumin (BSA).⁸³ BSA-glycan conjugates were then printed on epoxy slides. The multivalent presentation of glycans and the local spacing between the glycans is predictable and it is dictated by the spacing of glycans on BSA carrier. In contrast, the density or spacing of glycans directly attached to glass are often impossible to measure. There is an even distribution of about 60 amines on the surface of BSA, out of which 54 are solvent exposed and available for glycan conjugation.⁸⁴ This was the first time the local spacing of glycans could be predictably controlled and high-binding avidity could be achieved.

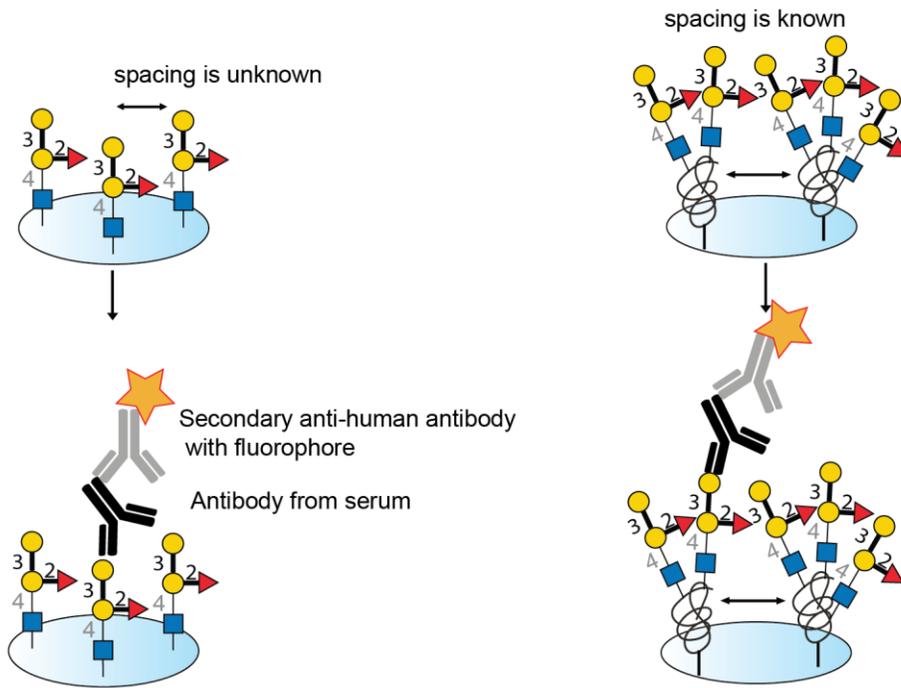
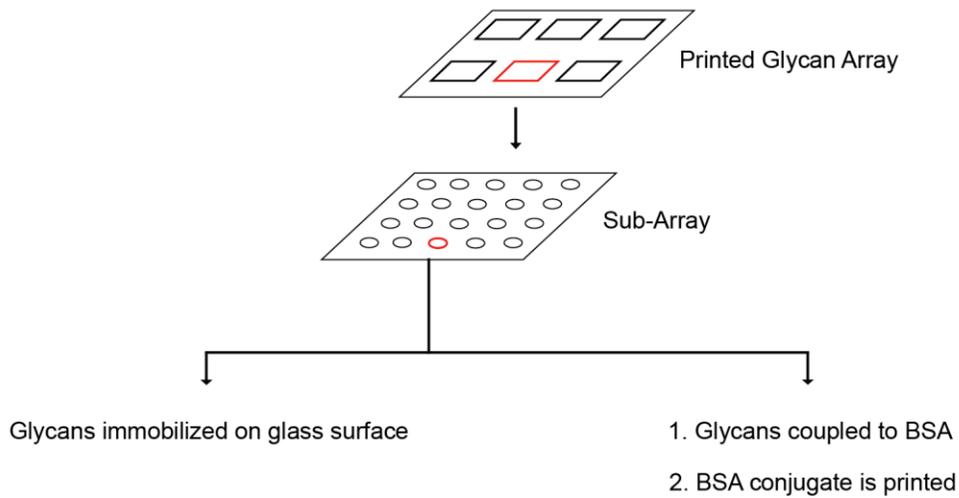


Figure 1-4: Design and workflow of printed glycan array. Glycans are covalently immobilized to a glass slide or conjugated to bovine serum albumin (BSA) and then attached to glass slides. Glycan binding proteins and antibodies are then detected using a secondary antibody conjugated to a fluorophore.

1.3.5. Bead based arrays

1.3.5.4. Luminex bead array

The discovery and development of new flow cytometric tools in the 1970s laid the groundwork for multiplex microsphere-based assays.⁸⁵ It was identified that fluorescently encoded microspheres coupled with an antigen can be used to measure antibody response using flow cytometry. Building on this technology, two decades later, Luminex Corporation launched the most widely used bead based multiplexing platform – the xMAP[®] microspheres. The xMAP system consists of 5.6 µm diameter polystyrene microspheres, each internally dyed with a fluorochrome that has a distinct emission profile.⁸⁶ The unique internal fluorescent “barcode” allows for simultaneous detection of multiple analytes in a complex mixture like the human serum. The Luminex platform was approved by FDA in 2008 for application in clinical laboratories and can screen large number of patient samples in one run.^{87, 88}

In 2011, Bovin and co-workers described the first Luminex-like assay with bead immobilized glycans to detect IgG and IgM antibodies to ABO antigens in human serum.⁸⁹ In 2017, the Gildersleeve group demonstrated the use of bona fide Luminex bead-based assay for the detection of anti-blood group antibodies in the serum of patients receiving prostate cancer vaccine (PROSTVAC-VF).⁹⁰ A previous study from the group suggested that measuring the serum IgM levels of blood-group A trisaccharide before receiving the vaccine could predict the overall survival rate of the cancer patients.⁹¹ The authors developed a multiplex assay by conjugating four neoglycoproteins – BG-A-tri antigen, alpha-Gal antigen, IgG, BSA to Luminex beads and measuring BG-A-tri IgM levels in human serum.⁹⁰

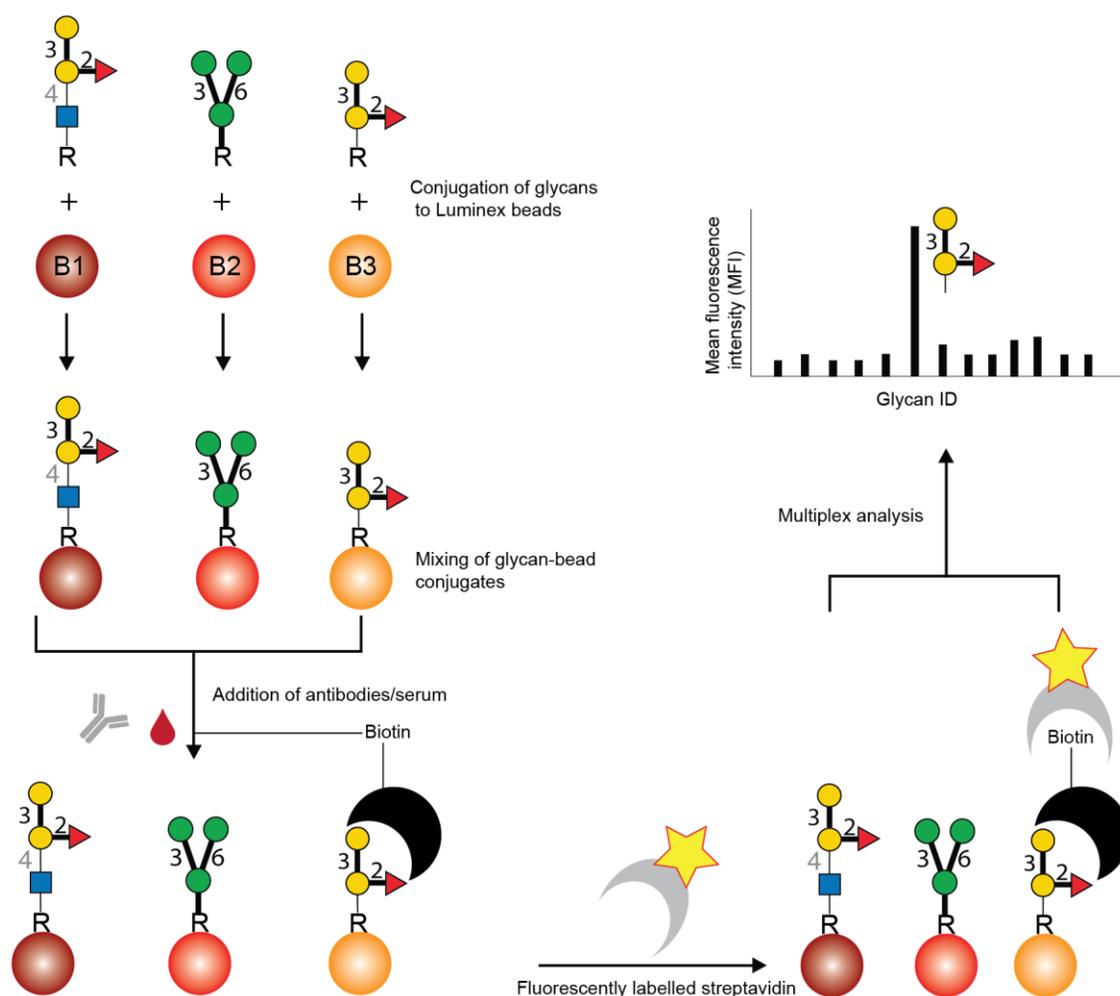


Figure 1-5: Design and workflow of Luminex glycan bead array. Each glycan was conjugated to one region specific Luminex bead, using 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride. After blocking the beads, the beads were probed with biotinylated anti-glycan antibodies and human serum. After washing, anti-glycan antibodies were detected using phycoerythrin labelled streptavidin (SAPE). The unbound SAPE was removed by washing, and beads were resuspended in wash buffer.

Within a year, She and co-workers developed a Luminex based Multiplex Glycan Bead Array (MGBA) consisting of 184 glycans (Figure 1-4). The authors demonstrated the utility of this high-throughput platform by screening 13 recombinant anti-glycan antibodies and 961 human serum samples for both naturally occurring IgG and IgM antibodies and anti-glycan antibody biomarkers

for ovarian cancer.⁹² More recently, the West group at the University of Alberta is developing a Luminex based ABO array consisting of different subtypes of the blood group antigens and profiling the IgG and IgM ABO antibodies in human plasma to evaluate the role of blood group subtypes in organ transplantation.⁹³

1.3.6. Next generation sequencing (NGS) based arrays

1.3.6.5. DNA-encoded glycan array

30 years ago, Lerner and Benner first proposed the concept of conjugating small molecule building blocks to DNA tags. This concept is now popularly known as DNA encoded libraries (DEL).^{94, 95} With the advent of high-throughput DNA sequencing technologies, pipettes were put into action and the idea of DEL transformed into a powerful drug discovery tool employed in pharmaceutical companies and academic labs.⁹⁶⁻¹⁰⁰ Mixing of these libraries with an immobilized target protein, amplification of the DNA barcode by polymerase chain reaction followed by next generation sequencing allowed for the detection of binder molecules.^{101, 102} Several research groups have previously reported different bioconjugation strategies for the development of DNA encoded glycan libraries and their application in vaccine-design. Chevolot and co-workers were the first ones to synthesize and use DNA-glycan conjugates in a microarray format to probe glycan-lectin interactions.¹⁰³

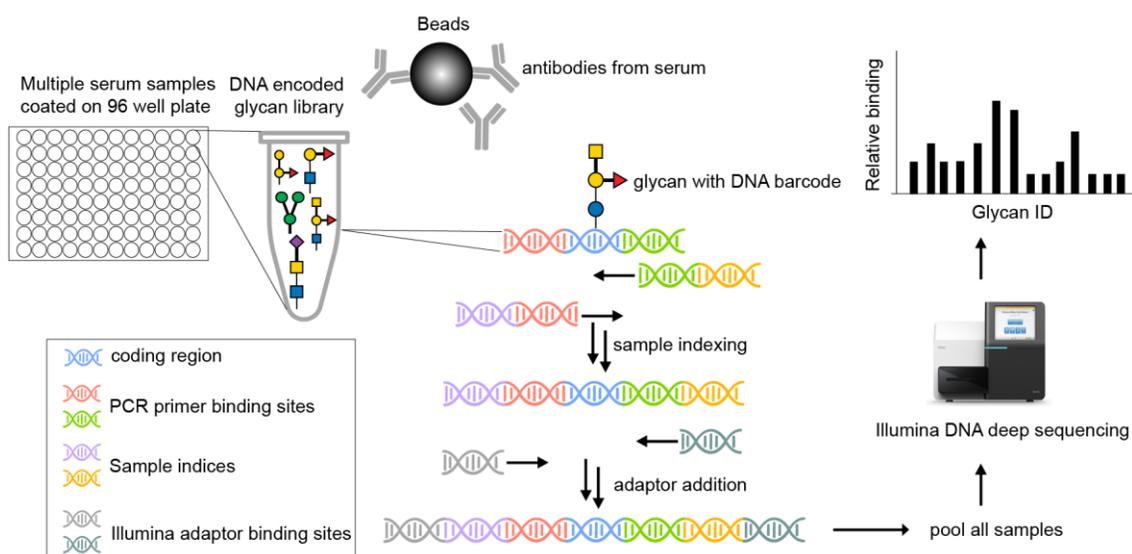


Figure 1-6: Design and workflow of next generation DNA encoded sequencing based array.^a Song and co-workers developed the next-generation glycan microarray (NGGM) based on artificial DNA coding of glycan structures. In this approach, a glycan library is presented as a mixture of glycans and glycoconjugates, each of which is coded with a unique oligonucleotide sequence (code). The glycan mixture is interrogated by anti-glycan antibodies followed by the separation of unbound coded glycans. The DNA sequences that identify individual bound glycans are quantitatively sequenced (decoded) by next-generation sequencing (NGS) technology, and copy numbers of the DNA codes represent relative binding specificities of corresponding glycan structures to anti-glycan antibodies.

In the 2000s, Linhardt and co-workers developed a singleplex assay based on quantitative polymerase chain reaction (qPCR) for detecting glycan-lectin interactions.^{104,105} However, it was only in 2019 that Song and co-workers successfully leveraged the power of NGS to develop the first DNA-glycan multiplex assay called Next Generation Glycan Microarray (NGGM) (Figure 1-5).¹⁰⁶

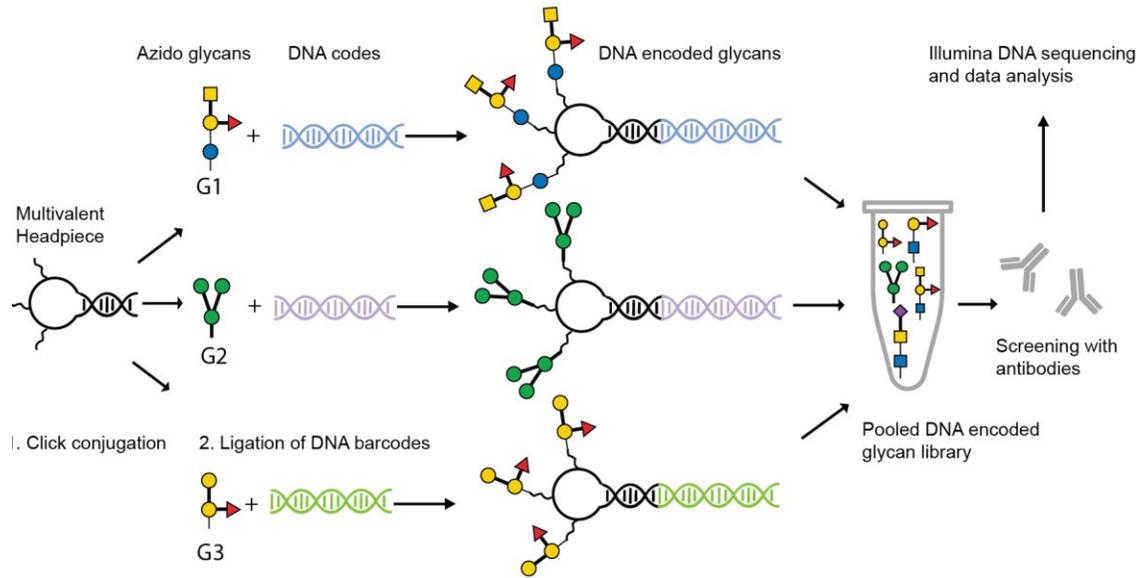


Figure 1-7: DNA encoded glycan array on multivalent headpiece Peng Wang and co-workers developed a DNA-encoded glycan array using a unique headpiece conjugation-code ligation (HCCL) strategy where single or multiple copies of glycans were attached to a headpiece linked to a unique DNA barcode. A library of 50 glycan epitopes was screened against anti-glycan antibodies and bound glycans were decoded by NGS.

Among other glycans, the NGGM array consisted of different types of human blood group antigens and their specificities were validated against four commercial blood group antibodies – anti-A, anti-B, anti-Le^x and anti-Le^y. A 2020 pre-print by Wang and co-workers reported the development of a novel DNA encoded glycan library (DEGL) that allows for both monovalent and multivalent display of glycans on unique DNA barcodes (Figure 1-6). This DEGL consisted of glycans of biological importance such as blood group antigens, gangliosides, 2-3 and 2-6 linked sialic acids and were tested against both anti-glycan IgG and IgM antibodies – VK9 (mouse, IgG binds to globoside H), anti-A (mouse, IgM binds to blood group antigen A) among others.¹⁰⁷ However, it is important to note that none of these NGS DNA-glycan arrays were used to profile anti-glycan antibodies in a more complex environment like the human serum.

1.3.7. Phage-display based glycan microarray

Since its inception in 1985, phage display has primarily been used for discovery of ligands and selection of ligand-binding proteins. Several examples of the utility of such “liquid arrays” in serodiagnostic applications are: Displaying proteins such as antibodies on phage for discovery of antibodies that detect leukocyte surface antigens¹⁰⁸. Display of libraries of peptides enabled discovery of short peptides that bind to antibodies in sera of patients with cancer.¹⁰⁹ cDNA libraries has been used for the diagnosis of lung cancer.¹¹⁰

If libraries of glycans could be displayed on phage and encoded in its genome, the concept of phage display could be extended from protein antigens to glycan antigens.

So far, only 3 groups have reported the development of glycan functionalized phage. DeElisa¹¹¹ and Aebi¹¹² groups developed a “glycophage” microarray based on M13 phagemid particles by cloning the glycan biosynthesis machinery into *E. coli* and by displaying the glycosylation acceptor peptide sequence on the pIII protein of phagemid (Figure 1-7). Both groups observed that only 1 in 100 phagemid particles carried a glycan biosynthesized by the bacterial host. The displayed glycopeptides were often truncated. DeLisa and co-workers solved the latter problem by employing truncated pIII instead of full-length pIII.¹¹³ However, even in the optimized construct, the authors could not quantify the density of glycans on the different glycophages due to the low level of phage glycosylation.

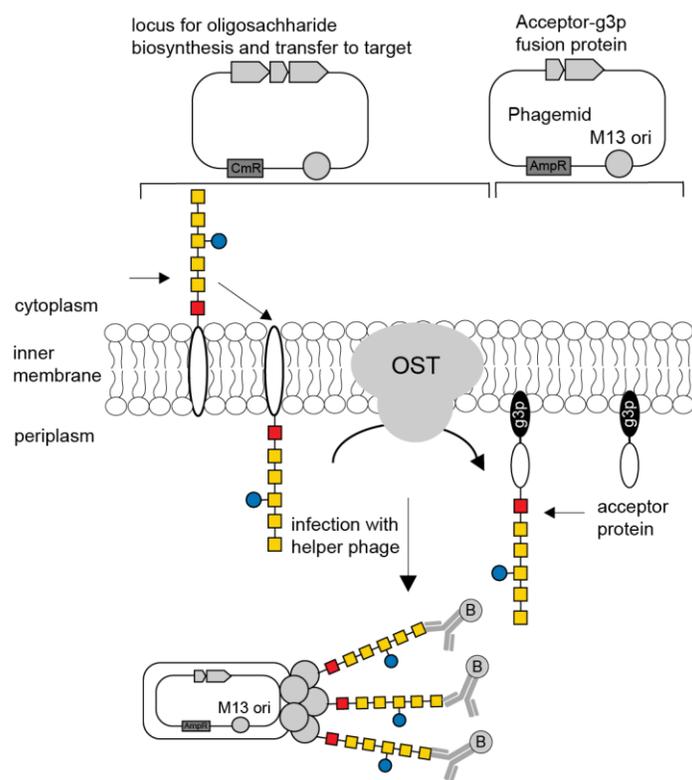


Figure 1-8. Illustration of glycoarray Phage particles were engineered to display the glycan present on the surface of *C. jejuni* and the glycoarray constructs were validated by screening of *C. jejuni* and *P. aeruginosa* glycan-specific antibodies in the human serum.

Another critical limitation of this approach is that it can only be employed for monovalent display of glycans. The DeLisa group demonstrated the utility of the glycoarray display for antibody testing by probing with serum antibodies against the *P. aeruginosa* O-11 O-antigen.

The authors used a detection system like that of traditional PGA assays by employing a fluorescently labelled secondary antibody and observed strong binding to glycoarrays displaying the O-11 antigen but not to any other glycoarrays or control glycans.

1.3.8. Genetically encoded multivalent phage-display enabled liquid glycan array

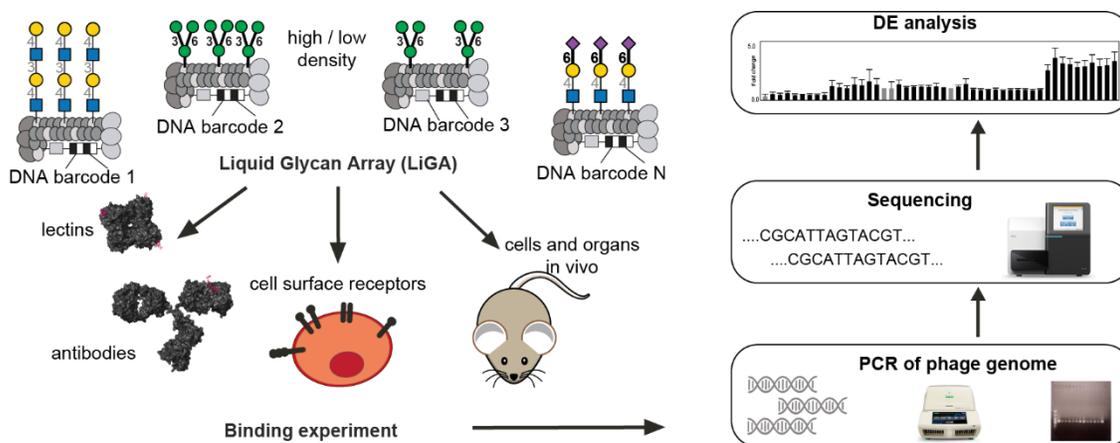


Figure 1-9: Illustration of Liquid Glycan Array (LiGA) and experimental workflow of a typical LiGA experiment

The genetically encoded multivalent glycan array on M13 phage was developed and recently published by our group.¹¹⁴ Unlike the ‘glycophage display’ system that employs the biosynthesis of glycans, LiGA decouples DNA encoding and glycan display from biosynthesis. Chemical manufacturing of LiGA components allows the repurposing of many chemical approaches previously employed in the construction of traditional PGAs. LiGA is built on filamentous M13 phage particles with silent DNA barcodes inside the phage genome. The authors chemically conjugated glycans to a subset of ~2700 copies of major coat protein pVIII to produce a multivalent display of ~150-1500 copies of glycans. LiGA was used to screen lectins, cells and *in-vivo*. Figure 1-9 adapted from the publication shows the summary of results obtained by screening LiGA with different glycan binding proteins.

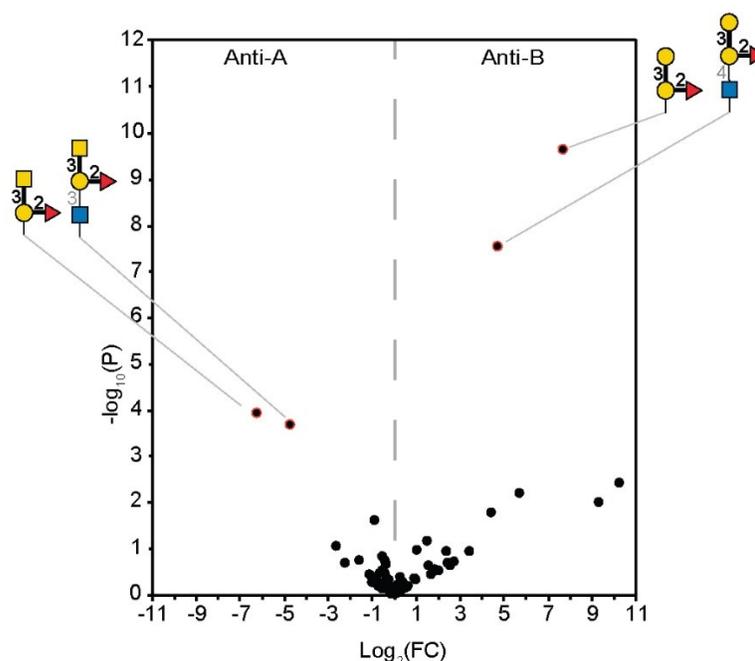


Figure 1-11: Screening of LiGA on anti-A and anti-B antibodies. Anti-A IgM and anti-B IgM were coated on polystyrene wells and binding of LiGA was performed (n = 5 and 6 independent biological replicates). The data was analyzed using EdgeR differential enrichment analysis using anti-A and anti-B as two sets. The volcano plot describes fold change (FC) difference and p-value of glycans enriched on anti-A (FC < 0) and anti-B (FC > 0) antibodies. Red circles denote four DNA barcodes that exhibited enrichment with FDR < 0.05. The structures of the glycans associated with these barcodes are known blood group glycans targeted by anti-A and anti-B antibodies.

1.4. Thesis overview

In my project, I demonstrated the utility of LiGA for specific detection carbohydrate binding antibodies in human serum. Application of LiGA to serological samples was not shown in the previous publication from our lab. In chapter 2, I will discuss the design and construction of a multiplex liquid glycan array (mLiGA) conjugated to the 3 blood group antigens – A, B and O glycans and its validation using ELISA and PFU assay. In chapter 3, I will discuss the functional performance of mLiGA by PCR and Next generation sequencing technology (NGS) and

calculation of a Z' -factor to evaluate the robustness of mLiGA assays. In chapter 4, I will show the application of mLiGA to specifically detect blood group antibodies from human serum samples and comparison with other existing platforms such as neoglycoprotein array and Luminex bead array.

Chapter 2: Synthesis and validation of ABO LiGA array

2.1. Introduction

In this chapter, I describe the construction of an ABO liquid glycan array (ABO LiGA) for detection of specific anti-glycan antibodies in humans. As described in chapter 1, there are ~70 publications from 33 research groups that already demonstrated the utility of ELISA arrays, glass microarrays and bead-based arrays for profiling antibodies in human serum. ELISA-style assays based on micro-well plates are not amenable to large-scale multiplexing and miniaturization. Glass-based glycan arrays can profile 100's of glycans with millimolar sample volume but they require specialized microarray reader instruments that are not available in clinical diagnostic laboratories. The QC of printed array or components is not possible, with one exception¹¹⁵ (MALDI-based QC of the glycan-conjugated BSA in neoglycoprotein array manufacturing). The Luminex strategy allows for encoding and tracking of different glycan structures in a multiplex fashion. However, the number of glycans that can be tested with this method is limited to ~100. Based on the overview of prior technologies in Chapter 1, an ideal platform for serological test and detection of glycan binding antibodies in the human serum should have the following properties: i. Reading by instruments already available in the clinic such as FACS, PCR and DNA sequencers. ii. A simple standardized workflow and built-in quality control procedures (QC). iii. Ability to display and test glycans with different linkages and densities. The same ideal standardization highlights the shortcomings of the existing platforms. The Liquid Glycan Array (LiGA) platform developed and published¹¹⁴ by our lab has the potential to fulfill these requirements. It is based on genetic barcoding of M13 virions with silent DNA barcodes in the genome of the phage. 2700 copies of outer phage coat protein allow

for multivalent display of glycans (between 10-1500 copies per virion). The DNA barcodes inside the phage with theoretical diversity of ~12 billion variants (Figure 2-1) can encode a vast number of variables such as composition and density of glycans and, as shown in this chapter it can encode replicates and glycan concentration. LiGA was validated using lectins and cell-based targets. In my project, I took it one step ahead and demonstrated the ability of LiGA to detect antibodies against carbohydrates in a more complex environment of human serum that contains billions¹¹⁶ of diverse antibodies.

As this chapter deals with the synthesis of ABO glycan array, I will briefly review several traditional chemical strategies used for glycan array synthesis and the rationale behind using ELISA and PFU assays to perform the quality control (QC) phage-glycan conjugates.

2.1.1. Robustness and reproducibility of the assay

ABO LiGA is a profiling/screening tool that can measure the enrichment of specific glycan by specific antibody from a mixture of glycans. Like any diagnostics or profiling/screening tool, LiGA should be characterized by statistical criteria that describe the ability of LiGA to distinguish true from false measurements. The traditional high throughput screening assays in which target-binding of individual compounds is tested in parallel are evaluated using a Z' statistical parameter that evaluates the robustness of an assay.¹¹⁷ In 1999, DuPont Pharmaceuticals introduced this parameter to HTS field to standardize and quantify the evaluation of robust assay as the one assay condition that gives reproducible results every single time an experiment is repeated. Z' -score is the measure of such reproducibility, and it is calculated from variability of signal in a positive and negative control samples across N technical replicates. In LiGA binding assays, there are different parameters - the amount of ABO LiGA used for screening, wash stringency, human error, and PCR bias that influence the robustness and

reproducibility. To allow for quantification of robustness and reproducibility, I introduced multiple technical replicates within the same bulk assay. Instead of attaching one glycan to one phage with SDB, I mix ten clones each with a unique SDB to create a multiplexed SDB (MSDB) and then conjugate the MSDB mixture to the glycan (Figure 2-1d). Having ten clones attached to a single glycan allowed to measure ten replicates of binding response of the same glycan to antibodies in the same bulk assay. These replicates can be used to evaluate the Z' -score (robustness) of ABO LiGA. Similar DNA-based encoding of replicates in one bulk assay has already been used by Krusemark and co-workers to calculate Z' -score of assays that used DNA-encoded libraries (DELs).¹¹⁸ To know the Z' -score for the interaction of a serum sample from a patient 'X' with LiGA, a set of 'positive' and 'negative' controls is needed which is unknown for a complex mixture of the human serum. The only known identity of the serum sample is the individual's ABO blood type and that knowledge allows for the calculation of Z' -score. It is thus logical to build an MSDB-ABO reagent that can be spiked into any sero-LiGA and serve as an internal reference.

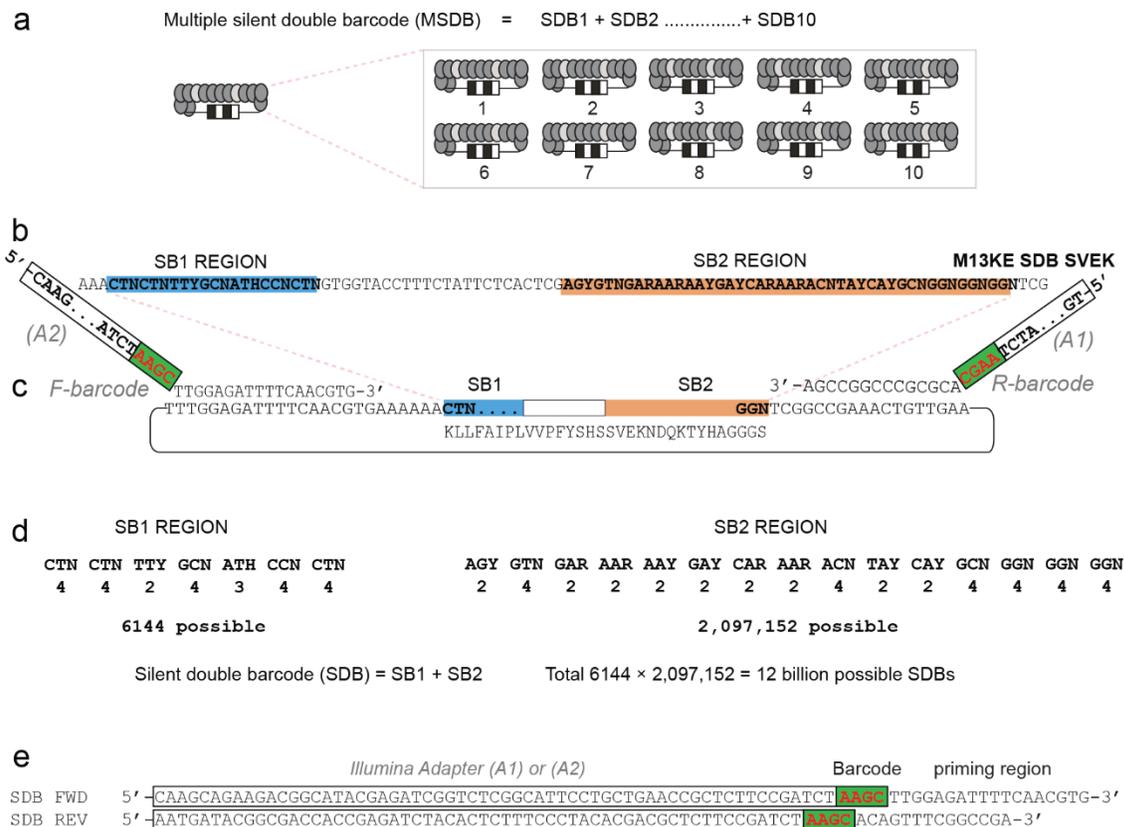


Figure 2-1: Scheme of silent double barcode (SDB) and multiple SDB (MSDB) **a**, Illustration showing that MSDB is prepared by mixing individually cloned SDB stocks **b**, Each SDB clone contains two unique barcoding sites in M13-SDB vector termed silent barcode 1 and 2 (SB1 and SB2). These regions are located in pIII gene of M13 genome. **c**, Hybridization of Illumina seq primers to SDB vector **d**, Theoretical analysis of SDB diversity **e**, Sequences of the Illumina sequencing primers.

2.1.2. Conjugation strategies

Conjugating glycans to a scaffold is the first step towards creating a glycan array. Traditionally this conjugation is done by immobilizing glycans on a solid support such as a surface of glass slide or agarose bead either by physical adsorption or covalent attachment to the reactive groups on the surface. In a typical serological assay, the array is then incubated with serum or purified glycan binding antibodies (GBA). Subject to wash procedure that removes non-specifically bound proteins, and the binding response to each glycan component of the array is measured using a fluorescently labelled secondary antibody. There are comprehensive reviews that discuss previously published physical and chemical methods of immobilizing glycans.^{36, 58, 59, 119-124} One of the most commonly used covalent immobilization strategies is the reaction of amine functionalized glycans with glass or agarose bead surfaces that display *N*-hydroxysuccinimide (NHS) esters.¹²⁵⁻¹²⁸ The alternative to the amine-NHS chemistry is the reaction between thiol functionalized glycans and maleimide groups printed on the glass slide.¹²⁹ Both and amine and thiol functionalized glycans can also react with epoxy coated glass slide. All three types of reactive groups (activated esters, epoxides and maleimides) require access to a library of glycans with a definitive reactive group (amine or thiol). The source of these glycans is either synthetic, semi-synthetic or biological. In the latter case, natural glycans from the surfaces of cells can be used as source of reagents for glycan array as long as the appropriate linker with reactive moiety (e.g., amine) can be introduced into these glycans.^{130, 131}

Preparation of LiGA requires bio-orthogonal conjugation between glycan and proteins of the M13 virion. First generation LiGA uses glycans with azido-alkyl linker as the glycan source and these glycans react with cyclooctyne moiety on the surface of the phage. In my project, I used azido blood group glycans Tri-AN3 (trisaccharide blood group A), Tri-BN3 (trisaccharide blood group B) and Di-N3 (disaccharide blood group O) that were prepared by Dr. Todd

Lowary's group at the University of Alberta. I also used ~70 glycopHage conjugates prepared from azidoglycans by Consortium of Functional Glycomics (CFG). Majority of these glycans have been validated extensively in glass and bead arrays. Most of them have also been tested in LiGA platform using traditional lectins as model reagents.

2.1.3. Glycan presentation

The density at which glycans are displayed on the surface of the carrier influences both the binding specificity as well as the affinity for glycan binding proteins.^{84, 125, 132} All antibodies have multivalent binding sites.¹³³ For example, classical IgG type contains two binding sites, whereas the IgM class of antibodies exist in the serum as a pentamer with ten binding sites and are capable of forming a multivalent complex with up to 10 antigens.⁸⁴ Antigen presentation and density is especially important for certain classes of antigens like carbohydrates since the K_d value of a binding equilibrium (affinity) between a monomeric glycan ligand and individual glycan binding receptors is usually in the high micromolar to millimolar range.¹³⁴ But once the ligands and receptors are presented to one another as multiple copies, the “avidity” of multiple binding events can be orders of magnitude lower (better) than monovalent affinity. There are several studies that focus on the role of density in not only avidity¹³⁴⁻¹³⁷ but also selectivity.¹³⁸⁻¹⁴² Even though most of these interactions were investigated with lectins, similar behaviour was observed for anti-glycan antibodies in a study by Gildersleeve and co-workers.⁸⁴ In this study, systematic variation of glycan density on microarrays enabled differentiating subpopulations of serum antibodies that are not detectable using a single glycan density.

Based on well-established importance of multivalent presentation of glycans in serological assays, I conjugated ABO glycans to phage at 3 different densities to test the role of

glycan presentation denoted as low [90-120], medium [600-800] and high [1500-1800]. The numbers in square brackets indicate the number of glycan particles per phage and as we showed in our published report¹¹⁴, these numbers can be used to calculate the average spacing between glycans on the bacteriophage surface. The construction and validation of these conjugates are discussed later in this chapter.

2.2. Quality control of glycophage conjugates and components of LiGA

2.2.1. Enzyme Linked Immunosorbent Assay (ELISA)

Since the discovery of phage-display technology in 1985, it has been widely used for the epitope mapping of monoclonal antibodies, therefore, there exists a wide repertoire of technologies to infer phage-antibody interactions.¹⁴³ Traditional epitope mapping by phage display is identical to any phage panning campaign: it employs repeated rounds of target guided selection (here: antibody) and phage propagation. Phage ELISA is often used to assess the efficiency of the rounds of biopanning. In this ELISA, the phage library from different rounds of selection is incubated on antibody coated plate and, after multiple washes, the phage population bound to the antibody is detected by a secondary anti-M13-HRP that recognizes the major coat protein and a suitable HRP chromogenic substrate. A more intense color of wells with selected libraries compared to that of control (e.g., unselected, naïve) libraries indicates that phage pools is enriched in clones that bind to the target antibody. In LiGA, we repurposed traditional phage-ELISA to evaluate the functional integrity of individual glycosylated phage constructs. In my project, ELISA was also a convenient method to optimize the concentration of purified anti-A and anti-B antibody to be coated on a plate. I also used this method to fine tune the concentration of the ABO LiGA library to be used for binding experiments. In this chapter, I will discuss the

functional validation studies of ABO LiGA using ELISA. However, it is important to note that ELISA can only track the response (binding) of individual glycosylated clones or a global response of a population of clones. But it cannot decouple responses of different phage clones in a library of mixed phage clones. Such responses can be measured by plaque forming assay (PFU, this chapter) and sequencing assays (next chapter).

2.2.2. Plaque forming Units (PFU) and PFU assays

In 1917, virologist Felix d'Herelle developed a method to quantify bacteriophages called “plaque forming assay”.¹⁴⁴ In this method, agar plates seeded with phage samples and their specific bacterial host cells are mixed in molten soft agar and poured onto a nutrient rich layer that can help the bacteria to grow. The rationale behind the plaque forming assay is that a single lytic phage particle interacts with a strain of bacterium such as *E.coli*, which causes the host bacterium to lyse and release the phage progeny.¹⁴⁴ Where phages are absent, the bacteria grow to the stationary phase and form a confluent and opaque “lawn” in the soft agar overlay. In areas where phages are present, the phage progeny released from each infected bacterium lyses or slow down the growth of neighboring bacteria and produce a growing zone of released phages, which eventually becomes visible to the naked eye as a clear circular area or “plaque” in the otherwise opaque lawn. The plaques are then counted, and number of molecules (particles) of phage is calculated as plaque forming units (PFU/mL) of the assayed sample. Despite 100 years of age, the PFU assay is still widely employed because it has an “infinite dynamic range”: on its low end it can detect one molecule (particle) of phage in the entire assay volume. In a typical assay with 100 μ L volume, this corresponds to 10^4 molecule/L = $1/6 \cdot 10^{-23+4}$ moles/L = $1.6 \cdot 10^{-18}$ molar

concentration (i.e. 1.6 attomolar). On its upper end PFU assay can detect phage at the limit of solubility (10^{13} particles/mL) once the sample is serially diluted to yield ~100 particles per plate.

PFU assay can be further upgraded to be a gene reporter assay as well. Recently, our group constructed phages that transduce three reporter genes: galactosidase¹⁴⁵, mNeonGreen (G) or mCherry (R) reporters within the *E.coli* host cells.¹¹⁴ This was a formal upgrade from a classical two color blue-white screen to a four color BWGR screen. An example of such screen was published by our group: The tracer clones were conjugated to four different glycans or controls and were used to track and evaluate the binding of glycans to glycan binding proteins or cells¹¹⁴. In my project, I conjugated ABO glycans to phage with lac-Z reporter that encodes for β -galactosidase. I used PFU assay as a proxy for deep sequencing and for optimization of binding assays.

2.3. Results and discussion

2.3.1. Construction and characterization of MSDB phage-ABO glycoconjugates

Each set of MSDB consisted of 10 phage clones each with a unique DNA barcode uniformly mixed by matching the titer of the phage stock. ABO glycans were then chemically conjugated to this mixture to yield MSDB-ABO phage-glycoconjugates. I used synthetic trisaccharide antigens A (blood group A) and B (blood group B) and disaccharide antigen H (blood group O) glycans with alkyl-azido linker (Figure 2-2) prepared by Dr. Todd Lowary's group at the University of Alberta and as described in previous publication¹¹⁴. The glycans were ligated to phage in a two-step process using strain promoted azide-alkyne cycloaddition (SPAAC). This method was used as an alternative to the Cu-activated azide-alkyne cycloaddition

(CuAAC). Ligation of glycans to phage using CuAAC was tested in the publication that describes LiGA technology and the authors observed that this method reduced the number of infective M13 phage particles, indicating damage either to the viral capsid or the enclosed DNA.

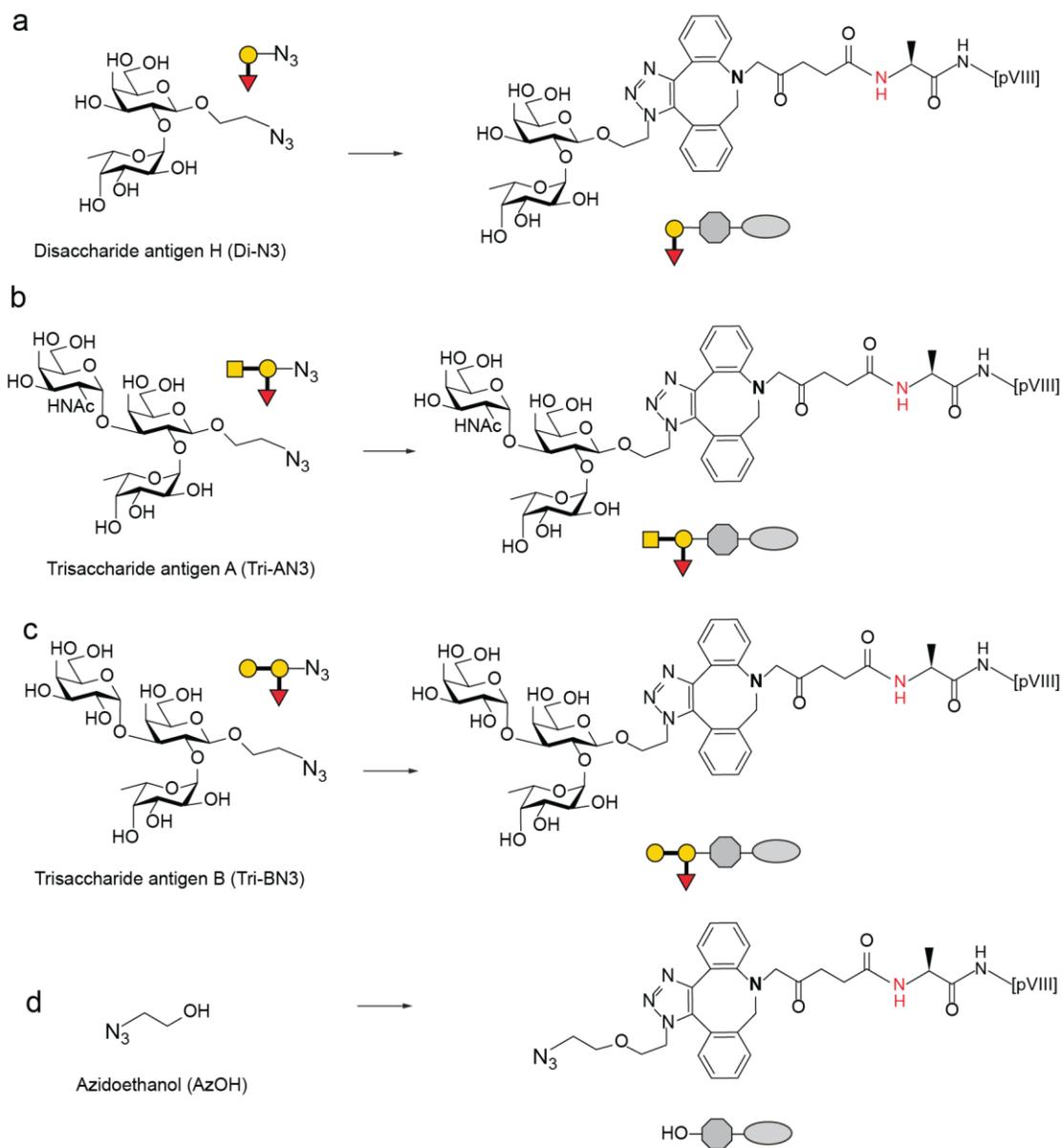


Figure 2-2: Structures of ABO glycans before and after conjugation to phage **a**, Disaccharide H **b**, Trisaccharide A **c**, Trisaccharide B **d**, Azidoethanol

In contrast, SPAAC did not decrease the number of infectious particles (Figure 2-3 adapted from the publication¹¹⁴). Based on the published procedure, I ligated the ABO glycans to MSDB phage using a two-step procedure (Figure 2-4a).

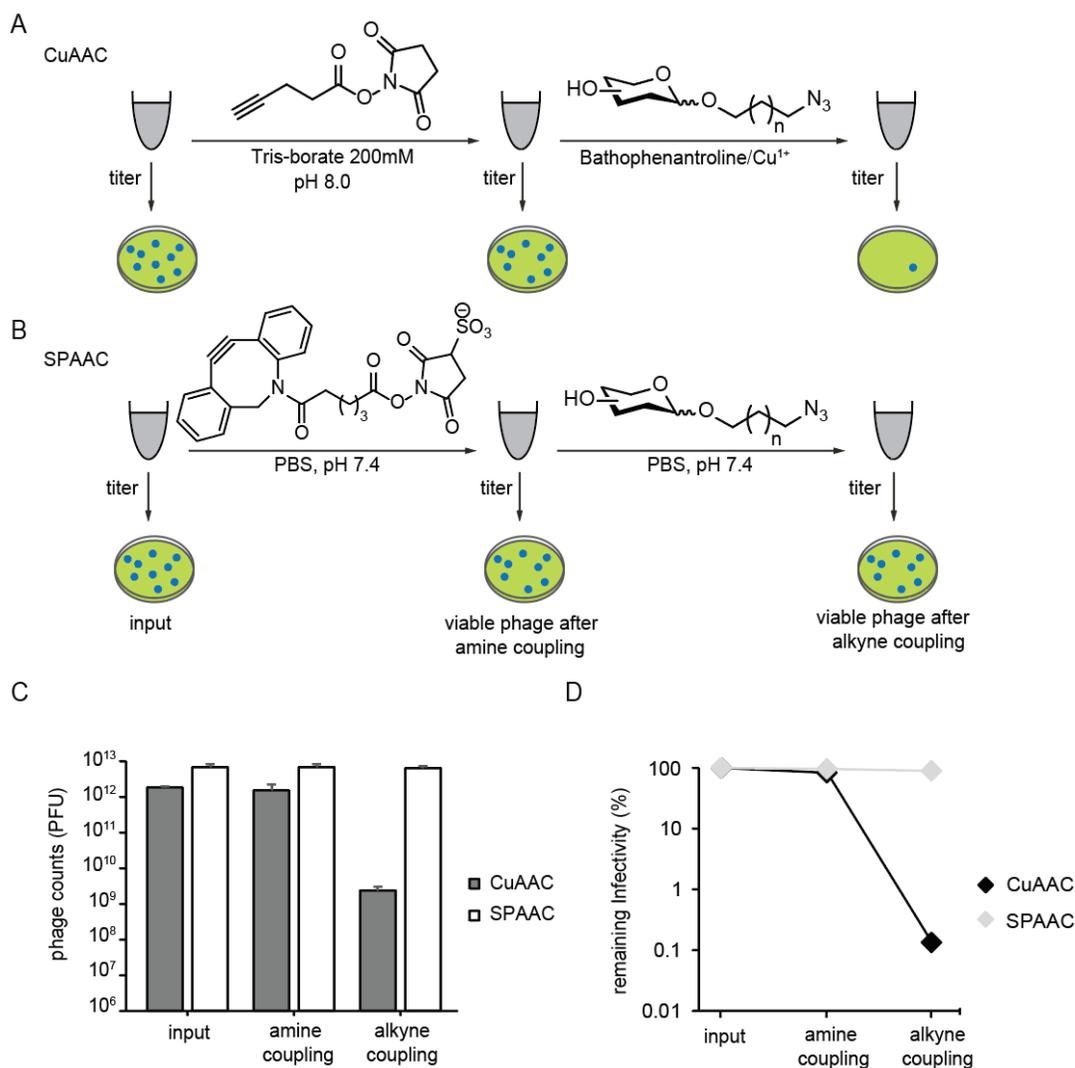


Figure 2-3: Toxicity of the chemical glycosylation to phage **a**, General steps of the procedure to incorporate the glycans via CuAAC. **b**, Glycosylation of phages using copper-free (SPAAC) chemistry. **c**, Phage titers as total counts of plaque-forming units (PFU) from chemical glycosylation experiments, n=3. **d**, Phage viability results represented as percentage of remaining infectivity observed after first and second coupling steps amine and alkyne respectively, relative to the unmodified phage mixture (input).

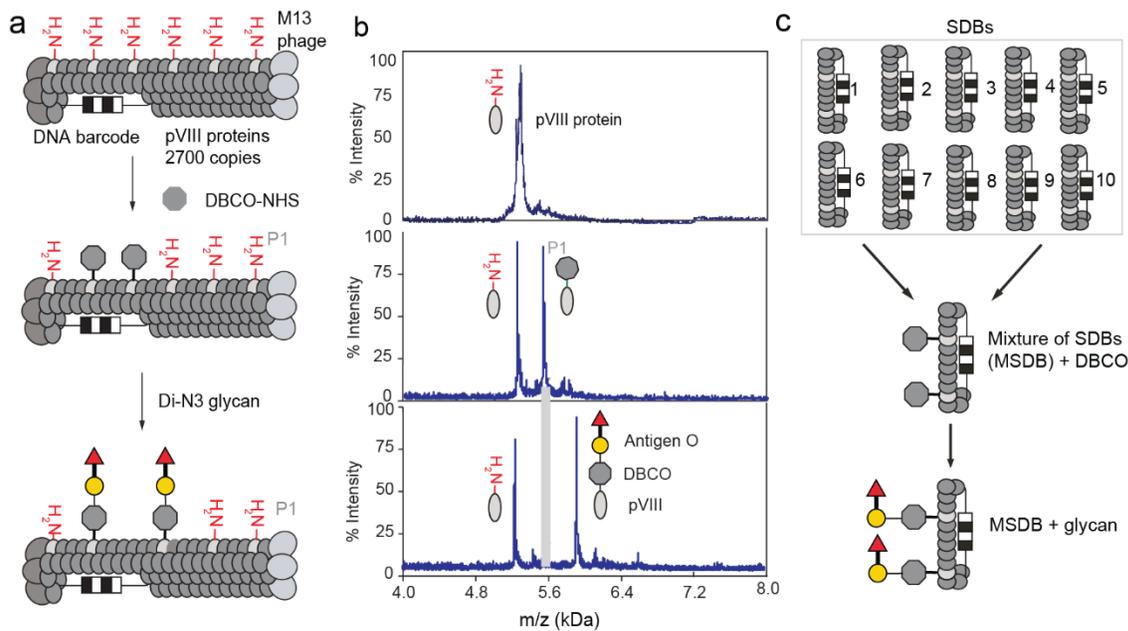


Figure 2-4: Synthesis and characterization of ABO phage glycoconjugates **a**, Representation of two-step chemical glycosylation of phage with disaccharide antigen H **b**, MALDI mass spectrometry characterization of starting material (protein pVIII), alkyne-functionalized product (DBCO–pVIII, P1) and glycoconjugate product (P2). **c**, Representation of synthesis of MSDB glycan conjugates.

In the first step, N-terminus of phage pVIII protein was acylated with dibenzocyclooctyne-N-hydroxysuccinimidyl ester (DBCO–NHS) followed by SPAAC ligation with ABO glycans with azido-alkyl linker. Subsequent addition of 2 mM azido-glycan ligated the A, B, O glycans or azidoethanol to pVIII. I monitored each step of the reaction using matrix assisted laser desorption-ionization (MALDI) time of flight (TOF) mass spectrometry. Disappearance of pVIII-DBCO peak confirmed a complete conversion of DBCO to pVIII-glycan conjugate.

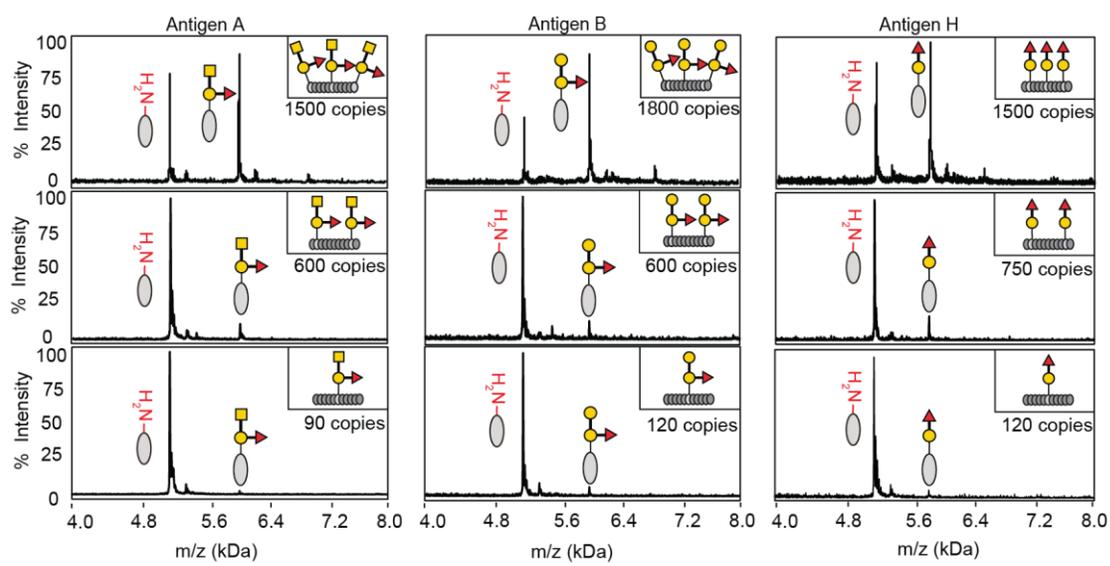


Figure 2-5: MALDI-TOF spectra of MSDB conjugated to ABO glycans with varying densities

The peak intensities semi-quantitatively represented the number of glycans on phage particles or what fraction of a total of 2700 copies of pVIII has been modified by DBCO and then by glycan. I conjugated A, B and O glycans each with ~50% density yielding phage particles with ~1500 copies of glycans per MSDB. This modification density was achieved by controlling the reaction conditions and concentration of DBCO-NHS ester. For example, by changing the concentration of DBCO-NHS from 1.5 mM to 1 mM and 0.5 mM, I was able to produce conjugates Tri-AN3-[1500], Tri-AN3-[600], Tri-AN3-[90], Tri-BN3-[1800], Tri-BN3-[600], Tri-BN3-[120], Di-N3-[1500], Di-N3-[750], Di-N3-[120]. Numbers in square brackets indicate the number of glycans per phage particle (Figure 2-5).

2.3.2. Validation of ABO LiGA using ELISA

To confirm the functional integrity of ABO glycans, I designed a sandwich type enzyme linked immunosorbent assay (ELISA) and measured the binding of individually glycosylated MSDB populations to anti-A and anti-B antibodies. The blood group antibodies used in this assay are readily available reagents used in clinical laboratories for blood typing. The MSDB phage-ABO glycoconjugates mimicked the presentation of blood group antigens on erythrocytes and were tested with murine monoclonal anti-A and anti-B IgM antibodies kindly donated by Dr. Lori West at the University of Alberta. I performed an antibody dilution study to determine an optimal coating concentration of antibodies on 96 well plates (Figure 2-6b). In binding to anti-A antibody coated wells, phage displaying antigen A trisaccharide exhibited the highest signal while phage conjugated to antigen O disaccharide exhibited 8-10x lower binding. Similarly, phage displaying antigen B trisaccharide exhibited the highest binding with anti-B antibody and a 15-17x lower binding with disaccharide O. Both A and B phages show a significantly lower binding to plates coated with bovine serum albumin (BSA). Since both anti-A and anti-B antibodies exhibited optimal signal-to-noise for 1:1000 and 1:2000 dilutions, determination of dose response of glycosylated phage particles was performed using these two dilutions. The EC₅₀ of phage particles was 5-8 X 10⁶ PFU when tested with both anti-A and anti-B antibodies (Figure 2-6c).

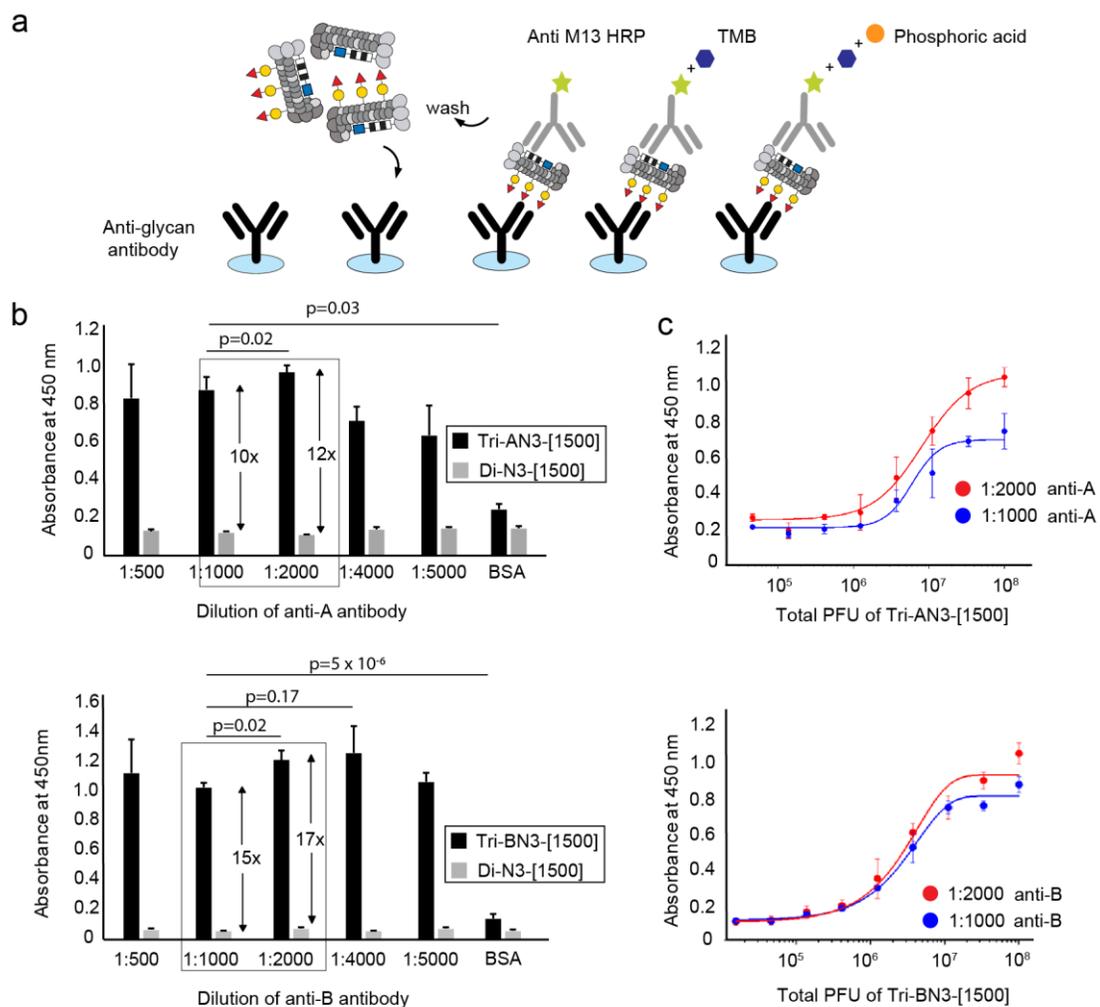


Figure 2-6: Functional validation of ABO phage glyco-conjugates using ELISA **a**, Representation of the phage-ELISA assay **b**, Dilution studies to determine the optimum coating concentration of anti-A (top) and anti-B (bottom) antibodies: Antibodies were coated at 5 different dilutions and tested with phage glycosylated with antigen A and O. Indicated p-values are calculated using two tailed t test **c**, EC50 of ABO phage particles was determined using dose titration curve experiment.

Thus, the results of ABO ELISA helped me determine the optimum coating concentration of antibodies as well as the minimum amount of glycosylated phage particles to be added to wells for detectable binding. Based on antibody dilution screen, antibodies were diluted 1:1000 to

ensure that the wells have optimal coating. In all LiGA experiments, the input concentration of LiGA was calculated to contain at least 10^6 copies of every glycosylated clone.

2.3.3. Functional validation of MSDB phage-ABO glycoconjugates using PFU assay

The PFU assay is a convenient method to further fine tune the performance of a LiGA binding assay before DNA sequencing. I used PFU assay in my project to validate the binding of individually glycosylated ABO LiGA clones as well as mixture of clones to purified anti-A and anti-B IgM antibodies (Fig. 2.7). Clones conjugated to Tri-AN3 (Tri-AN3[1500]) and Tri-BN3 (Tri-BN3[1500]) were each tested separately for their ability to bind on anti-A and anti-B antibody coated wells. The recoveries for Tri-AN3[1500] and Tri-BN3[1500] anti-A and anti-B antibodies respectively, were higher than that on BSA coated wells (Fig. 6b). MSDB ABO LiGA library that contained 5 sets of phage conjugates – Tri-AN3[1500], Tri-BN3[1500], Di-N3[1500], azidoethanol and non-conjugated “blank” phages were mixed together and tested on anti-A and anti-B antibodies and an isotype mouse IgM antibody as a control. The recovery of each subpopulation of phage was calculated as $\text{PFU}(\text{output}) \div \text{PFU}(\text{input})$. The library recovery on anti-A antibody was 10000 higher than that on the isotype antibody, whereas the recovery on anti-B antibody was 8000 higher when compared to that of the isotype antibody.

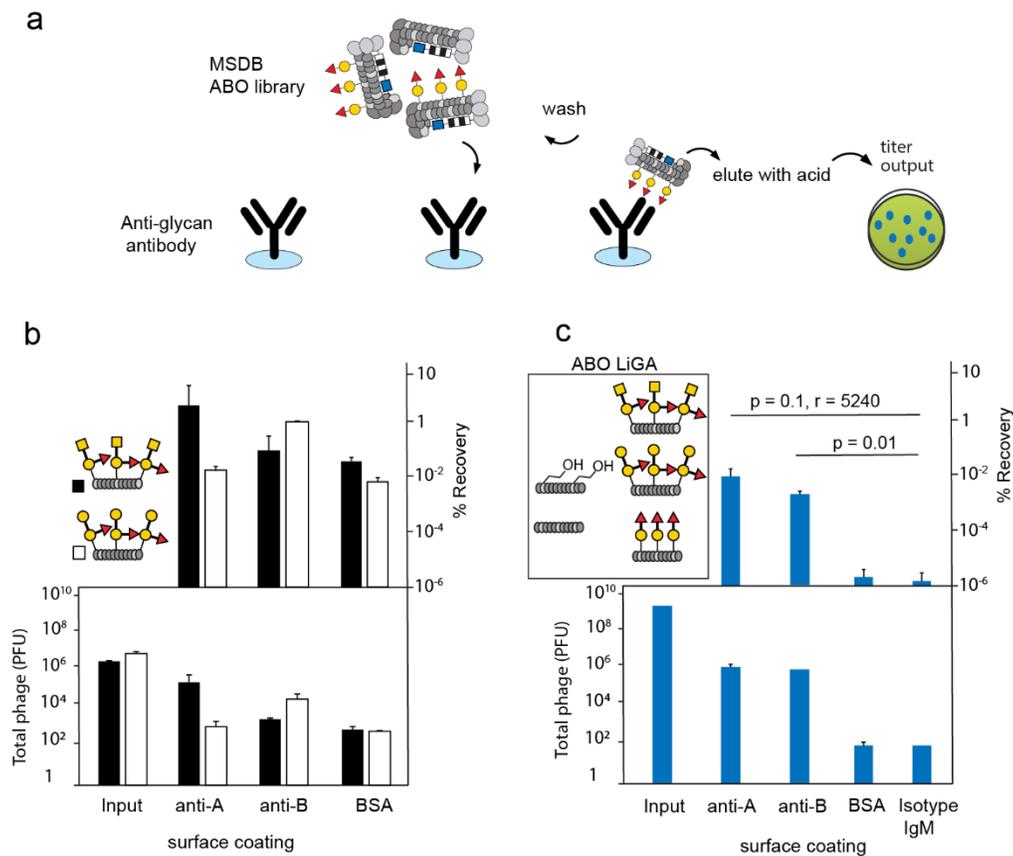


Figure 2.7: Functional validation of MSDB phage-ABO glycoconjugates using plaque forming assay **a**, Graphical illustration of phage binding assay workflow **b**, Recoveries of phage clones glycosylated with Tri-AN3 and Tri-BN3 on anti-A, anti-B antibodies and BSA coated wells. **c**, Recovery of MSDB ABO LiGA library on anti-A, anti-B antibodies, isotype mouse IgM and BSA coated wells

2.4. Conclusion

Based on MALDI, ELISA and PFU assay, I validated that phage clones conjugated to trisaccharide antigens A and B bind to purified anti-A and anti-B IgM antibodies respectively, as expected. Next step is to use polymerase chain reaction (PCR) and deep sequencing to

characterize binding of glycosylated MSDB samples and mixtures to antibodies and serum samples. These steps are discussed in detail in Chapters 3 and 4.

2.5. Materials and methods

PBS buffer contains 137 mM NaCl, 10 mM Na₂HPO₄, 2.7 mM KCl with pH adjusted to 7.4 after preparation. Solutions used for phage work were sterilized through 0.22 µm filters. Murine monoclonal anti-A and anti-B antibodies were generous gifts from Dr. Lori West (University of Alberta, CA). Trisaccharide A, B and disaccharide O was provided by Dr. Todd Lowary (University of Alberta, CA and Academia Sinica, Taiwan). Azido ethanol was kindly donated by Dr. James Paulson (The Scripps Research Institute, USA). MS-MALDI-TOF spectra were recorded on AB Sciex Voyager Elite MALDI, mass spectrometer equipped with MALDI-TOF pulsed nitrogen laser (337nm) (3ns pulse - up to 300 µJ/pulse) operating in Full Scan MS in either positive or negative ionization modes. All ELISA measurements were done on BioTek Cytation™ 5 microplate reader and imaging software.

2.5.1. Construction of silent distal barcode (SDB) M13 phage library

This protocol was performed by Dr. Nicholas Bennett in our lab and it is described in the publication.¹⁴⁶ A library of degenerate codons, termed, silent double barcodes (SDB) was created in the phage genome at a position proximal to the gIII cloning site starting from a M13KE vector containing the stuffer sequence CAG TTT ACG TAG CTG CAT CAG GGT GGA GGT corresponding to the peptide QFT*LHQ, with * representing an amber stop codon. The insert fragment was PCR amplified using the primers **P1** and **P2** and the vector fragment was PCR amplified using primers **P3** and **P4**:

Name Sequence (5'→3'):

P1GAGATTTTCAACGTGAAAAAACTNCTNNTTYGCNATHCCNCTNGTGGTACCTTTCT
ATTCTCA

P2 TTAAGACTCCTTATTACGCAGTA

P3 TTGCTAACATACTGCGTAATAAG

P4 TTTTTTCACGTTGAAAATCTC

P5GTGGTACCTTTCTATTCTCACTCGAGYGTNGARAARAAYGAYCARAARACNTAY
CAYGCNNGGNGGNGGNT-CGGCCGAAACTGTTGAAAG

P6 CGAGTGAGAATAGAAAGGTAC

PCR was performed using 50 ng phage dsDNA with 1 mM dNTPs, 0.5 μM primers, 0.5 μL Phusion High Fidelity DNA polymerase in 1x PCR buffer (NEB #B0518S) in a total volume of 50 μL. The temperature cycling protocol was performed as follows:

- a) 98 °C 3 min,
- b) 98 °C 30 s,
- c) 60 °C 30 s,
- d) 72 °C 4 min,
- e) repeat b) - d) for 35 cycles,
- f) 72 °C 10 min,
- g) 4 °C hold.

PCR amplified fragments were treated with restriction enzyme DPN 1 (NEB #R0176S) and then gel purified. NEBuilder Hifi DNA assembly was then carried out following the manufacturer protocols by mixing 100 ng of vector, 4 ng insert, 10 μ L of NEBuilder Hifi DNA assembly master mix (NEB #E2621S), and deionized H₂O up to a total volume of 20 μ L. The resulting ligated DNA was transformed into *E. coli* K12 ER2738 and propagated overnight at 37 °C. The overnight culture was then centrifuged to separate bacteriophage from host cells. Deep sequencing of the resulting SB1 QFT*LHQ cloning vector with 6,144 theoretical sequence combinations in the leader region is available at: <https://48hd.cloud/file/20161105-68OOooIC-NB>. We observed that many of the 6,144 theoretically possible sequences were eliminated due to grow disadvantage and top 200 sequences occupied ~90% of the available diversity (Figure S1). The cloning of a degenerate codon library of the SVEKNDQKTYHAGGG peptide to produce SB2 region was conducted as follows. Primers **P5** and **P2** were annealed to and amplified by PCR following the 35-cycle protocol described above to produce a dsDNA insert. The vector SB1 QFT*LHQ was PCR amplified using primers **P4** and **P6**. PCR fragment were processed using NEBuilder Hifi DNA assembly kit (NEB #E2621S) as described above. The resulting ligated DNA was transformed into electrocompetent cells *E. coli* SS320 (Lucigen). The resulting overnight culture was centrifuged to remove host cells and incubated with 5% PEG-8000, 0.5 M NaCl for 8 h at 4 °C, followed by 15 min centrifugation at 13000 g to concentrate released phage. PEG precipitated phage were re-suspended in PBS-Glycerol 50% and stored at -20 °C. M13-SDB-SVEKY is a library of chemically identical phage with 6,144 theoretical redundant sequence combinations SB1 region (~200 practically observed), and 2.1×10^6 theoretical redundant sequence combinations in SB2 region yielding between 108 to 1010 possible sequence combinations. Sequence of the vector containing M13-SDB-SVEKY library is available on GeneBank

(#MN865131). Deep sequencing of the resulting library is available at: <https://48hd.cloud/file/20161215-6700ooOO-NB>

2.5.2. SDB clone isolation and amplification

Isolation of SDB clones was performed by Dr. Nicholas Bennett, Susmita Sarkar, Jasmine Maghera and Mirat Sojitra in our lab. For my own project, I amplified the clones previously isolated in the lab. The M13-SDB-SVEKY library described in the previous section was used to isolate each monoclonal silently encoded phage. A 10 μ L aliquot of phage was diluted and plated at a density of 100 plaques per plate. Single colonies were manually picked, and individually transferred into a clean 1.7 mL plastic tube containing 0.5 mL of PBS-Glycerol 50% and incubated at room temperature for 30 min. The tubes were then placed in 55 °C heating block for 10 min to inactivate any remaining bacterial cells. After the incubation, 20 μ L sample of each colony suspension was amplified for 4.5 h in 5 mL of LB supplemented with a 0.5 mL of log phase *E. coli* K12 ER2738. After amplification the phage clones were collected from the culture supernatant by centrifugation at 4500 g for 10 min. Next, bacterial cell pellet and culture supernatant were processed separately. The supernatant was incubated with 5% PEG-8000, 0.5 M NaCl for 8 h at 4 °C, followed by 15 min centrifugation at 13000 g to precipitate the viral particles. The phages were re-suspended into 1 mL PBS-Glycerol 50%, titered, and stored at -20 °C until further use. The bacterial cell pellet was processed for phage-DNA extraction using GeneJET Plasmid Miniprep kit (Thermo Fisher, #K0502). For SDB identification, a sample of 400 ng of the phage DNA was submitted for Sanger sequencing. We selected the clones that contained three or more base pair substitutions from one another (i.e., Hamming distance (H) \geq 3).

H₃ permits correction of any point mutations that may have arisen during the analysis by deep sequencing.

2.5.3. Monitoring phage viability in CuAAC reaction

This protocol was performed by Susmita Sarkar in our lab and it was described in a publication.¹¹⁴ A tube containing 150 μ L of 3.4×10^{13} PFU/mL of phage clone SBD3 in PBS pH 7.4 was concentrated using 0.5 mL Amicon filter (Sigma, # UFC5010) with 10 kDa MW cut off, and resuspended in Tris-borate 200 mM pH 8.0, to a final volume of 150 μ L. The phage solution was now tittered by plaque forming assay and then divided into triplicates containing 50 μ L of phage solution. Each aliquot was reacted with NP (1mM final concentration) at room temperature 1 h. The reaction mix was then passed through Zeba Spin Desalting Columns 40K (Thermo-Fisher, #87766) to eliminate unreacted linker following these steps: first the column was washed with 0.5 mL PBS, then, the phage mix was loaded onto the resin and centrifuged at 1000 \times g for 1 min. After filtration, 10 μ L sample of phage mix was used for analysis by MALDI-TOF spectrometry and plaque forming assay. The CuAAC reaction was implemented following literature procedures with minor modifications. Bathophenanthroline/Cu¹⁺ catalyst was prepared as follows: glass vial (4 mL) was charged with 10 mg of Cu₂SO₄·5H₂O and bathophenanthroline sulfonate (64.4 mg) (GFS Chemicals Inc.) dissolved in 0.2 M Tris HCl, pH 8.0 buffer (1 mL). Copper powder (~50 mg) was added; the vial was closed with a rubber septa and purged with argon. The vial was rotated for 2 h; the reduction of copper II to copper I by metallic copper was indicated by the appearance of a dark green color. To the filtered phage solution was added azido-glycan JB-1 up to a final concentration of 2 mM and ~1 mg of copper powder. The phage solution was now degassed by 30 min sonication in argon atmosphere. The glycosylation reaction was

then initiated by adding 3 μL of 130 Bathophenanthroline/ Cu^{1+} catalyst. The tube was closed under argon atmosphere and incubated 8 h at 4 $^{\circ}\text{C}$. The mixture was then passed through Zeba Spin column as described above. The viability of the recovered phages was analyzed by plaque forming assay and the chemical modification was analyzed by MALDI-TOF.

2.5.4. Ligation of ABO glycans to M13 phage

A solution of SDB phage clone (10^{12} - 10^{13} PFU/mL in PBS) was combined with DCBO-NHS (20 mM in DMF) to afford a 0.2-2.0 mM concentration of DCBO-NHS in reaction mixture, which typically yields 5-50% of pVIII modification after 45 min of incubation. After conjugation of DBCO-NHS, each clone was individually purified on a ZebaTM column following the manufacturer instructions. Solutions of azido-glycans (10 mM stock in Nuclease Free H₂O) were added to the filtrates to afford a 2 mM concentration of glycan-azide and the solutions were further incubated overnight at 4 $^{\circ}\text{C}$. All chemical reactions were verified and quantified by MALDI-TOF as described on the previous sections. If reactions were incomplete and residual pVIII-DBCO peak was detected in MALDI, we added additional amount of azido glycan and extended the incubation. If reactions were completed, the conjugates were purified by ZebaTM column and stored at 4 $^{\circ}\text{C}$ or supplemented by glycerol and stored as 50% glycerol stock at -20 $^{\circ}\text{C}$. LiGA mixture was prepared by combining these stock solutions.

2.5.5. Characterization of phage-ABO glycoconjugates by MALDI-TOF MS

The sinapinic acid matrix was formed by deposition of two layers. Layer 1 was prepared as 10 mg/mL solution of sinapinic acid (Sigma, #D7927) in acetone-methanol (4:1). Layer 2 was prepared as 10 mg/mL solution of sinapinic acid in acetonitrile:water (1:1) with 0.1% TFA. In a

typical sample preparation, 2 μL of phage solution in PBS was combined with 4 μL of layer 2, then a 1:1 mixture of (layer 1): (layer 2+phage) was deposited in that order onto the MALDI inlet plate ensuring that layer 1 is completely dry before adding layer 2+phage. The spots were washed with 10 μL of water with 0.1% TFA to remove salt ions from the PBS. To estimate the ratio of modified to unmodified pVIII, we fit and plotted the data using MatLab.

2.5.6. Binding of phage-ABO glycoconjugates to anti-A and anti-B measured by

ELISA

Blood group reagents Anti-A and Anti-B murine monoclonal antibody (Obtained from Immucor, Inc. and generously donated to me by Dr. Lori West at the University of Alberta) was dissolved in PBS at final dilution of 1:1000 and 1:2000. 50 μL of the solution was added to each well of 96 well plate (Corning®, #CLS3369). The plate was covered with sealing tape (Thermo Scientific, #15036) and incubated overnight at 4 °C. The following day, the wells were washed 3 times by adding washing buffer (200 μL , 0.1% Tween-20 in PBS) in the wells and discarding the solution by inverting the plate on top of a paper towel. Thereafter, blocking solution (100 μL , 20 $\mu\text{g}/\mu\text{L}$ BSA in PBS) was added to the wells and incubated for 1 h at rt. The solution was discarded by inverting the plate, the wells were then washed three times with washing buffer. Thereafter, solutions of LiGA clones (50 μL in PBS) was added to the wells and incubated for 1 h at rt. The solution was discarded by inverting the plate, the wells were washed three times washing buffer (200 μL 0.1% Tween-20 in PBS) and probed with anti-M13-HRP conjugate.

2.5.7. Construction of Multiplex Silent Distal Barcodes (MSDBs)

There are two ways to construct MSDBs. In the first method, an MSDB was prepared by mixing 10 phage clones each with a unique SDB. Individual phage clones amplified in 25 mL flasks were characterized by deep sequencing and mixed by matching the titer of the phage stock. The MSDB was then used to conjugate blood group antigens. Each MSDB was assigned a two-letter identifier (e.g., ZA) and a “dictionary” (e.g., ZA.xlsx, Figure 2-8), a table that describes the correspondence between the DNA barcodes and the glycan conjugated to the respective MSDB. In the second method, 10 individual clones were amplified separately in 3 mL volume and mixed together before purification of phage by PEG precipitation. The initial set of MSDBs described in this chapter were constructed using the first method, however, since the second method utilizes lesser volume of reagents (amplification in 3 mL tubes as opposed to 25 mL flasks in the first method), moving forward, I used the second method to construct all my MSDBs.

	B	C	D	E
...SDB Position...		Sequence.....	Alphanum..	# Modification.....
SDB181	[7:27 52:96]	TTATTATTCGCAATTCCTTTAAGTGTGGAGAAAAACGACCAAAAAGACCTACCATGCTGGGGGGGGT	Tri-AN3-1	1 Gala1-3[Fuca1-2]Galb-Sp
SDB12	[7:27 52:96]	CTTCTGTTTCGCATACCTCTAAGTGTGGAGAAGAATGATCAGAAGACTTATCATGCGGGTGGAGGT	Tri-AN3-2	2 Gala1-3[Fuca1-2]Galb-Sp
SDB16	[7:27 52:96]	CTGCTGTTTCGCAATCCCGCTGAGTGTGGAGAAGAATGATCAGAAGACTTATCATGCGGGTGGAGGT	Tri-AN3-3	3 Gala1-3[Fuca1-2]Galb-Sp
SDB70	[7:27 52:96]	CTACTGTTTCGCAATCCCGCTCAGTGTGAAAAAAAACGATCAAAAAACGTATCATGCTGGTGGAGGT	Tri-AN3-4	4 Gala1-3[Fuca1-2]Galb-Sp
SDB11	[7:27 52:96]	CTACTATTTCGCAATCCCGCTGAGTGTGGAGAAGAATGATCAGAAGACTTATCATGCGGGTGGAGGT	Tri-AN3-5	5 Gala1-3[Fuca1-2]Galb-Sp
SDB25	[7:27 52:96]	CTGCTTTTCGCAATACCTCTAAGTGTGGAGAAGAATGATCAGAAGACTTATCATGCGGGTGGAGGT	Tri-AN3-6	6 Gala1-3[Fuca1-2]Galb-Sp
SDB45	[7:27 52:96]	CTGCTGTTTCGCAATCCCTCTGAGCCTGGAAAAAATGACCAAAAAACCTACCATGCAGGGGGGGGA	Tri-AN3-7	7 Gala1-3[Fuca1-2]Galb-Sp
SDB48	[7:27 52:96]	TTATTATTCGCAATTCCTTTAAGCGTAGAGAAAAACGACCAAGAAGACTTATCATGCGGGGAGGAGGT	Tri-AN3-8	8 Gala1-3[Fuca1-2]Galb-Sp
SDB49	[7:27 52:96]	CTACTGTTTCGCTATCCCGCTGAGTGTGGAAAAAGAATGATCAGAAGACTTATCATGCGGGTGGGGG	Tri-AN3-9	9 Gala1-3[Fuca1-2]Galb-Sp
SDB50	[7:27 52:96]	CTGCTCTTTTCGCAATCCCGCTGAGTGTGCAGAGAAGAACGACCAAAAAACATATCACGCGGGGGGGGA	Tri-AN3-10	10 Gala1-3[Fuca1-2]Galb-Sp

Figure 2-8: Snapshot of ZA.xlsx

2.5.8. Binding of ABO glycosylated clones to different dilutions of anti-A and anti-B antibodies measured by ELISA

Murine monoclonal anti-A and anti-B IgM antibodies (produced by Immucor, Inc. and generously donated by Dr. Lori West at the University of Alberta) were each dissolved in PBS

to yield the dilutions – 1:500, 1:1000, 1:2000, 1:4000 and 1:5000. 50 µL of the dilutions was added to each well of 96 well plate (Corning®, #CLS3369). The plate was covered with sealing tape (Thermo Scientific, #15036) and incubated overnight at 4 °C. The following day, the wells were washed 3 times by adding washing buffer (200 µL, 0.1% Tween-20 in PBS) in the wells and discarding the solution by inverting the plate on top of a paper towel. Thereafter, blocking solution (100 µL, 20 µg/uL BSA in PBS) was added to the wells and incubated for 1 h at rt. The solution was discarded by inverting the plate, the wells were then washed three times with washing buffer. Thereafter, solutions of antigens A and O and antigens B and O (50 µL in PBS) were added to the wells coated with anti-A and anti-B antibody respectively and incubated for 1.5 h at rt. The solution was discarded by inverting the plate, the wells were washed three times with washing buffer (200 µL 0.1% Tween-20 in PBS) and probed with anti-M13-HRP conjugate. 100 µL of TMB ELISA substrate (Thermo Scientific, #34028) was added to each well and incubated for 5-10 min until a deep blue color developed. The solutions were then quenched with 100 µL 1 M phosphoric acid and absorbance at 450 nm was recorded.

2.5.9. Dose titration study of ABO glycosylated clones measured by ELISA

Anti-A and anti-B antibodies were each dissolved in PBS to yield final dilutions of 1:1000 and 1:2000. Plates were prepared as described in previous section **1.4.9**. Antigen A and antigen B glycosylated phage clones of concentrations - 10^4 , 10^5 , 10^6 , 10^7 , 10^8 PFU were added to wells coated with anti-A and anti-B antibody respectively and incubated for 1.5 h at RT. The solution was discarded by inverting the plate, the wells were washed three times with washing buffer (200 µL 0.1% Tween-20 in PBS) and probed with anti-M13-HRP conjugate. 100 µL of TMB ELISA substrate (Thermo Scientific, #34028) was added to each well and incubated for 5-10 min until a

deep blue color developed. The solutions were then quenched with 100 μ L 1 M phosphoric acid and absorbance at 450 nm was recorded using BioTek Cytation 5 Plate reader.

2.5.10. Generation of MALDI plots using MATLAB script

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

addpath (fullfile('/Volumes/Data/!! OUTLINES/Jasmine -
LiGA/MALDI/Mirat/untitledfolder'));

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

manual = 0;

filename = 'Phage+Lac_finish reaction.txt';

names = 'Lac';

full_temp = 'Galb1-4G1cb-S2';

XLIM1 = 3000;

XLIM2 = 8000;

%range around the peak for gaussian fitting and baseline fit
RANGE = 255;

fid = fopen(filename,'r');

disp([names ' glycan is in ' filename ' ' fgetl(fid)...
char(10) fgetl(fid)]);

FORM = '%f %f %*[^\\n\\r]';

AllVar = textscan(fid,FORM);

mass = AllVar{1};

intensity = AllVar{2};

disp(['Read ' num2str( size(mass, 1)) ' lines' ]);

fclose(fid);

IX = find(mass>XLIM1 & mass<XLIM2);
```

```

plot_mass = mass(IX);
plot_intensity = intensity(IX);
YMIN = min(plot_intensity);
YMAX = max(plot_intensity);
YH = YMAX-YMIN; 272

SCALE = 1000/YH;
plot_intensity = plot_intensity*SCALE;
YH = YH*SCALE;
YMIN = min(plot_intensity);
YMAX = max(plot_intensity);
margin = 0.1;
h = figure(1);
set(h, 'Units', 'normalized', 'position', [0.2 0.6 0.45 0.25] )
% find the full name based on abbreviated name
%%%%%%%%%% PLOT MALDI TRACE AND MAKE IT PRETTY
%%%%%%%%%%

figure(1);
hold off;
plot(1,1);
hold on;
plot(plot_mass, plot_intensity, '-k');
% add full name of glycan to the title
title([names ' - ' full_temp]);
set(gca, 'yscale', 'lin', 'xscale', 'lin', ...
'TickDir', 'out');
ylabel('intensity');
xlabel('M/z');
xlim([4000 8000]);
ylim([YMIN - margin*YH YMAX + margin*YH]);

```

```

drawnow;
hold on;
IX2 = find ( plot_intensity == ...
max(plot_intensity(plot_mass>5000 & plot_mass<5500)) );
pVIII = plot_mass(IX2(1));
% define the RANGE and make sure it doesn't exceed the
boundaries 273

if IX2+RANGE > size(plot_intensity,1)
H = size(plot_intensity,1);
else
H = IX2+RANGE;
end
if IX2-RANGE < 1
L = 1;
else
L = IX2-RANGE;
end
initial_mass = pVIII;
initial_int = max( plot_intensity(L:H) );
fo = fitoptions('Method','NonlinearLeastSquares',...
'StartPoint',[initial_int initial_mass 1 1 1]);
ft = fittype('a1*exp(-((x-b1)/c1)^2)+e1*x+f1','options',fo);
f = fit(plot_mass(L:H),plot_intensity(L:H),ft);
X = plot_mass(L:H);
MASS_P8 = f.b1;
MASS_P8_DBCO = MASS_P8 + 315;
HEIGHT_P8 = f.a1;
MAX_P8 = max(f(X));
plot(X, f(X), '-r');

```

```

plot(X, f.e1*X + f.f1, '-b');
line(MASS_P8*[1 1], [MAX_P8, MAX_P8-HEIGHT_P8], 'color','r');
text(MASS_P8*1.01, MAX_P8,...
['M(pVIII)=' num2str(round(MASS_P8)) char(10)...
'H(pVIII)=' num2str(round(HEIGHT_P8))],...
'VerticalAlignment', 'bottom',... 274

'HorizontalAlignment','left');
text(pVIII, YMIN, 'pVIII', ...
'FontWeight', 'bold',...
'VerticalAlignment', 'top',...
'HorizontalAlignment','center');
line(MASS_P8_DBCO*[1 1], YMIN + [0 0.3*f.a1], 'color', 'r');
text(MASS_P8_DBCO, YMIN, ['pVIII' char(10) '+DBCO'], ...
'FontWeight', 'bold',...
'VerticalAlignment', 'top',...
'HorizontalAlignment','center');

%%%%%%%%%% EXTRACT THE FULL NAME AND MAKE GLYCAN OBJECT
%%%%%%%%%%

sugar = GlycanLeaf.createObj(full_temp);
hold on; plot(1,1);
drawGlycan('input',sugar.String, 'XY', [7500 YMAX],...
'spacing', 150, 'angle', pi);
TEXT = ['MW=' num2str(round(sugar.MW)) char(10)];
DEBUG = 0;
[MW(char(10)),b,c] = sugar.MW;
if DEBUG
for j = 1:numel(sugar.glycans)
TEXT = [TEXT sugar.glycans{j} ...
': ' num2str(round(c(j) )) char(10)];

```

```

end
end
text(7500, 0.90*YMAX, TEXT, 'HorizontalAlignment', 'left',...
'VerticalAlignment', 'top');
% set(gca,'xtick',[], 'ytick',[]) 275

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
MASS_P8_DBCO_GLY = MASS_P8_DBCO + MW(char(10));
line(MASS_P8_DBCO_GLY*[1 1], YMIN + [0 0.1*HEIGHT_P8], 'color',
'b');
text(MASS_P8_DBCO_GLY, YMIN, '*',...
'FontWeight', 'bold',...
'VerticalAlignment', 'bottom',...
'HorizontalAlignment','center');
text(MASS_P8_DBCO_GLY, YMIN, 'P+glycan',...
'FontWeight', 'bold',...
'VerticalAlignment', 'top',...
'HorizontalAlignment','center');
drawnow;
repeating = 1;
j=0;
while repeating>0
if manual
choice = questdlg('define a peak?',...
'define a peak?',...
'define a peak','no mo peaks','escape','define a peak');
switch choice
case 'escape', return;
case 'no mo peaks', repeating = 0; continue;

```

```
case 'define a peak',
[x]=ginput(1);
j = j+1;
lineX(j) = x(1);
end
else
j = j+1;
if j==1
% find the main peak
lineX(j) = MASS_P8_DBCO_GLY;
276

IX =
(strcmp('Neu5Ac',sugar.glycans)
```

```

strcmp('Neu5,9Ac2',sugar.glycans) |...
strcmp('Neu5Gc',sugar.glycans) |...
strcmp('KDN',sugar.glycans) );
repeating = sum(IX);
elseif j==2 && repeating
IX = ~(strcmp('Neu5Ac',sugar.glycans) |...
strcmp('Neu5,9Ac2',sugar.glycans) |...
strcmp('Neu5Gc',sugar.glycans) |...
strcmp('KDN',sugar.glycans) );
Nsialo = sum( ~IX );
[~,~,c]= sugar.MW;
asialoMW = sum(c(IX)) - 18*(sum(IX)-1);
MASS_P8_DBCO_GLY = MASS_P8_DBCO + asialoMW;
lineX(j) = MASS_P8_DBCO_GLY;
line(MASS_P8_DBCO_GLY*[1 1], YMIN + [0 0.1*HEIGHT_P8],
'color', 'b');
text(MASS_P8_DBCO_GLY, YMIN, '*',...
'FontWeight', 'bold',...
'VerticalAlignment', 'bottom',...
'HorizontalAlignment','center');
text(MASS_P8_DBCO_GLY, YMIN,...
['-' num2str(Nsialo) 'xSialo'] ,...
'FontWeight', 'bold',...
'VerticalAlignment', 'top',... 277

'HorizontalAlignment','center');
repeating = 0;
end

```

```

% find the index of data nearest to the mouse click
IX2 = max ( find( plot_mass < lineX(j) ) );

% define the RANGE and make sure it doesn't exceed the
boundaries

if IX2+RANGE > size(plot_intensity,1)
H = size(plot_intensity,1);
else
H = IX2+RANGE;
end

if IX2-RANGE < 1
L = 1;
else
L = IX2-RANGE;
end

initial_mass = lineX(j);
initial_int = max( plot_intensity(L:H) );
fo = fitoptions('Method','NonlinearLeastSquares',...
'StartPoint',[initial_int initial_mass 1 1 1]);
ft = fittype('a1*exp(-((x-b1)/c1)^2)+e1*x+f1','options',fo);
f = fit(plot_mass(L:H),...
plot_intensity(L:H),ft);
X = plot_mass(L:H);
MASS(j) = f.b1;
HEIGHT(j) = f.a1;
MAX(j) = max(f(X)); 278

plot(X, f(X), '-r');
plot(X, f.e1*X + f.f1, '-b');

```

```

line(MASS(j)*[1 1], [MAX(j), MAX(j)-HEIGHT(j)], 'color','r');
Delta(j) = round(abs(MASS(j)-MASS_P8_DBCO));
Ratio(j) = round(100*HEIGHT(j) / (HEIGHT(j) + HEIGHT_P8));
if j==1
text(MASS(j)+200, MAX(j)*0.7,...
['M=' num2str(round(MASS(j))) char(10)...
'dM=' num2str(Delta(j)) char(10)...
'R=' num2str(Ratio(j)) '%' char(10)],...
'VerticalAlignment', 'bottom',...
'HorizontalAlignment','left');
elseif j==2
text(MASS(j), MAX(j) + (MAX(j)-HEIGHT(j))*0.1,...
['M=' num2str(round(MASS(j))) char(10)...
'dM=' num2str(Delta(j)) char(10)...
'R=' num2str(Ratio(j)) '%' char(10)],...
'VerticalAlignment', 'bottom',...
'HorizontalAlignment','right');
end
end
end

```

Chapter 3: Validation of MSDB ABO LiGA by deep sequencing

3.1. Introduction

In this chapter, I discuss the use of PCR and next-generation sequencing (NGS) technology to check the integrity and functional performance of the multiplexed liquid glycan array technology m-LiGA. The m-LiGA is an upgraded variant of previously published LiGA¹¹⁴, in which multiple DNA barcodes encode the same glycan structure and the same glycan density and introduce multiple *in-situ* measurements for every binding event. I demonstrate how this multiplexing makes it possible to calculate Z'-factor of the binding assay and characterize the confidence of the LiGA assay. As this chapter deals with sequencing and calculation of Z'-factor, I will briefly introduce different DNA sequencing technologies, DNA encoding strategies and common biases encountered in NGS based assays.

3.1.1. Sequencing technologies

3.1.1.1. Sanger sequencing

Developed in 1977 by the two-time Nobel laureate Fredrick Sanger, Sanger sequencing is based on the “chain termination method”. In this method, the DNA sequence of interest is amplified using a special type of polymerase chain reaction called chain termination PCR. Along with the deoxyribonucleotide (dNTP) used in a standard PCR, chain termination PCR involves addition of dideoxyribonucleotide (ddNTPs) that terminates the elongation of DNA strands. Since ddNTPs lack a 3'-OH group that is critical in the formation of a phosphodiester bond, DNA polymerase ceases the extension of the template strand. In

automated Sanger sequencing, all ddNTPs are mixed in a single reaction and each dNTP has a unique radioactive or fluorescent label. The chain terminated oligonucleotides are further run in a single capillary gel electrophoresis where DNA fragments are separated based on size. The output is read in the form of a chromatogram which shows the fluorescent peak of each nucleotide along the length of the DNA strand.

Since the discovery of the phage-display platform in 1985, Sanger sequencing was the only available tool for identification of binder sequences.¹⁴⁷⁻¹⁵¹ The implementation of Sanger Sequencing to identify binders from a DNA-encoded library was first described by Neri and co-workers. After PCR amplification of the enriched library compounds, concatamers with coding sequences were generated and ligated into a puC19 vector. A representative number of the resulting colonies were Sanger sequenced which revealed the frequencies of codes present in the library before and after the selection experiment.¹⁴⁹ In a phage-display based selection experiment, Sanger sequencing is usually used to analyze the DNA sequence of individual clones. It can be repeated for separately isolated clones to sequence <100 different clones. With the help of automated colony picking and automated DNA preparation, it can be used to sequence rarely up to 1000 phage clones. More importantly, even for a library with less than 1000 phage clones, this process is labor-intensive as it involves the isolation of DNA from individual phage clones.

3.1.1.2. Roche 454 sequencing

Launched by 454 Life Sciences in 2005 and later acquired by Roche in 2007, the 454-sequencing method was the first next-generation DNA sequencing technology to reach the market. The 454 method is based on detection of release of the pyrophosphate group in the form of luminescence mediated by an enzymatic reaction. Library preparation for sequencing begins with fractionation of DNA templates into smaller fragments and ligation of short adaptors onto the ends of the DNA fragments. Biotinylated DNA library is immobilized on streptavidin beads containing sequences complimentary to that of the library and PCR amplified in an emulsion of water-in-oil mixture. These beads are then sequenced in parallel wells and the light emitted during the incorporation of nucleotides is recorded by a camera. The intensity of the light corresponds to the number of nucleotides of the same type that have been incorporated.

Most of the early work in DNA-encoded libraries and phage-display technologies were done using the Roche 454 sequencing method.¹⁵²⁻¹⁵⁵ This technology allows for pooling of multiple DNA samples from different experiments and processing and sequencing in the same run, thus, reducing the sequencing cost per experiment. This process is known as multiplexing which is achieved by using primers with short, standardized DNA sequences (“barcodes”), that allow segregating independent samples during the processing of sequencing data. For example, Sidhu and co-workers sequenced 22 phage screening experiments in one run by using barcoded primers.¹⁵⁴ These primers enable the analysis of specific sequences present in a library before and after a selection experiment and enrichment against a target of interest. In 2009, Arap, Pasqualini and co-workers demonstrated the utility

of next generation sequencing in phage-display library selection using the 454 sequencing method and did a thorough side-by-side comparison with the Sanger Sequencing technology. The authors performed *in vivo* biopanning of a phage CX₇C library on a terminally ill patient. Phage clones recovered from biopsies of various organs and tissues were sequenced through Sanger or Roche 454 sequencing. Sanger sequencing identified 3,840 sequences, while 454 identified 319,361 sequences. The time required to sequence 3,840 clones by Sanger sequencing was ~15 days. The authors extrapolated the cost and time required to sequence 100,000 phage clones by Sanger as 9,898 hrs (412 days) and \$338,884. Analysis of 100,000 phage clones by 454 sequencing required only 74.8 hrs and \$1,307. Though the 454 method is now discontinued, other NGS platforms like Ion Torrent and Illumina have further dramatically reduced the sequencing cost, time and increased the amount of data that can be obtained.

3.1.1.3. Ion Torrent Sequencing

Ion Torrent Sequencing is where chemistry meets electrical engineering. This technology translates the nucleotide sequence information stored in a DNA to digital information (0,1) on a semiconductor chip. The sensors on the chip detect the release of hydrogen (H⁺) ions during addition of dNTPs to the template strand by DNA polymerase.¹⁵⁶ The semiconductor chips are made of a dense array of more than a million micro-wells. Each micro-well contains an ion sensitive transistor and space for one bead with a clonal population of DNA. The change in pH during DNA polymerization is measured by transistors in real time, which is then converted into a voltage. The dNTPs are sequentially

added to the semiconductor chips containing the DNA template. Hydrogen ions are released if the nucleotide is complementary to the DNA strand and the ion sensor is triggered. The sensor detects an electrical signal that is proportional to the number of bases incorporated. No signal is detected if the added dNTP is not complementary to the template nucleotide since no polymerization occurs. In Ion Torrent platform, amplification of DNA sequence occurs by emulsion PCR. Ion sphere particles (ISPs) contain sequence complementary to that of adaptor sequence on the DNA template. PCR amplification of DNA library fragments in microdroplets results in population of beads containing a monoclonal population of single DNA fragment. The microdroplets are then loaded onto proton-sensing wells where sequencing is primed from a specific location on the adapter sequence.¹⁵⁷ Our group optimized the protocol for performing phage-display library selection using Ion Torrent platform. This optimized method was used to demonstrate the utility of Ion Torrent to identify growth bias in phage libraries and characterize the parasitic sequences in phage screening experiments.^{158, 159}

3.1.1.4. Illumina deep sequencing

The Illumina DNA sequencing platform is based on the ‘sequencing by synthesis’ (SBS) chemistry. The sample preparation for Illumina sequencing begins by ligating specialized adapters to both ends of the DNA fragment to be sequenced. The sample is then loaded into a glass flow cell where single-stranded DNA fragments complementary to that of the adapters are hybridized on the surface of the flow cell. Each fragment is amplified into distinct monoclonal clusters through a process called bridge amplification. Sequencing

reagents along with fluorescently labelled nucleotides at the 3'-OH position are added to the flow cell. The 3'-OH blocking group allows only one nucleotide to be incorporated by DNA polymerase. After the nucleotide incorporation, fluorescence intensity is recorded, and the blocking group is chemically removed to incorporate the next base.

Several groups have demonstrated the utility of Solexa/Illumina platform in phage-display library screening and selection experiments.¹⁶⁰⁻¹⁷¹ Using Illumina, Fischer and co-workers assessed the quality and diversity of sequences in single chain antibody libraries developed using different strategies. They also monitored the evolution of antibody sequences after each round of phage panning and identified enriched antibody sequences against the target antigen.¹⁶⁰ Dunnen and co-workers first demonstrated that Illumina sequencing can successfully identify binder sequences after first round of selection. The authors employed Illumina NGS to characterize Ph.D.-7 phage display library before and after several rounds of biopanning on osteoblast cells.¹⁶¹ More recently, Börner and co-workers did a comparative study using phage-display screening combined with Illumina and Sanger sequencing to identify peptide based binders for polypropylene. In this study, the authors showed that after just one round of panning Illumina identified arginine rich polypropylene binders that were not evident in Sanger sequencing.¹⁶⁵

In 2012, our group developed a protocol that converted a library of plasmids isolated from the phage library to double stranded DNA sequences suitable for Illumina sequencing using one-step PCR.¹⁷² Our Liquid Glycan Array platform also uses Illumina deep sequencing to decode the structure and density of the recognized glycans.¹¹⁴ In this chapter, I will discuss the validation of ABO LiGA with deep sequencing in further detail.

3.1.2. DNA encoding strategies

Multiplex deep sequencing is a powerful approach in which multiple DNA samples can be mixed together and sequenced in a single run using high-throughput sequencing platforms instead of analyzing one sample at a time. Multiplexing in sequencing is possible because of DNA tags that can identify specific sequences in a mixture and assign them back to original samples. Known as DNA barcodes, or barcodes for short, these DNA tags are a distinctive DNA sequence commonly used to identify a molecule of interest or for encoding information in a larger region of DNA.¹⁷³ For example, combinatorial libraries like ribosome display libraries, phage display libraries, DNA-encoded libraries (DELs) comprise a large variety of protein variants and/or small molecules individually connected to DNA sequences that code for them.¹⁷⁴

Every barcode design relies on the simple combinatorial rule that, with a given number of bases ‘a’ (typically 4 in DNA barcodes) and defined length of the sequence ‘b’, the total number of combinations is a^b . This rule identifies the minimum required length of the barcode that can be used to generate a sufficient number of barcoded primers for all samples. Several barcoding strategies described in literature differ in the strategy for selection of barcodes out of all possible combinations.¹⁷⁵⁻¹⁸⁶ DNA barcode libraries can be grouped into two categories, i.) randomly generated DNA libraries and ii.) rationally designed libraries. In randomly generated libraries, oligos are physically assembled in pools whereas rationally designed libraries are designed using computer simulations and then manufactured.¹⁷³ Commercially available barcoded primers from Illumina belong to the *in silico* category.¹⁸⁷ Hamming distance and Levenshtein distance, based on the popular Hamming¹⁸⁸ and Levenshtein¹⁸⁹

codes respectively, are the two common ways of rationally computing the sequence difference between DNA barcodes to ensure that each barcode can be identified unambiguously even if the PCR or sequencing procedure introduces random errors into such barcode. The rationale behind these codes and strategies for error correction are beyond the scope of this thesis and can be found elsewhere in literature.^{183, 188-190}

In 2015, our group developed phage-displayed peptide libraries with chemical post-translational modifications encoded using a “silent barcoding” approach.¹⁴⁶ This approach leverages the degeneracy of codons to translate chemically identical peptide sequence from genetically distinct DNA sequence. The authors used silent barcodes to trace specific sequences in a mixed modified library. Our group has demonstrated the utility of silent barcoding in several projects.^{114, 191-197} The LiGA platform is also based on silent encoding of glycans where each phage clone is attached to a distinct glycan. The concept of MSDBs used in my project is also based on silent encoding of same glycan using multiple DNA barcodes.

3.1.3. Errors and biases in DNA sequencing and NGS screening

Illumina deep sequencing uses a “sequencing by synthesis” approach where one end of the DNA molecule is attached to a chip and nucleotides are chemically added to the immobilized DNA strand.¹⁹⁸ This chemical addition of new nucleotides on a growing DNA strand can introduce various errors: “insertion” of a nucleotide at an incorrect position and “deletion” of a nucleotide from a desired position, giving rise to “indel” errors, addition of a

nucleotide other than the intended one yielding a “substitution” error, and sometimes, loss of chemical reactivity at the end of growing strand leading to “termination”. Growing strands of DNA in any given spot on the chip may undergo one or more of these errors leading to single or multiple types of variations for one target sequence. To overcome these errors, Illumina introduced paired-end sequencing where both ends of a target region are sequenced and the reads are aligned.¹⁹⁹ Paired-end sequencing produces twice the number of reads compared to single-read data which helps in detecting the indel variants.

Along with sequencing errors, factors such as selection parameters, library size, sample handling, sample processing and PCR amplifiability can affect the outcome of a screening experiment. To ensure a higher success rate of a screening campaign, it is important to implement statistics that would take all possible variables into consideration. Several groups have worked on the development of statistical methods for the evaluation and comparison of HTS experiments. Some of these methods include the use of negative binomial distribution²⁰⁰, Poisson distribution^{201, 202}, Z’ score metric^{118, 203} and more recently, machine learning approaches^{204, 205}. In my project, I use Z’ score as a statistical parameter to evaluate the robustness of LiGA-based assay to profile carbohydrate binding blood group antibodies. Calculation of Z’ score in MSDB ABO LiGA experiments is shown in the results and discussion section of this chapter.

3.1.4. PCR bias

NGS is usually preceded by PCR amplification of DNA strands with target sequence. Depending on the PCR method and DNA sequence, each DNA molecule is replicated by a factor of about 1.6 - 1.8.²⁰⁶ PCR in itself is understood as a “high-fidelity” process, however, it is not error-free. Previous studies have shown that PCR is known to preferentially amplify

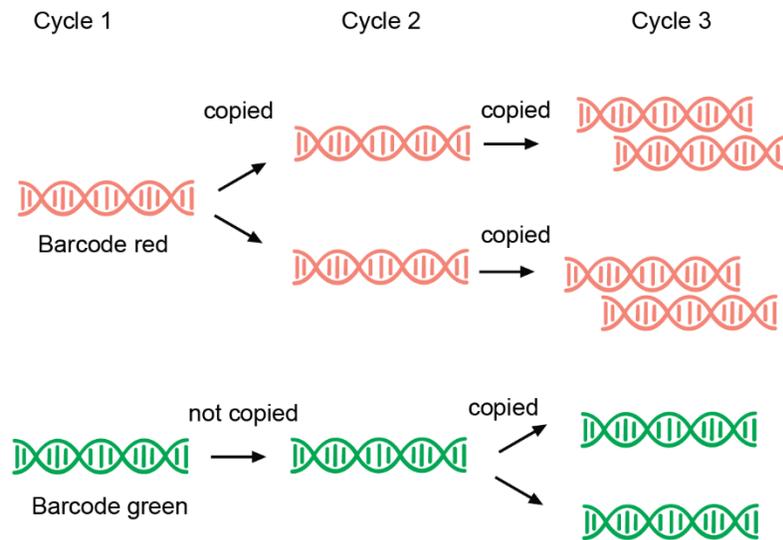


Figure 3-1: Illustration of PCR stochastic bias

some sequences over others, which can distort the distribution of copy number of individual sequences within a mixture.²⁰⁷⁻²⁰⁹ For example, Strauss and co-workers showed that PCR bias is related to copy number of the template, and the bias is particularly severe for sequences with low copy numbers in the initial mixture.²¹⁰ This phenomenon is known as PCR stochastic bias. It is based on the concept that every molecule in PCR undergoes replication with a probability less than 1. For example, if probability of amplification is 0.9, PCR will fail to replicate one out of ten molecules amplified every cycle. This is not a problem for

high-copy number templates, but it can be concerning when certain sequences in a mixture are present in very low copy numbers. For example, consider the red barcode (Figure 3-1) that undergoes replication in the first cycle. In cycle 2 there are two copies and if it undergoes replication again, cycle 3 will have 4 copies. In contrast, the green barcode (Figure 3-1) that fails to get amplified in both cycles will just have one copy at the end of cycle 3. Even if in subsequent cycles, both barcodes get amplified equally, the red barcode will appear at a copy number of four times more than that of green barcode. Strauss and co-workers demonstrated PCR stochastic bias by performing a serial dilution-PCR experiment. The authors arbitrarily chose a DNA pool with 7,373 sequences and diluted to different copy numbers ranging from 8 to 113. Each dilution was PCR amplified using a two-step PCR. First step was done with Illumina PCR primers with overhangs and second step installed the Illumina adapters. Post-PCR Illumina sequencing of the dilution series showed that lower the copy number, lower was the sequencing coverage. This shows that PCR stochastic bias can significantly affect the outcome of a screening experiment where low copy number samples are involved. In my project, having a mixture of 10 clones encoding the same glycan allows us to better understand the different biases originating from PCR and sequencing and their effect on low-copy number phage clones in a LiGA-based diagnostics assay.

3.2. Results and discussion

3.2.1. Design and validation of a semi-nested PCR for amplification of MSDB library

PCR amplification of phage samples recovered from a binding assay is an important step that must be performed before Illumina sequencing. Amplification of phage samples obtained from m-LiGA binding assay on anti-A and anti-B antibodies was done using the PCR protocol published¹¹⁴ by our lab. The primers used in this protocol amplifies the Silent double barcode (SDB) region as well as append the Illumina adapters to the barcode region in a “one-step” process. However, using this protocol, amplification of phage samples that had a pre-PCR titer of 10^4 copies was not successful (Figure 3-4a). Experiments performed by lab member Dr. Alexey Atrazhev revealed that PCR amplification of low copy phage samples using primers with Illumina adapters resulted in the formation of “primer-dimers” and no detectable amplification of target sequence was observed. To overcome this challenge, Dr. Atrazhev designed a “semi-nested” PCR method for phage amplicons with titer $<10^4$ PFU. Traditionally, a nested PCR involves an initial set of thermal cycles with a pair of primers spaced relatively far apart from the target sequence. Then, a small amount of output from the first round of PCR is used as a template for the second step and is amplified using a set of primers that are located between the outer primers. When one of the two inner primers is used as an outer primer in the initial amplification, it is known as “semi-nested” PCR. The first step of our protocol involves a quantitative PCR (qPCR) assay since qPCR has higher sensitivity compared to traditional PCR and allows for easy real-time quantification of phage titers even before performing a PFU assay.

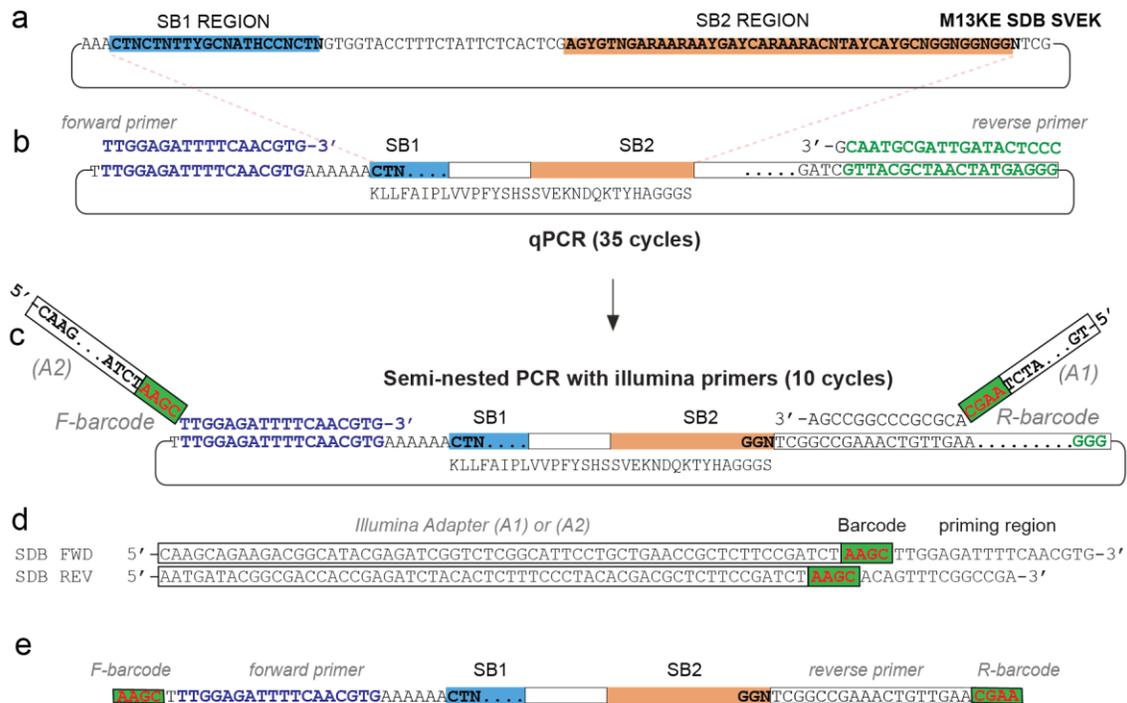


Figure 3-2: Scheme of qPCR and semi-nested PCR amplification **a**, Each SDB clone contains two unique sites in M13-SDB vector termed silent barcode (SB1) and (SB2 regions) **b**, first step qPCR with SDB forward primer without the Illumina overhangs and an outer reverse primer **c**, Second step semi-nested PCR with inner Illumina primers **d**, Sequences of the Illumina sequencing primers **e**, Region visible after Illumina sequencing of the second-step PCR product

In contrast to the protocol previously published by our lab, since this method involves two steps, I will henceforth refer to the semi-nested PCR as “two-step PCR”. To validate the two-step protocol and evaluate its limit of detection, I performed a serial dilution qPCR consisting of LiGA copies ranging from $10 - 10^8$ PFU (Figure 3-3). qPCR quantification cycle (Cq) values revealed that phage amplicons with copy number as low as 100 could be detected using the designed qPCR assay.

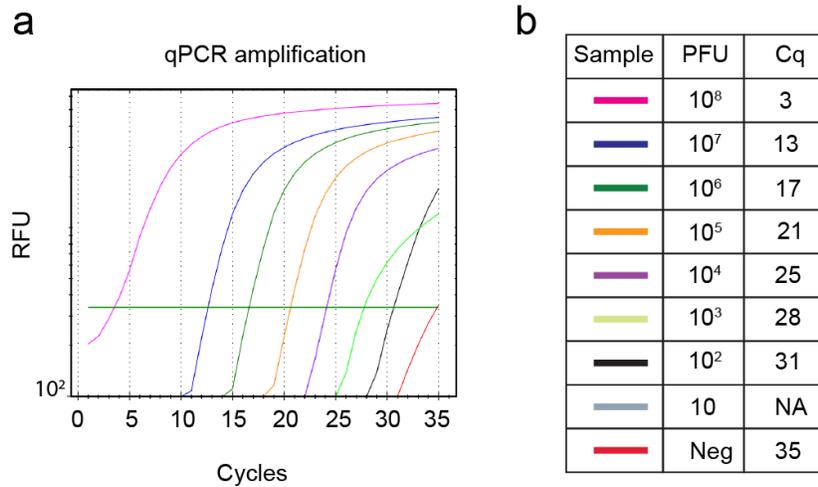


Figure 3-3: Serial dilution-qPCR of an MSDB ABO LiGA library **a**, qPCR amplification curve of different dilutions of a naïve MSDB library ranging from 10 to 10⁸ copies of phage calculated by PFU assay. **b**, Sample concentrations and their corresponding C_q values

As recovery of LiGA on control targets such as BSA or blank wells falls in the range of 10⁴ – 10³ PFU, all binding experiments in my project is done using two-step PCR. For example, Figure 3-4 shows the result of a binding experiment done with MSDB ABO LiGA. Two-step PCR resulted in successful amplification of samples with a copy number of 5000.

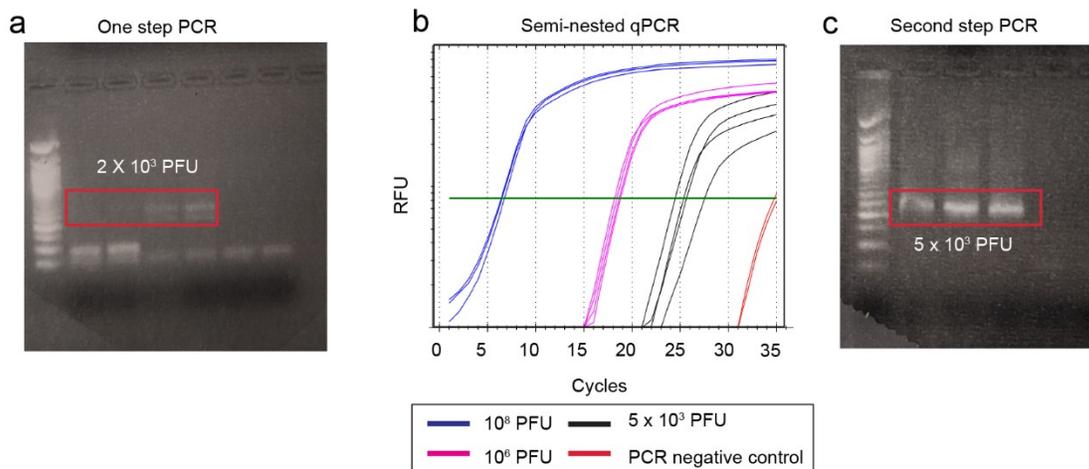


Figure 3-4: Comparison between one step and two step semi-nested PCR for low copy number phages **a**, Gel electrophoresis image showing bands (highlighted in red) after one-step PCR amplification of phage with 2000 copies as PCR template **b**, qPCR amplification curve of phage samples recovered after a panning experiment **c**, Gel electrophoresis image showing bands after second-step semi nested PCR of low copy number phage with Illumina sequencing primers

3.2.2. Comparison of sequencing performance of ABO MSDB LiGA amplified by one vs. two-step PCR

This experiment was performed to systematically investigate the effect of two different PCR methods on the composition of an MSDB mixture after Illumina sequencing. MSDB ABO LiGA consisting of 5 sets of MSDBs, each conjugated to an ABO glycan or azidoethanol and non-conjugated phages were mixed together. The total copy number of the PCR template was chosen to be 10^6 since it can be successfully amplified by both one and two-step PCR. Parts per million (PPM) reads of individual clones in a MSDB mixture is shown in Fig. 3.5 for both one and two-step PCR.

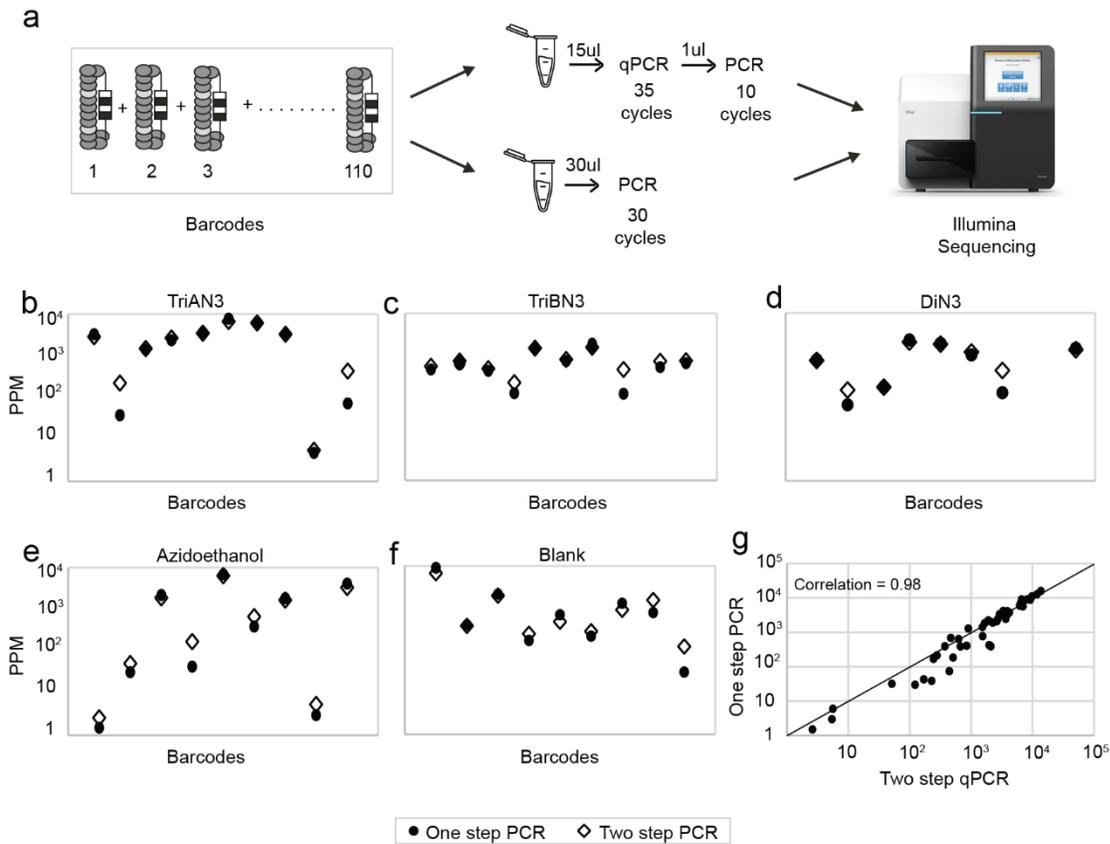


Figure 3-5: Comparison of Illumina deep sequencing results after one and two-step PCR
a, Illustration showing the sequencing workflow after PCR of a mixture consisting of 5 sets of MSDBs with 10^5 - 10^6 templates used as input, each conjugated to either **b**, blood group antigen A (Tri-AN3), **c**, blood group antigen B (Tri-BN3), **d**, blood group antigen H (Di-N3), or **e**, azidoethanol, **f**, unconjugated “blank” phages **g**, MSDB ABO library PCR amplified using one step and two-step PCR showed little deviation of each clone’s contribution to the total library sequenced with correlation coefficient of 0.98.

Since individual clones in a MSDB mixture were mixed uniformly based on the phage titer in a 1:1:1...:1 ratio, the “ideal” expectation would be to see uniformity in clones by Illumina sequencing too. However, as shown in Figure 3-5, this was not true for all the MSDB

mixtures. This could be because of biases introduced during sample processing, PCR and Illumina sequencing. However, further QC studies such as qPCR of individual phage clones in an MSDB could be potentially performed to better understand the library composition.

While comparing one-step to the two-step PCR, small deviation from one-step PCR was observed with a correlation coefficient of 0.98.

3.2.3. Screening m-LiGA on anti-A and anti-B antibodies

In this experiment, binding of ABO LiGA to purified anti-A and anti-B IgM antibodies was tested using a focused library consisting of 5 sets of MSDBs. Each set of MSDB consisted of 10 phage clone with a unique DNA barcode and conjugated individually to Tri-AN3, Tri-BN3, Di-N3 blood group glycans or azidoethanol (Figure 2-2 in Chapter 2 that describes glycan structures). Non-conjugated “blank” and azidoethanol MSDBs were used as controls. The individual MSDB glycoconjugates were mixed together to yield MSDB ABO LiGA. Selection of phage with binding activity to a target protein depends on the strategy used to immobilize the target protein on a surface. I designed this experiment to test the binding of MSDB ABO LiGA to anti-A and anti-B antibodies immobilized using two methods – i.) passive adsorption of antibodies on polystyrene 96 well plates (plate panning) and ii.) incubation with Protein-L magnetic beads that are known to bind to mouse IgM antibodies (bead panning).

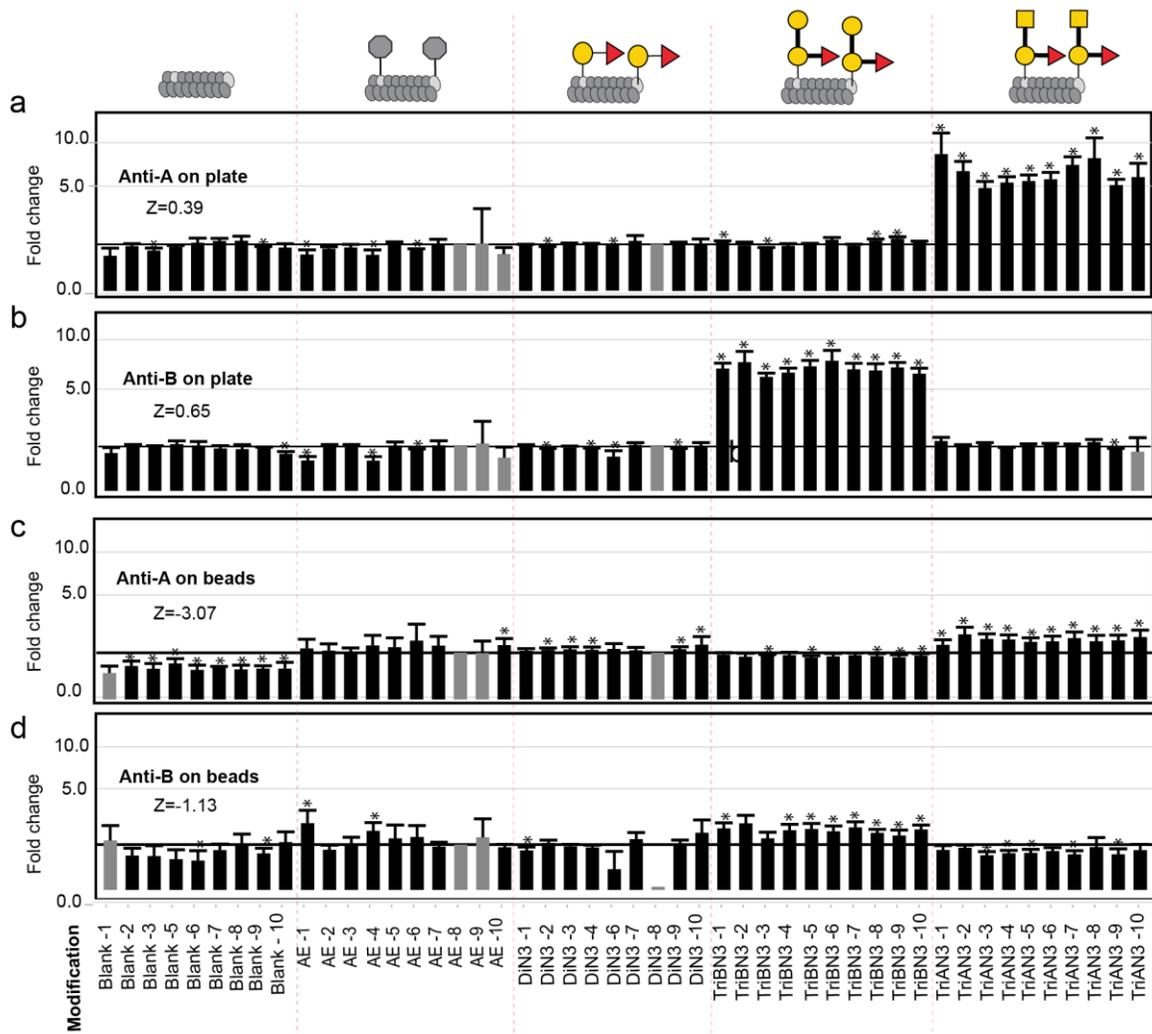


Figure 3-6: Deep sequencing analysis and Z' score calculation of m-LiGA with purified anti-A and anti-B IgM antibodies coated on **a,b**, plates and **c,d**, beads

Plate panning was done manually using a multichannel pipette, whereas bead panning was done using an automated magnetic bead handling system (Kingfisher). In plate panning, bound phages were recovered by elution with 10 mM HCl and in bead panning, boiling the beads released DNA of the bound phage. DE analysis confirmed the binding of Tri-AN3 and Tri-BN3 MSDBs to anti-A and anti-B antibodies adsorbed on the plate surface respectively

(Figure 3-6). However, this enrichment was dramatically attenuated antibodies were coated on Protein L magnetic beads. The enrichment of Tri-AN3 on bead-immobilized anti-A or Tri-BN3 on bead immobilized anti-B was statistically significant but the fold enrichment (FC) factor decreased from ~5 to < 2. Binding results were compared using the Z' score, commonly employed in the evaluation of HTS assay robustness. Plate panning yielded Z' score >0.5 (0.57) for anti-B antibody and >0 (0.39) for anti-A antibody, indicating robust signal by HTS metrics. Z' scores for bead panning on both anti-A and anti-B antibodies were <1, which means no significant difference in binding was observed between the test and the control MSDBs. This may be due to weak or insufficient binding of anti-A and anti-B IgM antibodies to Protein L magnetic beads. An SDS PAGE assay can be done in the future to further investigate this. A limitation of bead panning is that only 12 samples can be tested at once on a 12-channel Kingfisher instrument, in contrast to plate panning where 96 samples can be assayed simultaneously. Based on the results obtained from this experiment and a higher throughput when compared to beads, all future experiments were done by coating purified antibodies or serum on plates.

3.2.4. Screening a complex LiGA library on anti-A and anti-B antibodies

In the previous experiment, I screened a focused library consisting of MSDBs conjugated to ABO glycans. In this experiment, I mixed MSDB ABO LiGA with another LiGA library composed of 65 different glycopHage constructs (produced by lab member Mirat Sojitra).

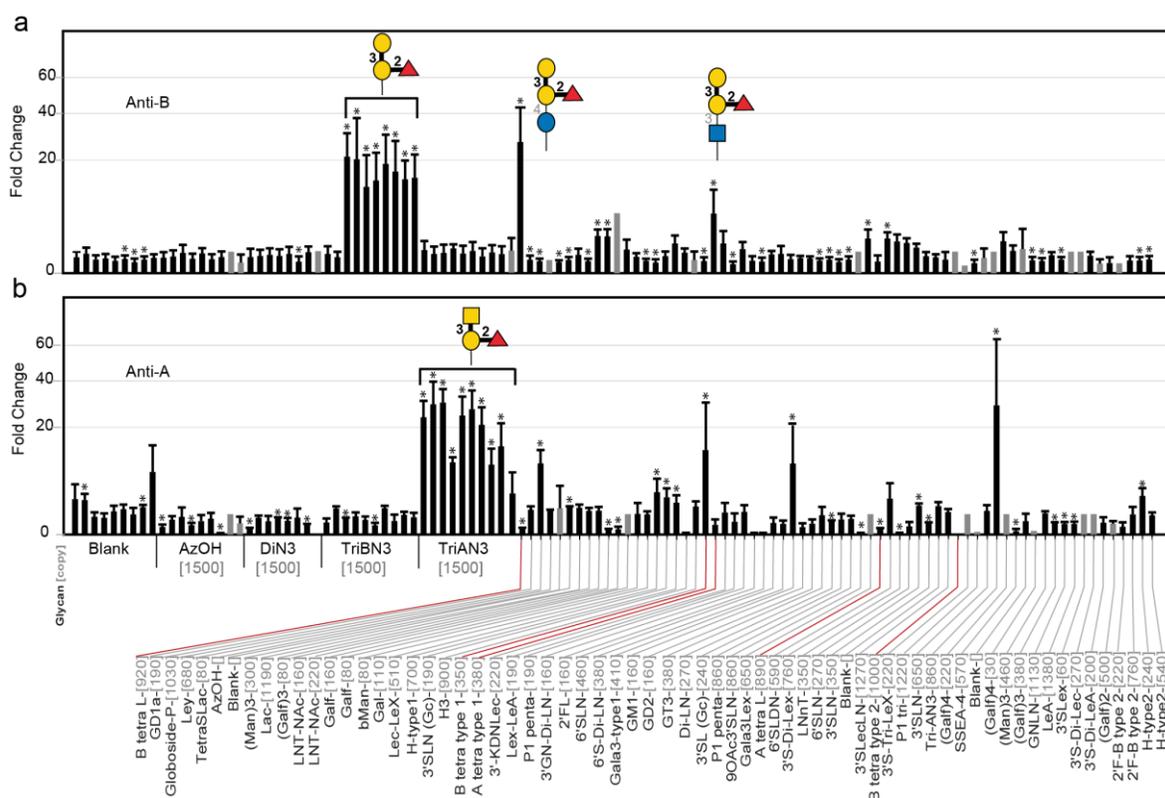


Figure 3-7: Screening m-LiGA library mixed with LiGA-65 on **a**, anti-A and **b**, anti-B antibody. Red lines highlight ABO tetrasaccharide subtypes present in LiGA library.

This experiment was done to understand the behaviour of ABO MSDBs in a complex library consisting of glycans that are not known to bind to anti-A and anti-B antibodies along with A and B tetrasaccharide antigens belonging to subtype I and II. Tri-AN3[1500] MSDB showed significant binding when tested on anti-A antibody but not on anti-B antibody. Similarly, Tri-BN3 [1500] was enriched when tested on anti-B antibody but not on anti-A antibody. This shows that the specificity of Tri-AN3 and Tri-BN3 MSDBs towards their respective target antibodies is preserved even in the mixture with ~70 glycosylated phage clones that display non-ABO epitopes.

Along with Tri-BN3 [1500], B tetrasaccharide [920] also showed significant binding on anti-B antibody. B tetra type 1 [350] showed a modest yet significant enrichment, whereas no binding was observed for B tetra type 2 [1000]. For anti-A antibody, none of the A tetrasaccharide subtypes in the library showed any significant binding. Some groups have studied the correlation of different blood group subtypes with anti-A and anti-B IgG and IgM antibodies. For example, using a neoglycoprotein array, Gildersleeve and co-workers demonstrated that type 2 blood group tetrasaccharides show better correlation with ABO IgG and IgM antibodies in human serum than the trisaccharide determinants. The West lab used a printed ABO glycan microarray to study the different subtype specific blood group antibodies in human serum. For a more direct comparison of the results of my experiment with other similar studies, factors such as the density of the conjugated glycans, array design and detection methods would have to be considered. Subtype specific antibody binding studies are beyond the scope of this thesis but can be potentially investigated in the future with an ABO LiGA consisting of different subtype glycans conjugated at different densities.

3.2.5. Avidity based response of m-LiGA

As described in the publication from our group, LiGA measured density dependent binding of glycans to purified lectins and lectins on cells *in vitro* and *in vivo*. To systematically investigate the importance of density in my project, I created 3 sets of MSDBs for each blood group antigen.

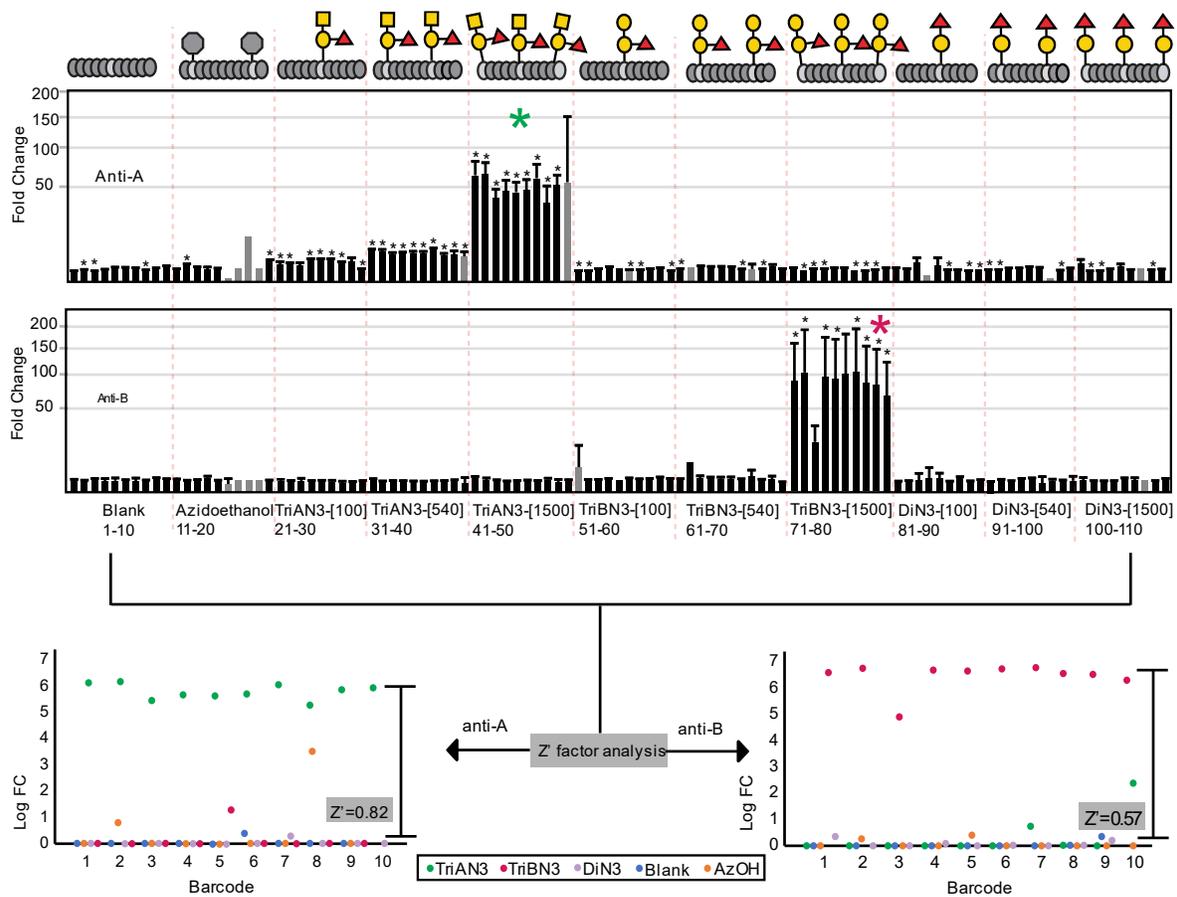


Figure 3-8: Density scan of m-LiGA on a, anti-A and b, anti-B antibodies c, Z' score analysis of MSDB ABO LiGA screening on anti-A antibody d, Z' score analysis of MSDB ABO LiGA on anti-B antibody.

Each set of MSDB was conjugated to blood group glycan at either low (90-120 copies), medium (500-700 copies) or high (1500-1800 copies) density. MSDB phage decorated with high copy number of Tri-AN3 and Tri-BN3 exhibited stronger binding to anti-A and anti-B antibodies respectively, when compared to MSDBs that contained low or medium copy numbers (Figure 3-8). Tri-AN3 [1500] in anti-A screening had a Z' score of 0.82 whereas Tri-BN3 [1500] in anti-B screening had a score of 0.57. Since a Z'-score > 0.5 indicates that

an assay is robust by HTS metrics, it can be concluded that MSDB ABO LiGA can decouple both affinity- and avidity-based observations.

3.3. Conclusion

Based on PCR and Illumina sequencing, I evaluated the functional performance of MSDB ABO LiGA as well as calculated Z' factor to determine the robustness of a LiGA-based diagnostic assay. In the next chapter, I demonstrate the utility of MSDB-ABO LiGA to profile blood group antibodies from human serum samples and comparison with two other existing glycan array platforms.

3.4. Materials and methods

3.4.1. Binding assay of mLiGA on 96 well plates

A 1:1000 dilution of murine monoclonal IgM anti-A and anti-B antibodies (produced by Immucor Inc. and donated by Lori West at the University of Alberta) was added to a 96-well plate (Corning, CLS3369), which was sealed with tape (Thermo Scientific, 15036) and incubated overnight at 4 °C. The following day, wells were washed three times with washing buffer, incubated with blocking solution (1 h at room temperature) and washed three times with washing buffer. Solutions of m-LiGA was added to the wells (50 µl, 10¹⁰ PFU/mL in PBS) and incubated for 1.5 h at room temperature. Wells were washed three times with washing buffer. To elute bound phage, 50 µl HCl (pH 2.0) was added to the wells and incubated for 9 min at room temperature, and the content of each well was transferred to a

microcentrifuge tube containing 25 μ L 5 \times Phusion HF buffer (NEB, M0530S). The neutralized solution was used for titering and as template for the PCR reaction.

3.4.2. Binding assay of mLiGA on Protein L beads

A 30 μ L suspension of Protein L magnetic beads (ThermoFisher, 88850) was rinsed with wash buffer (1 mL). 1 μ L of anti-A and anti-B antibody was added to the bead suspension in PBS buffer (1 mL). The mixture was incubated overnight at 4 $^{\circ}$ C with mixing. The following day, the bead suspension and other reagents were added to a 96-well Deepwell plate (Thermo Fisher, 95040450), and the KingFisher Duo Prime Purification System was used to perform the experiment with the following steps: (1) wash buffer (1 mL), (2) blocking buffer (1 mL), (3) binding with m-LiGA (1 mL), (4) wash buffer (1 mL), (5) wash buffer (1 mL) and (6) PBS buffer. Beads were transferred to microcentrifuge tubes and centrifuged (500g for 1 min). Pelleted beads were resuspended in nuclease-free water (30 μ L); 2 μ L of the suspension was used in the PFU assay. The remaining solution was incubated at 90 $^{\circ}$ C for 15 min and centrifuged at 21,000g for 10 min, and 25 μ L of the supernatant was used as template for the PCR reaction.

3.4.3. One-step PCR Protocol and Illumina sequencing

This protocol is adapted from the publication¹¹⁴. 25 mL of DNA template solution after bead panning procedure in Nuclease free water and 15 mL of eluted phage after plate panning procedure was amplified in total volume of 50 mL with 1x Phusion[®] buffer, 50 mM

of each dNTPs, 500 mM MgCl₂, 1 mM forward barcoded primer, 1 mM reverse barcoded primer and one unit of Phusion® High-Fidelity DNA Polymerase (NEB, #M0530S). In amplification of naïve libraries, volume of template (phage solution) was 2 µL. Cycling was performed using the following thermocycler settings: a) 98 °C 3 min, b) 98 °C 10 s, c) 50 °C 20 s, d) 72 °C 30 s, e) repeat b)-d) for 10 cycles, f) 98 °C 10 s, g) 72 °C 30s, h) repeat f)-g) for 20 cycles, h) 72 °C 5 min, i) 4 °C hold. The PCR products were quantified by 2% (w/v) agarose gel in Tris-Borate-EDTA buffer at 100 volts for ~35 min using a low molecular weight DNA ladder as standard (NEB, #N3233S). PCR products that contain different indexing barcodes were pooled allowing 10 ng of each product in the mixture. The mixture was purified by eGel, quantified by Qubit (Thermo Fisher) and sequenced using the Illumina NextSeq paired-end 500/550 High Output Kit v2.5 (2x75 Cycles). Data was automatically uploaded to BaseSpace™ Sequence Hub. Processing of the data is described in section “*Processing of Illumina data*”.

3.4.4. Two-step semi-nested PCR

15 mL of eluted phage after plate panning procedure was used as qPCR template and amplified in total volume of 50 µL with 1x Phusion® buffer, 50 mM of each dNTPs, 500 mM MgCl₂, 1 mM forward qPCR primer, 1 mM reverse qPCR primer and one unit of Phusion® High-Fidelity DNA Polymerase (NEB, #M0530S). In amplification of naïve libraries, volume of template (phage solution) was 2 µL. Bio-Rad C1000 qPCR machine was used for DNA amplification. Cycling was performed using the following thermocycler settings: a) 95 °C 3 min, b) 95 °C 10 s, c) 58 °C 30 s, d) 72 °C 20 s (with picture), e) 80 °C

20 s (with picture) f) repeat b)-e) for 34 cycles, g) Melt curve 65 °C – 95 °C (with picture) h) end. qPCR data was analyzed using Bio-Rad CFX Manager software. For the second-step semi-nested PCR, 1-5 µL of the qPCR product was used as DNA template and amplified in total volume of 50 µL with 1x Phusion® buffer, 50 mM of each dNTPs, 50 mM MgCl₂, 1 mM forward qPCR primer, 1 mM reverse qPCR primer and one unit of Phusion® High-Fidelity DNA Polymerase (NEB, #M0530S). Amplification was performed using the following thermocycler settings: a) 95 °C 3 min, b) 95 °C 10 s, c) 58 °C 30 s, d) 72 °C 20 s e) repeat b-d for 9 cycles f) 72 °C 5 min g) hold 4 °C. The PCR products were quantified by 2% (w/v) agarose gel in Tris-Borate-EDTA buffer at 100 volts for ~35 min using a low molecular weight DNA ladder as standard (NEB, #N3233S).

3.4.5. Processing of Illumina data

This protocol is adapted from the publication¹¹⁴. The Gzip compressed FASTQ files were downloaded from BaseSpace™ Sequence Hub. The files were converted into tables of DNA sequences and their counts per experiment, essentially as previously described¹⁵⁸. Briefly, FASTQ files were parsed into separate files based on unique multiplexing barcodes within the reads. Reads that did not contain an identifiable multiplex barcode were discarded. Reads that contained a Phred=0 quality score in any position were also discarded. Additional quality control was based on: (i) mapping the forward and reverse primer regions allowing no more than one base substitution each, (ii) alignment of the forward and reverse read-end overlap, allowing no mismatches between forward and reverse read in the overlap region. Reads that did not match criteria (i) and (ii) were discarded. The two ends of read-pairs that

pass the filtering criteria were joined and trimmed to the DNA sequences located between the priming regions; the reads were organized in a tab-delimited text file containing the unique DNA sequences and their copy numbers. Technical replicates were often combined in the same file. Using an SDB-lookup table, we mapped DNA sequences to SDB and translated them to glycans using a LiGA-specific lookup table (“LiGA dictionary”). Any reads with one substitution away from the expected SDB sequence were assigned to the parent SDB sequence. Abbreviated names of glycans are based on those used on CFG website:(<http://www.functionalglycomics.org/static/consortium/resources/resourcecored2.shtml>). The files with DNA reads, raw counts, and mapped glycans were uploaded to <http://ligacloud.ca/> server. Each experiment has a unique alphanumeric name (e.g., **20180711-87YOrdRB-JM**) and unique static URL: for example, <http://ligacloud.ca/searchLibInfo?f=0&b=0&d=20180711-87YOrdRB-JM>.

Chapter 4: Testing of MSDB ABO LiGA on human serum samples

4.1. Introduction

In this chapter, I used MSDB ABO LiGA for the specific detection of blood group antibodies in human serum samples. Several glycan array platforms have been developed and used for detecting carbohydrate binding antibodies in human serum. These platforms have already been covered in Chapter 1.

Glycan arrays differ in ligand presentation, glycan origin (isolated from natural sources or chemically synthesized), assay conditions, detection method, and immobilization strategies - on flat surfaces (printed glycan array, ELISA), microspheres (Luminex bead array), DNA (DNA encoded glycan array), M13 phage (glycophage, LiGA) all of which contribute to the affinity, avidity and selectivity of binding. Given the diversity of the glycan array platforms, there is a crucial need for standardization and comparison between different glycan array formats. To date, only a few groups have compared different glycan array platforms. Of these, most studies have compared glass based glycan arrays to the traditional ELISA and only investigated the effect of sample size, dynamic range and sensitivity in signal detection.²¹¹⁻²¹⁴ A cross-platform comparison between different glass based glycan array formats was performed for the first time by the Mahal group in 2014.²¹⁵ The authors compared glycan arrays developed by 6 different groups – CFG, Huflejet, Gildersleeve, Joshi, Reichardt and Pieters. These arrays varied in size, density and composition. Each of the six research groups analyzed five biotinylated plant lectins (*concanavalin A*, *Helix*

pomatia agglutinin, Maackia amurensis lectin I, Sambucus nigra agglutinin and wheat germ agglutinin) and the results were compared using a “universal threshold” method. The strongest binding glycans were consistently detected across all glycan array platforms, as long as the epitope was present. However, this was not true for weakly binding glycans. The method by which “positive” signals are assigned on an array is known as thresholding. The authors found that a difference in thresholding values between different glycan arrays had a significant effect on the interpretation of results especially for weaker binders. Spacing and local concentration of glycans was another determining factor in detection of weak binders. For example, the authors observed that binding response of WGA to its monomeric ligand β -GlcNAc was only detected in densely conjugated neoglycoprotein array and not in the traditional glycan array printed on a glass slide. It is important to note that in this study, all six microarrays had the same detection system – using a fluorescently labelled streptavidin. With the emergence of newer technologies like the Luminex bead array and the DNA encoded glycan array that have a completely different detection system, there is a need for an improved method of cross-platform standardization.

Given the diversity and complexity in construction and detection methods of different glycan array platforms, systematic comparison between the different glycan array formats is a topic for doctoral studies in itself. In my project, I first used MSDB ABO LiGA for specific detection of blood group antibodies in human serum samples and then compared to three other glycan array platforms – traditional array printed on a glass slide, Luminex bead ABO glycan array (developed by the Lori West group at the University of Alberta) and

neoglycoprotein array (developed and published²¹⁶ by the Gildersleeve group at the NIH). The results of this study are covered in results section.

4.2. Experimental design

Serum samples used in this study were obtained from the West group at the University of Alberta and the Gildersleeve group at the NIH. A total of 20 samples from healthy volunteers belonging to different ABO blood types were analysed in this experiment. MSDB ABO LiGA consisted of 11 sets of MSDBs (a total of 110 clones) and each ABO glycan was displayed at 3 different densities – [1500 copies], [540 copies] and [120 copies]. LiGA-65 library (made by lab member Mirat Sojitra) composed of 65 different glycoprotein constructs. Sera samples were tested with both MSDB ABO LiGA and LiGA-65 mixed with MSDB ABO LiGA. Traditionally, blood group antibodies are known to belong to the IgM class, however, studies from the Gildersleeve²¹⁷ and West groups have also shown the presence of serum IgG antibodies to blood group determinants. In my project, I designed an assay for the detection of blood group antibodies belonging to both the IgG and the IgM class in the same experiment. This was done by coating the wells of 96-well polystyrene plates with a secondary human IgG or IgM antibody that could bind to antibodies in human serum. Sera were diluted 1:100 to ensure that majority of the antibodies are detected.

Serum samples obtained from the West lab were analyzed using 3 platforms - LiGA, glass array and Luminex bead array whereas samples obtained from the Gildersleeve group were analyzed using 4 platforms – LiGA, glass array, Luminex bead array and the neoglycoprotein array. Glass array and Luminex array experiments were done by the West lab and neoglycoprotein array experiments were done by the Gildersleeve group.

4.3. Results and discussion

4.3.1. Detection of blood group antibodies in human serum tested by MSDB ABO LiGA

Figure 4-1 shows the binding response of blood group glycans in the ABO LiGA mixture to IgM (Figure 4-1a) and IgG (Figure 4-1b) antibodies in the serum sample. IgG antibody signals to both A and B trisaccharides correlated to the blood type with a much lower background signal compared to the IgM signals. This may be due to multiple (10) binding sites in IgM as compared to 2 sites in IgG. Another important observation was that serum antibodies to blood groups did not correlate perfectly with the blood type. For example, individuals with blood type A would be “expected” to have anti-B antibodies, similarly blood type B individuals would have anti-A antibodies and individuals with blood type AB would not have either of the two antibodies. However, in this assay, both A (Tri-AN3[1500]) and B (Tri-BN3[1500]) trisaccharides showed binding to IgM antibodies in some individuals with blood type A. Few of the same individuals showed little to no signal for IgG antibodies. Individuals with blood-type B showed stronger IgM signal to Tri-AN3 [1500] and a modest response to Tri-AN3[540] and Tri-AN3[120]. Similar response was seen for IgG antibodies. All blood type O individuals showed binding to both antigen A (Tri-AN3[1500]) and antigen B (Tri-BN3[1500]) in the IgM class. For the same O type individuals, IgG signals to Tri-AN3[1500] were the strongest. Another important observation in this experiments was density-dependent binding of certain ABO glycans to blood group antibodies. In theory, individuals with blood type AB are not expected to either anti-A or anti-B antibodies.

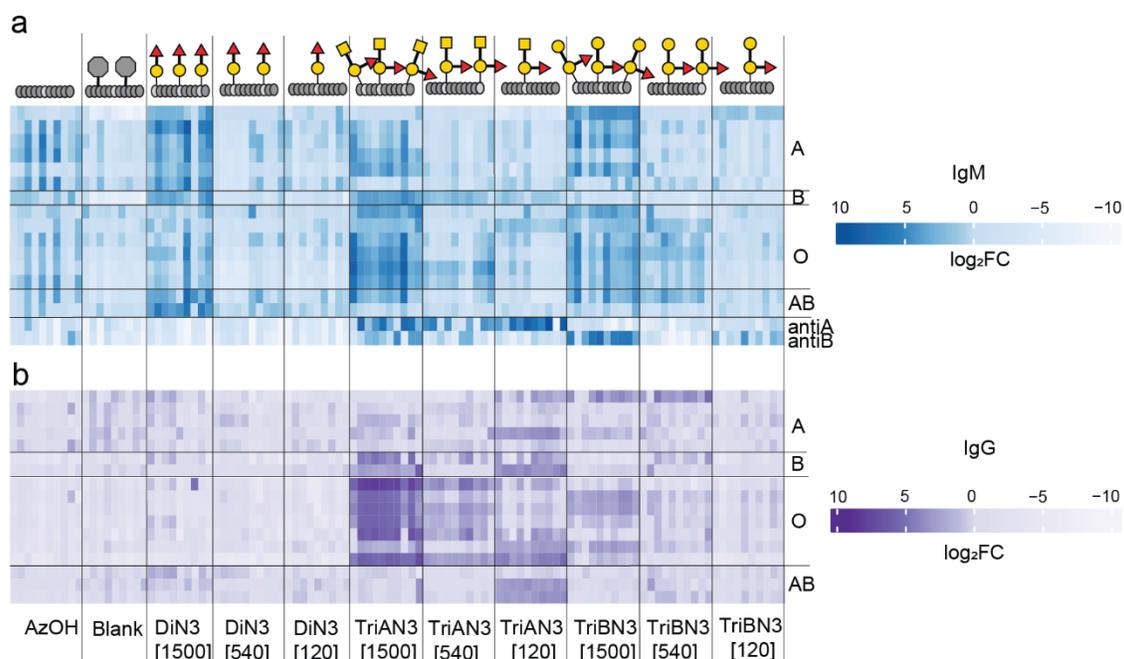


Figure 4-1: Heatmaps of sera antibody response to ABO LiGA **a**, IgM antibodies **b**, IgG antibodies. Subjects are clustered by blood-type in the rows. Tri-AN3, Tri-BN3 and Di-N3 glycans were conjugated at 3 densities and are organized in columns

In this assay, 2 out of 3 individuals of the AB blood type showed IgG binding to Tri-AN3[120] as opposed to IgM binding to Tri-AN3[1500]. Preference for Tri-AN3[120] in the IgG class was also seen in certain O type individuals.

4.3.2. Detection of blood group antibodies in human serum tested by ABO LiGA mixed with LiGA-65

ABO LiGA was mixed with LiGA-65 comprising 65 glycoconjugates produced from 56 different glycans. This experiment was done to evaluate the binding response of serum

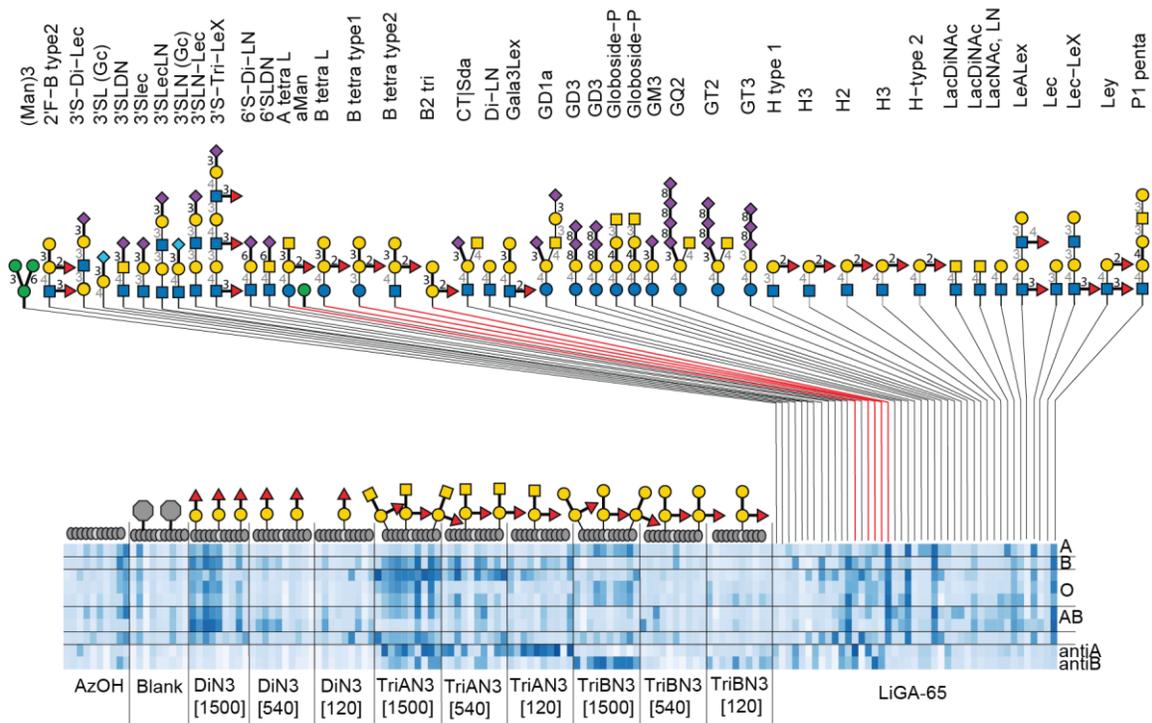


Figure 4-2: Heatmaps of sera IgM signals to ABO LiGA mixed with LiGA-65 Subjects are clustered by blood-type in the rows. Glycans highlighted in red represent the ABO blood group glycans in the LiGA-65 mixture. Heatmap shows comparison of log₂ FC values for different glycans obtained by DE analysis.

antibodies to the focused ABO LiGA in a complex mixture comprising majorly non-ABO glycans. 5 glycans in the LiGA-65 mixture belonged to the A and B subtypes – A tetra L (Tetrasaccharide A), B tetra L (Tetrasaccharide B), B tetra type I, B tetra type II and B2 tri (Trisaccharide B). These glycans are highlighted in red in Figure 4-2. Along with A and trisaccharides, purified anti-A and anti-B IgM antibodies showed strongest binding to A tetra L and B tetra type 1, consistent with our observation in the publication.¹¹⁴ As expected, anti-B antibody also showed some binding to B tetra type 1.

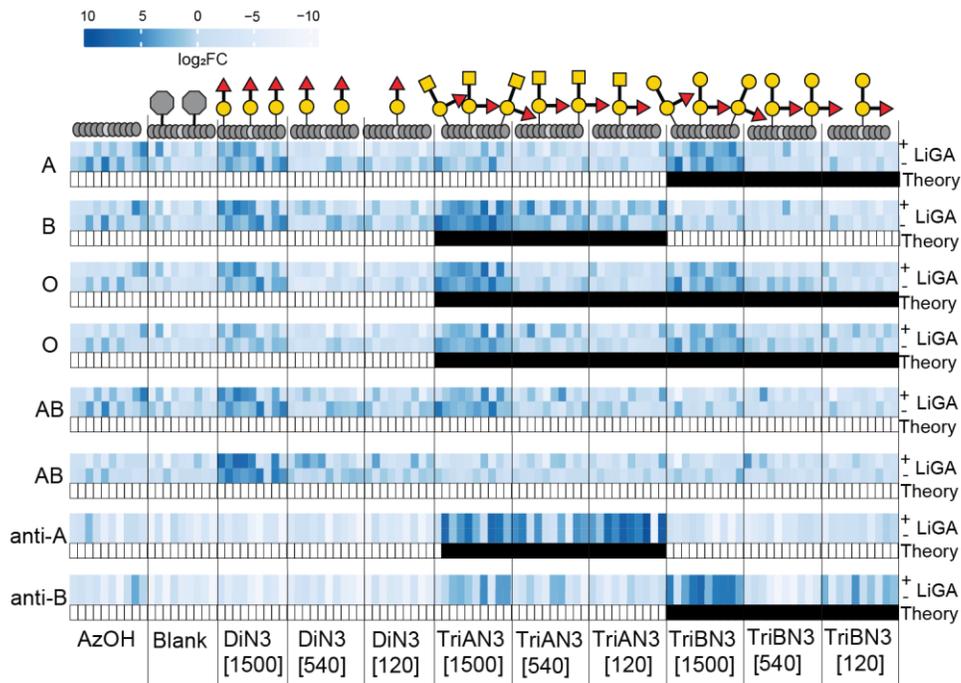


Figure 4-3: Comparison of IgM signals to ABO LiGA with ABO LiGA+LiGA-65 mixture. Black boxes indicate expected correlation of bloodtype to their respective antibodies.

Some non-specific binding to blood group tetrasaccharides were also observed. For example, anti-A antibody showed modest binding to B tetra L and B tetra type 2 and anti-B antibody to A tetra L. However, from the ABO LiGA mixture, only Tri-AN3 showed binding to anti-A antibody and Tri-BN3 to anti-B antibody. No detectable binding was observed to the A and B tetrasaccharides in most serum samples, except in two individuals with blood type O who showed some binding to B tetra type L and B tetra type 2. There was a general agreement in binding response when serum samples from the same individuals tested with the focused blood group library ABO LiGA were compared with signals obtained from ABO LiGA + LiGA-65 (Figure 4-3). This was observed even across different densities. For example, the individual with blood type B showed strong binding to Tri-AN3 [1500] and

modest binding to Tri-AN3 [540] and Tri-AN3 [120] when tested with ABO LiGA. The same pattern was observed when the sample was tested with LiGA-65.

4.3.3. Comparison of MSDB ABO LiGA with glass, neoglycoprotein and Luminex bead array

In this experiment, I compared the results of sera testing obtained by glass, Luminex and neoglycoprotein arrays with ABO LiGA (Figure 4-4). Binding responses to glycans other than the A and B trisaccharides and O disaccharide in the glass, Luminex and neoglycoprotein arrays were excluded from this comparison since my project only deals with the ABO di- and trisaccharides. Signal from the glass, Luminex and neoglycoprotein arrays were recorded as mean fluorescence intensity (MFI) and LiGA signals were recorded as $\log_2(\text{fold change})$. Binding of Tri-AN3 [540], Tri-AN3 [120] and Tri-BN3[540], Tri-BN3[120] in ABO LiGA were excluded from this comparison since Tri-AN3 [1500] and Tri-BN3 [1500] showed the strongest binding signal to both purified IgM anti-A and anti-B antibodies as well as ABO antibodies in sera. There was a general agreement between glycan-sera interactions measured by LiGA and binding of serum antibodies to glass, Luminex and neoglycoprotein arrays. Some individuals consistently showed a stronger IgM response across all arrays compared to their IgG response

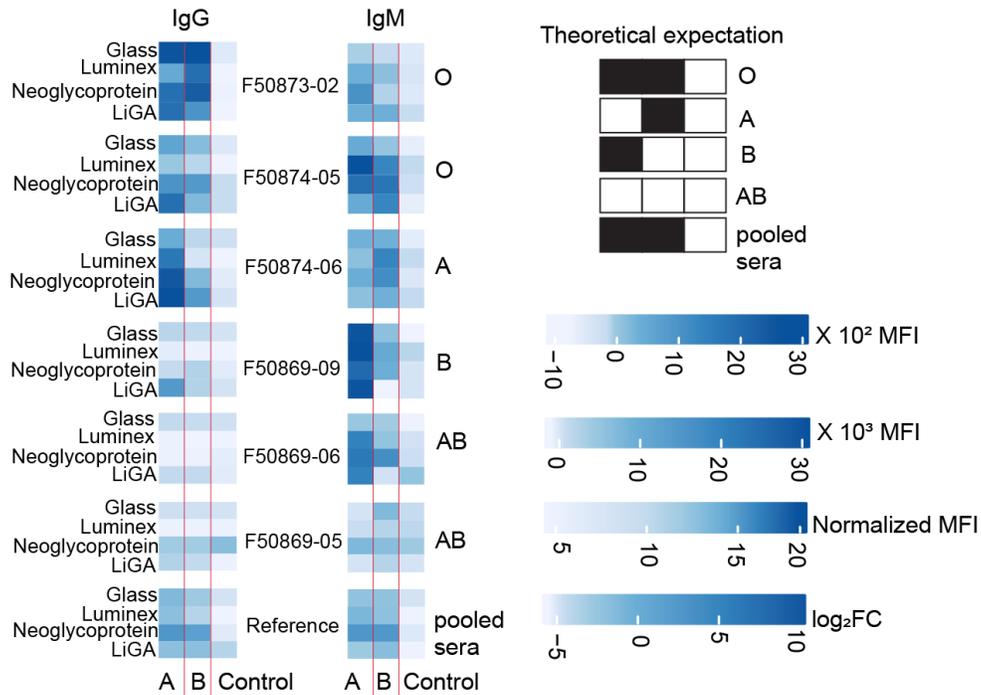


Figure 4-4: Heatmaps showing comparison of IgG and IgM signals tested by glass array, luminex array, neoglycoprotein array and ABO LiGA. Signals for glass, Luminex and neoglycoprotein arrays were recorded in mean fluorescence intensity (MFI) and log₂(Fold Change) for ABO LiGA. Black and white boxes indicate the “expected” antibody response for the ABO blood types.

For example, individuals with blood type AB are not expected to have any anti-A or anti-B antibodies but, in this experiment, binding was observed to both A and B trisaccharides in glass, Luminex and neoglycoprotein arrays and only to A trisaccharide in LiGA. There were some cross-platform differences that were observed between LiGA and other arrays. For example, the individual with blood type B showed a strong IgM signal to the A trisaccharide across all arrays and a modest signal to B trisaccharide for arrays other than LiGA. IgG

antibodies in the pooled serum, which is a mixture of ten sera samples obtained from individuals with different blood types showed non-specific binding to the control azidoethanol conjugated MSDBs as opposed to no binding seen in the control BSA conjugates for other arrays. This maybe due to the bulky DBCO linker used for glycan conjugation. The effect of linker on binding response can be potentially investigated in the future by conjugating ABO glycans to MSDBs with a linker other than DBCO, for e.g., sugars functionalized with carboxy groups.

4.4. Conclusion

In summary, I demonstrated the utility of LiGA for specific detection of IgG and IgM blood group antibodies from a complex mixture like the human serum. In addition, I also compared ABO LiGA to other existing glycan array platforms like the glass-based glycan array, neoglycoprotein array and Luminex bead array. However, I only focused on antibodies binding to the blood group trisaccharides. The West and the Gildersleeve group have reported the presence of subtype specific antibodies in human serum. Future experiments would involve constructing a panel of ABO LiGA consisting of all the 18 blood group subtypes and testing on human serum. I believe that my project laid the groundwork for LiGA based testing of antibodies in serum and can be expanded to applications like screening of antibodies in diseased human serum samples.

4.5. Materials and methods

4.5.1. Serum samples

All serum samples were generous gifts from the West group at the University of Alberta and the Gildersleeve group at the NIH. Table 4.1 summarizes information regarding the serum samples. Samples from the West group were collected from healthy volunteers at the University of Alberta hospital. Some samples contained heparin as an anticoagulant in them. Healthy donor sera from the Gildersleeve group were purchased from Valley Biomedical Products and Services (Winchester, VA). The reference pooled human sera was purchased from SeraCare Life Sciences (Milford, MA). Samples were stored in -20°C prior to use.

4.5.2. Microarray fabrication and binding assay

4.5.2.1. ABO glass array

This experiment was performed by Jean Pearcey from the West Lab at the University of Alberta. The glycan array consisted of blood group antigens A and B subtypes I-VI, A trisaccharide, B trisaccharide, O disaccharide and α Gal antigen and BSA conjugates as control. The chemistry, fabrication and characterization of ABO glass array has been previously described.²¹⁸⁻²²¹ The protocol described here is adapted from the publication²²². Microarray slides were printed at Engineering Arts LLC. Plasma samples were diluted (100 μ l at 1:100) in blocking buffer and incubated for 30 min at 37°C.

Sample ID	Blood type	Age	Sex	Antibody titer	Group
F50873-02	O	49	M	NA	Gildersleeve
F50869-05	AB	39	F	NA	Gildersleeve
F50874-05	O	20	F	NA	Gildersleeve
F50876-06	O	46	F	NA	Gildersleeve
F50873-03	A	42	M	NA	Gildersleeve
F50874-06	A	21	F	NA	Gildersleeve
F50869-06	AB	30	F	NA	Gildersleeve
F50866-03	B	31	M	NA	Gildersleeve
F50869-09	B	30	F	NA	Gildersleeve
541	O	44	F	High	West
655	A	34	F	Medium	West
688	AB	45	M	None	West
791	O	33	M	Medium	West
794	B	41	M	High	West
1029	O	45	F	Medium	West
1221	O	27	M	High	West
1446	O	22	F	High	West
1471	A	NA	M	High	West

Table 4.1: Information regarding serum samples used in this chapter

Bound antibodies were detected using fluorochrome-conjugated goat antihuman IgM or IgG secondary antibodies (109-505-008 DyLight 549 AffiniPure Goat Anti-Human IgG, Fc (γ) Fragment Specific; 109-495-129 Dylight 649 Affinipure Goat Anti-human IgM, Fc (5 μ) Fragment Specific, Jackson ImmunoResearch) at predetermined dilutions (0.3 μ g/ml) in blocking buffer. These affinity-chromatography purified secondary antibodies have been reported to bind specifically to the Fc fragment of the heavy chain of human IgG (all subclasses) and IgM.²³ Microarray slides were scanned using Nimblegen MS200 (Roche) at 5- μ m resolution and analyzed using ImaGene software (Biodiscovery). Normalized mean fluorescence intensities (MFI) were derived by subtracting local background fluorescence for individual spots and BSA-only spots; averages of triplicates were reported. Microarray printing protocols for dispensing BSA conjugates of ABH subtype I-VI antigens and α Gal were optimized based on spot morphology and antigen density using monoclonal antibodies specific for A and B subtype structures and α Gal.

4.5.2.2. Neoglycoprotein array

This experiment was performed by the Gildersleeve group at the NIH. Fabrication of the neoglycoprotein array used in this assay is described in the publication^{223, 224}. The protocol described here is adapted from the preprint²²⁵. The microarray contained 816 array components and included a variety of human glycans (N-linked glycans, O-linked glycans, and glycan portions of glycolipids), non-human glycans, glycopeptides, and glycoproteins. Each array component was printed in duplicate to produce a full array, and 8 copies of the full array were printed on each slide. Prior each experiment, each microarray slide was

scanned in an InnoScan 1100 AL fluorescence scanner to check for any defects and missing print spots. The slides were fitted with an 8-well module (Sigma-Aldrich) to allow 8 independent assays on each slide. In the assay, arrays were blocked with 3% BSA in PBS buffer (400 μ L/well) overnight at 4 °C, then washed six times with PBST buffer (PBS with 0.05% v/v Tween 20). Serum samples diluted at 1:50 in 3% BSA and 1% HSA in PBST were added onto each slide (100 μ L/well). To minimize technical variations, all samples were assayed in duplicate on separate slides. After agitation at 100 rpm for 4 hours at 37 °C, slides were washed six times with PBST (200 μ L/well). The bound serum antibodies were detected by incubating with Cy3 anti-Human IgG and DyLight 647 anti-human IgM (Jackson ImmunoResearch) at 3 μ g/mL in PBS buffer with 3% BSA and 1% HSA (100 μ L/well) under agitation at 37 °C for 2 hours. Slides were covered with aluminum foil to prevent photobleaching. After washing with PBST eight times (200 μ L/well), the slides were removed from the modules and soaked in PBST for 5 min prior to being dried by centrifugation at 1000 rpm ($112 \times g$) for 10 minutes. Slides were then scanned with an InnoScan 1100 AL (Innopsys) at 5 μ m resolution. The photomultiplier tube (PMT) settings were the same for all experiments to limit unintentional signal variation. Slides were scanned at “high” and “low” PMT settings (for the 532 nm laser, high pmt = 5 and low = 1; for the 635 nm laser, high = 25 and low = 9) to increase the dynamic range and appropriately scale-saturated components. The fluorescence intensity of each array spot was quantified with GenePix Pro 7 software (Molecular Devices). Any features marked as missing or defective in the prescan were excluded from further analysis. The local background corrected median was used for data analysis, and spots with intensity lower than 150 RFU (1/2 the typical IgM background) were set to 150.

4.5.2.3. ABO Luminex array

This experiment was performed by the West group at the University of Alberta. A and B antigens of subtypes I-VI, trisaccharide A, trisaccharide B, disaccharide O obtained from Dr. Todd Lowary's lab were conjugated to Luminex xMAP[®] beads and quantified using monoclonal ABO antibodies. BSA and α Gal antigen were coupled as negative and positive controls, respectively. The detailed protocol on coupling antigens to Luminex beads can be found in the Luminex xMAP cookbook²²⁶. IgG and IgM antibodies in sera with specificities for ABO glycans were probed using PE-conjugated Goat Anti-Human IgG and PE-conjugated Goat Human IgM respectively. IgG and IgM signals were recorded as mean fluorescence intensities (MFI).

4.5.2.4. MSDB ABO LiGA array

Construction and characterization of MSDB ABO LiGA is described in detail in Chapter 2. 1500, 540 or 120 copies of Tri-AN3, Tri-BN3 or Di-N3 was conjugated to each MSDB. Azidoethanol conjugated MSDB and non-conjugated blank MSDB were used as negative controls. The MSDBs were mixed together in a 1:1:1.....:1 ratio based on the phage titer obtained by PFU assay. LiGA-65 consisting of phage clones with 65 glycans was constructed and characterized by lab member Mirat Sojitra. MSDB ABO LiGA was mixed with LiGA-65 uniformly based on titer. AffiniPure Goat Anti-Human, Fc γ fragment specific IgG (Cat# 109-005-008) and AffiniPure Goat Anti-Human, Fc5 μ fragment specific IgM (Cat # 109-005-043) were purchased from Jackson ImmunoResearch Laboratories. IgG and IgM antibodies were diluted to 20 μ g/ml in 1X PBS and 50 μ l of the dilution solution was added

to a 96-well plate (Corning, CLS3369), which was sealed with tape (Thermo Scientific, 15036) and incubated in at room temperature for 2 h. After 2 h, wells were washed two times with washing buffer, incubated with blocking solution (1% BSA, 1 h at RT) and washed three times with washing buffer. Serum samples were diluted (1:100) in 1X PBS and 50 μ l was added to wells and incubated overnight at 4 °C. Next day, wells were washed two times with wash buffer. Solutions of MSDB ABO LiGA or MSDB ABO LiGA with LiGA-65 (50 μ l 10^8 PFU/ml in PBS) were added to the wells and incubated for 1.5 h at RT. Wells were washed three times with washing buffer. To elute bound phage, 50 μ l HCl (pH 2.0) was added to the wells and incubated for 9 min at room temperature, and the content of each well was transferred to a microcentrifuge tube containing 25 μ l 5 \times Phusion HF buffer (NEB, M0530S). The neutralized solution was used for titering and as template for PCR reaction.

4.5.3. PCR protocol and Illumina sequencing

PCR and Illumina sequencing was performed using the methods described in Chapter 3.

4.5.4. Generation of heatmaps

Output .txt file after DE analysis (described in Chapter 3) was exported as a .csv file. $\text{Log}_2(\text{FC})$ values were used for plotting the heatmap in R studio. I modified the open source script obtained from Bioconductor Complex Heatmap Library²²⁷ as follows:

```

library(tidyr)
library(ggplot2)
library(ComplexHeatmap)
library(circlize)

trial.df <- read.csv(file = "Gildersleeve IgM LiGA avg.csv",
check.names = FALSE)

trial.df$Barcode <- factor(trial.df$Barcode, levels =
unique(trial.df$Barcode), ordered = TRUE)

long.trial <- gather(trial.df, key = Sample, value = logFC,
"869 05":"874 05")

dist <- dist(t(trial.df[ ,-1]) , diag=TRUE)

hc <- hclust(dist)

plot (hc)

mat = (as.matrix(t(trial.df[ ,-1])))

rownames (mat) = colnames(trial.df)[-1]

colnames (mat) = trial.df$Barcode

col_fun = colorRamp2(c(300, 2000, 5000, 10000, 20000, 30000),
c("#eff3ff", "#c6dbef", "#9ecae1", "#6baed6", "#3182bd",
"#08519c")) #BLUE

pdf("Gildersleeve IgM LiGA avg.pdf", width = 10)

Heatmap (mat, column_order = trial.df$Barcode, col = col_fun)

#width = unit(15, "cm"), height = unit(8, "cm"))

#print(heatmap)

```

References:

1. Apweiler, R., On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1999**, *1473* (1), 4-8.
2. Matzinger, P., Tolerance, Danger, and the Extended Family. *Annu. Rev. Immunol.* **1994**, *12* (1), 991-1045.
3. Gallucci, S.; Matzinger, P., Danger signals: SOS to the immune system. *Curr. Opin. Immunol.* **2001**, *13* (1), 114-119.
4. Medzhitov, R.; Janeway, C. A., Innate Immunity: The Virtues of a Nonclonal System of Recognition. *Cell* **1997**, *91* (3), 295-298.
5. Kiwamoto, T.; Katoh, T.; Evans, C. M.; Janssen, W. J.; Brummet, M. E.; Hudson, S. A.; Zhu, Z.; Tiemeyer, M.; Bochner, B. S., Endogenous airway mucins carry glycans that bind Siglec-F and induce eosinophil apoptosis. *J. Allergy Clin. Immunol.* **2015**, *135* (5), 1329-1340.e9.
6. Ge, X. N.; Ha, S. G.; Greenberg, Y. G.; Rao, A.; Bastan, I.; Blidner, A. G.; Rao, S. P.; Rabinovich, G. A.; Sriramarao, P., Regulation of eosinophilia and allergic airway inflammation by the glycan-binding protein galectin-1. *Proceedings of the National Academy of Sciences* **2016**, *113* (33), E4837-E4846.
7. Butvilovskaya, V. I.; Smoldovskaya, O. V.; Feyzkhanova, G. U.; Filippova, M. A.; Pavlushkina, L. V.; Voloshin, S. A.; Rubina, A. Y., Modification of Anti-Glycan IgG and IgM Profiles in Allergic Inflammation. *Mol. Biol.* **2018**, *52* (4), 548-555.

8. Carlin, A. F.; Uchiyama, S.; Chang, Y.-C.; Lewis, A. L.; Nizet, V.; Varki, A., Molecular mimicry of host sialylated glycans allows a bacterial pathogen to engage neutrophil Siglec-9 and dampen the innate immune response. *Blood* **2009**, *113* (14), 3333-3336.
9. Khatua, B.; Ghoshal, A.; Bhattacharya, K.; Mandal, C.; Saha, B.; Crocker, P. R.; Mandal, C., Sialic acids acquired by *Pseudomonas aeruginosa* are involved in reduced complement deposition and siglec mediated host-cell recognition. *FEBS Lett.* **2010**, *584* (3), 555-561.
10. Sterner, E.; Flanagan, N.; Gildersleeve, J. C., Perspectives on Anti-Glycan Antibodies Gleaned from Development of a Community Resource Database. *ACS Chemical Biology* **2016**, *11* (7), 1773-1783.
11. Cheung, N.-K. V.; Cheung, I. Y.; Kramer, K.; Modak, S.; Kuk, D.; Pandit-Taskar, N.; Chamberlain, E.; Ostrovskaya, I.; Kushner, B. H., Key role for myeloid cells: Phase II results of anti-GD2 antibody 3F8 plus granulocyte-macrophage colony-stimulating factor for chemoresistant osteomedullary neuroblastoma. *Int. J. Cancer* **2014**, *135* (9), 2199-2205.
12. Yu, A. L.; Gilman, A. L.; Ozkaynak, M. F.; London, W. B.; Kreissman, S. G.; Chen, H. X.; Smith, M.; Anderson, B.; Villablanca, J. G.; Matthay, K. K.; Shimada, H.; Grupp, S. A.; Seeger, R.; Reynolds, C. P.; Buxton, A.; Reisfeld, R. A.; Gillies, S. D.; Cohn, S. L.; Maris, J. M.; Sondel, P. M., Anti-GD2 Antibody with GM-CSF, Interleukin-2, and Isotretinoin for Neuroblastoma. *New Engl. J. Med.* **2010**, *363* (14), 1324-1334.
13. Houghton, A. N.; Mintzer, D.; Cordon-Cardo, C.; Welt, S.; Fliegel, B.; Vadhan, S.; Carswell, E.; Melamed, M. R.; Oettgen, H. F.; Old, L. J., Mouse monoclonal IgG3

antibody detecting GD3 ganglioside: a phase I trial in patients with malignant melanoma.

Proceedings of the National Academy of Sciences **1985**, 82 (4), 1242-1246.

14. Scott, A. M.; Lee, F.-T.; Hopkins, W.; Cebon, J. S.; Wheatley, J. M.; Liu, Z.; Smyth, F. E.; Murone, C.; Sturrock, S.; MacGregor, D.; Hanai, N.; Inoue, K.; Yamasaki, M.; Brechbiel, M. W.; Davis, I. D.; Murphy, R.; Hannah, A.; Lim-Joon, M.; Chan, T.; Chong, G.; Ritter, G.; Hoffman, E. W.; Burgess, A. W.; Old, L. J., Specific Targeting, Biodistribution, and Lack of Immunogenicity of Chimeric Anti-GD3 Monoclonal Antibody KM871 in Patients With Metastatic Melanoma: Results of a Phase I Trial. *Journal of Clinical Oncology* **2001**, 19 (19), 3976-3987.

15. Livingston, P. O.; Hood, C.; Krug, L. M.; Warren, N.; Kris, M. G.; Brezicka, T.; Ragupathi, G., Selection of GM2, fucosyl GM1, globo H and polysialic acid as targets on small cell lung cancers for antibody mediated immunotherapy. *Cancer Immunol., Immunother.* **2005**, 54 (10), 1018-1025.

16. Lindholm, L., Immunohistological Detection of Fucosyl-G, ni Ganglioside in Human Lung Cancer and Normal Tissues with Monoclonal Antibodies¹. *Cancer Res.* **1989**, 49, 1300-1305.

17. Nilsson, O.; Brezicka, F.; Holmgren, J.; Sörenson, S.; Svennerholm, L.; Yngvason, F.; Lindholm, L., Detection of a ganglioside antigen associated with small cell lung carcinomas using monoclonal antibodies directed against fucosyl-GM1. *Cancer Res.* **1986**, 46 (3), 1403-1407.

18. Scott, A. M.; Tebbutt, N.; Lee, F.-T.; Cavicchiolo, T.; Liu, Z.; Gill, S.; Poon, A. M. T.; Hopkins, W.; Smyth, F. E.; Murone, C.; Macgregor, D.; Papenfuss, A. T.; Chappell, B.; Saunder, T. H.; Brechbiel, M. W.; Davis, I. D.; Murphy, R.; Chong, G.;

- Hoffman, E. W.; Old, L. J., A Phase I Biodistribution and Pharmacokinetic Trial of Humanized Monoclonal Antibody Hu3s193 in Patients with Advanced Epithelial Cancers that Express the Lewis-Y Antigen. *Clin. Cancer. Res.* **2007**, *13* (11), 3286-3292.
19. Posey, J. A.; Khazaeli, M. B.; Bookman, M. A.; Nowrouzi, A.; Grizzle, W. E.; Thornton, J.; Carey, D. E.; Lorenz, J. M.; Sing, A. P.; Siegall, C. B.; LoBuglio, A. F.; Saleh, M. N., A Phase I Trial of the Single-Chain Immunotoxin SGN-10 (BR96 sFv-PE40) in Patients with Advanced Solid Tumors. *Clin. Cancer. Res.* **2002**, *8* (10), 3092-3099.
20. Pai, L. H.; Wittes, R.; Setser, A.; Willingham, M. C.; Pastan, I., Treatment of advanced solid tumors with immunotoxin LMB-1: An antibody linked to Pseudomonas exotoxin. *Nat. Med.* **1996**, *2* (3), 350-353.
21. Ross, H. J.; Hart, L. L.; Swanson, P. M.; Rarick, M. U.; Figlin, R. A.; Jacobs, A. D.; McCune, D. E.; Rosenberg, A. H.; Baron, A. D.; Grove, L. E., A randomized, multicenter study to determine the safety and efficacy of the immunoconjugate SGN-15 plus docetaxel for the treatment of non-small cell lung carcinoma. *Lung Cancer* **2006**, *54* (1), 69-77.
22. Boeck, S.; Haas, M.; Laubender, R. P.; Kullmann, F.; Klose, C.; Bruns, C. J.; Wilkowski, R.; Stieber, P.; Holdenrieder, S.; Buchner, H.; Mansmann, U.; Heinemann, V., Application of a Time-Varying Covariate Model to the Analysis of CA 19-9 as Serum Biomarker in Patients with Advanced Pancreatic Cancer. *Clin. Cancer. Res.* **2010**, *16* (3), 986-994.
23. Hess, V.; Glimelius, B.; Grawe, P.; Dietrich, D.; Bodoky, G.; Ruhstaller, T.; Bajetta, E.; Saletti, P.; Figer, A.; Scheithauer, W., CA 19-9 tumour-marker response to

- chemotherapy in patients with advanced pancreatic cancer enrolled in a randomised controlled trial. *The lancet oncology* **2008**, *9* (2), 132-138.
24. Sawada, R.; Sun, S.-M.; Wu, X.; Hong, F.; Ragupathi, G.; Livingston, P. O.; Scholz, W. W., Human monoclonal antibodies to sialyl-Lewisa (CA19. 9) with potent CDC, ADCC, and antitumor activity. *Clin. Cancer. Res.* **2011**, *17* (5), 1024-1032.
25. Oriol, R., GENETIC CONTROL OF THE FUCOSYLATION OF ABH PRECURSOR CHAINS. EVIDENCE FOR NEW EPISTATIC INTERACTIONS IN DIFFERENT CELLS AND TISSUES. *Eur. J. Immunogenet.* **1990**, *17* (4-5), 235-245.
26. Clausen, H.; Hakomori, S.-I., ABH and Related Histo-Blood Group Antigens; Immunochemical Differences in Carrier Isotypes and Their Distribution. *Vox Sang.* **1989**, *56* (1), 1-20.
27. Ravn, V.; Dabelsteen, E., Tissue distribution of histo-blood group antigens. Review article. *APMIS* **2000**, *108* (1), 1-28.
28. Clausen, H.; Holmes, E.; Hakomori, S., Novel blood group H glycolipid antigens exclusively expressed in blood group A and AB erythrocytes (type 3 chain H). II. Differential conversion of different H substrates by A1 and A2 enzymes, and type 3 chain H expression in relation to secretor status. *J. Biol. Chem.* **1986**, *261* (3), 1388-1392.
29. Nosaka, M.; Ishida, Y.; Tanaka, A.; Hayashi, T.; Miyashita, T.; Kaminaka, C.; Eisenmenger, W.; Furukawa, F.; Kimura, A., Aberrant Expression of Histo-blood Group A Type 3 Antigens in Vascular Endothelial Cells in Inflammatory Sites. *Journal of Histochemistry & Cytochemistry* **2008**, *56* (3), 223-231.
30. Jeyakanthan, M.; Tao, K.; Zou, L.; Meloncelli, P. J.; Lowary, T. L.; Suzuki, K.; Boland, D.; Larsen, I.; Burch, M.; Shaw, N.; Beddows, K.; Addonizio, L.; Zuckerman,

- W.; Afzali, B.; Kim, D. H.; Mengel, M.; Shapiro, A. M. J.; West, L. J., Chemical Basis for Qualitative and Quantitative Differences Between ABO Blood Groups and Subgroups: Implications for Organ Transplantation. *Am. J. Transplantation* **2015**, *15* (10), 2602-2615.
31. Watkins, W., The ABO blood group system: historical background. *Transfusion medicine* **2001**, *11* (4), 243-265.
32. Gooi, H.; Feizi, T.; Kapadia, A.; Knowles, B.; Solter, D.; Evans, M., Stage-specific embryonic antigen involves α 1 \rightarrow 3 fucosylated type 2 blood group chains. *Nature* **1981**, *292* (5819), 156-158.
33. Kannagi, R.; Nudelman, E.; Levery, S. B.; Hakomori, S.-i., A series of human erythrocyte glycosphingolipids reacting to the monoclonal antibody directed to a developmentally regulated antigen SSEA-1. *J. Biol. Chem.* **1982**, *257* (24), 14865-14874.
34. Magnani, J. L.; Nilsson, B.; Brockhaus, M.; Zopf, D.; Stepkowski, Z.; Koprowski, H.; Ginsburg, V., A monoclonal antibody-defined antigen associated with gastrointestinal cancer is a ganglioside containing sialylated lacto-N-fucopentaose II. *J. Biol. Chem.* **1982**, *257* (23), 14365-14369.
35. Magnani, J. L.; Smith, D. F.; Ginsburg, V., Detection of gangliosides that bind cholera toxin: direct binding of ¹²⁵I-labeled toxin to thin-layer chromatograms. *Anal. Biochem.* **1980**, *109* (2), 399-402.
36. Oyelaran, O.; Gildersleeve, J. C., Glycan arrays: recent advances and future challenges. *Curr. Opin. Chem. Biol.* **2009**, *13* (4), 406-413.
37. Liao, B.; Tan, B.; Teo, G.; Zhang, P.; Choo, A.; Rudd, P. M., Shotgun Glycomics Identifies Tumor-Associated Glycan Ligands Bound by an Ovarian Carcinoma-Specific Monoclonal Antibody. *Scientific Reports* **2017**, *7* (1), 14489.

38. Song, X.; Lasanajak, Y.; Xia, B.; Heimbürg-Molinaro, J.; Rhea, J. M.; Ju, H.; Zhao, C.; Molinaro, R. J.; Cummings, R. D.; Smith, D. F., Shotgun glycomics: a microarray strategy for functional glycomics. *Nat. Methods* **2011**, *8* (1), 85-90.
39. Byrd-Leotis, L.; Liu, R.; Bradley, K. C.; Lasanajak, Y.; Cummings, S. F.; Song, X.; Heimbürg-Molinaro, J.; Galloway, S. E.; Culhane, M. R.; Smith, D. F., Shotgun glycomics of pig lung identifies natural endogenous receptors for influenza viruses. *Proceedings of the National Academy of Sciences* **2014**, *111* (22), E2241-E2250.
40. van Diepen, A.; van der Plas, A.-J.; Kozak, R. P.; Royle, L.; Dunne, D. W.; Hokke, C. H., Development of a *Schistosoma mansoni* shotgun O-glycan microarray and application to the discovery of new antigenic schistosome glycan motifs. *Int. J. Parasitol.* **2015**, *45* (7), 465-475.
41. Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Mickum, M. L.; Ashline, D. J.; Prasad, B. V.; Estes, M. K.; Reinhold, V. N., Human milk contains novel glycans that are potential decoy receptors for neonatal rotaviruses. *Molecular & cellular proteomics* **2014**, *13* (11), 2944-2960.
42. Aydin, S., A short history, principles, and types of ELISA, and our laboratory experience with peptide/protein analyses using ELISA. *Peptides* **2015**, *72*, 4-15.
43. Engvall, E.; Perlmann, P., Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. *Immunochemistry* 1971, *8* (9), 871-874.
44. Van Weemen, B. K.; Schuurs, A. H. W. M., Immunoassay using antigen-enzyme conjugates. *FEBS Lett.* **1971**, *15* (3), 232-236.
45. Serology testing for COVID-19 antibodies <https://www.cdc.gov/coronavirus/2019-ncov/lab/serology-testing.html>

46. EUA authorized serology test performance <https://www.fda.gov/medical-devices/coronavirus-disease-2019-covid-19-emergency-use-authorizations-medical-devices/eua-authorized-serology-test-performance>
47. MacMullan, M. A.; Ibrayeva, A.; Trettner, K.; Deming, L.; Das, S.; Tran, F.; Moreno, J. R.; Casian, J. G.; Chellamuthu, P.; Kraft, J.; Kozak, K.; Turner, F. E.; Slepnev, V. I.; Le Page, L. M., ELISA detection of SARS-CoV-2 antibodies in saliva. *Scientific Reports* **2020**, *10* (1), 20818.
48. Buchs, J.-P.; Nydegger, U. E.. Development of an ABO-ELISA for the quantitation of human blood group anti-A and anti-B IgM and IgG antibodies. *Journal of Immunological Methods* 1989, *118* (1), 37–46.
49. Rieben, R.; Buchs, J.; Fluckiger, E.; Nydegger, U., Antibodies to histo-blood group substances A and B: agglutination titers, Ig class, and IgG subclasses in healthy persons of different age categories. *Transfusion* **1991**, *31* (7), 607-615.
50. Dotan, N.; Altstock, R.; Schwarz, M.; Dukler, A., Anti-Glycan Antibodies as Biomarkers for Diagnosis and Prognosis. *LUPUS* **2006**, *15* (7), 442-450.
51. Kamm, F.; Strauch, U.; Degenhardt, F.; Lopez, R.; Kunst, C.; Rogler, G.; Franke, A.; Klebl, F.; Rieders, F., Serum anti-glycan-antibodies in relatives of patients with inflammatory bowel disease. *PLOS ONE* **2018**, *13* (3), e0194222.
52. Rejchrt, S.; Drahošová, M.; Kopáčová, M.; Cyrany, J.; Doua, T.; Pintér, M.; Bureš, J., Antilaminaribioside and antichitobioside antibodies in inflammatory bowel disease. *Folia Microbiol.* **2008**, *53* (4), 373-376.

53. Menge, T.; Lalive, P. H.; Von Büdingen, H.-C.; Cree, B.; Hauser, S. L.; Genain, C. P., Antibody responses against galactocerebroside are potential stage-specific biomarkers in multiple sclerosis. *J. Allergy Clin. Immunol.* **2005**, *116* (2), 453-459.
54. Lolli, F.; Mulinacci, B.; Carotenuto, A.; Bonetti, B.; Sabatino, G.; Mazzanti, B.; D'Ursi, A. M.; Novellino, E.; Pazzagli, M.; Lovato, L.; Alcaro, M. C.; Peroni, E.; Pozo-Carrero, M. C.; Nuti, F.; Battistini, L.; Borsellino, G.; Chelli, M.; Rovero, P.; Papini, A. M., An N-glycosylated peptide detecting disease-specific autoantibodies, biomarkers of multiple sclerosis. *Proceedings of the National Academy of Sciences* **2005**, *102* (29), 10273-10278.
55. Dotan, I.; Fishman, S.; Dgani, Y.; Schwartz, M.; Karban, A.; Lerner, A.; Weishauss, O.; Spector, L.; Shtevi, A.; Altstock, R. T.; Dotan, N.; Halpern, Z., Antibodies Against Laminaribioside and Chitobioside Are Novel Serologic Markers in Crohn's Disease. *Gastroenterology* **2006**, *131* (2), 366-378.
56. Wang, D.; Liu, S.; Trummer, B. J.; Deng, C.; Wang, A., Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells. *Nat. Biotechnol.* **2002**, *20* (3), 275-281.
57. Huflejt, M. E.; Vuskovic, M.; Vasiliu, D.; Xu, H.; Obukhova, P.; Shilova, N.; Tuzikov, A.; Galanina, O.; Arun, B.; Lu, K.; et al.. Anti-carbohydrate antibodies of normal sera: Findings, surprises and challenges. *Molecular Immunology* 2009, *46* (15), 3037–3049.
58. Gao, C.; Wei, M.; McKittrick, T. R.; McQuillan, A. M.; Heimburg-Molinaro, J.; Cummings, R. D., Glycan Microarrays as Chemical Tools for Identifying Glycan Recognition by Immune Proteins. *Frontiers in Chemistry* **2019**, *7*.

59. Geissner, A.; Seeberger, P. H., Glycan Arrays: From Basic Biochemical Research to Bioanalytical and Biomedical Applications. *Annu. Rev. Anal. Chem* **2016**, *9* (1), 223-247.
60. Huflejt, M. E.; Vuskovic, M.; Vasiliu, D.; Xu, H.; Obukhova, P.; Shilova, N.; Tuzikov, A.; Galanina, O.; Arun, B.; Lu, K.; et al.. Anti-carbohydrate antibodies of normal sera: Findings, surprises and challenges. *Molecular Immunology* 2009, *46* (15), 3037–3049.
61. Khasbiullina, N. R.; Shilova, N. V.; Navakouski, M. J.; Nokel, A. Y.; Blixt, O.; Kononov, L. O.; Knirel, Y. A.; Bovin, N. V., The Repertoire of Human Antiglycan Antibodies and Its Dynamics in the First Year of Life. *Biochemistry (Moscow)* **2019**, *84* (6), 608-616.
62. Heidepriem, J.; Dahlke, C.; Kobbe, R.; Santer, R.; Koch, T.; Fathi, A.; Seco, B.; Ly, M.; Schmiedel, S.; Schwinge, D.; Serna, S.; Sellrie, K.; Reichardt, N.-C.; Seeberger, P.; Addo, M.; Loeffler, F., Longitudinal Development of Antibody Responses in COVID-19 Patients of Different Severity with ELISA, Peptide, and Glycan Arrays: An Immunological Case Series. *Pathogens* **2021**, *10* (4), 438.
63. Robinson, W. H.; Digennaro, C.; Hueber, W.; Haab, B. B.; Kamachi, M.; Dean, E. J.; Fournel, S.; Fong, D.; Genovese, M. C.; De Vegvar, H. E. N.; Skriner, K.; Hirschberg, D. L.; Morris, R. I.; Muller, S.; Pruijn, G. J.; Van Venrooij, W. J.; Smolen, J. S.; Brown, P. O.; Steinman, L.; Utz, P. J., Autoantigen microarrays for multiplex characterization of autoantibody responses. *Nat. Med.* **2002**, *8* (3), 295-301.
64. Robinson, W. H.; Fontoura, P.; Lee, B. J.; De Vegvar, H. E. N.; Tom, J.; Pedotti, R.; Digennaro, C. D.; Mitchell, D. J.; Fong, D.; Ho, P. P. K.; Ruiz, P. J.; Maverakis, E.; Stevens, D. B.; Bernard, C. C. A.; Martin, R.; Kuchroo, V. K.; Van Noort, J. M.; Genain, C. P.; Amor, S.; Olsson, T.; Utz, P. J.; Garren, H.; Steinman, L., Protein

microarrays guide tolerizing DNA vaccine treatment of autoimmune encephalomyelitis.

Nat. Biotechnol. **2003**, *21* (9), 1033-1039.

65. Wang, D.; Lu, J., Glycan arrays lead to the discovery of autoimmunogenic activity of SARS-CoV. *Physiol. Genomics* **2004**, *18* (2), 245-248.

66. Wang, D.; Bhat, R.; Sobel, R. A.; Huang, W.; Wang, L.-X.; Olsson, T.; Steinman, L., Uncovering Cryptic Glycan Markers in Multiple Sclerosis (MS) and Experimental Autoimmune Encephalomyelitis (EAE). *Drug Dev. Res.* **2014**, n/a-n/a.

67. Comstock, L. E.; Kasper, D. L., Bacterial Glycans: Key Mediators of Diverse Host Immune Responses. *Cell* **2006**, *126* (5), 847-850.

68. Thirumalapura, N. R.; Morton, R. J.; Ramachandran, A.; Malayer, J. R., Lipopolysaccharide microarrays for the detection of antibodies. *J. Immunol. Methods* **2005**, *298* (1-2), 73-81.

69. Parthasarathy, N.; Deshazer, D.; England, M.; Waag, D. M., Polysaccharide microarray technology for the detection of *Burkholderia pseudomallei* and *Burkholderia mallei* antibodies. *Diagn. Microbiol. Infect. Dis.* **2006**, *56* (3), 329-332.

70. Parthasarathy, N.; Saksena, R.; Kováč, P.; Deshazer, D.; Peacock, S. J.; Wuthiekanun, V.; Heine, H. S.; Friedlander, A. M.; Cote, C. K.; Welkos, S. L.; Adamovicz, J. J.; Bavari, S.; Waag, D. M., Application of carbohydrate microarray technology for the detection of *Burkholderia pseudomallei*, *Bacillus anthracis* and *Francisella tularensis* antibodies. *Carbohydr. Res.* **2008**, *343* (16), 2783-2788.

71. Feinberg, H.; Taylor, M. E.; Razi, N.; McBride, R.; Knirel, Y. A.; Graham, S. A.; Drickamer, K.; Weis, W. I., Structural Basis for Langerin Recognition of Diverse

Pathogen and Mammalian Glycans through a Single Binding Site. *J. Mol. Biol.* **2011**, *405* (4), 1027-1039.

72. Blixt, O.; Hoffmann, J.; Svenson, S.; Norberg, T., Pathogen specific carbohydrate antigen microarrays: a chip for detection of Salmonella O-antigen specific antibodies.

Glycoconjugate J. **2008**, *25* (1), 27-36.

73. Baader, J.; Klapproth, H.; Bednar, S.; Brandstetter, T.; Ruhe, J.; Lehmann, M.; Freund, I., Polysaccharide microarrays with a CMOS based signal detection unit.

Biosensors and Bioelectronics **2011**, *26* (5), 1839-1846.

74. Wang, D.; Carroll, G. T.; Turro, N. J.; Koberstein, J. T.; Kovc, P.; Saksena, R.;

Adamo, R.; Herzenberg, L. A.; Herzenberg, L. A.; Steinman, L., Photogenerated glycan arrays identify immunogenic sugar moieties of Bacillus anthracis exosporium. *Proteomics*

2007, *7* (2), 180-184.

75. Borchers, A. T.; Keen, C. L.; Huntley, A. C.; Gershwin, M. E., Lyme disease: A rigorous review of diagnostic criteria and treatment. *J. Autoimmun.* **2015**, *57*, 82-115.

76. Chuang, P.-K.; Hsiao, M.; Hsu, T.-L.; Chang, C.-F.; Wu, C.-Y.; Chen, B.-R.;

Huang, H.-W.; Liao, K.-S.; Chen, C.-C.; Chen, C.-L.; Yang, S.-M.; Kuo, C. W.; Chen,

P.; Chiu, P.-T.; Chen, I. J.; Lai, J.-S.; Yu, C.-D. T.; Wong, C.-H., Signaling pathway of

globo-series glycosphingolipids and β 1,3-galactosyltransferase V (β 3GalT5) in breast

cancer. *Proceedings of the National Academy of Sciences* **2019**, *116* (9), 3518-3523.

77. Wang, C. C.; Huang, Y. L.; Ren, C. T.; Lin, C. W.; Hung, J. T.; Yu, J. C.; Yu,

A. L.; Wu, C. Y.; Wong, C. H., Glycan microarray of Globo H and related structures for

quantitative analysis of breast cancer. *Proceedings of the National Academy of Sciences*

2008, *105* (33), 11661-11666.

78. Danishefsky, S. J.; Shue, Y.-K.; Chang, M. N.; Wong, C.-H., Development of Globo-H Cancer Vaccine. *Acc. Chem. Res.* **2015**, *48* (3), 643-652.
79. Huang, Y. L.; Hung, J. T.; Cheung, S. K. C.; Lee, H. Y.; Chu, K. C.; Li, S. T.; Lin, Y. C.; Ren, C. T.; Cheng, T. J. R.; Hsu, T. L.; Yu, A. L.; Wu, C. Y.; Wong, C. H., Carbohydrate-based vaccines with a glycolipid adjuvant for breast cancer. *Proceedings of the National Academy of Sciences* **2013**, *110* (7), 2517-2522.
80. Walker, L. M.; Huber, M.; Doores, K. J.; Falkowska, E.; Pejchal, R.; Julien, J.-P.; Wang, S.-K.; Ramos, A.; Chan-Hui, P.-Y.; Moyle, M.; Mitcham, J. L.; Hammond, P. W.; Olsen, O. A.; Phung, P.; Fling, S.; Wong, C.-H.; Phogat, S.; Wrin, T.; Simek, M. D.; Principal Investigators, P. G.; Koff, W. C.; Wilson, I. A.; Burton, D. R.; Poignard, P., Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **2011**, *477* (7365), 466-470.
81. Adams, E. W.; Ratner, D. M.; Bokesch, H. R.; McMahon, J. B.; O'Keefe, B. R.; Seeberger, P. H., Oligosaccharide and Glycoprotein Microarrays as Tools in HIV Glycobiology. *Chem. Biol.* **2004**, *11* (6), 875-881.
82. Mouquet, H.; Scharf, L.; Euler, Z.; Liu, Y.; Eden, C.; Scheid, J. F.; Halper-Stromberg, A.; Gnanapragasam, P. N. P.; Spencer, D. I. R.; Seaman, M. S.; Schuitemaker, H.; Feizi, T.; Nussenzweig, M. C.; Bjorkman, P. J., Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. *Proceedings of the National Academy of Sciences* **2012**, *109* (47), E3268-E3277.
83. Manimala, J. C.; Roach, T. A.; Li, Z.; Gildersleeve, J. C., High-Throughput Carbohydrate Microarray Analysis of 24 Lectins. *Angew. Chem.* **2006**, *118* (22), 3689-3692.

84. Oyelaran, O.; Li, Q.; Farnsworth, D.; Gildersleeve, J. C., Microarrays with Varying Carbohydrate Density Reveal Distinct Subpopulations of Serum Antibodies. *Journal of Proteome Research* **2009**, *8* (7), 3529-3538.
85. Herzenberg LA, Parks D, Sahaf B, Perez O, Roederer M, Herzenberg LA. The history and future of the fluorescence activated cell sorter and flow cytometry: a view from Stanford. *Clin Chem*. 2002;48(10):1819-1827.
86. Graham, H.; Chandler, D. J.; Dunbar, S. A., The genesis and evolution of bead-based multiplexing. *Methods* **2019**, *158*, 2-11.
87. Kettman, J. R.; Davies, T.; Chandler, D.; Oliver, K. G.; Fulton, R. J., Classification and properties of 64 multiplexed microsphere sets. *Cytometry* **1998**, *33* (2), 234-243.
88. Elshal, M.; McCoy, J., Multiplex bead array assays: Performance evaluation and comparison of sensitivity to ELISA☆. *Methods* **2006**, *38* (4), 317-323.
89. Pochechueva, T.; Chinarev, A.; Spengler, M.; Korchagina, E.; Heinzelmann-Schwarz, V.; Bovin, N.; Rieben, R., Multiplex suspension array for human anti-carbohydrate antibody profiling. *Analyst* **2011**, *136* (3), 560-569.
90. Lucas, J. L.; Tacheny, E. A.; Ferris, A.; Galusha, M.; Srivastava, A. K.; Ganguly, A.; Williams, P. M.; Sachs, M. C.; Thurin, M.; Tricoli, J. V.; Ricker, W.; Gildersleeve, J. C., Development and validation of a Luminex assay for detection of a predictive biomarker for PROSTVAC-VF therapy. *PLOS ONE* **2017**, *12* (8), e0182739.

91. Campbell, C. T.; Gulley, J. L.; Oyelaran, O.; Hodge, J. W.; Schlom, J.; Gildersleeve, J. C., Serum Antibodies to Blood Group A Predict Survival on PROSTVAC-VF. *Clin. Cancer. Res.* **2013**, *19* (5), 1290-1299.
92. Purohit, S.; Li, T.; Guan, W.; Song, X.; Song, J.; Tian, Y.; Li, L.; Sharma, A.; Dun, B.; Mysona, D.; Ghamande, S.; Rungruang, B.; Cummings, R. D.; Wang, P. G.; She, J.-X., Multiplex glycan bead array for high throughput and high content analyses of glycan binding proteins. *Nature Communications* **2018**, *9* (1).
93. Halpin, A.; Zhou, J.; Pearcey, J.; Lowary, T. L.; Cairo, C. W.; Daskhan, G.; Motyka, B.; Maier, S.; West, L. J.. OR34 Modernizing ABO antibody detection: Development of a bead-based ABO antibody detection assay. *Human Immunology* **2018**, *79*, 40.
94. Brenner, S.; Lerner, R. A., Encoded combinatorial chemistry. *Proceedings of the National Academy of Sciences* **1992**, *89* (12), 5381-5383.
95. Lerner, R. A.; Brenner, S., DNA-Encoded Compound Libraries as Open Source: A Powerful Pathway to New Drugs. *Angew. Chem. Int. Ed.* **2017**, *56* (5), 1164-1165.
96. Salamon, H.; Klika Škopić, M.; Jung, K.; Bugain, O.; Brunschweiger, A., Chemical Biology Probes from Advanced DNA-encoded Libraries. *ACS Chemical Biology* **2016**, *11* (2), 296-307.
97. Goodnow, R., DNA-Encoded Library Technology (DELTA) After a Quarter Century. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **2018**, *23* (5), 385-386.
98. Machutta, C. A.; Kollmann, C. S.; Lind, K. E.; Bai, X.; Chan, P. F.; Huang, J.; Ballell, L.; Belyanskaya, S.; Besra, G. S.; Barros-Aguirre, D.; Bates, R. H.; Centrella, P. A.; Chang, S. S.; Chai, J.; Choudhry, A. E.; Coffin, A.; Davie, C. P.; Deng, H.; Deng,

J.; Ding, Y.; Dodson, J. W.; Fosbenner, D. T.; Gao, E. N.; Graham, T. L.; Graybill, T. L.; Ingraham, K.; Johnson, W. P.; King, B. W.; Kwiatkowski, C. R.; Lelièvre, J.; Li, Y.; Liu, X.; Lu, Q.; Lehr, R.; Mendoza-Losana, A.; Martin, J.; McCloskey, L.; McCormick, P.; O’Keefe, H. P.; O’Keefe, T.; Pao, C.; Phelps, C. B.; Qi, H.; Rafferty, K.; Scavello, G. S.; Steinginga, M. S.; Sundersingh, F. S.; Sweitzer, S. M.; Szewczuk, L. M.; Taylor, A.; Toh, M. F.; Wang, J.; Wang, M.; Wilkins, D. J.; Xia, B.; Yao, G.; Zhang, J.; Zhou, J.; Donahue, C. P.; Messer, J. A.; Holmes, D.; Arico-Muendel, C. C.; Pope, A. J.; Gross, J. W.; Evindar, G., Prioritizing multiple therapeutic targets in parallel using automated DNA-encoded library screening. *Nature Communications* **2017**, *8* (1), 16081.

99. Song, M.; Hwang, G. T., DNA-Encoded Library Screening as Core Platform Technology in Drug Discovery: Its Synthetic Method Development and Applications in DEL Synthesis. *J. Med. Chem.* **2020**, *63* (13), 6578-6599.

100. Yuen, L. H.; Franzini, R. M., Achievements, Challenges, and Opportunities in DNA-Encoded Library Research: An Academic Point of View. *ChemBioChem* **2017**, *18* (9), 829-836.

101. Neri, D.; Lerner, R. A., DNA-Encoded Chemical Libraries: A Selection System Based on Endowing Organic Compounds with Amplifiable Information. *Annual Review of Biochemistry* **2018**, *87* (1), 479-502.

102. Flood, D. T.; Kingston, C.; Vantourout, J. C.; Dawson, P. E.; Baran, P. S., DNA Encoded Libraries: A Visitor's Guide. *Isr. J. Chem.* **2020**, *60* (3-4), 268-280.

103. Chevolut, Y.; Bouillon, C.; Vidal, S.; Morvan, F.; Meyer, A.; Cloarec, J.-P.; Jochum, A.; Praly, J.-P.; Vasseur, J.-J.; Souteyrand, E., DNA-Based Carbohydrate

Biochips: A Platform for Surface Glyco-Engineering. *Angewandte Chemie International Edition* 2007, 46 (14), 2398–2402.

104. (1) Kwon, S. J.; Lee, K. B.; Solakyildirim, K.; Masuko, S.; Ly, M.; Zhang, F.; Li, L.; Dordick, J. S.; Linhardt, R. J.. Signal Amplification by Glyco-qPCR for Ultrasensitive Detection of Carbohydrates: Applications in Glycobiology. *Angewandte Chemie* 2012, 124 (47), 11970–11974.

105. Kwon, S. J.; Jeong, E. J.; Yoo, Y. C.; Cai, C.; Yang, G.-H.; Lee, J. C.; Dordick, J. S.; Linhardt, R. J.; Lee, K. B., High Sensitivity Detection of Active Botulinum Neurotoxin by Glyco-Quantitative Polymerase Chain-Reaction. *Anal. Chem.* **2014**, 86 (5), 2279-2284.

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

107. Kondengaden, S. M.; Zhang, J.; Zhang, H.; Parameswaran, A.; Kondengadan, S. M.; Pawar, S.; Puthengot, A.; Sunderraman, R.; Song, J.; Polizzi, S. J.; Wen, L.; Wang, P. G., DNA Encoded Glycan Libraries as a next-generation tool for the study of glycan-protein interactions. Cold Spring Harbor Laboratory: 2020.

108. Bi, Q.; Cen, X.; Wang, W.; Zhao, X.; Wang, X.; Shen, T.; Zhu, S., A protein microarray prepared with phage-displayed antibody clones. *Biosensors and Bioelectronics* **2007**, 22 (12), 3278-3282.

109. Cekaite, L.; Haug, O.; Myklebost, O.; Aldrin, M.; Østenstad, B.; Holden, M.; Frigessi, A.; Hovig, E.; Sioud, M., Analysis of the humoral immune response to

immunoselected phage-displayed peptides by a microarray-based method. *Proteomics* **2004**, *4* (9), 2572-2582.

110. Zhong, L.; Hidalgo, G. E.; Stromberg, A. J.; Khattar, N. H.; Jett, J. R.; Hirschowitz, E. A., Using Protein Microarray as a Diagnostic Assay for Non–Small Cell Lung Cancer. *Am. J. Respir. Crit. Care Med.* **2005**, *172* (10), 1308-1314.

111. Celik, E.; Fisher, A. C.; Guarino, C.; Mansell, T. J.; DeLisa, M. P., A filamentous phage display system for N-linked glycoproteins. *Protein Sci.* **2010**, *19* (10), 2006-2013.

112. Dürr, C.; Nothaft, H.; Lizak, C.; Glockshuber, R.; Aebi, M., The Escherichia coli glycoprobe display system. *Glycobiology* **2010**, *20* (11), 1366-1372.

113. Çelik, E.; Ollis, A. A.; Lasanajak, Y.; Fisher, A. C.; Gür, G.; Smith, D. F.; DeLisa, M. P., Glycoarrays with engineered phages displaying structurally diverse oligosaccharides enable high-throughput detection of glycan–protein interactions. *Biotechnology journal* **2015**, *10* (1), 199-209.

114. Sojitra, M.; Sarkar, S.; Maghera, J.; Rodrigues, E.; Carpenter, E. J.; Seth, S.; Ferrer Vinals, D.; Bennett, N. J.; Reddy, R.; Khalil, A.; Xue, X.; Bell, M. R.; Zheng, R. B.; Zhang, P.; Nycholat, C.; Bailey, J. J.; Ling, C.-C.; Lowary, T. L.; Paulson, J. C.; Macauley, M. S.; Derda, R., Genetically encoded multivalent liquid glycan array displayed on M13 bacteriophage. *Nat. Chem. Biol.* **2021**, *17* (7), 806-816.

115. Zhang, Y.; Gildersleeve, J. C., General procedure for the synthesis of neoglycoproteins and immobilization on epoxide-modified glass slides. *Methods in molecular biology (Clifton, N.J.)* **2012**, *808*, 155-165.

116. Fanning, L. J.; Connor, A. M.; Wu, G. E., Development of the Immunoglobulin Repertoire. *Clin. Immunol. Immunopathol.* **1996**, *79* (1), 1-14.

117. Zhang, J. H.; Chung, T. D.; Oldenburg, K. R., A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **1999**, *4* (2), 67-73.
118. Denton, K. E.; Wang, S.; Gignac, M. C.; Milosevich, N.; Hof, F.; Dykhuizen, E. C.; Krusemark, C. J., Robustness of In Vitro Selection Assays of DNA-Encoded Peptidomimetic Ligands to CBX7 and CBX8. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **2018**, *23* (5), 417-428.
119. Fukui, S.; Feizi, T.; Galustian, C.; Lawson, A. M.; Chai, W., Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate-protein interactions. *Nat. Biotechnol.* **2002**, *20* (10), 1011-1017.
120. Rillahan, C. D.; Paulson, J. C., Glycan microarrays for decoding the glycome. *Annual review of biochemistry* **2011**, *80*, 797-823.
121. Park, S.; Gildersleeve, J. C.; Blixt, O.; Shin, I., Carbohydrate microarrays. *Chem. Soc. Rev.* **2013**, *42* (10), 4310-4326.
122. Palma, A. S.; Feizi, T.; Childs, R. A.; Chai, W.; Liu, Y., The neoglycolipid (NGL)-based oligosaccharide microarray system poised to decipher the meta-glycome. *Curr. Opin. Chem. Biol.* **2014**, *18*, 87-94.
123. Song, X.; Heimbürg-Molinari, J.; Cummings, R. D.; Smith, D. F., Chemistry of natural glycan microarrays. *Curr. Opin. Chem. Biol.* **2014**, *18*, 70-77.
124. Song, X.; Heimbürg-Molinari, J.; Smith, D. F.; Cummings, R. D., Glycan microarrays of fluorescently-tagged natural glycans. *Glycoconjugate J.* **2015**, *32* (7), 465-473.

125. Blixt, O.; Head, S.; Mondala, T.; Scanlan, C.; Huflejt, M. E.; Alvarez, R.; Bryan, M. C.; Fazio, F.; Calarese, D.; Stevens, J.; Razi, N.; Stevens, D. J.; Skehel, J. J.; Van Die, I.; Burton, D. R.; Wilson, I. A.; Cummings, R.; Bovin, N.; Wong, C. H.; Paulson, J. C., Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proceedings of the National Academy of Sciences* **2004**, *101* (49), 17033-17038.
126. Song, X.; Lasanajak, Y.; Xia, B.; Heimburg-Molinaro, J.; Rhea, J. M.; Ju, H.; Zhao, C.; Molinaro, R. J.; Cummings, R. D.; Smith, D. F., Shotgun glycomics: a microarray strategy for functional glycomics. *Nat. Methods* **2011**, *8* (1), 85-90.
127. Pereira, C. L.; Geissner, A.; Anish, C.; Seeberger, P. H., Chemical Synthesis Elucidates the Immunological Importance of a Pyruvate Modification in the Capsular Polysaccharide of *Streptococcus pneumoniae* Serotype 4. *Angew. Chem. Int. Ed. Engl.* **2015**, *54* (34), 10016-9.
128. Lee, H. Y.; Chen, C. Y.; Tsai, T. I.; Li, S. T.; Lin, K. H.; Cheng, Y. Y.; Ren, C. T.; Cheng, T. J.; Wu, C. Y.; Wong, C. H., Immunogenicity study of Globo H analogues with modification at the reducing or nonreducing end of the tumor antigen. *J. Am. Chem. Soc.* **2014**, *136* (48), 16844-53.
129. Wang, L.; Cummings, R. D.; Smith, D. F.; Huflejt, M.; Campbell, C. T.; Gildersleeve, J. C.; Gerlach, J. Q.; Kilcoyne, M.; Joshi, L.; Serna, S.; Reichardt, N. C.; Parera Pera, N.; Pieters, R. J.; Eng, W.; Mahal, L. K., Cross-platform comparison of glycan microarray formats. *Glycobiology* **2014**, *24* (6), 507-17.
130. Song, X.; Ju, H.; Lasanajak, Y.; Kudelka, M. R.; Smith, D. F.; Cummings, R. D., Oxidative release of natural glycans for functional glycomics. *Nat. Methods* **2016**, *13* (6), 528-534.

131. Song, X.; Heimbürg-Molinaro, J.; Smith, D. F.; Cummings, R. D., Derivatization of free natural glycans for incorporation onto glycan arrays: derivatizing glycans on the microscale for microarray and other applications (ms# CP-10-0194). *Current protocols in chemical biology* **2011**, 3 (2), 53-63.
132. Dam, T. K.; Brewer, C. F., Lectins as pattern recognition molecules: the effects of epitope density in innate immunity. *Glycobiology* **2010**, 20 (3), 270-279.
133. Mian, I. S.; Bradwell, A. R.; Olson, A. J., Structure, function and properties of antibody binding sites. *J. Mol. Biol.* **1991**, 217 (1), 133-51.
134. Toone, E., The cluster glycoside effect. *Chem Rev* **2002**, 102, 555-578.
135. Mammen, M.; Choi, S. K.; Whitesides, G. M., Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed.* **1998**, 37 (20), 2754-2794.
136. Brewer, C. F.; Miceli, M. C.; Baum, L. G., Clusters, bundles, arrays and lattices: novel mechanisms for lectin–saccharide-mediated cellular interactions. *Curr. Opin. Struct. Biol.* **2002**, 12 (5), 616-623.
137. Lee, R. T.; Lee, Y. C., Affinity enhancement by multivalent lectin–carbohydrate interaction. *Glycoconjugate J.* **2000**, 17 (7), 543-551.
138. Mortell, K. H.; Weatherman, R. V.; Kiessling, L. L., Recognition specificity of neoglycopolymers prepared by ring-opening metathesis polymerization. *Journal of the American Chemical Society* **1996**, 118 (9), 2297-2298.
139. Gestwicki, J. E.; Cairo, C. W.; Strong, L. E.; Oetjen, K. A.; Kiessling, L. L., Influencing receptor–ligand binding mechanisms with multivalent ligand architecture. *Journal of the American Chemical Society* **2002**, 124 (50), 14922-14933.

140. Smith, E. A.; Thomas, W. D.; Kiessling, L. L.; Corn, R. M., Surface plasmon resonance imaging studies of protein-carbohydrate interactions. *Journal of the American Chemical Society* **2003**, *125* (20), 6140-6148.
141. Wolfenden, M. L.; Cloninger, M. J., Mannose/glucose-functionalized dendrimers to investigate the predictable tunability of multivalent interactions. *Journal of the American Chemical Society* **2005**, *127* (35), 12168-12169.
142. Roseman, D. S.; Baenziger, J. U., The Mannose/N-Acetylgalactosamine-4-SO₄ Receptor Displays Greater Specificity for Multivalent than Monovalent Ligands. *J. Biol. Chem.* **2001**, *276* (20), 17052-17057.
143. Barbas, C. F.; Burton, D. R.; Silverman, G. J., *Phage display: a laboratory manual*. CSHL Press: 2004.
144. Anderson, B.; Rashid, M. H.; Carter, C.; Pasternack, G.; Rajanna, C.; Revazishvili, T.; Dean, T.; Senecal, A.; Sulakvelidze, A., Enumeration of bacteriophage particles: Comparative analysis of the traditional plaque assay and real-time QPCR- and nanosight-based assays. *Bacteriophage* **2011**, *1* (2), 86-93.
145. Karttunen, J.; Shastri, N., Measurement of ligand-induced activation in single viable T cells using the lacZ reporter gene. *Proceedings of the National Academy of Sciences* **1991**, *88* (9), 3972-3976.
146. Tjhung, K. F.; Kitov, P. I.; Ng, S.; Kitova, E. N.; Deng, L.; Klassen, J. S.; Derda, R., Silent Encoding of Chemical Post-Translational Modifications in Phage-Displayed Libraries. *Journal of the American Chemical Society* **2016**, *138* (1), 32-35.
147. Needels, M. C.; Jones, D. G.; Tate, E. H.; Heinkel, G. L.; Kochersperger, L. M.; Dower, W. J.; Barrett, R. W.; Gallop, M. A., Generation and screening of an

oligonucleotide-encoded synthetic peptide library. *Proceedings of the National Academy of Sciences* **1993**, *90* (22), 10700-10704.

148. Nielsen, J.; Brenner, S.; Janda, K. D., Synthetic methods for the implementation of encoded combinatorial chemistry. *Journal of the American Chemical Society* **1993**, *115* (21), 9812-9813.

149. Melkko, S.; Scheuermann, J.; Dumelin, C. E.; Neri, D., Encoded self-assembling chemical libraries. *Nat. Biotechnol.* **2004**, *22* (5), 568-74.

150. Halpin, D. R.; Harbury, P. B., DNA display II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. *PLoS Biol.* **2004**, *2* (7), E174.

151. Gartner, Z. J.; Tse, B. N.; Grubina, R.; Doyon, J. B.; Snyder, T. M.; Liu, D. R., DNA-templated organic synthesis and selection of a library of macrocycles. *Science* **2004**, *305* (5690), 1601-5.

152. Mannocci, L.; Zhang, Y.; Scheuermann, J.; Leimbacher, M.; De Bellis, G.; Rizzi, E.; Dumelin, C.; Melkko, S.; Neri, D., High-throughput sequencing allows the identification of binding molecules isolated from DNA-encoded chemical libraries. *Proceedings of the National Academy of Sciences* **2008**, *105* (46), 17670-17675.

153. Glanville, J.; Zhai, W.; Berka, J.; Telman, D.; Huerta, G.; Mehta, G. R.; Ni, I.; Mei, L.; Sundar, P. D.; Day, G. M.; Cox, D.; Rajpal, A.; Pons, J., Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. *Proc Natl Acad Sci U S A* **2009**, *106* (48), 20216-21.

154. Ernst, A.; Gfeller, D.; Kan, Z.; Seshagiri, S.; Kim, P. M.; Bader, G. D.; Sidhu, S. S., Coevolution of PDZ domain–ligand interactions analyzed by high-throughput phage display and deep sequencing. *Molecular BioSystems* **2010**, *6* (10), 1782-1790.
155. Lee, S. M.; Lee, E. J.; Hong, H. Y.; Kwon, M. K.; Kwon, T. H.; Choi, J. Y.; Park, R. W.; Kwon, T. G.; Yoo, E. S.; Yoon, G. S.; Kim, I. S.; Ruoslahti, E.; Lee, B. H., Targeting bladder tumor cells in vivo and in the urine with a peptide identified by phage display. *Mol. Cancer Res.* **2007**, *5* (1), 11-9.
156. Glenn, T. C., Field guide to next-generation DNA sequencers. *Mol Ecol Resour* **2011**, *11* (5), 759-69.
157. Quail, M. A.; Smith, M.; Coupland, P.; Otto, T. D.; Harris, S. R.; Connor, T. R.; Bertoni, A.; Swerdlow, H. P.; Gu, Y., A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* **2012**, *13* (1), 341.
158. Matochko, W. L.; Cory Li, S.; Tang, S. K. Y.; Derda, R., Prospective identification of parasitic sequences in phage display screens. *Nucleic Acids Res.* **2014**, *42* (3), 1784-1798.
159. Matochko, W. L.; Derda, R., Next-generation sequencing of phage-displayed peptide libraries. *Methods Mol Biol* **2015**, *1248*, 249-66.
160. Ravn, U.; Didelot, G.; Venet, S.; Ng, K.-T.; Gueneau, F.; Rousseau, F.; Calloud, S.; Kosco-Vilbois, M.; Fischer, N., Deep sequencing of phage display libraries to support antibody discovery. *Methods* **2013**, *60* (1), 99-110.
161. 't Hoen, P. A. C.; Jirka, S. M. G.; ten Broeke, B. R.; Schultes, E. A.; Aguilera, B.; Pang, K. H.; Heemskerk, H.; Aartsma-Rus, A.; van Ommen, G. J.; den Dunnen, J. T.,

Phage display screening without repetitious selection rounds. *Anal. Biochem.* **2012**, *421* (2), 622-631.

162. Ngubane, N. A. C.; Gresh, L.; Ioerger, T. R.; Sacchettini, J. C.; Zhang, Y. J.; Rubin, E. J.; Pym, A.; Khati, M., High-Throughput Sequencing Enhanced Phage Display Identifies Peptides That Bind Mycobacteria. *PLoS ONE* **2013**, *8* (11), e77844.

163. Hu, D.; Hu, S.; Wan, W.; Xu, M.; Du, R.; Zhao, W.; Gao, X.; Liu, J.; Liu, H.; Hong, J., Effective Optimization of Antibody Affinity by Phage Display Integrated with High-Throughput DNA Synthesis and Sequencing Technologies. *PLOS ONE* **2015**, *10* (6), e0129125.

164. Liu, G. W.; Livesay, B. R.; Kacherovsky, N. A.; Cieslewicz, M.; Lutz, E.; Waalkes, A.; Jensen, M. C.; Salipante, S. J.; Pun, S. H., Efficient Identification of Murine M2 Macrophage Peptide Targeting Ligands by Phage Display and Next-Generation Sequencing. *Bioconjugate Chemistry* **2015**, *26* (8), 1811-1817.

165. Juds, C.; Schmidt, J.; Weller, M. G.; Lange, T.; Beck, U.; Conrad, T.; Börner, H. G., Combining Phage Display and Next-Generation Sequencing for Materials Sciences: A Case Study on Probing Polypropylene Surfaces. *Journal of the American Chemical Society* **2020**, *142* (24), 10624-10628.

166. Huang, H.; Sidhu, S. S., Studying Binding Specificities of Peptide Recognition Modules by High-Throughput Phage Display Selections. In *Network Biology: Methods and Applications*, Cagney, G.; Emili, A., Eds. Humana Press: Totowa, NJ, 2011; pp 87-97.

167. Villequey, C.; Kong, X.-D.; Heinis, C., Bypassing bacterial infection in phage display by sequencing DNA released from phage particles. *Protein Engineering, Design and Selection* **2017**, *30* (11), 761-768.

168. Ivarsson, Y.; Arnold, R.; McLaughlin, M.; Nim, S.; Joshi, R.; Ray, D.; Liu, B.; Teyra, J.; Pawson, T.; Moffat, J.; Li, S. S.-C.; Sidhu, S. S.; Kim, P. M., Large-scale interaction profiling of PDZ domains through proteomic peptide-phage display using human and viral phage peptidomes. *Proceedings of the National Academy of Sciences* **2014**, *111* (7), 2542-2547.
169. Vaisman-Mentesh, A.; Wine, Y., Monitoring Phage Biopanning by Next-Generation Sequencing. In *Phage Display: Methods and Protocols*, Hust, M.; Lim, T. S., Eds. Springer New York: New York, NY, 2018; pp 463-473.
170. Nannini, F.; Senicar, L.; Parekh, F.; Kong, K. J.; Kinna, A.; Bughda, R.; Sillibourne, J.; Hu, X.; Ma, B.; Bai, Y.; Ferrari, M.; Pule, M. A.; Onuoha, S. C., Combining phage display with SMRTbell next-generation sequencing for the rapid discovery of functional scFv fragments. *mAbs* **2021**, *13* (1), 1864084.
171. Davey, N. E.; Seo, M. H.; Yadav, V. K.; Jeon, J.; Nim, S.; Krystkowiak, I.; Blikstad, C.; Dong, D.; Markova, N.; Kim, P. M.; Ivarsson, Y., Discovery of short linear motif-mediated interactions through phage display of intrinsically disordered regions of the human proteome. *The FEBS Journal* **2017**, *284* (3), 485-498.
172. Matochko, W. L.; Chu, K.; Jin, B.; Lee, S. W.; Whitesides, G. M.; Derda, R., Deep sequencing analysis of phage libraries using Illumina platform. *Methods* **2012**, *58* (1), 47-55.
173. Lyons, E.; Sheridan, P.; Tremmel, G.; Miyano, S.; Sugano, S., Large-scale DNA Barcode Library Generation for Biomolecule Identification in High-throughput Screens. *Sci Rep* **2017**, *7* (1), 13899.

174. Gironda-Martínez, A.; Donckele, E. J.; Samain, F.; Neri, D., DNA-Encoded Chemical Libraries: A Comprehensive Review with Successful Stories and Future Challenges. *ACS Pharmacology & Translational Science* **2021**.
175. Bonaldo, M. F.; Lennon, G.; Soares, M. B., Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res.* **1996**, *6* (9), 791-806.
176. Meyer, M.; Stenzel, U.; Myles, S.; Prüfer, K.; Hofreiter, M., Targeted high-throughput sequencing of tagged nucleic acid samples. *Nucleic Acids Res.* **2007**, *35* (15), e97.
177. Parameswaran, P.; Jalili, R.; Tao, L.; Shokralla, S.; Gharizadeh, B.; Ronaghi, M.; Fire, A. Z., A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Res.* **2007**, *35* (19), e130.
178. Frank, D. N., BARCRAWL and BARTAB: software tools for the design and implementation of barcoded primers for highly multiplexed DNA sequencing. *BMC Bioinformatics* **2009**, *10*, 362.
179. Ben-Dor, A.; Karp, R.; Schwikowski, B.; Yakhini, Z., Universal DNA tag systems: a combinatorial design scheme. *J. Comput. Biol.* **2000**, *7* (3-4), 503-19.
180. Liu, W.; Wang, S.; Gao, L.; Zhang, F.; Xu, J., DNA sequence design based on template strategy. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (6), 2014-8.
181. Hamady, M.; Walker, J. J.; Harris, J. K.; Gold, N. J.; Knight, R., Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat. Methods* **2008**, *5* (3), 235-7.
182. Craig, D. W.; Pearson, J. V.; Szelinger, S.; Sekar, A.; Redman, M.; Corneveaux, J. J.; Pawlowski, T. L.; Laub, T.; Nunn, G.; Stephan, D. A.; Homer, N.; Huentelman, M.

- J., Identification of genetic variants using bar-coded multiplexed sequencing. *Nat. Methods* **2008**, 5 (10), 887-93.
183. Buschmann, T.; Bystrykh, L. V., Levenshtein error-correcting barcodes for multiplexed DNA sequencing. *BMC Bioinformatics* **2013**, 14, 272.
184. Kracht, D.; Schober, S., Insertion and deletion correcting DNA barcodes based on watermarks. *BMC Bioinformatics* **2015**, 16, 50.
185. Buschmann, T., DNABarcodes: an R package for the systematic construction of DNA sample tags. *Bioinformatics* **2017**, 33 (6), 920-922.
186. Wang, B.; Zheng, X.; Zhou, S.; Zhou, C.; Wei, X.; Zhang, Q.; Wei, Z., Constructing DNA Barcode Sets Based on Particle Swarm Optimization. *IEEE/ACM Trans Comput Biol Bioinform* **2018**, 15 (3), 999-1002.
187. Barcoded Illumina primers
(<https://icom.illumina.com/download/summary/ATZRuMiBPkukcRQOJ792Xg>)
188. Hamming, R. W., Error detecting and error correcting codes. *The Bell system technical journal* **1950**, 29 (2), 147-160.
189. Levenshtein, V. I. In *Binary codes capable of correcting deletions, insertions, and reversals*, Soviet physics doklady, Soviet Union: 1966; pp 707-710.
190. Bystrykh, L. V., Generalized DNA barcode design based on Hamming codes. *PLoS One* **2012**, 7 (5), e36852.
191. Ekanayake, A. I.; Sobze, L.; Kelich, P.; Youk, J.; Bennett, N. J.; Mukherjee, R.; Bhardwaj, A.; Wuest, F.; Vukovic, L.; Derda, R., Genetically Encoded Fragment-Based Discovery from Phage-Displayed Macrocyclic Libraries with Genetically Encoded

Unnatural Pharmacophores. *Journal of the American Chemical Society* **2021**, *143* (14), 5497-5507.

192. Wong, J. Y. K.; Mukherjee, R.; Miao, J.; Bilyk, O.; Triana, V.; Miskolzie, M.; Henninot, A.; Dwyer, J. J.; Kharchenko, S.; Iampolska, A.; Volochnyuk, D. M.; Lin, Y.-S.; Postovit, L.-M.; Derda, R., Genetically-encoded discovery of proteolytically stable bicyclic inhibitors for morphogen NODAL. *Chemical Science* **2021**, *12* (28), 9694-9703.

193. Lima, G. M.; Atrazhev, A.; Sarkar, S.; Sojitra, M.; Reddy, R.; Macauley, M. S.; Monteiro, G.; Derda, R., DNA-Encoded Multivalent Display of Protein Tetramers on Phage: Synthesis and *In Vivo* Applications. *bioRxiv* **2021**, 2021.02.20.432100.

194. Vinals, D. F.; Kitov, P. I.; Tu, Z.; Zou, C.; Cairo, C. W.; Lin, H. C.-H.; Derda, R., Selection of galectin-3 ligands derived from genetically encoded glycopeptide libraries. *Peptide Science* **2019**, *111* (1), e24097.

195. Ng, S.; Bennett, N. J.; Schulze, J.; Gao, N.; Rademacher, C.; Derda, R., Genetically-encoded fragment-based discovery of glycopeptide ligands for DC-SIGN. *Biorg. Med. Chem.* **2018**, *26* (19), 5368-5377.

196. Chou, Y.; Kitova, E. N.; Joe, M.; Brunton, R.; Lowary, T. L.; Klassen, J. S.; Derda, R., Genetically-encoded fragment-based discovery (GE-FBD) of glycopeptide ligands with differential selectivity for antibodies related to mycobacterial infections. *Organic & Biomolecular Chemistry* **2018**, *16* (2), 223-227.

197. Triana, V.; Derda, R., Tandem Wittig/Diels–Alder diversification of genetically encoded peptide libraries. *Organic & Biomolecular Chemistry* **2017**, *15* (37), 7869-7877.

198. Sequencing Technology| Sequencing by Synthesis
<https://www.illumina.com/science/technology/next-generation-sequencing/sequencing-technology.html>
199. Paired end vs. Single-read sequencing
<https://www.illumina.com/science/technology/next-generation-sequencing/plan-experiments/paired-end-vs-single-read.html>
200. Buller, F.; Zhang, Y.; Scheuermann, J.; Schäfer, J.; Bühlmann, P.; Neri, D., Discovery of TNF inhibitors from a DNA-encoded chemical library based on diels-alder cycloaddition. *Chem. Biol.* **2009**, *16* (10), 1075-1086.
201. MacConnell, A. B.; Paegel, B. M., Poisson statistics of combinatorial library sampling predict false discovery rates of screening. *ACS combinatorial science* **2017**, *19* (8), 524-532.
202. Kuai, L.; O’Keeffe, T.; Arico-Muendel, C., Randomness in DNA encoded library selection data can be modeled for more reliable enrichment calculation. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **2018**, *23* (5), 405-416.
203. Faver, J. C.; Riehle, K.; Lancia Jr, D. R.; Milbank, J. B.; Kollmann, C. S.; Simmons, N.; Yu, Z.; Matzuk, M. M., Quantitative comparison of enrichment from DNA-encoded chemical library selections. *ACS combinatorial science* **2019**, *21* (2), 75-82.
204. Komar, P.; Kalinic, M., Denoising DNA encoded library screens with sparse learning. *ACS Combinatorial Science* **2020**, *22* (8), 410-421.
205. Amigo, J.; Rama-Garda, R.; Bello, X.; Sobrino, B.; de Blas, J.; Martín-Ortega, M.; Jessop, T. C.; Carracedo, Á.; Loza, M. I. G.; Domínguez, E., tagFinder: a novel tag analysis methodology that enables detection of molecules from DNA-encoded chemical

- libraries. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **2018**, *23* (5), 397-404.
206. Heckel, R.; Mikutis, G.; Grass, R. N., A Characterization of the DNA Data Storage Channel. *Scientific Reports* **2019**, *9* (1), 9663.
207. Ruijter, J.; Ramakers, C.; Hoogaars, W.; Karlen, Y.; Bakker, O.; Van den Hoff, M.; Moorman, A., Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* **2009**, *37* (6), e45-e45.
208. Pan, W.; Byrne-Steele, M.; Wang, C.; Lu, S.; Clemmons, S.; Zahorchak, R. J.; Han, J., DNA polymerase preference determines PCR priming efficiency. *BMC Biotechnol.* **2014**, *14* (1), 1-17.
209. Warnecke, P. M.; Stirzaker, C.; Melki, J. R.; Millar, D. S.; Paul, C. L.; Clark, S. J., Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res.* **1997**, *25* (21), 4422-4426.
210. Chen, Y.-J.; Takahashi, C. N.; Organick, L.; Bee, C.; Ang, S. D.; Weiss, P.; Peck, B.; Seelig, G.; Ceze, L.; Strauss, K., Quantifying molecular bias in DNA data storage. *Nature Communications* **2020**, *11* (1), 3264.
211. de Jager, W.; Rijkers, G. T., Solid-phase and bead-based cytokine immunoassay: a comparison. *Methods* **2006**, *38* (4), 294-303.
212. Galanina, O. E.; Mecklenburg, M.; Nifantiev, N. E.; Pazynina, G. V.; Bovin, N. V., GlycoChip: multiarray for the study of carbohydrate-binding proteins. *Lab Chip* **2003**, *3* (4), 260-5.

213. Willats, W. G.; Rasmussen, S. E.; Kristensen, T.; Mikkelsen, J. D.; Knox, J. P., Sugar-coated microarrays: a novel slide surface for the high-throughput analysis of glycans. *Proteomics* **2002**, *2* (12), 1666-71.
214. Wang, C. C.; Huang, Y. L.; Ren, C. T.; Lin, C. W.; Hung, J. T.; Yu, J. C.; Yu, A. L.; Wu, C. Y.; Wong, C. H., Glycan microarray of Globo H and related structures for quantitative analysis of breast cancer. *Proc Natl Acad Sci U S A* **2008**, *105* (33), 11661-6.
215. Wang, L.; Cummings, R. D.; Smith, D. F.; Huflejt, M.; Campbell, C. T.; Gildersleeve, J. C.; Gerlach, J. Q.; Kilcoyne, M.; Joshi, L.; Serna, S.; Reichardt, N.-C.; Parera Pera, N.; Pieters, R. J.; Eng, W.; Mahal, L. K., Cross-platform comparison of glycan microarray formats. *Glycobiology* **2014**, *24* (6), 507-517.
216. Zhang, Y.; Gildersleeve, J. C., General procedure for the synthesis of neoglycoproteins and immobilization on epoxide-modified glass slides. *Methods Mol Biol* **2012**, *808*, 155-65.
217. Muthana, S. M.; Gildersleeve, J. C., Factors Affecting Anti-Glycan IgG and IgM Repertoires in Human Serum. *Scientific Reports* **2016**, *6* (1), 19509.
218. Meloncelli, P. J.; Lowary, T. L., Synthesis of ABO histo-blood group type I and II antigens. *Carbohydr. Res.* **2010**, *345* (16), 2305-2322.
219. Meloncelli, P. J.; Lowary, T. L., Synthesis of ABO histo-blood group type V and VI antigens. *Aust. J. Chem.* **2009**, *62* (6), 558-574.
220. Meloncelli, P. J.; West, L. J.; Lowary, T. L., Synthesis and NMR studies on the ABO histo-blood group antigens: Synthesis of type III and IV structures and NMR characterization of type I–VI antigens. *Carbohydr. Res.* **2011**, *346* (12), 1406-1426.

221. Jeyakanthan, M.; Tao, K.; Zou, L.; Meloncelli, P.; Lowary, T.; Suzuki, K.; Boland, D.; Larsen, I.; Burch, M.; Shaw, N., Chemical basis for qualitative and quantitative differences between ABO blood groups and subgroups: implications for organ transplantation. *Am. J. Transplantation* **2015**, *15* (10), 2602-2615.
222. Bentall, A.; Jeyakanthan, M.; Braitch, M.; Cairo, C. W.; Lowary, T. L.; Maier, S.; Halpin, A.; Motyka, B.; Zou, L.; West, L. J.; Ball, S., Characterization of ABH-subtype donor-specific antibodies in ABO-A-incompatible kidney transplantation. *Am. J. Transplantation n/a* (n/a).
223. Campbell, C. T.; Zhang, Y.; Gildersleeve, J. C., Construction and use of glycan microarrays. *Curr Protoc Chem Biol* **2010**, *2* (1), 37-53.
224. Xia, L.; Gildersleeve, J. C., The Glycan Array Platform as a Tool to Identify Carbohydrate Antigens. *Methods Mol Biol* **2015**, *1331*, 27-40.
225. Butler, D. L.; Gildersleeve, J. C., Abnormal antibodies to self-carbohydrates in SARS-CoV-2 infected patients. *bioRxiv* **2020**.
226. Luminex xMAP Cookbook https://cdn2.hubspot.net/hub/128032/file-213083097-pdf/Luminex-xMap_Cookbook.pdf
227. <https://www.bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html>