UNIVERSITY OF CALIFORNIA SAN DIEGO

Systems methods for the discovery of coordinated host proteome and glycoproteome response in infectious diseases using mass spectrometry

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

in

Bioinformatics and Systems Biology

by

James Terrance Sorrentino

Committee in charge:

Professor Nathan E. Lewis, Chair
Professor Jeffrey D. Esko, Co-Chair
Professor Steven P. Briggs
Professor Victor Nizet
Professor Elizabeth A. Winzeler

2023

The Dissertation of James Terrance Sorrentino is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

I dedicate this thesis to my family for always being by my side.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge the many people who have supported my growth as a scientist and helped me achieve this goal.

I would like to thank my advisor, Nathan E. Lewis for his continued encouragement of scientific exploration and learning. Prof. Lewis has provided a powerful mentorship that has helped me tackle problems I would not have been able to overcome alone. I am incredibly grateful to have been a student under Prof. Lewis and I will always look forward to discussing new ideas and new developments.

I would like to thank my dissertation committee for all their support and guidance. My Co-chair Jeffrey D. Esko for the immense amount of time and advice he has given me throughout this process. Prof. Esko has been a wonderful collaborator who has provided wisdom and knowledge that has been crucial to this dissertation. I would like to thank Prof. Elizabeth Winzeler for letting me join her lab for a few crucial months in the beginning of my career at UCSD. The study I had conducted during my rotation with Prof. Winzeler was one of my first exposures to systems biology. I would like to thank Prof. Victor Nizet for his expert insight into human bacterial pathogens. This domain specific knowledge was valuable in the development of this project. I would also like to thank Prof. Steven P. Briggs for providing thoughtful questions and comments about my research. This has put an emphasis on preparedness that is an important quality of a great researcher.

I would like to thank my mentor Alejandro Gomez Toledo for helping me grow as an independent scientist and building my confidence as a researcher. Alejandro has been an incredible research partner and guide. He has continually offered support and was always available when I wanted to discuss new ideas or ask for help. Without him the research conducted here would not

be possible. He has set an example for how great collaborations should happen and how to maintain them. Alejandro is an exemplary instructor and I hope I can teach others the way he has taught me. I will cherish our many hours we have shared discussing science and life throughout the years we have worked together. I am honored to have shared this research with you and thank you for everything.

I would like to thank all the members of the Lewis and Esko labs that have supported me throughout my time at UCSD. I would like to thank Ben Kellman for being one of my first bioinformatics mentors and great friend. And to Ben, thank you for introducing me to Alejandro, a connection that has had an everlasting positive impact on my life. I would also like to thank Austin W.T. Chiang for helping shape my scientific process and helping me integrate into the Lewis lab. Thank you to the many other members of both labs for making our work collaborative, rewarding, and fun.

I would like to thank my undergraduate mentors from the Holman lab at UCSC. Thank you, Prof. Holman, for giving me the opportunity to learn about research early into my undergraduate degree. Thank you to Steven Perry for teaching me how to properly use a pipette and introducing me to mass spectrometry.

I would like to thank my parents and family for their love and support. Specifically, I would like to thank my father for instilling values in me that have helped me be a graceful collaborator and team player. And I would like to thank my mother for showing me the humanity of the world and pushing me to pursue work that can hopefully reach and help many people. I would like to thank them for having the patience to sit with me and hear about my research while it lies outside of their interests. My family has continued to believe in me and has helped me push forward with my progress.

I would like to thank my wife and life partner, Heather Sorrentino, for her continued support throughout our lives together. Heather has shown me unconditional love and empathy during both strenuous and enjoyable times. She has taught me how to stay grounded and pushes me to achieve my goals whenever my resilience falters. I look up to her every day. I truly would not have been able to complete this research without her.

Finally, I would like to thank all my friends in Los Angeles, Santa Cruz, San Diego, and more. You have all been like family to me and have helped me enjoy the times I have taken away from work. I will forever be grateful for our camping trips, beach days, ski trips, and celebrations.

Chapter 1, in full, is an amalgamation of the published material as it appears in "Proteomic atlas of organ vasculopathies triggered by Staphylococcus aureus sepsis" in *Nature Communications* 2019 and "Vascular Proteome Responses Precede Organ Dysfunction in a Murine Model of Staphylococcus aureus Bacteremia" in *Msystems* 2022. The dissertation author was the primary investigator and/or author on these papers. I would like to thank each co-author that worked to bring these manuscripts to fruition, Alejandro Gomez Toledo, Gregory Golden, Alex Rosa Campos, Hector Cuello, Claire Morris, Chelsea Painter, Victor Nizet, Nissi Varki, Jeffrey Smith, Christofer Karlsson, Johan Malmström, Nathan E. Lewis, and Jeffrey Esko.

Chapter 2, in part, is taken from the manuscript in preparation for submission entitled "Pathogen-driven degradation of therapeutic antibodies in murine models of Group A streptococcal infection". Additionally, chapter 2, in part, is taken from the manuscript in preparation for submission entitled "Quantifying glycopeptides and assessing the heterogeneity of glycosylation in circulating glycoproteins". The dissertation author was the primary investigator and/or author on these papers. I would like to thank each co-author that worked to bring this manuscript to fruition, Alejandro Gomez Toledo, Eleni Bratanis, Erika Velásquez, Sounak

Chowdhury, Christofer Karlsson, Nathan Lewis, Jeffrey D Esko, Mattias Collin, Oonagh Shannon, and Johan Malmström.

Chapter 3, in part, is taken from the manuscript in preparation for submission entitled "Quantifying glycopeptides and assessing the heterogeneity of glycosylation in circulating glycoproteins". The dissertation author was a primary investigator and author of this paper. I would like to thank each co-author that worked to bring this manuscript to fruition, Alejandro Gomez Toledo, Nathan Lewis, Jeffrey D Esko, and Johan Malmström.

VITA

2017 University of California Santa Cruz
Bachelor of Science, Bioengineering - Molecular Concentration
Minor, Bioinformatics

2023 University of California San Diego
Doctor of Philosophy, Bioinformatics and Systems Biology

PUBLICATIONS

Sorrentino, J. T., Golden, G. J., Morris, C., Painter, C. D., Nizet, V., Campos, A. R., Smith, J. W., Karlsson, C., Malmström, J., Lewis, N. E., Esko, J. D., & Gómez Toledo, A. (2022). "Vascular Proteome Responses Precede Organ Dysfunction in a Murine Model of Staphylococcus aureus Bacteremia." *MSystems*, 7(4).

Sorrentino, J. T. *, Gómez Toledo, A.*, Esko, J. D., Malmström, J., & Lewis, N. E. "Quantifying glycopeptides and assessing the heterogeneity of glycosylation in circulating glycoproteins" *in preparation for submission*

Gómez Toledo, A.*, Bratanis, E.*, Velásquez, E., Chowdhury, S., **Sorrentino, J. T.**, Karlsson, C., Lewis, N. E., Esko, J. D., Collin, M., Shannon, O., & Malmström, J. "Pathogen-driven IgG degradation in murine models of Group A streptococcal infection" *in preparation for submission*

Zhang Y., Krishnan S., Bao B., Chiang A. W. T., **Sorrentino J. T.**, Schinn S. M., Kellman B. P., & Lewis N. E. (2023) "Preparing glycomics data for robust statistical analysis with GlyCompareCT" *STAR Protocols, accepted*. bioRxiv doi: 10.1101/2022.05.31.494178

Martino, C.*, Kellman, B.P.*, Sandoval, D.R.*, Clausen, T.M., Marotz, C., Song, S.J., Wandro, S., Zaramela, L., Benítez, R.A.S., Zhu, Q., Armingol, E., Vázquez-Baeza, Y., McDonald, D., **Sorrentino, J.**, Taylor, B., Belda-Ferre, P., Liang, C., Zhang, Y., Schifanella, L., Klatt, N.R., Havulinna, A.S., Jousilahti, P., Huang, S., Haiminen, N., Parida, L., Kim, H.C., Swafford, A.D., Zengler, K., Cheng, S., Inouye, M., Niiranen, T., Jain, M., Salomaa, V., Esko, J.D., Lewis, N.E., Knight, R. (2020) "Bacterial modification of the host glycosaminoglycan heparan sulfate modulates SARS-CoV-2 infectivity." *bioRxiv*. DOI: 10.1101/2020.08.17.238444.

Yamaguchi, A., Hoorebeke, C., Tourdot, B. E., Perry, S. C., Lee, G., Rhoads, N., Rickenberg, A., Green, A. R., **Sorrentino, J.**, Yeung, J., Li, C., Freedman, J. C., Holman, T. R., & Holinstat, M. (2023) "Fatty acids negatively regulate platelet function through formation of noncanonical 15-lipoxygenase-derived eicosanoids" *Pharmacology Research & Perspectives*, 11(1), e01056. https://doi.org/10.1002/PRP2.1056

Perry, S. C., van Hoorebeke, C., **Sorrentino, J.**, Bautista, L., Akinkugbe, O., Conrad, W. S., Rutz, N., & Holman, T. R. (2022) "Structural basis for altered positional specificity of 15-lipoxygenase-

1 with 5S-HETE and 7S-HDHA and the implications for the biosynthesis of resolvin E4" *Archives of Biochemistry and Biophysics*, 727, 109317. https://doi.org/10.1016/J.ABB.2022.109317

Chiang, A. W. T., Baghdassarian, H. M., Kellman, B. P., Bao, B., **Sorrentino, J. T.**, Liang, C., Kuo, C. C., Masson, H. O., & Lewis, N. E. (2021). "Systems glycobiology for discovering drug targets, biomarkers, and rational designs for glyco-immunotherapy." *Journal of Biomedical Science* 2021 28:1, 28(1), 1–15. https://doi.org/10.1186/S12929-021-00746-2

Golden, G. J., Toledo, A. G., Marki, A., **Sorrentino, J. T.**, Morris, C., Riley, R. J., Spliid, C., Chen, Q., Cornax, I., Lewis, N. E., Varki, N., Le, D., Malmström, J., Karlsson, C., Ley, K., Nizet, V., & Esko, J. D. (2021). "Endothelial Heparan Sulfate Mediates Hepatic Neutrophil Trafficking and Injury during Staphylococcus aureus Sepsis." *MBio*, 12(5).

Bao, B., Kellman, B. P., Chiang, A. W. T., Zhang, Y., **Sorrentino, J. T.**, York, A. K., Mohammad, M. A., Haymond, M. W., Bode, L., & Lewis, N. E. (2021). "Correcting for sparsity and interdependence in glycomics by accounting for glycan biosynthesis." *Nature Communications* 2021 12:1, 12(1), 1–14. https://doi.org/10.1038/s41467-021-25183-5

Gómez Toledo, A., **Sorrentino, J. T.**, Sandoval, D. R., Malmström, J., Lewis, N. E., & Esko, J. D. (2021). "A Systems View of the Heparan Sulfate Interactome." *The Journal of Histochemistry and Cytochemistry : Official Journal of the Histochemistry Society*, 69(2), 105–119. https://doi.org/10.1369/0022155420988661

Liang, C., Chiang, A. W. T., Hansen, A. H., Arnsdorf, J., Schoffelen, S., **Sorrentino, J. T.**, Kellman, B. P., Bao, B., Voldborg, B. G., & Lewis, N. E. (2020). "A Markov model of glycosylation elucidates isozyme specificity and glycosyltransferase interactions for glycoengineering." *Current Research in Biotechnology,* 2, 22–36. https://doi.org/10.1016/J.CRBIOT.2020.01.001

Kellman, B. P., Zhang, Y., Logomasini, E., Meinhardt, E., Godinez-Macias, K. P., Chiang, A. W. T., **Sorrentino, J. T.**, Liang, C., Bao, B., Zhou, Y., Akase, S., Sogabe, I., Kouka, T., Winzeler, E. A., Wilson, I. B. H., Campbell, M. P., Neelamegham, S., Krambeck, F. J., Aoki-Kinoshita, K. F., & Lewis, N. E. (2020). "A consensus-based and readable extension of Linear Code for Reaction Rules (LiCoRR)." *Beilstein Journal of Organic Chemistry* 16:215, 16(1), 2645–2662. https://doi.org/10.3762/BJOC.16.215

Toledo, A. G., Golden, G., Campos, A. R., Cuello, H., **Sorrentino, J.**, Lewis, N., Varki, N., Nizet, V., Smith, J. W., & Esko, J. D. (2019). "Proteomic atlas of organ vasculopathies triggered by Staphylococcus aureus sepsis." *Nature Communications* 2019 10:1, 10(1), 1–13. https://doi.org/10.1038/s41467-019-12672-x

Fong SH, Carlin DE, Ozturk K; **2018 UCSD Network Biology Class**, Ideker T. (2019) "Strategies for Network GWAS Evaluated Using Classroom Crowd Science." *Cell Systems* 8(4):275-280. doi: 10.1016/j.cels.2019.03.013.

*Co-corresponding authors

PATENTS

LEWIS NATHAN (US); CHIANG WAN-TIEN (US); LIANG CHENGUANG (US); **SORRENTINO JAMES T** (US) 2022 METHOD OF MEASURING COMPLEX CARBOHYDRATES UNIV CALIFORNIA (US) WO/2022/026944

ABSTRACT OF THE DISSERTATION

Systems methods for the discovery of coordinated host proteome and glycoproteome response in infectious diseases using mass spectrometry

by

James Terrance Sorrentino

Doctor of Philosophy in Bioinformatics and Systems Biology

University of California San Diego, 2023

Professor Nathan E. Lewis, Chair
Professor Jeffrey D. Esko, Co-Chair

Organisms exhibit a vast catalog of mechanisms to target, capture, and disable foreign agents that have infiltrated the body. One major driver of immunogenic response to infection is the production, interaction, and regulation of proteins at the blood-tissue interface. These proteins have several key features that enable them to combat harmful microorganisms. Constituent factors of host proteins such as composition, 3D structure, post-translational modifications, and interaction partners allow for both innate nonspecific response and targeted adaptive response in the fight against pathogens. Mass spectrometry, the analytical technique to measure the mass-to-charge ratio of ions, has long been used for the identification, characterization, and quantification of proteins *in vivo* and *ex vivo*. Recent advances have taken advantage of this straightforward experimental design to measure complex post-translational modifications of proteins, including but not limited to glycosylation. Here I have chosen to focus on a major population of the blood-tissue interface during infection, the glycoproteins and their interaction partners.

In the first chapter, I describe our large-scale mass spectrometry based proteomic studies of the blood-tissue interface in murine models of infection. In these studies, I have characterized the largest-to-date time-resolved proteomic atlas of the blood-tissue interface during methicillin-resistant *Staphylococcus aureus* bacteremia. I introduce the idea that major blood and vascular remodeling during infection is organ-specific, highly compartmentalized, synchronously coordinated, and significantly correlated with the progression of the disease. Additionally, this chapter highlights the importance of the blood-tissue interface during infectious disease and displays many proteomic and glycoproteomic changes which represent vascular degradation, subsequent organ failure, and death.

In the second chapter, I describe mechanisms by which the blood-tissue interface is impacted by pathogens exploiting proteolytic and glycan degradation of immunoglobulins. Using

in vivo murine models of group A streptococcal infection, we showcase that human immunoglobulins and a variant of mouse immunoglobulin is specifically degraded after bacterial dissemination to organs. These results stimulate concern for the evaluation of antibody-based therapies in current animal models of streptococcal infections. Finally, I introduce the relationship of glycosylation and antibody function during immunogenic response to infection and highlight the importance of immunoglobulin G which will be a major topic of the third chapter.

In the third chapter, I introduce our novel pipeline for identification and quantification of glycopeptides using MS1 & MS2 levels of mass spectrometry data. We have chosen to expand our quantification methods of the glycoproteome to gain greater insight of the heterogeneity of the blood-tissue interface during infection. Here we have leveraged several advancements in the analysis of mass spectrometry data to expand the scope of glycopeptide identification and strengthen the quantification of these biological units. I will describe how I have built a computational pipeline which links the discovery of new glycopeptides from large traditional mass spectrometry datasets to identification and quantification of these glycopeptides in data independent acquisition (DIA) mass spectrometry experiments. Additionally, I analyze the heterogeneity found in the glycosylation of circulating proteins in the blood-tissue interface in models of infection and cell systems used for recombinant antibody production.

INTRODUCTION

Since the advent of microbiology, advances in the understanding of microorganisms and their role in illness have caused annual mortality from infectious disease to plummet. By the work of Koch, Pasteur, and many others since the nineteenth century, the prevailing theory is that small, microscopic organisms known as pathogens can invade living hosts and lead to disease. Advances in public health measures to combat bodily invasion, such as vaccination, water chlorination, dairy pasteurization, and other antimicrobial measures and treatments have dramatically decreased our chances of contracting fatal bacterial infections (Doron & Gorbach, 2008, Opal, 2010). However, it is believed that the morbidity associated with bacterial infection is significantly contributed to by the infected host's immune response (De Backer et al., 2014, Delano et al., 2016, Hotchkiss et al., 2016, Levi & Van Der Poll, 2017, Van Der Poll et al., 2017). This dysregulated host response, sepsis, can lead to life-threatening organ dysfunction and multiple organ failure (Bone et al., 1989 & 1992, Lelubre & Vincent, 2018, Singer et al., 2016, Vincent et al., 1996). It is estimated that globally 48.9 million cases of sepsis and 11.0 million sepsis-related deaths were recorded in 2017, highlighting the lack of recent advances in tracking prognostic biomarkers and care for those who contract a bacterial infection and become critically ill. Sepsis, which is caused by underlying infection, is still considered as a major cause of health loss globally (Kaukonen et al., 2014, Liu et al., 2014, Rhee et al., 2017, Rudd et al., 2020, Weng et al., 2018). Additionally, the increase in antibiotic resistant bacteria or "superbugs" is a cause to search for new antimicrobial interventions and create more effective patient support measures (Stryjewski et al., 2014, Urban-Chmiel et al., 2022).

Over the last 50 years the innate immune response has been increasingly described as a complex system of multiple interacting parts and feedback loops (Beutler & Rietschel, 2003, Marshall, 2014, Opal, 2010). Early studies of murine endotoxemia, a model of released bacterial products in the mouse bloodstream (i.e. lipopolysaccharide, LPS), implicated hundreds of cytokines, membrane proteins, intracellular proteins, transcription factors, coagulation factors, and other proteins, steroids, and signaling molecules based on each molecules ability to alter lethality when inhibited or administrated (Beutler & Rietschel, 2003, Ince et al., 2016, Lewis et al., 2016, Poli-de-Figueiredo et al., 2008). Following initial detection of a foreign microbe, coagulatory and inflammatory cascades are processes that evolved to protect the host; however, the details of these lifesaving functions are muddied by mechanisms tailored to the composition of the pathogen. Additionally, dysregulated immune responses can often turn into a life-threatening response. Unfortunately, sepsis treatments aimed at inhibition of systemic inflammation have little clinical success. Sadly, the conclusion of hundreds of failed clinical trials in sepsis have determined that host response varies by the infecting microorganism and patient health status, thus it would and has proven useless to target any single host-derived mediator molecule (Marshall 2008 & 2014). Interestingly, there is a lack of dedicated databases or repositories to store data derived from preclinical sepsis research. This is crucially important because anyone is at risk of getting sepsis introducing immense heterogeneity in sepsis research and study results. Future research continues to address the need for molecular subtyping of sepsis and more effectively translate basic research into clinically useful information.

As advertised above, deep coverage omics research of models of monomicrobial sepsis are advantageous for assessing the progression of experimental sepsis due to their robustness and reproducibility. In such a controlled environment, bug-specific molecular and temporal

relationships can be tracked and well defined. *Staphylococcus aureus* is one of the most common commensal bacterium and human pathogens in the world. Astonishingly, around 30% of the human population is colonized with *S. aureus* with no effect. However, *S. aureus* remains the leading cause of gram-positive bacteremia (Kim et al., 2014, Tong et al., 2015). Worse is the rise of antibiotic resistance in variants of *S. aureus*, such as, MRSA (methicillin-resistant *S. aureus*) (Stryjewski et al., 2014). In the United States, *S. aureus* is the largest cause of hospital-acquired infectious disease mortality, attributed to its resistance of common antibiotics (Tong et al., 2015). *S. aureus* sepsis leads to pathogen accumulation primarily in the liver where major coagulation and inflammation can cause tremendous endothelial dysfunction and organ damage (Kolaczkowska et al., 2015). *S. aureus* sepsis does not only affect the liver, but it may also damage other organs and have a profound effect on the circulatory system of the organisms. Elucidating the temporal dynamics of vascular remodeling and tissue damage in MRSA sepsis is an important question addressed in our studies.

Similarly, *Streptococcus pyogenes* (GAS), the human pathogen responsible for strep throat, has an interesting range of diverse clinical manifestations and several characterized virulence factors that modulate both innate and adaptive immune responses. Previous research suggests that *S. pyogenes* infections give rise to this immense heterogeneity due to the specific repertoire of virulence factors being expressed by the bacteria while in proximity to the specific state of the host proteome. Remarkably, *S. pyogenes* is able to evade the immune system and secure nutrients from the host using a variety of surface proteins which inhibit phagocytosis and promote adhesion to host extracellular matrix proteins (Rohde & Cleary, 2016, Wollein Waldetoft & Raberg 2014). Secreted virulence factors can prevent the normal function of immunoglobulin G (IgG) through nonimmune Fc-mediated binding of IgG antibodies and severely destabilize the circulating

immune environment (Happonen et al., 2019, Mitchell, 2003). Recently *in vivo* activity of the *S. pyogenes* virulence factor repertoire has been explored using novel targeted quantitative mass spectrometry methods to focus on IgG cleavage at different infection sites. The results demonstrate that the combined *in vivo* effect of IgG cleavage by IdeS is dependent on the microenvironments associated with different disease and infection types (Karlsson et al., 2018, Wenig et al., 2004). However, our endeavors using the same murine model have led us to describe the effects of pathogen-driven glycan degradation of IgGs both endogenous and exogenously administered.

All preclinical models of disease have advantages and disadvantages. Sepsis is a highly heterogeneous condition which makes translating lab results into clinical initiatives extremely difficult (Seok et al., 2013). However, simplified models such as, the LPS model (endotoxin driven sepsis), the CLP model (polymicrobial sepsis through cecal ligation and puncture), and models of monomicrobial sepsis using peritoneal or intravenous inoculation have been important in expanding our knowledge of the molecular basis of sepsis (Kim et al., 2014, Weidner et al., 2016). Traditionally, sepsis studies using preclinical murine models have solely focused on circulating host-derived inflammatory molecules, such as TNF or IL-1 (Ince et al., 2016), ignoring the synergistic properties of molecular crosstalk at the blood-tissue interface. Our study was not to compare the pathology of multiple sepsis models nor to identify single prognostic markers without contextual function. We conducted our proteomics and glycoproteomics assessments of MRSA-driven sepsis and GAS immune system evasion in order to establish molecular and temporal relationships between vascular dysfunction and organ dysfunction. Although the injection of a large bolus of bacteria to induce sepsis in mice might differ in many aspects from the relatively slower development of human sepsis, vascular dysfunction and glycocalyx remodeling in the mouse have many similarities to what occurs in humans.

A key factor in the progression from sepsis to multiple organ failure is the degradation of the vascular endothelium. Endothelial cells make up the lining of the blood vessels and act as the primary interface between circulating fluids and the parenchyma of organs. This vascular barrier is responsible for regulating the homeostasis of the circulatory system through vasomotor control, solute transport and osmosis (Ince et al., 2016). Additionally, the endothelial surface layer is decorated with a variety of proteins, lipids, glycans, and other biomolecules that participate in immunological functions. For example, under normal conditions, endothelial cells present a negatively charged surface densely coated in glycoproteins, proteoglycans, and glycosaminoglycans (GAGs) known as the vascular glycocalyx (VG). Upon inflammatory stimuli, which can be caused by pathogen exposure, tremendous shedding and remodeling of the VG may occur to expose adhesion molecules and release cytokines. The presentation of membrane bound cytokines may bind leukocyte chemokine receptors in order to recruit immune cells to areas of pathogen sensing. GAG chains such as heparan sulfate (HS), a GAG which can harbor many negatively charged sulfate groups, are typical components of the VG and play a role in immune response as well. Neutrophil-derived myeloperoxidase (MPO) is an enzyme that can interact with the negatively charged HS and reduce the charge repulsion between leukocytes and endothelial cells and thus promote leukocyte adhesion. Another abundant endothelial surface GAG, hyaluronan, has been hypothesized to have a variety of functions at both homeostasis and during inflammation (Marki et al., 2015). Two interesting purposes related to the work put for here are worth highlighting. Several *in vivo* and *in vitro* experiments show that serum GAGs, including hyaluronan, were able to help the endothelial surface layer recover after enzymatic treatment and that there may be a CD44-hyaluronan interaction assisting in neutrophil arrest under inflammatory conditions (Henry & Duling, 1999, Marki et al., 2015).

Under bacteremia, the blood-tissue interface exhibits major tendencies toward extensive coagulation. This blockage and thrombosis can cause serious organ damage during advanced stages of sepsis (Levi & Van Der Poll, 2017). Early studies of baboon models of Gram-negative sepsis have reported several alterations in the extrinsic pathway of coagulation marked by decreased tissue factor pathway inhibition. Unfortunately these findings proved to be disappointing as results from sepsis clinical trials of exogenous administration of Tissue Factor Pathway Inhibitor (TFPI) have not satisfied criteria for human relevance (Tang et al., 2007, Taylor et al., 2001). Coagulation is not a process that is exclusively and spontaneously happening in the open space of blood vessels. The endothelial cell surface is remodeled during inflammation and infection to present an environment ripe for platelet deposition. The interaction of platelets with the endothelium is in part carried out by the glycoprotein Von Willebrand Factor (VWF). This glycoprotein is both secreted by platelets and by the endothelium during an inflammatory response of the blood-tissue interface. Several platelet membrane glycoprotein receptors exist which may interact with endothelial VWF and anchor platelets to any surface of modified endothelium. Platelet aggregation is closely related to the formation of intravascular thrombosis. Enzymes such as thrombin can cleave another blood glycoprotein, fibrinogen, to form fibrin-based blood clots (Engelmann & Massberg, 2013, Schmugge et al., 2003). As stated, these intrusive structures can lead to tissue necrosis and organ dysfunction. There is not any standardized way at the molecular level to know what stage of blood-tissue response to sepsis other than metrics based on the general health and appearance of humans and animals. Understanding the temporal relationship of circulating vascular markers is an area ripe for research.

Circulating plasma proteins during inflammatory and infectious conditions are being extensively explored for early identification of high-risk sepsis patients. Acute phase response, the

extensive systemic reaction to infection, inflammation or tissue injury briefly described above, is followed by a number of changes that make up the adaptive immune response in sepsis. The complement proteins of the innate immune system work in conjunction with proteins found in the adaptive response, namely antibodies (Alberts et al., 2002). Immunoglobulin (Ig) production, modification, and secretion transform the blood plasma during sepsis. Igs circulate in the hopes to find and bind bacterial antigens via their variable Fab-domains and induce phagocytosis via binding of their constant domain (Fc) to Fc-receptors on phagocytic cells. These steps are important in triggering leukocyte activation, cytokine release, and antibody-dependent cellular toxicity (ADCC). Additionally, non-immune cells such as endothelial cells are also known to express Fc-receptors further complicating the interplay of the innate and adaptive immune response (Ferrara et al., 2011). Notably, the vascular endothelium can also harbor membrane Igs which could be released into the circulating plasma at the time of shedding. Many viral and bacterial pathogens express virulent FC-binding factors to circumvent the host immunity during infection. For example, studies combining targeted cross-linking mass spectrometry (TX-MS) and molecular dynamics (MD) simulation have shown that a bacterial GAS virulence factor, M1 protein, can strongly interact with human immunoglobulin G (IgG1-4). This non-immune Fc-mediated interaction was shown to prevent human IgGs from opsonization essentially masking the Fc-receptor binding site (Khakzad et al., 2021). This destabilization is theorized to protect the bacteria from phagocytic killing. Many factors can modulate the binding affinity of FC-recognizing proteins and FC-domains. In general these factors include IgG isotype (IgG1-4 in humans), the expression levels of the FC-receptors at the cell surface, and the presence of posttranslational modifications (PTMs). To elaborate on glycosylation, all IgGs are decorated with a single N-glycan chain on the CH2-domain of the FC-regions. IgG subtypes may carry different

glycosylation signatures due to their varying capacity for glycosyltransferase access. Biochemical and structural studies have demonstrated that the switch between an "open" and "closed" FC-conformations correlates with degree of glycosylation. Total removal of the glycan significantly diminishes all FC-dependent effector functions, both *in vitro* and *in vivo* (Kao et al., 2017). To date, most studies have focused on the role of glycosylation in mediating IgG1-FC dependent functions. The importance of this PTM for the structure-and-function of other IgG subtypes such as IgG2, IgG3 and IgG4 remains to be elucidated. Similarly, the role of FC-glycosylation for the interaction between antibodies and viral/bacterial FC-binding proteins warrants further investigation.

Donated human immunoglobulins from healthy individuals are known to have incredible therapeutic benefit in those combating illness. Intravenous immunoglobulins (IVIGs) infusions can treat disorders such as systemic lupus erythematosus (SLE), immunothrombocytopenia (ITP) and autoimmune hemolytic anemia (AIHA) (Almizraq et al., 2021, Majer & Hyde, 1988, Zandman-Goddard et al., 2005). In sepsis the therapeutic benefit of IVIGs remains unclear due to the complexity of both the disease and the IVIG formulation. Currently IVIG pools are derived from thousands of donors and their therapeutic effects are thought to be exerted through many FAB and FC-mediated mechanisms. Regardless, in an application of IVIGs for treating severe streptococcal toxic-shock syndrome (STSS) caused by group A streptococcus (GAS) infections, patients who received IVIGs had significantly higher 30-day survival rates and lower Acute Physiology and Chronic Health Evaluation (APACHE) II scores compared to untreated patients (Kaul et al., 1999). Additional studies have shown that specific IgG glycopatterns can be evoked through specific immunological priming which raise the possibility of novel immunotherapeutic approaches by targeting IgG glycosylation to elicit specific immune responses (Lee et al., 2016,

Mahan et al., 2016). Finally, *in vivo* modification of glycan features on IgGs can offer promising modulation of autoimmune functions unlocking new therapeutic potential (Pagan et al., 2018).

Antibodies remain as one of the most prevalent approved biopharmaceuticals (Walsh, 2018). This is because of the unsurpassed versatility of these biomolecules to treat a variety of diseases, ranging from uncomfortable to deadly. As previously stated, antibody glycosylation can impact efficacy and effectiveness which makes glycan heterogeneity a prime consideration in recombinant production cell lines. Chinese hamster ovary (CHO) cells offer a high yielding and predictable yet flexible model to produce glycoprotein therapeutics due to the baseline mammalian secretory system. Through genetic manipulation or process engineering glycan heterogeneity on recombinantly produced antibodies can be tailored to meet specific needs (Kotidis et al., 2023). This tool further enables researchers to deconvolve biological activity arising from a diverse set of glycoproteins.

Today, mass spectrometry-based proteomics and glycoproteomics have matured into a powerful tool that can be used to contextualize protein form and function in order to reveal synergistic properties. Heterogeneity of post-translational modifications, specifically glycosylation, can help explain unique observed interactions between proteins and within proteins themselves. Glycosylation has multiple levels of heterogeneity (Caval et al., 2021); intuitively macro-heterogeneity can refer to the presence or absence of a sugar moiety being attached to an eligible amino acid while micro-heterogeneity can describe the chemical makeup of that sugar. These measures are important to consider because they may impact the pharmacokinetics or pharmacodynamics of a candidate therapeutic glycoprotein, as previously stated. However, many glycoproteins have been found to be decorated with sugars of various compositions at various amino acids simultaneously. This meta-heterogeneity is now being used to describe a correlation

between the selective regulation of individual glyco-sites in a protein and that protein's structure-function relationship. Several factors introduce difficulty in measuring glycan heterogeneity. At the experimental level related to the mass spectrometer, co-elution under reverse-phase condition, lower sensitivity of modified peptides due to poorer ionization, and lack of sensitivity or detection of low abundant or unexpected glycopeptide glycoforms introduce challenges to the researcher. At the data analysis level, critical considerations and perspectives to account for the interdependence of glycan structures and glycoform populations make heterogeneity measures tough to interpret, but desperately needed.

To date, mass spectrometry-based approaches have been revolutionary in accurately identifying and quantifying biomolecules in complex mixtures. Tandem mass spectrometry paired with liquid chromatography (LC-MS/MS) is a common platform for measuring peptides and glycopeptides generated from enzymatic digestion of serum and other fluids (Mann et al., 2003, Tabang et al., 2021). LC-MS/MS gives researchers the ability to understand the differences in relative abundance of proteins and glycoproteins in different samples in a high throughput fashion. Briefly, LC-MS/MS is a bottom-up approach that usually starts with a protein extraction and digestion from a biological sample followed by separation via liquid chromatography. This separated flow of biomolecules is then ionized and separated by their mass to charge ratio, fragmented, and measured. Finally, this data is used to query relevant databases to identify the biomolecules and deconvolve the measures. There are several varieties of the LC-MS/MS methodology, each offering unique machine setups and experimental parameters. As one can imagine these methodology differences can arise from the exact model of the machinery, separation columns, style of ionization (electrospray ionization or matrix-assisted laser desorption/ionization), ionization energies, acquisition methods (DDA, DIA, MRM) and many

others (Wolf-Yadlin et al., 2016). Additionally, many upstream biochemical steps can be taken to isolate a subpopulation of biomolecules that the researcher is interested in. For our work, we employ several biochemical enrichments to interrogate specific proteomes and glycoproteomes. Due to these factors, each experimental setup impacts the considerations that must be made in the data processing and analysis. Software development in the field of proteomics and glycoproteomics is specialized to account for the interesting properties that arise in the data when measuring these types of biomolecules.

For our work we have used a variety of data acquisition approaches and parameter settings to measure peptides and glycopeptides; exact experimental setups are described in the methodology of each chapter. Most importantly we have toggled between both data dependent and data independent acquisition modes for LC-MS/MS in order to satisfy the needs of our objectives. Data dependent acquisition (DDA) encompasses a strategy that is most commonly used when performing high throughput proteomics studies. In brief, this method selects only certain peptides for fragmentation during the second stage of MS based on signal intensity (typically highest abundance). Alternatively, a sliding window approach to fragmentation can be taken known as sequential window acquisition of all theoretical mass spectra (SWATH-MS). DIA-SWATH-MS uses broad isolation windows to measure almost all MS-detectable peptides in a biological sample (Bruderer et al., 2017). Window size becomes an important parameter when discussing topics of data analysis for DIA-SWATH-MS. In order to efficiently identify and quantify either peptides or glycopeptides from DIA-SWATH-MS data, a spectral library is normally used. This library contains the individual patterns of fragmentation for a variety of moieties expected to be present in the biological sample. As the mass to charge ratio window widens for fragmentation the MS spectra that are acquired become more complexly multiplexed. High quality libraries can be

generated from extensive DDA data pertaining to the biological samples. DDA offers near peptide or glycopeptide specific fragmentation spectra enabling confident patterns to be recorded in the spectral library. However, because glycopeptides may have a variety of closely related glycans appear at a single glycosylation site, they offer a unique challenge in spectral library generation for DIA-SWATH-MS measure.

Glycan heterogeneity and structural interdependence becomes an important variable in the analysis of large glycoproteomics datasets. This variation is attributed to the possibility for a suite of glycans to appear at a single glycosylation site and that the abundance of a single glycan at this site can be overshadowed by another. The shared intermediate biosynthetic steps used to build a glycan cause the fragmentation patterns of these glycopeptides to have large structural similarities (Bao & Kellman et al., 2021). Due to this difficulty, the identification of glycans and the space of glycan heterogeneity may be underestimated with current approaches. Techniques such as molecular networking can be implemented to leverage spectral similarities between related molecules within mass spectrometry (MS) data (Bouslimani et al., 2014, Wang et al., 2016). While some software only uses select scans from an MS experiment to identify biomolecules, molecular networking approaches draw spectral correlations and use scan alignment built from all measured scans in the experiment to identify related molecules. This network of relatedness allows for algorithmic prediction of what unidentified biomolecules can be annotated as.

Increasingly networks or graphs are used to analyze large, high-dimensional data sets in the field of biology. Leveraging correlations of abundance, subcellular location, interaction partners, and biological function are useful in separating specific context and mechanisms of action being displayed at the blood-tissue interface. Because this space has many cell types and innumerable feedback loops we set out to build strategies that highlight a few distinct yet

connected mechanisms. Formulating features of our datasets into pairwise relationships and network language allowed for the (1) map of the temporal molecular changes in the blood-tissue interface during MRSA sepsis, (2) evaluation of antibody destabilization in multiple murine models of GAS infection, and (3) the quantification of glycopeptides and glycan heterogeneity using DIA-SWATH-MS.

Specifically, we have developed several computational methods when combined with mass spectrometry based proteomics and glycoproteomics that have furthered the understanding of coordinated host response at the blood-tissue interface in monomicrobial sepsis. We have focused on using systems biology strategies to create (1) a multilayer correlation network analysis for multicompartment sepsis proteomics and (2) a molecular networking workflow for the generation of hybrid spectral libraries to be used in DIA glycoproteomics and glycan heterogeneity analysis.

In chapter 1, I will discuss my methods for interfacing and interrogating large-scale mass spectrometry proteomic studies of MRSA driven sepsis. I will outline the **COR**relative **S**tratification of **P**r**O**teome **T**rajectorie**S** (COR-SPOTS) method used to examine the molecular crosstalk and temporal changes found in murine plasma, vascular cell surface, and parenchymal proteomes after infection. A molecular atlas of this MRSA-driven sepsis model is put forth and several biomolecules are proposed which indicate vascular failure precedes organ damage. This approach can be extended to other studies of monomicrobial murine sepsis and generate multiple higher-order molecular networks which can easily be analyzed together. We envision extensions of this work to aid in stratification and subtyping of sepsis based on molecular markers coming from vascular components. There is hope to incorporate many omics datasets to these networks to build multilayer molecular libraries (MMLs) of sepsis.

In chapter 2, I apply protein abundance and compositional glycan analysis methods to detail the IgG destabilization caused by the bacteria *S. pyogenes* (GAS). When combined with analysis of the blood proteome over the course of infection, the virulence factors EndoS and IdeS were found to specifically target IgGs for deglycosylation and degradation. The data indicates that only one isotope of murine IgG is degraded by this mechanism while all human isotypes are degraded in the subcutaneous infection model. Interestingly, different bacterial delivery to the murine body turns off the EndoS/IdeS IgG degradation and may offer an alternative model of IVIG administration in a GAS murine model.

In chapter 3, I have built a DIA-SWATH-MS data processing and analysis workflow utilizing several available mass spectrometry software tools including GNPS molecular networking, Skyline, and MSstats. We have looked at various glycoproteins involved in infectious disease but primarily focused on antibodies. This reproducible pipeline was used to analyze several recombinant antibodies produced in CHO cells and human and murine IgGs. Quantification and summarization of glycopeptides using MS2 measures revealed new diversity in glycan species in the FC region. Additionally, glycan feature-centric (ex. galactosylation, sialylation, fucosylation etc.) summarization and quantification based on MS2 measures displayed the utility of large window analysis in the DIA-SWATH-MS method. Glycosite heterogeneity analysis is presented as a tool for fast and scalable analysis of important site-specific glycan features that can play a role in glycoprotein function in infection.

This introduction is, in part, adapted from "Vascular Proteome Responses Precede Organ Dysfunction in a Murine Model of *Staphylococcus aureus* Bacteremia" in *mSystems*, 2022 by James T. Sorrentino, Gregory J. Golden, Claire Morris, Chelsea D. Painter, Victor Nizet, Alexandre Rosa Campos, Jeffrey W. Smith, Christofer Karlsson, Johan Malmström, Nathan E.

Lewis, Jeffrey D. Esko, and Alejandro Gómez Toledo; "Pathogen-driven degradation of therapeutic antibodies in murine models of Group A streptococcal infection" *in preparation* by Alejandro Gómez Toledo, Eleni Bratanis, Erika Velásquez, James T. Sorrentino, Christofer Karlsson, Nathan E. Lewis, Jeffrey D. Esko, Mattias Collin, Oonagh Shannon, and Johan Malmström; and "Quantifying glycopeptides and assessing the heterogeneity of glycosylation in circulating glycoproteins" *in preparation* by James T. Sorrentino, Alejandro Gómez Toledo, Jeffrey D. Esko, Johan Malmström, and Nathan E. Lewis. The dissertation author was the primary investigator and/or author on these papers.

Chapter 1: Organ-specific signatures of proteome remodeling during MRSA infection

1.1 Abstract

Dysregulated host response to microbial infection has the dramatic potential for lasting organ damage, organ failure, and death. In a host, the blood-tissue interface is the majorly affected compartment during bacteremia. Luckily, this compartment offers a measurable domain for prognosis and a targetable space for therapeutic intervention. However, time-resolved proteomic studies to chart the blood-tissue interface in monobacterial systemic infections are rarely organized due to the complexities around compartment-specific protein measures and tightly controlled infection models. To measure the vascular cell surface proteome, we have designed two mass spectrometry-based studies utilizing targeted labeling by perfusion in a murine model of methicillin-resistant *Staphylococcus aureus* (MRSA). We were able to quantify thousands of proteins across the plasma, parenchymal, and vascular compartments of multiple organs in a time-resolved fashion. We used various network-based strategies to connect and describe the blood-tissue interface atlas of proteome response in *S. aureus* bacteremia. This analysis revealed that vascular surfaces are primed in an organ-specific manner and their proteome response precedes bacterial invasion and leukocyte infiltration into organs. Astonishingly, these changes also preceded well-established cellular and biochemical correlates of systemic coagulopathy and tissue dysfunction enabling novel prognostic measures of organ distress.

1.2 Introduction

The mammalian circulatory system has developed unique structure, morphology, and composition across the various vasculature found in the body. The highly specialized cell types and tissues found in the vasculature are adapted to the anatomical and histological constraints of their encompassing organs (Aird, 2007). Additionally, the dense coating of carbohydrates, glycoproteins, and proteoglycans, known as the glycocalyx, makes up a considerable part of the structure and function of these specialized vascular cells (Reitsma et al., 2007). At the blood-tissue interface, the responsibility of the endothelial cells and their glycocalyx include maintaining vascular homeostasis and several aspects of native immunity (Zhang et al., 2018). In order to tag, purify, and measure proteins that reside on the endothelial surface layer and those that are normally exposed to the vascular flow we applied a strategy based on systemic biotinylation of murine tissues, streptavidin affinity chromatography and high-resolution liquid chromatography with tandem mass spectrometry (LC-MS/MS). This approach was simple and effective, enabling selective labeling of structures closely associated with the vasculature to be interfaced with shotgun proteomics to dissect the vascular cell surface proteome in bacteremia.

Bacteremia and subsequent sepsis induce dramatic perturbations in the plasma and organ proteomes of both humans and mice (Lapek et al., 2018, Malmstrom et al., 2016, Michalik et al., 2020, Pimienta et al. 2019, Wozniak et al., 2020). However, due to the complexity of the blood-tissue interface, the use of animal models in this type of research has been the subject of debate due to several uncorrelated response mechanisms with clinical human sepsis. Still, most of the knowledge of the molecular basis of this disease has been derived from animal models of endotoxin (LPS) driven sepsis and the polymicrobial sepsis through cecal ligation and puncture (CLP) model (Kim et al., 2014). These models have been shown to share many similarities of

vascular and glycocalyx dysregulation in humans and these models remain useful (Weidner et al., 2016). In our studies we have chosen to use a monobacterial murine infection model in order to assess the progression of experimental sepsis with an easily pinpointed time course and a reproducible bacterial inoculation. This experimental structure enables the tracking of the molecular and temporal relationship between vascular dysfunction and organ dysfunction.

During bacteremia and subsequent sepsis, shedding of glycocalyx components and remodeling of vascular surfaces have been reported to fuel dysregulated inflammatory loops (Fischetti & Tedesco, 2009, Iba & Levy, 2019). After shedding/remodeling, plasma bound components, sometimes referred to as disease associated molecular patterns (DAMPs), can attribute to activation of the coagulation and complement systems, while glycocalyx reconstitution can promote leukocyte recruitment and tissue extravasation (Finigan, 2009, Golden et al., 2021, Marki et al., 2015). Due to the known functions of the endothelial surface in multiple host defenses and for its accessibility by soluble plasma proteins, the vascular compartment is an attractive therapeutic target to prevent multiple organ failure (MOF) in sepsis (Uchimido et al., 2019).

Interestingly, it remains elusive as to why certain organs undergo failure more often than others during a systemic infection. However, observance of sepsis inducing tissue-specific protein changes in the plasma of both humans and mice stimulates the hypothesis that the vasculature is primed to react to changing environmental cues in the blood-tissue interface in an organ specific manner (Rossaint & Zarbock, 2015, Slany et al., 2016, Wiesinger et al., 2013). Tracking the spatiotemporal relationships between vascular and parenchymal tissues at the molecular level is key to understanding the early events of systemic inflammation. For example, identification of a disrupted angiopoietin (Ang)/Tie2 axis and restabilization through therapeutic intervention reduces vasculopathy and inflammation in preclinical models of sepsis (Han et al., 2016,

Leligdowicz et al., 2018). However, methodological limitations have prevented the study of vascular proteome changes in sepsis and the suitability of in vitro cellular systems has been questioned in multiple studies (Cleuren et al., 2019, Durr et al., 2004).

To quantitatively explore the proteome of the blood-tissue interface we used multiple methods of LC-MS/MS. Both studies generated label-free molecular atlas of the murine plasma, parenchymal, and vascular space in several organs during MRSA. However, the latter study consisted of the selected time course and utilized a data independent acquisition (DIA) workflow based on Sequential Window Acquisition of all Theoretical Mass Spectra (DIA-SWATH-MS) while the first study was posed at a single time point and used a traditional data dependent acquisition workflow. In our two studies we developed a perfusion method of short perfusion times in ice-cold buffers to preserve tissue integrity and deliver biotin tags specifically to perivascular proteins normally accessible to the blood flow. This method was assessed by both histology and mass spectrometry in uninfected mice as well as our murine model of MRSA bacteremia.

During characterization of the blood-tissue interface we investigated our hypothesis that the vascular proteome response precedes organ dysfunction and predisposes the organism to organ failure in MRSA bacteremia. We conducted a longitudinal proteomics analysis of the plasma, parenchymal, and vascular cell surface proteomes of four relevant organs (heart, kidney, liver, spleen). For this unprecedented amount of vascular MRSA bacteremia data, we developed a network-based computational framework to integrate the data across organs, compartments, and time points. This generated evidence for the presence of novel temporally coordinated proteome signatures which were significantly correlated with organ specific bacterial invasion and damage. Surprisingly, we identify vascular alterations that precede well-established cellular and biochemical markers of sepsis and propose the vascular cell surface proteome as a determinant of

the susceptibility of the organs to develop coagulopathy and tissue dysfunction. We are hopeful that our framework will help determine novel molecular classifications of sepsis subtypes in the context of how specific pathogens impact the organ vasculatures.

1.3 Materials and Methods

To gather a large-scale near-complete proteomic atlas of the murine plasma, organ, and vascular surface during *S. aureus* infection several reproducible procedures were performed. Mice were inoculated with a consistent bolus of an *S. aureus* strain originally isolated from an outbreak in Houston, Texas which caused severe invasive disease in adolescents. These mice were then euthanized at 6, 12, and 24 hours post infection. Animals were either immediately subjected to the chemical perfusion assay or utilized for blood collection and parenchymal tissue collection; multiple organs were collected from both groups. Histological analysis, immunofluorescence, blood chemistry, and bacterial colony forming units assays were used to verify a consistent and prominent infection and proper identification of the vascular cell surface. The aim of the chemical perfusion with a biotin tag coupled with LC-MS/MS was to selectively measure proteins at the vascular surface with either DDA or DIA methodologies. Several common statistical procedures were performed to quantify and describe differential abundance of proteins and enrichment of biological pathways. A novel correlative higher-order network analysis was performed in order to integrate the proteomic signals across the analyzed organs and compartments.

Bacterial preparation and infection. *S. aureus* strain USA300/TCH1516 was grown at 37 °C in liquid cultures of Todd-Hewitt broth (THB, Difco) with agitation (200 rpm), and later incubated in 5 mL of fresh THB overnight. Roughly, 400μL of the overnight culture was inoculated into 6 mL of fresh THB and incubated until an OD600 = 0.4 was reached. Bacteria were centrifuged, washed twice, and resuspended in PBS. 8–10-week-old C57Bl/6 mice were

intravenously infected through the retroorbital sinus with 5 × 107 CFU of the bacteria culture in 100µL PBS, or just with 100µL PBS in the control group. Animals were euthanized at 6h, 12h or 24h using isoflurane. One group of animals was immediately subjected to chemical biotinylation perfusions as described below, and another group was subjected to cardiac puncture to collect blood samples. After these procedures, multiple organs were collected from both groups (liver, kidney, heart and spleen). All animals were housed in individual ventilated cages in vivaria approved by the Association for Assessment and Accreditation of Laboratory Animal Care at the School of Medicine, UC San Diego. All experiments followed relevant guidelines and regulations consistent with standards and procedures approved by the UC San Diego Institutional Animal Care and Use Committee (protocol #S99127 and #S00227M).

Systemic chemical perfusions. Chemical perfusions were performed as previously reported (Toledo et al., 2019). Briefly, mice were subjected to a median sternotomy and the left ventricle of the heart was punctured with a 25-gauge butterfly needle (BD Vacutainer). A small cut was made in the right atrium to allow draining of perfusion solutions. All ice-cold perfusion reagents were infused using a perfusion pump (Fisher scientific). Blood was washed out with PBS for 5 min at a rate of 5 mL/min. A solution containing 100mM EZ-link Sulfo-NHS- biotin (Thermo Fisher) in PBS, pH 7.4 was perfused at a rate of 3 ml/min for 10 min. Finally, a quenching solution (50mM Tris-HCl, pH 7.4) was perfused at 3 ml/min for 5 min. Control animals were perfused with PBS only.

Organ preparations. Collected organs were homogenized in a buffer containing 5M urea, 0.25M NaCl and 0.1% SDS. Samples were briefly centrifuged for 5 min and the supernatant was transferred to new tubes. Protein was quantified by standard BCA assay (Thermo Scientific) as per manufacturer instructions. Samples were stored at −80 °C until further analysis. In parallel, a piece

of each organ was homogenized in 1 mL ice-cold PBS, and the samples were plated on Todd-Hewitt Agar for CFU analysis.

Blood Chemistry and Coagulation Assays. Whole blood was collected via cardiac puncture and placed in a pro-coagulant serum tube (BD Microtainer #365967) for 4 hours at room temperature. Serum was isolated by spinning the tubes at 2000xg and collecting the supernatant. All samples were frozen and thawed once before analysis. Blood chemistry parameters were measured on a Cobas 8000 automated chemistry analyzer (Roche) with a general coefficient of variance of <5%. All samples were frozen and thawed no more than two times before analysis. Blood coagulation factor assays were performed as previously described (Wang et al., 2001).

Bacterial Colony Forming Units Counts. Organs of interest were harvested at the indicated time points and placed in a 2mL tube (Sarstedt #72.693.005) containing 1 mL ice cold PBS and 1.0 mm diameter Zirconia/Silica beads (Biospec Products #11079110z). Samples were homogenized using a MagNA Lyzer (Roche) for 2 minutes at 6000 rpm. Whole blood was collected via cardiac puncture and placed in an EDTA tube (BD Microtainer #365974). An aliquot of each organ or blood sample was serially diluted in PBS and plated on Todd-Hewitt Agar to enumerate CFU.

Histological analysis and immunofluorescence. Tissues subjected to H&E stains were first fixed in 10% Buffered Formalin Phosphate (Fisher Chemical) for 24 hours, followed by submersion in 70% ethanol for at least 24 hours. Processing, paraffin embedding, 3 μm section creation, and H&E staining were performed using standard procedures. For immunofluorescence of perfused tissues, organs were harvested immediately following systemic Sulfo-NHS-Biotin perfusion and fixed in ice cold PBS + 4% paraformaldehyde for 18-24 hours with gentle end-over-end agitation. Fixed organs were placed in 40% sucrose solution overnight and incubated until

sucrose saturation. Saturated organs were then submerged in Optimal Cutting Temperature compound (OCT) (Sakura) and flash frozen in cassettes submerged in 2-Methylbutane chilled with dry ice. 20 µm sections were permeabilized and stained for one hour with 1 µg/mL Streptavidin Alexa Fluor 488 (Invitrogen) and 20 µg/mL Isolectin B4 Alexa Fluor 594 (Thermo Fisher Scientific), followed by mounting medium containing DAPI (Thermo Fisher Scientific). Parallel tissue sections were stained using standard H&E staining.

LC–MS/MS DDA analysis. Dried peptides were reconstituted with 2% acetonitrile, 0.1% formic acid, and quantified by modified BCA peptide assay (Thermo Fisher Scientific). Equal peptide amounts derived from each sample were injected and analyzed by LC-MS/MS using a Proxeon EASY nanoLC system (Thermo Fisher Scientific) coupled to a Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific). Peptides were separated using an analytical C18 Acclaim PepMap column (0.075 × 500 mm, 2 µm; Thermo Scientific) equilibrated with buffer A (0.1% formic acid in water) and eluted in a 93-min linear gradient of 2–28% solvent B (100% acetonitrile) at a flow rate of 300 nL/min. The mass spectrometer was operated in positive data-dependent acquisition mode. MS1 spectra were measured with a resolution of 70,000, an automated gain control (AGC) target of 1e6 and a mass range from 350 to 1700 m/z. Up to 12 MS2 spectra per duty cycle were triggered, fragmented by higher energy collisional dissociation (HCD), and acquired with a resolution of 17,500 and an AGC target of 5e4, an isolation window of 1.6 m/z and a normalized collision energy of 25. Dynamic exclusion was enabled with duration of 20 s.

2D-LC–MS/MS analysis. Dried samples were reconstituted in 0.1 M ammonium formate pH ~10 and analyzed by 2D-LC–MS/MS using a 2D nanoACQUITY Ultra Performance Liquid Chromatography (UPLC) system (Waters corp., Milford, MA) coupled to a Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific). Peptides were loaded onto the first-dimension column,

XBridge BEH130 C18 NanoEase (300 µm × 50 mm, 5 µm) equilibrated with solvent A (20 mM ammonium formate, pH 10, first-dimension pump) at 2 µL/min. The first fraction was eluted at 17% of solvent B (100% acetonitrile) for 4 min and transferred to the second dimension Symmetry C18 trap column 0.180 × 20 mm (Waters corp., Milford, MA) using a 1:10 dilution with 99.9% second dimensional pump solvent A (0.1% formic acid in water) at 20 µL/min. Peptides were then eluted from the trap column and resolved on the analytical C18 Acclaim PepMap column (0.075 × 500 mm, 2 µm particles; Thermo Scientific) at low pH by increasing the composition of solvent B (100% acetonitrile) from 1 to 38% (non-linear) over 96 min at 300 nL/min. Subsequent fractions were generated with increasing concentrations of solvent B. The first-dimension fractions were eluted at 19.5, 22, 26, and 65% solvent B, respectively. The mass spectrometer was operated in positive data-dependent acquisition mode. MS1 spectra were measured with a resolution of 70,000, an AGC target of 1e6 and a mass range from 350 to 1700 m/z. Up to 12 MS2 spectra per duty cycle were triggered, fragmented by HCD, and acquired with a resolution of 17,500 and an AGC target of 5e4, an isolation window of 1.6 m/z and a normalized collision energy of 25. Dynamic exclusion was enabled with duration of 20 s.

DDA MS data analysis. MS Raw.files were processed in the MaxQuant platform (Tyanova & Temu & Cox, 2016) (version 1.6.1.0) and searched by the Andromeda search engine (Cox et al., 2011) against the mouse UniProt FASTA database (downloaded 06–02–2017) and against a common contaminant database. Search parameters were set as follows: enzyme, trypsin/LysC with up to 2 potential missed cleavages; fixed modifications, carbamidomethyl on cysteines; variable modifications, oxidation of methionine and acetylation of protein N-terminus; minimum peptide length, 7. The false discovery rate (FDR) for both peptide and protein identifications was set to 1% and was calculated by searching the MS/MS data against a reversed

decoy database. Allowed mass deviation for precursor ions was set to 5 ppm and for peptide fragments was set to 20 ppm. Label-free quantifications (LFQ) was based on a minimum of 2 counts, minimum number of neighbors was set to 3 and average number of neighbors was 6. The match between runs option was applied with a match time window of 0.7 min and an alignment time window of 20 minutes. Some statistical analysis of proteomics results were conducted in the Perseus statistical suite (version 1.6.5.0) of the MaxQuant computational platform (Tyanova et al., 2016). MaxQuant results were imported into Perseus and identified hits were filtered based on the number of peptides (>2) and number of MS/MS scans for each peptide (>2). Missing values were addressed by requiring a cut-off corresponding to 75% valid values in at least one group (infected + biotin, uninfected + biotin, or PBS controls). Remaining missing values were imputed from the normal distribution using a width of 0.3 and a downshift of 1.8. Control samples from PBS-perfused samples were used to subtract the background, keeping protein identifications in labeled samples displaying at least a 2-fold change enrichment.

DIA-SWATH-MS. DIA-SWATH-MS analysis was performed on a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-nLC 1200 ultra-high-performance liquid chromatography system (Thermo Fisher Scientific). Peptides were trapped on a pre-column (PepMap100 C18 3 µm; 75µm x 2cm, Thermo Fisher Scientific) and separated on an EASY-Spray column (ES803, column temperature 45 °C, Thermo Fisher Scientific). Equilibrations of columns and sample loading were performed as per manufacturer's guidelines. Solvent A (0.1 % formic acid), and solvent B (0.1 % formic acid, 80% acetonitrile) were used to run a linear gradient from 5 % to 38 % over 120 min at a flow rate of 350 nl/min. A DIA method was implemented using a schedule of 44 variable acquisition windows as previously reported

(Bruderer et al., 2017). The mass range for MS1 was 350-1,650 m/z with a resolution of 120,000 and a resolution of 30,000 for MS2 with a stepped NCE of 25.5, 27 and 30.

DIA-SWATH-MS data analysis. An in silico spectral library was generated for the reference Mus musculus proteome (EMBL-EBI RELEASE 2020_04: 22,295 entries) using deep neural networks as implemented in DIA-NN (v1.7.10) (Demichev et al., 2020). For search space reduction a list of previously MS detectible mouse peptides were compiled from the Peptide Atlas Project (Desiere et al., 2006), and from two published large-scale DIA proteomics studies (Malmstrom et al., 2016, Zhong et al., 2020). The compiled library with a final number of 667,455 precursors was used for DIA data extraction with a protein q-value of 0.01 and RT-profiling enabled. Data was normalized using the cyclic loess normalization method in the NormalyzerDE tool.

Statistical Analysis. All statistical methods were implemented using Python 3.6.10. Proteomics results from the DIA-SWATH-MS analysis were filtered using a one-way ANOVA followed by a Benjamini–Hochberg procedure to control for a false discovery rate (FDR) < 0.10. Statistically significant identifications were further subjected to principal component analysis (PCA). Proteins were given a standardized score using a z-score normalization. Proteomics results were separately analyzed using Welch's t-test to generate volcano plots and heatmaps.

Functional Enrichment Analysis. Functional enrichment analysis of differentially abundant proteins was performed through the Database for Annotation, Visualization, and Integrated Discovery (DAVID). DAVID was run using default settings with the thresholds count ≥ 2 and EASE ≤ 0.1 (Huang et al., 2009). Visualization of enriched Genome Ontology (GO) terms was produced using code adapted for treemap visualizations from the Web tool REVIGO

(Supek et al., 2011). Additional pathway and GO term analysis were conducted using the online version of Metascape and the BINGO-plug-in within the Cytoscape computational platform.

Non-negative matrix factorization. Plasma DIA proteomics was analyzed using non-negative matrix factorization (NMF). Following pseudocode of Lee and Seung's multiplicative update rule, the distance between the dataset matrix (D) and the dot product of hypothetical partitions (W and H) was minimized using a coordinate descent algorithm (Lee & Seung, 1999). The two resulting matrices were constrained during the optimization to allow for better interpretability. First, the predictions in the matrix H must fall within the range 0.0 to the maximum observed log2(relative abundance). Second, the distributions in the W matrix must fall in the range 0.0 to 1.0 and must sum across signatures to 1.0. In order to determine the number of predicted proteome signatures, k, for each value k, 200 random initializations were made. The average Pearson correlation between the ground truth hierarchy learned from D and the weights hierarchy learned from W and average distance scores (||D - WH||F) are plotted for visual inspections at each k. Cluster number k was finally selected using the elbow selection method as well as meaningful investigation of correlation and proteome signature content.

Clustering.

Abundance Clustering: Significant proteins were grouped based on their relative abundances using k-means clustering. K selection was done through the visualization of distances to cluster centers across k values up to 20 and selecting based on the elbow method (Tarca et al., 2007).

Functional Clustering: Protein identifications were parsed through the medium confidence (PPA>0.4) STRING-DB to build preliminary networks based on physical or functional associations. These networks were subjected to the Louvain method for community detection, a

modularity based clustering algorithm. The Louvain clustering was implemented using the Python packages NetworkX and python-Louvain with default settings (Blondel et al., 2008). To circumvent the heuristic property of the Louvain algorithm, ~1000 parallel random starts were initially performed. Using the combined results from these 1000 iterations, a matrix of cluster clarity was built to store the average number of times two proteins appear in the same community, divided by the number of iterations. Using this community clarity matrix, a k-means clustering was finally performed with k being the average number of communities found across the 1000 runs.

Cluster Annotations: The Jaccard similarity index was also used to determine the overlap of proteins contained between two clusters by quantifying the intersection over the union. A ranked list of functional associations for each cluster were defined by using a hypergeometric test (SciPy v1.4.1). Using the Tau-b statistic from the Kendall rank correlation test (SciPy v1.4.1), functional terms between two clusters were compared by ranked enriched terms across GO terms, Reactome, and KEGG pathways.

Higher-order Molecular Networking. All significant protein identifications from all datasets (i.e., different time points, organs, and compartments) were parsed through the functional clustering pipeline described in the previous section and outlined in the supplemental schematic (Supplemental Fig. 1.2). Each of these higher-order community nodes contained a variable number of proteins (ranging from just a few to hundreds of proteins), operating within similar biological processes. Finally, we computed pairwise spearman correlations over all combinations of detected community nodes, which resulted in detection of significant correlations that were used to define linking edges between the nodes. Benjamini–Hochberg procedure was used to control for a false discovery rate (FDR) < 0.10. The multi-edge was then constructed using three pieces of relational

information. Contained in the multi-edge are the scores: the proportion of significant pairwise spearman correlations, the Jaccard similarity index, and the Tau-b statistics previously defined. Visualization and analysis of the network layers were conducted through Cytoscape (Shannon et al., 2003).

Data Availability. We provide the raw data on MassIVE, a community resource developed by the NIH-funded Center for Computational Mass Spectrometry. Data can be found at: ftp://massive.ucsd.edu/MSV000089030/

Code Availability. We provide the code library in Python described in this work through Github: https://github.com/LewisLabUCSD/CORSPOTS. We provide jupyter notebooks in python to generate our figures and analysis.

1.4 Results

To fully calibrate our experimental methods and data processing, we strategically produced two explorations of the murine blood-tissue proteome during MRSA-induced bacteremia. Initially we display specific targeting of the vascular cell surface proteome through perfusion with a Sulfo-NHS-biotin tag and select four relevant organs for further temporal study. Also, we demonstrate the power of community-based clustering methods for detection of phenomenon in quantified proteomes mapped to the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database. In the second iteration of the blood-tissue proteomic atlas in MRSA bacteremia, we make four major improvements: (i) increase quantitative ability by using a data independent acquisition (DIA) mass spectrometry approach, (ii) dramatically increase the number of samples generated to assess the organ proteomes at four different timepoints, (iii) harvest plasma and parenchymal proteomes for each of the samples, (iv) and build on our previous network analysis to produce a higher-ordered molecular stratification of the multi-compartment proteomic datasets.

Targeted *in vivo* biotinylation of vascular cell surface proteins. The vascular cell surface proteome is in direct contact with varying cell types, signaling molecules, proteins, and debris that fill the circulatory system during a systemic bacterial challenge. During microbial infection, such as MRSA bacteremia, the blood-tissue interface offers a ripe space for tracking disease progression and targeted intervention. However, due to the intrinsic nature of a septic response it is quite difficult to recapitulate the changes in the blood-tissue interface in vitro. Thus, we have used vascular tagging methods, previously reported by Rybak et al. 2005, to identify the dynamics of proteins exposed to the vascular flow in a murine model of MRSA sepsis (Fig. 1.1). Either infected or healthy wildtype C57BL/6J mice were subjected to terminal systemic perfusion with sulfo-N-hydroxysuccinimide (NHS)-biotin and specificity of tagging was verified by labeling and visualization of biotinylated material.

During the first leg of this study, multiple organs from uninfected mice were harvested post perfusion, cryosectioned, and stained with fluorophore-conjugated streptavidin. Biotinylated proteins were detected in the liver around the hepatic central veins and the microvasculature, but not in the parenchyma. Additionally, in the kidney, strong streptavidin reactivity was observed in the glomerular compartments and proximal tubule. In the remaining organs (heart, brain), it seemed biotinylated proteins were localized primarily within the microvasculature with negligible leakage. Co-staining with Isolectin B4 (IB4), a glycoprotein with specific binding to terminal alpha-linked galactose residues on the carbohydrates lining the endothelial lumen, was performed to determine if the biotin labeling was specifically associated with endothelial cell surface (Fig. 1.2). Intracellular staining was low and partial overlap between streptavidin and IB4 staining was observed in all tissues confirming the inability of sulfo-NHS-biotin to penetrate the cell membrane.

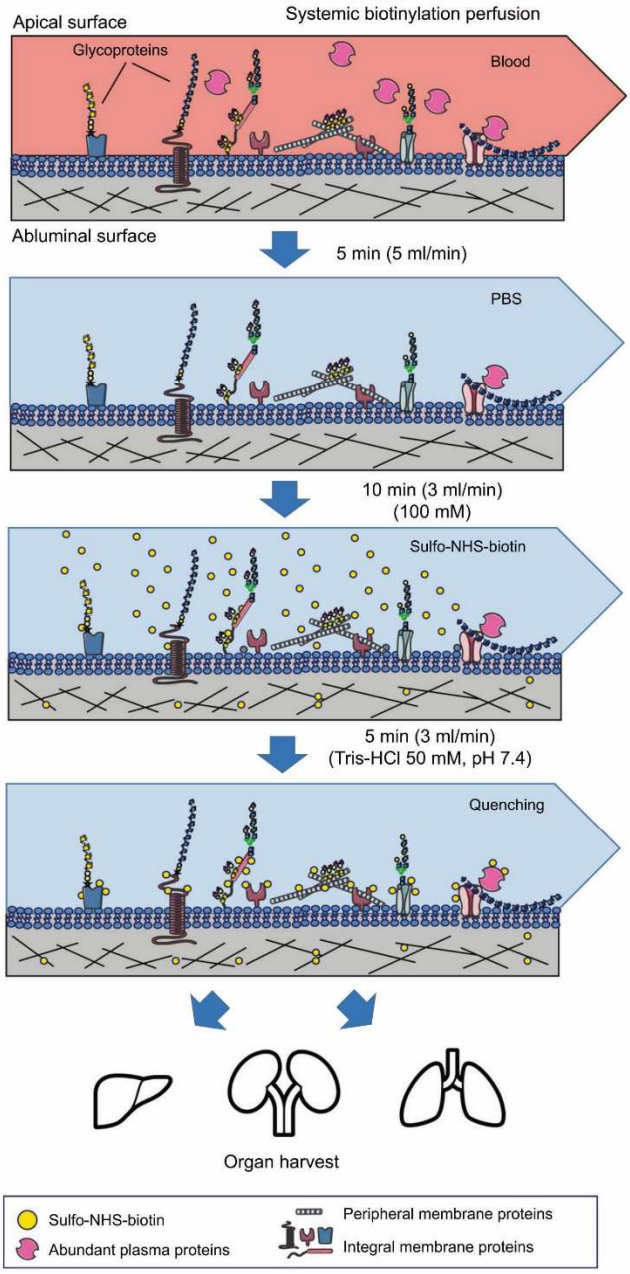


Figure 1. 1. Systemic vascular proteome tagging using biotinylation perfusion

Workflow for in vivo biotinylation of vascular antigens. Animals were first perfused with saline (PBS) to remove blood, followed by biotinylation using an isotonic solution of sulfo-NHS-biotin. Unreacted NHS-groups were quenched by perfusion with a Tris-HCl buffer (pH 7.4). All buffers were kept ice-cold, and the perfusion times were kept as short as possible to minimize potential tissue damage and disruption. After biotinylation, multiple organs were harvested and preserved for histological analysis, or immediately homogenized and subjected to proteomics analysis.

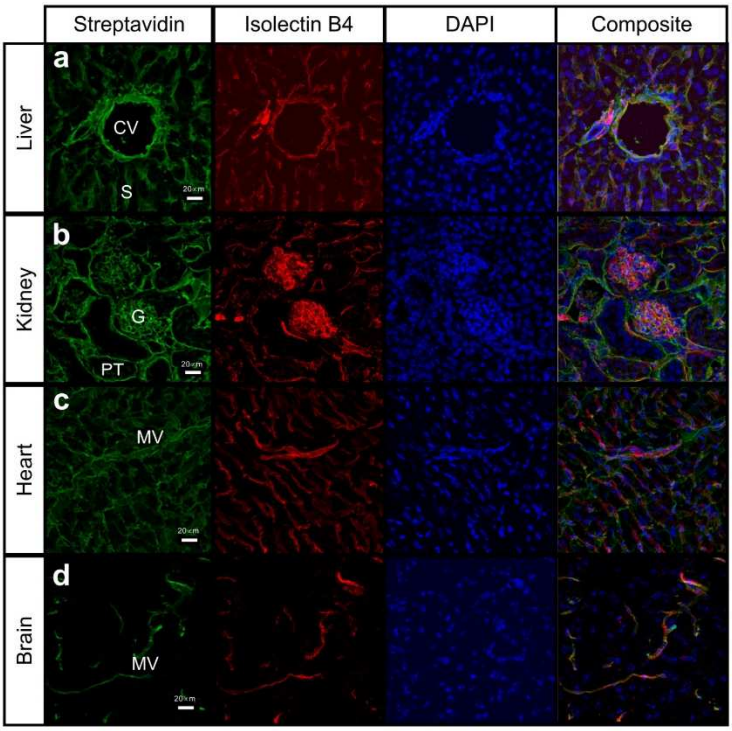


Figure 1. 2. Histological examination of vascular biotinylation

Protein biotinylation is primarily associated with vascular compartments. Murine tissues from animals perfused with sulfo-NHS-biotin were excised and subjected to cryosectioning, followed by histological analysis using fluorescently labeled streptavidin. Cryosections from liver (a), kidney (b), heart (c), and brain (d) were imaged using confocal microscopy. Most of the streptavidin reactivity was closely associated with vascular tissue structures such as the liver sinusoid or the kidney glomerular microvasculature. Tissue slides were also co-stained with IB4 to visualize the endothelial lumen. Partial colocalization between streptavidin and IB4 stains indicated incorporation of biotin into the endothelial glycocalyx but also in the nearby extracellular matrix and the vascular extracellular space. Histological analysis was conducted in biological triplicates, but only representative slides are shown. CV: hepatic central veins, S: sinusoids, G: glomeruli, PT: proximal tubules, MV: microvasculature. Scale bar, 20 μm

Functional commonalities and organ-specific processes found through network clustering of vascular cell surface proteins. To quantitatively validate the systemic labeling of the vascular cell surface proteome, we selected two organs (liver, kidney) to be subject to the streptavidin immunoprecipitation + LC-MS/MS workflow. In this specific experiment peptide digests were analyzed through an online 2D-LC-MS/MS pipeline, as described in the Methods

section, to enhance the proteome coverage through deep fractionation. Matched control tissue from PBS-perfused animals was subject to the same workflow to correct for background proteome signatures represented by non-biotinylated proteins which may non-specifically interact with streptavidin beads. In order to quantitatively account for this phenomenon we only considered peptides which display at least a 2-fold enrichment in measured intensity when compared to the respective PBS-perfused organ homogenate. Two additional data filters were applied to further increase confidence in quantified relative abundance of proteins. Briefly, a positive protein identification would need to be represented by at least 2 unique peptide spectra matches (> 2 unique PSMs) and each unique peptide hit would need at least 2 identified matched spectra.

409 unique vascular cell surface proteins were identified in the liver while 533 proteins were identified in the kidney. There was a high degree of overlap in the protein identifications from liver and kidney (138) and hypergeometric enrichment tests on associated Gene Ontology (GO) terms reveal significant enrichment for proteins located in the plasma membrane and extracellular matrix (Fig. 1.3a). Additionally, these shared proteins contained classical endothelial markers, VE-cadherin, endoglin, vascular cell adhesion molecule 1, and intercellular adhesion molecule 1, as well as, integral transmembrane proteins such as, cadherins and cadherin-like proteins, integrins, collagens, laminins, proteoglycans, cellular receptors, enzymes, and a large repertoire of receptor-type protein tyrosine phosphatases.

A network analysis was conducted in order to characterize the potential protein-protein associations amongst the unique vascular proteomes that could be altered under infection. All identified proteins in each tissue were independently mapped to their corresponding indices in the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database. From this mapping, a network was drawn based on high confidence protein–protein associations (STRING

association scores >0.07). The Louvain method was used to identify communities (or clusters) displaying a higher density of interconnected nodes than expected by random chance (Blondel et al., 2008). These clusters were further segregated via force-directed visualization algorithms, and subjected to functional enrichment analysis. Several results from the Louvain clustering showed a clear pattern of functional commonalities between the biotinylated tissues reassuring the power of the workflow. Interestingly, several distinct pathways were revealed to have strong organ specific mappings.

Liver samples were immediately recognized by protein and pathway identifications belonging to the hepatic reticuloendothelial system and those involved in lipoprotein metabolism and uptake. Multiple scavenger receptors such as C-type lectin family 4 F (Clec4F), Scavenger Receptor Class B Member 1 (Scarb1), stabilin 1 (Stab1), Stab2, and asialoglycoprotein receptor 2 (Asgr2), as well as, low density lipoprotein receptor (Ldlr), angiopoietin-related protein 3 (Angptl3), hepatic triglyceride lipase precursor (Lipc), fatty acid binding protein 1 (Fabp1), and LDLR associated protein 1 (Ldlrap1) were unique to these samples. Approximately 33% of liver protein identifications clustered into 4 distinct Louvain communities. Each community was represented by a distinct set of pathway enrichments which highlighted distinct functions of this organ (Fig. 1.3b-e). In Fig. 1.3c potential vascular biomarkers of scavenging are represented, while in Fig. 1.3d potential biomarkers of lipoprotein signaling are shown.

Kidney samples were enriched for proteins such as, angiotensin converting enzyme (Ace), Ace2, glutamyl aminopeptidase (Enpep), and renin 2 (Ren2) which are directly involved in the regulation of blood pressure and fluid balance, specifically components of the renin-angiotensin system. Approximately 31% of kidney protein identifications clustered into 5 distinct Louvain communities. Again, each of these communities (Fig. 1.3f-j) contained several functional layers

of kidney tissue. Multiple communities for blood vessel regulation and smooth muscle contraction were identified and biomarkers noted (Fig. 1.3g-h). Additionally, biomarkers were noted in a community revolving around glutathione metabolism which also surfaced (Fig. 1.3j).

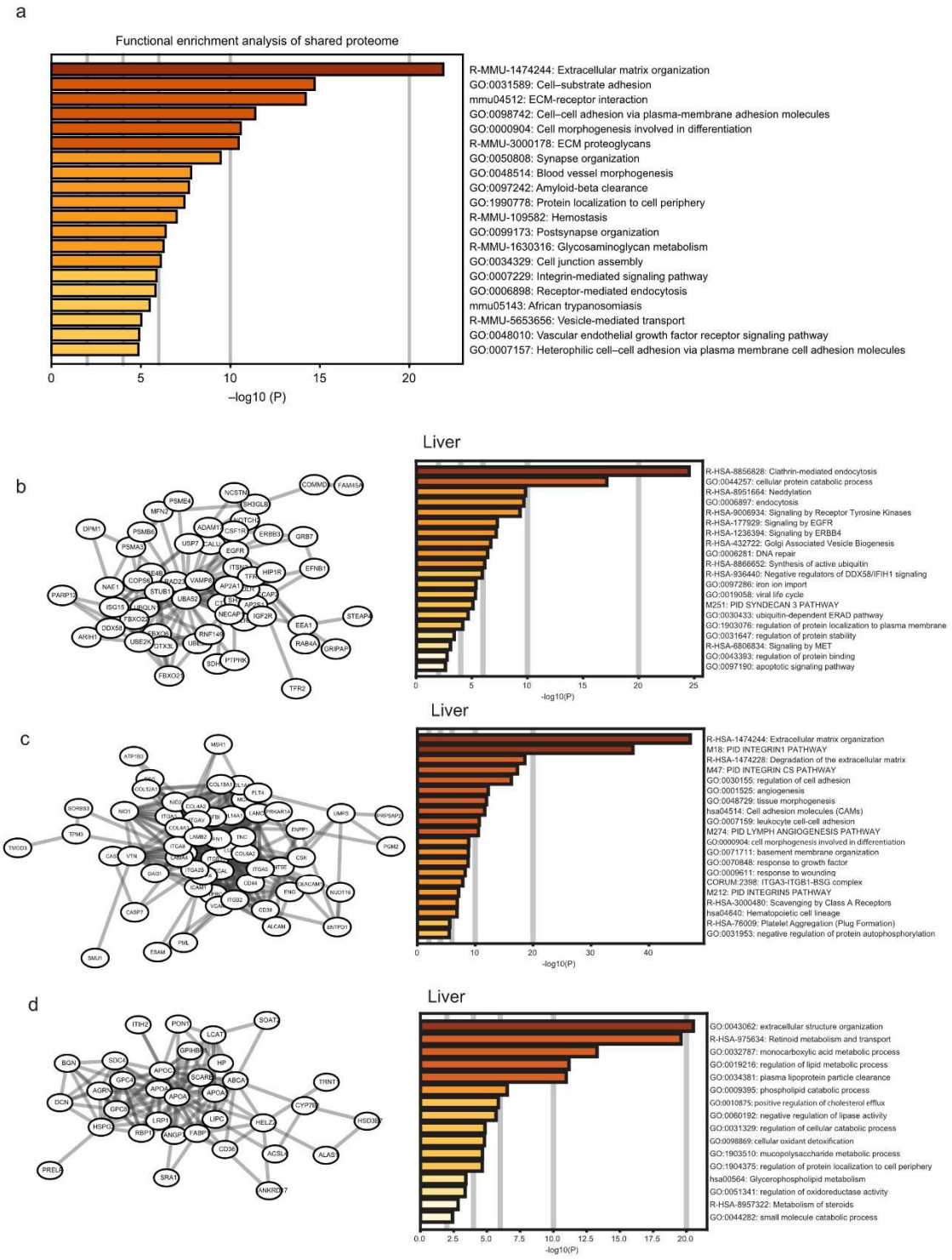


Figure 1. 3. Functional enrichment of biotinylated vascular proteins

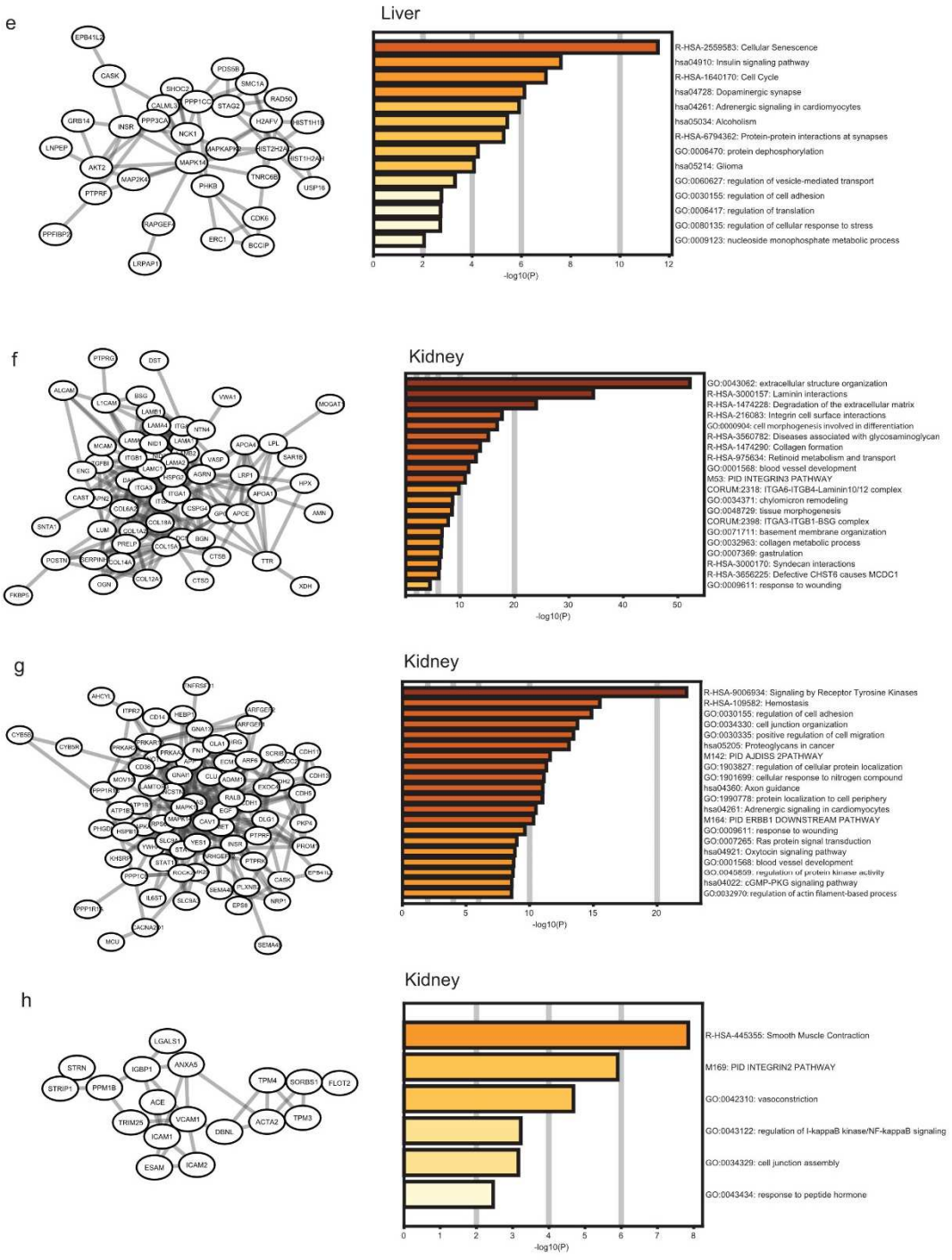


Figure 1. 3. Functional enrichment of biotinylated vascular proteins

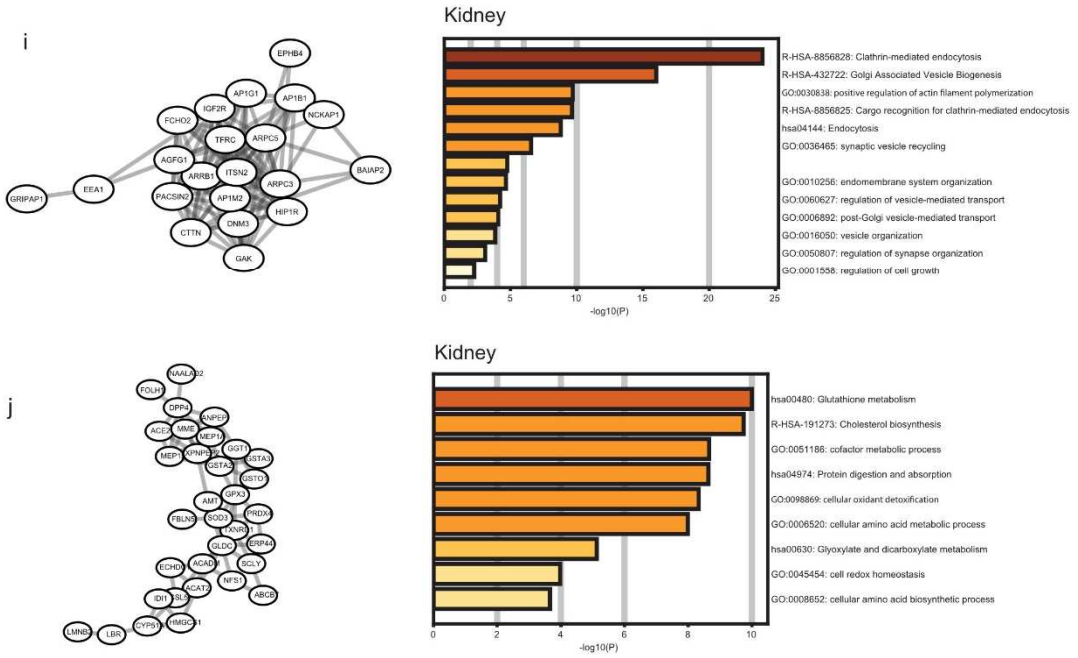


Figure 1. 3. Functional enrichment of biotinylated vascular proteins

Functional enrichment analysis of proteins located in the cell membrane and extracellular matrix in both liver and kidney (a). Proteins identified through proteomics analysis were used to generate a protein-protein association network through the STRING database. Protein interactions were limited to high confidence physical or functional associations (association score > 0.7). The final network was subjected to Louvain clustering to identify highly interconnected communities. The identified clusters were further segregated via force-directed visualization algorithms and subjected to functional enrichment analysis using the web-based version of Metascape. Liver samples were segregated into 4 different clusters (b-e) whereas kidney proteins were clustered into 5 different communities (f-j). Some of these clusters were enriched in specific functions and organ specific biological pathways.

Assessment of accessible vascular surface proteins in various organs during MRSA sepsis. As previously shown, the biotinylation perfusion assay was able to selectively label protein targets within the vascular cell surface of several organs for mass spectrometry measure. Thus, this workflow was used to render a molecular snapshot of the septic organ vasculatures during sepsis triggered by MRSA bacteremia. For each experiment wild type C57BL/6J mice were intravenously infected through the retroorbital route with 5 × 107 colony-forming units (cfu) of

MRSA, a model that induces lethality within 48 h post-infection (Kim et al., 2014), or with PBS as a control. We focused on the early pre-mortality stages of the disease 24 h post-infection in order to capture vascular changes that precede organ failure. Briefly, for each experiment, 3 MRSA-infected and 3 uninfected mice were anesthetized and subjected to the systemic biotinylation protocol at 24 h post-infection. In addition, one infected and one uninfected mouse were perfused only with saline to account for potential background signals. Five different organs were selected for proteomics analysis based on their clinical relevance: liver, kidney, heart, brain, and white adipose tissue (WAT). In total, we performed four independent experiments resulting in a total of 160 LC-MS/MS runs. All samples were analyzed using the DDA approach in a single run LC-MS/MS and strict bioinformatic and statistical criteria were used to focus on robust and reproducible changes, as specified in the Methods section.

After protein identification, quantification, and normalization, each individual organ proteome was assessed for changes using a two-way analysis of variance (ANOVA) with a permutation-based false discovery rate correction for multiple test comparisons. From this analysis there were several hundred robust protein alterations found across the organ set (liver: 272, kidney: 300, heart: 275, WAT: 185, and brain: 85). In liver, kidney, and heart, there were many more proteins altered between infected and non-infected samples than brain and WAT, a piece of selection criteria for the next study. These LFQ-intensities were visualized with a principal component analysis (PCA) to determine if the total organ protein patterns could differentiate infection clearly (Fig. 1.4a). Liver, kidney, and brain had clear meaningful separations whereas WAT and heart samples did not. 57 differentially abundant proteins were shared across all of the organs. Interestingly their patterns of expression varied on an individual basis. Some proteins exhibited consistent changes across organs while others had a scattered expression. The shared

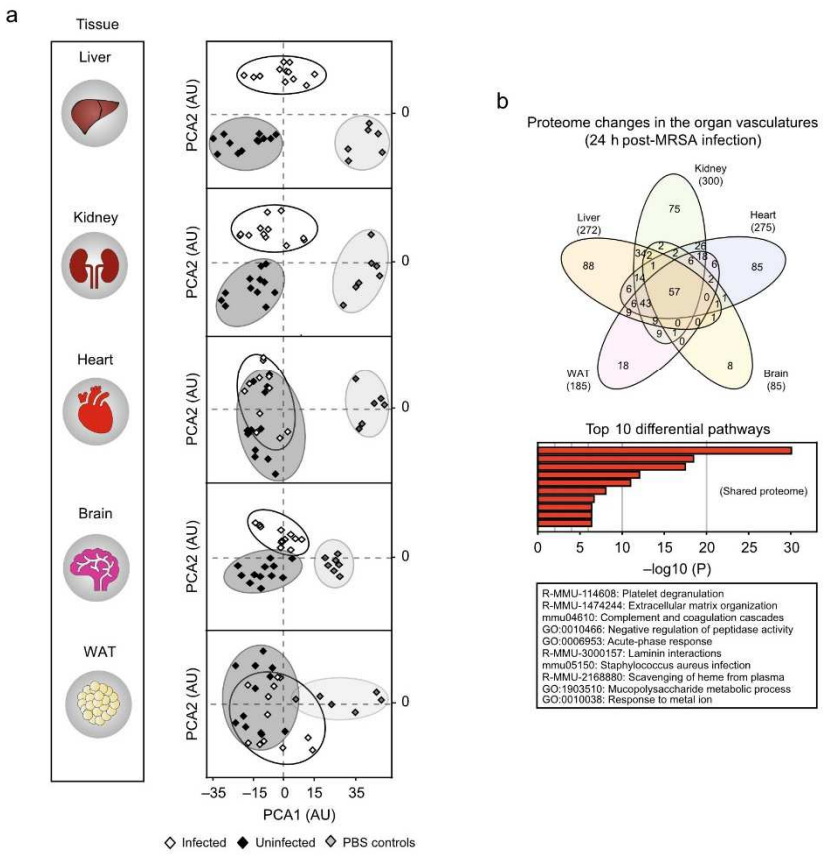


Figure 1. 4. Remodeling of the vascular surfaces during MRSA-sepsis

Principal component analysis (PCA) of the vascular proteins differentially regulated during MRSA-sepsis (infected n = 12, uninfected n = 12) segregated the liver, kidney, and brain tissues into infected and uninfected groups, but was less specific for WAT and heart stratification (a). Venn diagrams of significant protein hits across the organs revealed that a total of 57 proteins were shared among all examined tissues, whereas other targets were changing in a tissue-specific fashion. Functional enrichment analysis using the Metascape tool indicates that biological processes related to coagulation, acute phase responses and ion hemostasis are highly enriched in the shared proteome across all organs (b).

proteome was enriched for components of several key biological pathways associated with coagulation, complement, ECM-remodeling, and *Staphylococcus aureus* infection. Tissue-specific proteins were identified by significance in abundance being in exclusively one organ. Unsurprisingly liver, kidney, and heart had the most tissue-specific infection dependent markers (liver: 88, kidney: 75, heart: 85, WAT: 18, and brain: 8) (Fig. 1.4b). In the liver, neutrophil

degranulation was the most enriched biological process, reflecting the presence of multiple neutrophil-derived proteins such as myeloperoxidase (MPO), an essential factor for neutrophil antimicrobial responses. Contrastingly, heart-specific processes were highly enriched in proteins related to muscle contraction and markers of hypertrophic cardiomyopathy (HCM), reflecting potential ongoing heart failure. These two seemingly different timelines, with heart rapidly approaching markers for organ failure, became an interesting focal point of the next study.

Resolution of selected organs and time course for DIA study of MRSA sepsis. Due to our previous investigation of the vascular proteome and our goals of covering large variable proteomes we have assessed the mass spectrometry method used. We have found that a crucial component to consider is the data acquisition technique of the LC-MS/MS quantification pipeline we are using. We have transitioned to using data independent acquisition (DIA) for our proteomic studies of organs and plasma in MRSA induced sepsis due to the increased resolution and accuracy of the method. DIA-SWATH-MS uses broad isolation windows to measure almost all MS detectable peptides in a biological sample. This results in an increased number of protein identifications and more robust quantification compared to approaches based on data-dependent acquisition (DDA) (Bruderer et al., 2017). We evaluated the sensitivity and precision of using the DIA approach by performing a study of blood plasma proteins in a murine model of MRSA derived sepsis (0hr, 6hr, 12hr, 24hr). Compared with traditional data dependent acquisition (DDA), DIA detected 95.6% (195/204) of previously detected proteins and increased the number of detected proteins by ~3 fold (598 vs. 204). Additionally, we performed a simple differential abundance analysis for 3 comparisons (Fig. 1.5). The DIA approach performed astoundingly well, capturing ~3 fold more significant differences in all 3 comparisons.

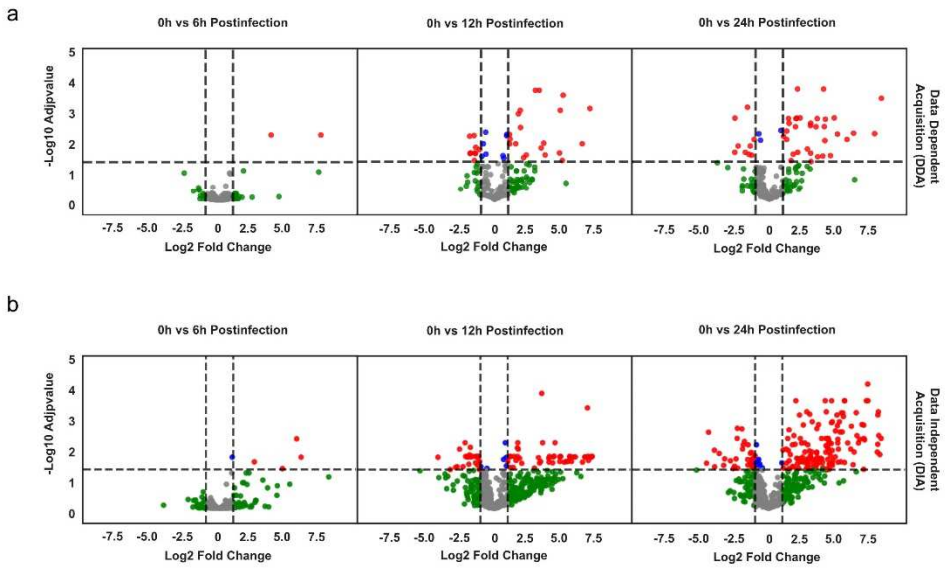


Figure 1. 5. Comparison of data acquisition techniques for proteomics of septic blood

Differential abundance analysis for 3 time points post-infection with MRSA vs. uninfected in a murine model of sepsis from two separate experiments. Data dependent acquisition (DDA) LC-MS/MS experiment conducted on plasma proteome for septic mice (a). Data independent acquisition (DIA) LC-MS/MS experiment conducted on plasma proteome for septic mice (b). An alpha of 0.05 was chosen for consideration of significance. Coloring thresholds were set at a fold change of >2 and adjusted P value of <0.05.

As shown in Fig. 1.5b, the plasma proteome is significantly altered during infection, and clustering analysis segregates the samples into early and late responses (Fig. 1.6a). We used a non-negative matrix factorization (NMF) to uncover the major protein changes driving the early and late responses (Lee & Seung, 1999). Three NMF clusters or "protein signatures" were successfully identified and condensed into discrete proteome trajectories (i.e., temporal intensity patterns) (Fig. 1.6b). The first signature was defined as the early (6h) inflammation and acute-phase response axis being represented by an accumulation of inflammatory proteins, and acute-phase reactants (APRs), including the C-reactive protein (CRP), the serum amyloid protein 2 (SAA2), and soluble CD14,

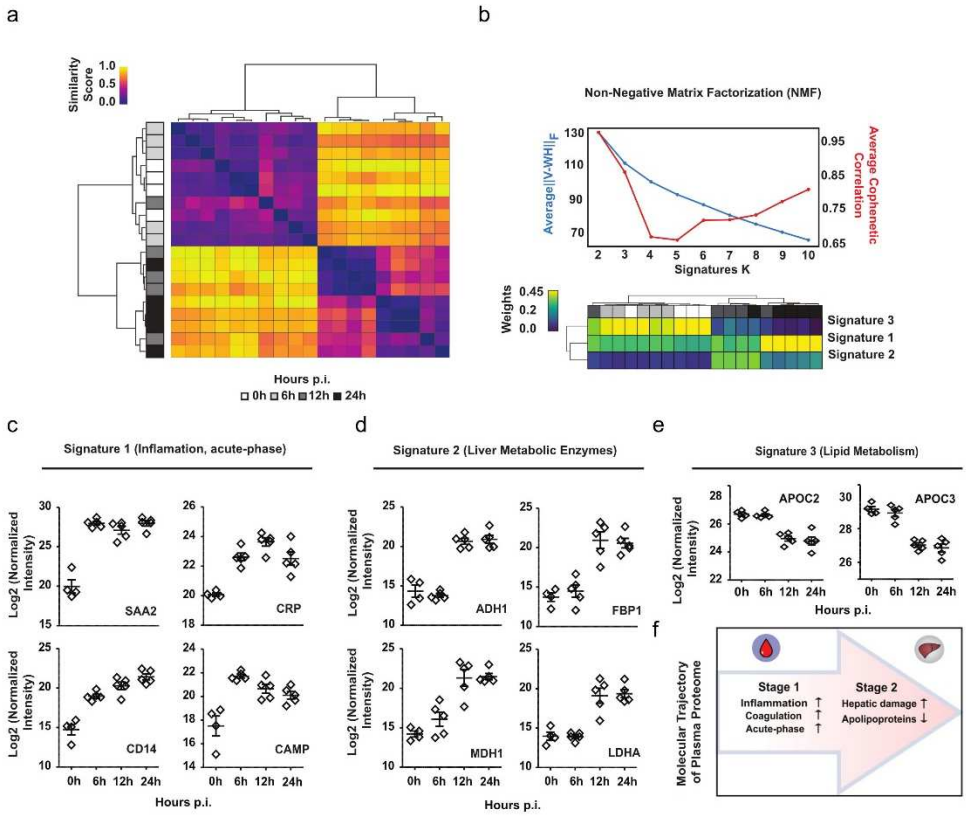


Figure 1. 6. Time-resolved analysis of plasma proteome trajectories in a murine model of Staphylococcus aureus sepsis

Protein abundances in murine plasma at 0 h (n = 5), 6 h (n = 5), 12 h (n = 5), and 24 h (n = 5) postinfection and captured by proteomics analysis using data-independent acquisition (DIA) mass spectrometry. Hierarchical clustering of the time-course plasma samples resolved by the DIA-SWATH-MS proteomic analysis (a). Nonnegative matrix factorization (NMF) analysis of the plasma proteome changes converges into three major protein signatures with distinct temporal behavior. The top graph depicts the model fit for the NMF computed at k = 2 to 10 across 200 random initializations. The bottom heatmap represents the distribution of the three major protein signatures selected from the top scoring model (k = 3) (b). Representative normalized intensity values for the top proteins in the NMF signature 1 (c), signature 2 (d), and signature 3 (e). Schematic depiction of the two-staged model of the plasma proteome trajectory over the course of infection (f).

as well as antimicrobial factors from the cathelicidins family (Fig. 1.6c). The other two signatures were able to split the late stage (12h, 24h) response into distinct characterizations of the occurrence of acute liver failure. One was represented by the plasma accumulation of liver metabolic enzymes (Fig. 1.6d), whereas the final signature identified the depletion of hepatic factors involved in lipid

metabolism (Fig. 1.6e). Based on these data-driven patterns, the plasma proteome analysis reflected a two-staged model of the progression of sepsis, each stage associated with specific inflammatory and metabolic alterations.

Four major organs were selected for measurement of both parenchymal and vascular surface proteins using the DIA-SWATH-MS protocol (liver, kidney, heart, and spleen). We found these organs are clinically significant, highly vascularized, and have specific pathological changes taking place in the model of *S. aureus* sepsis, as previously reported in Toledo et al., 2019. Thousands of proteins were quantified to track the progression of parenchymal and vascular cell surface response during systemic infection. Principal component analysis (PCA) of the significantly dynamic proteomes (tested by a one-way ANOVA) for each of these compartments across the four organs immediately showed both shared and organ-specific proteome trajectories (Fig. 1.7). All four organ vascular surfaces displayed a dramatic remodeling in the 6h post infection proteomes when compared the parenchymal tissue. The parenchymal proteome responses were highly heterogeneous across the organs, displaying different temporal patterns. For example, the liver tissue was governed by early changes and a significant separation of the samples at 0h and 6h, whereas there were no notable differences between the 0h and 6h heart tissue proteomes, indicating a more delayed response. The kidney and the spleen were also quick responders, although the spleen showed the largest parenchymal proteome changes of all organs, with >2000 proteins significantly altered during sepsis. All vascular proteome datasets displayed a distinct separation of the 0h and 6h time points, and clustering of the 12h and 24h time points, independently of the organ. Despite this similar behavior, the number of dysregulated proteins that were shared across the organ vascular fractions at 6h was very low (41, ~1.4%). These findings

suggest that sepsis induces early changes in the vascular proteome of all organs, but each proteome response is most likely shaped by organ specific factors.

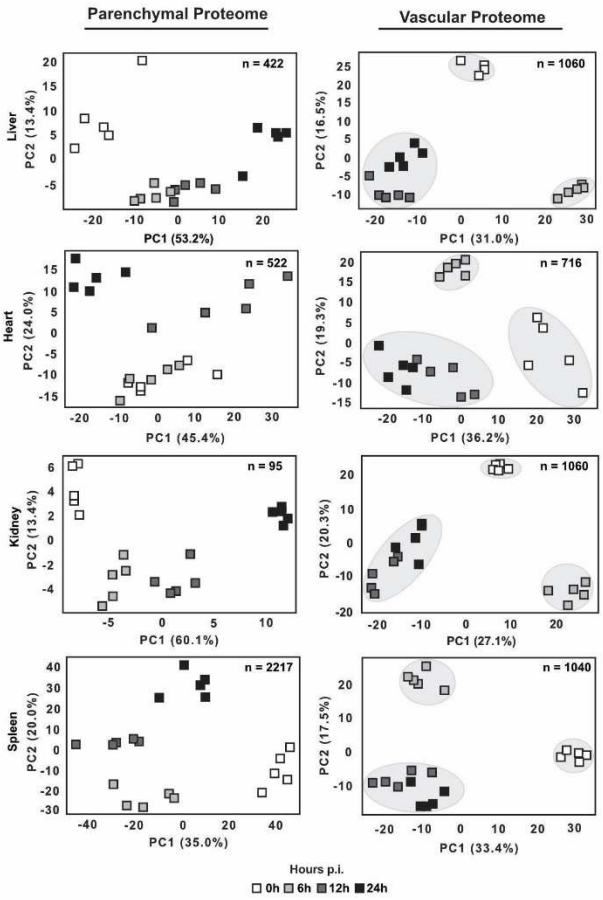


Figure 1. 7. Principal-component analysis (PCA) of significantly altered parenchymal and vascular proteins during infection

Proteins identified displaying statistically significant temporal patterns using a one-way ANOVA (FDR, <0.10) were subject to PCA analysis to identify time-dependent variability for liver parenchymal proteins (n = 422), heart parenchymal proteins (n = 522), kidney parenchymal proteins (n = 95), spleen parenchymal proteins (n = 2,217), liver vascular proteins (n = 1,060), heart vascular proteins (n = 716), kidney vascular proteins (n = 1,060), and spleen vascular proteins (n = 1,040).

Over 3500 proteins were accurately quantified across the vasculature of the organs resulting in a substantial increase in the number of identifications compared with our previous study. Additionally, over 2,900 proteins display statistically significant perturbation in the vascular cell surface proteome, making up a large fraction of the proteins that reside in this compartment

(Fig. 1.8a-b). Using a k-means method coupled to a silhouette analysis of cluster clarity, it was found that septic vascular cell surface proteome has a more complex kinetics than the plasma proteome, displaying clusters with similar temporal behaviors across the organs, and protein clusters exhibiting organ-specific changes (Fig. 1.8c). Four clusters (heart cluster-5, spleen cluster-2, liver cluster-3 and kidney cluster-3; Blue) shared a significant number of proteins (~15%), and common biological pathways, including neutrophil and platelet degranulation, activation of coagulation and complement, and regulation of leukocyte-endothelial interactions. The abundance of all proteins in these clusters was linearly increased over the time of infection. Additionally, three other clusters (spleen cluster-6, liver cluster-1 and kidney cluster-1; Green) also shared many proteins (~14%), as well as pathways related to phagocytosis and the adaptive immune response, particularly receptor mediated interactions. Alternatively, the abundance of all proteins in these clusters dramatically dropped at 6h post infection, suggesting the rapid downregulation and/or shedding of a large fraction of the vascular proteome. We mapped these down-regulated proteins to the plasma compartment and found that the levels of some of them but not all were simultaneously increased in plasma, suggesting complex interaction between tissue source, shedding/downregulation, and systemic clearance. In sharp contrast to these general responses across the organs, evidence of tissue-specific changes was also detected. For example, some specific clusters (heart cluster-3, spleen cluster-5, liver cluster-4 and kidney cluster-4; Orange) shared only a few proteins (~2%), suggesting the induction of organotypic responses. Functional enrichment linked these clusters to ion transport and homeostasis (heart cluster-3), metal sequestration by antimicrobial proteins (spleen cluster-5), cellular responses to labile heme (liver cluster-4) and mitochondrial dysfunction (kidney cluster-4). Notably, despite the marked

differences in their protein contents, these clusters exhibited synchronous proteome trajectories (Fig. 1.8d-h).

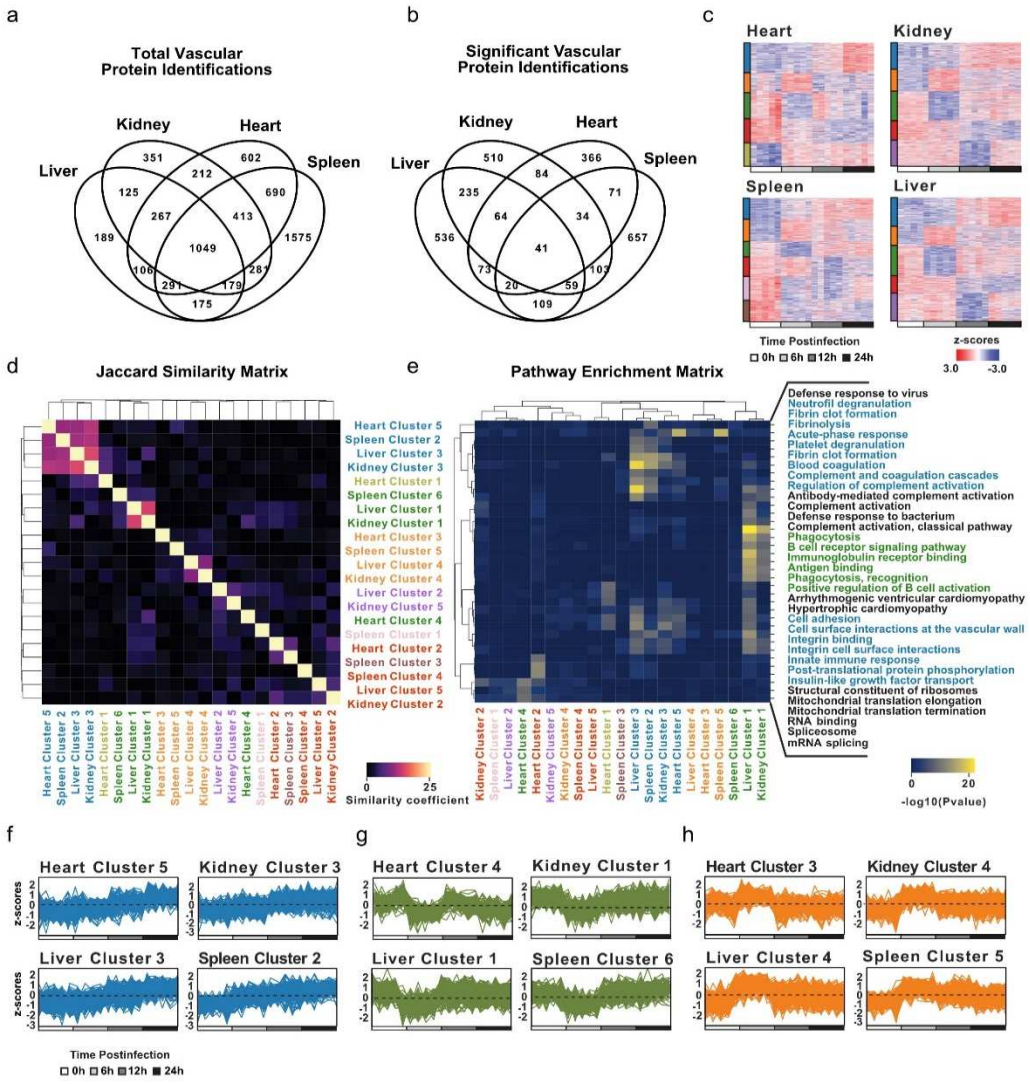


Figure 1. 8. Coordinated vascular proteome responses triggered by S. aureus sepsis

Total number of proteins identified through systemic vascular tagging across liver, kidney, heart, and spleen at 0 h, 6 h, 12 h, and 24 h postinfection (n = 4 to 5 per group) (a). Significantly altered proteins using a one-way ANOVA (FDR, <0.10) across the four organs and the intersections between them (b). Heatmaps showing the k-means clustering of significant protein identifications in the heart, kidney, spleen, and liver. Clusters with the same proteome trajectory (i.e., temporal intensity patterns) are coded with the same color as represented in the bars to the left of the heatmaps (c). Jaccard similarity matrix (intersection over the union) representing the percent overlap of the protein identifications across vascular clusters. Blue and green squares refer to subsets of vascular clusters displaying a high overlap of proteins, and the orange squares represent a subset of vascular clusters displaying low overlap. All cluster colors are according to the proteome trajectories (d). Hierarchical clustering of the main biological pathways obtained from functional enrichment analysis of the vascular clusters colored according to P value (e). Trace representations of the normalized protein abundances (Z-scores) for the vascular clusters that increase over time (f), decrease at 6 h (g), or increase at 6 h postinfection (h).

Higher-order molecular networking of multi-compartment proteomics data in MRSA sepsis. Inference of the specific links between vascular failure and organ failure in sepsis relies on integrating vascular remodeling and tissue remodeling into a single projection. However, meaningful projection of temporal data from multi-level datasets poses a major challenge to systems biologists. To reconstruct the disease trajectory of sepsis using the host proteome response as a readout, we developed a novel computational pipeline to integrate the data across time points, organs, and tissue compartments. Correlative Stratification of Proteome Trajectories (COR-SPOTS) was designed as a higher-order networking approach to store, organize and integrate multiple data layers based on temporal and functional relationships between the data points (see Methods section). Significantly dysregulated proteins from each organ and compartment (plasma, vascular, and parenchymal) were parsed through STRING-DB to group proteins based on functional associations (von Mering et al., 2003). Highly interconnected clusters of nodes were then identified through the Louvain algorithm for community detection and collapsed into new single nodes. We then computed ~24 million pairwise spearman correlations over all combinations of detected community nodes, which resulted in ~3.7 million significant correlations (false discovery rate; FDR<0.10) that were used to define the linking edges of the network. Approximately an equal number of significant positive and negative correlations were observed. Other similarity relationships between the nodes such as the content overlap and the functional enrichment were also stored within the network, resulting in a compiled and fully searchable small database containing all proteomics results and metadata from this study, which can be found on the online repository the Network Data Exchange (NDEx) with UUID: 45474980-9d56-11eb-9e72-0ac135e8bacf (Fig. 1.9a). To demonstrate the utility of COR-SPOTS to dissect molecular changes across the blood-tissue interfaces, we ranked the nodes based on highest inter-community

correlation and selected the hepatic vascular, parenchymal and plasma nodes which were top scoring. Proteins stored within these highly correlated nodes displayed a bimodal network topology characterized by the presence of two distinct anti-correlated clusters of proteins (Fig. 1.9b). Cluster 1 contained most of the plasma and parenchymal proteins linked to triglyceride and phospholipid metabolism, whereas cluster 2 was enriched in vascular proteins linked to coagulation. This was consistent with an overall model tendency for dominant coagulopathic phenotype in the liver as previously reported. Inspection of the cardiac nodes in the network showed a distinct separation of the parenchymal nodes from the vascular nodes, distinguishing the heart from the other organ proteomes (Fig. 1.9a). We hypothesize that this disparity in the cardiac protein response was linked to a difference in the kinetics of the disease trajectory of sepsis in the heart, and perhaps linked to the presence or absence of major cellular events, such as bacterial invasion and/or leukocyte infiltration of the organs.

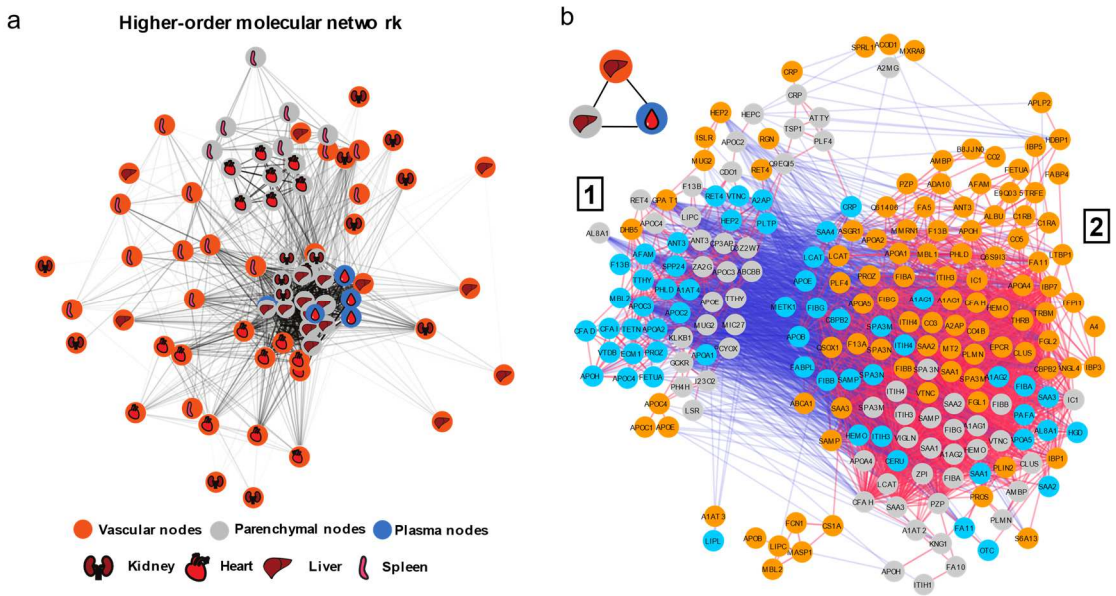


Figure 1. 9. Higher-order molecular networking reveals vascular proteome perturbations

Complete network analysis of all proteome changes across organs and tissue compartments visualized by the edge-weighted spring embedded layout, where the edge distance and shade represent the proportion of significant pairwise correlations between two nodes. Each node color represents the compartment of origin with tissue represented by gray fill color, vascular represented by orange fill color, and plasma represented by blue fill color (a). Functional cross talk between liver parenchymal, vascular, and plasma proteins. Three intercorrelated higher-order nodes were selected, and their pairwise protein correlations were extracted. The subnetwork contains 2 major anticorrelated clusters of proteins. Edges are colored using a blue, white, and red continuous scale ranging from −1.0 to 1.0 (b).

Vascular cell surface proteome changes precede canonical markers of organ damage. Rapidly noticeable in the first hours after systemic infection with MRSA (6hr samples), many proteins involved in lipid metabolism were immediately diminished from the hepatic vasculature. This included multiple apolipoproteins, angiopoietin-like proteins, and the glycosylphosphatidylinositol-anchored high-density lipoprotein-binding protein 1 (HDBP1). While these changes slightly registered in the liver parenchyma and plasma during various points of infection, these changes were more delayed when compared with the vascular proteome. On the other hand, proteins involved in coagulation were rapidly accumulated in the hepatic vasculature. This included deposition of several platelet glycoproteins (GPIX, GP1BB, GP1BA, GPV), indicating abundant intravascular platelet deposition, together with an increase in the levels of Von Willebrand Factor (VWF) and fibrinogen. Most of these changes were specific to the vascular proteome and were only weakly captured in the parenchyma and plasma compartments towards the later time points (12h and 24h). Notably, the levels of Tissue Factor Pathway Inhibitor 1 (TFPI1) at 6h post infection were rapidly depleted, strongly suggesting a prothrombotic priming of the liver vasculature very early during infection (Fig. 1.10a). Organ failure based on the increased levels of AST and ALT in circulation was not apparent until 12h post-infection (Fig. 1.10c), which coincided with the occurrence of massive thrombosis and tissue necrosis in the liver towards the late time points (Fig. 1.10d). Further confirming that *S. aureus* bacteremia results in systemic coagulopathy, the levels of D-dimers in plasma were significantly increased at 24h post infection, and widespread consumption of multiple coagulation factors was also confirmed by enzyme activity assays (Fig. 1.10b). Additionally, most animals lost significant amounts of weight over time due to metabolic alterations and/or changes in body fluid distribution (Fig. 1.10c). Taken

together, the data suggest a global reprogramming of the liver by shutting down metabolic functions as the level of intravascular coagulation and vasculopathy increases.

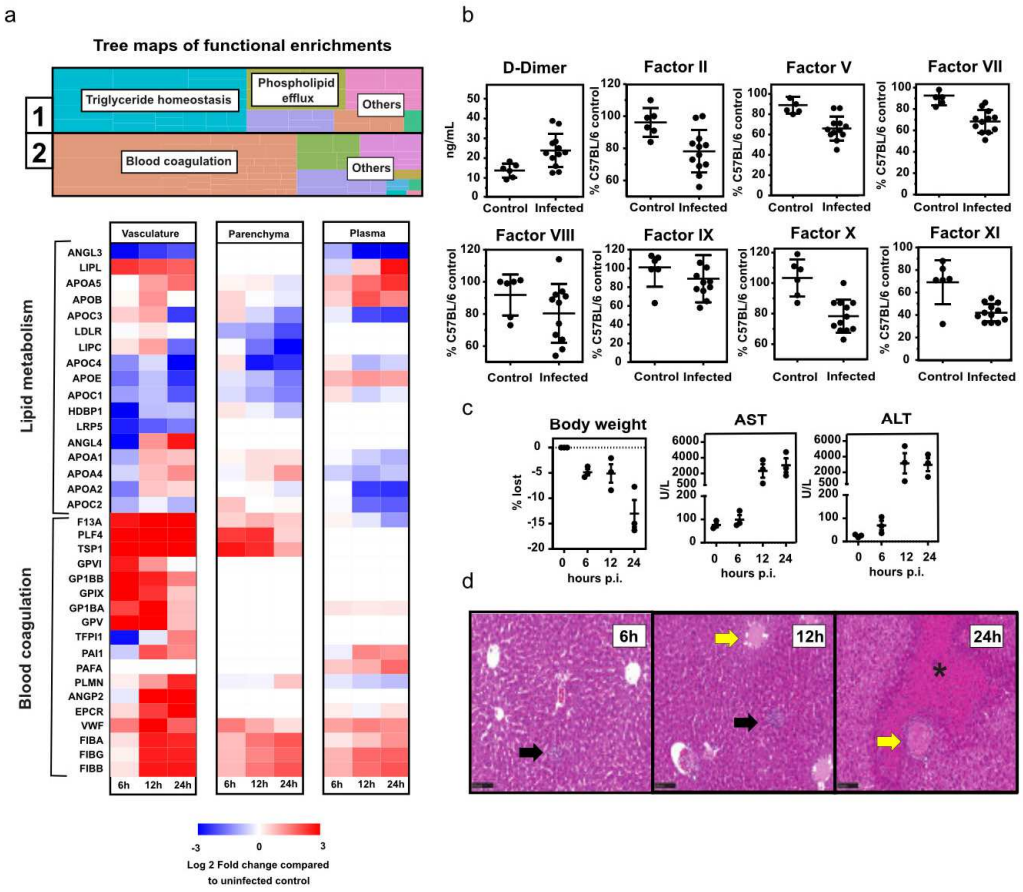


Figure 1. 10. Network markers of vascular proteome perturbations precede coagulopathy and hepatic damage

Functional enrichment analysis of the 2 major protein clusters from higher-order networking visualized by treemaps (top). Perturbation of lipid metabolic proteins and upregulation of blood coagulation components are represented by compartmentalized heatmaps of the liver parenchymal tissue, vasculature, and blood plasma. Log2-fold change values are compared with uninfected proteome measurements (0 h) for each compartment (bottom) (a). Enzyme activity assays measuring blood coagulation factors present in murine plasma at 24 h postinfection (b). Changes in body weight, AST, and ALT (c). Representative images of the hematoxylin and eosin stain of liver tissue over the time course. Leukocyte infiltration was noticed as early as 6 h postinfection (black arrows) and vascular occlusion and thrombosis as early as 12 h postinfection (yellow arrows); necrotic areas are marked with an asterisk (*) (d).

The heart vascular proteome showed rapid remodeling 6hr post infection; however, no bacteria were found in the heart until 12h after inoculation. This is in stark contrast to the bacterial burden of the organs which show presence of bacteria in the liver, kidney, and spleen already by

6h post infection. Since neutrophils are the primary responders to the presence of bacterial infiltrates in the tissues, we assessed the levels of myeloperoxidase (MPO) and neutrophil elastase (ELANE), two granule proteins that are specific to neutrophil degranulation and formation of extracellular traps. The levels of MPO and ELANE in the liver were rapidly raised upon infection both in the parenchyma and vascular proteomes, whereas their deposition in the cardiac parenchyma and vasculature was first observed after 12h (Fig. 1.11a). Histological analysis confirmed that neither bacterial micro abscesses nor immune infiltrates were observable in heart tissue until 12h post infection (Fig. 1.11b). Taken together, these data suggests that in this model of *S. aureus* sepsis, bacterial dissemination and leukocyte infiltration are temporally delayed in the heart compared to the other organs, opening an interesting window to capture very early molecular changes in the vasculature as a response to systemic inflammation. Next, we focused on the changes in the heart vasculature at 6h to identify molecular alterations that precede bacterial and leukocyte invasion of the tissue (Fig. 1.11c). Multiple vascular adhesion proteins were rapidly induced during infection, such as CADM1, ICAM1 and VCAM1. Upregulation of junctional endothelial proteins such as JAM3 was also observed, together with an increase in the levels of ADAM17, a protease that has been linked to most of the early shedding of the vascular surfaces upon inflammation. Changes in the glycocalyx were also noticed, including upregulation of heparan sulfate proteoglycans such as GPC4, and very robust induction of the hyaluronic acid synthase HAS1. In fact, early upregulation of HAS1 was consistently observed across all vascular surfaces, directly implicating HAS1 in glycocalyx remodeling during sepsis (Fig. 1.11d).

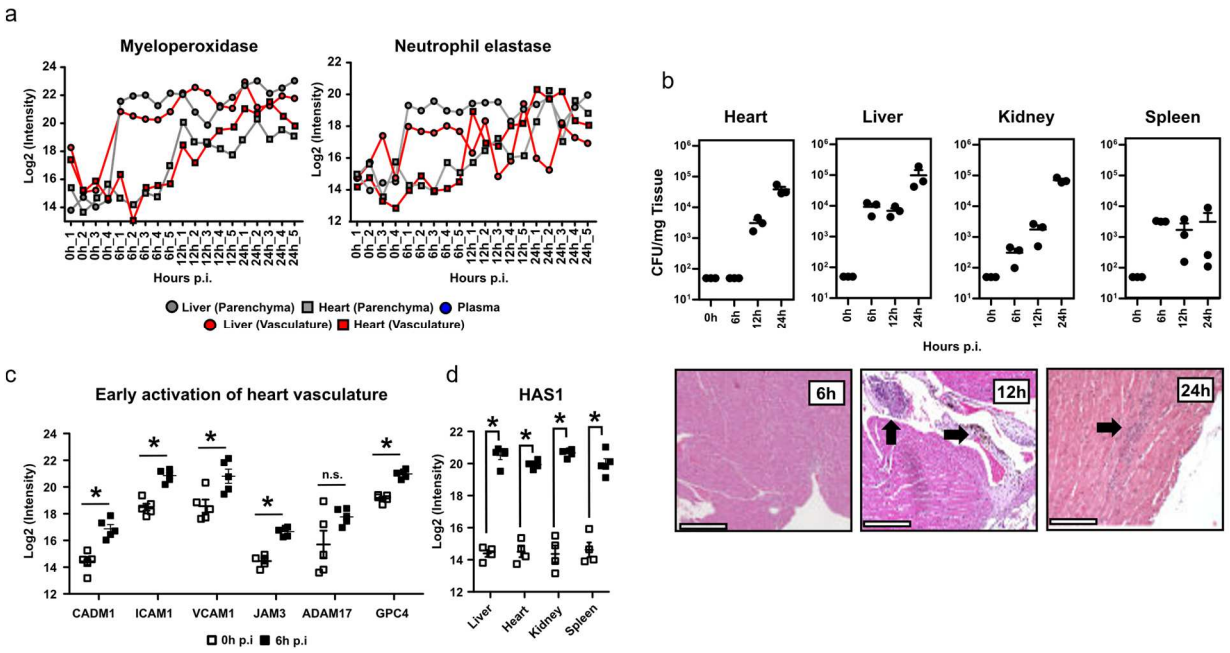


Figure 1. 11. Bacterial invasion and neutrophil infiltration are delayed in the heart compared with other organs

Proteomic measurements of myeloperoxidase (MPO) and neutrophil elastase (ELANE) for both liver and heart vascular and parenchymal proteomes. Liver measurements are represented by filled circles; heart measurements are represented by filled squares. Tissue was represented by a gray fill color, and vascular measurements by orange fill color (a). Organs were harvested, homogenized, and plated to quantify the bacterial CFU present over the experiment time course. Representative images of histological analysis of heart tissue over the experiment time course. Bacterial and leukocyte infiltration was first noticed at 12 h and 24 h postinfection (black arrows) (b). Cardiac vascular proteins change very early during infection. p.i, postinfection; n.s., not significant (c). Hyaluronan synthase 1 (HAS1) was significantly increased in the vascular surface of all four organs (d).

Acute phase proteins (APPs) are inflammation markers that exhibit rapid and substantial changes in serum concentration during inflammation. Most APPs are synthesized in the liver as a response to both local and systemic inflammation, and some of them get associated with the endothelium, where they exert multiple functions. We used COR-SPOTS to map the dynamics of APPs during infection across organs and compartments. Although the dynamics of the APPs in plasma followed a linear increase, as has been previously described, when mapping the levels of APPs associated with parenchymal and vascular compartments, we noticed at least two distinct patterns. Type I APPs experienced rapid induction in all organs and compartments, and steadily

accumulated over time. This pattern was characteristic of APPs such as SAA1, SAA2, HPT and LBP (Fig. 1.12a). However, we also identified a second group or Type II APPs, whose levels were compartmentalized and sensitive to the disease status of the organs. For example, A1AG2 and SAMP levels in liver parenchyma and vascular fractions followed the same kinetics as in plasma (Fig. 1.12a). Interestingly, their levels in cardiac parenchyma and vasculature remained unchanged at 6h, followed by a linear increase at 12h and 24h. Other APPs such as SAA3 followed a similar trend, except for a temporal difference between the cardiac vascular proteome response (up at 12h) and the cardiac parenchymal response (up at 24h) (Fig. 1.12a). CD14 was also increased in the liver and in the heart vasculature by 6h, but increased cardiac parenchymal levels were only detected by 24h (Fig. 1.12a). Plasma levels of hyaluronic acid (HA), megadalton acidic polysaccharide involved recruitment of activated neutrophils to the liver during inflammatory responses (McDonald et al., 2008), were elevated more than 6-fold in MRSA-infected samples compared to controls at 24 h post-infection. Notably, PRG4, a lubricating and anti-inflammatory protein which binds HA, displayed the largest induction of all detected liver proteins, in all liver replicates, and across all independent experiments at 24hr postinfection. PRG4 is abundantly expressed in synovial fluid and its glycosylation plays an important role in the lubrication of articular joints (Das et al., 2019). Although extra-articular expression of PRG4 is known to occur, its physiological role or involvement in liver disease has never been reported. Looking at the dynamics of PRG4 across organs and compartments, it became apparent that vascular accumulation starts early in the liver and increases linearly, but association with the heart vasculature could not be detected until the late time points, while PRG4 remained undetected in the heart parenchyma at all investigated time points (Fig. 1.12a). Finally, we also observed that in contrast to the increasing levels of many vascular adhesion molecules and APPs, the most dramatic

downregulation taken place in the cardiac vasculature at 6h was the shedding of many immunoglobulins (IgGs), probably loosely associated with the vascular wall, without changing IgG total plasma levels (Fig. 1.12b). All these changes are consistent with a massive remodeling of the vascular proteome during sepsis, even before bacteria or infiltrating immune cells are observed. These changes are expected to significantly alter the physicochemical properties of the vascular surfaces, and their organotypic responses to systemic infection.

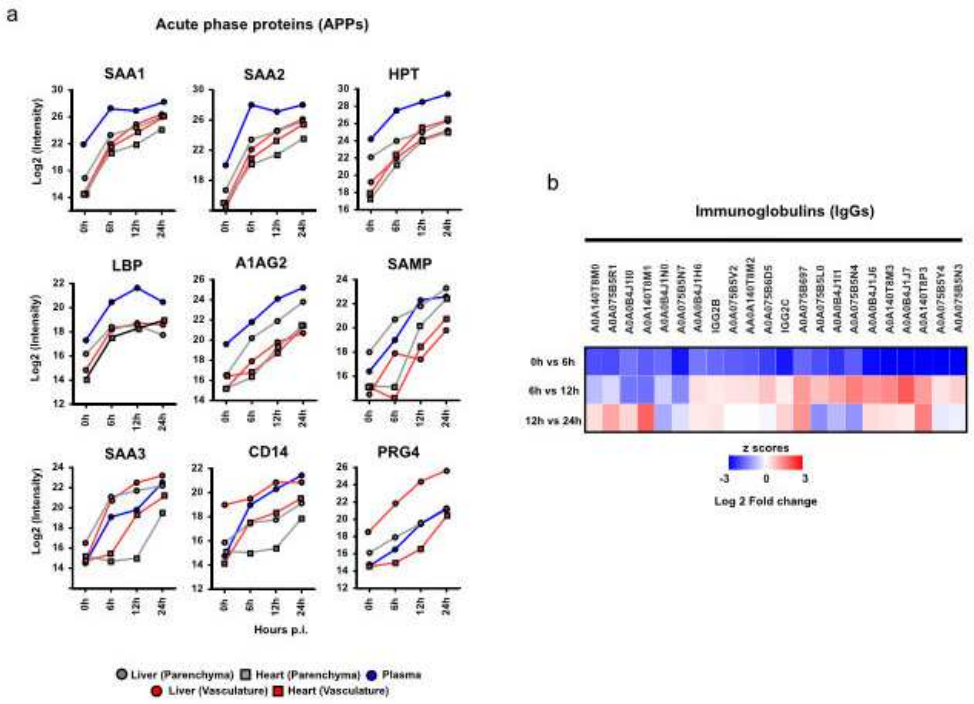


Figure 1. 12. Global versus compartmentalized changes of acute-phase proteins and immunoglobulins during MRSA sepsis

Levels of acute phase proteins (APPs) in liver, heart, and plasma. Liver measurements are represented by filled circles; heart measurements are represented by filled squares. Tissue was represented by a gray fill color, and vascular measurements by orange fill color. Blue filled circles for plasma measurements (a). A total of 23 immunoglobulin proteins were found to dramatically decrease in abundance as early as 6 h postinfection. Log2-fold change values are compared with the previous time point (b).

1.5 Discussion

The blood-tissue interface remains as one of the unique body compartments which offer a direct measurable landscape for the initial interactions during a systemic infection and subsequent response. In our reports we have built upon current techniques of perfusion-based tagging assays to streamline a large study of the blood tissue interface during broad MRSA dissemination in a murine model. This fully automated *in vivo* chemical labeling and proteomics workflow enabled us to capture and dissect changes in the blood-tissue interface at the resolution of the vascular glycocalyx. Because this murine model of infection rapidly dissolves into a septic shock, we were able to hypothesize mechanisms by which organ vascular surfaces are primed to deal with bacteremia and which failures lead to necrosis of the organ tissue.

Proteomics profiling of multiple healthy murine organs confirmed the presence of a complex proteome landscape embedded in the vascular cell surface and glycocalyx which had both shared and organ specific compositions. The protein families accessible to the vascular flow were largely comprised of adhesion molecules (including many classical endothelial markers), receptor tyrosine kinases, enzymes, scavenger receptors, proteoglycans and plasma proteins. Profiling molecular changes in the vascular compartment can aid in the search of novel markers in body fluids by defining molecular targets undergoing remodeling at the vascular wall. Having this pre-knowledge of what kind of candidate molecules to look for might facilitate the application of targeted proteomics assays to the plasma proteome, to increase sensitivity and coverage.

Understanding the relationships between vascular failure and organ damage has important implications for sepsis therapeutics and diagnostics. Even though many patients quickly respond to standard care, an unacceptably high fraction of cases still progress into intractable shock responses with a high risk for long term morbidity and deadly outcomes. Vascular stabilization as

a therapeutic goal in sepsis has been discussed for years, although few preclinical studies have been conducted to assess the impact of vasculoprotective approaches on relevant sepsis outcomes, and so far, none of them have succeeded in making a translational jump to application. Part of the problem lies on the enormous heterogeneity of the organ vasculatures, and a lack of detailed knowledge on the molecular composition of the vascular surfaces and how they are affected by sepsis.

Here, we demonstrate that sepsis triggers a series of well-coordinated proteome changes across the blood-tissue interfaces of the organs, but the nature of those changes and their specific temporal patterns are largely organotypic. These organ-specific patterns seem to be partially linked to the occurrence of major sepsis "checkpoints", such as systemic vascular activation or the actual invasion of the organs by pathogens and infiltrating immune cells. Interestingly, the liver undergoes a rapid reprogramming of its vascular proteome by downregulating lipid metabolic functions and becoming primed towards a prothrombotic and anti-fibrinolytic state, long before signs of organ dysfunction and actual thrombosis are observable. This finding is remarkable, firstly because it suggests that vascular proteome alterations can have a yet unrealized potential as a source of early diagnostic and prognostic markers of sepsis. Secondly, it also opens the possibility that timely therapeutic interventions targeting the vasculature might prevent irreparable organ damage. For example, we observed that hepatic thrombosis is preceded by early depletion of vascular TFPI, upregulation of VWF and intravascular platelet deposition. Alterations in the extrinsic pathway of coagulation were also reported in baboon models of gram-negative sepsis, and exogenous TFPI administration rescues the organs from coagulation-induced damage. However, there can be difficulties in translating some of the findings from preclinical models to therapy (Tang et al., 2007, Taylor et al., 2001). For example, results from clinical trials to evaluate

the efficacy and safety of Tifacogin, (recombinant TFPI) to treat sepsis have so far been disappointing, emphasizing the complexity of the host response to systemic infection.

Many vascular proteins were found changing in the hepatic vasculature during infection. One of them in particular, PRG4, was consistently displaying large fold changes and dominating the global pattern of proteome changes in the liver post 24hr after infection. PRG4 is a heavily modified glycoprotein clearly deposited onto the superficial zone of the articular joints, acting as a lubricating gel to reduce friction between the cartilages in the synovial cavities. Due to the large amount of negatively charged carbohydrates decorating its mucin-domain, PRG4 can impact the adhesion properties of multiple synovial cell populations, bacteria and immune cells. Interestingly, PRG4 is also expressed in the liver but a role in basic liver physiology or pathology has remained elusive. We have shown that PRG4 is rapidly induced during MRSA-bacteremia, which results in its secretion and vascular deposition around the major blood vessels and within the sinusoidal microvasculature of the liver. Notably, in previous studies we have seen hepatic mRNA expression of PRG4 is induced 12hr post infection and PRG4 staining was also detected in association with Ly6G + neutrophil populations swarming the edges of large necrotizing thrombi (Supplemental Fig. 1.1). In this context, it is conceivable that vascular deposition of PRG4 during liver inflammation might affect the adhesion properties of the glycocalyx and influence how immune cells such as neutrophils interact with the activated endothelium. In fact, human peripheral polymorphonuclear granulocytes (PMNs) have been shown to circulate carrying a PRG4 coating that is partially dependent on L-selectin binding. Activation with TNF-α leads to L-selectin shedding from the surface of the PMNs and to reduced binding of PRG4. Since PRG4 can also regulate NF-κB signaling by engaging TLR2 and TLR4 in vitro, it may also exert an additional anti-inflammatory role during liver infection *in vivo* (Al-Sharif et al., 2015, Das et al., 2019). A

mechanistic understanding of the role of PRG4 during acute hepatic injury, and especially during MRSA sepsis, warrants further research.

The host response to sepsis is overwhelmingly complex since it encompasses crosstalk between multiple systems, and a time-dependent staging, where both pro- and anti-inflammatory components might be operating at different time points. Predicting the disease trajectory of sepsis is therefore a most sought-after clinical goal. Here we show that the host proteome response to sepsis is a powerful molecular readout of the septic response. Using a novel networking approach, COR-SPOTS, we were able to identify very early vascular changes that even precede bacterial and leukocyte infiltration of the tissue. Some of these responses have been previously reported, such as the upregulation of vascular adhesion molecules and glycocalyx components, whereas other were completely unexpected, such as the compartmentalization of the acute-phase reaction and the massive shedding of vasculature-associated IgGs. Future work should explore how these early vascular events are coordinated, for example by focusing on transcriptional and/or posttranslational mechanisms of proteome regulation.

Finally, the accumulating data suggests that sepsis entails a series of heterogeneous responses to systemic infection, with distinct molecular trajectories and disparate clinical outcomes. The integration of this data across compartments through COR-SPOTS, has allowed us to identify markers that are informative of processes happening in other tissues or compartments in *S. aureus* sepsis. However, this strategy can be exploited in numerous cases. For example, identification of vascular markers leaking out to plasma could potentially be measured using more sensitive and targeted approaches directly on clinical material. Additionally, construction of concise molecular readouts of the host response during systemic inflammation can be used to

compare the impact of individual pathogens and/or virulence factors to the molecular heterogeneity of sepsis.

1.6 Acknowledgements

This research was supported by grant P01HL131474 to J.D.E. and J.W.S. N.L. and J.S. We also acknowledge support from NIGMS R35 GM119850 and grant T32GM008806 to J.T.S.. J.W.S. was supported by R01 GM107523. G.G. was supported by Microbial Sciences Graduate Research Fellowship Awards 1-F17GG and 1-F18GG. J.M. would like to acknowledge support from grant 2017.0271 from the Wallenberg Foundation.

1.7 Author Information

J.T.S., J.D.E., and A.G.T. conceived the project. G.J.G., C.M., C.D.P., A.R.C., C.K., H.C., and A.G.T. conducted the mouse infection, perfusion, organ isolation, and mass spectrometry experiments. C.K. constructed the in-silico library for DIA-SWATH-MS and analyzed the DIA data. J.T.S. developed COR-SPOTS. J.T.S., V.N., A.R.C., J.W.S., J.M., N.E.L., J.D.E., H.C., N.V., and A.G.T. interpreted the data. J.T.S. and A.G.T. prepared the figures. J.T.S., J.D.E., and A.G.T. wrote the manuscript with significant input from all authors.

Chapter 1, in full, is an amalgamation of the published material as it appears in "Proteomic atlas of organ vasculopathies triggered by *Staphylococcus aureus* sepsis" in *Nature Communications* 2019 and "Vascular Proteome Responses Precede Organ Dysfunction in a Murine Model of *Staphylococcus aureus* Bacteremia" in *Msystems* 2022. The dissertation author was the primary investigator and/or author on these papers.

1.8 Supplemental Materials

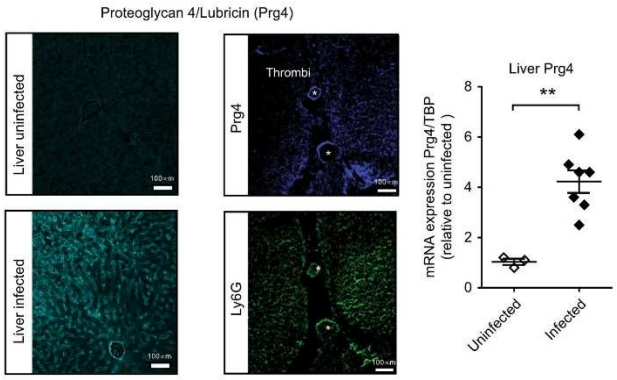


Figure 1.S 1. MRSA infection increases expression and deposition of Prg4 along the central veins and the sinusoidal microvasculature

Prg4-immunoreactivity was found along the central veins and the sinusoidal microvasculature and at the edges of large necrotic thrombi in association with Ly6G + neutrophils. qPCR analysis in a separate cohort of mice (infected livers n = 7, uninfected controls n = 3) demonstrated increased expression of hepatic Prg4-mRNA levels already at 12 h post-infection. Data are represented as mean ± SD. ***$p < 0.001$, **$p < 0.01$ and *$p < 0.05$ by two-sided Student's t-test. Scale bar, 100µm.

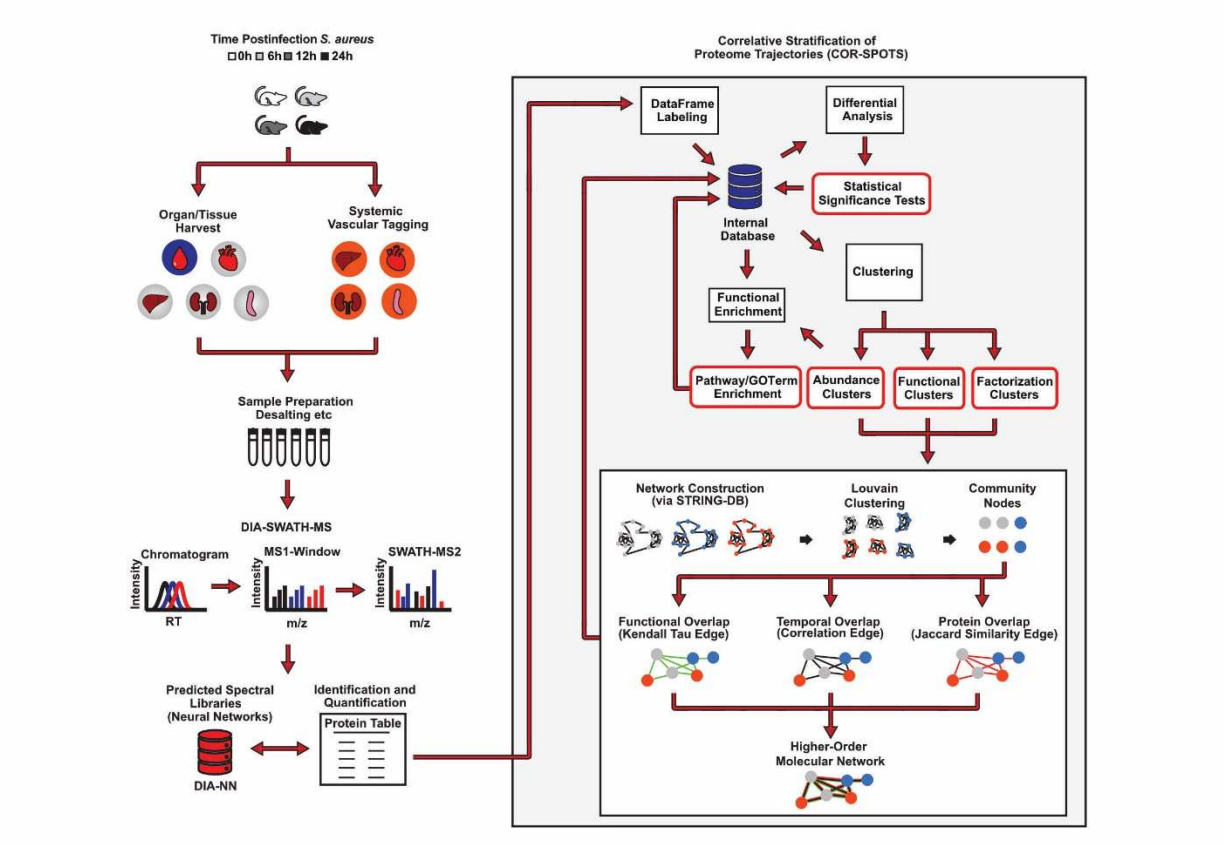


Figure 1.S 2. Experimental design and analytical workflow employed for COR-SPOTS.

Chapter 2: Evaluation of antibody-based therapies in various murine models of group A streptococcal infection

2.1 Abstract

The blood-tissue interface holds many significant molecular cross communications between organs and tissues. Several plasma proteins undergo glycosylation changes that are hypothesized to be impactful to binding capacity and signal transduction. Endogenous antibodies are glycosylated serum proteins that are responsible for the identification and neutralization of pathogenic bacteria and viruses. Elucidating the role of immunoglobulin glycosylation in the build-up of a protective immune response against bacterial pathogens is thus of critical importance to the successful development of new antimicrobial strategies. Our collaborators have previously found that pathogen-driven immunoglobulin G (IgG) degradation is a common feature of group A streptococcus (GAS) infection. Interestingly, this is more pronounced in throat and skin infections compared to systemic blood infections. The choice of a suitable preclinical model to test the efficacy of both passive and active immunization therapies remains an open question in GAS infections. We demonstrate that various models of GAS infection have significantly altered IgG degradation. GAS preclinical models are impacted by the route of bacterial inoculation and the local vs systemic character of the inoculation mode. We utilized mass spectrometry-based capture of site-specific glycosylation and novel measures of glycan heterogeneity to show that GAS virulence factors directly impact human immunoglobulins (IVIGs) in a murine model of sepsis. Understanding the role and how to control glycosylation in both endogenous and exogenously administered IgGs in establishment of immunological memory will inform vaccination regimes and facilitate tuning of passive immunization therapies for GAS infections.

2.2 Introduction

Streptococcal related diseases range from mild skin infections and uncomplicated tonsillitis to life-threatening invasive diseases, such as necrotizing fasciitis and sepsis (Bessen et al., 1996, Bessen et al., 2015). Group A streptococcus (GAS) is an obligate human pathogen that is responsible for more than half a million deaths worldwide per year (Carapetis et al., 2005). GAS infections display significant heterogeneity regarding tissue tropism, disease severity, and occurrence of post-infectious sequelae (Cole et al., 2011, Henningham, et al., 2012, Walker et al., 2014). The variable disease course paired with a high incidence rate have stimulated the development of murine models of GAS infection. Interestingly, GAS infections have several known mechanisms to avoid host immune response. GAS surface proteins can bind to the IgG FC-region and prevent phagocytosis via FCγ-receptor expressing immune cells (Happonen et al., 2019). This "non-immune" IgG binding by the bacteria is dependent on the local IgG concentration suggesting a variable range of outcomes from GAS mediated adaptive immune function dysregulation (Nordenfelt et al., 2012). More directly, GAS is able to counteract adaptive immunity by specifically destroying IgGs. GAS secretes the protease called the IgG-degrading enzyme of *S. pyogenes* (IdeS). This protease specifically cleaves IgGs at the hinge region separating the antigen-binding Fabs from the effector function–promoting FC-region interfering with phagocytic killing of GAS bacteria (Von Pawel-Rammingen et al., 2002). Additionally, GAS releases another protease, SpeB, which degrades antibodies (as well as other protein targets) at multiple sites, reducing the capacity of IgG to elicit downstream FC-dependent effector functions (Collin & Olsen, 2001). Another major tool used by the GAS bacterium is the secretion of endoglycosidases (EndoS) to extensively degrade FC N-glycan from IgGs and severely impact FC-mediated protective functions (Allhorn et al., 2008, Collin & Olsen, 2001). However, some

factors involved in human infections do not really contribute to disease in the murine setting. This has given rise to an extensive literature of effective preclinical models with differences that range from subtle to substantial in order to measure various GAS infection properties (Watson et al., 2016). Properly immunizing mice to have long-lasting protection from severe GAS infection is further complicated by route of immunization and bacterial inoculation. Unfortunately, we have seen that IgG glycosylation is severely missed as a topic of murine model attributes to consider when studying GAS infection. We show that degradation of specific murine IgG-subtypes, and severe FC-deglycosylation of both endogenous and exogenously administered IgGs is overly abundant in murine models of GAS infection.

Most animal models of blood infections do not recapitulate all the aspects of the disease in humans, where bacterial colonization starts locally and slowly disseminates into the bloodstream. Thus subcutaneous inoculation can be a tool to evaluate a synonymous dysregulated systemic inflammatory response. However, the goal of an active immunization study may require assessment of initial immune response and protective efficacy of circulating stimulated antibodies. For this instance an intraperitoneal (IP) infection offers direct introduction of bacteria to the blood-tissue interface. Given the benefits of each route of bacterial inoculation, it is crucial to identify how FC-region glycan modulation is affected by the activation or repression of virulence factors due to mode of infection. Despite some of the obvious shortcomings of using mouse models of infection, they still constitute invaluable tools to study GAS pathogenesis, and to test the ability of candidate antigens to elicit an immune response.

While the antigen binding region remains crucially important to the neutralizing capacity of protective antibodies against GAS, the FC-region of the antibodies is required for activation of FC-receptors in various immune cell populations. In all IgGs, the FC-region is decorated with a

single N-glycan chain on the CH2-domain. This glycosylation has been implicated in the confirmation of the FC-region and plays an important role in the affinity of FC-domains to various FC-receptor subtypes (Ferrara et al., 2011, Kao et al., 2017). Removal of the glycan significantly diminishes all FC-dependent effector functions, both *in vitro* and *in vivo*, and composition of the glycan can modulate binding and cellular activation. For example, lack of N-glycan core fucosylation increases the affinity of IgG1 towards FCγRIIIa by ~ 50 fold, which correlates with a significant enhancement in antibody-dependent cellular cytotoxicity (ADCC) in natural killer (NK) cells (Ferrara et al., 2011). Additionally, it has been shown that spontaneous control of HIV is associated with decreased FC-galactosylation, fucosylation, and sialylation and in turn a significant enhancement of FCγRIIIa binding and more effective control of viral replication (Ackerman et al., 2013). During microbial infections, pathogens can express virulent FC-binding factors to circumvent the host immunity during infection. Some pathogens even utilize enzymatic removal of the glycan chain to completely abrogate all FC-dependent effector functions. Therefore, establishment of a protective immune response against pathogens may be enhanced by outlining the importance of IgG glycosylation and IgG subtype specific attributes.

Intravenous immunoglobulins (IVIGs) infusion is a passive immunization strategy used to treat autoimmune disorders such as systemic lupus erythematosus (SLE) and immunothrombocytopenia (ITP); IVIGs are even used prevent infections in immunocompromised (Almizraq et al., 2021, Castro & Dorfmueller, 2021, Kaul et al., 1999, Majer & Hyde, 1988, Sriskandan et al., 2006, Zandman-Goddard et al., 2005). However, treatment of sepsis using IVIG infusions have been wildly inconclusive. This variable outcome is impart due to the complexity of the current IVIG formulations, the lack of understanding of their mechanisms of action, the use of different dosages across studies, and poor patient selection. A notable exception is the treatment

of streptococcal toxic-shock syndrome (STSS) arising from severe group A streptococcus (GAS) infections. IVIG administration in STSS patients has been associated with significantly higher 30-day survival rates and lower Acute Physiology and Chronic Health Evaluation (APACHE) II scores compared to untreated patients. IVIG therapy enhances neutralization of bacterial mitogenicity and reduces T cell production of interleukin-6 and tumor necrosis factor alpha (Castro & Dorfmueller, 2021, Kaul et al., 1999, Sriskandan et al., 2006). Interestingly, current successful IVIG formulations consist of heterogeneous IgG pools derived from thousands of donors. It is hypothesized that the therapeutic effects of these IVIG formulations are exerted through multiple FAB and FC-mediated mechanisms. It has been shown that increasing the content of sialylated FC-glycoforms through affinity pulldowns or chemoenzymatic alterations results in a ~10-fold enhancement in IVIG anti-inflammatory activities in both arthritis and ITP murine models (Pagan et al., 2018). Alternatively, recent active immunization and immunological priming studies have shown that specific IgG glyco-patterns can be evoked in an antigen-specific way (Washburn et al., 2015). Taken together, IgG glycosylation is a relevant candidate for novel immunotherapeutic approaches.

The plasma glycoproteome offers a suitable readout to monitor systemic alterations in the body. Unfortunately, the analytical challenge of separating causal signals in the heterogeneous pool of carbohydrate and protein components has severely limited advancements in this area. Focus on the immunoglobulin family allows for the identification of discrete patterns of glycoproteome remodeling during GAS infection. This pathogen-driven proteolytic and glycan degradation of both endogenous (murine) and exogenously administered (IVIGs) antibodies is important in preclinical studies of GAS and other microbial pathogens. Interference with the evaluation of antibody-based therapies in murine models of GAS infection could undermine the

protective FC-mediated effector functions of the therapies. Here we provide a complete overview of IgG degradation in experimental models of group A streptococcal bacteremia for use as a reference in preclinical studies.

2.3 Materials and Methods

Here multiple experimental models of *S. pyogenes* are tested to gather novel insight on the pathogenic destruction of Immunoglobulin G (IgG). Additionally, it has been crucial to consider glycosylation in the assessment of antibody treatments for *in vivo* models of *S. pyogenes* bacteremia. By targeting unique peptides and associated glycosylation sites this methodology presents a systematic evaluation of evasion mechanisms tied to the IdeS/EndoS activation. Various stages/models of *S. pyogenes* infection exhibit important unique properties of pathogenesis. Glycosylation site heterogeneity analysis offers new insights into population-based evaluation of post-translational modifications.

Bacteria and culture conditions. *Streptococcus pyogenes* AP1 (from the Collection of the World Health Organization Collaborating Center for Reference and Research on Streptococci, Prague, Czech Republic) were grown in Todd–Hewitt broth, supplemented with 0.2 % yeast extract (THY, BD diagnostic), overnight (o/n) at 37°C and 5% CO2. Methicillin-resistant *Staphylococcus aureus* (MRSA USA300 TCH1516) was grown as previously reported (Sorrentino et al., 2022).

Bacterial infections. All animal use and procedures were approved by the local Malmö/Lund Institutional Animal Care and Use Committee, ethical permit number 03681-2019. *S. pyogenes* AP1 was grown to logarithmic phase in THY medium (37°C, 5% CO2). Bacteria were washed and resuspended in sterile PBS. Nine-week-old female C57BL/6J mice (Janvier, Le Genest-Saint-Isle, France) were infected with 50 µl (2x105 cfu) bacteria by subcutaneous injection

on the right flank, or 100 µl (1x107 cfu) by intraperitoneal injection. Control groups were similarly injected with sterile saline (n= 9). Mice were rehydrated subcutaneously with saline 24 h after infection. Body weight and general symptoms of infection were monitored regularly. Mice were sacrificed and multiple organs were collected. *S. aureus* intravenous infection was conducted as previously reported in Toledo et al., 2019 (n=5/time point). Blood was taken by cardiac puncture and collected in tubes containing sodium citrate (MiniCollect tube, Greiner Bio-One).

Intravenous immunoglobulin (IVIG) treatments. IVIG (Octagam, 1mg/ g of body weight) was given by intraperitoneal administration 6h prior to subcutaneous (n=10) or intraperitoneal (n=10) infection. Control groups were similarly injected with sterile saline (n= 10). Body weight and general symptoms of infection were monitored regularly. Mice were sacrificed 24h post infection, and multiple organs were harvested. Blood was taken by cardiac puncture and collected in tubes containing sodium citrate (MiniCollect tube, Greiner Bio-One).

Plasma and organ preparations. Citrated blood collected from infected and control mice was centrifuged (2000 rcf, 10 min) to obtain platelet free plasma. Plasma was aliquoted and stored at -80°C. Collected organs were homogenized (MagnaLyzer, Roche) in PBS using sterile silica beads (1 mm diameter, Techtum). Skin samples were obtained by punch biopsy (10 mm, Acu punch). Degree of bacterial dissemination was determined by serial dilution and plating organ homogenates onto blood agar plates. Colony forming units were counted following o/n incubation (37°C, 5% CO2), and are presented as cfu/g of tissue. Remaining organ homogenates were centrifuged (20000 rcf, 10 min) and supernatants immediately transferred and aliquoted into fresh vials. All samples were stored at −80 °C until further analysis. Protein concentration was determined using standard BCA assay (Thermo Scientific) according to manufacturer instructions.

Patient samples. Patient swab samples from those who were diagnosed with GAS tonsillitis were obtained from patients (>8 yr old) at Laurentiikliniken and Skåne University Hospital (SUS), both in Lund, Sweden. A follow-up tonsillar swab sample (3–5 d later) was taken from a subset of patients (n = 5) treated with antibiotics. Wound swabs from the local infection site of patients clinically diagnosed with GAS sepsis and necrotizing fasciitis were obtained from SUS during surgical intensive care. Full sample collection, handling, and ethical protocols are outlined in previous publication (Naegeli et al., 2019).

Flow cytometry of blood cells. Citrated blood collected from infected and control mice was diluted with HEPES buffer containing Mouse BD Fc-block (BD Pharmingen). Total leucocytes were gated according to characteristic forward and side scatter and platelets were identified using anti-CD41 FITC (BD Pharmingen). Antibodies were diluted 1:200 and incubated (15 min, RT). Samples were lysed using 1-step Fix/Lyse Solution (e-Bioscience), washed (500 rcf, 5 min) cellular pellets were resuspended in PBS. The samples were analyzed using an Accuri Plus C6 Flow Cytometer (BD Biosciences), and the data was analyzed using C6 Software (BD Biosciences).

IgG western blots. Samples were normalized by protein concentration, determined using standard BCA assay, and separated on SDS-PAGE (Mini-protean TGX stain-free gels, 4-15% acrylamide, BioRad) under reducing conditions. Proteins were transferred to PVDF membranes using the Trans-Blot Turbo kit (Bio-Rad) according to the manufacturer's instructions. Membranes were blocked with 5% (wt/vol) skimmed milk in PBST (PBS, 0,1% Tween 20) followed by incubation with rabbit anti-mouse IgG (1:2500, Bio-rad). The membranes were washed, followed by incubation with a secondary antibody (goat anti-rabbit HRP-conjugated antibody; Bio-Rad).

The membranes were developed using Clarity Western ECL substrate (Bio-Rad) and visualized with the ChemiDoc MP Imager.

IgG pulldowns. IgGs were purified in a 96 well plate setup using the Protein G AssayMAP Bravo (Agilent) technology, according to the manufacturer's instructions. Briefly, 10µl plasma was diluted with PBS to a final volume of 100 µl and applied to pre-equilibrated Protein G columns. Columns were washed with PBS and eluted in 0.1M glycine (pH2). The final pH was neutralized with 1M Tris, and saved until further use.

Trypsin digestion and peptide desalting. Protein samples were resuspended in 8M urea and reduced with 5 mM Tris(2-carboxyethyl)phosphine hydrochloride, pH 7.0 for 45 min at 37 °C, and alkylated with 25 mM iodoacetamide (Sigma) for 30 min at RT, followed by dilution with 100 mM ammonium bicarbonate to a final urea concentration below 1.5 M. Proteins were digested by incubation with trypsin (1/100, w/w, Sequencing Grade Modified Trypsin, Porcine; Promega) for at least 9 h at 37 °C. Digestion was stopped using 10% trifluoracetic acid (Sigma) to pH 2 to 3. Peptide clean-up was performed by C18 reversed-phase spin columns according to manufacturer instructions (Silica C18 300 Å Columns; Harvard Apparatus). Solvents were removed using a vacuum concentrator (Genevac, miVac) and samples were resuspended in 50 µl HPLC-water (Fisher Chemical) with 2% acetonitrile and 0.2% formic acid (Sigma).

LC-MS/MS for proteomics analysis. Peptide analyses (corresponding to 1 µg protein) were performed on a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific) connected to an EASY-nLC 1200 ultra-HPLC system (Thermo Fisher Scientific). Peptides were trapped on precolumn (PepMap100 C18 3 µm; 75 µm × 2 cm; Thermo Fisher Scientific) and separated on an EASY-Spray column (ES903, column temperature 45 °C; Thermo Fisher Scientific). Equilibrations of columns and sample loading were performed per manufacturer's guidelines.

Mobile phases of solvent A (0.1% formic acid), and solvent B (0.1% formic acid, 80% acetonitrile) were used to run a linear gradient from 5% to 38% over 90 min at a flow rate of 350 nl/min. The variable window data independent acquisition (DIA) method is described by Bruderer et al. (Bruderer et al., 2017). The data dependent acquisition (DDA) method was the manufacturer's default for 'high sample amount'. LC-MS performance was quality controlled with yeast protein extract digest (Promega). MS raw data was stored and managed by openBIS (v20.10.0) and converted to centroid indexed mzMLs with ThermoRawFileParser (v1.2.1).

IgG glycoproteomics analysis. Purified IgG glycopeptides were analyzed on a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific) connected to an EASY-nLC 1200 ultra-HPLC system (Thermo Fisher Scientific). Peptides were trapped on a precolumn (PepMap100 C18 3 μm; 75 μm × 2 cm; Thermo Fisher Scientific) and separated on an EASY-Spray column (Thermo Fisher Scientific). Mobile phases of solvent A (0.1% formic acid), and solvent B (0.1% formic acid, 80% acetonitrile) were used to run a linear gradient from 4 to 45% over 60 min. MS scans were acquired in data-dependent mode with the following settings, 60,000 resolution @ m/z 400, scan range m/z 600-1800, maximum injection time of 200 ms, stepped normalized collision energy (SNCE) of 15 and 35%, isolation window of 3.0 m/z, data-dependent HCD-MS/MS was performed for the ten most intense precursor ions.

Proteomics data analysis. DIA MS files were processed using DIA-NN, version 1.8 and searched against the reference proteome (EMBL-EBI RELEASE 2019_04) EMBL mouse using DIA-NNs library-free mode using default parameters. Data was normalized using the cyclic loess normalization method in the NormalyzerDE tool (Willforss et al., 2019). All statistical methods were implemented in Python 3.6.10. The proteomics results were filtered using a one-way analysis of variance (ANOVA) followed by a Benjamini-Hochberg correction to control for a false

discovery rate (FDR) of 0.10. Statistically significant identifications were subjected to principal-component analysis (PCA). Protein measurements were also standardized using a Z-score normalization. Proteomics results were analyzed separately using Welch's t test to generate volcano plots and heatmaps. Functional enrichment analysis of differentially abundant proteins was performed through Metascape (Zhou et al., 2019). Visualization of enriched Genome Ontology (GO) terms was produced using code adapted for TreeMap visualizations from the Web tool Revigo (Supek et al., 2011).

Glycoproteomics data analysis. Raw files were searched in the Byonic software integrated as a node in Proteome Discoverer. Files were searched against a UniProt human and mouse protein database using the default search strategy: enzyme: trypsin with a maximum of two missed tryptic cleavages per peptide, up to one glycan per peptide as a 'rare' variable modification, up to 10/20 ppm deviation of the observed precursor/product ion masses from the expected values, up to one Met oxidation (+15.994 Da) per peptide (variable 'common' modification), together with a predefined plasma glycan database of mouse and human structures and a decoy and contaminant database available in Byonic. Identifications with a Byonic score >100 were considered positive hits, and their spectra were manually validated. Total glycopeptide counts and unique glycopeptide counts were harvested from IgG enriched and depleted samples. Semi-tryptic glycopeptide variants were collapsed into consensus glycopeptides and glycoforms were represented as a percent of total glycoform intensity for each glycosylation site. Heatmaps represent each glycosylation distribution present at a unique glycosylation site.

Glycan heterogeneity analysis. Simpson's diversity index was adapted to quantify heterogeneity of each unique glycosylation site. This Heterogeneity Index (H-index) is visualized by a heatmap with values ranging from 0.0 to 1.0 with 1.0 being equal frequency for any glycoform

that appears on that glycosylation site. In the field of ecology, the Simpson's diversity index is measured by the equation $1 - \sum \left(\frac{n}{N}\right)^2$, with n being the total number of organisms of a particular species and N being the total number of organisms of all species. Here we have adapted these values to represent n being the number of spectral counts of a particular unique glycoform and N being the number of total spectral counts for all unique glycoforms at a glycosylation site.

2.4 Results

Tracking the abundance of peptides and glycopeptides using mass spectrometry can reveal infection induced glycoproteome remodeling at high resolution. In a previous study by our collaborators, the accumulation of deglycosylated IgG in the body fluids and tissues of GAS-infected patients was reported and partially the accumulation of deglycosylated IgG in the body fluids and tissues of GAS-infected patients (Naegeli et al., 2019). In this previous study, IgG deglycosylation was monitored through targeted selected reaction monitoring (SRM) mass spectrometry (MS), for high sensitivity. However, this came at the expense of losing track of potential changes occurring on other IgG isoforms. To generate a more comprehensive view of all ~30 different glycoforms that normally modify the 4-5 different IgG subtypes circulating in mouse plasma, we took advantage of the versatility of label-free glycoproteomics analysis, using data-dependent acquisition (DDA) mass spectrometry and stepped high-energy collisional dissociation (sHCD), which has been shown to effectively resolves site-specific IgG glycopeptide differences(Chandler et al., 2019). Statistical analysis of this data can drive identification compartment localized and/or protein targeted mechanisms of immune system avoidance by bacteria. Specifically, here we reinforce that the human pathogen *Streptococcus pyogenes* harbors immunomodulating virulence factors that directly work on circulation IgGs. However, we also show that this proteolytic and glycan degradation of IgGs should be considered a significant

variable in murine models of GAS infection. Interestingly this immunomodulation can be investigated by changing the experimental design of infection. We show that systemic GAS infection will result in degraded endogenous and exogenous IgGs under the presence of EndoS.

Proteolytic degradation of murine antibodies in a subcutaneous model of GAS infection. In order to slowly disseminate GAS bacteria to the bloodstream we elected to use a model of GAS skin infection which when uncontrolled develops into bacteremia. Mice were subcutaneously inoculated with the highly virulent M1 strain AP1, which is known to cause local skin infection that becomes fully systemic by 24 hours, and results in ~50% mortality by 36h p.i.. Proteomics analysis of murine plasma, liver, and kidney proteome over the course of infection: 0h, 12h, 24h, 36h post infection (p.i.) indicates that GAS bacteremia induces changes in the abundance of several hundred plasma and tissue proteins. The proteomics data revealed that disease progression was associated with significant changes in the levels of several proteins linked to coagulation and inflammation, and with alterations in the abundance of factors involved in cellular and humoral responses. Principal component analysis (PCA) of plasma proteome alterations stratifies the samples into two defined stages that correlate with the timing of bacterial dissemination from the skin to the organs (Fig. 2.1a). The early stage was driven by the activation of the acute phase response and other inflammatory proteome changes, whereas the late stage was characterized by the plasma accumulation of several markers of tissue damage, neutrophil proteins, and Proteoglycan 4 (PRG4), a protein that has recently been reported to accumulate both in the plasma and on the injured vasculature of a murine model of *Staphylococcal* (*S. aureus*) sepsis (Toledo et al., 2019, Sorrentino et al., 2022). Surprisingly, levels of murine IgG3 were reduced in all 3 proteomes by 36h p.i. (Fig. 2.1b). Quantification of individual IgG subtypes (IgG1, IgG2b, IgG2c and IgG3) in plasma and liver tissues shows that IgG depletion is mostly restricted to IgG3,

and the abundances of the other subtypes do not change significantly over the course of infection (Fig. 2.1c). Western blot analysis further confirms overall reduction in total plasma and organ associated IgGs over time in subcutaneous GAS infection. It is also noted that there is a decrease in the apparent molecular weight of the heavy chains in plasma, skin, kidney, and liver, indicating additional changes in the IgG structure (Supplemental Fig. 2.1a). This IgG3 phenotype seems specifically linked to GAS infection since *Staphylococcus aureus* bacteremia does not induce similar temporal changes in the murine IgG levels, neither in the plasma nor in the organs (Supplemental Fig. 2.1b).

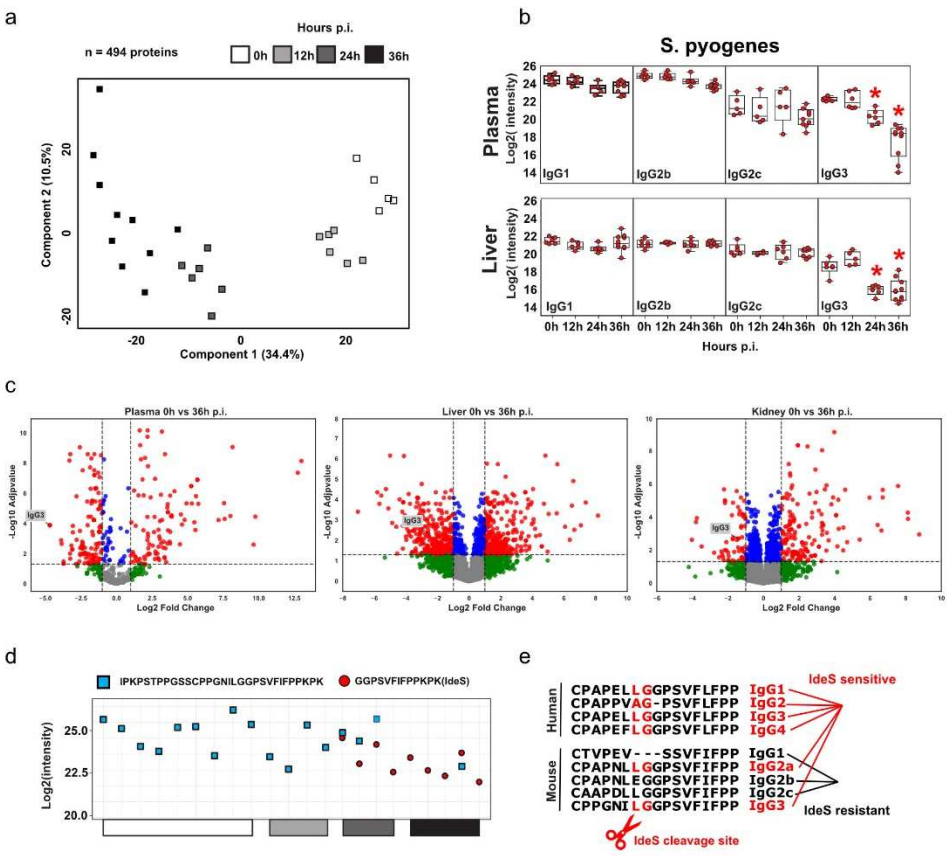


Figure 2. 1. IdeS mediates changes in murine IgG homeostasis in a model of local to systemic GAS infection

Principal component analysis of murine plasma proteins significantly altered during GAS infection (n = 5 mice/ time point). Statistical significance was assessed by analysis of variance (ANOVA) with a permutation-based false discovery rate (FDR) for multiple test correction (a). Mass spectrometric measurements of murine IgG subclasses in plasma and liver homogenates during the course of S. pyogenes (b). Volcano plots displaying the most significantly up- or downregulated plasma, liver, and kidney proteins at 36h p.i. compared to uninfected controls (c). Plasma levels of intact and IdeS-cleaved tryptic peptides of the hinge region of murine IgG3 at different time points during the infection (d). Sequence alignment of the IdeS-specific hinge regions across human and mouse IgG subclasses (e).

The known GAS virulence factor capable of degrading IgGs is a streptococcal cysteine protease known as IdeS. Interestingly this protease has a high specificity for targeting the IgG hinge region to cleave the FC from the FAB domain. This cleavage quickly ships IgG fragments for complete clearance from circulation by cytosolic and cellular machinery. Murine IgG3 is structurally and functionally different from the others. IgG3 has increased abilities for self-

oligomerization, targeting polyvalent antigens, and activating complement. In order to evaluate whether IdeS activity is linked to IgG3 depletion we measured the occurrence of intact tryptic peptides derived from the IgG3 hinge region in infected plasma. Hinge region peptides degraded by trypsin digest were found in uninfected as well as infected samples up to 12h p.i. However, hinge region peptide fragments with IdeS-specific proteolytic cleavages within their sequence began to appear and increase in abundance by 24h and 36h p.i.. This occurrence was reflected by the decrease in intact hinge region peptides, clearly suggesting that IdeS activity is indeed responsible for the reduction of IgG3 (Fig. 2.1d). Repeated efforts to measure IdeS directly in plasma turned unsuccessful, indicating that it circulates at very low levels. All human IgGs are IdeS sensitive, which emphasizes species differences between human and murine IgGs, and their vulnerability to GAS-specific virulence factors (Fig. 2.1e). This should be kept in mind for the subsequent studies of exogenously administered antibodies in murine models of GAS infection.

Glycan degradation of murine antibodies in a subcutaneous model of GAS infection. As previously mentioned, western blot analysis of murine IgGs revealed a ~2kDa reduction in molecular weight, which approximately matches the mass of the expected N-linked glycan in the IgG FR-region. In order to determine if GAS triggers deglycosylation of murine IgGs, collection of IgGs from uninfected and infected blood at various time points (12h, 24h, 36h) post-infection were collected for mass spectrometry analysis. Plasma IgG was purified on protein G columns using an automated liquid handling platform, digested into peptides, and subjected to glycoproteomics analysis. A glycoproteomics workflow to quantify glycopeptides was implemented. This allowed for the measure of IgG N-linked glycan distribution in a subtype specific manner. By tracking the quantified measure of glycopeptides derived from murine IgGs, we have seen that rapid deglycosylation occurs across all murine IgGs. The truncated products

reveal themselves as glycopeptide fragments from the FC-region containing a single N-acetyl-glucosamine (GlcNAc) linked to the asparagine residue, with or without a core fucose (Fuc) (Fig. 2.2a). Initially, murine IgGs in uninfected plasma are seen to be modified with a few variants of complex-type biantennary N-linked glycans, ranging in degrees of fucosylation, galactosylation, and sialylation. As soon as 12h p.i. The abundance of these extended structures begins to decline across all IgG subtypes, with a reflection of the truncated products increasing. Over the course of infection, the relative abundance of all intact IgG glycoforms was decreased (IgG1: 45.4% reduction, IgG2b: 41.1% reduction, IgG2c: 49.0% reduction, IgG3: 27.1% reduction) across all subtypes already by 12h p.i. This reduction became more pronounced by 24h p.i., and paralleled by the accumulation (average of 10.5 log2 Fold Change) of the characteristic EndoS degradation products (HexNAc1, HexNAc1Fuc1), resulting in an almost complete (IgG1: 77.5% reduction, IgG2b: 94.7% reduction, IgG2c: 90.8% reduction, IgG3: 97.0% reduction%) deglycosylation of the entire murine IgG pool by 36h p.i. (Fig. 2.2b). Traditional compositional analysis of major N-linked glycan structures (fucosylation, galactosylation, sialylation, and amount of bisecting GlcNAc) reveals that diversity of the murine IgG glycan pool significantly decreases in complexity (Fig. 2.2b). Implementing a measure of species diversity, the Simpson diversity index, we can see that the IgG glycoform heterogeneity is severely decreased by 36h p.i. (Fig. 2.2c). Most GAS strains are known to express EndoS, a bacterial endoglycosidase that specifically binds to IgGs and removes the FC-region oligosaccharide chain (Collin & Olsen, 2001, Naegeli et al., 2019). However, similarly to IdeS, multiple attempts to measure EndoS in infected mouse plasma by mass spectrometry were also unsuccessful.

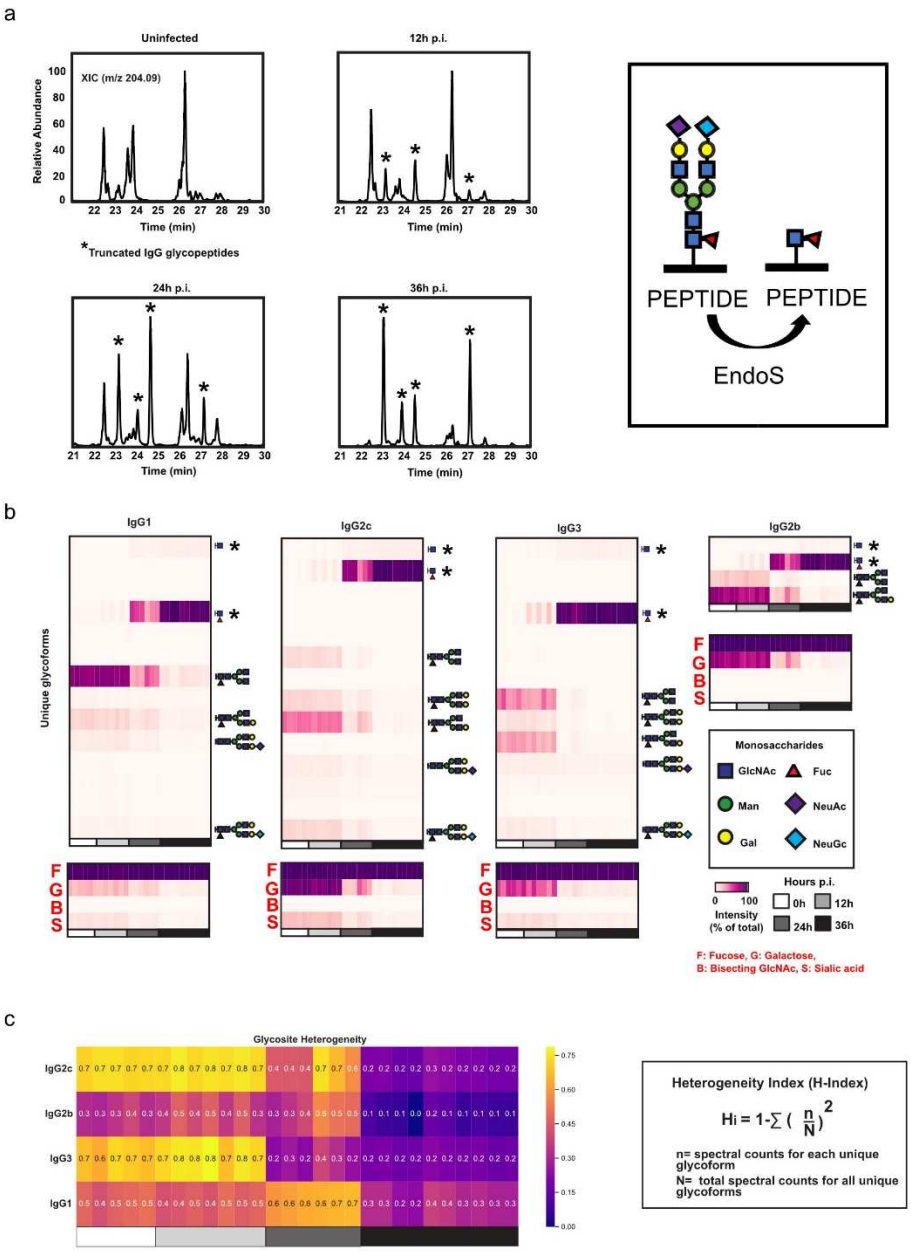
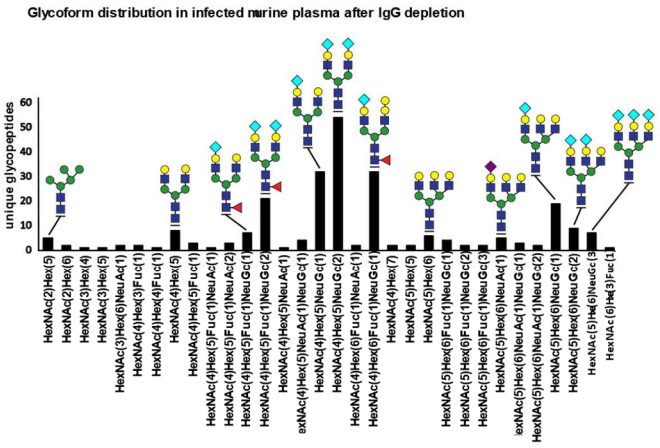


Figure 2. 2. Glycosylation analysis of murine IgGs during local to systemic GAS infection

Representative time-resolved elution profiles of murine tryptic IgG glycopeptides based on extracted ion chromatograms (XIC) of the N-acetylglucosamine (GlcNAc) oxonium ion m/z 204.09, generated upon glycopeptide fragmentation. Stars denote differential chromatographic peaks compared to uninfected controls (a). Relative quantification of the glycosylation pattern of IgG1, IgG2c, IgG3, and IgG2b across infected animals (n=5-10 animals/ time point). Truncated glycoforms are highlighted with stars. The glycan composition of the most abundant glycoforms are drawn (b). FC-region glycosylation heterogeneity as measured by an adapted version of the Simpson Diversity Index. Values range from 0-1.0 with 1.0 referring to equal probability of selecting any glycoform that is present at the specific glycosylation site (c).

Reduction of glycoform diversity in GAS infected murine plasma is specific to IgGs.

We suspect EndoS is the major virulence factor in deglycosylation of the pool of murine IgGs as EndoS is known to be highly specific towards human and murine antibodies (Collin & Olsen, 2001, Naegeli et al., 2019). In order to determine if other blood glycoproteins were experiencing trimming of N-linked glycans in the subcutaneous GAS infection model we measure glycopeptides from depleted plasma. Briefly, we passed plasma collected from each infection timepoint through Protein G columns and collected the flow-through (plasma depleted of IgGs) and subjected these fractions to the mass spectrometry based glycoproteomics analysis. There were 246 unique glycopeptides found over 48 of the most abundant plasma glycoproteins where we quantified site-specific N-glycan abundance. By tracking the spectral counts of these 246 unique glycopeptides we found that the murine plasma N-glycome was very diverse even under infection conditions. We were able to determine that GAS-mediated glycan truncation is exclusively associated with the murine IgGs due to the lack of appearance of glycopeptides containing a single N-acetyl-glucosamine (GlcNAc) linked to the asparagine residue, with or without a core fucose (Fuc) (Fig. 2.3). The murine plasma N-glycome remained saturated with bi- and tri-antennary complex-type structures, with a minor amount of high-mannose structures.



Glycoproteomics identifications in infected murine plasma after IgG depletion

Proteins	Unique glycopeptides	Glycopeptide counts
Afamin	3	76
Alpha-1-acid glycoprotein 1	9	679
Alpha-1-acid glycoprotein 2	6	269
Alpha-1-antitrypsin 1-3	3	165
Alpha-2-macroglobulin-P	2	23
Angiotensinogen	1	5
Antithrombin-III	4	44
Apolipoprotein C-IV	2	8
Beta-2-glycoprotein 1	2	39
Carboxylesterase 1C	4	212
Carboxypeptidase B2	3	323
Carboxypeptidase N subunit 2	1	16
Ceruloplasmin	4	35
Coagulation factor V	1	2
Complement C4-B	1	2
Complement factor H	4	69
Complement factor I	2	27
Corticosteroid-binding globulin	5	43
Epidermal growth factor receptor	1	7
Fetuin-B	1	71
Fibrinogen beta chain	15	895
Fibronectin	2	84
H-2 class I histocompatibility antigen	1	103
Haptoglobin	31	1386
Hemopexin	6	246
Immunoglobulin heavy constant mu	6	119
Inter alpha-trypsin inhibitor heavy chain 4	11	406
Inter-alpha-trypsin inhibitor heavy chain H1	2	7
Isoform 2 of Leukemia inhibitory factor receptor	4	81
Isoform D of Proteoglycan 4	1	2
Kininogen-1	6	567
Lipopolysaccharide-binding protein	1	2
Monocyte differentiation antigen CD14	2	19
Murinoglobulin-1	15	377
Phosphatidylcholine-sterol acyltransferase	1	7
Phosphatidylinositol-glycan-specific phospholipase D	5	46
Plasma kallikrein	3	18
Pregnancy zone protein	40	2901
Protein Z-dependent protease inhibitor	1	5
Prothrombin	4	145
Selenoprotein P	1	4
Serine protease inhibitor A3F	3	68
Serine protease inhibitor A3K	14	754
Serine protease inhibitor A3M	1	4
Serum amyloid P-component	2	41
Serum paraoxonase/arylesterase 1	2	32
Transthyretin	2	68
Vitronectin	5	55

Glycoproteins: 48, unique glycopeptides: 246

Figure 2. 3. Glycoform distribution of glycopeptides derived from non-IgG glycoproteins

Non-IgG glycoproteins were isolated for glycoproteomics after plasma de-erichment of IgGs via protein G column. Glycopeptides were measured in the plasma of GAS-infected mice at 36 p.i. Hex: Hexose, HexNAc: N-acetylglucosamine, Fuc: fucose, NeuAc: N-acetylneuraminic acid, NeuGc: N-glycolylneuraminic acid.

Therapeutic intravenous immunoglobulins (IVIGs) are degraded in GAS infected plasma from a subcutaneous model of infection with dependance on EndoS. It is important to know the possibility of interference with evaluation of antibody-based therapies in our model of GAS infection. After investigation and conclusion that murine GAS bacteremia results in targeted degradation of endogenous IgG3 and unspecific deglycosylated IgG FC-glycans, we investigated the dynamics of exogenously administered IgGs for treatment of GAS bacteremia. This issue is important due to the fact that FC-region glycosylation is linked to FC receptor activation and immune system functions. Deglycosylation of the FC-region could result in underestimation of protective FC-mediated effector functions that are glycan dependent. In order to determine if infected murine plasma induced deglycosylation of murine IgGs exclusively we purchased human intravenous immunoglobulins (IVIGs) to administer at different stages of infection. These IVIGs were a pharmaceutical-grade IgG mixture derived from thousands of human donors, and therefore representative of different IgG subclasses and glycosylation patterns. In order to track the rate of deglycosylation we exploited the fact that mouse and human IgGs have different amino acid sequences and can be separately quantified through mass spectrometry. Infected mouse plasma was harvested 36h p.i. with GAS and incubated overnight with an aliquot of IVIGs. We found truncated human FC-glycopeptides are abundantly detected in infected mouse plasma after *ex vivo* overnight incubation with IVIGs (Fig. 2.4a-c). To investigate if GAS mediated deglycosylation of IVIGs takes place *in vivo*, we pretreated mice with literature based preclinical experimental therapeutic doses of the antibodies. Mass spectrometry analysis of glycopeptides in murine plasma after pretreatment with IVIGs confirms that human IgG glycopeptides are in circulation at therapeutic doses and outweigh the endogenous murine IgG pool (Fig. 2.4d). Challenging the pretreated mice with GAS via a subcutaneous route of infection results in 50-70% of the injected

human IgGs being targeted by the glycosidase activity *in vivo*, becoming partially degraded by 36h p.i. (Fig. 2.4e-f).

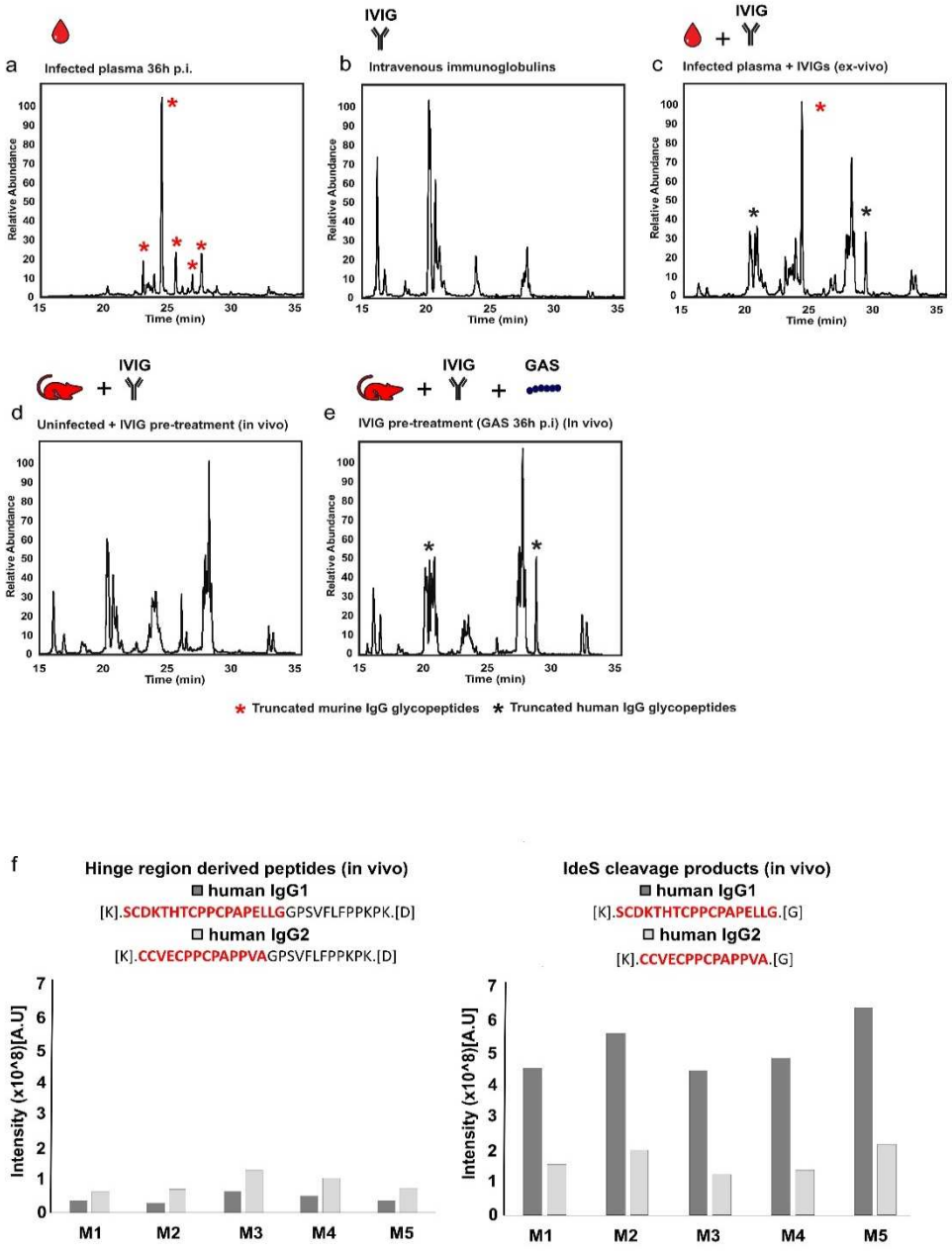


Figure 2. 4. Deglycosylation of human IVIGs and IdeS activity contributes to IVIG degradation in vivo

Representative elution profiles of IgG glycopeptides based on extracted ion chromatograms (XIC) of the N-acetylglucosamine (GlcNAc) oxonium ion m/z 204.09, and purified from infected mouse plasma 36h p.i. (a), pharmaceutical-grade IVIGs (b), infected mouse plasma spiked with IVIG and incubated ex-vivo (c), plasma from IVIG-treated mice (d), IVIG-treated mice challenged with wildtype GAS (e). Red stars denote truncated mouse IgGs; black stars denote truncated human IgGs. Mass spectrometric intensity of tryptic peptides derived from intact and IdeS-cleaved human IgG hinge regions in murine plasma upon GAS infection (f).

To determine if exogenously administered IVIGs were being targeted for degradation by the EndoS-IdeS deglycosylation and clearance system, we used a GAS variant with a genetic manipulation of EndoS expression to knock out the virulence factor. Subcutaneous inoculation with EndoS-mutant GAS resulted in negligible difference on the host plasma proteome response to infection and does not affect bacterial clearance from the skin or from secondary organs (Supplemental Fig. 2.2). Overall disease progression remains unaltered, based on the significant weight losses experienced by all animals, independently of genotype or IVIG treatment administration (Supplemental Fig. 2.2). A bacterial challenge with an EndoS-mutant reduces both the truncated murine and human IgG glycopeptides to base levels, and significantly increases the amounts of extended glycan structures on human IgG1 and IgG2, strongly confirming the specific role of EndoS in driving the glycan degradation (Fig. 2.5).

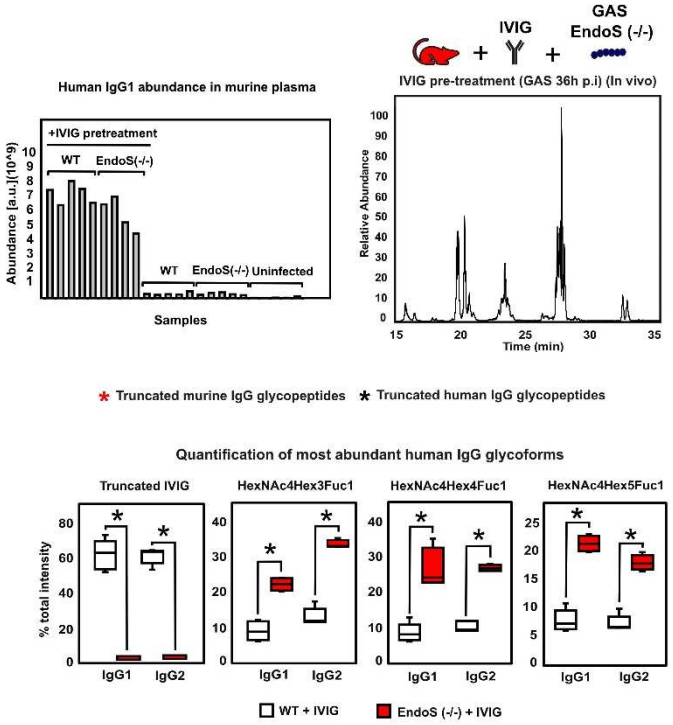


Figure 2. 5. EndoS drives deglycosylation of human IVIGs

Representative elution profiles of IgG glycopeptides based on extracted ion chromatograms (XIC) of the N-acetylglucosamine (GlcNAc) oxonium ion m/z 204.09, and purified from IVIG-treated mice challenged with EndoS (KO) GAS. Red stars denote truncated mouse IgGs; black stars denote truncated human IgGs. Quantification of human IgG1 and IgG2 FC glycopeptides upon infection with wildtype vs EndoS (KO) GAS strains.

Intraperitoneal GAS infection model loses EndoS activity for both endogenous and exogenous IgGs. Notably, another bacteremia inoculation model is delivery of the pathogen via intraperitoneal (IP) infection. In order to investigate if the EndoS-IdeS deglycosylation and degradation mechanism was consistent across both inoculations (subcutaneous and IP) of mouse models pretreated with IVIGs we performed glycoproteomics analysis on GAS IP infection plasma samples at 0h and 24h p.i.. Mice were inoculated with the AP1 strain intraperitoneally, to completely bypass the local skin stage of the infection and test host microenvironment impacts. Surprisingly, both endogenous and exogenously administered IgGs were not significantly deglycosylated in any fashion (Fig. 2.6a-d). Additionally, tracking human IgG peptides revealed

that IVIGs were not being targeted for degradation in circulating plasma. In sharp contrast to the results obtained in the subcutaneous model, pre-treatment of the IP model with IVIG substantially reduced the levels of splenic colonization, with a similar trend also observed for other organs such as the liver, but without having a significant effect on weight loss (Fig. 2.6e).

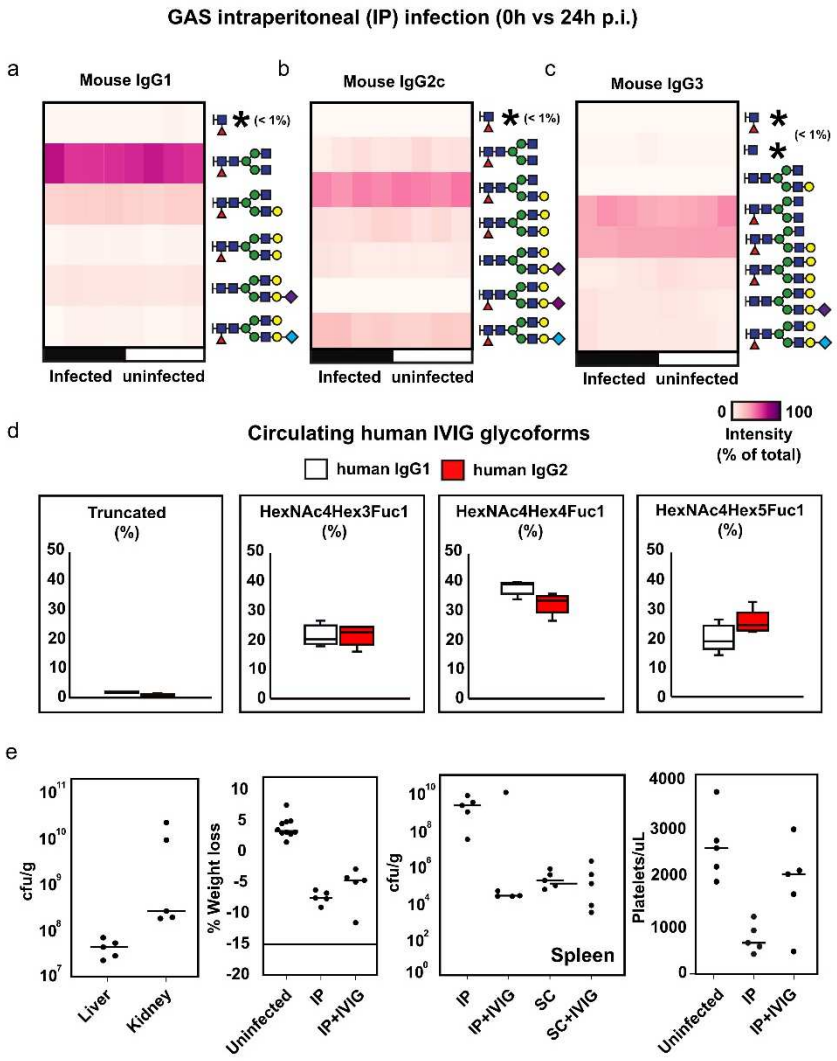


Figure 2. 6. EndoS is not catalytically active in the murine model of intraperitoneal (IP) GAS infection

Glycopeptide analysis of most abundant murine plasma IgG1 (a), IgG2c (b) and IgG3 © glycoforms before and after IP-infection, and glycopeptide quantification of human IVIGs circulating in mouse plasma at 24h p.i. (n=5 mice / condition) (d). Weight measurements of intraperitoneal (IP) infected mice at 24h p.i. with or without IVIG pretreatment (6h before infection). Splenic and hepatic bacterial burden in IP vs subcutaneous (SC) infection models and IVIG treatment effects. Stabilizing effect of IVIGs in platelet counts (e).

Human GAS sepsis patients do not exhibit EndoS targeted deglycosylation in acute plasma. We had access to three GAS-positive septic patients both during and post-recovery. Both acute and convalescent sera were collected during the course of the disease for each patient. Interestingly, the level of truncated IgG N-glycans in both the acute and convalescent sera was below 1% (Fig. 2.7), in stark contrast to the complete IgG deglycosylation in the subcutaneous GAS mouse model. Previous studies have shown that local IgG deglycosylation can be observed in throat swab samples from GAS-infected tonsillitis patients, but it has been more difficult to detect in the plasma of systemically infected individuals with GAS sepsis.

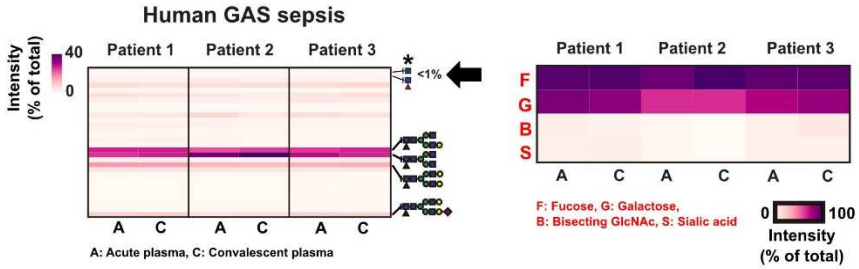


Figure 2. 7. Glycosylation analysis of human IgG in 3 patients with GAS sepsis

Relative quantification of the glycosylation pattern of IgGs across 3 patients with GAS sepsis from either acute phase and convalescent plasma. IgGs were collapsed as glycopeptide counts were similar across each subtype.

2.5 Discussion

In this investigation of a murine model of severe GAS infection and administration of human IVIGs for treatment, we have illuminated several factors that are important for consideration when testing for IgG mediated anti-inflammatory and protective effects. Far from being a generic syndrome, there is growing evidence that the septic course is largely heterogeneous, reflecting the complex interaction between a myriad of host and pathogen factors. In previous work I have shown multiple examples of host proteome response directly combating

foreign pathogens and inflammatory and coagulation cascades; however, here I show that the infecting agent also actively combats host mechanisms during GAS sepsis.

Interestingly, work conducted in collaboration with Dr. Oonagh Shannon, a leading authority in platelet biology at Lund University, suggests that the pattern of vascular and organ dysfunction in multiple infection models are quite different. For example, compared to the *S. aureus*, the *E.coli* model is dominated by a localized glomerular phenotype, whereas the GAS model has a more diffused pattern, with signs of widespread disseminated intravascular coagulation (DIC) in most organs, as well as changes in IgG homeostasis (Shannon 2020). Thus, it is very likely that alternative mechanisms of vascular failure might be evoked in these other bacterial models. Here we are able to demonstrate that even within a single model of GAS infection multiple factors should be considered in evaluation of IgG mediated protection. We show that species specific IgGs are subject to protease and glycosidase activity and that infection inoculation routes interfere with the evaluation of antibody-based therapies in our model. For example, we have shown the targeted degradation of murine IgG3 in a subcutaneous GAS infection. This is not recapitulated in the MRSA model we have previously interrogated and may be facilitated by the GAS virulence factor IdeS. After monitoring the degradation of exogenously administered IVIGs by both GAS infected plasma and subcutaneous GAS infection of the mice, we further confirmed the specific IgG degradation. Using mass spectrometry to track the tryptic peptide fragments, we saw that IdeS is the primary suspect due to the appearance of FC hinge region fragments. Using this data, we suspect and show that murine IgG3 and all human IgGs administered to our subcutaneous model would experience IdeS cleavage and clearance. This is a crucial mechanism to control for when testing human antibodies in a subcutaneous murine GAS model. Furthermore, we have shown the clear removal of FC-region glycans modulated by the GAS virulence factor

EndoS. This glycan degradation was unique to both murine and human antibodies and did not extend to other plasma glycoproteins. We have seen that factoring IgG deglycosylation and subsequent degradation has been severely missed in a review of current models of GAS infection and GAS induced sepsis. Given our results of both an EndoS-mutant GAS infection model and an IP infection model we have proposed multiple testing systems that directly combat proteolytic and glycosidic removal of administered human antibodies. Finally, we demonstrated that acute plasma in clinical samples derived from 3 patients which experienced severe GAS infection there was seemingly no deglycosylation of natural antibodies. Follow up measures of convalescent plasma from the same patients once again showed no deglycosylation. This may extend the argument that the proteolytic and glycan degradation of murine IgG3 and administered IVIGs in the subcutaneous model of GAS infection is not properly replicating what may be seen in the hospital.

Previous reports have suggested that FC-fragments derived from IVIGs have a therapeutic effect in children suffering from autoimmune diseases (Hoffman & Enk, 2019). The important role of IVIGs in the treatment landscape requires special attention to the preclinical models used to assess biological availability, efficacy and effectiveness. As we have previously shown, the blood-tissue interface undergoes rapid and expansive proteomic changes during bacteremia. This compartment-wide remodeling exposes new proteomic motifs to the blood-tissue interface that can be anchored to a cell or free floating. Passive immunization through the administration of exogenous IVIGs offers a therapeutic prospect for personalized fine-tuning of interaction with the newly exposed protein motifs during bacteremia. However, two major questions remain to be interrogated. First, how disease specific virulence factors differentially impact host IgGs vs. human IVIGs in a murine model of GAS infection. Our results highlight the fact that mouse and human IgGs are significantly different as to their structures, subtype distributions, and sensitivity to

bacterial proteolytic factors, which should be kept in mind when translating findings from murine models to human settings. Second, can genetic manipulation of both model systems and IVIGs be utilized to understand the role of FC-glycosylation in binding and activation of both host and pathogen IgG-binding factors that are relevant to GAS infection. We have briefly touched this point here by investigation of GAS induced sepsis via an EndoS-mutant strain of the bacteria. In the murine EndoS-mutant GAS model we have seen the return of wild-type glycosylation patterns for both murine and human IgG glycopeptides isolating the role of EndoS in performing the IgG glycan degradation. This is an important issue since pathogen-driven glycan truncation may result in inconclusive results obfuscating the effectiveness of exogenously administered IVIGs in the most common murine model of GAS infection. Future prospectives point towards the modulation of specific glycan epitopes on IVIGs to target enhanced anti-inflammatory benefits of glyco-engineered IVIGs. The fact that sialic acid engineered IVIGs also demonstrate an immunomodulatory effect in murine models of autoimmune disorders makes us to hypothesize that some of these effects are specifically mediated by IgG-subtypes with selective FC-glycopatterns. More importantly, these anti-inflammatory benefits might be also relevant to treat other severe immune disorders such as bacterial sepsis. The molecular basis for this enhancement and the potential contribution of other rare IVIG FC-glycoforms have not been fully elucidated. Therefore, we hope to use enrichment methods to target more defined IgG glycan constructs from IVIG populations based on glycosylation status. to. We hope to interrogate some of the FC-dependent IVIG beneficial effects that may protect against organ damage and vascular failure in GAS sepsis.

2.6 Acknowledgements

This research was supported by grant P01HL131474 to J.D.E. and by funding from NIGMS (R35 GM119850 N.E.L., T32GM008806 J.T.S) and the Novo Nordisk Foundation (NNF20SA0066621 N.E.L.). AGT, CK, SC and JM were supported by the Wallenberg foundation (WAF grant number 2017.0271), the Swedish research council (grant number 2019-01646 and 2018-05795), the Viral and Bacterial Adhesin Network Training (ViBrANT) Program funded by the European Union's HORIZON 2020 Research and Innovation Program under the Marie Sklodowska-Curie Grant Agreement No 765042 and Alfred Österlunds Foundation.

2.7 Author Information

J.T.S. and A.G.T. conceived the glycoproteomics project. A.G.T., E.B., O.S., and J.M. conceived the GAS investigation. J.T.S. developed the analysis for proteomics and glycoproteomics. A.G.T, E.B., E.V., and S.C. conduct the mouse infection, treatment, purification, and mass spectrometry. C.K. constructed the in-silico library for DIA-SWATH-MS and quantified the DIA data. A.G.T. quantified the DDA glycoproteomics data. A.G.T., E.B., E.V., S.C., J.T.S., C.K., N.E.L., J.D.E., M.C., O.S., and J.M. interpreted the data and wrote the manuscript with significant input from all authors.

Chapter 2 is adapted from "Pathogen-driven IgG degradation in murine models of Group A streptococcal infection", a collaborative manuscript in preparation by Alejandro Gomez Toledo, Eleni Bratanis, Erika Velásquez, Sounak Chowdhury, James T Sorrentino, Christofer Karlsson, Nathan E Lewis, Jeffrey D Esko, Mattias Collin, Oonagh Shannon, and Johan Malmström. The dissertation author was a primary investigator and author on this paper.

2.8 Supplemental Materials

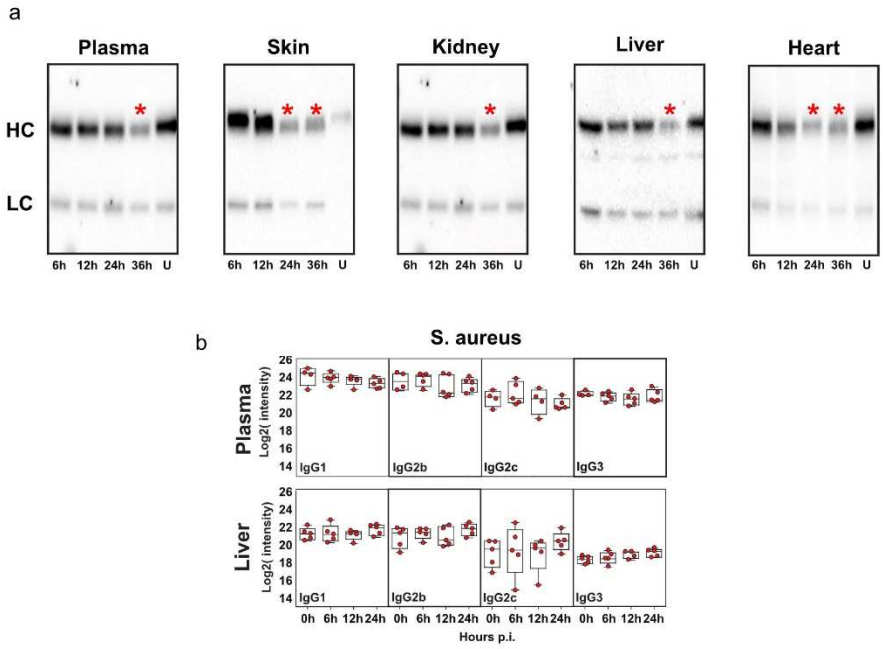


Figure 2.S 1. Distinct degradation of IgGs in GAS samples vs. lack thereof in S. aureus samples

Western blot analysis of IgG levels in plasma, skin, kidney, liver and heart samples across the course of infection (a). Mass spectrometric measurements of murine IgG subclasses in plasma and liver homogenates during S. aureus infections (b). Red stars denote statistically significant changes.

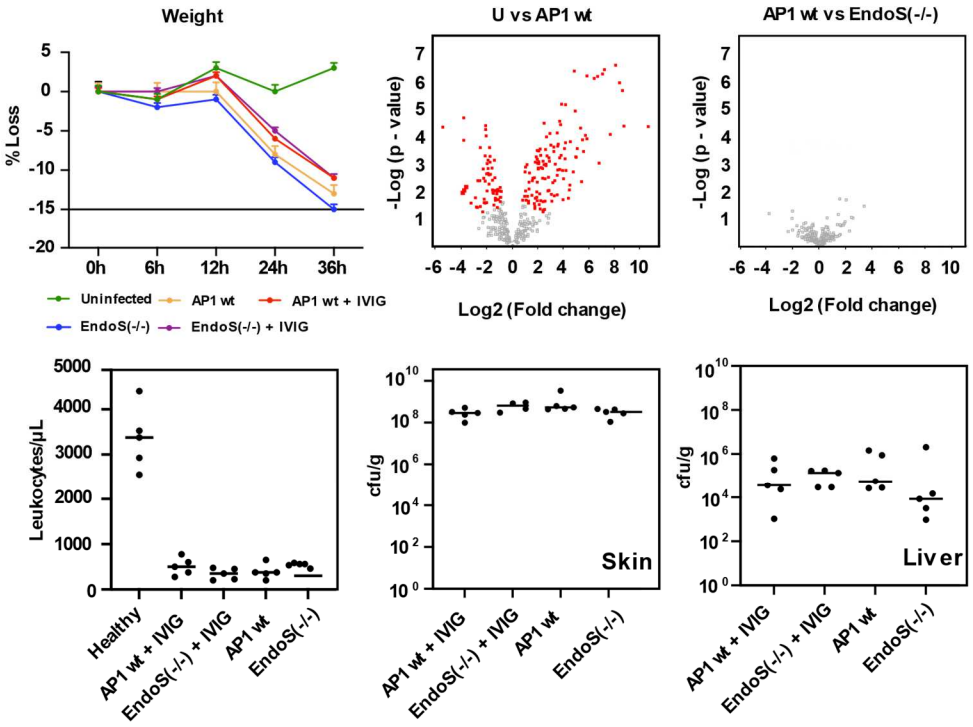


Figure 2.S 2. GAS virulence in the subcutaneous infection model is not impacted by IVIG treatments or EndoS-inactivation

Weight measurements of infected mice during the course of infection (n=5 animals/ time point). Volcano plots of changes in plasma proteins between uninfected and infected mice with wildtype bacteria, and between animals infected with wildtype vs EndoS (KO) GAS strain at 36h p.i.. Plasma leukocyte counts. Bacterial burden in skin, and liver.

Chapter 3: A network-based strategy for annotating and quantifying glycopeptides and assessing the heterogeneity of glycosylation in circulating glycoproteins

3.1 Abstract

The predominant functional units of a healthy organism are distinct protein molecules modified for important biological tasks. Post translational modifications (PTMs), such as glycans, impact protein folding, stability, viability, and interactions giving rise to both microscopic and macroscopic diversity. Glycans impact biological functions, and several major biotherapeutics that harbor N-linked glycans, such as monoclonal antibodies. Antibodies are functional circulating glycoproteins crucial for the organism's immune response. Glycoproteomics approaches identify and quantify glycan heterogeneity present on proteins and illuminate their purpose as functional PTMs. However, unique properties of glycans hinder high throughput glycoproteomics and efficient data analysis. Importantly, glycan biosynthesis is determined by several factors including metabolite availability, proximal amino-acid sequence, enzymatic activity. Experimentally, glycans cause chromatographic challenges and spectral challenges during mass spectrometry protocols. To accelerate glycoproteomics data analysis, we developed an analysis protocol that leverages molecular networking to aid in more rapid and precise annotation and quantification of glycans using data independent acquisition-mass spectrometry (DIA-MS). This approach creates dual purpose datasets used to extract intensities and quantify glycopeptides as well as glycan features. We demonstrate this method by assessing five antibody glycoproteomics datasets that reveal the utility of biodiversity indices in the quantification of context specific glycan traits. These

results showcase differences in pooled donor antibody therapy and recombinantly produced antibodies, non-canonical antibody glycans resulting from glycoengineering, and sepsis-induced glycan features.

3.2 Introduction

Glycoproteomics is a field of study for investigating the systems-wide biological context of peptides with site-specific glycosylation. Hypotheses within this field are typically drawn from measurements via liquid chromatography tandem mass spectrometry (LC-MS/MS) experiments and are analyzed using a variety of software tools designed to account for both the peptide and glycan information in the mass spectra. The analysis of glycopeptides poses numerous challenges including (i) limited knowledge of the glycan compositions available to different peptide sequences, (ii) high similarity of chemical properties of glycans with differing composition, causing co-elution of glycopeptides, (iii) limited capture of elution peak intensity due to many co-eluting glycopeptides, and (iv) co-fragmentation of glycopeptides with subtle differences in composition. Currently no existing method leverages properties of glycan heterogeneity to address these challenges; moreover, many current approaches focus on single glycan species and do not leverage population-based metrics. By explicitly tracking and quantifying ions resulting from fragmentation of co-eluting glycopeptides we can measure data-driven glycopeptide features paired with full glycan composition measures using existing LC-MS/MS methods. Our method of glycopeptide analysis is specifically designed to measure population traits of glycoforms appearing at glycosylation sites and is the first method that successfully uses extra-large acquisition windows for fragmenting co-eluting glycopeptides.

Glycopeptide identification is crucial to glycoproteomics. Existing software tools have employed a variety of strategies to search patterns of MS/MS spectra and connect them to either

N- or O-linked glycopeptides (Dallas et al., 2012, Hu et al., 2016). MS/MS spectra contain both peptide and glycan fragments, thus confounding efforts to identify injected biomolecules from their molecular makeup. Many peptide identification tools have been adapted to accommodate glycopeptide identification, such as Byonic, MSFragger, Skyline, Proteome Discoverer, and Protein Prospector. Additionally, new standalone software has been developed for general application and/or specific needs, such as pGlyco, MAGIC, SweetNET, and several others, as has been previously reviewed (Bollineni et al., 2018, Kong et al., 2022, Lee et al., 2016). Commonly, the strategy for glycopeptide identification includes using search algorithms, such as Mascot to score matches in a predefined or user-defined database of peptide/glycopeptide fragmentation patterns. However, glycans are immensely diverse and have several intrinsic properties that make it difficult to feasibly catalog the whole population.

Glycan heterogeneity poses a primary challenge in glycopeptide identification. The heterogeneity emerges since glycan biosynthesis is a sequential procedure of adding and subtracting monosaccharides to a chain anchored onto a peptide backbone (Toustou et al., 2021). The glycan chain is made up of a combination of several distinct monosaccharides and has the unique ability to branch out from a linear structure. Size, length, and complexity of the glycan result from several genetically and environmentally regulated glycosyltransferases and glycosidases that work together simultaneously (Varki et al., 2022). Therefore, it is challenging to computationally generate the total glycopeptide repertoire of an experiment in an unbiased fashion. However, due to the shared biosynthesis network and subtle differences in structure of related glycans, glycopeptides derived from the same protein or biological context can be very similar in size and composition. When given enough MS/MS spectra of related samples, it can be useful to explore the overlapping spectral information to extend annotation of known features to similar

unknown features. In metabolomics, spectral molecular networking has been developed to align MS/MS spectra from many experiments to identify pairs of features that are closely related in mass and composition (Bandeira et al., 2007, Wang et al., 2016). The resulting molecular networks help visualize and identify similarities in a large dataset and propagate spectral annotations. We have adapted this to glycopeptides by combining traditional database searching and spectral molecular networking to enhance glycopeptide identification.

Glycopeptide quantification poses challenges that reduce accuracy of relative abundance measures within samples. Due to subtle microheterogeneity that is more complex than that of an unmodified peptide, quantification suffers from large numbers of missing values for rarer glycopeptides. For tandem mass spectrometry, biomolecules are ionized and separated in the first mass analyzer and traditional MS1 quantification of these precursor ions is used in various data acquisition approaches for glycopeptides. The first survey scan for MS precursor abundance happens in a prefixed timing given by the duty cycle. Unfortunately subtly related glycopeptides concurrently elute off on traditional separation columns and less abundant species may be missed in these surveys. There are several experimental techniques to increase the number of survey scans done during a tandem mass spectrometry experiment. One such technique is to employ an unbiased strategy for sending packets of ions to the second mass analyzer using the technique sequential window acquisition of all theoretical mass spectra (SWATH-MS). Briefly, this data independent acquisition (DIA) approach does not specifically target precursor ions for secondary fragmentation and thus can perform many more survey scans of the MS spectra at an eluting peak. However, this methodology creates more complex MS/MS spectra to interrogate due to not isolating single precursor ions for fragmentation. Additionally, the duty cycle is a by-product of window size selection as narrower windows behave more closely to traditional data dependent acquisition

(DDA) approaches. Therefore, as the window size broadens, the number of MS/MS scans across an eluting peak will increase, but MS/MS fragmentation patterns become increasingly multiplexed. Researchers have developed many strategies for deconvolving the MS/MS fragmentation patterns associated with eluting modified peptides. One such strategy is to use a rigorously constructed spectral library that consists of a map of spectral relationships defining a precursor ion. This map can be built using information gathered from several DDA experiments and can be augmented using algorithms predicting possible spectral features of precursors. Spectral library construction for DIA glycoproteomics is an area of active research with tools transferring techniques from proteomics to enable similar methods of quantification. These developments are exciting, but uncommon. Only a few fully outlined workflows have been published such as DIALib, GproDIA, and Glyco-DIA. Once a satisfactory spectral library is built, it can be used for MS2 level quantification. Software tools such as Skyline (Adams et al., 2020), are designed to leverage information from spectral libraries to identify and quantify modified peptides from their MS2 measures. However, MS2 quantification can be approach in other ways to better account for the unique properties of glycopeptides. During the SWATH-MS method, as window size increases and complexity builds, it can be infeasible to deconvolve these mixed MS/MS patterns. Interestingly, co-eluting glycopeptides often share structural features that are important for biological functions. For example, recombinantly produced IgG1 therapeutics with biantennary galactosylated glycans in the FC region have been shown to increase binding to the FcγRIII receptor for effector function and have increased half-life and safety (Eon-Duval et al., 2012). These structures can contain either 1 or two galactose monosaccharides and would not be distantly separated during chromatography. Understanding the relative abundance of glycopeptide features is important for glycan-related disease mechanisms or glycosylated protein treatments. By

modifying a spectral library to harbor spectral relationships of glycan features, the search space drastically decreases. For example, instead of attempting to quantify two galactosylated glycopeptides, a spectral library can be constructed to quantify fragments only pertaining to galactosylation. For that reason, closely related glycopeptides can be collapsed to shared features that can be measured with high accuracy. This feature-based metric can be then used for the differential analysis between samples. Flexibly in defining features of interest will allow researchers to link populations of glycopeptides to specific biological mechanisms of action.

In this study we introduce a bioinformatics workflow that can be coupled with traditional glycoproteomics workflows or can be tied to a new SWATH-MS variant. This novel approach utilizes large isolation windows for fragmentation to offer a feature-centric MS2 quantification, deemed XL-DIA. Using population-based heterogeneity analysis, borrowed from the field of ecology and evolution, we introduce easily interpretable statistics for glycopeptide data. Additionally, differential analysis facilitated through MSstats (Choi et al., 2014) enabled scoring and assessment of differential glycopeptide features across samples in each dataset. Using over 130 data files from multiple DDA experiments, we showcase the power of molecular networking by expanding the annotation of human and mouse IgG glycosylation in a subtype specific fashion. These experiments ranged over commercially available antibodies, IVIG donor samples, glycoengineered CHO produced antibodies, and sepsis models. This expanded annotation found features and glycopeptides distributions that correlated to different subtypes of IgG. Finally, we interrogated human plasma samples from group C/G streptococcus sepsis patients using the XL-DIA methodology. At the glycopeptide population level, human sepsis derived IgGs show a stark difference in fucosylation, galactosylation, and high mannose changes in acute phase vs. convalescent plasma.

3.3 Materials and Methods

The data collected here aim to introduce new methods for capture and measurement of site-specific glycan heterogeneity. The analysis of the most common antibody, Immunoglobulin G (IgG), successfully showcases advanced understanding of the scope of FC region glycosylation with subtype specific trends displayed in multiple biological contexts. Successfully implementing a glycan-centric network walk for molecular networking analysis and glycopeptide identification proves to be fruitful for both traditional quantification methods and new feature-centric approaches. A novel XL-DIA experimental protocol enhanced our ability to measure import glycan features occurring in a site-specific manner. The flexibility of the workflow leaves room for expansion to identify other complex glycan features that may be of interest in new experiments. Each dataset was analyzed using heterogeneity statistics analogous to those used in the field of ecology and evolution. This new way of quantifying glycan differences identified several important disease-driven changes and is applicable to other datasets gathered using the methodology presented here.

Glycoengineered antibody production and purification. Glycosyltransferase knockouts and knockins for the CHO cells expressing Herceptin were conducted at the Denmark Technical University. These lines were derived from the CHO-S cell line (Gibco Cat. # A11557-01), and they were generated and verified according to the procedures described previously (Amann et al., 2019). Cells were cultured in CD CHO medium (Gibco 10743-029) supplemented with 8 mM l-glutamine (Lonza BE17-605F) and 2 mL/L of anti-clumping agent (Gibco 0010057AE) according to the Gibco guidelines. The day prior to transfection, cells were washed and cultured in exponential phase in a medium not supplemented with an anti-clumping agent. At the day of transfection, viable cell density was adjusted to 800,000 cells/mL in 125 mL shake flasks (Corning

431143) containing 30 mL medium only supplemented with 8 mM l-glutamine. Plasmids encoding for Herceptin, were used for transient transfections. For each transfection, 30 µg plasmid was diluted in OptiPro SFM (Gibco 12309019) to a final volume of 750 uL. Separately, 90 uL FuGene HD reagent (Promega E2311) was diluted in 660 uL OptiPro SFM. The plasmid/OptiPro SFM mixture was added to the FuGENE HD/OptiPro SFM mixture and incubated at room temperature for 5 min. The resultant 1.5 mL plasmid/lipid mixture was added dropwise to the cells. Supernatants containing model protein were harvested after 72 h by centrifugation of cell culture at 1000g for 10 min and stored at −80 °C until purification and N-glycan analysis.

Commercial antibodies and IVIG samples. Monoclonal antibodies Avastin (Roche), Erbitux (Merck), Xolair (Novartis), Herceptin (Roche) and IVIG (Octapharma). Roughly, 10 micrograms IgGs were resuspended in 8M urea and reduced with 5 mM Tris(2-carboxyethyl)phosphine hydrochloride, pH 7.0 for 45 min at 37 °C, and alkylated with 25 mM iodoacetamide (Sigma) for 30 min at RT, followed by dilution with 100 mM ammonium bicarbonate to a final urea concentration below 1.5 M. Proteins were digested by incubation with trypsin (1/100, w/w, Sequencing Grade Modified Trypsin, Porcine; Promega) for at least 9 h at 37 °C. Digestion was stopped using 10% trifluoracetic acid (Sigma) to pH 2 to 3. Peptide clean-up was performed by C18 reversed-phase spin columns according to manufacturer instructions (Silica C18 300 Å Columns; Harvard Apparatus). Solvents were removed using a vacuum concentrator (Genevac, miVac) and samples were resuspended in 50 µl HPLC-water (Fisher Chemical) with 2% acetonitrile and 0.2% formic acid (Sigma).

Bacteria and culture conditions. *Streptococcus pyogenes* AP1 (from the Collection of the World Health Organization Collaborating Center for Reference and Research on Streptococci, Prague, Czech Republic) were grown in Todd–Hewitt broth, supplemented with 0.2 % yeast

extract (THY, BD diagnostic), overnight (o/n) at 37°C and 5% CO2. Methicillin-resistant *Staphylococcus aureus* (MRSA USA300 TCH1516) was grown as previously reported (Sorrentino et al., 2022).

Bacterial infections. All animal use and procedures were approved by the local Malmö/Lund Institutional Animal Care and Use Committee, ethical permit number 03681-2019. *S. pyogenes* AP1 was grown to logarithmic phase in THY medium (37°C, 5% CO2). Bacteria were washed and resuspended in sterile PBS. Nine-week-old female C57BL/6J mice (Janvier, Le Genest-Saint-Isle, France) were infected with 50 µl (2x105 cfu) bacteria by subcutaneous injection on the right flank, or 100 µl (1x107 cfu) by intraperitoneal injection. Control groups were similarly injected with sterile saline (n= 9). Mice were rehydrated subcutaneously with saline 24 h after infection. Body weight and general symptoms of infection were monitored regularly. Mice were sacrificed and multiple organs were collected. *S. aureus* intravenous infection was conducted as previously reported in Toledo et al., 2019 (n=5/time point). Blood was taken by cardiac puncture and collected in tubes containing sodium citrate (MiniCollect tube, Greiner Bio-One).

Patient samples. Patients with S. pyogenes bacteremia in 2018 to 2020 were prospectively included in the study after oral and written consent were obtained. Acute-phase serum was collected within 5 days after hospital admission, and convalescent-phase serum was collected after 4 to 6 weeks. Medical records of patients were reviewed to obtain clinical and epidemiological parameters. To our knowledge, none of the patients had a history of previous S. pyogenes infection during adulthood. The concentration of the immunoglobulins in serum was determined at the Department of Clinical Chemistry in Skåne, Sweden (Neergaard et al., 2022). Patients with S. dysgalactiae bacteremia were identified through the Laboratory for Clinical Microbiology in the county of Skåne, Sweden between 2017 and 2018. Patients were included after informed and

written consent. Acute sera were obtained <5 days after admission and at follow-up 6 weeks after admission convalescent sera were obtained. Epidemiological, clinical parameters and microbiological results for each patient were collected through the medical records (Bläckberg et al., 2021).

IgG pulldowns. IgGs were purified in a 96 well plate setup using the Protein G AssayMAP Bravo (Agilent) technology, according to the manufacturer's instructions. Briefly, 10µl plasma was diluted with PBS to a final volume of 100 µl and applied to pre-equilibrated Protein G columns. Columns were washed with PBS and eluted in 0.1M glycine (pH2). The final pH was neutralized with 1M Tris and saved until further use.

Trypsin digestion and peptide desalting. Protein samples were resuspended in 8M urea and reduced with 5 mM Tris(2-carboxyethyl)phosphine hydrochloride, pH 7.0 for 45 min at 37 °C, and alkylated with 25 mM iodoacetamide (Sigma) for 30 min at RT, followed by dilution with 100 mM ammonium bicarbonate to a final urea concentration below 1.5 M. Proteins were digested by incubation with trypsin (1/100, w/w, Sequencing Grade Modified Trypsin, Porcine; Promega) for at least 9 h at 37 °C. Digestion was stopped using 10% trifluoracetic acid (Sigma) to pH 2 to 3. Peptide clean-up was performed by C18 reversed-phase spin columns according to manufacturer instructions (Silica C18 300 Å Columns; Harvard Apparatus). Solvents were removed using a vacuum concentrator (Genevac, miVac) and samples were resuspended in 50 µl HPLC-water (Fisher Chemical) with 2% acetonitrile and 0.2% formic acid (Sigma).

IgG DDA glycoproteomics analysis. Purified IgG glycopeptides were analyzed on a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific) connected to an EASY-nLC 1200 ultra-HPLC system (Thermo Fisher Scientific). Peptides were trapped on a precolumn (PepMap100 C18 3 µm; 75 µm × 2 cm; Thermo Fisher Scientific) and separated on an EASY-

Spray column (Thermo Fisher Scientific). Mobile phases of solvent A (0.1% formic acid), and solvent B (0.1% formic acid, 80% acetonitrile) were used to run a linear gradient from 4 to 45% over 60 min. MS scans were acquired in data-dependent mode with the following settings, 60,000 resolution @ m/z 400, scan range m/z 600-1800, maximum injection time of 200 ms, stepped normalized collision energy (SNCE) of 15 and 35%, isolation window of 3.0 m/z, data-dependent HCD-MS/MS was performed for the ten most intense precursor ions.

IgG XL-DIA glycoproteomics analysis. Purified IgG glycopeptides were analyzed on a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific) connected to an EASY-nLC 1200 ultra-HPLC system (Thermo Fisher Scientific). Peptides were trapped on a precolumn (PepMap100 C18 3 μm; 75 μm × 2 cm; Thermo Fisher Scientific) and separated on an EASY-Spray column (Thermo Fisher Scientific). Mobile phases of solvent A (0.1% formic acid), and solvent B (0.1% formic acid, 80% acetonitrile) were used to run a linear gradient from 4 to 45% over 60 min. MS data was acquired in data-independent mode, with 2 sequential DIA isolation windows, spanning from 600-1200Da, and 1200-1800Da, followed by an MS1 survey scan. Normalized collision energy (SNCE) of 15%, 25% or 35% was used for fragmentation.

Glycoproteomics data analysis. Raw files were searched in the Byonic software integrated as a node in Proteome Discoverer. Files were searched against a UniProt human and mouse protein database using the default search strategy: enzyme: trypsin with a maximum of two missed tryptic cleavages per peptide, up to one glycan per peptide as a 'rare' variable modification, up to 10/20 ppm deviation of the observed precursor/product ion masses from the expected values, up to one Met oxidation (+15.994 Da) per peptide (variable 'common' modification), together with a predefined plasma glycan database of mouse and human structures and a decoy and contaminant

database available in Byonic. Identifications with a Byonic score >100 were considered positive hits, and their spectra were manually validated.

Molecular Networking. Raw spectra files were processed to MS/MS peak lists (.mgf files) using mscovert with the continuous wavelet transformation (CWT) method for peak selection. Peak lists were subjected to filtering to maintain only MS/MS spectra that contained an oxonium ion characteristic of HexNAc containing glycopeptides denoted by the 204.09 m/z peak. For each species, human and mouse, pooled datasets were independently created and analyzed separately. 80 human IgG DDA samples and 51 mouse IgG DDA samples from various biological contexts were considered for analysis. Molecular networks were generated using the online molecular networking workflow, METABOLOMICS-SNETS-V2, at the Global Natural Products Social Molecular Networking Server (GNPS) (http://gnps.ucsd.edu/). The Precursor Ion Mass Tolerance and Fragment Ion mass Tolerance used for consensus spectrum creation were set to 0.02 Da. The minimum cosine score between pairs of consensus spectrum to create a network edge was 0.7 and the minimum number of matched fragment ions to maintain that edge was set to 6. Network visualization and rudimentary analysis were done in Cytoscape. Node and edge lists were further analyzed using our python workflow.

Network walk for Annotation Propagation. After initial identification of glycopeptides search using Byonic's Glycopeptide Search, results files were used for seed annotation of the molecular networks inspired by our previous work on *SweetNET* (Nasir et al., 2016). This seed matching had an absolute mass tolerance of 0.1 Da and allowed for the addition or removal of a single H+ ion that may appear during initial ionization of the glycopeptide. Then the propagation algorithm searches for nodes which are connected to annotated nodes that have an acceptable mass difference for annotation transfer. An acceptable mass difference is a mass that is within 0.1 Da of

the annotated node mass +/- the mass summation of monosaccharide differences. For example, a mass difference of +15.995 could define a transferred glycopeptide annotation that has one less fucose (-146.058) and one more hexose (+162.053). The algorithm allows the user to specify the maximum possible number of changes which are allowed; for this work this was set to 2. Additionally, proximal nodes that are exactly 1 Da difference had exact glycopeptide annotation transferred to account for ionization differences (1H+). The propagation algorithm iterates over newly annotated nodes until there are no longer any annotation transfers and the network walk is complete.

Precursor library and Hybrid Spectral library generation. From the annotated molecular networks glycopeptide hits were transformed into precursor ions expected to have the possibility of appearing with our experimental setup. First, all possible ionization states between +1 and +4 are formed. The list is then filtered for experimental parameters and curated using N-linked glycan biosynthesis reaction rules (Bao & Kellman et al., 2021, Yu et al., 2016). Additionally, glycopeptide hits from the molecular networks are used to generate feature-centric spectral libraries for the XL-DIA analysis. To generate an entry in the spectral library all glycoforms observed are collected and organized into a substructure network using GlyCompare (Bao & Kellman et al., 2021). This substructure network is an acyclic graph that connects nodes via an edge if the child is a compositional substructure of the parent. For example, the galactosylated node HexNAc(4)Hex(5)Fuc(1) is a child of the sialylated parent node HexNAc(4)Hex(5)Fuc(1)NeuAc(1). Using the annotated glycoforms as a starting point, GlyCompare infers each compositional substructure required to create the full network. Therefore, this network can represent any theoretical glycopeptide fragmentation pattern. In this way, with augmentation with specific characteristic oxonium ions, an entire theoretical spectral library is

made. To substantiate the fragmentation patterns in the theoretical spectral library, the pooled DDA dataset is used. Each MS/MS consensus spectrum that is linked to an annotated glycopeptide precursor is investigated and the relative peak areas for theoretical fragments are added as information in the spectral library. This complete spectral library can be used for traditional DIA data analysis workflows. However, we further annotate this spectral library to contain entries that represent major glycan features, such as fucosylation, galactosylation, sialylation, and high-mannose structures. The user defines the criteria that a fragment ion must contain to be considered for the quantification of the glycan feature. For example, fucosylation must be quantitated by fragments that contain one or more fucose monosaccharides. Then the spectral library entries are searched for all fragment ions, which come from many precursor ions, that meet the defined criteria and are collapsed into a superclass precursor ion representing the defined feature. Thus, an entry in the feature-centric spectral library contains all observable oxonium and glycopeptide fragment ions that can appear for precursors in the superclass.

XL-DIA data analysis. The resulting feature-centric spectral library was designed to be compatible with the open-source mass spectrometer data analysis software Skyline. Once the library is loaded, known explicit retention time windows are defined for each feature. Skyline parameters are also tuned to facilitate analysis of 600 Da window DIA data files. After results are imported into Skyline, peak picking adjustments are made to synchronize the integration of peaks across selected explicit retention time windows for each feature. Transition results are exported and prepared for feature quantification using the MSstats R package. Only transitions with feature specificity are used in feature quantification. The peak area for a transition found in the two 600 Da windows is summed. MSstats summarization is run without imputation with “globalStandards” normalization. Using this normalization setting allows for specific usage of transitions that

quantify total glycopeptide abundance in each sample. Here we have elected to use the glycopeptide fragment common to all structures the Y1 (peptide + HexNAc(1)) for normalization. Differential feature abundance analysis was performed using MSstats groupComparisons method.

Glycan heterogeneity analysis. Simpson's diversity index was adapted to quantify heterogeneity of each unique glycosylation site. This Heterogeneity Index (H-index) is visualized by a heatmap with values ranging from 0.0 to 1.0 with 1.0 being equal frequency for any glycoform that appears on that glycosylation site. In the field of ecology, the Simpson's diversity index is measured by the equation $1 - \sum \left(\frac{n}{N}\right)^2$, with n being the total number of organisms of a particular species and N being the total number of organisms of all species. Here we have adapted these values to represent n being the number of spectral counts of a particular unique glycoform and N being the number of total spectral counts for all unique glycoforms at a glycosylation site.

3.4 Results

Experimental and conceptual evaluation of the determinants of glycan heterogeneity is useful in understanding many biological mechanisms such as immune response to a pathogen. In this study we have leveraged novel strategies of molecular networking with traditional and unconventional mass spectrometry to measure the scope and intensity of glycosylation on Immunoglobulin G (IgG). Additionally, context-specific IgG glycan feature distributions were determined to be affected by both species of origin and subtype sequence. Specifically, by measuring IgG FC region glycosylation in a variety of scenarios we more rigorously evaluated blood plasma response in septic samples from both human and mouse. Sharing heterogeneity measures and tactics from the fields of ecology and evolution we have proposed several statistical quantifications and visualizations of glycan response to bacteremia. Certainly, measures of site-

specific glycan heterogeneity can profoundly influence our understanding of modified peptide interactions with homeostatic and perturbed molecular mechanisms.

Development of a novel bioinformatics pipeline to assess site-specific glycan heterogeneity using mass spectrometry. As the synergistic relationships of protein-glycan partnerships are further studied, it is important to develop new experimental measures of site-specific glycosylation that can be expanded to populations of glycoproteins. Many glycoproteins have been found to have the capacity to harbor various glycans, many of which carrying distinct impactful features such as core fucosylation etc. The glycoprotein diversity can be described to have multiple levels of characteristic glycan heterogeneity (Fig. 3.1). Traditionally, macro-heterogeneity refers to the presence or absence of a sugar moiety being attached to an eligible amino acid while micro-heterogeneity describes the chemical makeup of that sugar. The newly coined meta-heterogeneity of a glycoprotein describes the decoration of various amino acids with a diverse population of glycans simultaneously (Caval et al., 2021).

In this study we focus on IgGs, which can harbor both FAB and FC region N-linked glycosylation. Specific to the work here, multiple IgG subtypes' FC region glycosylation was investigated. Even focusing on such a narrow task the levels of glycoprotein diversity present major challenges for mass spectrometry based measures (Fig. 3.2). For example the FC region is known to carry a wide variety of glycans, but due to structural similarities it is hard to measure the total glycan repertoire with current glycopeptide identification software. Similarly because the glycans in this region share characteristic motifs they are chemically similar making it difficult to quantify each glycan properly. Finally, because glycans are made via a biosynthetic network of glycosyltransferases it is hard to account for impacts of interdependent structures with traditional statistics. To improve our understanding of the scope of glycan populations residing at a

glycosylation site and to quantify population-based statistics we have built an experimental and bioinformatic workflow inspired by the DIA-SWATH-MS protocol (Bruderer et al., 2017) (Fig. 3.3). Briefly, many DDA glycoproteomics experiments were conducted on IgGs enriched from a variety of conditions ranging from recombinantly produced to human plasma from septic patients. This pooled dataset had the IgG glycopeptides annotated with existing software and was then subjected to a molecular networking algorithm to link related spectra. This similarity network was then subjected to an annotation propagation algorithm of our design to expand the glycopeptide annotations found in our pooled dataset. The details of the workflow are described in the Material and Methods section. This expanded network annotation algorithm is designed specifically to work with the output of the GNPS server. The constraints for the annotation propagation are openly allowing for glycan conformations that do not conflict with current glycan biosynthesis or degradation. However, new structure-based constraints are a possible expansion of the workflow. The precursor ion list generated after annotation propagation can be easily used in Skyline for traditional site-specific glycopeptides quantification. Additionally, the spectral library containing transitions which represent user-defined glycan features can be used in quantification of samples acquired in an XL-DIA experiment. Each piece of quantified data, glycopeptide and glycan feature measures, are visualized and represented in statistics and plots that showcase glycan micro-, macro-, and meta-heterogeneity.

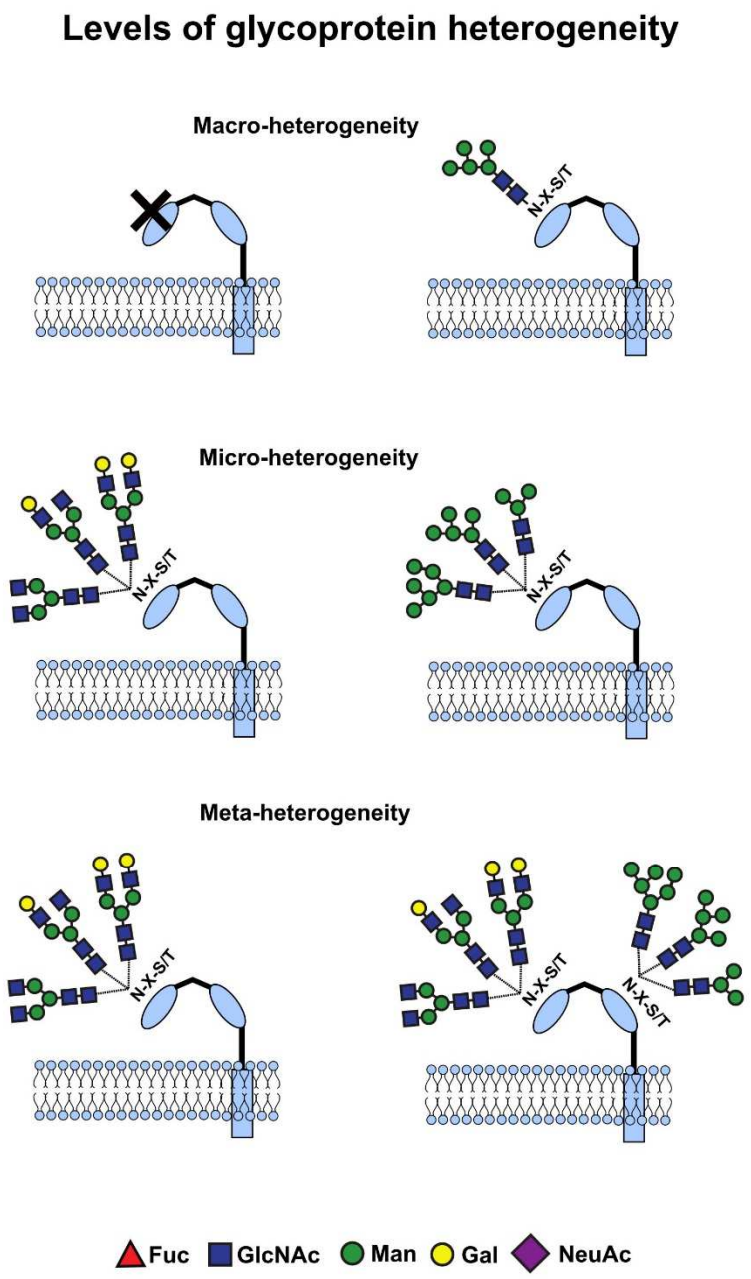


Figure 3. 1. Glycoprotein heterogeneity

Cartoon representation of the three levels of glycoprotein glycan heterogeneity referred to in the text. Macro-heterogeneity refers to presence or absence of glycosylation regardless of composition. Micro-heterogeneity refers to the distribution of full length glycans that may appear at a glycosylation site. Meta-heterogeneity refers to the diverse mixture of glycans and glycan features that may appear across a glycoprotein.

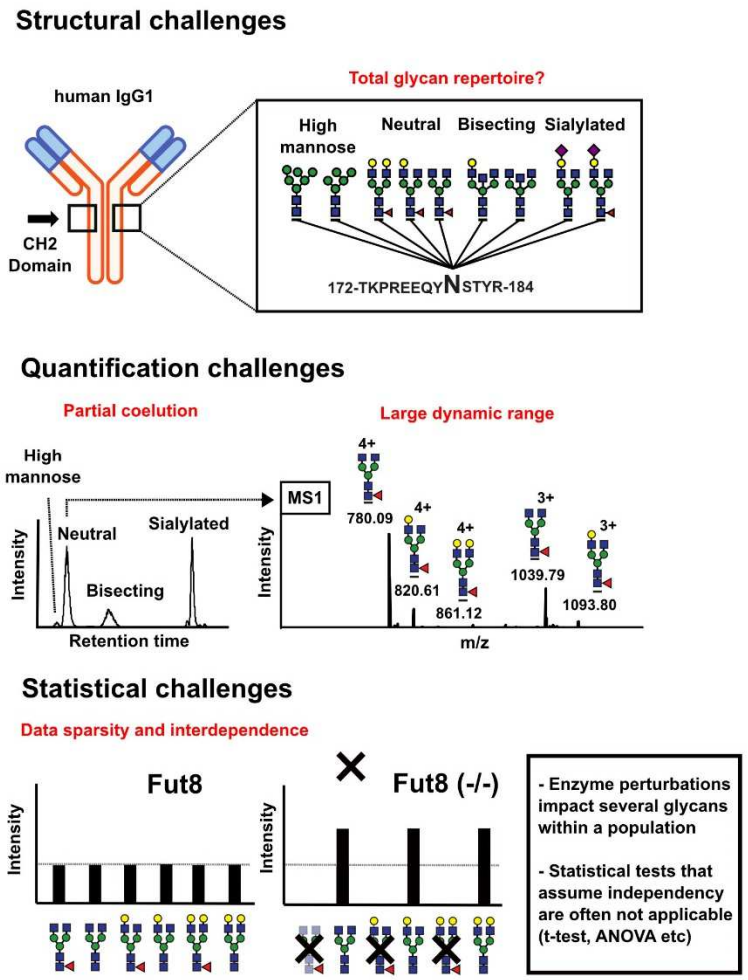


Figure 3. 2. Analytical challenges for mass spectrometry based glycoproteomics

Cartoon representation of the structural, quantitative, and statistical challenges of site-specific glycoprotein measures. Structural diversity with a dynamic repertoire able to sense environmental cues contributes to unexpected glycans. Partial co-elution of similar glycan structures with a large range of observable ions can be difficult to deconvolve. Lack of independence in glycan appearance impacts assumptions used for statistical testing.

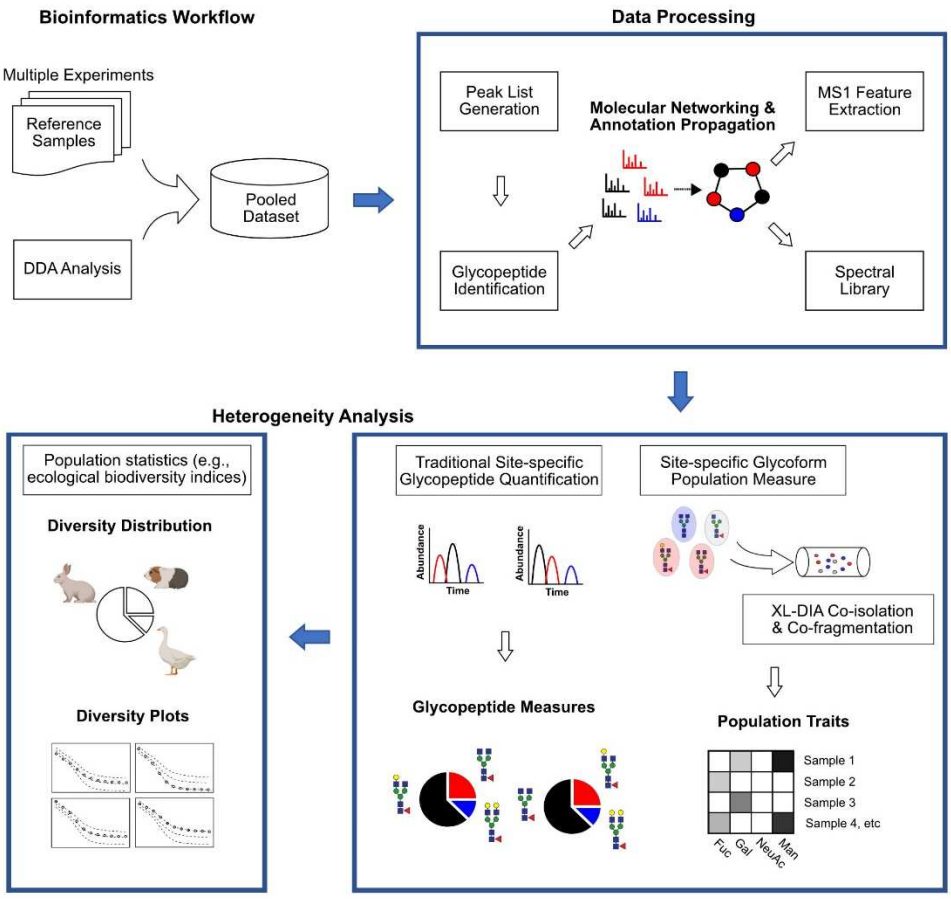


Figure 3. 3. Bioinformatic pipeline design and analytical workflow employed for site-specific glycoproteomics.

Immunoglobulin G harbors an FC region with several unique features across multiple biological contexts. For this study of IgG, we track the glycopeptides derived from the FC region on the antibody. IgG is broken into four classes IgG1-4 respectively. When the FC region is fully trypsinized in the human population of IgGs we are unable to distinguish between IgG3 and IgG4 as they are isobaric twins. Therefore, these two subtypes are analyzed together as human IgG3/4. To determine if molecular networking achieved novel meaningful glycopeptide annotations which were missed by Byonic, 80 human IgG enriched samples and 51 mouse IgG enriched samples acquired using DDA mass spectrometry were subjected to the molecular networking workflow on the GNPS server at https://gnps.ucsd.edu/. The result was two molecular networks, one for each species, which represented the spectrum-to-spectrum alignments as a similarity network. Where nodes are unique spectrum, or MS/MS m/z patterns, and edges are the connection of two unique spectrums that have a cosine similarity greater than 0.7. The human IgG network had 12259 nodes and 18062 edges made from 418242 spectra. Similarly, the mouse IgG network was composed of 6568 nodes and 8865 edges made from 227172 spectra. For the glycopeptide annotation propagation algorithm, first the Byonic glycopeptide identifications were mapped to each network. The human IgG network had 1056 nodes annotated by Byonic glycopeptide representing 38496 spectra. The mouse IgG network had 441 nodes annotated representing 15163 spectra. After the annotation propagation with flexibility of up to 2 monosaccharide changes for an annotation transfer (see Materials and Methods) the human network had 1677 nodes annotated representing 77192 spectra and the mouse network had 789 nodes annotated representing 29540 spectra (~ 2-fold increase in spectra covered for each species). An example of a transferred annotation is shown in Fig 3.4a. The spectrum representing the 3+ precursor ion of IgG1 FC region HexNAc(3)Hex(3) (shown in red) was newly identified using the

annotation propagation algorithm. Finally, the similarity network was separated into subnetworks that represented IgG subtypes showcased for human IgGs in Fig. 3.4b-d. From these similarity subnetworks IgG FC region glycopeptide precursor lists with associated fragmentation patterns ar3e harvested for MS1 quantification and spectral library design for standard DIA and XL-DIA experiments.

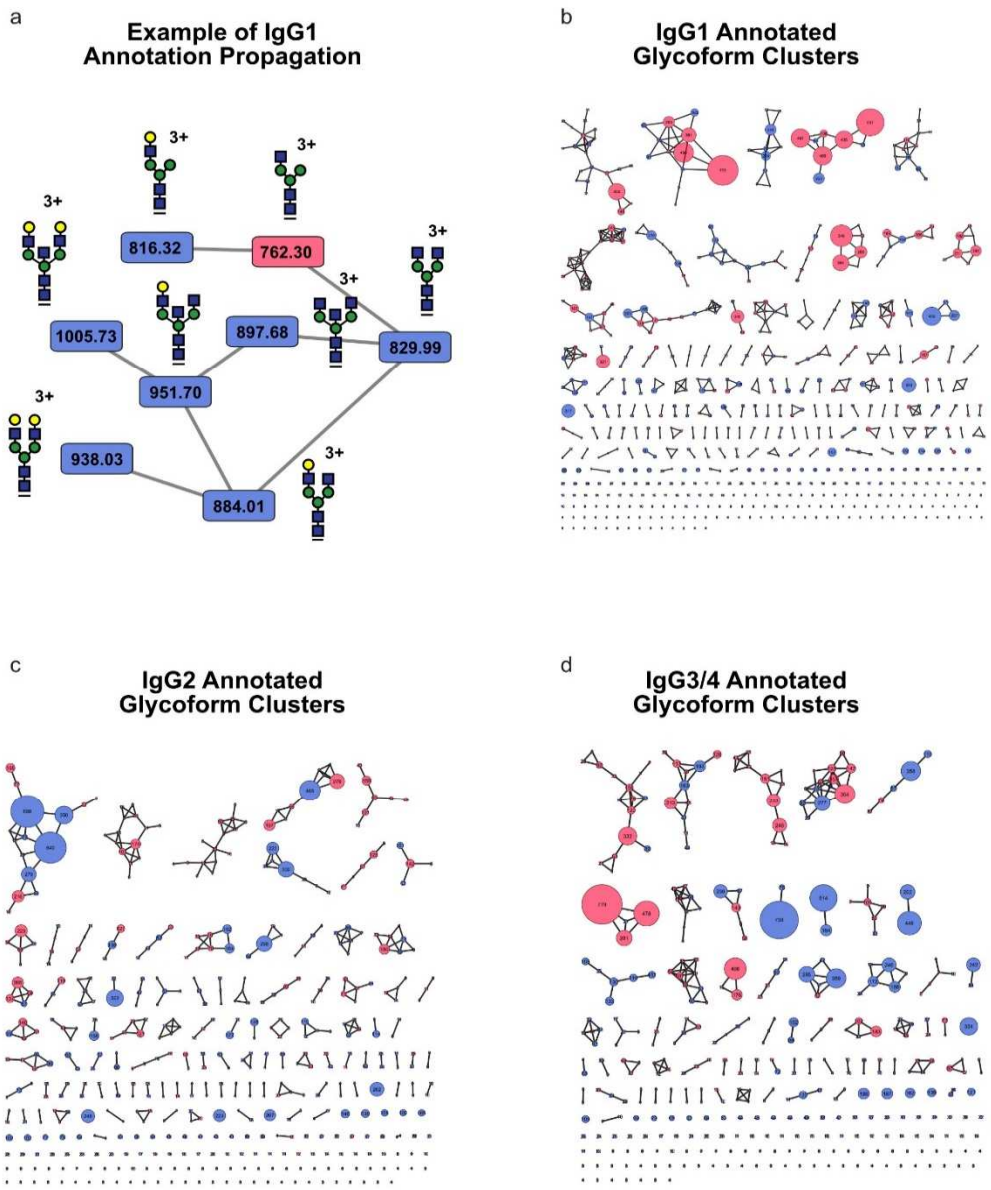
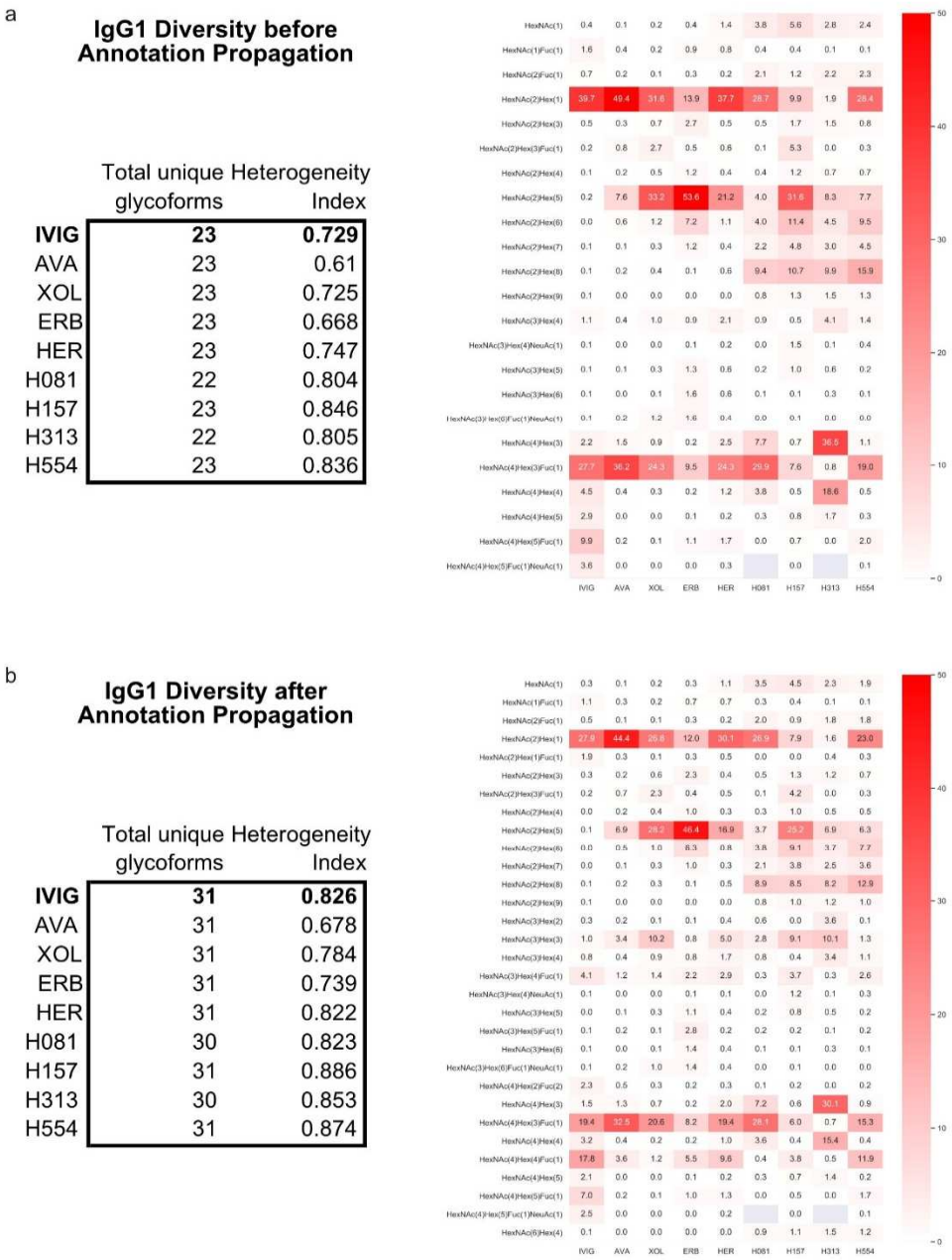


Figure 3. 4. Spectral relationships define new and unexpected glycan structures found at the FC region of human IgG

Human IgG molecular networks across 80 DDA LC-MS/MS experiments show newly identified FC region glycoforms. A newly annotated precursor ion (HexNAc(3)Hex(3)) found after propagating the Byonic annotations using the network walk algorithm in IgG1 (a). Entire annotated IgG molecular networks, IgG1 (b), IgG2 (c), IgG3/4 (d). IgG3/4 peptide backbones are isobaric peptides. Blue indicates previous annotation by the glycopeptide identification from Byonic. Red indicates a new annotation propagated by the network walk. Size of nodes is proportional to the number of spectra found to belong to the spectral cluster identified by MS/MS of a unique precursor ion.

The scope of IgG diversity is better captured with precursors generated after the annotation propagation. Using the IgG subtype specific precursor lists and associated representative spectra we were able to quantify FC region glycopeptides from IgG1 antibodies in a variety of contexts. Using a range of donated IVIG plasma, purchased on-the-market drugs (Avastin, Erbitux, Xolair, Herceptin), and glycoengineered Chinese Hamster Ovary cells (CHO) produced Herceptin. We showcase the increased granularity of glycan quantification enabled by our method. In total 30 unique IgG1 glycopeptides were identified with Byonic while 60 unique IgG1 glycopeptides were identified using Byonic + annotation propagation. In order to show the quantitative power of these novel glycopeptides, precursor ion lists were imported into Skyline and results from a DDA experiment containing our previously described IgG1 samples were analyzed. First the Byonic only precursor list found 23 unique IgG1 glycopeptides with TIC intensity greater than 1%. The distribution of IgG1 glycopeptide abundances as seen in the heatmap was represented by the heterogeneity index (H-index) defined in chapter 2 (Fig 3.5a). The Byonic + annotation propagation precursor ion list found 31 unique IgG1 glycopeptides in the same samples with the same parameters (Fig. 3.5b). Interestingly, IVIG derived IgG1 was categorized to be more diverse than all on-the-market IgG1 therapeutics only when we accounted for the newly annotated glycan structures. Also it can be seen that newly annotated glycan structures such as HexNAc(3)Hex(3) or HexNAc(3)Hex(5)Fuc(1) or HexNAc(3)Hex(4)Fuc(1) contribute significantly to the diversity measure of the IgG1 glycan population. These structures have been seen to occur on IgG (found on the Expasy database GlyConnect) (Alocci et al., 2019) and are biosynthetically achievable under homeostatic conditions. Finally, our four recombinantly produced Herceptin analogues harvested from glycoengineered CHO cell lines have perturbations to their glycan diversity that is better captured by the new scope of glycan possibilities. These lines

labeled H081, H157, H313, and H554 correspond to a galactosyltransferase knockout, sialyltransferase knockin, a fucosyltransferase knockout, and a wildtype CHO system, respectively. The stark changes in glycan distribution correspond nicely to the status of the CHO cell line. For example, lack of fucosylated glycans appearing in the FC region of Herceptin in the fucosyltransferase knockout line.



a

IgG1 Diversity before Annotation Propagation

	Total unique glycoforms	Heterogeneity Index
IVIG	**23**	**0.729**
AVA	23	0.61
XOL	23	0.725
ERB	23	0.668
HER	23	0.747
H081	22	0.804
H157	23	0.846
H313	22	0.805
H554	23	0.836

b

IgG1 Diversity after Annotation Propagation

	Total unique glycoforms	Heterogeneity Index
IVIG	**31**	**0.826**
AVA	31	0.678
XOL	31	0.784
ERB	31	0.739
HER	31	0.822
H081	30	0.823
H157	31	0.886
H313	30	0.853
H554	31	0.874

Figure 3. 5. Increase in coverage of possible glycans appearing at the FC region of IgG1 using annotation propagation

IgG1 diversity is given by the number of glycans that can appear at the FC region and their distribution. By increasing the scope of glycans measured from 30 to 60 our annotation propagation gives a better view of the FC region heterogeneity. Before the annotation propagation 23 glycans contributed 1% or more to total glycopeptide TIC intensity from MS1 measures and the distribution is given by the heatmap (a). After annotation propagation 31 glycans contributed 1% or more to total glycopeptide TIC intensity from MS1 measures (b). Heterogeneity was measured by implementing the Simpson's diversity index (see Materials and Methods).

Unconventional mass spectrometer data acquisition useful for feature-centric site-specific IgG glycosylation measures and traditional quantitation. While it is extremely useful to utilize our expanded list of IgG precursor ions to quantify the differences in glycopeptides appearing at the FC region of IgG subtypes in differing conditions, it does not come without challenge. Due to the inherent nature of the DDA approach during LC-MS/MS, only select precursor ions are shipped to the second machine for MS/MS fragmentation and thus the number of identifications is limited usually to highly abundant precursor ions. Therefore, the standard using this method is to quantify using MS1 measures. However, we are most interested in glycan motifs with functional relevance in many of our studies and thus would not only want to quantify the most abundant glycans, but instead capture the abundance of a glycan motif directly. That is why we have implemented a DIA-SWATH-MS protocol for the measurement of glycan feature-centric spectral libraries. Using this approach, we can fragment and measure the byproducts of each precursor ion corresponding to a glycopeptide in an unbiased manner. Naturally this approach helps us gather and measure the fragment ions from multiple co-eluting glycans. It is true that there is a major challenge to then deconvolve these fragmentation windows because they contain several precursor ions, but for our purposes, deconvolution is not a goal. Instead, we quantitate the fragment ions corresponding to a glycan motif, enabling high resolution measures of the functional components of the glycan. By tracing the identifying oxonium ion corresponding to freely ionized HexNAc monosaccharide (ms2@m/z 204.09) we can understand how MS2 quantitative ability is affected by what is sent for fragmentation (Fig. 3.6). During a DDA experiment neutral glycan species eluting across a ~2 min peak will have about ~50-75 MS/MS scans which can be used for fragment quantitation resulting in poor capture of the peak (Fig. 3.6a). Agreeably the same level of MS2 quantitation is available using standard window size (30Da) in a DIA-SWATH-MS

approach (Fig. 3.6b). However, because we are not trying to deconvolve the fragment patterns for exact glycopeptide quantification, we can widen the fragmentation window to 200Da or even 600Da, deemed XL-DIA. This enables us to have thousands of MS/MS scans over the course of a ~2-minute peak. Greatly enhancing the quantitative ability of the MS/MS data (Fig. 3.6c-d).

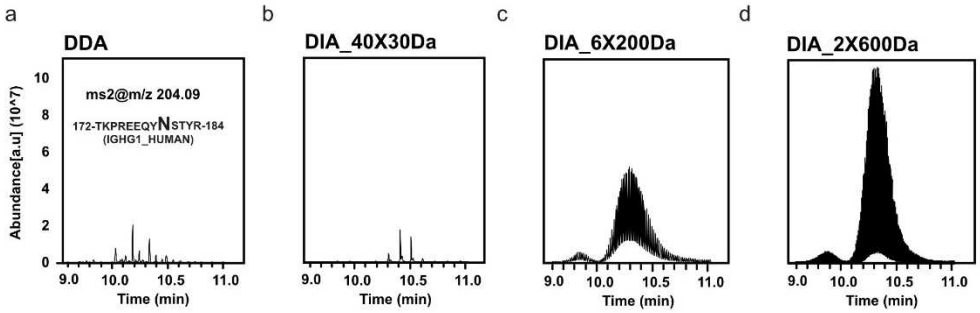


Figure 3. 6. Intensity of glycopeptide fragment ions increase as the isolation window range opens during SWATH-MS

Tracking the characteristic 204.09 ion shows that neutral biantennary glycoforms elute during a small retention time window in mass spectrometry experiments of human plasma derived IgG1 from IVIGs. Varying the window size for MS/MS fragmentation during the SWATH-MS protocol increases the number of scans that are taken during a ~ 2 min elution peak resulting in a larger amount of fragment ions measured from glycopeptide precursors. This contrasts the number of MS/MS scans from a DDA experiment (a) a standard 30Da window SWATH-MS experiment (b) and two larger window sizes for the SWATH-MS experiment, 200Da (c) and 600Da (d).

While proceeding with 600Da windows for MS/MS fragmentation enables more quantitative scans, there is a non-trivial question of which fragment ions to use for a glycan motif measure. Following a glycosylation site from human IgG1 derived from donated IVIG plasma using XL-DIA, we can see that the neutral species of glycans will co-elute around a 15-18 retention time (Fig 3.7a-b). Many precursor ions will be gathered together for MS/MS fragmentation in a 600Da window (Fig 3.7c). Two cycles of 600Da acquisitions are required to cover our glycopeptide-specific mass range of 600-1800m/z. Interestingly, most precursor ions fall into the first window (Fig. 3.7c); however, a non-zero amount of precursor ions do fall in the second 600Da cycle as seen by the fragment ions measured for each window (Fig. 3.7d-e). For this study we

wanted to track the glycan features corresponding to fucosylation, galactosylation, sialylation, and high mannose. In order to link MS/MS measured fragment ions to these aforementioned glycan features we implemented a substructure motif networking approach from our previous work on the glycomics analysis tool GlyCompare (Bao & Kellman et al., 2021). First all glycans which appear on our annotated glycopeptides from the annotation propagation algorithm we harvested for each IgG subtype. Separately each IgG subtype glycan population was decomposed into all possible compositional substructures and linked together if the substructure was a direct descendant of the parent node as defined in GlyCompare. For example, human IgG1 contained 30 FC region glycans annotated from Byonic (blue) and 30 additional glycans after the annotation propagation (red). Additionally, the substructure network was augmented with 19 glycan structures (light blue) that would complete the substructure decomposition to the lone HexNAc(1) on the peptide backbone (Fig 3.8). The resulting directed acyclic graph can be used to theoretically fragment any glycan from a precursor ion derived from the FC region of human IgG1. This is done by finding the precursor ion's glycan and traveling down the network towards the final leaf (HexNAc(1)). Each node that is traveled through is given the corresponding ionization traits (possible charge states, associated oxonium ions etc.) and stored in a theoretical spectral library. Then this theoretical spectral library is augmented with experimentally derived fragmentation patterns from our DDA pooled dataset (80 human file, 51 mouse files) and the final measurable fragment ions remain. Each fragment ion is then linked to a glycan feature by a user defined rule and the corresponding fragment ions are then considered to be representative of that glycan feature (see Materials and Methods). For example, when comparing Avastin to IgG1 from IVIG therapy, fragment ions corresponding to galactosylation will be lowered due to less galactosylated glycans in the FC region of Avastin (Fig. 3.9a).

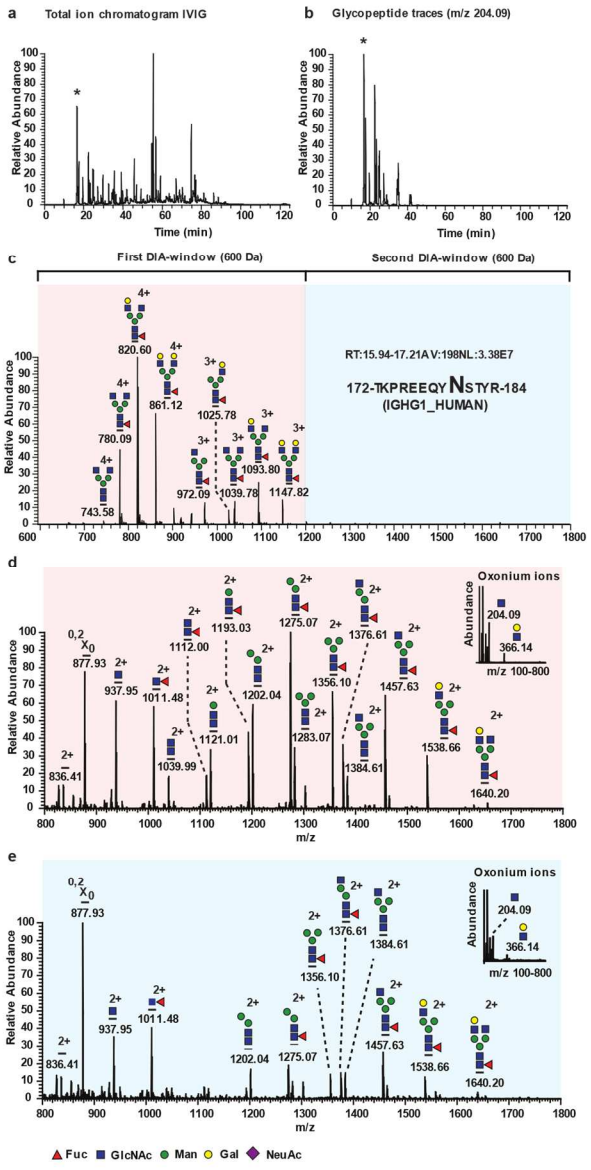


Figure 3. 7. A 2 cycle 600Da fragmentation window for SWATH-MS properly captures glycopeptides features in XL-DIA

The total ion chromatogram displays the relative intensity of biomolecules eluting within a retention time window (a). By tracking the 204.09 oxonium ion indicating a glycopeptide is within the peak we can see that the typical glycopeptides derived from IgG1 elute off of our column between 10-40 minutes (b). The neutral glycan species peak is denoted by the star. By implementing a 2 cycle 600Da window scheme for SWATH-MS we can see that a variety of neutral glycan species precursor ions are captured for fragmentation in the first window (c). The MS/MS of these precursor ions gives rise to a mixture of overlapping fragment ions with different glycan features and distinct m/z (d). Additionally, the second 600Da also measures a set of overlapping fragment ions as some neutral glycan species precursor ions may appear with a higher m/z (e).

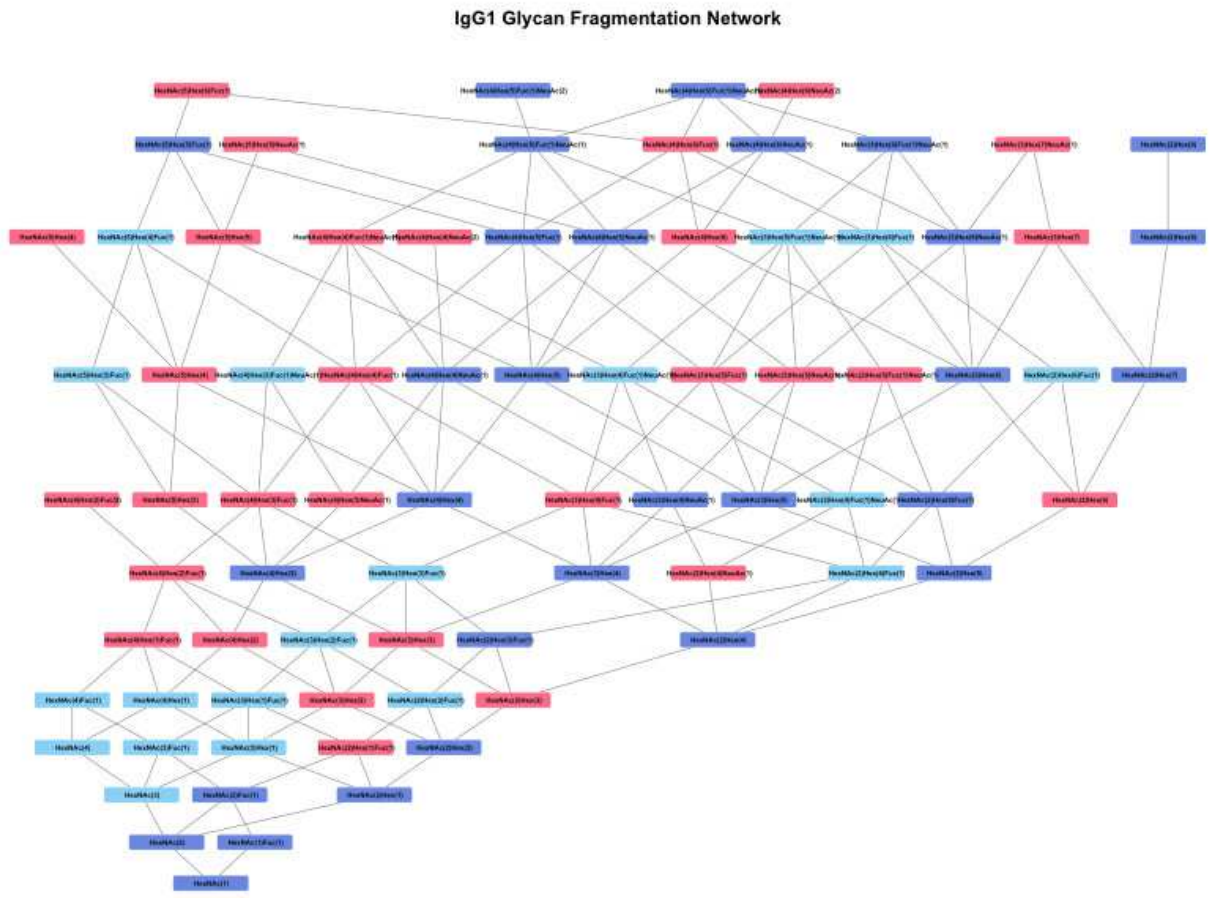


Figure 3. 8. Leveraging biosynthetic glycosylation rules for possible IgG1 FC region glycans results in a graph of fragmentation

Construction of a directed acyclic graph of possible glycan fragmentation was done with the help of the glycomics analysis tool GlyCompare. GlyCompare took the glycan list from the precursor ions available in the FC region of human IgG1 and through motif decomposition built the substructure network. Visualized in color are the origins of the nodes in the network. Blue indicates an original Byonic identified IgG1 FC region glycan, red indicates a glycan identification after annotation propagation, and light blue indicates a glycan addition by GlyCompare to complete the network. The network concludes with the HexNAc(1) glycan that can appear as the most truncated glycan at this specific peptide backbone.

Many different glycan motifs are known to function as components of biological processes involving proteins; however, it is challenging to robustly link a glycan motif to a function because glycans are so densely related in structure. To create an independent measure of glycan motifs we have used the XL-DIA method to directly measure the glycan motifs coming from multiple glycans

that reside at a specific glycosylation site. For human IgG1, our feature-centric spectral library for fragment identification and feature quantification contained 20 unique fragments measuring fucosylation, 38 unique fragments for galactosylation, 10 unique fragments for sialylation, and 2 unique fragments for high mannose structures. For human IgG2, our feature-centric spectral library for fragment identification and feature quantification contained 16 unique fragments measuring fucosylation, 26 unique fragments for galactosylation, 26 unique fragments for sialylation, and 1 unique fragment for high mannose structures. For human IgG3/4, our feature-centric spectral library for fragment identification and feature quantification contained 11 unique fragments measuring fucosylation, 38 unique fragments for galactosylation, 15 unique fragments for sialylation, and 2 unique fragments for high mannose structures. Using the fragment ions derived from the aforementioned fragments we measured MS2 quantification of fucosylation, galactosylation, sialylation, and high mannose. This population-based approach was compatible with the assumptions of standard statistical analysis for differential abundance analysis between samples. We found that glycan feature quantification using MS/MS scans closely correlated with the percent contribution of glycopeptides to features found from MS survey scans (Fig 3.9b). As expected IgG1 derived from human IVIG contained more glycopeptides with glycosylation and sialylation than the other IgG1 therapeutics, while Erbitux had higher FC region high mannose structures. Additionally, using MSstats, an R package for statistical analysis of quantitative mass spectrometry data, it was found that statistically significant differential analysis can be performed between the MS2 feature level measurements (Fig. 9c).

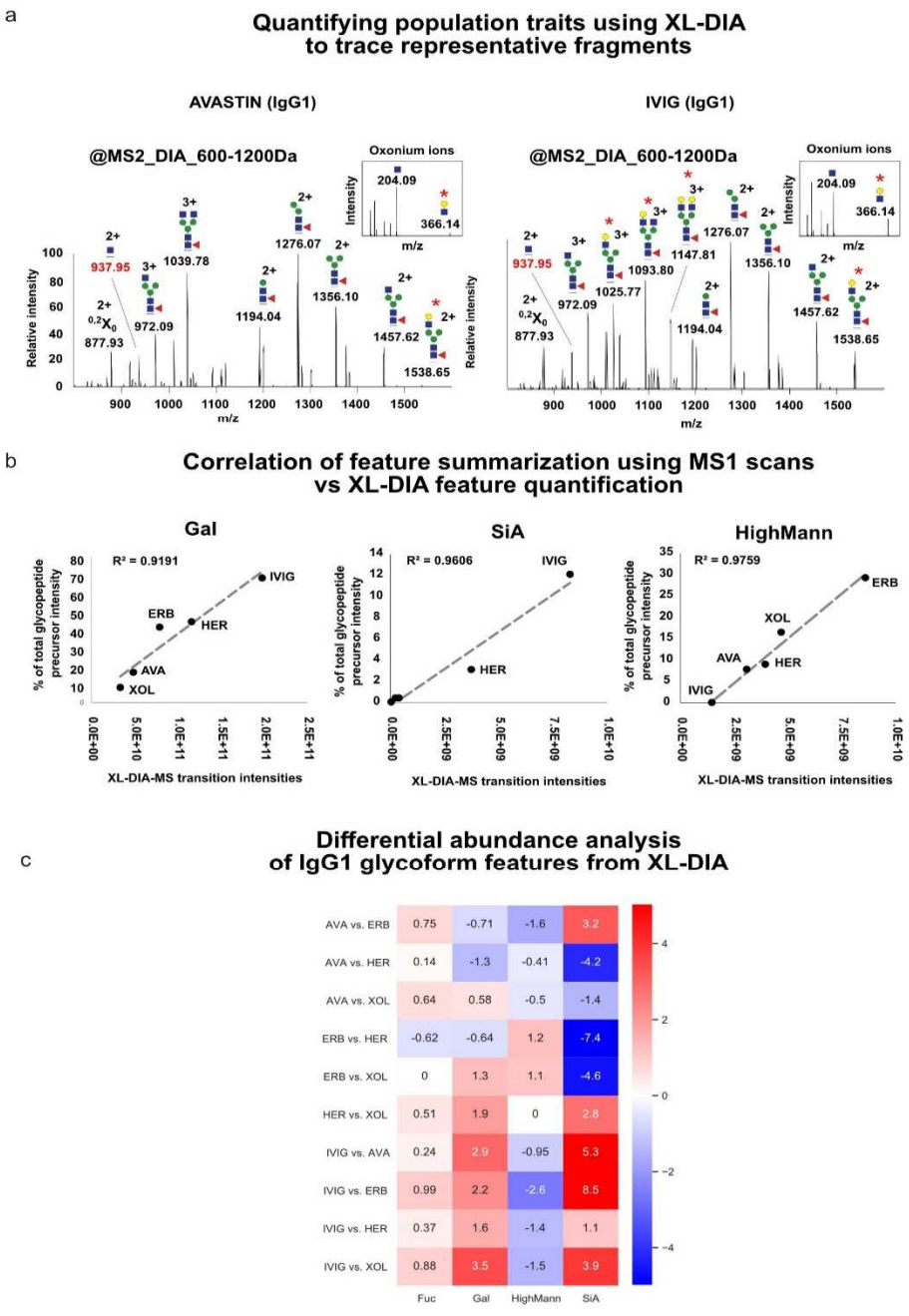


Figure 3. 9. MS2 level quantification of glycan features are robust, and measures maintain statistical independence

Using XL-DIA to compare the 600-1200Da window of fragmentation for neutral glycan species of either commercially available Avastin or IVIG derived IgG1 shows that several more glycopeptide fragment ions contain galactosylation in IVIG derived IgG1 (a). Quantified glycan features using MS2 measures correlated with MS1 measured % contribution to total glycopeptide TIC intensity (b). MS2 measures of glycan features from XL-DIA experiment of various human IgG1 compared using differential abundance analysis in MSstats (c).

Finally, the XL-DIA methodology and data analysis presented here off the bonus of still maintaining the MS survey scans in order to perform abundant glycopeptide quantification on MS1 measures. Using the human IgG1 precursor ion list we have quantified the glycopeptides from the FC region in a DDA experiment and an XL-DIA experiment to compare the results. 25 glycopeptides reached a threshold of TIC intensity greater than 1% in the DDA while 20 glycopeptides reached the threshold in the XL-DIA experiment (Fig. 3.10a). However, in both experiments the same unique glycopeptides were found, but possibly due to better coverage from more survey scans in the XL-DIA experiment the percent contribution of glycopeptide to the total amount of glycopeptide measured was different. It was also seen that correlation of major glycopeptides in each experimental setup was very high. For example both the high mannose structure HexNAc(2)Hex(5) and the galactosylated and fucosylated structure HexNAc(4)Hex(4)Fuc(1) had R-squared values above 0.8 (Fig. 10b).

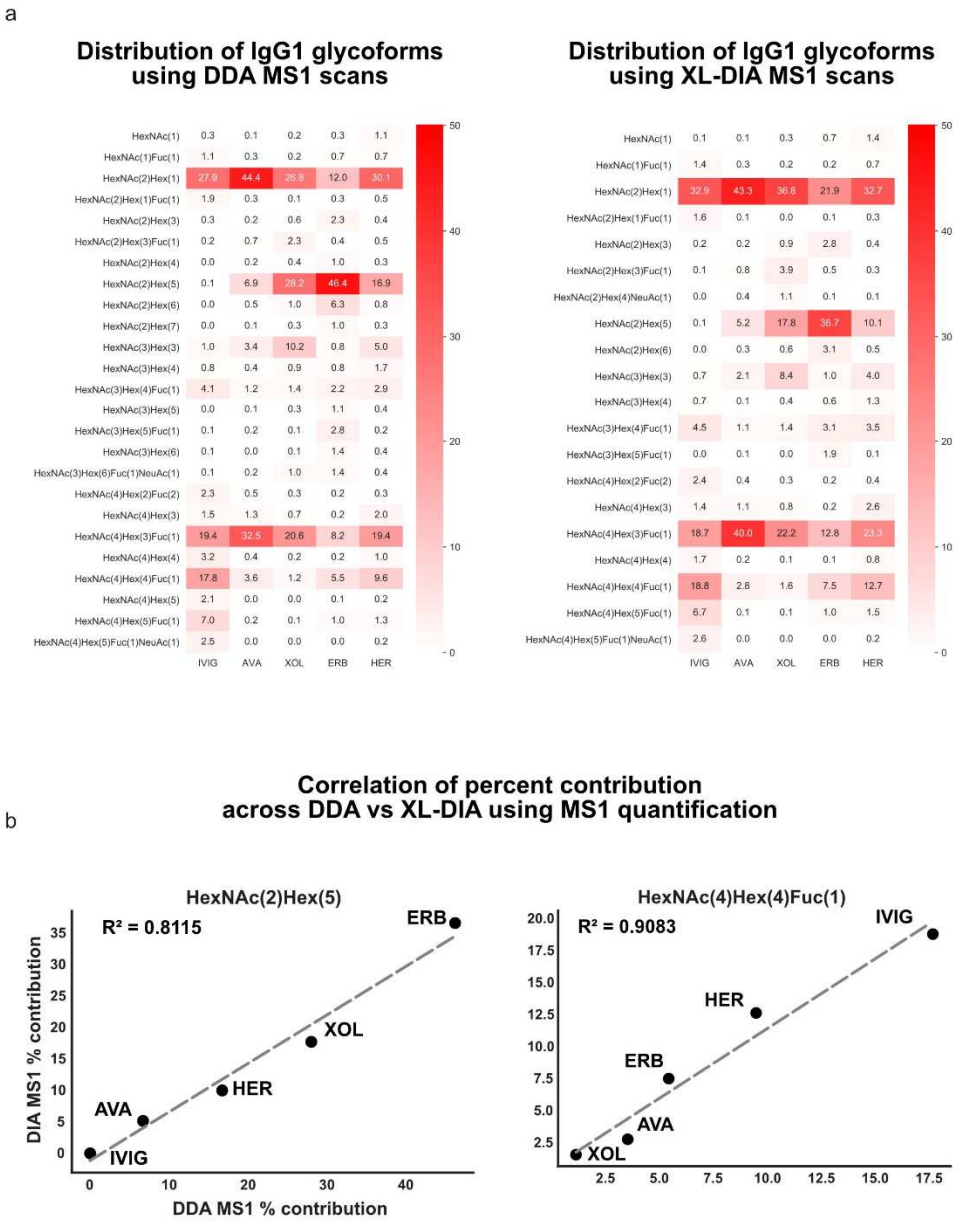


Figure 3. 10. XL-DIA maintains useful MS1 survey scans for unique glycopeptide quantitation

Comparative analysis of glycopeptides derived from the FC region of IgG1 from 5 different production systems shows that DDA MS1 measures and XL-DIA MS1 measures are comparable (a). Two major glycopeptides are showcased as having high correlation between % contribution to TIC intensity from survey scans from both DDA and DIA experiments (b).

Group C/G human septic blood plasma has unique glycan features of IgG FC region. Next, we applied our XL-DIA methodology to 16 human group C/G streptococcus sepsis patients' plasma (32 samples). First IgG pulldowns were done to ensure robust measurement of patient IgG populations in both acute phase plasma taken at time of hospitalization and convalescent plasma take 2 weeks are recovery. Using 600Da windows for MS/MS acquisition we thousands of scans for 204.09 containing elution peaks at retention times across 5 to 25 minutes as is the expected IgG glycopeptide window in our setup. The data was then interrogated in Skyline on molecule mode using our hybrid feature-centric spectral libraries for IgGs. All fragment ions intensities corresponding to glycan features defined in the library we collected over the 5-to-25-minute retention time window. For each packet of collected fragment ion intensities, we summed their amounts across the 2 cycles of fragmentation (600-1200 m/z window and 1200-1800 m/z window). Then these values we collapsed into feature intensities by summing the unique fragment ions which represent each feature. Finally, these feature intensities were normalized using dataProcess function in MSstats set to 'globalStandards' using the HexNAc(1) glycopeptide as the global standard across each sample. We were able to quantify fucosylation, galactosylation, sialyation, and high mannose structures across IgG1, IgG2, and IgG3/4 from each of these samples (Fig. 3.11a). By investigating the glycan feature patterns across acute phase plasma and convalescent plasma a sub-population of patients was identified that had an increase in galactosylation and sialyation in the IgG1 FC region after recovery (Fig. 3.11b). This patient sub-population, which consisted of 6 of the 16 group C/G sepsis patients, was analyzed for glycan feature patterns across IgG1, IgG2, and IgG3/4. Principal components analysis was used to investigate how well the variation in IgG FC region glycosylation could separate this sub-population. IgG1, IgG2, and IgG3/4 FC regions all showed independent power to identify these patients (Fig. 3.12a-c).

Interestingly, the FC region glycan feature distribution is not consistent across these IgG subtypes (Fig. 3.11a) which raises more questions about the fine-tuning of this region during sepsis.

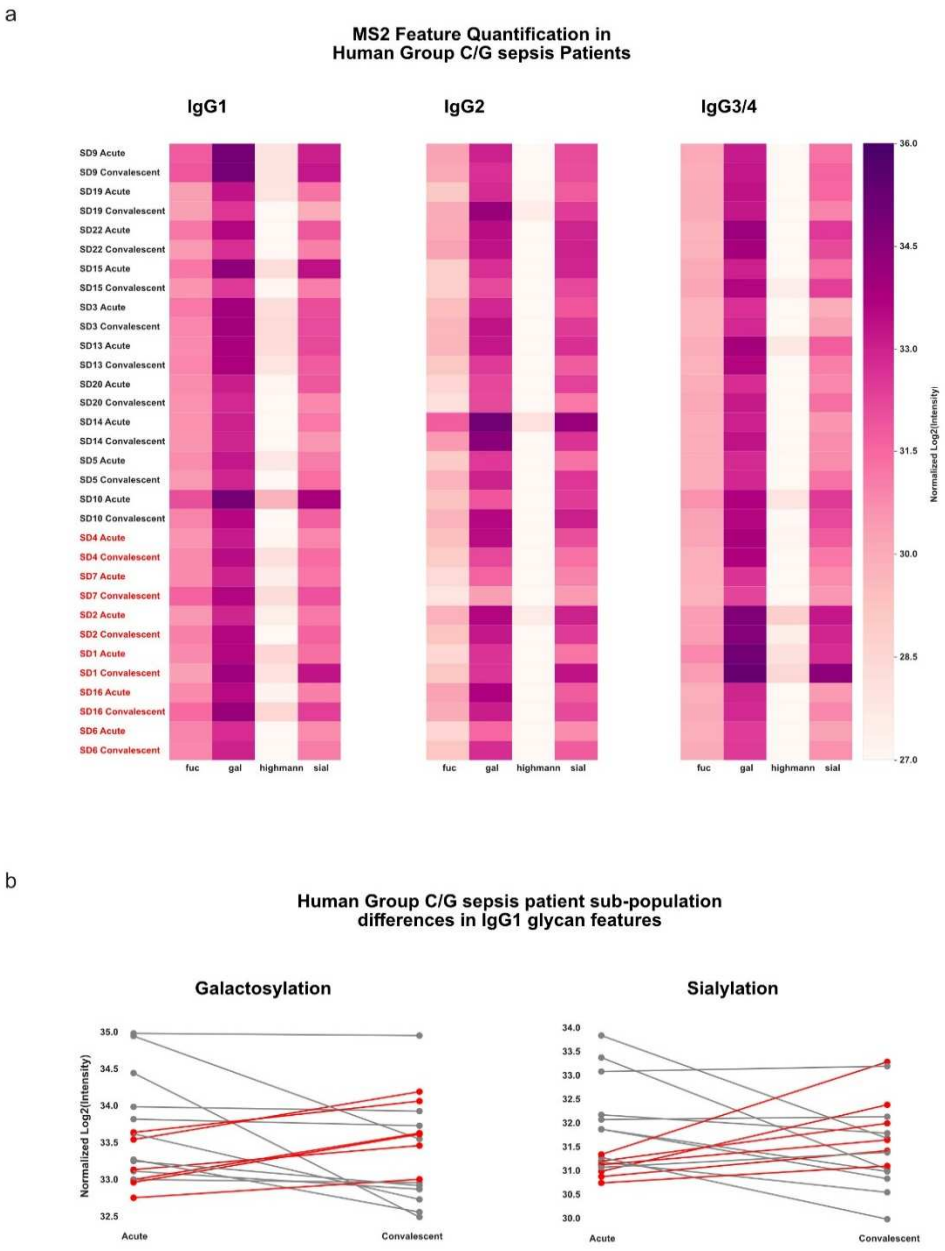


Figure 3. 11. XL-DIA analysis of human group C/G streptococcus sepsis patient plasma from acute phase and convalescent plasma

Human septic patients had blood plasma collected during hospitalization with group C/G streptococcus sepsis and 2 weeks after recovery. IgG subtype specific FC region glycosylation patterns as tracked by mass spectrometry reveal a large heterogeneous population of modifications (a). A sub-population of patients were identified that had an increase in both galactosylation and sialylation from acute phase to convalescent plasma (b). Sub-population marked in red identifiers.

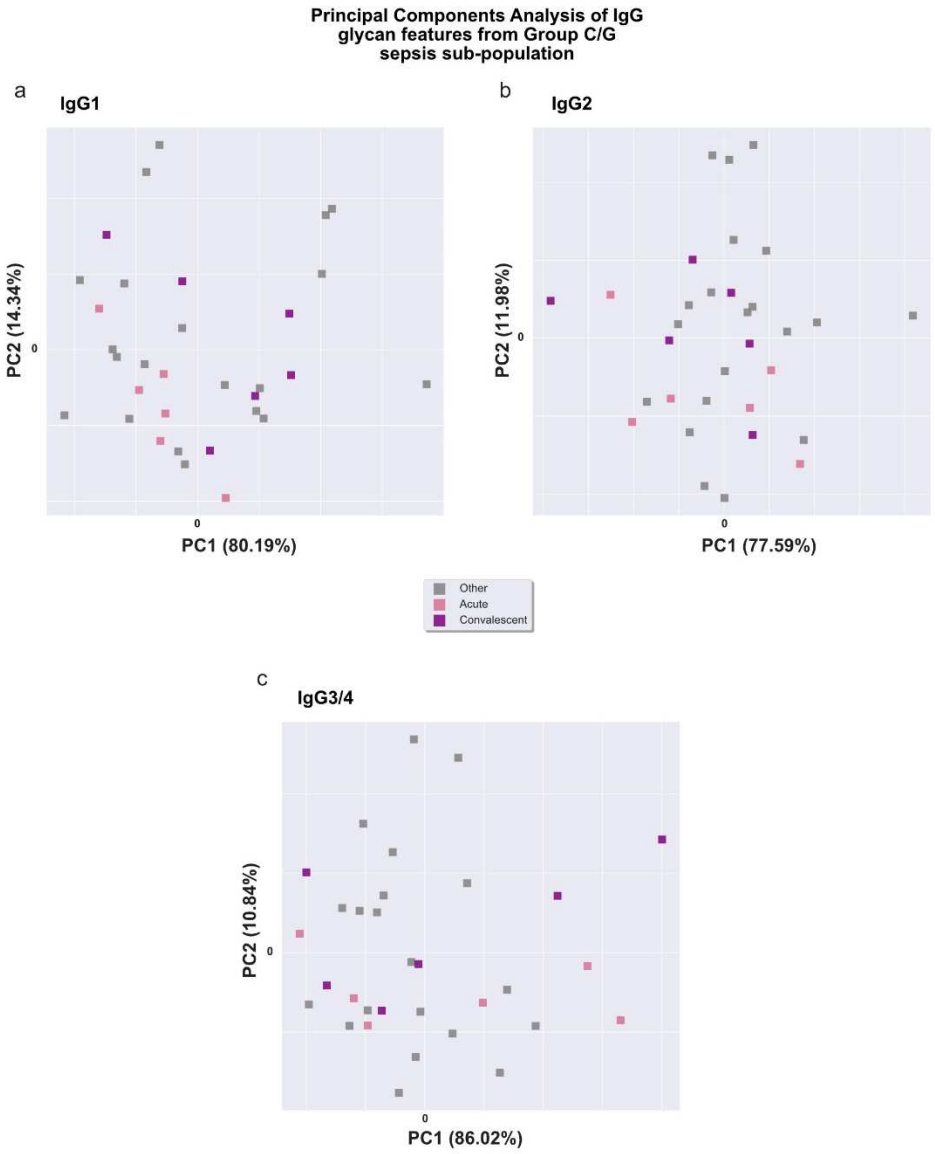


Figure 3. 12. Principal components analysis of XL-DIA glycan feature quantification from human group C/G streptococcus patient plasma

Each glycan feature pattern from the FC region of IgG subtypes was subject to principal components analysis. For IgG1 (a), IgG2 (b), and IgG3/4 (c), the patient sub-population defined by increased galactosylation and sialylation in convalescent plasma were colored depending on sample derivation (pink and purple) while others were not (gray).

3.5 Discussion

In this investigation of the glycosylation patterns in the FC region of IgG subtypes using mass spectrometry-based approaches, we have successfully shown that leveraging a large, pooled dataset of DDA LC-MS/MS experiments targeted at IgG in several conditions can broaden the measurable scope of glycans that appear in this region. Additionally, we have shown that opening the isolation window during SWATH-DIA glycoproteomics can capture the fragment ions related to several functional glycan features as independent measures and raise the measurable intensity of these ions. By implementing a molecular networking protocol, we have roughly doubled the identification power of our DDA pooled dataset in its ability to annotate spectral clusters and relationships. Finally, we have shown that our XL-DIA method and data analysis is comparable to traditional glycopeptide quantification and expands our understanding of glycan heterogeneity in multiple datasets.

Recent advances in glycomics and glycoproteomics has deduced that glycan features, either defined as compositional traits or terminal motifs, can appear with differing frequency across various biological samples (Bao & Kellman et al., 2021). The fine-tuned regulatory mechanisms that a cell can exhibit in homeostatic conditions or under distress gives rise to the context specific PTMs that many studies have explored for insight and utility. Specifically, glycan biosynthesis can be perturbed during cellular response and can result in changes to micro-, macro-, and meta-heterogeneity. As demonstrated here, classical glycan identification tools can benefit greatly from being presented with a diverse set of biological samples directed at perturbed glycan biosynthesis.

Spectral similarities defined by highly related MS/MS patterns have illuminated many features of complex data that comes from LC-MS/MS studies. Here we present a workflow adapted

and customized to work in tandem with glycopeptide identification done by Byonic and molecular networking facilitated by the GNPS server. We have employed a propagation algorithm that has built on our previous work for novel node annotation in molecular networks for glycoproteomics spectra. Our annotation propagation algorithm offers a flexible range of user defined parameters for acceptable mass differences and the addition or subtraction of a variable number of monosaccharides. Additionally, the annotated spectra can be transformed to a spectral library useful for the identification and quantification of glycopeptides in DDA and DIA datasets. This spectral library is the product of combining a theoretically generated fragmentation library with experimentally observed MS/MS patterns. Finally, these spectral relationships allow us to measure glycan features using SWATH-MS. These advancements help combat the challenges presented in glycoproteomics ranging from the scope of site-specific glycosylation to the statistical interpretation of the data.

While our annotation propagation algorithm takes the possible biosynthetic steps into account when transferring glycopeptide annotations to new spectral clusters it can be theorized that not all proposed glycans will appear in the wild. Biological constraints, such as the structural context of the peptide backbone for which a glycan is placed do limit the glycan possibilities. For example, location of □-sheets and α-helices flanking a glycopeptide backbone can impact the glycosylation of that sequence in a model of bacterial N-glycosylation (Silverman and Imperiali, 2016). While these statistical tendencies can be derived, a certain level of manual structure curation will always be needed; however, leveraging big glycomics data and distribution based structural constraints can offer assistance. Large data repositories documenting site-specific glycosylation events and references to possible glycan structures viewed experimentally contain a vast wealth of knowledge that can help streamline the annotation propagation and supplement the label transfer

with measures of likelihood. Databases such as GlyConnect and UniCarbKB contain thousands of topological glycan tendencies and associated peptide backbones (Alocci et al., 2019, Campbell et al., 2014). Searchable knowledgebases such as GlyGen (York et al., 2020) have begun to sprout up that make all of this information accessible programmatically through Application Programming Interfaces (APIs). It would therefore be an exciting expansion to integrate biosynthetic rules and structural tendencies when scoring the possibility of a glycopeptide annotation transfer in a molecular network.

Another advancement brought by using the XL-DIA method is the direct measure of glycan properties, either defined as topological motifs, features, or substructures, to account for the shared similarity between glycans when performing statistical tests based on the assumption of independence. Due to the hierarchical nature of glycan biosynthesis, a sequential network of additions and subtractions to build a glycan with the ability to end at any point, full length glycopeptides share a heavy dependance assumption on each other. This means that when the elimination of one full length glycopeptide takes place, either through glycoengineering or in response to stimuli, the impact is distributed in the abundance of all other possible glycopeptides. Current work in this field treats each glycan as an independent unit which ignores the property of shared features among glycans as explained above. This view is restricted to correlation analysis between glycan pairs with very similar structures, or a local view. Here we present a new way to measure and analyze differences between each glycan. This analysis is better able to capture global trends in the possible glycan landscape by aggregating the interrelation between glycopeptides with similar structure. While this adaptation may be made in traditional glycoproteomics by transforming the measures into a % contribution of glycopeptides to a glycan feature, it is valuable to be able to make this measure directly using XL-DIA. This direct measure of glycan features

may also be useful in our previous work estimating the trends in glycan biosynthesis flux us glycan substructure decomposition (Bao & Kellman et al., 2021). There we had explored the product-substrate ratios of glycan biosynthesis networks through inference of substructure abundance coming from full length glycan measures. It is conceivable that direct measurements of substructures at the MS2 level would substantiate these analyses.

Finally, in an exploratory analysis of human group C/G streptococcus sepsis patients we have displayed tat XL-DIA and find subtle differences the glycosylation patterns in the FC region of IgGs in acute phase and convalescent plasma. Sepsis is a complex disease that suffers greatly in the identification of robust molecular markers that can subtype patient populations that would benefit from specific therapeutic intervention. With glycosylation being an immense and impactful biological phenomenon, it is a natural avenue for the search for bio identifiers. Interestingly, antibody glycosylation patterns are altered during the onset of the disease and may maintain a differential composition even after the patient has recovered. There is very little known about the implications of an altered plasma glycome and how it can affect future infections and challenges. We hope that by expanding our ability to measure the scope of glycoprotein heterogeneity we can use circulating glycoprotein measures as tools for stratifying patients and inspire others to explore further.

3.6 Acknowledgements

This research was supported by grant P01HL131474 to J.D.E. and by funding from NIGMS (R35 GM119850 N.E.L., T32GM008806 J.T.S) and the Novo Nordisk Foundation (NNF20SA0066621 N.E.L.). AGT, CK, SC and JM were supported by the Wallenberg foundation (WAF grant number 2017.0271), the Swedish research council (grant number 2019-01646 and 2018-05795), the Viral and Bacterial Adhesin Network Training (ViBrANT) Program funded by

the European Union's HORIZON 2020 Research and Innovation Program under the Marie Sklodowska-Curie Grant Agreement No 765042 and Alfred Österlunds Foundation.

3.7 Author Information

J.T.S. and A.G.T. conceived the project. N.E.L and J.M. supervised the research. A.G.T. conducted the sample preparation, IgG purification, mouse infections, and mass spectrometry experiments. J.T.S. and A.G.T. analyzed the DDA glycoproteomics datasets. J.T.S. and A.G.T. conceived the molecular networking and XL-DIA methodology. J.T.S. implemented the annotation propagation algorithm. J.T.S. developed the hybrid spectral libraries. J.T.S. and A.G.T. analyzed the XL-DIA data. J.T.S., A.G.T., J.D.E., J.M., and N.E.L. interpreted the data and wrote the manuscript with significant input from all authors.

Chapter 3 is adapted from "Quantifying glycopeptides and assessing the heterogeneity of glycosylation in circulating glycoproteins", a manuscript in preparation by James T Sorrentino, Alejandro Gomez Toledo, Jeffrey D Esko, Johan Malmström, and Nathan E Lewis. J.T.S. and A.G.T. contributed equally. The dissertation author was a primary investigator and author on this paper.

References

Ackerman, M. E., Crispin, M., Yu, X., Baruah, K., Boesch, A. W., Harvey, D. J., Dugast, A. S., Heizen, E. L., Ercan, A., Choi, I., Streeck, H., Nigrovic, P. A., Bailey-Kellogg, C., Scanlan, C., & Alter, G. (2013). Natural variation in Fc glycosylation of HIV-specific antibodies impacts antiviral activity. The Journal of Clinical Investigation, 123(5), 2183–2192. https://doi.org/10.1172/JCI65708

Adams, K. J., Pratt, B., Bose, N., Dubois, L. G., St. John-Williams, L., Perrott, K. M., Ky, K., Kapahi, P., Sharma, V., Maccoss, M. J., Moseley, M. A., Colton, C. A., Maclean, B. X., Schilling, B., & Thompson, J. W. (2020). Skyline for Small Molecules: A Unifying Software Package for Quantitative Metabolomics. Journal of Proteome Research, 19(4), 1447–1458. https://doi.org/10.1021/ACS.JPROTEOME.9B00640

Aird, W. C. (2007). Phenotypic Heterogeneity of the Endothelium. Circulation Research, 100(2), 158–173. https://doi.org/10.1161/01.RES.0000255691.76142.4A

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). The Adaptive Immune System. https://www.ncbi.nlm.nih.gov/books/NBK21070/

Allhorn, M., Olin, A. I., Nimmerjahn, F., & Colin, M. (2008). Human IgG/FcγR Interactions Are Modulated by Streptococcal IgG Glycan Hydrolysis. PLOS ONE, 3(1), e1413. https://doi.org/10.1371/JOURNAL.PONE.0001413

Almizraq, R. J., & Branch, D. R. (2021). Efficacy and mechanism of intravenous immunoglobulin treatment for immune thrombocytopenia in adults. Annals of Blood, 6(0). https://doi.org/10.21037/AOB-20-87

Alocci, D., Mariethoz, J., Gastaldello, A., Gasteiger, E., Karlsson, N. G., Kolarich, D., Packer, N. H., & Lisacek, F. (2019). GlyConnect: Glycoproteomics Goes Visual, Interactive, and Analytical. Journal of Proteome Research, 18(2), 664–677. https://doi.org/10.1021/ACS.JPROTEOME.8B00766

Al-Sharif, A., Jamal, M., Zhang, L. X., Larson, K., Schmidt, T. A., Jay, G. D., & Elsaid, K. A. (2015). Lubricin/proteoglycan 4 binding to CD44 receptor: A mechanism of the suppression of proinflammatory cytokine-induced synoviocyte proliferation by lubricin. Arthritis and Rheumatology, 67(6), 1503–1513. https://doi.org/10.1002/ART.39087

Amann, T., Hansen, A. H., Kol, S., Hansen, H. G., Arnsdorf, J., Nallapareddy, S., Voldborg, B., Lee, G. M., Andersen, M. R., & Kildegaard, H. F. (2019). Glyco-engineered CHO cell lines producing alpha-1-antitrypsin and C1 esterase inhibitor with fully humanized N-glycosylation profiles. Metabolic Engineering, 52, 143–152. https://doi.org/10.1016/J.YMBEN.2018.11.014

Bandeira, N., Tsur, D., Frank, A., & Pevzner, P. A. (2007). Protein identification by spectral networks analysis. Proceedings of the National Academy of Sciences of the United States of America, 104(15), 6140–6145. https://doi.org/10.1073/PNAS.0701130104

Bao, B., Kellman, B. P., Chiang, A. W. T., Zhang, Y., Sorrentino, J. T., York, A. K., Mohammad, M. A., Haymond, M. W., Bode, L., & Lewis, N. E. (2021). Correcting for

sparsity and interdependence in glycomics by accounting for glycan biosynthesis. Nature Communications 2021 12:1, 12(1), 1–14. https://doi.org/10.1038/s41467-021-25183-5

Bessen, D. E., McShan, W. M., Nguyen, S. V., Shetty, A., Agrawal, S., & Tettelin, H. (2015). Molecular Epidemiology and Genomics of Group A Streptococcus. Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases, 33, 393. https://doi.org/10.1016/J.MEEGID.2014.10.011

Bessen, D. E., Sotir, C. M., Readdy, T. L., & Hollingshead, S. K. (1996). Genetic correlates of throat and skin isolates of group A streptococci. The Journal of Infectious Diseases, 173(4), 896–900. https://doi.org/10.1093/INFDIS/173.4.896

Beutler, B., & Rietschel, E. T. (2003). Innate immune sensing and its roots: The story of endotoxin. Nature Reviews Immunology, 3(2), 169–176. https://doi.org/10.1038/nri1004

Bläckberg, A., de Neergaard, T., Frick, I. M., Nordenfelt, P., Lood, R., & Rasmussen, M. (2021). Lack of Opsonic Antibody Responses to Invasive Infections With Streptococcus dysgalactiae. Frontiers in Microbiology, 12. https://doi.org/10.3389/FMICB.2021.635591

Bollineni, R. C., Koehler, C. J., Gislefoss, R. E., Anonsen, J. H., & Thiede, B. (2018). Large-scale intact glycopeptide identification by Mascot database search. Scientific Reports 2018 8:1, 8(1), 1–13. https://doi.org/10.1038/s41598-018-20331-2

Bone, R. C., Balk, R. A., Cerra, F. B., Dellinger, R. P., Fein, A. M., Knaus, W. A., Schein, R. M. H., & Sibbald, W. J. (1992). Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. Chest, 101(6), 1644–1655. https://doi.org/10.1378/chest.101.6.1644

Bone, R. C., Fisher, C. J., Clemmer, T. P., Slotman, G. J., Metz, G. A., & Balk, R. A. (1989). Sepsis syndrome: A valid clinical entity. Critical Care Medicine, 17(5), 389–393. https://doi.org/10.1097/00003246-198905000-00002

Bouslimani, A., Sanchez, L. M., Garg, N., & Dorrestein, P. C. (2014). Mass spectrometry of natural products: current, emerging and future technologies. Nat. Prod. Rep., 31(6), 718–729. https://doi.org/10.1039/c4np00044g

Bruderer, R., Bernhardt, O. M., Gandhi, T., Xuan, Y., Sondermann, J., Schmidt, M., Gomez-Varela, D., & Reiter, L. (2017). Optimization of experimental parameters in data-independent mass spectrometry significantly increases depth and reproducibility of results. Molecular and Cellular Proteomics, 16(12), 2296–2309. https://doi.org/10.1074/mcp.RA117.000314

Campbell, M. P., Peterson, R., Mariethoz, J., Gasteiger, E., Akune, Y., Aoki-Kinoshita, K. F., Lisacek, F., & Packer, N. H. (2014). UniCarbKB: building a knowledge platform for glycoproteomics. Nucleic Acids Research, 42(Database issue). https://doi.org/10.1093/NAR/GKT1128

Carapetis, J. R., Steer, A. C., Mulholland, E. K., & Weber, M. (2005). The global burden of group A streptococcal diseases. The Lancet. Infectious Diseases, 5(11), 685–694. https://doi.org/10.1016/S1473-3099(05)70267-X

Castro, S. A., & Dorfmueller, H. C. (2021). A brief review on Group A Streptococcus pathogenesis and vaccine development. Royal Society Open Science, 8(3). https://doi.org/10.1098/RSOS.201991

Čaval, T., Heck, A. J. R., & Reiding, K. R. (2021). Meta-heterogeneity: Evaluating and Describing the Diversity in Glycosylation Between Sites on the Same Glycoprotein. Molecular & Cellular Proteomics : MCP, 20, 100010. https://doi.org/10.1074/MCP.R120.002093

Chandler, K. B., Mehta, N., Leon, D. R., Suscovich, T. J., Alter, G., & Costello, C. E. (2019). Multi-isotype Glycoproteomic Characterization of Serum Antibody Heavy Chains Reveals Isotype- and Subclass-Specific N-Glycosylation Profiles. Molecular & Cellular Proteomics : MCP, 18(4), 686–703. https://doi.org/10.1074/MCP.RA118.001185

Choi, M., Chang, C. Y., Clough, T., Broudy, D., Killeen, T., MacLean, B., & Vitek, O. (2014). MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. Bioinformatics (Oxford, England), 30(17), 2524–2526. https://doi.org/10.1093/BIOINFORMATICS/BTU305

Cleuren, A. C. A., van der Ent, M. A., Jiang, H., Hunker, K. L., Yee, A., Siemieniak, D. R., Molema, G., Aird, W. C., Ganesh, S. K., & Ginsburg, D. (2019). The in vivo endothelial cell translatome is highly heterogeneous across vascular beds. Proceedings of the National Academy of Sciences of the United States of America, 116(47), 23618–23624. https://doi.org/10.1073/PNAS.1912409116

Cole, J. N., Barnett, T. C., Nizet, V., & Walker, M. J. (2011). Molecular insight into invasive group A streptococcal disease. Nature Reviews. Microbiology, 9(10), 724–736. https://doi.org/10.1038/NRMICRO2648

Collin, M., & Olsén, A. (2001). EndoS, a novel secreted protein from Streptococcus pyogenes with endoglycosidase activity on human IgG. The EMBO Journal, 20(12), 3046. https://doi.org/10.1093/EMBOJ/20.12.3046

Collin, M., & Olsén, A. (2001). Effect of SpeB and EndoS from Streptococcus pyogenes on human immunoglobulins. Infection and Immunity, 69(11), 7187–7189. https://doi.org/10.1128/IAI.69.11.7187-7189.2001/ASSET/E10BF2F2-D6FF-4627-9DC5-7C4928A8D045/ASSETS/GRAPHIC/II1110721002.JPEG

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., & Mann, M. (2011). Andromeda: A peptide search engine integrated into the MaxQuant environment. Journal of Proteome Research, 10(4), 1794–1805. https://doi.org/10.1021/PR101065J/SUPPL_FILE/PR101065J_SI_002.ZIP

Dallas, D. C., Martin, W. F., Hua, S., & Bruce German, J. (2013). Automated glycopeptide analysis—review of current state and future directions. Briefings in Bioinformatics, 14(3), 361–374. https://doi.org/10.1093/BIB/BBS045

Das, N., Schmidt, T. A., Krawetz, R. J., & Dufour, A. (2019). Proteoglycan 4: From Mere Lubricant to Regulator of Tissue Homeostasis and Inflammation: Does proteoglycan 4 have the ability to buffer the inflammatory response? BioEssays, 41(1). https://doi.org/10.1002/BIES.201800166

De Backer, D., Cortes, D. O., Donadello, K., & Vincent, J. L. (2014). Pathophysiology of microcirculatory dysfunction and the pathogenesis of septic shock. Virulence, 5(1), 73–79. https://doi.org/10.4161/viru.26482

de Neergaard, T., Bläckberg, A., Ivarsson, H., Thomasson, S., Kumra Ahnlide, V., Chowdhury, S., Khakzad, H., Bahnan, W., Malmström, J., Rasmussen, M., & Nordenfelt, P. (2022). Invasive Streptococcal Infection Can Lead to the Generation of Cross-Strain Opsonic Antibodies. Microbiology Spectrum, 10(6). https://doi.org/10.1128/SPECTRUM.02486-22

Delano, M. J., & Ward, P. A. (2016). Sepsis-induced immune dysfunction: can immune therapies reduce mortality? J. Clin. Invest., 126(1), 23–31. https://doi.org/10.1172/jci82224

Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S., & Ralser, M. (2020). DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. Nature Methods, 17(1), 41–44. https://doi.org/10.1038/S41592-019-0638-X

Desiere, F., Deutsch, E. W., King, N. L., Nesvizhskii, A. I., Mallick, P., Eng, J., Chen, S., Eddes, J., Loevenich, S. N., & Aebersold, R. (2006). The PeptideAtlas project. Nucleic Acids Research, 34(Database issue). https://doi.org/10.1093/NAR/GKJ040

Doron, S., & Gorbach, S. L. (2008). Bacterial Infections: Overview. International Encyclopedia of Public Health, 273. https://doi.org/10.1016/B978-012373960-5.00596-7

Durr, E., Yu, J., Krasinska, K. M., Carver, L. A., Yates, J. R., Testa, J. E., Oh, P., & Schnitzer, J. E. (2004). Direct proteomic mapping of the lung microvascular endothelial cell surface in vivo and in cell culture. Nature Biotechnology, 22(8), 985–992. https://doi.org/10.1038/NBT993

Engelmann, B., & Massberg, S. (2013). Thrombosis as an intravascular effector of innate immunity. Nature Reviews Immunology, 13(1), 34–45. https://doi.org/10.1038/nri3345

Eon-Duval, A., Broly, H., & Gleixner, R. (2012). Quality attributes of recombinant therapeutic proteins: An assessment of impact on safety and efficacy as part of a quality by design development approach. Biotechnology Progress, 28(3), 608–622. https://doi.org/10.1002/BTPR.1548

Ferrara, C., Grau, S., Jaĝer, C., Sondermann, P., Bruñker, P., Waldhauer, I., Hennig, M., Ruf, A., Rufer, A. C., Stihle, M., Umanã, P., & Benz, J. (2011). Unique carbohydrate-carbohydrate interactions are required for high affinity binding between FcγRIII and antibodies lacking core fucose. Proceedings of the National Academy of Sciences of the United States of America, 108(31), 12669–12674. https://doi.org/10.1073/PNAS.1108455108/SUPPL_FILE/PNAS.1108455108_SI.PDF

Finigan, J. H. (2009). The coagulation system and pulmonary endothelial function in acute lung injury. Microvascular Research, 77(1), 35–38. https://doi.org/10.1016/J.MVR.2008.09.002

Fischetti, F., & Tedesco, F. (2009). Cross-talk between the complement system and endothelial cells in physiologic conditions and in vascular diseases.

Https://Doi.Org/10.1080/08916930600739712, 39(5), 417–428. https://doi.org/10.1080/08916930600739712

Golden, G. J., Toledo, A. G., Marki, A., Sorrentino, J. T., Morris, C., Riley, R. J., Spliid, C., Chen, Q., Cornax, I., Lewis, N. E., Varki, N., Le, D., Malmström, J., Karlsson, C., Ley, K., Nizet, V., & Esko, J. D. (2021). Endothelial Heparan Sulfate Mediates Hepatic Neutrophil Trafficking and Injury during Staphylococcus aureus Sepsis. MBio, 12(5). https://doi.org/10.1128/MBIO.01181-21

Han, S., Lee, S. J., Kim, K. E., Lee, H. S., Oh, N., Park, I., Ko, E., Oh, S. J., Lee, Y. S., Kim, D., Lee, S., Lee, D. H., Lee, K. H., Chae, S. Y., Lee, J. H., Kim, S. J., Kim, H. C., Kim, S., Kim, S. H., … Koh, G. Y. (2016). Amelioration of sepsis by TIE2 activation-induced vascular protection. Science Translational Medicine, 8(335). https://doi.org/10.1126/SCITRANSLMED.AAD9260

Happonen, L., Hauri, S., Svensson Birkedal, G., Karlsson, C., de Neergaard, T., Khakzad, H., Nordenfelt, P., Wikström, M., Wisniewska, M., Björck, L., Malmström, L., & Malmström, J. (2019). A quantitative Streptococcus pyogenes–human protein–protein interaction map reveals localization of opsonizing antibodies. Nature Communications, 10(1). https://doi.org/10.1038/s41467-019-10583-5

Henningham, A., Barnett, T. C., Maamary, P. G., & Walker, M. J. (2012). Pathogenesis of Group A Streptococcal Infections. Discovery Medicine, 13(72), 329–342.

Henry, C. B. S., & Duling, B. R. (1999). Permeation of the luminal capillary glycocalyx is determined by hyaluronan. American Journal of Physiology - Heart and Circulatory Physiology, 277(2 46-2). https://doi.org/10.1152/AJPHEART.1999.277.2.H508/ASSET/IMAGES/LARGE/AHEA40822006Y.JPEG

Hoffmann, J. H. O., & Enk, A. H. (2019). High-dose intravenous immunoglobulin in skin autoimmune disease. Frontiers in Immunology, 10(JUN), 1090. https://doi.org/10.3389/FIMMU.2019.01090/BIBTEX

Hotchkiss, R. S., Moldawer, L. L., Opal, S. M., Reinhart, K., Turnbull, I. R., & Vincent, J. L. (2016). Sepsis and septic shock. Nat. Rev. Dis. Primers., 2, 16045. https://doi.org/10.1038/nrdp.2016.45

Hu, H., Khatri, K., Klein, J., Leymarie, N., & Zaia, J. (2016). A review of methods for interpretation of glycopeptide tandem mass spectral data. Glycoconjugate Journal, 33(3), 285–296. https://doi.org/10.1007/S10719-015-9633-3/FIGURES/2

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols, 4(1), 44–57. https://doi.org/10.1038/NPROT.2008.211

Iba, T., & Levy, J. H. (2019). Derangement of the endothelial glycocalyx in sepsis. Journal of Thrombosis and Haemostasis, 17(2), 283–294. https://doi.org/10.1111/JTH.14371

Ince, C., Mayeux, P. R., Nguyen, T., Gomez, H., Kellum, J. A., Ospina-Tascón, G. A., Hernandez, G., Murray, P., & De Backer, D. (2016). The endothelium in sepsis. Shock, 45(3), 259–270. https://doi.org/10.1097/SHK.0000000000000473

Junhee Seok, H. Shaw Warren, Alex, G. C., Michael, N. M., Henry, V. B., Xu, W., Richards, D. R., McDonald-Smith, G. P., Gao, H., Hennessy, L., Finnerty, C. C., López, C. M., Honari, S., Moore, E. E., Minei, J. P., Cuschieri, J., Bankey, P. E., Johnson, J. L., Sperry, J., … Tompkins, R. G. (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. Proceedings of the National Academy of Sciences of the United States of America, 110(9), 3507–3512. https://doi.org/10.1073/pnas.1222878110

Kao, D., Lux, A., Schaffert, A., Lang, R., Altmann, F., & Nimmerjahn, F. (2017). IgG subclass and vaccination stimulus determine changes in antigen specific antibody glycosylation in mice. European Journal of Immunology, 47(12), 2070–2079. https://doi.org/10.1002/EJI.201747208

Karlsson, C. A. Q., Järnum, S., Winstedt, L., Kjellman, C., Björck, L., Linder, A., & Malmström, J. A. (2018). Streptococcus pyogenes Infection and the Human Proteome with a Special Focus on the Immunoglobulin G-cleaving Enzyme IdeS. Molecular & Cellular Proteomics, 17(6), 1097–1111. https://doi.org/10.1074/MCP.RA117.000525

Kaukonen, K. M., Bailey, M., Suzuki, S., Pilcher, D., & Bellomo, R. (2014). Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand, 2000-2012. JAMA, 311(13), 1308–1316. https://doi.org/10.1001/JAMA.2014.2637

Kaul, R., McGeer, A., Norrby-Teglund, A., Kotb, M., Schwartz, B., O'Rourke, K., Talbot, J., Low, D. E., Allen, U. D., Archibald, S. D., Brunham, R. R., Cameron, D. W., Desruisseaux, D. B., Embil, J., Fuller, J. G., Jutras, P., Lam, A. S. C., Lamothe, F., Loo, V. G., … Wobeser, W. (1999). Intravenous immunoglobulin therapy for streptococcal toxic shock syndrome--a comparative observational study. The Canadian Streptococcal Study Group. Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America, 28(4), 800–807. https://doi.org/10.1086/515199

Khakzad, H., Happonen, L., Karami, Y., Chowdhury, S., Bergdahl, G. E., Nilges, M., Van Nhieu, G. T., Malmstrom, J., & Malmstrom, L. (2021). Structural determination of Streptococcus pyogenes M1 protein interactions with human immunoglobulin G using integrative structural biology. PLOS Computational Biology, 17(1), e1008169. https://doi.org/10.1371/JOURNAL.PCBI.1008169

Kim, H. K., Missiakas, D., & Schneewind, O. (2014). Mouse models for infectious diseases caused by Staphylococcus aureus. Journal of Immunological Methods, 410, 88–99. https://doi.org/10.1016/J.JIM.2014.04.007

Kolaczkowska, E., Jenne, C. N., Surewaard, B. G. J., Thanabalasuriar, A., Lee, W. Y., Sanz, M. J., Mowen, K., Opdenakker, G., & Kubes, P. (2015). Molecular mechanisms of NET formation and degradation revealed by intravital imaging in the liver vasculature. Nature Communications 2015 6:1, 6(1), 1–13. https://doi.org/10.1038/ncomms7673

Kong, S., Gong, P., Zeng, W. F., Jiang, B., Hou, X., Zhang, Y., Zhao, H., Liu, M., Yan, G., Zhou, X., Qiao, X., Wu, M., Yang, P., Liu, C., & Cao, W. (2022). pGlycoQuant with a

deep residual network for quantitative glycoproteomics at intact glycopeptide level. Nature Communications 2022 13:1, 13(1), 1–17. https://doi.org/10.1038/s41467-022-35172-x

Kotidis, P., Donini, R., Arnsdorf, J., Hansen, A. H., Voldborg, B. G. R., Chiang, A. W. T., Haslam, S., Betenbaugh, M., Jimenez del Val, I., Lewis, N. E., Krambeck, F., & Kontoravdi, C. (2023). CHOGlycoNET: Comprehensive glycosylation reaction network for CHO cells. Metabolic Engineering. https://doi.org/10.1016/J.YMBEN.2022.12.009

Lapek, J. D., Mills, R. H., Wozniak, J. M., Campeau, A., Fang, R. H., Wei, X., van de Groep, K., Perez-Lopez, A., van Sorge, N. M., Raffatellu, M., Knight, R., Zhang, L., & Gonzalez, D. J. (2018). Defining Host Responses during Systemic Bacterial Infection through Construction of a Murine Organ Proteome Atlas. Cell Systems, 6(5), 579-592.e4. https://doi.org/10.1016/J.CELS.2018.04.010

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. Nature, 401(6755), 788–791. https://doi.org/10.1038/44565

Lee, J., Boutz, D. R., Chromikova, V., Joyce, M. G., Vollmers, C., Leung, K., Horton, A. P., DeKosky, B. J., Lee, C. H., Lavinder, J. J., Murrin, E. M., Chrysostomou, C., Hoi, K. H., Tsybovsky, Y., Thomas, P. V., Druz, A., Zhang, B., Zhang, Y., Wang, L., … Georgiou, G. (2016). Molecular-level analysis of the serum antibody repertoire in young adults before and after seasonal influenza vaccination. Nature Medicine 2016 22:12, 22(12), 1456–1464. https://doi.org/10.1038/nm.4224

Lee, L. Y., Moh, E. S. X., Parker, B. L., Bern, M., Packer, N. H., & Thaysen-Andersen, M. (2016). Toward Automated N-Glycopeptide Identification in Glycoproteomics. Journal of Proteome Research, 15(10), 3904–3915. https://doi.org/10.1021/ACS.JPROTEOME.6B00438

Leligdowicz, A., Richard-Greenblatt, M., Wright, J., Crowley, V. M., & Kain, K. C. (2018). Endothelial Activation: The Ang/Tie Axis in Sepsis. Frontiers in Immunology, 9(APR). https://doi.org/10.3389/FIMMU.2018.00838

Lelubre, C., & Vincent, J. L. (2018). Mechanisms and treatment of organ failure in sepsis. Nature Reviews Nephrology 2018 14:7, 14(7), 417–427. https://doi.org/10.1038/s41581-018-0005-7

Levi, M., & van der Poll, T. (2017). Coagulation and sepsis. Thromb. Res., 149, 38–44. https://doi.org/10.1016/j.thromres.2016.11.007

Lewis, A. J., Seymour, C. W., & Rosengart, M. R. (2016). Current Murine Models of Sepsis. Surgical Infections, 17(4), 385. https://doi.org/10.1089/SUR.2016.021

Liu, V., Escobar, G. J., Greene, J. D., Soule, J., Whippy, A., Angus, D. C., & Iwashyna, T. J. (2014). Hospital deaths in patients with sepsis from 2 independent cohorts. JAMA, 312(1), 90–92. https://doi.org/10.1001/JAMA.2014.5804

Mahan, A. E., Jennewein, M. F., Suscovich, T., Dionne, K., Tedesco, J., Chung, A. W., Streeck, H., Pau, M., Schuitemaker, H., Francis, D., Fast, P., Laufer, D., Walker, B. D., Baden, L., Barouch, D. H., & Alter, G. (2016). Antigen-Specific Antibody Glycosylation Is

Regulated via Vaccination. PLOS Pathogens, 12(3), e1005456. https://doi.org/10.1371/JOURNAL.PPAT.1005456

MAJER, R. V., & HYDE, R. D. (1988). High-dose intravenous immunoglobulin in the treatment of autoimmune haemolytic anaemia. Clinical and Laboratory Haematology, 10(4), 391–395. https://doi.org/10.1111/J.1365-2257.1988.TB01186.X

Malmström, E., Kilsgård, O., Hauri, S., Smeds, E., Herwald, H., Malmström, L., & Malmström, J. (2016). Large-scale inference of protein tissue origin in gram-positive sepsis plasma using quantitative targeted proteomics. Nature Communications, 7. https://doi.org/10.1038/NCOMMS10261

Mann, M., Hendrickson, R. C., & Pandey, A. (2003). Analysis of Proteins and Proteomes by Mass Spectrometry. Https://Doi.Org/10.1146/Annurev.Biochem.70.1.437, 70, 437–473. https://doi.org/10.1146/ANNUREV.BIOCHEM.70.1.437

Marki, A., Esko, J. D., Pries, A. R., & Ley, K. (2015). Role of the endothelial surface layer in neutrophil recruitment. Journal of Leukocyte Biology, 98(4), 503–515. https://doi.org/10.1189/JLB.3MR0115-011R

Marshall, J. C. (2008). Sepsis: rethinking the approach to clinical research. Journal of Leukocyte Biology, 83(3), 471–482. https://doi.org/10.1189/JLB.0607380

Marshall, J. C. (2014). Why have clinical trials in sepsis failed? Trends in Molecular Medicine, 20(4), 195–203. https://doi.org/10.1016/J.MOLMED.2014.01.007

McDonald, B., McAvoy, E. F., Lam, F., Gill, V., De La Motte, C., Savani, R. C., & Kubes, P. (2008). Interaction of CD44 and hyaluronan is the dominant mechanism for neutrophil sequestration in inflamed liver sinusoids. Journal of Experimental Medicine, 205(4), 915–927. https://doi.org/10.1084/JEM.20071765/VIDEO-1

Michael E. Watson, J., Neely, M. N., & Caparon, M. G. (2016). Animal Models of Streptococcus pyogenes Infection. Streptococcus Pyogenes: Basic Biology to Clinical Manifestations. https://www.ncbi.nlm.nih.gov/books/NBK333421/

Michalik, S., Sundaramoorthy, N., Murr, A., Depke, M., Völker, U., Bröker, B. M., Aamot, H. V., & Schmidt, F. (2020). Early-Stage Staphylococcus aureus Bloodstream Infection Causes Changes in the Concentrations of Lipoproteins and Acute-Phase Proteins and Is Associated with Low Antibody Titers against Bacterial Virulence Factors. MSystems, 5(1). https://doi.org/10.1128/MSYSTEMS.00632-19

Mitchell, T. J. (2003). The pathogenesis of streptococcal infections: From Tooth decay to meningitis. Nature Reviews Microbiology, 1(3), 219–230. https://doi.org/10.1038/nrmicro771

Naegeli, A., Bratanis, E., Karlsson, C., Shannon, O., Kalluru, R., Linder, A., Malmström, J., & Collin, M. (2019). Streptococcus pyogenes evades adaptive immunity through specific IgG glycan hydrolysis. The Journal of Experimental Medicine, 216(7), 1615. https://doi.org/10.1084/JEM.20190293

Nasir, W., Toledo, A. G., Noborn, F., Nilsson, J., Wang, M., Bandeira, N., & Larson, G. (2016). SweetNET: A bioinformatics workflow for glycopeptide MS/MS spectral analysis.

Journal of Proteome Research, 15(8), 2826–2840. https://doi.org/10.1021/ACS.JPROTEOME.6B00417/ASSET/IMAGES/LARGE/PR-2016-00417Q_0009.JPEG

Nordenfelt, P., Waldemarson, S., Linder, A., Mörgelin, M., Karlsson, C., Malmström, J., & Björck, L. (2012). Antibody orientation at bacterial surfaces is related to invasive infection. The Journal of Experimental Medicine, 209(13), 2367–2381. https://doi.org/10.1084/JEM.20120325

Opal, S. M. (2010). A Brief History of Microbiology and Immunology. Vaccines: A Biography, 31. https://doi.org/10.1007/978-1-4419-1108-7_3

Pagan, J. D., Kitaoka, M., & Anthony, R. M. (2018). Engineered Sialylation of Pathogenic Antibodies In Vivo Attenuates Autoimmune Disease. Cell, 172(3), 564-577.e13. https://doi.org/10.1016/J.CELL.2017.11.041

Pimienta, G., Heithoff, D. M., Rosa-Campos, A., Tran, M., Esko, J. D., Mahan, M. J., Marth, J. D., & Smith, J. W. (2019). Plasma Proteome Signature of Sepsis: a Functionally Connected Protein Network. PROTEOMICS, 19(5), 1800389. https://doi.org/10.1002/PMIC.201800389

Poli-de-Figueiredo, L. F., Garrido, A. G., Nakagawa, N., & Sannomiya, P. (2008). Experimental models of sepsis and their clinical relevance. Shock, 30(SUPPL. 1), 53–59. https://doi.org/10.1097/SHK.0B013E318181A343

Reitsma, S., Slaaf, D. W., Vink, H., Van Zandvoort, M. A. M. J., & Oude Egbrink, M. G. A. (2007). The endothelial glycocalyx: Composition, functions, and visualization. Pflugers Archiv European Journal of Physiology, 454(3), 345–359. https://doi.org/10.1007/S00424-007-0212-8/FIGURES/2

Rhee, C., Dantes, R., Epstein, L., Murphy, D. J., Seymour, C. W., Iwashyna, T. J., Kadri, S. S., Angus, D. C., Danner, R. L., Fiore, A. E., Jernigan, J. A., Martin, G. S., Septimus, E., Warren, D. K., Karcz, A., Chan, C., Menchaca, J. T., Wang, R., Gruber, S., & Klompas, M. (2017). Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009-2014. JAMA - Journal of the American Medical Association, 318(13), 1241–1249. https://doi.org/10.1001/JAMA.2017.13836

Rohde, M., & Cleary, P. P. (2016). Adhesion and invasion of Streptococcus pyogenes into host cells and clinical relevance of intracellular streptococci. Streptococcus Pyogenes: Basic Biology to Clinical Manifestations. https://www.ncbi.nlm.nih.gov/books/NBK333420/

Rossaint, J., & Zarbock, A. (2015). Pathogenesis of Multiple Organ Failure in Sepsis. Critical Reviews™ in Immunology, 35(4), 277–291. https://doi.org/10.1615/CRITREVIMMUNOL.2015015461

Rudd, K. E., Johnson, S. C., Agesa, K. M., Shackelford, K. A., Tsoi, D., Kievlan, D. R., Colombara, D. V., Ikuta, K. S., Kissoon, N., Finfer, S., Fleischmann-Struzek, C., Machado, F. R., Reinhart, K. K., Rowan, K., Seymour, C. W., Watson, R. S., West, T. E., Marinho, F., Hay, S. I., … Naghavi, M. (2020). Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. The Lancet, 395(10219), 200–211. https://doi.org/10.1016/S0140-6736(19)32989-7

Ryback, J. N., Ettorre, A., Kaissling, B., Giavazzi, R., Neri, D., & Elia, G. (2005). In vivo protein biotinylation for identification of organ-specific antigens accessible from the vasculature. Nature Methods 2005 2:4, 2(4), 291–298. https://doi.org/10.1038/nmeth745

Schmugge, M., Rand, M. L., & Freedman, J. (2003). Platelets and von Willebrand factor. Transfusion and Apheresis Science, 28(3), 269–277. https://doi.org/10.1016/S1473-0502(03)00046-6

Shannon, O. (2021). The role of platelets in sepsis. Research and Practice in Thrombosis and Haemostasis, 5(1), 27–37. https://doi.org/10.1002/RTH2.12465

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research, 13(11), 2498–2504. https://doi.org/10.1101/GR.1239303

Silverman, J. M., & Imperiali, B. (2016). Bacterial N-Glycosylation Efficiency Is Dependent on the Structural Context of Target Sequons. The Journal of Biological Chemistry, 291(42), 22001. https://doi.org/10.1074/JBC.M116.747121

Singer, M., Deutschman, C. S., Seymour, C., Shankar-Hari, M., Annane, D., Bauer, M., Bellomo, R., Bernard, G. R., Chiche, J. D., Coopersmith, C. M., Hotchkiss, R. S., Levy, M. M., Marshall, J. C., Martin, G. S., Opal, S. M., Rubenfeld, G. D., Poll, T. Der, Vincent, J. L., & Angus, D. C. (2016). The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA, 315(8), 801–810. https://doi.org/10.1001/jama.2016.0287

Slany, A., Bileck, A., Kreutz, D., Mayer, R. L., Muqaku, B., & Gerner, C. (2016). Contribution of human fibroblasts and endothelial cells to the hallmarks of inflammation as determined by proteome profiling. Molecular and Cellular Proteomics, 15(6), 1982–1997. https://doi.org/10.1074/mcp.M116.058099

Sorrentino, J. T., Golden, G. J., Morris, C., Painter, C. D., Nizet, V., Campos, A. R., Smith, J. W., Karlsson, C., Malmström, J., Lewis, N. E., Esko, J. D., & Gómez Toledo, A. (2022). Vascular Proteome Responses Precede Organ Dysfunction in a Murine Model of Staphylococcus aureus Bacteremia. MSystems, 7(4). https://doi.org/10.1128/MSYSTEMS.00395-22/ASSET/CA05F0A2-F9D3-4D50-A4C5-AA591BC4F75D/ASSETS/IMAGES/MEDIUM/MSYSTEMS.00395-22-F006.GIF

Sriskandan, S., Ferguson, M., Elliot, V., Faulkner, L., & Cohen, J. (2006). Human intravenous immunoglobulin for experimental streptococcal toxic shock: bacterial clearance and modulation of inflammation. Journal of Antimicrobial Chemotherapy, 58(1), 117–124. https://doi.org/10.1093/JAC/DKL173

Streb, M., Klett, M., Lindbergh, G., Banzatti, A., Pontoppidan, K. M., Salyk, C., Iop, A., Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Stryjewski, M. E., & Corey, G. R. (2014). Methicillin-resistant staphylococcus aureus: An evolving pathogen. Clinical Infectious Diseases, 58(SUPPL. 1). https://doi.org/10.1093/CID/CIT613

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLOS ONE, 6(7), e21800. https://doi.org/10.1371/JOURNAL.PONE.0021800

Tabang, D. N., Ford, M., & Li, L. (2021). Recent Advances in Mass Spectrometry-Based Glycomic and Glycoproteomic Studies of Pancreatic Diseases. Frontiers in Chemistry, 9, 579. https://doi.org/10.3389/FCHEM.2021.707387/BIBTEX

Tang, H., Ivanciu, L., Popescu, N., Peer, G., Hack, E., Lupu, C., Taylor, F. B., & Lupu, F. (2007). Sepsis-Induced Coagulation in the Baboon Lung Is Associated with Decreased Tissue Factor Pathway Inhibitor. The American Journal of Pathology, 171(3), 1066–1077. https://doi.org/10.2353/AJPATH.2007.070104

Tarca, A. L., Carey, V. J., Chen, X. wen, Romero, R., & Drăghici, S. (2007). Machine learning and its applications to biology. PLoS Computational Biology, 3(6). https://doi.org/10.1371/JOURNAL.PCBI.0030116

Taylor, J., Reinhart, K., & Vallet, B. (2001). Staging of the pathophysiologic responses of the primate microvasculature to Escherichia coli and endotoxin: Examination of the elements of the compensated response and their links to the corresponding uncompensated lethal variants. Critical Care Medicine, 29(7 SUPPL.). https://doi.org/10.1097/00003246-200107001-00026

Toledo, A. G., Golden, G., Campos, A. R., Cuello, H., Sorrentino, J., Lewis, N., Varki, N., Nizet, V., Smith, J. W., & Esko, J. D. (2019). Proteomic atlas of organ vasculopathies triggered by Staphylococcus aureus sepsis. Nature Communications 2019 10:1, 10(1), 1–13. https://doi.org/10.1038/s41467-019-12672-x

Tong, S. Y. C., Davis, J. S., Eichenberger, E., Holland, T. L., & Fowler, V. G. (2015). Staphylococcus aureus Infections: Epidemiology, Pathophysiology, Clinical Manifestations, and Management. Clinical Microbiology Reviews, 28(3), 603. https://doi.org/10.1128/CMR.00134-14

Toustou, C., Walet-Balieu, M. L., Kiefer-Meyer, M. C., Houdou, M., Lerouge, P., Foulquier, F., & Bardor, M. (2022). Towards understanding the extensive diversity of protein N-glycan structures in eukaryotes. Biological Reviews, 97(2), 732–748. https://doi.org/10.1111/BRV.12820

Tyanova, S., Temu, T., & Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nature Protocols 2016 11:12, 11(12), 2301–2319. https://doi.org/10.1038/nprot.2016.136

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., & Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature Methods 2016 13:9, 13(9), 731–740. https://doi.org/10.1038/nmeth.3901

Uchimido, R., Schmidt, E. P., & Shapiro, N. I. (2019). The glycocalyx: A novel diagnostic and therapeutic target in sepsis. Critical Care, 23(1), 1–12. https://doi.org/10.1186/S13054-018-2292-6/TABLES/1

Urban-Chmiel, R., Marek, A., Stępień-Pyśniak, D., Wieczorek, K., Dec, M., Nowaczek, A., & Osek, J. (2022). Antibiotic Resistance in Bacteria—A Review. Antibiotics, 11(8). https://doi.org/10.3390/ANTIBIOTICS11081079/S1

Van Der Poll, T., Van De Veerdonk, F. L., Scicluna, B. P., & Netea, M. G. (2017). The immunopathology of sepsis and potential therapeutic targets. Nat. Rev. Immunol., 17(7), 407–420. https://doi.org/10.1038/nri.2017.36

Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., & Seeberger, P. H. (2022). Essentials of Glycobiology. Essentials of Glycobiology. https://doi.org/10.1101/9781621824213

Vincent, J. L., Moreno, R., Takala, J., Willatts, S., De Mendonça, A., Bruining, H., Reinhart, C. K., Suter, P. M., & Thijs, L. G. (1996). The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive Care Med., 22(7), 707–710. https://doi.org/10.1007/bf01709751

von Mering, C., Huynen, M., Jaeggi, D., Schmidt, S., Bork, P., & Snel, B. (2003). STRING: a database of predicted functional associations between proteins. Nucleic Acids Research, 31(1), 258–261. https://doi.org/10.1093/NAR/GKG034

Von Pawel-Rammingen, U., Johansson, B. P., & Björck, L. (2002). IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. The EMBO Journal, 21(7), 1607. https://doi.org/10.1093/EMBOJ/21.7.1607

Walker, M. J., Barnett, T. C., McArthur, J. D., Cole, J. N., Gillen, C. M., Henningham, A., Sriprakash, K. S., Sanderson-Smith, M. L., & Nizet, V. (2014). Disease manifestations and pathogenic mechanisms of Group A Streptococcus. Clinical Microbiology Reviews, 27(2), 264–301. https://doi.org/10.1128/CMR.00101-13

Walsh, G. (2018). Biopharmaceutical benchmarks 2018. Nature Biotechnology 2018 36:12, 36(12), 1136–1145. https://doi.org/10.1038/nbt.4305

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen, D. D., Watrous, J., Kapono, C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W. T., Crüsemann, M., Boudreau, P. D., Esquenazi, E., Sandoval-Calderón, M., … Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nature Biotechnology 2016 34:8, 34(8), 828–837. https://doi.org/10.1038/nbt.3597

Wang, Y., Tan, J., Sutton-Smith, M., Ditto, D., Panico, M., Campbell, R. M., Varki, N. M., Long, J. M., Jaeken, J., Levinson, S. R., Wynshaw-Boris, A., Morris, H. R., Le, D., Dell, A., Schachter, H., & Marth, J. D. (2001). Modeling human congenital disorder of glycosylation type IIa in the mouse: conservation of asparagine-linked glycan-dependent

functions in mammalian physiology and insights into disease pathogenesis. Glycobiology, 11(12), 1051–1070. https://doi.org/10.1093/GLYCOB/11.12.1051

Washburn, N., Schwabb, I., Ortiz, D., Bhatnagar, N., Lansing, J. C., Medeiros, A., Tyler, S., Mekala, D., Cochran, E., Sarvaiya, H., Garofalo, K., Meccariello, R., Meador, J. W., Rutitzky, L., Schultes, B. C., Ling, L., Avery, W., Nimmerjahn, F., Manning, A. M., … Bosques, C. J. (2015). Controlled tetra-Fc sialylation of IVIg results in a drug candidate with consistent enhanced anti-inflammatory activity. Proceedings of the National Academy of Sciences of the United States of America, 112(11), E1297–E1306. https://doi.org/10.1073/PNAS.1422481112/SUPPL_FILE/PNAS.201422481SI.PDF

Weidner, C., Steinfath, M., Opitz, E., Oelgeschläger, M., & Schönfelder, G. (2016). Defining the optimal animal model for translational research using gene set enrichment analysis. EMBO Molecular Medicine, 8(8), 831–838. https://doi.org/10.15252/EMMM.201506025

Weng, L., Zeng, X. ying, Yin, P., Wang, L. jun, Wang, C. yao, Jiang, W., Zhou, M. geng, Du, B., Zhao, M. yan, Yang, S. lin, Ma, X. chun, Sun, Y. ni, Yin, Y. jie, Liu, D. xin, Hu, Z. jie, Huo, Y., Jiang, L., Zhang, Q., An, Y. zhong, … Hu, X. yun. (2018). Sepsis-related mortality in China: a descriptive analysis. Intensive Care Medicine, 44(7), 1071–1080. https://doi.org/10.1007/S00134-018-5203-Z

Wenig, K., Chatwell, L., Von Pawel-Rammingen, U., Björck, L., Huber, R., & Sondermann, P. (2004). Structure of the streptococcal endopeptidase IdeS, a cysteine proteinase with strict specificity for IgG. Proceedings of the National Academy of Sciences of the United States of America, 101(50), 17371–17376. https://doi.org/10.1073/PNAS.0407965101

Wiesinger, A., Peters, W., Chappell, D., Kentrup, D., Reuter, S., Pavenstädt, H., Oberleithner, H., & Kümpers, P. (2013). Nanomechanics of the Endothelial Glycocalyx in Experimental Sepsis. PLOS ONE, 8(11), e80905. https://doi.org/10.1371/JOURNAL.PONE.0080905

Willforss, J., Chawade, A., & Levander, F. (2019). NormalyzerDE: Online Tool for Improved Normalization of Omics Expression Data and High-Sensitivity Differential Expression Analysis. Journal of Proteome Research, 18(2), 732–740. https://doi.org/10.1021/ACS.JPROTEOME.8B00523/SUPPL_FILE/PR8B00523_SI_010.ZIP

Wolf-Yadlin, A., Hu, A., & Noble, W. S. (2016). Technical advances in proteomics: new developments in data-independent acquisition. F1000Research 2016 5:419, 5, 419. https://doi.org/10.12688/f1000research.7042.1

Wollein Waldetoft, K., & Råberg, L. (2014). To harm or not to harm? On the evolution and expression of virulence in group A streptococci. Trends in Microbiology, 22(1), 7–13. https://doi.org/10.1016/j.tim.2013.10.006

Wozniak, J. M., Mills, R. H., Olson, J., Caldera, J. R., Sepich-Poore, G. D., Carrillo-Terrazas, M., Tsai, C. M., Vargas, F., Knight, R., Dorrestein, P. C., Liu, G. Y., Nizet, V., Sakoulas, G., Rose, W., & Gonzalez, D. J. (2020). Mortality Risk Profiling of Staphylococcus aureus Bacteremia by Multi-omic Serum Analysis Reveals Early Predictive and

Pathogenic Signatures. Cell, 182(5), 1311-1327.e14. https://doi.org/10.1016/J.CELL.2020.07.040

York, W. S., Mazumder, R., Ranzinger, R., Edwards, N., Kahsay, R., Aoki-Kinoshita, K. F., Campbell, M. P., Cummings, R. D., Feizi, T., Martin, M., Natale, D. A., Packer, N. H., Woods, R. J., Agarwal, G., Arpinar, S., Bhat, S., Blake, J., Castro, L. J. G., Fochtman, B., … Zhang, W. (2020). GlyGen: Computational and Informatics Resources for Glycoscience. Glycobiology, 30(2), 72–73. https://doi.org/10.1093/GLYCOB/CWZ080

Yu, C. Y., Mayampurath, A., Zhu, R., Zacharias, L., Song, E., Wang, L., Mechref, Y., & Tang, H. (2016). Automated glycan sequencing from tandem mass spectra of N-linked glycopeptides. Analytical Chemistry, 88(11), 5725. https://doi.org/10.1021/ACS.ANALCHEM.5B04858

Zandman-Goddard, G., Levy, Y., & Shoenfeld, Y. (2005). Intravenous immunoglobulin therapy and systemic lupus erythematosus. Clinical Reviews in Allergy & Immunology, 29(3), 219–228. https://doi.org/10.1385/CRIAI:29:3:219

Zhang, X., Sun, D., Song, J. W., Zullo, J., Lipphardt, M., Coneh-Gould, L., & Goligorsky, M. S. (2018). Endothelial cell dysfunction and glycocalyx – A vicious circle. Matrix Biology, 71–72, 421–431. https://doi.org/10.1016/J.MATBIO.2018.01.026

Zhong, C. Q., Wu, J., Qiu, X., Chen, X., Xie, C., & Han, J. (2020). Generation of a murine SWATH-MS spectral library to quantify more than 11,000 proteins. Scientific Data, 7(1). https://doi.org/10.1038/S41597-020-0449-Z

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., Benner, C., & Chanda, S. K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature Communications, 10(1). https://doi.org/10.1038/S41467-019-09234-6