

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Protein Engineering Approaches to Biomarker Identification and Biosensor Development

Permalink

<https://escholarship.org/uc/item/9d59d9xg>

Author

Sen, Sanjana

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Protein Engineering Approaches to Biomarker Identification
and Biosensor Development

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Molecular Biology and Biochemistry

by

Sanjana Sen

Dissertation Committee:
Professor Gregory A. Weiss, Chair
Professor Jennifer A. Prescher
Professor Thomas L. Poulos

2021

Chapter 2 © American Society for Microbiology
Chapter 3 © Sanjana Sen, Emily C. Sanders, Alicia M. Santos, Keertna Bhuvan, Derek Y.
Tang, Aidan A. Gelston, Joni L. Ricks-Oddie and Gregory A. Weiss
Chapter 5 © American Chemical Society
Chapter 6 © Emily C. Sanders, Sanjana Sen, Aidan A. Gelston, Alicia M. Santos, Xuan Luo,
Colin L. Raston and Gregory Weiss
All other materials © Sanjana Sen

DEDICATION

To

my parents,
Pratik, and Billy

*“The truth is we don’t know what we don’t know.
We don’t even know the questions we need to ask in order to find out, but when we learn one tiny little
thing, a dim light comes on in a dark hallway, and suddenly a new question appears.”*

– Yaa Gyasi, **Transcendent Kingdom**

TABLE OF CONTENTS

| | |
|---|-------------|
| LIST OF FIGURES | v |
| LIST OF TABLES | vii |
| ACKNOWLEDGEMENTS..... | viii |
| CURRICULUM VITAE..... | xi |
| ABSTRACT OF THE DISSERTATION..... | xiv |
| Antibody-dependent regulation of COVID-19 disease outcomes | 1 |
| 1.1 Abstract..... | 1 |
| 1.2 Introduction..... | 1 |
| 1.3 The Role of Neutralization in Disease Outcomes | 4 |
| 1.4 The Role of Immunological Memory in COVID-19 Progression | 8 |
| 1.5 Deregulated IgG Responses Lead to COVID-19 Severity..... | 12 |
| 1.6 Deregulated Ab Timing and the Pro-Inflammatory Response..... | 13 |
| 1.7 Altered Post-Translational Modifications Can Drive Disease Outcomes | 17 |
| 1.8 Conclusions and Discussion..... | 18 |
| 1.8 References | 22 |
| Predicting COVID-19 Severity with a Specific Nucleocapsid Antibody plus Disease Risk Factor Score 29 | |
| 2.1 Abstract..... | 29 |
| 2.2 Introduction..... | 30 |
| 2.3 Results..... | 31 |
| 2.4 Discussion..... | 42 |
| 2.5 Materials and Methods:..... | 46 |
| 2.6 Supporting Figures and Tables..... | 53 |
| 2.7 References | 62 |
| Evidence for Deleterious Antigenic Imprinting in SARS-CoV-2 Immune Response | 66 |
| 3.1 Abstract..... | 66 |
| 3.2 Importance..... | 66 |
| 3.3 Introduction..... | 67 |
| 3.4 Results and Discussion..... | 68 |
| 3.5 Materials and Methods..... | 76 |

| | |
|---|------------|
| 3.6 References | 96 |
| Development towards a real-time insulin biosensor for Type-1-Diabetes disease management..... | 98 |
| 4.1 Abstract..... | 98 |
| 4.2 Introduction..... | 99 |
| 4.3 Results and Discussion..... | 103 |
| 4.4 Conclusions and Future Directions..... | 121 |
| 4.5 Materials and methods..... | 123 |
| 4.6 References | 137 |
| Photostable and Proteolysis-Resistant Förster Resonance Energy Transfer-Based Calcium Biosensor | 141 |
| 5.1 Abstract..... | 141 |
| 5.2 Introduction..... | 142 |
| 5.3 Experimental Section..... | 145 |
| 5.4 Fabrication of the Membrane-Encapsulated Optical Fiber (MEOF) | 147 |
| 5.5 Characterization of the Filter-Membrane | 149 |
| 5.6 Twitch-2B Ca ²⁺ Probe | 151 |
| 5.7 The Development of a Photostable Twitch-2B FRET Construct | 154 |
| 5.8 Characterization of MEOF Probes with the STC Construct..... | 155 |
| 5.9 Protection from Proteolytic Enzymes | 157 |
| 5.10 Conclusion..... | 158 |
| 5.11 Materials and Methods | 159 |
| 5.12 References..... | 169 |
| Under 5 Minute Immunoblot Assays by Vortex Fluidic Device Acceleration..... | 172 |
| 6.1 Abstract..... | 172 |
| 6.2 Introduction..... | 172 |
| 6.3 Results and Discussion..... | 174 |
| 6.4 Materials and Methods..... | 180 |
| 6.5 References | 190 |
| Summary and Conclusions | 192 |
| 7.1 References | 197 |

LIST OF FIGURES

| | |
|---|-----|
| Figure 1-1. Factors impacting Ab-dependent immune regulation of COVID-19..... | 3 |
| Figure 1-2. Neutralization can protect against SARS-CoV-2 infection | 6 |
| Figure 1-3. Antigenic Imprinting (AIM) can result in protective and detrimental effects. | 9 |
| Figure 2-1. Predicted SARS-CoV-2 epitopes examined by phage ELISA..... | 33 |
| Figure 2-2 Mapping COVID-19 patient antibody responses with phage-displayed SARS-CoV-2 epitopes | 35 |
| Figure 2-3. Patients with α Ep9 Abs have more severe disease..... | 37 |
| Figure 2-4. Correlation between disease severity and risk factors in patients with α Ep9 Abs..... | 39 |
| Figure 2-5. Association of inflammatory cytokine and tissue damage markers in patients with α Ep9 Abs.. | 41 |
| Figure 2-6. Quality control ELISA (QC ELISA) for phage-displayed, epitope candidates..... | 53 |
| Figure 2-7. Epitope homology of SARS-CoV-2 with four phylogenetically related coronaviruses known to infect humans | 54 |
| Figure 2-8. COVAM data showing the variation in IgG seroreactivity of patient plasma..... | 55 |
| Figure 2-9. Variation in IgM seroreactivity of patient plasma | 56 |
| Figure 2-10. Comparison of disease severity and clinical parameters of patients with α Ep9 Abs..... | 57 |
| Figure 2-11. Schematic of plasmid map for phage-displayed or recombinant expression epitopes..... | 58 |
| Figure 2-12. eGFP-FLAG, eGFP-Ep9, and full-length N protein purity assessed by 10% SDS-PAGE | 59 |
| Figure 3-1. Potential OAS epitopes for binding α Ep9 Abs suggested by bioinformatics and tested by phage ELISA | 69 |
| Figure 3-2. Cross-reactive Ab binding to both Ep9 and EpNeu, and EpNeu epitope prediction..... | 73 |
| Figure 3-3. Early upregulation of α Ep9 IgGs..... | 93 |
| Figure 3-4. Expression of phage-displayed and eGFP-fused potential AIM epitopes | 93 |
| Figure 3-5. Optimization of assay to determine cross-reactivity of α Ep9 Ab to Ep9 and EpNeu | 94 |
| Figure 3-6. Linear and structural epitope mapping prediction of Influenza A H3N2 Neuraminidase..... | 95 |
| Figure 4-1. Insulin receptor structure and biosensor constructs..... | 104 |
| Figure 4-2. Expression and characterization of InR constructs in CHO cells..... | 106 |
| Figure 4-3. Humalog®-dependent change in FRET signal of CHO cell expressed InR C4 | 109 |
| Figure 4-4. Phage display and insulin binding of InR constructs. A and C..... | 111 |
| Figure 4-5. Improving insulin binding using site-directed library selections | 112 |
| Figure 4-6. E. coli-expressed L1 domain disrupts InR solubility and insulin binding | 114 |
| Figure 4-7. Designing compact, soluble, and specific insulin binding constructs | 116 |
| Figure 4-8. A two-protein FRET Assay design..... | 119 |
| Figure 5-1. Schematic of the FRET calcium ion probe | 145 |
| Figure 5-2. Fabrication of the membrane-encapsulated optical fiber (MEOF)..... | 147 |

| | |
|---|-----|
| Figure 5-3. Examining the permeability of the PEGDMA 2000 hydrogel..... | 150 |
| Figure 5-4. SEM images of PTFE and the filter-membrane..... | 151 |
| Figure 5-5. Properties of the filter-membrane..... | 151 |
| Figure 5-6. Efficacy of Twitch2B to sense calcium ions | 153 |
| Figure 5-7. STC Ca ²⁺ response in a fluorescence plate reader normalized by Alexa Fluor 488 peak intensity | 154 |
| Figure 5-8. Ca ²⁺ sensing with the MEOF probe | 156 |
| Figure 5-9. Sensor Isolation from Proteinase K..... | 158 |
| Figure 5-10. Swelling studies of PEGDMA 2000 | 163 |
| Figure 5-11. Spectroscopic Benchtop System | 164 |
| Figure 5-12. GFP encapsulation within the MEOF | 165 |
| Figure 5-13. Selective permeability of the filter-membrane | 166 |
| Figure 5-14. FRET properties of non-photobleached and photobleached Twitch2B samples..... | 166 |
| Figure 5-15. Nucleotide sequence encoding the STC protein | 167 |
| Figure 5-16. Plasmid Map of the gene encoding the STC protein in the pRSETB vector | 167 |
| Figure 5-17. Production of STC..... | 168 |
| Figure 6-1. Optimization of the VAIA method | 175 |
| Figure 6-2. Generalization of VAIA to three common immunoassay formats..... | 177 |
| Figure 6-3. Generalizability of VAIA with different blocking agents | 178 |
| Figure 6-4. Clinical potential of VAIA with biofluids..... | 179 |
| Figure 6-5. SDS-PAGE analysis of eGFP-FLAG expression and purification..... | 183 |
| Figure 6-6. Images of HSA-phage VAIAs with varying speeds and assay times..... | 184 |
| Figure 6-7. Images of eGFP-FLAG VAIAs..... | 185 |
| Figure 6-8. Images of conventional eGFP-FLAG IAs..... | 186 |
| Figure 6-9. Image of VAIA assay completed in a variety of biofluids | 188 |
| Figure 6-10. Images of patient plasma VAIAs..... | 188 |

LIST OF TABLES

| | |
|--|-----|
| Table 2-1. Phage-displayed putative epitopes of SARS-CoV-2..... | 60 |
| Table 2-2. Demographics and clinical characteristics of COVID-19 patients categorized by α Ep9 Abs response | 61 |
| Table 3-1. Potential “original” epitopes targeted by α Ep9 Abs | 91 |
| Table 3-2. Primers used to subclone potential original epitopes | 92 |
| Table 4-1. Gibson Assembly primers and template used for CR library Cloning..... | 134 |

ACKNOWLEDGEMENTS

I would like to thank the American Chemical Society for their permission to include published materials into Chapter 5 of my dissertation. Additionally, the investigations described in Chapter 2 is adapted from a manuscript published in mSphere, (DOI:10.1128/mSphere.00203-21). Chapters 3 and 6 are adapted from submitted manuscripts, currently in review. Thank you to all the co-authors for authorizing the use of these manuscripts towards my thesis. The work outlined in my dissertation would not have been possible without the contributions of my co-authors, whose names and contributions are listed before each Chapter. I would like to also thank the Juvenile Diabetes Research Foundation (2-SRA-2017-330-Q-R), Helmsley Charitable Trust (2018PG-T1D008) UCI COVID-19 Basic, Translational and Clinical Research Fund (CRAFT), the Allergan Foundation, and UCOP Emergency COVID-19 Research Seed Funding and UCI Graduate Division (Public Impact Fellowship) for their financial support of the research described in the dissertation. The investigations described in chapters 2 and 3 were born as a result of COVID-19 pandemic that impacted the lives of millions around the globe. I would like to extend my sympathies to those who has lost loved ones and express my deep gratitude to the 186 patients who donated their blood samples to make our research possible.

I would like to express my gratitude to Prof. Gregory Weiss for handing me a key into the world of protein engineering and chemical biology, and making me feel at home there. Thank you for taking a chance on a wide-eyed student who had never before purified a protein. Thank you for creating an immensely supportive and enriching environment for me to grow as a scientist, a mentor and a person. Your optimism and enthusiasm have inspired to me tackle many an obstacle with excitement and confidence. Your passion for learning has taught me to observe and appreciate science beyond the boundaries of a subject or a lab bench. For that lesson and many more that I take with me, I am truly grateful.

I would like to thank the members of my thesis advisory committee, Dr. Jennifer Prescher and Dr. Thomas Poulos, for their time and invaluable insights on navigating some very challenging projects. I have always walked away from our meetings with a fresh perspective and renewed enthusiasm for my research. Thank you as well to Prof. Chris Hughes and Prof. Michelle Digman for their very helpful insights during my pre-advancement committee meeting. Thank you as well to Dr. Wenqi Wang for his guidance in conducting cell-based assays.

During my PhD I have been incredibly fortunate to have the opportunity to collaborate with Prof. Elliot Botvinick's Lab. Prof. Botvinick, I am truly grateful for your mentorship and encouragement through a very challenging and exciting project. Thank you for the many science and non-science conversations that really helped provide new perspectives. Thank you for welcoming me into your lab, involving me in outreach and letting me be an extended part of the BEAMS family. I would also like to thank all the members of the Botvinick lab for being so welcoming and helping bridge the gap between proteins and

bioengineering. Dat Nguyen, thank you for the many lunch hours spent pouring over data and helping me better understand the amazing work you do. Thank you as well to Danielle Behrens, Toni Wilkinson, and Micah Lawrence for making the collaboration seamless, it has been a real pleasure working with you all. I would also like to thank Rachel Gurlin, for her continued friendship and support.

I would like to thank all the members of the Weiss lab, each of you have made the past 5 years of my life memorable and meaningful. Thank you, Dr. Sudipta Majumdar, for your continued mentorship and support. I am so grateful for your invaluable guidance that ran the gamut from teaching me how to purify my very first proteins to advice about the future. Thank you for being so patient while I learned the ropes, and always making time for me. I would like to thank my lab mentors, Dr. Joshua Britton, Dr. Luz Meneghini, Dr. Mark Richardson, Dr. Mariam Iftikhar, Dr. Josh Smith and visiting researcher, Dr. Bjarne Vincents for their continued friendship and guidance. Each of you has made a significant impact in shaping my development as a scientist, and I am very grateful for that. I am also very grateful to have had the opportunity to work closely with recent graduates, Dr. Kristin Gabriel and Dr. Emily Sanders, during the pandemic. I have learned so much from both of you during the experience. The work we did was real testament to the collaborative spirit and enthusiasm of the team. I am so excited to see you both soar! I would also like to thank members of the Martin lab, Prescher Lab, Nowick Lab, Edinger Lab and Wang Lab for sharing their equipment and expertise to help meet our research goals.

To the insulin biosensor protein engineering team, The InRmy, thank you so much for your incredible hard work and for weathering the many challenges of this project. Together, we learned to be creative, optimistic and better scientists. I am so grateful to you all for celebrating the wins and shouldering the burdens. I am so proud of the amazing undergraduates that have been a part of this project, Chris (Toto) Totoiu, Katrin (Buffy) Safronyuk, Linette Medina, Sahil Ahuja, and Derek Tang. I have watched you grow into strong and incredible scientists, with bright futures ahead. A special thank you to Keertna Bhuvan for her immense dedication to the project, for contributing many creative ideas, and staying late into the night to troubleshoot assays. I am so proud of the work you have all done. Thank you as well to Kim Weiss for all her work on the project at a time when I really needed it, you have been a great source of support and encouragement during this journey. I would like to thank visiting members, Emily Crawley and Dr. Bjarne Vincents, as well as lab alumni Jessica N. Pham for their significant contributions to the project. Thank you as well to Brian Miller and Teodora Nedic it has been a great pleasure working together. I am so excited to see the new and amazing directions you take the project!

To Dr. Rebekah Dyer, Dr. Gaetano Speciale, and Dr. Madison Fletcher, your friendship over the years has meant so much to me. I am so grateful to you for your mentorship and support in everything related to the science and beyond. I have learned, and continue to learn so much from you. Thank you for the many, many times that each of you has dropped everything to lend a hand, listen to an idea, or give a pep talk. Thank you as well for the Hallmark movies, comfort food and amazing desk art!

To the mentors and fellows at the Research Translational Group (RTG), thank you so for teaching me about science the beyond the bench. The skills learned here have been invaluable to shaping the way I approach science problems. I'd like to especially thank Dr. Lisa Stutts. Lisa you have been an amazing and inspiring mentor. Thank you so much for your immense support in so many endeavors. I'd also like to thank Dr. Stephen Ro, for his mentorship and guidance.

Thank you to all the AWIS@UCI members for helping create a community that I know will last well beyond my time at UCI. Rebekah, I am so glad we got to build this together. Thank you Dr. Celia Goudling, Dr. Aimee Edinger, and Dr. Lisa Stutts for your continued mentorship throughout the process of building this organization. Thank you GPS-STEM, Prof. David Fruman, Harinder Singh and Emma Flores for providing us graduate students such a vital resource for preparing us for life after a PhD. I look forward to putting to used everything I have learned. Thank you as well to Dr. Amy Hopkin, for being a wonderful and extremely supportive mentor. I would also like to thank Dr. Aleksander Hinek and Dr. Gary Sweeney for their mentorship and support through the years. They have played a significant role in shaping my future as a scientist.

To my CMB cohort, you have been my home, away from home. We've been a family from the day I arrived in SoCal, and you will all stay with me where ever I go. Lily, thank you so much for everything you do, you have the biggest heart of any one I know. To the friends that have lasted almost a lifetime Pooja, Penny, Carol, Phuong, Esther, Sana and Uma thank you for cheering me on from different time zones, staying up late into the night to call, travelling far and wide just to hang out for a weekend.

I would like to thank my parents, Animesh and Sathi, who are responsible for where I am today and for the opportunities that I had along the way. Thank you for a lifetime of unwavering support, and showing me what it means to be resilient. Thank you for teaching me how to turn bumpy rides into fun adventures and to never stop exploring.

I would like to thank my husband, Pratik, who has been by my side since I started my journey at the lab bench as an undergraduate. Thank you for staying up late with me for so many nights and for learning a lot of molecular biology along the way. I am so glad to have shared this journey with you. Thank you for supporting me every step of the way. I really could not have done this without you. I am so excited for what life has in store for us next!

CURRICULUM VITAE

SANJANA SEN

EDUCATION

| | |
|---|-------------|
| PhD, Molecular Biology and Biochemistry, University of California, Irvine | 2015 – 2021 |
| Masters of Science (MSc), Laboratory Medicine & Pathobiology, University of Toronto | 2009 – 2011 |
| Bachelor of Science (BSc) High Distinction, Human Biology, University of Toronto | 2004– 2008 |

RESEARCH EXPERIENCE

PhD Candidate, Molecular Biology and Biochemistry
UC, Irvine, California
PI: Prof. Gregory Weiss
2016 –2021

Project 1 (Published) | Using phage display, identified epitope associated with COVID-19 disease severity. Developed a risk score for predicting disease severity in a subset of COVID-19 patients.

Project 2 (in review)| Identified the mechanistic involvement of anti-SARS-COV-2 antibody in disease severity. Used bioinformatics server tools to predict putative cross-reactive epitopes and developed ELISA-based assays to validate epitope predictions for Ab cross reactivity.

Project 3| Used protein engineering techniques towards developing an insulin biosensor for type-1 diabetes disease management. Designed mutagenesis libraries and selection strategies towards increasing target binding and solubility. Devised high throughput binding assays using ELISA and FRET-based screening platform.

Project 4 (published)| Developed a photostable variant of a real-time FRET calcium sensor using genetically-encoded bioconjugation tags. Optimized and characterized soluble expression and bioconjugation of variant.

Project 5 (in review)| Developed assay design and binding targets for testing rapid immunoblot testing platform. Validated and tested multiple assay strategies and conditions.

Project 6 (published)| Optimized NHS-bioconjugation reaction conditions to achieve monovalent biotinylation of insulin for comparisons against novel conjugation methods. Characterized bioconjugation efficiency using mass spectrometry.

Graduate Research Assistant, Cell and Molecular Biology
York University, Toronto
PI: Prof. Gary Sweeney
2012 – 2014

- Conducted investigations of multiple projects examining the therapeutic potential of the hormone adiponectin in insulin resistance. Research resulted in two co-authored publications.

Clinical Research Intern, Cardiology 2012
Apollo Gleneagles Hospital, India
PI: Dr. Rabindra Nath Chakraborty

- Conducted retrospective correlation study, using the medical history of patients ≤ 40 yrs diagnosed with coronary artery disease (CAD) between 2007 and 2011

MSc. Graduate Researcher, Laboratory Medicine and Pathobiology 2009 – 2011
University of Toronto,
PI: Prof. Aleksander Hinek

Project 1 (published) | Established novel molecular signaling pathway involved in vascular and dermal integrity in patients with Costello syndrome. Tested Phase II clinical pharmaceuticals as potential therapeutics

Project 2 (published) | Investigated the role of insulin in extracellular matrix production and in promoting vascular integrity.

HONORS AND AWARDS

- UCI Public Impact Distinguished Fellowship 2019-2020
- Beall Applied Innovation Business Pitch competition: 2nd place 2019
- NSF I-Corps BioEntrepreneurship Fellowship 2019
- UCI Beall Applied Innovation Fellowship 2018 – 2019
 - *Completed >30 life science technology transfer and prior art reports towards commercialization of UCI intellectual property*
- American Society of Biochemistry & Molecular Biology 2018
 - *Travel Award*
- Laurence Becker Symposium: Poster Prize 2010
- University of Toronto Scholarship for Academic Excellence 2005 – 2007
- The Gerald J. Laurendeau Memorial Scholarship 2005
- Queen Elizabeth II Aiming for the Top Scholarship 2004 – 2006

PROFESSIONAL DEVELOPMENT AND COURSES

Graduate Professional Success (GPS)-STEM 2016 – 2021
Graduate Student & Post Doc Council 2018 – 2019
The Business of Science Certificate Program 2019
Introduction to Data Science, UCI Continuing Education 2018
Regulatory Requirements for Pharmaceutical Products, UCI continuing education 2018

COMMUNITY OUTREACH AND LEADERSHIP

Association of Women in Science at UC, Irvine (AWIS@UCI) 2018 – 2020
Founder and President
The Savvy Diabetics, Juvenile Diabetes Research Foundation 2017 – 2019
Juvenile Diabetes Research Foundation, Orange County Chapter 2017 – 2019
Canada-Korea Diabetes Research Association 2013 – 2014
Ethnic Minority Health Initiative 2011 – 2014

POSTERS AND PRESENTATIONS

- Towards the development of a real-time insulin biosensor: Engineering the insulin receptor. Poster Presentation. American Chemical Society National Meeting. Aug 2019, San Diego, California
- Real-time Insulin Detection: Engineering the Insulin Receptor. Poster Presentation. American Society of Biochemistry and Molecular Biology Experimental Biology Conference. Apr 2018, San Diego, California
- Real-time Insulin Detection: Engineering the Insulin Receptor. Oral Presentation. Juvenile Diabetes Research Foundation Chapter Meeting. Feb 2018, Newport Beach, California
- Elastogenesis: New Developments and Roles in Regenerative Medicine. Oral Presentation. SickKids Hospital Cardio-Pulmonary Medicine: Seminar Series. Oct. 2010, Toronto, Ontario

PUBLICATIONS

* authors contributed equally

- **Sen S**, Sanders EC; Santos AM, Bhuvan K, Tang DY, Gelston, AA, Weiss, GA. Evidence for Deleterious Antigen Imprinting in SARS-CoV-2 Immune Response . *In review*
- Sanders EC, **Sen S**, Gelston AA, Santos AM, Bhuvan K, Tang DY; Luo X, Raston CL, Weiss GA. Under Five Minute Immunoblot Assays by Vortex Fluidic Device Acceleration. *In review*.
- **Sen S***, Sanders EC*, Gabriel KN*, Miller BM, Isoda HM, Salcedo GS, Garrido JE, Dyer RP, Nakajima R, Jain A, Santos AM, Bhuvan K, Tifrea DF, Ricks-Oddie JL, Felgner PL, Edwards RA, Majumdar S, Weiss GA. Predicting COVID-19 Severity with a Specific Nucleocapsid Antibody plus Disease Risk Factor Score. *mSphere*. 6:e00203-21.
- Nguyen D, Behrens DM*, **Sen S***, Najdahmadi A, Pham JN, Speciale G, Lawrence MM, Majumdar S, Weiss GA, Botvinick EL. A Photo-stable and Proteolysis-Resistant FRET-Based Calcium Biosensor, *Anal Chem*, 2020, 92(11): 7683-7689.
- Richardson M*, Gabriel K*, Garcia J, Ashby S, Dyer R, Kim J, Lau C, Hong J, Le Tourneau R, **Sen S**, Narel D, Katz B, Ziller J, Majumdar S, Collins P, Weiss G. Pyrocinchonimides Conjugate to Amine Groups on Proteins via Imide Transfer. *Bioconjug Chem*, 2020, 31(5):1449-1462.
- Liu Y, **Sen S**, Wannaiampikul S, Rengasamy P, Hoo RLC, Isserlin R, Bader GD, Tungtrongchitr R, Deshaies Y, Xu A and Sweeney G. Metabolomic profiling in liver uncovers lysophospholipid metabolism as an important target of adiponectin action in mice. *Biochem J*. 2015, 469(1):71-82.
- Vu V, Liu, Y, **Sen S**, Xu A, Sweeney G. Tissue specific delivery of adiponectin gene to skeletal muscle using ultrasound targeted microbubble gene delivery is sufficient to improve whole body glucose homeostasis. *Am J Physiol*, 2013, 304(2): E168-E175.
- Shi J, Wang A, **Sen S**, Wang Y, Kim H, Mitts TF, Hinek A. Insulin Promotes Elastin Production by Human Aortic Smooth Muscle Cells and Fibroblasts from Skin and Heart. *Am J Pathol*. 2012, 180(2):715-26.
- **Sen S***, Bunda S*, Shi J, Wang A, Mitts TF, Hinek A. Retinoblastoma Protein Modulates the Inverse Relationship Between Cellular Proliferation and Elastogenesis. *J Biol Chem*. 2011, 286: 36580-36591.

ABSTRACT OF THE DISSERTATION

Protein Engineering Approaches to Biomarker Identification and Biosensor Development

by

Sanjana Sen

Doctor of Philosophy in Molecular Biology & Biochemistry

University of California, Irvine, 2021

Professor Gregory A. Weiss, Chair

Numerous factors, such as genetics, medical history, environment, and lifestyle contribute to patient responses to disease conditions. Approaches in precision medicine focus on examining disease in the context of overall patient health, so that prevention and effectively tailored treatments result in improved outcomes. This thesis describes investigations of disease characterization and tool development geared towards uses in precision medicine.

The COVID-19 pandemic has opened new avenues of investigation for the scientific community. The most important of which is understanding why patients have such diverse responses to infection. Here, we develop an M13-bacteriophage display-based screening platform to understand how antibody responses play a role in foreshadowing and driving disease severity in a subset of patients. We further characterize disease response in the context of co-morbidities and viral infection history and find that both factors can negatively impact disease severity in patients with a specific antibody response. Additionally, sensitivity of these patients to rising levels of cytokines suggest specific routes of preemptive treatments that may be beneficial. We, additionally, develop a platform for rapid (<5

mins) point-of-care detection, which can be used to assist in multi-analyte testing or in urgent triaging.

Precision medicine is also important in addressing changes in disease response over time. Real-time automated monitoring of disease markers can help improve disease management and reduce adverse side effects. Using techniques in protein engineering, we harness the specific binding interactions of native proteins to develop biosensors for managing highly dynamic disease conditions. Firstly, we engineer the human insulin receptor towards developing an optical insulin biosensor to improve automated type-1-diabetes disease management. The investigation focuses on the development of soluble, compact insulin-binding variants. We further test various assay designs for detecting insulin-dependent optical signaling. Secondly, we engineer a photo-stable version of a previously developed FRET-based calcium sensor, Twitch 2B. The new version of this sensor demonstrates improved stability and performance within an optical-fiber-based implantable device. Such a sensor can be applied to early diagnosis of hypocalcemia during blood transfusions.

Taken together, these investigations explore a wide range of protein engineering techniques across multiple expression platforms. By harnessing native protein-protein interactions, we build prognostic and disease management tools that can contribute to advancing the field of precision medicine. Additionally, we develop optical detection platforms and rapid antibody screening methods that can be used to dynamically monitor and explore crucial molecular interactions in disease.

CHAPTER 1

Antibody-dependent regulation of COVID-19 disease outcomes

1.1 Abstract

COVID-19 patients present a broad range of disease outcomes. Emerging investigations of COVID-19 antibody (Ab) profiles and their downstream Fc-receptor functions provide insights into the underlying mechanisms that can help predict and treat severe disease. This review, discusses the various aspects of COVID-19 Ab regulation and the molecular mechanisms that lead to severe disease. Specifically, Ab neutralization capacity, epitope-selection, immunological history, and the regulation of Ab subclasses, titer, timing and post-translational modifications are examined here. The combined effect of multiple factors are thought to tip the balance in favor of either protection or deregulation, thereby resulting in diverse outcomes ranging from asymptomatic to fatal. Additionally, we examine the role of Abs in characterizing disease pathways and predicting outcomes towards goals of early intervention and treatment.

1.2 Introduction

The COVID-19 pandemic has infected over 211 million people and led to the deaths of over 4.4 million globally, as of August 20, 2021¹. The virus responsible for the disease, SARS-CoV-2, belongs to the *Orthocoronaviridae* family of enveloped, single-stranded, positive-sense RNA viruses that cause respiratory, enteric, and systemic diseases in humans and animals. SARS-CoV-2 is the most virulent of the three highly pathogenic human-host coronaviruses. The other two include the severe acute respiratory syndrome (SARS) and the Middle East respiratory syndrome (MERS), which led to the 2002-2003 pandemic and 2012

epidemic, respectively. All three virulent strains belong to the subfamily of *Betacoronaviruses*. Other highly prevalent and less virulent human CoVs (hCoVs) include the *Betacoronaviruses* HKU-1 and OC43 and *Alphacoronavirus* 229E and NL63². These hCoVs range in amino acid sequence similarity to each other, with SARS being the most homologous to SARS-CoV-2³. The collective understanding of hCoV infection can help determine roles of immune cross-talk from previous hCoVs exposures and identify pathways that are unique to SARS-CoV-2.

Compared to other hCoV infections, SARS-CoV-2-infected patients present with a dramatic range of symptoms. Investigations comparing disease severity loosely categorize patients into groups of asymptomatic, mild, moderate, and severe⁴⁻⁶. Approximately 40-45% of the patient population are asymptomatic after infection⁷, while another 40% of patients presented with an array of mild symptoms such as anosmia, diarrhea, and symptoms akin to the common cold, requiring no medical care or out-patient only care. About 15% of patients are hospitalized due to moderate systems including respiratory distress, pneumonia, cardiac arrhythmia, blood clotting, liver and kidney distress. An additional 5% of patients are severely ill and develop further exacerbated disease symptoms over time. These critically ill patients commonly present with acute respiratory distress syndrome (ARDS) or organ failure, requiring ICU or intubated ventilator assistance⁸. While certain factors and co-morbidities, such as age, sex, chronic inflammation, and metabolic disease, and cancers have been associated with increased risk of severe disease⁹⁻¹², they are insufficient in predicting disease outcomes and the mechanisms underlying their influence are not yet known^{13,14}.

Elucidation of molecular mechanisms involved in driving disease outcomes can highlight the importance of SARS-CoV-2 antibodies (Abs) in mediating disease outcomes^{4-6,15-18}.

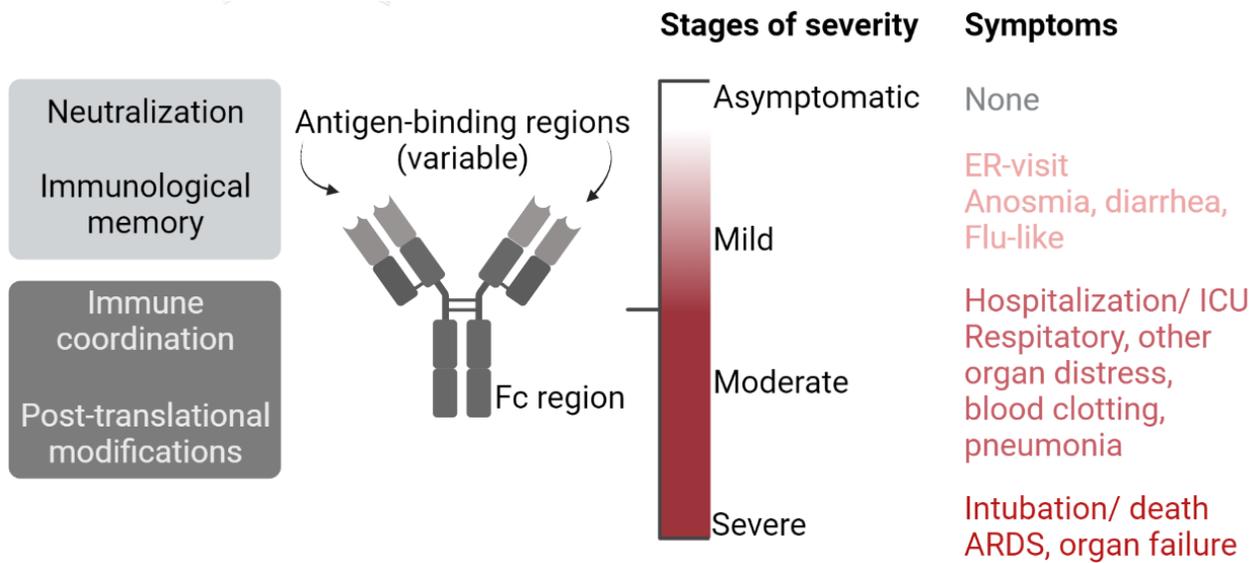


Figure 1-1. Factors impacting Ab-dependent immune regulation of COVID-19. Schematic represents the structure of an anti-SARS-CoV-2 Ab. Variable regions of the Ab are shown in light gray and constant regions are shown in dark gray (Middle). Factors that are dependent on antigen-Ab interactions via the variable domain (light gray box) and the Fc domain (dark grey box) are depicted on the left. Evidence shows that these factors influence Ab-dependent immune responses to SARS-CoV-2 and underlie the many diverse symptoms and outcomes observed in patients^{4-6,15-18}. Variability in disease outcomes is depicted on the right.

Like SARS, SARS-CoV-2 initiates viral entry by engaging the human ACE2 receptor using its spike surface glycoprotein (S). Abs that disrupt this host-virus interaction are termed neutralizing Abs (nAbs) and can protect against infections. The potential for direct protection by nAbs has resulted in the rapid development of neutralizing monoclonal Abs (mAbs) that target the S protein. There are currently six mAbs in the U.S., South Korean and Indian markets, with several in phase II/III trials¹⁹. While only the variable complement determining regions (CDRs) of the Ab are required for viral neutralization, the antigen-Ab complex also promotes binding interactions with the Abs' constant Fc region and Fc-receptors. Such Fc-receptor activation triggers downstream immune mechanisms that are

necessary for elimination of the virus and infected cell^{6,20-22}. Together with Cytotoxic T lymphocytes (CTL) responses, coordinated efforts between neutralization and Fc-mediated functions are important in protecting against viral disease. Deregulation of these pathways, however, can result in hyperinflammation and increased disease severity²³⁻²⁵.

Both neutralizing and non-neutralizing Abs activate Fc-mediated immune responses and have been shown to have both protective^{22,26,27} and detrimental^{4,5,17,21} effects on patient outcomes. While neutralization capacity can explain some components of COVID-19 disease response, recently investigations identify the importance of Ab characteristics beyond neutralization. In this review, we explore the different aspects Ab-mediated regulation that are implicated in COVID-19 and drive diverse disease outcomes (Figure 1-1).

1.3 The Role of Neutralization in Disease Outcomes

Neutralization is the most direct pathway to inhibit viral infection and epitope-selection plays a very important role. All neutralizing Abs have been shown to target epitopes on the S protein²⁸. The receptor binding domain (RBD) of the S protein directly contacts the ACE2 receptor on host cells and comprises the epitope for 90% of all known nAbs against SARS-CoV-2²⁸. The functional S protein exists as a trimer, with the RBD between two disulfide forming cysteines (C336 and C525) (Figure 1-2B). A short linear region within the RBD, termed the receptor binding motif (RBM), contains the majority of residues that contact ACE2. Structural analysis of RBD and nAb interactions identify five binding clusters²⁹. While some of these residues directly outcompete ACE2 binding, others inhibit receptor binding by destabilizing the prefusion conformation of the S protein or by causing steric hindrance through virion aggregation (Figure 1-2). Interestingly, due to the increased

sequence diversity in the RBM of SARS-CoV-2 and SARS-CoV-1, cross-reactive nAbs often target regions outside the RBM resulting in allosteric binding interactions^{30,31 32,33}.

With new variants spreading within populations³⁴⁻³⁷, the generation of Abs that target regions outside the RBM or at multiple neutralizing regions of S protein are more likely to effectively protect against viral infection. Niu et al. report that four Abs targeting different binding clusters could simultaneously bind a single RBD²⁹. The additional steric hindrance provided by the same Ab binding different RBDs within a S trimer makes a strong case for polyclonal or multiple mAb cocktail treatments, and highlights the importance of neutralizing Ab diversity early in infection. Currently, the REGN-COV2's mAb cocktail, authorized for emergency use by the FDA, contains two recombinant Abs, REGN10933 and REGN10987 that bind two different clusters on the RBD. While effective treatments, different COVID-19 strains have been shown to escape neutralization by either one or both of the Regeneron Abs^{29,38,39}. Though the prevalence of those strains are low (0.1-0.3% of sequences), identifying epitopes, binding competition, and mechanisms of nAbs can aid in the development of multi-target therapies effective against a wider set of current and future COVID-19 strains.

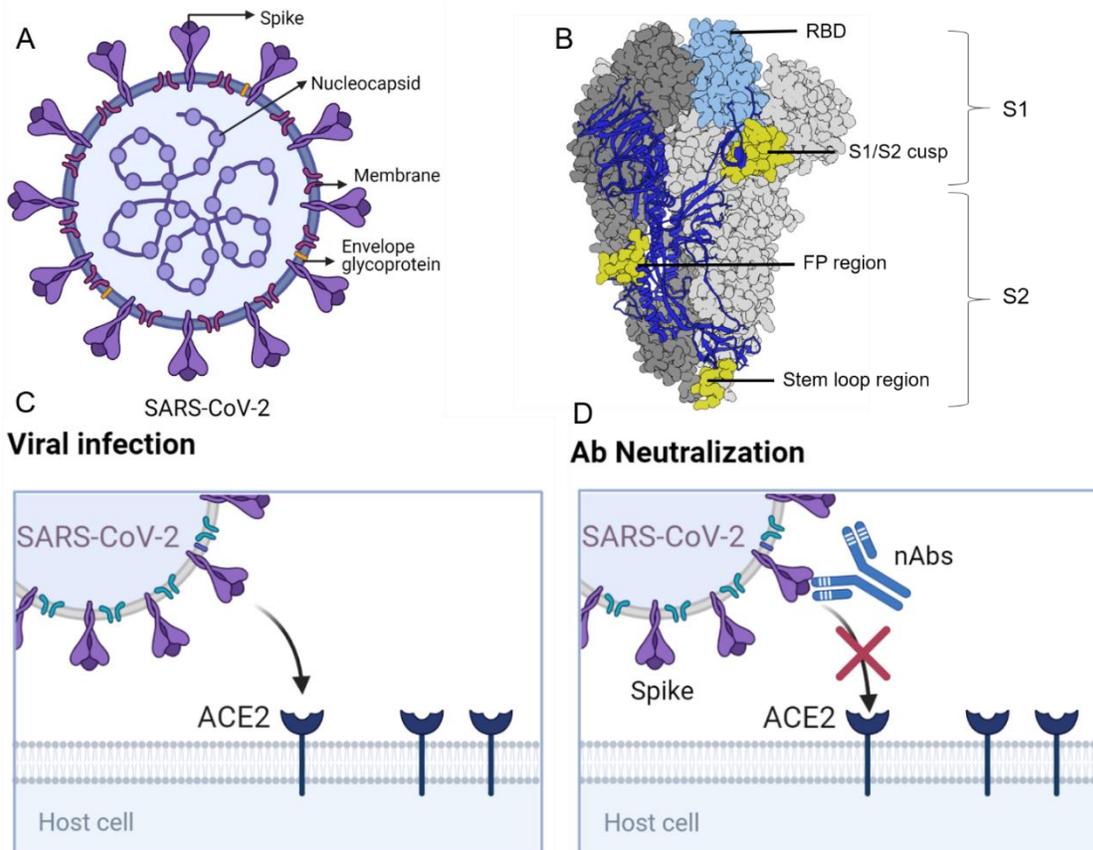


Figure 1-2. Neutralization can protect against SARS-CoV-2 infection. **A.** Schematic representation of the SARS-CoV-2 virus depicting its four structural proteins. The surface glycoprotein spike (S), exists as a trimer and is targeted by antibodies for neutralization. Abs target the SARS-CoV-2 proteins, nucleocapsid (N), membrane (M) and envelope (E). **B.** The spike (S) protein the surface glycoprotein that consists of two main domains the S1 and S2 and exists stably as a prefusion trimer conformation (PDB: 6VXX). The S1 domain contains the receptor binding domain (RBD), shown in blue. **C.** The RBD directly engages the human host cell receptor ACE-2, and its conformation changes to permit cell fusion and viral entry into the host cells⁴⁰. **D.** Abs that bind the S protein and hinder its interactions with the ACE2 receptor are called neutralization Abs (nAbs). While the RBD contains the majority of neutralization targeted sites, other sites can allosterically hinder ACE-2 binding or inhibit cell fusion. Sites targeted outside the RBD for neutralization ^{6,17,41,42} are depicted in yellow in panel B.

While investigations of nAbs have primarily focused on the RBD, nAbs can also inhibit viral entry by targeting regions outside the S1 domain, like the stem loop region in the S2 domain^{6,32,43,44}. A second non-RBD neutralization region resides in cleavage site at the S1/S2 junction targeted by the endoprotease, furin ^{41,42}. Abs that target furin cleavage sites are

likely to block access to the endoprotease and prevent the cleavage necessary for S protein conformational change and subsequent viral entry⁴⁵. Regions outside the RBD are more highly conserved amongst hCoVs, and the S2 protein greater conserved regions than S1⁴⁶ (Figure 1-2B). Therefore, neutralizing Abs targeting conserved regions are less likely to be mutated in new variants.

Interestingly, however, nAbs targeting a region just upstream of the furin cleavage site and downstream of the RBD, at residues 553-579 (Figure 1-2B) are associated with increased disease severity^{6,17}. The observation indicates that neutralization alone is not always sufficient to ensure disease protection. This association further suggests that Ab-mediated immune pathways and outcomes both protective and detrimental may be epitope-specific. The presence of high titers of neutralizing α S Abs in severe disease has been confirmed in numerous studies to-date^{6,21,47}, suggesting that high neutralization potency rather than titers of nAbs is correlated with increased survival²¹. While nAbs can have both neutralization and Fc-mediated functions, the neutralizing functions of SAR-CoV-2 Abs are the primary means of prophylactic and early infection protection before viral peak replication²⁷. However, beyond viral peak replication, Fc effector functions can be more important in mediating disease outcomes²⁷.

Since Fc-mediated effector functions can be activated, or deregulated, independent of Ab neutralization capabilities, the high titers of nAbs in severe disease are likely employed for Fc-dependent activities. Non-neutralizing Abs are typically categorized as non-protective, not only due to their lack of neutralization capabilities, but also because they have been associated with Ab-dependent disease enhancement (ADE)⁴⁴. ADE is a phenomenon in which antigen and non-neutralizing Abs engage with FcRs to promote live virus

internalization, leading to enhanced viral infection and hyperactivation of the complement system. First reported by Hawkes⁴⁸ in 1964, ADE has been described in many viral infections including HIV, Dengue fever, respiratory syncytial virus (RSV), and Zika virus²³⁻²⁵. Hallmarks of ADE have been observed in coronaviruses, including SARS and MERS, and literature has suggested, though not proven, that ADE may be of concern in COVID-19^{44,49,50}. High titers of non-neutralizing Abs to anti-nucleocapsid (α N) Abs, especially Abs targeting the epitopes in the RNA binding region 152-172, have been associated with increased disease severity. In contrast, non-neutralizing α SARS-CoV-2 Abs have also been shown to confer disease protection by activating Ab-dependent cellular cytotoxicity (ADCC). ADCC is a process in which FcR activation on immune effector cells programs them to target and kill virally infected cells²². Such protection by non-neutralizing Abs has also been observed in HIV^{51,52}, Ebola⁵³ and Influenza^{54,55}. With both neutralizing and non-neutralizing Abs associated with disease severity and protection. Exploring what factors that determine different immune pathways in different disease outcomes may help better predict, categorize and treat disease symptoms.

1.4 The Role of Immunological Memory in COVID-19 Progression

In addition to *de novo* epitope selection by Abs for the purposes of viral neutralization, Abs may target specific epitopes on a new pathogen due to its previous exposure to a similar antigen. Increasing evidence suggests that such immunological memory shapes Ab repertoire in response to SARS-CoV-2 and may influence disease outcomes⁵⁶. Antigenic imprinting (AIM) is a phenomenon in which memory B-cells (MBCs) generated against antigens from a primary infection are re-propagated in response to a new and similar, but not identical, infection with homologous antigens⁵⁷ (Figure 1-3).

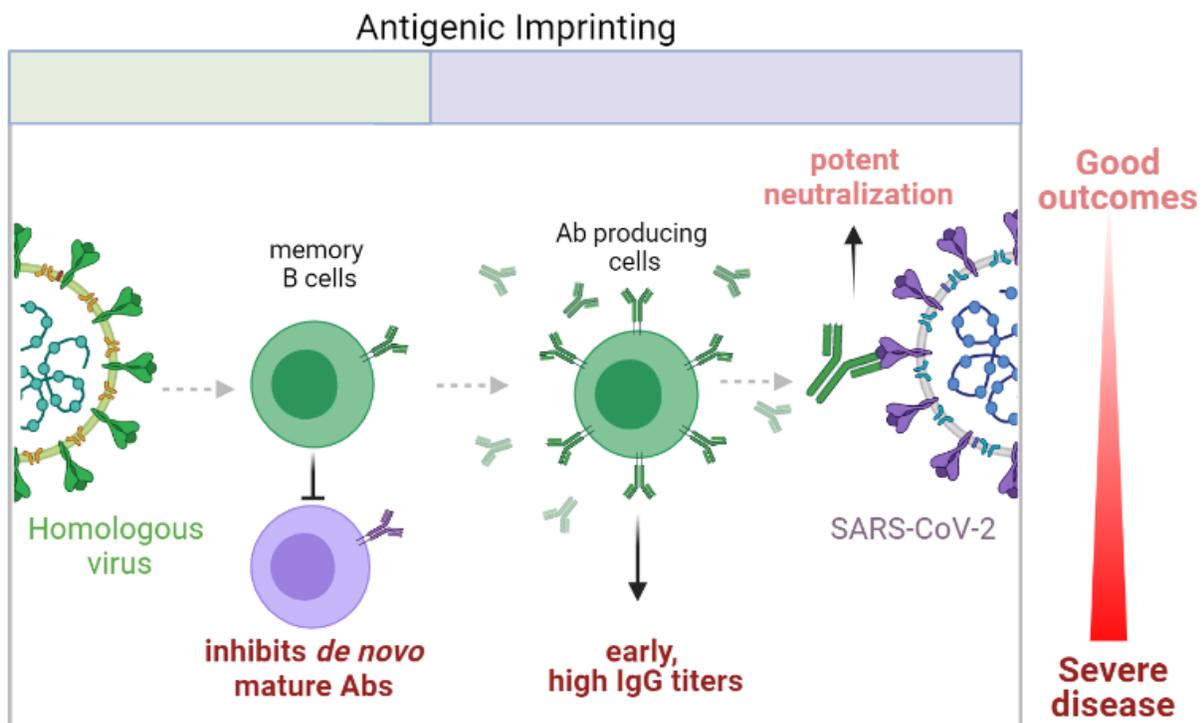


Figure 1-3. Antigenic Imprinting (AIM) can result in protective and detrimental effects. The schematic represents the phenomenon of antigenic imprinting (AIM). AIM occurs when antibodies generated against a previous infection are boosted in response to a new similar, but not identical pathogen. If the imprinted Abs are potently neutralizing (nAbs) then AIM can be protective by preventing viral entry and subsequent infection⁵⁸. If, however, the initial boost of IgG is too robust (high titers), it can instead lead to a loss of immune coordination and cause inflammation-related severe disease. Continued affinity maturation of memory B-cells (MBCs) from previous disease exposures can outcompete new Abs, thereby inhibiting the generation of specific, protective Abs⁵⁹. The balance between protective (light pink) and harmful mechanisms (dark red) may determine disease outcomes.

Outcomes of AIM can vary depending on whether the imprinted Abs can effectively neutralize the new pathogen. Ideally, AIM can generate an early mature humoral response resulting in rapid viral clearance. However, if viral neutralization is ineffective, AIM can be detrimental. Abs complexed with non-neutralized virus can aid in live viral internalization into host immune cells via the Abs' interactions with FcRs. Subsequent increases in viral burden and activation of hyper-inflammatory pathways can result; the particularly dire outcomes are termed as ADE^{24,57}. Additionally, AIM can hinder the *de novo* generation of Abs

specific for the new pathogen through outcompeting new Abs at the germinal center by high-affinity binding imprinting B cells⁵⁸ (Figure 1-3). Therefore, identifying epitopes and neutralization capability of Abs that target them is particularly important in the context of AIM. The hallmarks of AIM include 1) a high prevalence of memory B-cells (MBCs), 2) an early upregulation of affinity-matured Abs such as IgG and IgM 3) the prevalence of cross-reactive Abs with a pre-exposed pathogen^{18,58,60}. In cases where the outcome is detrimental, complement activation, cytokine storm markers, or high viral titers are also indicative of AIM-induced ADE⁵⁷.

Studies suggest that previous exposure to endemic hCoVs can improve disease outcomes via early upregulation of cross-reactive, neutralizing IgGs against conserved regions of SARS-CoV-2⁶¹. Epitope cross-reactivity analyses identify the S2 stem helix loop region as a vulnerable site for hCoV Ab cross-reactivity^{32,61,62}. Upregulated levels of cross-reactive Abs to this region and memory B-cells from endemic hCoVs were observed in SARS-CoV-2 convalescent individuals, suggesting a possible protective role⁶³. It is thought that these cross-reactive Abs originated from either HKU-1 or OC43 and has been shown to induce activity of Fc-mediated activation of receptors, FcγR2A and FcγR2B⁶³. Similarly, SARS-COV-2 patients were also shown to upregulate Abs cross-reactive with OC43, targeting the S2 furin cleavage site³², though the impact of these Abs on disease severity is not yet known. Cross-reactive Abs for the S2 epitopes were only present at low prevalence and low titers in pre-pandemic individuals, suggesting that the memory B-cell in serum likely boost the original Abs against endemic infections upon encountering SARS-CoV-2.

While some cross-reactive Abs from pre-pandemic individuals have benign effects against SARS-CoV-2, others suggest a role in disease enhancement⁶⁴. Several hallmarks of

AIM have been observed in COVID-19 patients with severe disease. Firstly, several studies report higher levels of α S and α N IgG titers and low viral load in severity impacted patients relative to patients with mild disease, within the first two weeks post-symptom onset (PSO)⁶⁵⁻⁶⁸. Conversely, severity has also been associated with high viral titers, suggesting different severe disease pathways⁶⁹⁻⁷¹. These high Abs titers are upregulated early after the onset of symptoms, within the first week, and have been shown to persist for up to 40 days⁷². Secondly, B-cell profile analyses demonstrated that patients with severe disease have greater MBC responses against the S protein and ORF8⁶⁸. Boosted levels of MBCs were also shown to have a greater affinity for hCoV antigens, providing strong evidence that the Abs originated from a previous exposure to endemic hCoVs infections⁵⁹. A similar boost is observed for α ORF1 Abs in COVID-19 patients pre-exposed to endemic hCoVs³². The mechanisms governing when antigenic imprinting can be protective or detrimental in SARS-CoV-2 infections are still unclear. However, some studies suggest that early upregulation of these previously generated (or “backboosted”) Abs may play a role in deregulating Fc-receptor functions and lead to disease severity. This topic will be discussed in the following section 4.59. Therefore, the balance between the protective effects of Ab neutralization and the harmful effects of high Ab titers and inhibited de novo Ab generation could determine the overall impact of specific AIM Abs on disease outcomes (Figure 3-1).

While related coronaviruses have been the main focus for examining Ab cross-reactivity, some studies suggest that conserved antigens from more distantly related viruses could potentially also generate cross-reactive Abs. For example, within a CoVID-19 patient cohort, hospitalized patients were more likely to have a history of human herpesvirus 5 or Herpes Simplexvirus-1 (HSV-1) infections than patients with mild symptoms³² and cross-

reactive Abs against these viruses have also been identified and associated with severe disease⁷³. The latter study, screened SARS-CoV-2 Abs for cross-reactivity to peptides from 1,276 strains of human viruses characterized an immune signature based on cross-reactive viral epitopes that can predict critical illness in COVID-19 patients based on their infection history⁷³. Work detailed in later chapters of the thesis predicts SARS-CoV-2 disease severity in a subset of patients by detecting an α N Ab that cross-reacts with an epitope from an influenza A strain⁵. A broad understanding of conserved epitopes in common human viruses could allow predictions of broader range of disease outcomes.

1.5 Deregulated IgG Responses Lead to COVID-19 Severity

In addition to Ab cross-talk, another main hallmark of AIM is an early mature upregulation of mature Abs. Interestingly, early IgG upregulation has been correlated with severe COVID-19 disease^{68,74,755,76}. Longitudinal comparisons of distinct serological patterns and downstream responses demonstrated two different patterns of disease severity – one that eventually leads to convalescence despite severe symptoms and the other to fatality⁴. The severe disease pathway that leads to recovery shows characteristics of detrimental AIM, such as early upregulation of α S-IgG1 and robust downstream inflammatory activation. After the initial early IgG1 upregulation within the first week, IgG3 and IgA1/2 are subsequently upregulated in the severe patients that recover compared moderately affected patients whose IgG and IgA levels were rising at a slower rate⁴. Interestingly, IgG3 Abs are known to potently induce effector functions and excessive or prolonged upregulation can lead to hyperinflammatory responses⁷⁷ (Figure 1-4A).

While patients with moderate symptoms show a steady and sustained climb in IgG levels, a delayed and transient response results in a high chance of mortality following severe symptoms⁴ (Figure 1-4B, 1-4B C). Though these severely impacted patients have an underdeveloped IgG response, they still upregulate IgA2 Abs in response to infection⁴. (Figure 1-4D) In COVID-19 patients, increased levels of IgA2 have been associated with organ injury and fatal outcomes attributed to the activation of hyperinflammatory pathways and neutrophil extracellular trap (NET), which together contribute to severe outcomes⁷⁸. In other studies, upregulation of α N IgG4 in severe COVID-19 has also been reported; IgG4 upregulation is usually observed in fibrotic diseases and can indicate lung injury and fibrosis⁶⁸. Taken together, the initial magnitude and length of time for IgG upregulation appears to be a marker of disease severity with varying outcomes. In the following sections, we further explore the role of IgG as a crucial molecular checkpoint for immune coordination and disease outcomes.

1.6 Deregulated Ab Timing and the Pro-Inflammatory Response

Although it is still unclear whether ADE occurs in COVID-19, research has shown that deregulated coordination of the immune response can impact disease severity in patients. IgGs induce immune effector functions by binding to low-affinity activating receptors Fc γ R2A, Fc γ R3A and Fc γ R3B, or the inhibitory receptor Fc γ R2B^{79,80}. The balance between FcR signaling in immune cells activates specific functional programs^{18,77}. Furthermore differences in the timing and prevalence of Ab subclasses with varied structures (ex. IgG1, IgG2, IgG3, IgG4) determine the overall intensities with which FcRs are engaged¹⁸. Downstream FcR and effector analyses of the patients with varied COVID-19 disease

outcomes and IgG profiles highlight the importance of Ab timing and balance in coordinating immune cellular effects.

In early IgG response patients, a similarly early α S IgG-mediated activation of the inhibitory receptor, Fc γ R2B is observed. This can, in turn, drive a broader overactivation of Fc γ Rs subtypes. Hyperactivation of Fc γ R2B in severely affected COVID-19 patients has been observed in multiple investigations^{81,82} and associated proinflammatory markers^{4,76,83}. Specifically, preferred Fc γ R2B activation over Fc γ R2A in severe COVID-19 results in monocyte reprogramming and the hyperproliferation of immature proinflammatory neutrophils (termed emergency myelopoiesis)⁸⁴. Fc γ R2B causes monocyte reprogramming by blocking the transcription of Fc γ R2A-mediated antiviral interferon stimulated gene (ISG) program and upregulating the expression of the proinflammatory calcium-binding S100A12 gene signature. While the upregulation of S100A12+ cells also occurs in patients with ARDS from other viral sources, the Fc γ R2B-mediated pathway in interferon signaling is unique to severe COVID-19 patients⁶⁵. This molecular mechanism may be the reason that disease trajectory in COVID-19 is especially sensitive to alterations in interferon signaling relative to other viral infections^{31,85-87}. Additionally, the already upregulated S100A12+ -mediated emergency myelopoiesis in metabolic comorbidities like diabetes^{88,89}, may further exacerbate disease symptoms in SARS-CoV-2 infected individuals (Figure 1-4D).

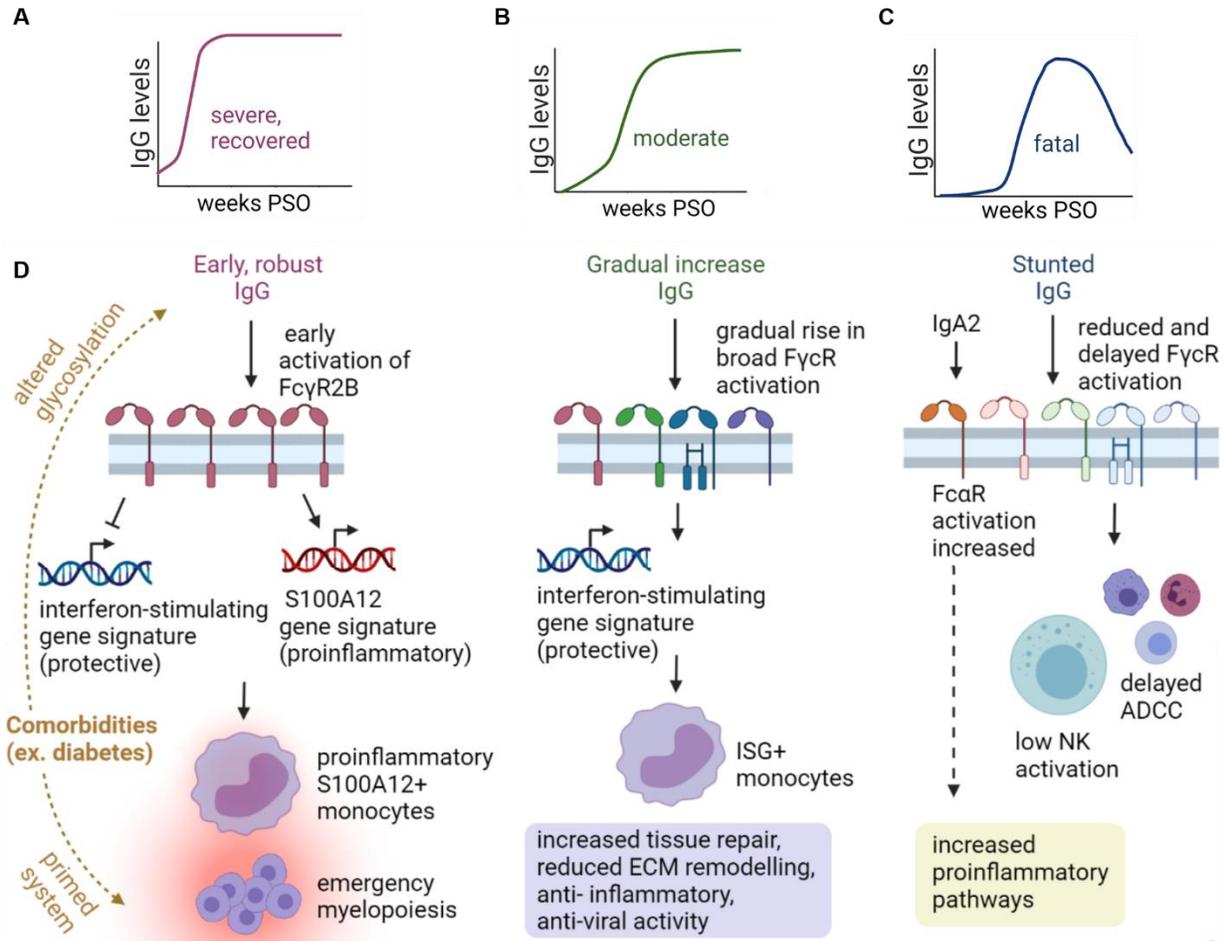


Figure 1-4 The timing and titers of IgGs are crucial to immune coordination. A-C. The schematic depicts different patient outcomes based on varying IgG trends over time. A robust, early IgG response results in severe disease with good prognosis for survival. A gradual IgG upregulation leads to moderate disease, whereas a delayed and stunted response results in severe disease with high risk of mortality. **D.** An early, robust IgG boost leads to an early activation of FcγR2B⁴, which in turn inhibits interferon stimulated gene (ISG+) signaling. Instead, proinflammatory genes (S100A12+) are upregulated^{65,84}. Such proinflammatory immune dysregulation in monocytes results in neutrophil proliferation (emergency myelopoiesis) and proinflammatory symptoms. Patients that have mild to moderate symptoms show increased ISG expression, high FcγR2A activation, coordinated effect functions involved in tissue repair and anti-viral activity^{27,65,81}. Conversely patients with stunted IgG responses show low FcγR activation and instead relies on IgA2 and its corresponding FcαR⁴. IgA2 hyperactivation can result in severe COVID-19 symptoms⁷⁸. Additionally, these patients also show stunted natural killer (NK) cell activities⁴. Therefore viral burden may have a greater impact on disease symptoms in this group. Changes in IgG glycosylation, specifically decreased fucosylation (termed afucosylation), have also been shown to affect FcγR activation profiles and is associated with severe disease^{18,20}. Patients with metabolic comorbidities like diabetes show a predisposition for afucosylation and emergency myelopoiesis and may exacerbate COVID-19 symptoms through these pathways¹⁸. PSO: post-symptom onset.

While the “robust and early” humoral response is well characterized in literature, less is known about the stunted immune response observed in patients with high mortality. In these patients initial levels of FcγR receptor activation are low and mirror that of moderate patients. This patient subset is, however, characterized by its αN-specific upregulation of IgA-specific Fc receptor (FcαR) at two weeks PSO. This αN IgA-mediated Fc activation has been shown to be a parameter for predicting disease survival⁴. While effector functions in these immune-stunted patients are delayed as well, by the third week PSO levels of most effector functions catch up to that of moderate and severe patient groups with high chances of survival. Natural killer (NK) cell-activation (ADNKA) levels, however, remain low over time⁴. NK dysfunction and exhaustion phenotypes have also been associated with increased COVID-19 severity⁹⁰⁻⁹², which may contribute to increased mortality in immune-stunted patients. Interestingly, a similarly blunted αS IgG response has been reported against the BioNT162b1 mRNA vaccine in multiple sclerosis patients treated with mAb treatment, ocrelizumab, targeting CD20 B-cells and in kidney transplant patients on immunosuppressant treatment⁹³⁻⁹⁵. Therefore, the stunted immune signature may represent patients with immune-suppressing comorbidities or those prescribed immunomodulating medications (Figure 1-4D).

Despite early similarities in Ab and FcγR profiles between patients with moderate disease and severe patients who do not recover, the long term effects are highly divergent. The IgG levels of patients with moderate disease climb steadily over time and match levels observed in patients with severe disease that eventually recover. Meanwhile these levels dramatically decline after the third week PSO in severe patients who eventually die from COVID-19⁴. Similar to the moderate symptom group, severe, but recovered patients also

demonstrate persistently elevated IgG levels over 4 weeks. While steady upregulation and sustained, high levels of IgGs appear to be important for survival⁴, other studies examining longer time points on more mildly infected patients demonstrate rapid decay in IgGs between 37-90 days PSO^{66,96,97}. Thus several subcategories of COVID-19 disease pathways exist that need to be carefully characterized (Figure 1-4 A-C).

1.7 Altered Post-Translational Modifications Can Drive Disease Outcomes

Another factor that regulates Ab function is post-translational modifications to the Fc region, which can also regulate the Ab's interactions with FcRs^{18,79,80}. PTMs at the IgG N-linked glycosylation site N297 are important for FcR selectivity and determining effector functions. The glycan composition at this site is highly variable, and the role of core fucosylation has been shown to impact disease outcomes^{20,98}. Comparisons of Fc-region post-translational modifications (PTMs) between COVID-19 patients with symptoms ranging from mild to severe report lower levels of Fc-fucosylation in α RBD IgG1 in severe patients. Longitudinal analysis of IgG1 has demonstrated that while afucosylation was stable for over 4 weeks, a significant decrease in Ab sialylation was also observed²⁰.

Interestingly, afucosylated Abs without a β 1,4-linked GlcNAc attached to the core β -mannose residue, termed F0N0, were particularly prevalent in severe male COVID-19 patients. In other studies, increased Fc galactosylation has been associated with age, male sex, and autoimmune diseases⁹⁹. The afucosylated Abs from severe COVID-19 patients increased activation of Fc γ R3A, where higher levels of afucosylation showed an approximately six-fold greater affinity for the receptor. Afucosylated Abs from severe patients upregulated natural kill (NK) cell activation and proinflammatory cytokine production, IL-6, TNF, and IL-1 β in monocytes and alveolar macrophages⁹⁸. While Ab

afucosylation-related hyperactivation FcγR3A has been previously associated with increased disease severity in dengue virus infections^{20,100,101}, the PTM signature is not as stable as SARS-CoV-2 and is observed to decrease over time¹⁰¹. Such longitudinal differences between different infections demonstrate that further points of regulation in immune coordination may determine disease progression and outcomes. The investigation of the role of Ab PTMs in COVID-19 is still limited, and future investigations may help answer questions of how risk factors such as age, sex, comorbidities alter the structure and functions of the Abs that are generated.

1.8 Conclusions and Discussion

Categorization by neutralization capability is insufficient for determining or treating disease outcomes. Ab neutralization may be most effective for prophylaxis²⁷ or in cases of low initial viral titer exposure^{102,103}. Rather than titers, a measure of nAb potency may be a better predictor of disease outcomes and help explain the differential protection observed²¹. As titers of Abs increase with severity, the increased binding between Abs and antigen results in concerted FcR aggregation and subsequent activation⁸⁰. Therefore, in severe disease, the balance between neutralization potency and Fc-mediated activity shifts towards hyperactivating downstream effector functions.

A central paradigm of innate immunity asserts that early antiviral responses occur first, and are then followed by pro-inflammatory actions. Patients hospitalized for severe COVID-19 seem to break the mold, where early pro-inflammatory responses and a delayed antiviral, IFN-mediated response are observed^{87,104}. The timing, titers, and subclasses of IgG in COVID-19 immune response plays a critical role in modulating disease outcomes. An early upregulation of IgGs and the subsequent hyperactivation of Fcγ2B, instead of the protective

Fcγ2A, results in upregulated proinflammatory pathways in monocytes and neutrophils^{4,65,84}. These patients present with severe disease with a high chance of recovery⁴. Immunological history, through phenomena like AIM, likely underlies the above described response⁵⁹. Similarly, changes in Ab PTMs like decreased fucosylation, which are prevalent in severely impacted SARS-CoV-2 patients, are also associated with the hyperactivation of another Fc-receptor, Fcγ3A⁹⁸. Therefore, disease severity in COVID-19 may be governed by factors that shunt Fc-receptor activation away from the protective, early antiviral Fcγ2A-mediated immune responses^{27,65,81}.

Additionally, the protective effects of FcγR-mediated responses in COVID-19 are apparent in patient cohorts with a stunted IgG response. Here, delayed and transient Fc-mediated effector functions resulting in weak NK activation is associated with patient mortality⁴. Since there have been conflicting observations of whether viral load correlates positively with disease severity, it is possible that the subset of patients with an inability to mount an appropriate Ab response succumb to the COVID-19 due to the effects of increased viral spreading. Furthermore, severity through this pathway may occur in patients that have immune suppression or are treated with immune-moderating medications⁹³⁻⁹⁵. Therefore, understanding mechanisms of action for different groups of patients can help predict and optimize therapeutic vaccine development.

Due to the clinical variability and molecular diversity of COVID-19 disease severity, a global test to predict disease severity is still challenging. Additionally, due to the complex molecular coordination and many points of regulation, a single biomarker snapshot at any given time PSO may not accurately depict immune skewing. In such cases, biomarker

signatures, or a combination of comorbidities and immune biomarkers may help determine risks of the severe disease down the line. While severity associated with early Ab upregulation may be more simple and practical to assay, for pathways characterized by delayed immune responses, predictions may be difficult. Here, early markers appear very similar to mild disease symptom and the delayed signature may only appear after the onset of critically illness. Investigations show that certain cross-reactive epitopes to detect the upregulation of imprinted Abs can be used to predict AIM-based disease outcomes^{5,73}. Studies of distinct early Ab, FcR, and effector biomarkers were have been used to decipher signatures of moderate, severe, and critical (low survival) patients. However, these signatures require testing that uses multi-biomarker platforms and will be more difficult to implement at the point-of-care ^{4,15}.

The main limitation of predicting disease severity with high sensitivity (low false negative rate) is that the mechanisms of most disease outcomes have not been identified. Furthermore, investigations attempting to characterize disease mechanisms typically group outcomes differently. For example, while some studies group patients' severity based on specific clinical presentations, others base the criteria on survival, yet others base it on the level of medical care received⁴⁻⁶. While this allows investigation of different parameters for disease severity, comparisons between otherwise similar investigations are often impossible. Standardization of disease outcome categories in future investigations in the field can aid in the facile discovery and validation of new pathways and predictive biomarkers of COVID-19-related disease outcomes.

In the following chapters, we further explore the role of α SARS-CoV-2 Abs in disease severity. In a quest to find correlations between Ab responses and disease outcomes, we empirically test panels of bioinformatics-predicted epitopes using a phage display binding platform. Our results identify and characterize disease markers of a subset of patients that respond to a specific epitope on the N protein. Using this Ab biomarker we develop a two-step test, including Ab detection and a disease risk factor score, to categorize patients with high chances of developing severe disease (Chapter 2). Next, we explore the role of immunological history in shaping disease outcomes in this group of high risk patients. The findings of our investigations show evidence of antigenic imprinting as an underlying mechanism of disease severity. Furthermore, our investigation provides insight into how epitope-specificity even in non-neutralizing Abs can be important in driving particular disease outcomes (Chapter 3). Additionally, we develop a rapid means of detecting plasma proteins using a device called the Vortex Fluidic Device (VFD), (Chapter 6). Such methods can be crucial to quick, high-throughput detection of specific biomarkers of COVID-19 disease severity, leading to early intervention.

Towards the same goals of rapid detection and early prevention, in Chapters 4 and 5, we expand our tool kit for analyzing protein-protein interactions by harnessing the binding interactions of native proteins to engineer real-time optical sensors. In Chapter 4, we develop platforms to assess the binding of insulin and insulin receptor variants towards the development of a biosensors that can improve type-1-diabetes management. In Chapter 5, we develop an improved calcium sensor, originally derived from Troponin C¹⁰⁵, for the real-time monitoring of fluctuating ions and early intervention in patients undergoing blood

transfusions. Together the investigations, work to use protein-based tools to improve personalized medical care.

1.8 References

1. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **20**, 533–534 (2020).
2. Llanes, A. *et al.* Betacoronavirus genomes: How genomic information has been used to deal with past outbreaks and the covid-19 pandemic. *Int. J. Mol. Sci.* **21**, 1–28 (2020).
3. Chen, Z., Boon, S. S., Wang, M. H., Chan, R. W. Y. & Chan, P. K. S. Genomic and evolutionary comparison between SARS-CoV-2 and other human coronaviruses. *J. Virol. Methods* **289**, 114032 (2021).
4. Zohar, T. *et al.* Compromised Humoral Functional Evolution Tracks with SARS-CoV-2 Mortality. *Cell* **183**, 1508-1519.e12 (2020).
5. Sen, S. R. *et al.* Predicting COVID-19 Severity with a Specific Nucleocapsid Antibody plus Disease Risk Factor Score. *mSphere* **6**, e00203-21 (2021).
6. Stoddard, C. I. *et al.* Epitope profiling reveals binding signatures of SARS-CoV-2 immune response in natural infection and cross-reactivity with endemic human CoVs. *Cell Reports* **35**, 109164 (2021).
7. Oran, D. P. & Topol, E. J. Prevalence of Asymptomatic SARS-CoV-2 Infection : A Narrative Review. *Annals of internal medicine* **173**, 362–367 (2020).
8. Pavelić, K., Kraljević Pavelić, S., Brix, B. & Goswami, N. A Perspective on COVID-19 Management. *J. Clin. Med.* **10**, 1586 (2021).
9. Tian, W. *et al.* Predictors of mortality in hospitalized COVID-19 patients: A systematic review and meta-analysis. *Journal of Medical Virology* **92**, 1875–1883 (2020).
10. Simonnet, A. *et al.* High Prevalence of Obesity in Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) Requiring Invasive Mechanical Ventilation. *Obesity* **28**, 1195–1199 (2020).
11. Lee, L. Y. W. *et al.* COVID-19 prevalence and mortality in patients with cancer and the effect of primary tumour subtype and patient demographics: a prospective cohort study. *Lancet Oncol.* **21**, 1309–1316 (2020).
12. Zhang, L. *et al.* Clinical Course and Mortality of Stroke Patients With Coronavirus Disease 2019 in Wuhan, China. *Stroke* **51**, 2674–2682 (2020).
13. Vaughan, L. *et al.* Relationship of socio-demographics, comorbidities, symptoms and healthcare access with early COVID-19 presentation and disease severity. *BMC Infect.*

- Dis.* **21**, 1–10 (2021).
14. Chatterjee, A. *et al.* Can predicting COVID-19 mortality in a European cohort using only demographic and comorbidity data surpass age-based prediction: An externally validated study. *PLoS One* **16**, e0249920 (2021).
 15. Atyeo, C. *et al.* Distinct Early Serological Signatures Track with SARS-CoV-2 Survival. *Immunity* **53**, 524-532.e4 (2020).
 16. Ng, K. W. *et al.* Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. *Science (80-.)*. **370**, 1339–1343 (2020).
 17. Amrun, S. N. *et al.* Linear B-cell epitopes in the spike and nucleocapsid proteins as markers of SARS-CoV-2 exposure and disease severity. *EBioMedicine* **58**, (2020).
 18. Chakraborty, S. *et al.* Proinflammatory IgG Fc structures in patients with severe COVID-19. *Nat. Immunol.* **22**, 67–73 (2021).
 19. Corti, D., Purcell, L. A., Snell, G. & Veessler, D. Tackling COVID-19 with neutralizing monoclonal antibodies. *Cell* **184**, 3086–3108 (2021).
 20. Larsen, M. D. *et al.* Afucosylated IgG characterizes enveloped viral responses and correlates with COVID-19 severity. *Science (80-.)*. **371**, abc8778 (2021).
 21. Garcia-Beltran, W. F. *et al.* COVID-19-neutralizing antibodies predict disease severity and survival. *Cell* **184**, 476-488.e11 (2021).
 22. Tso, F. Y. *et al.* Presence of antibody-dependent cellular cytotoxicity (ADCC) against SARS-CoV-2 in COVID-19 plasma. *PLoS One* **16**, e0247640 (2021).
 23. Khandia, R. *et al.* Modulation of Dengue/Zika Virus pathogenicity by antibody-dependent enhancement and strategies to protect against enhancement in Zika Virus infection. *Frontiers in Immunology* **9**, 597 (2018).
 24. Takada, A. & Kawaoka, Y. Antibody-dependent enhancement of viral infection: Molecular mechanisms and in vivo implications. *Reviews in Medical Virology* **13**, 387–398 (2003).
 25. Gunawardana, S. A. & Shaw, R. H. Cross-reactive dengue virus-derived monoclonal antibodies to Zika virus envelope protein: Panacea or Pandora's box? *BMC Infectious Diseases* **18**, 614 (2018).
 26. Earnest, J. T. *et al.* The mechanistic basis of protection by non-neutralizing anti-alphavirus antibodies. *Cell Rep.* **35**, 108962 (2021).
 27. Winkler, E. S. *et al.* Human neutralizing antibodies against SARS-CoV-2 require intact Fc effector functions for optimal therapeutic protection. *Cell* **184**, 1804-1820.e16 (2021).
 28. Piccoli, L. *et al.* Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell* **183**, 1024-1042.e21 (2020).

29. Niu, L., Wittrock, K. N., Clabaugh, G. C., Srivastava, V. & Cho, M. W. A Structural Landscape of Neutralizing Antibodies Against SARS-CoV-2 Receptor Binding Domain. *Front. Immunol.* **12**, 1427 (2021).
30. Huo, J. *et al.* Neutralization of SARS-CoV-2 by Destruction of the Prefusion Spike. *Cell Host Microbe* **28**, 445-454.e6 (2020).
31. Pinto, D. *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290–295 (2020).
32. Shrock, E. *et al.* Viral epitope profiling of COVID-19 patients reveals cross-reactivity and correlates of severity. *Science (80-.)*. **370**, eabd4250 (2020).
33. Wang, P. *et al.* Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* **593**, 130–135 (2021).
34. Hahn, G. *et al.* Genome-wide association analysis of COVID-19 mortality risk in SARS-CoV-2 genomes identifies mutation in the SARS-CoV-2 spike protein that colocalizes with P.1 of the Brazilian strain. *Genet. Epidemiol.* gepi.22421 (2021). doi:10.1002/gepi.22421
35. Imai, M. *et al.* Characterization of a new SARS-CoV-2 variant that emerged in Brazil. *Proc. Natl. Acad. Sci.* **118**, e2106535118 (2021).
36. Plante, J. A. *et al.* Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **575**, 1–19 (2020).
37. Xie, X. *et al.* Neutralization of SARS-CoV-2 spike 69/70 deletion, E484K and N501Y variants by BNT162b2 vaccine-elicited sera. *Nat. Med.* **27**, 620–621 (2021).
38. Harvey, W. T. *et al.* SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **19**, 409–424 (2021).
39. Starr, T. N. *et al.* Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science (80-.)*. **371**, 850–854 (2021).
40. Huang, Y., Yang, C., Xu, X. feng, Xu, W. & Liu, S. wen. Structural and functional properties of SARS-CoV-2 spike protein: potential antiviral drug development for COVID-19. *Acta Pharmacologica Sinica* **41**, 1141–1149 (2020).
41. Johnson, B. A. *et al.* Furin Cleavage Site Is Key to SARS-CoV-2 Pathogenesis. *bioRxiv Prepr. Serv. Biol.* (2020). doi:10.1101/2020.08.26.268854
42. Peacock, T. P. *et al.* The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets. *Nat. Microbiol.* 1–11 (2021). doi:10.1038/s41564-021-00908-w
43. Zamecnik, C. R. *et al.* ReScan, a Multiplex Diagnostic Pipeline, Pans Human Sera for SARS-CoV-2 Antigens. *Cell Reports Med.* **1**, 100123 (2020).
44. Wang, S. F. *et al.* Antibody-dependent SARS coronavirus infection is mediated by antibodies against spike proteins. *Biochem. Biophys. Res. Commun.* **451**, 208–214

- (2014).
45. Li, Y. *et al.* Linear epitopes of SARS-CoV-2 spike protein elicit neutralizing antibodies in COVID-19 patients. *Cellular and Molecular Immunology* **17**, 1095–1097 (2020).
 46. Shah, P., Canziani, G. A., Carter, E. P. & Chaiken, I. The Case for S2: The Potential Benefits of the S2 Subunit of the SARS-CoV-2 Spike Protein as an Immunogen in Fighting the COVID-19 Pandemic. *Front. Immunol.* **12**, 637651 (2021).
 47. Brouwer, P. J. M. *et al.* Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science (80-.).* **650**, 1–8 (2020).
 48. Hawkes, R. . Enhancement of the infectivity of arbovirus by specific antisera produced in domestic fowls. *Aust. J. Exp. Biol. Med. Sci.* **42**, 465–482 (1964).
 49. Lee, W. S., Wheatley, A. K., Kent, S. J. & DeKosky, B. J. Antibody-dependent enhancement and SARS-CoV-2 vaccines and therapies. *Nat. Microbiol.* **5**, 1185–1191 (2020).
 50. Bournazos, S., Gupta, A. & Ravetch, J. V. The role of IgG Fc receptors in antibody-dependent enhancement. *Nature Reviews Immunology* **20**, 633–643 (2020).
 51. Holl, V. *et al.* Nonneutralizing Antibodies Are Able To Inhibit Human Immunodeficiency Virus Type 1 Replication in Macrophages and Immature Dendritic Cells. *J. Virol.* **80**, 6177–6181 (2006).
 52. Mayr, L. M., Su, B. & Moog, C. Non-neutralizing antibodies directed against HIV and their functions. *Frontiers in Immunology* **8**, 1590 (2017).
 53. Saphire, E. O., Schendel, S. L., Gunn, B. M., Milligan, J. C. & Alter, G. Antibody-mediated protection against Ebola virus. *Nature Immunology* **19**, 1169–1178 (2018).
 54. Henry Dunand, C. J. *et al.* Both Neutralizing and Non-Neutralizing Human H7N9 Influenza Vaccine-Induced Monoclonal Antibodies Confer Protection. *Cell Host Microbe* **19**, 800–813 (2016).
 55. Padilla-Quirarte, H. O., Lopez-Guerrero, D. V., Gutierrez-Xicotencatl, L. & Esquivel-Guadarrama, F. Protective antibodies against influenza proteins. *Frontiers in Immunology* **10**, 1677 (2019).
 56. Aran, D., Beachler, D. C., Lanes, S. & Overhage, J. M. Prior presumed coronavirus infection reduces COVID-19 risk: A cohort study ☆. *J. Infect.* **81**, 923–930 (2020).
 57. Fierz, W. & Walz, B. Antibody Dependent Enhancement Due to Original Antigenic Sin and the Development of SARS. *Front. Immunol.* **11**, 1120 (2020).
 58. Brown, E. L. & Essigmann, H. T. Original Antigenic Sin: the Downside of Immunological Memory and Implications for COVID-19. *mSphere* **6**, e00056-21 (2021).
 59. Dugan, H. L. *et al.* Profiling B cell immunodominance after SARS-CoV-2 infection reveals antibody evolution to non-neutralizing viral targets. *Immunity* **54**, 1290-1303.e7 (2021).

60. Kohler, H. & Nara, P. A Novel Hypothesis for Original Antigenic Sin in the Severe Disease of SARS-CoV-2 Infection. *Monoclonal Antibodies in Immunodiagnosis and Immunotherapy* **39**, 107–111 (2020).
61. Ng, K. W. *et al.* Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. *Science (80-.)*. **370**, 1339–1343 (2020).
62. Wang, C. *et al.* A conserved immunogenic and vulnerable site on the coronavirus spike protein delineated by cross-reactive monoclonal antibodies. *Nat. Commun.* **12**, 1–15 (2021).
63. Song, G. *et al.* Cross-reactive serum and memory B cell responses to spike protein in SARS-CoV-2 and endemic coronavirus infection. *bioRxiv Prepr. Serv. Biol.* **12**, 2938 (2020).
64. Anderson, E. M. *et al.* Seasonal human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with protection. *medRxiv* **7**, 2020.11.06.20227215 (2020).
65. Combes, A. J. *et al.* Global absence and targeting of protective immune states in severe COVID-19. *Nature* **591**, 124–130 (2021).
66. Long, Q. X. *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat. Med.* **26**, 845–848 (2020).
67. Trinité, B. *et al.* SARS-CoV-2 infection elicits a rapid neutralizing antibody response that correlates with disease severity. *Sci. Rep.* **11**, 2608 (2021).
68. Guthmiller, J. J. *et al.* SARS-CoV-2 infection severity is linked to superior humoral immunity against the spike. *MBio* **12**, 1–13 (2021).
69. Fajnzylber, J. *et al.* SARS-CoV-2 viral load is associated with increased disease severity and mortality. *Nat. Commun.* **11**, 1–9 (2020).
70. Pujadas, E. *et al.* SARS-CoV-2 viral load predicts COVID-19 mortality. *The Lancet Respiratory Medicine* **8**, e70 (2020).
71. Westblade, L. F. *et al.* SARS-CoV-2 Viral Load Predicts Mortality in Patients with and without Cancer Who Are Hospitalized with COVID-19. *Cancer Cell* **38**, 661-671.e2 (2020).
72. Legros, V. *et al.* A longitudinal study of SARS-CoV-2-infected patients reveals a high correlation between neutralizing antibodies and COVID-19 severity. *Cell. Mol. Immunol.* **18**, 318–327 (2021).
73. Wang, L. *et al.* Serological Responses to Human Virome Define Clinical Outcomes of Italian Patients Infected with SARS-CoV-2. *medRxiv Prepr. Serv. Heal. Sci.* (2020). doi:10.1101/2020.09.04.20187088
74. Röltgen, K. *et al.* Defining the features and duration of antibody responses to SARS-CoV-2 infection associated with disease severity and outcome. *Sci. Immunol.* **5**, abe0240 (2020).

75. Brouwer, P. J. M. *et al.* Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science*. **369**, 643–650 (2020).
76. Batra, M. *et al.* Role of IgG against N-protein of SARS-CoV2 in COVID19 clinical outcomes. *Sci. Rep.* **11**, 3455 (2021).
77. Vidarsson, G., Dekkers, G. & Rispens, T. IgG subclasses and allotypes: From structure to effector functions. *Front. Immunol.* **5**, 520 (2014).
78. Staats, L. A. N. *et al.* IgA2 Antibodies against SARS-CoV-2 Correlate with NET Formation and Fatal Outcome in Severely Diseased COVID-19 Patients. *Cells* **9**, 2676 (2020).
79. Jefferis, R., Lund, J. & Pound, J. D. IgG-Fc-mediated effector functions: Molecular definition of interaction sites for effector ligands and the role of glycosylation. *Immunological Reviews* **163**, 59–76 (1998).
80. Mkaddem, S. Ben, Benhamou, M. & Monteiro, R. C. Understanding Fc receptor involvement in inflammatory diseases: From mechanisms to new therapeutic tools. *Frontiers in Immunology* **10**, 811 (2019).
81. Dhodapkar, K. M. *et al.* Selective blockade of the inhibitory Fc γ receptor (Fc γ RIIB) in human dendritic cells and monocytes induces a type I interferon response program. *J. Exp. Med.* **204**, 1359–1369 (2007).
82. He, X., Wang, M. & Wu, M. Fc γ RIIb blockage: a promising immunotherapy target for severe COVID-19. *Signal Transduct. Target. Ther.* **6**, 147 (2021).
83. Young, B. E. *et al.* Viral Dynamics and Immune Correlates of Coronavirus Disease 2019 (COVID-19) Severity. *Clin. Infect. Dis.* **14** (2020).
84. Schulte-Schrepping, J. *et al.* Severe COVID-19 Is Marked by a Dysregulated Myeloid Cell Compartment. *Cell* **182**, 1419–1440.e23 (2020).
85. Channappanavar, R. *et al.* Dysregulated Type I Interferon and Inflammatory Monocyte-Macrophage Responses Cause Lethal Pneumonia in SARS-CoV-Infected Mice. *Cell Host Microbe* **19**, 181–193 (2016).
86. Blanco-Melo, D. *et al.* Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* **181**, 1036–1045.e9 (2020).
87. Lee, J. S. *et al.* Immunophenotyping of covid-19 and influenza highlights the role of type I interferons in development of severe covid-19. *Sci. Immunol.* **5**, 1554 (2020).
88. Foell, D. *et al.* Proinflammatory S100A12 can activate human monocytes via toll-like receptor 4. *Am. J. Respir. Crit. Care Med.* **187**, 1324–1334 (2013).
89. Zhou, F. *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* **395**, 1054–1062 (2020).
90. Wilk, A. J. *et al.* A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat. Med.* **26**, 1070–1076 (2020).

91. Maucourant, C. *et al.* Natural killer cell immunotypes related to COVID-19 disease severity. *Sci. Immunol.* **5**, aebd6832 (2020).
92. van Eeden, C., Khan, L., Osman, M. S. & Tervaert, J. W. C. Natural killer cell dysfunction and its role in covid-19. *International Journal of Molecular Sciences* **21**, 1–17 (2020).
93. Gallo, A. *et al.* Preliminary evidence of blunted humoral response to SARS-CoV-2 mRNA vaccine in multiple sclerosis patients treated with ocrelizumab. *Neurol. Sci.* 1–4 (2021).
94. Achiron, A. *et al.* Humoral immune response to COVID-19 mRNA vaccine in patients with multiple sclerosis treated with high-efficacy disease-modifying therapies. *Ther. Adv. Neurol. Disord.* **14**, 1–8 (2021).
95. Grupper, A. *et al.* Reduced humoral response to mRNA SARS-CoV-2 BNT162b2 vaccine in kidney transplant recipients without prior exposure to the virus. *Am. J. Transplant.* **00**, 1–8 (2021).
96. Bruni, M. *et al.* Persistence of Anti-SARS-CoV-2 Antibodies in Non-Hospitalized COVID-19 Convalescent Health Care Workers. *J. Clin. Med.* **9**, 3188 (2020).
97. Ibarondo, F. J. *et al.* Rapid Decay of Anti-SARS-CoV-2 Antibodies in Persons with Mild Covid-19. *N. Engl. J. Med.* **383**, 1085–1087 (2020).
98. Hoepel, W. *et al.* High titers and low fucosylation of early human anti-SARS-CoV-2 IgG promote inflammation by alveolar macrophages. *Sci. Transl. Med.* **13**, eabf8654 (2021).
99. Dekkers, G., Rispen, T. & Vidarsson, G. Novel concepts of altered immunoglobulin G galactosylation in autoimmune diseases. *Front. Immunol.* **9**, 553 (2018).
100. Thulin, N. K. *et al.* Maternal Anti-Dengue IgG Fucosylation Predicts Susceptibility to Dengue Disease in Infants. *Cell Rep.* **31**, 107642 (2020).
101. Wang, T. T. *et al.* IgG antibodies to dengue enhanced for FcγRIIIA binding determine disease severity. *Science* **355**, 395–398 (2017).
102. Khoury, D. S. *et al.* Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat. Med.* 1–7 (2021). doi:10.1038/s41591-021-01377-8
103. Kawasuji, H. *et al.* Transmissibility of COVID-19 depends on the viral load around onset in adult and symptomatic patients. *PLoS One* **15**, e0243597 (2020).
104. Galani, I. E. *et al.* Untuned antiviral immunity in COVID-19 revealed by temporal type I/III interferon patterns and flu comparison. *Nat. Immunol.* **22**, 32–40 (2021).
105. Thestrup, T. *et al.* optimized ratiometric calcium sensors for functional in vivo imaging of neurons and t lymphocytes. *Artic. Nat. methods.* **11**, 175 (2014).

CHAPTER 2

Predicting COVID-19 Severity with a Specific Nucleocapsid Antibody plus Disease Risk Factor Score

Adapted with permission from *mSphere*.

DOI:10.1128/mSphere.00203-21

Author contributions: Sanjana R. Sen, Emily C. Sanders, Kristin N. Gabriel, Sudipta Majumdar, Phil L. Felgner, and Gregory A. Weiss designed research; Sanjana R. Sen, Emily C. Sanders, Kristin N. Gabriel, Brian M. Miller, Hariny M. Isoda, Gabriela S. Salcedo, Jason E. Garrido, Rebekah P. Dyer, Alicia M. Santos, Keertna Bhuvan, Rie Nakajima, and Aarti Jain performed research; Sanjana R. Sen, Emily C. Sanders, Kristin N. Gabriel, and Gregory A. Weiss analyzed data; Joni L. Ricks-Oddie advised on statistical analysis; Delia F. Tifrea and Robert A. Edwards collected patient samples and advised on patient clinical data analysis; and Sanjana R. Sen, Emily C. Sanders, Kristin N. Gabriel, Sudipta Majumdar, and Gregory A. Weiss wrote the manuscript.

2.1 Abstract

Effective methods for predicting COVID-19 disease trajectories are urgently needed. Here, ELISA and coronavirus antigen microarray (COVAM) analysis mapped antibody epitopes in the plasma of COVID-19 patients (n = 86) experiencing a wide-range of disease states. The experiments identified antibodies to a 21-residue epitope from nucleocapsid (termed Ep9) associated with severe disease, including admission to the ICU, requirement for ventilators, or death. Importantly, anti-Ep9 antibodies can be detected within six days post-symptom onset and sometimes within one day. Furthermore, anti-Ep9 antibodies correlate with various comorbidities and hallmarks of immune hyperactivity. We introduce a simple-to-calculate, disease risk factor score to quantitate each patient's comorbidities and age. For patients with anti-Ep9 antibodies, scores above 3.0 predict more severe disease outcomes with a 13.42 Likelihood Ratio (96.7% specificity). The results lay the groundwork for a new type of COVID-19 prognostic to allow early identification and triage of high-risk patients. Such information could guide more effective therapeutic intervention.

2.2 Introduction

The COVID-19 pandemic has triggered an ongoing global health crisis. More than 179 million confirmed cases and 3.9 million deaths have been reported worldwide as of June 25, 2021 ¹. The virus that causes COVID-19, severe acute respiratory syndrome coronavirus (SARS-CoV-2), belongs to the same family of viruses responsible for respiratory illness linked to recent epidemics – severe acute respiratory syndrome (SARS-CoV-1 termed SARS here) in 2002-2003 and Middle East respiratory syndrome (MERS) in 2012 ². The current and previous outbreaks suggest coronaviruses will remain viruses of concern for global health.

Many risk factors and comorbidities, including age, sex, hypertension, diabetes, and obesity, can influence COVID-19 patient outcomes ³. Analysis of patient immune parameters has linked disease severity to elevated levels of biomarkers for inflammation (c-reactive protein and cardiac troponin I), organ damage (aspartate aminotransferase, abbreviated AST, and hypoalbuminemia), immune hyperactivity (IL-6 and IL-10), and clotting (D-dimer) ⁴. Mortality in COVID-19 is often caused by multi-organ injury and severe pneumonia attributed to an excessive immune response, termed a cytokine storm ⁵. Given the rapid and wide spectrum of COVID-19 disease progression, a more precise prognostic linking disease risk factors and specific immune responses can potentially predict disease trajectories and guide interventions.

One hypothesis to explain differences in severity of COVID-19 implicates weakly binding, non-neutralizing antibodies (Abs) to SARS-CoV-2 proteins ⁶. However, the potential harm of these suboptimal Abs in COVID-19 patient outcomes remains ill-defined.

Furthermore, a recent review on antibody-dependent enhancement of SARS-CoV-2 stated, “At present, there are no known clinical findings, immunological assays or biomarkers that can differentiate any severe infection from immune-enhanced disease, whether by measuring antibodies, T cells or intrinsic host responses ⁷.” This conclusion inspired our study.

SARS-CoV-2 encodes four major structural proteins – spike (S), nucleocapsid (N), membrane (M), and envelope (E). The S, N, and M proteins from SARS elicit an Ab-based immune response ^{8,9}. The Ab response and its effects on disease progression in SARS-CoV-2 remain under investigation ^{10,11}. Bioinformatics has predicted >55 Ab binding epitope regions from SARS-CoV-2 ¹²⁻¹⁷. The epitopes for N, M or E proteins are less well-characterized than for S protein. Several studies have reported comprehensive epitope mapping of the antibody response to SARS-CoV-2 ¹⁸⁻²¹. Here, we sought to characterize epitopes from SARS-CoV-2 and their correlations with disease severity. ELISAs with phage-displayed epitopes (phage ELISAs) and coronavirus antigen microarray (COVAM) analysis ²² examined plasma samples from COVID-19 patients (n = 86). The results demonstrate that Abs to a specific epitope from N protein plus disease risk factors strongly correlate with COVID-19 disease severity.

2.3 Results

Design and production of candidate epitopes

Twenty-one putative SARS-CoV-2 epitopes were predicted through bioinformatics ¹²⁻¹⁴ and structure-based analysis. The candidate epitopes spanned the S, N, M, or E proteins and were on average 34 amino acids in length (Figure 2-1 and Table 2-1). These epitopes

were phage-displayed as fragments of the full-length protein and were likely unstructured. Here, epitope refers to the predicted region of the antigenic protein recognized by the antibody's paratope. The structure of S protein bound to a neutralizing antibody^{23,24} provided the starting point for 12 of these antibody epitopes. Epitopes were designed to potentially isolate even suboptimal Abs binding to small portions of these structural proteins; such suboptimal Abs were hypothesized to provide insight into disease severity. After display of each potential epitope on the surface of phage, the quality of the epitopes was evaluated by PCR, DNA sequencing, and QC ELISA (Figure 2-6). A total of 18 phage-displayed, putative epitopes passed quality control PCR, and were selected for further study.

Mapping epitope binding to anti-SARS-CoV-2 Abs

Plasma from COVID-19 patients was subjected to ELISAs with the phage-displayed SARS-CoV-2 epitopes (Figure 2-2A). Unless otherwise indicated (e.g., healthy controls), plasma refers to samples from PCR-verified, COVID-19 patients. In this initial assay, plasma was pooled, diluted 100-fold, and coated on a microtiter plate as the antigen of interest (3 pools of n = 5 patients per pool). Nonspecific interactions were blocked (ChonBlock), and phage-displayed epitopes were added for ELISA. The resultant data were normalized by signal from the corresponding negative control (phage without a displayed epitope). Seven candidate epitopes from the pooled patients were further investigated with a larger number of individual patient samples (n = 28) (Figure 2-2B). The strongest, reproducible binding was observed for three epitopes from M (Ep6), N (Ep9), and S (Ep20) proteins. Additional COVID-19 plasma samples were profiled for binding to these three epitopes (n = 86 total) (Figure 2-2B).

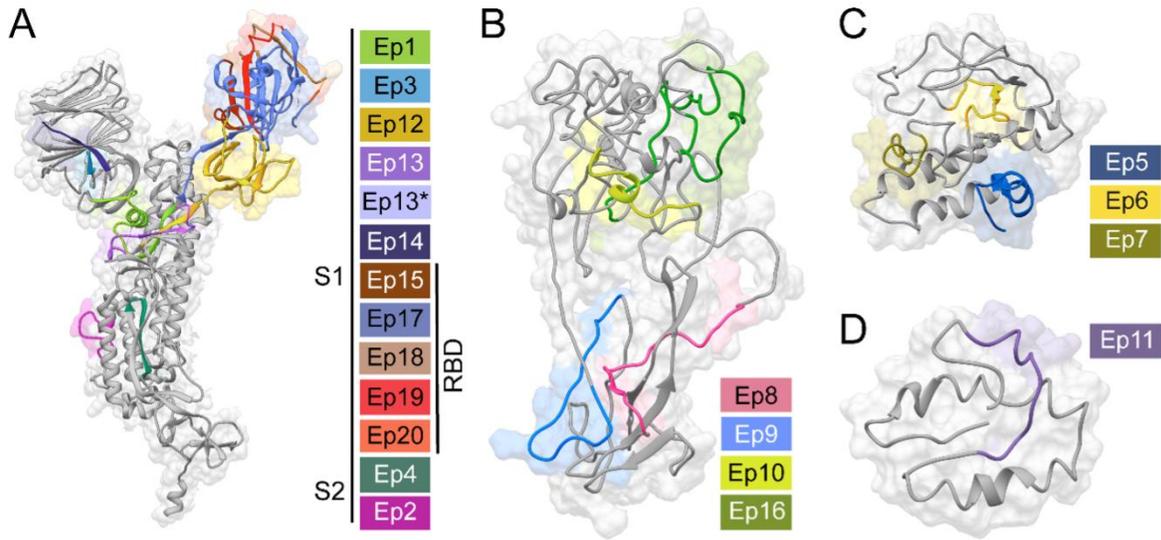


Figure 2-1. Predicted SARS-CoV-2 epitopes examined by phage ELISA. Structural models (gray) of the SARS-CoV-2 **A.** S, **B.** N, **C.** M, or **D.** E proteins illustrate our epitope design (colored). Sequence Ep13* has the mutation D614G, which increases the fitness of SARS-CoV-2²⁵⁻²⁷. The depicted structural models were derived from an S protein X-ray structure (PDB: 6VXX)²³ or computation modeling of N, M, and E proteins (Protein Gene Bank: QHD43423, QHD43419, and QHD43418, respectively)²⁸. **Table 2-1** provides sequences, sources, and rationale for selections.

Only the Ep9 epitope from N protein demonstrated robust, statistically significant antibody binding in 27% of patients (n = 186) (Figure 2-2B). Of these patients, 100 did not have corresponding health information and were not analyzed further in this report. To test non-phage displayed epitopes, dose-dependent binding of antibodies to Ep9 fused to eGFP (eGFP-Ep9) or to full-length N protein demonstrated that α Ep9 IgGs bound its antigen with $EC_{50} = 3.22$ nM (95% CI = 2.49 to 4.14 nM). This experiment examined plasma samples with the highest IgG response against the N protein in the COVAM assay. Patients without α Ep9 Abs had roughly the same level of binding to N protein as observed for α Ep9 Abs binding to

Ep9. However, such α Ep9 Abs appeared to add to N protein binding by antibodies; approximately two-fold increase in apparent antibody binding levels for N protein was observed, if the patient also has α Ep9 Abs (Figure 2-2C). In patients for whom longitudinal samples were available, the highest levels of α Ep9 Abs were observed at days 1 to 14 post-symptom onset (n = 11) and were detectable within 6 days (Figure 2-2D). In four of these patients, α Ep9 Abs persisted after day 14.

Cross-reactivity of α Ep9 Abs against orthologous epitopes from other coronaviruses

Next, the cross-reactivity of α Ep9 Abs was examined with Ep9-orthologs from four phylogenetically related coronaviruses known to infect humans (Figure 2-7A). Specifically, plasma with α Ep9 Abs (n = 3) and pooled plasma from healthy individuals (n = 5) were assayed. The Ep9 epitopes from SARS-CoV-2 and SARS have 90% amino acid sequence homology. Unsurprisingly, this high degree of similarity resulted in a cross-reactive Ep9 epitope, and a strong antibody response was observed to Ep9 epitopes from both viruses (Figure 2-2E). The coronaviruses, MERS, HKU-1, and NL63 have 52%, 43%, and 8% sequence homology to SARS-CoV-2 Ep9, respectively (Figure 2-7B). These more distantly related orthologs exhibited no cross-reactivity with the α Ep9 Abs. Furthermore, no response was observed to Ep9 in pooled plasma from healthy individuals.

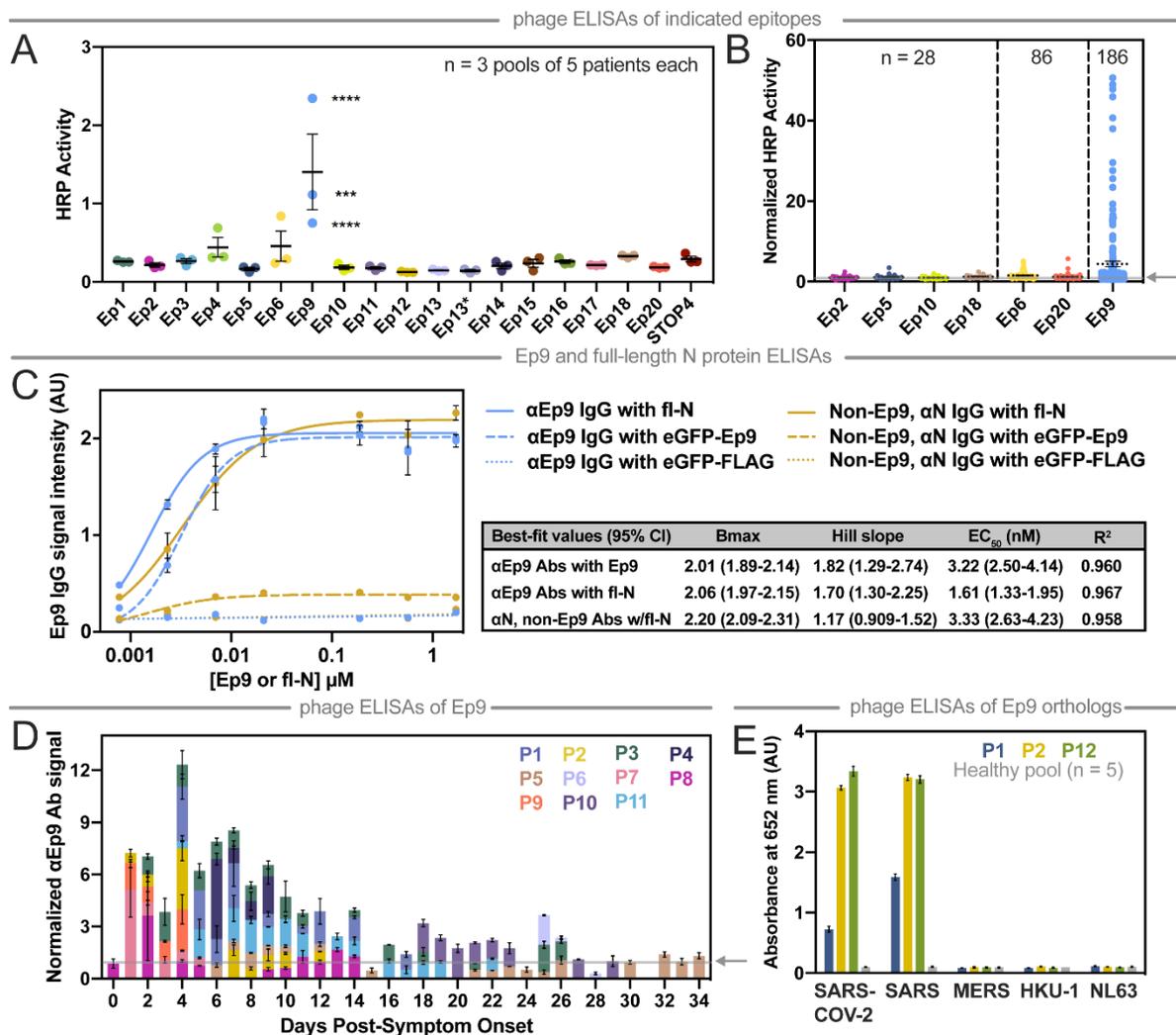


Figure 2-2 Mapping COVID-19 patient antibody responses with phage-displayed SARS-CoV-2 epitopes. **A.** This phage ELISA with the indicated epitopes (x-axis) examined plasma pooled from patients (n = 3 pools of 5 patients each, 2 technical replicates). STOP4 is the phage negative control. **B.** The epitopes with the highest signals were then further examined by ELISA with plasma from individual patients (n as indicated). **C.** This ELISA measures dose-dependent binding of α N IgGs from plasma pooled from five α Ep9 positive patients and five non-Ep9, α N positive patients to eGFP-Ep9 (dashed line), eGFP negative control (eGFP-FLAG, dotted line) or full-length N protein (fl-N, solid line). Varying doses of Ep9 or fl-N were immobilized on microtiter plates, and binding of pooled patient plasma (1:100) was detected using α -Fc IgG-HRP Abs (1:10,000). Pooled patients were matched by similar α N IgG binding signal in COVAM analysis (inset). Non-linear lines of best-fit for binding saturation are represented. Statistical comparisons of Bmax, Hill slope and EC₅₀ between groups, determines that binding of α Ep9 IgGs to fl-N or eGFP-Ep9, and non-Ep9, α N IgGs to fl-N are significantly different (p<0.0001). Error bars represent \pm SD. The data demonstrates that the

EC50 value of α Ep9 Abs is equal to the cumulative EC50 of all other α N Abs in patients lacking the α Ep9 Abs. In the presence of the α Ep9 Abs, the apparent binding levels of α N Abs against fl-N approximately doubles. **D.** With samples from individual patients (designated as P# and by color) collected at the indicated times, α Ep9 Abs were measured. The subset of patients shown here comprise all samples for which longitudinal data was available. **E.** Phage ELISA with samples from patients with strong α Ep9 Ab response (two from the longitudinal study and one from the patient population) examines cross-reactive binding to Ep9 or Ep9 orthologs from the indicated coronaviruses (x-axis, 3 technical replicates). The arrow on the y-axis and gray line (panels B and C) represents the negative control used for normalizing the data. Error bars represent SEM (panels A, B, and D) or range of two measurements (panel C).

The protein microarray COVAM analysis is a high-throughput serological test for SARS-CoV-2 Ab cross-reactivity with a panel of 67 antigens from 23 strains of 10 respiratory tract infection-causing viruses²². In this assay, each antigen was printed onto microarrays, probed with human plasma, and analyzed with an ArrayCam imager. COVAM distinguishes between IgG and IgM Abs binding to the full-length N protein (Figures 2-S3 and 2-S4, respectively). Thus, the COVAM analysis complemented the phage ELISA by expanding the scope of antigens surveyed and adding Ab serotype information. The ELISA and COVAM data both demonstrated that α Ep9 Abs were highly specific for lineage B betacoronaviruses, and were unlikely to be found in patients before their infection with SARS-CoV-2.

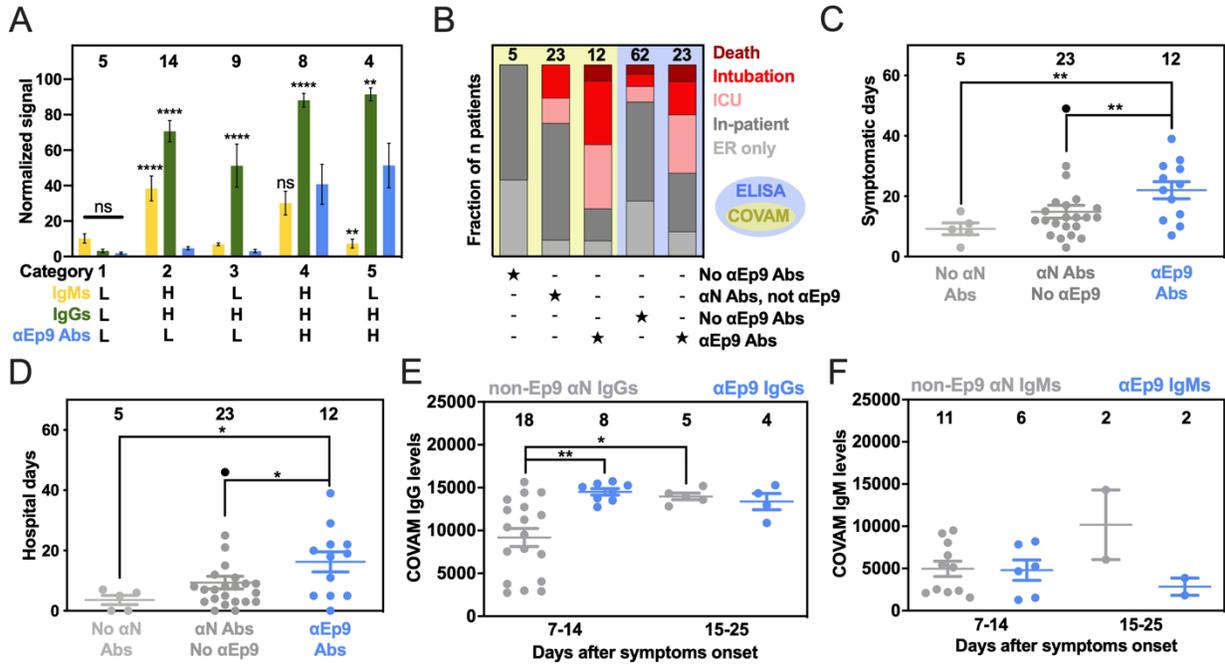


Figure 2-3. Patients with α Ep9 Abs have more severe disease. **A.** Normalized and categorized data from measurements by COVAM (IgMs in yellow, IgGs in green) and Ep9 phage ELISA (blue). ANOVA comparing COVAM to ELISA with Dunnett's multiple comparisons yields p-values of **<0.01, ****<0.0001, or ns: not significant. **B.** Disease severity (color) binned by antibody response (★) for COVAM (yellow) or ELISA (blue). Statistical analysis reveals significant differences between distributions of severe and non-severe disease comparing patient categories, $p < 0.01$ (Fisher's exact test) and $p < 0.001$ (Fisher's exact test) for COVAM and ELISA, respectively. Patients with α Ep9 Abs are **C.** symptomatic for longer durations and **D.** spend more days in the hospital than those with other α N Abs or no α N Abs. ANOVA with Tukey's multiple comparisons yields p-values of *<0.05 and **<0.01. One outlier (black) (ROUT = 0.1%) was omitted from statistical calculations for panels C and D. **E.** The α N IgG appear at high levels early in the course of disease only for α Ep9-positive patients, but are lower in non-Ep9, α N-positive patients. After >15 days post symptom onset, α N IgG levels increase for both groups of patients. **F.** However, IgM levels do not change significantly. Error bars depict SEM with the indicated number of patients (n, numbers above columns).

More severe disease and poorer outcomes for α Ep9 patients

Direct comparison of data with full-length N protein from COVAM and Ep9 phage ELISA (n = 40 patients assayed with both techniques) revealed five unique categories of

patients (Figure 2-3A). To enable this comparison, raw data from each assay was normalized as a percentage of the negative control. Category 1 consists of patients without Abs to the N protein. The next categories included patients with IgMs (Category 2) or IgGs (Category 3) binding to N protein, but not Ep9, termed non-Ep9 α N Abs. Category 4 included patients with α Ep9 Abs (both IgMs and IgGs). Category 5 patients had exclusively IgG α Ep9 Abs. The α Ep9 Abs were only found in patients with IgMs or IgGs against full-length N protein from the COVAM assay; the COVAM analysis thus independently corroborated the phage ELISAs (Figure 2-3A).

Interestingly, the patients with α Ep9 Abs suffered more prolonged illness and worse clinical outcomes compared to patients with non-Ep9 α N Abs or no α N Abs. In this study, severe COVID-19 cases were defined as resulting in death or requiring admission to the ICU or intubation. The fraction of severe COVID-19 cases was 2.5 times higher in α Ep9 Abs patients than non-Ep9 α N Abs patients (Figure 2-3B, yellow panel); the differences in proportions of severe and non-severe α N-positive patients with or without α Ep9 Abs were statistically significant ($p < 0.030$, Fisher's exact test). Patients without α N Abs (Category 1) had less severe symptoms. The α Ep9 Abs patients also had longer durations of symptoms and hospital stays relative to non-Ep9 α N Abs and no α N Abs patients (Figure 2-3C and 2-3D). A larger data set of patient plasma analyzed by phage ELISA confirmed this conclusion ($p < 0.0013$, Fisher's exact test) (Figure 2-3B, blue panel). Our data further demonstrated that asymptomatic COVID-19 patients ($n = 3$) also tested negative for α Ep9 Abs (Table 2-2). The data also revealed early seroconversion of α Ep9 IgGs (Figure 2-3E), but not α Ep9 IgMs (Figure 2-3F).

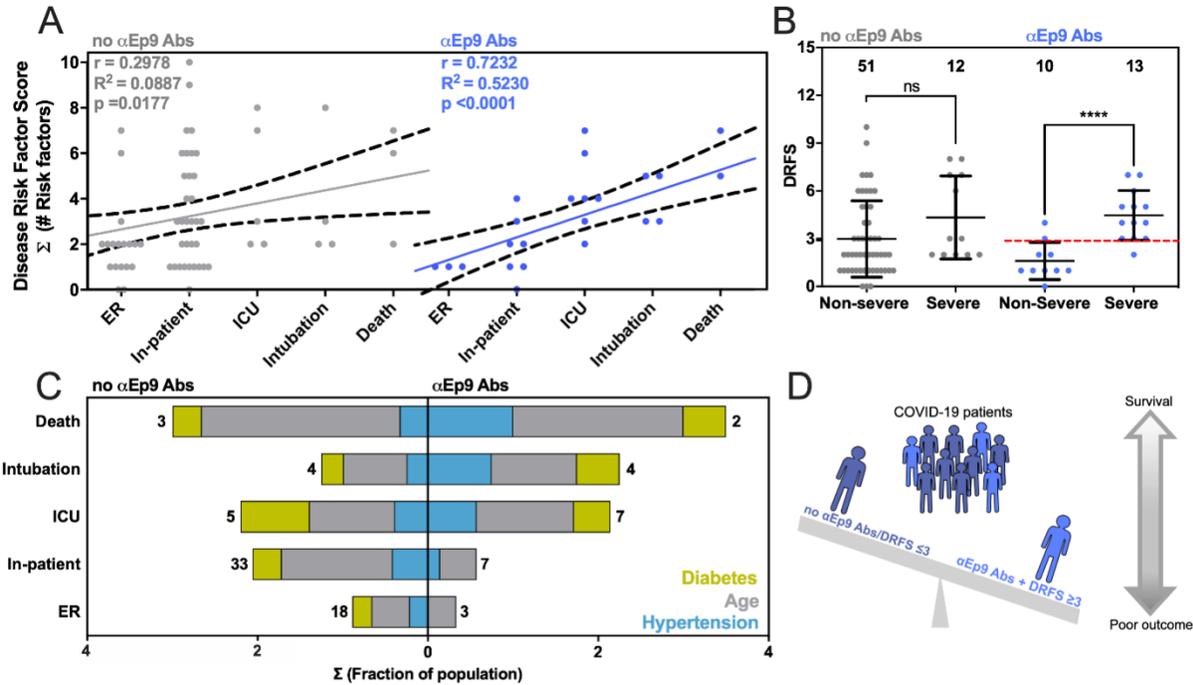


Figure 2-4. Correlation between disease severity and risk factors in patients with α Ep9 Abs. **A.** The relationship between DRFS and disease severity of COVID-19 patients with α Ep9 Abs (blue) or no α Ep9 Abs (gray). Each data point represents one patient. The solid lines indicate linear regression fits with 95% confidence intervals (dotted lines), and Pearson's r-value as noted. **B.** Correlation of disease severity with DRFS in patients with α Ep9 Abs. The data depicts a significant correlation between DRFS and disease severity in patients with α Ep9 Abs (blue), but not in patients lacking α Ep Abs (gray). In α Ep9 patients, a DRFS threshold of 3.0 can predict severe disease (red). Two-tailed, parametric t-tests were conducted to compare non-severe and severe disease outcomes of patients with and without α Ep9 Abs, where **** $p < 0.0001$. The error bars represent SD with the indicated n. **C.** The color-indicated risk factors (diabetes, hypertension, and age score) are depicted on the x-axis as the fractions of patients in each disease severity category (y-axis). Numbers indicate total patients (n) without α Ep9 Abs (left) or with α Ep9 Abs (right). The prevalence of risk factors (colors) increases with disease severity in patients with α Ep9 Abs, but not in patients without these Abs. **D.** Patients with α Ep9 Abs and DRFS ≥ 3 are predisposed to increased COVID-19 severity and poorer outcomes.

Strong correlation between disease severity and comorbidities in patients with α Ep9 Abs

We compared risk factors, clinical parameters, and disease outcomes among patients with α Ep9 Abs (n = 23) (Figures 2-4A and 2-S5). A *disease risk factor score* (DRFS) was

developed to evaluate the relationship between clinical preconditions and disease severity in patients with α Ep9 Abs. The DRFS quantified a patient's age, sex, and pre-existing health conditions associated with COVID-19 disease severity and mortality. Risk factors include hypertension, diabetes, obesity, cancer, and chronic conditions of the following: cardiac, cerebrovascular, kidney, and pulmonary ²⁹⁻³². Using the *age score* from the Charlson Comorbidity Index ³³ yields a patient's DRFS as:

$$DRFS = \Sigma (\# \text{ of risk factors}) + (\text{age score})$$

where each risk factor was valued as either 0 or 1 if absent or present, respectively. The DRFS of patients with α Ep9 Abs strongly correlated with COVID-19 disease severity (Pearson's $r = 0.72$, p -value < 0.0001 , and $R^2 = 0.52$) (Figure 2-4A). The correlation in patients without α Ep9 Abs was weak ($r = 0.30$, p -value = 0.089 , $R^2 = 0.018$) (Figure 2-4A). Amongst patients with α Ep9 Abs ($n = 23$), a $DRFS \geq 3$ determined disease severity with 92.3% sensitivity (1/13 false negatives) and 80% specificity (2/10 false positives) (Figure 2-4B). In the entire study cohort ($n = 86$), patients with α Ep9 Abs and $DRFS \geq 3$ ($n = 11$) have severe disease with a high degree of specificity (96.7%) and a sensitivity of 44%. Notably, DRFS predicted disease severity only for patients with α Ep9 Abs ($n = 23$), and patients without such Abs ($n = 63$) had no correlation with disease outcomes.

Examining key contributors to high DRFS, the presence of α Ep9 Abs correlated with more severe disease in patients who have hypertension, diabetes, or age > 50 years. Such correlation was not observed for patients lacking α Ep9 Abs (Figure 2-4C). Such risk factors were prevalent at roughly the same percentages in both populations of patients (Table 2-2). Thus, these risk factors were particularly acute for patients with α Ep9 Abs.

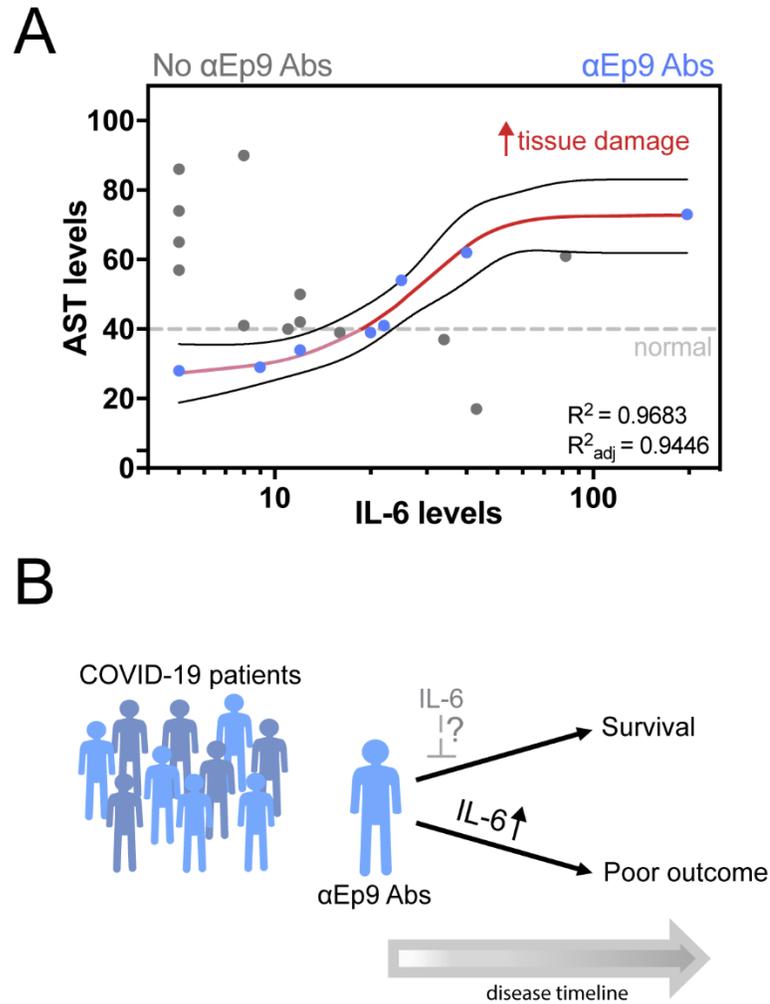


Figure 2-5. Association of inflammatory cytokine and tissue damage markers in patients with α Ep9 Abs. **A.** Association between the inflammatory cytokine, IL-6, and the tissue damage marker, aspartate transaminase (AST), shows a sigmoidal curve fit for patients with α Ep9 Abs, $R^2 = 0.9683$, Spearman's correlation coefficient = 1.0, $p < 0.0001$. **B.** Schematic of patients with α Ep9 Abs with increasing IL-6 levels leading to poor outcomes. We hypothesize patients with α Ep9 Abs could benefit from IL-6 inhibition early in the disease, such as monoclonal antibody drugs targeting IL-6 or its receptor (IL6R), to disrupt a cytokine storm and reduce severe outcomes.

High levels of inflammatory cytokine and tissue damage markers in patients with α Ep9 Abs

COVID-19 patients can have elevated serum concentrations of >20 inflammatory cytokines and chemokines ³⁴. However, information on the cytokine levels and the association with tissue damage and worse COVID-19 outcomes have been inconsistent ³⁴⁻³⁶. For patients with IL-6 concentrations measured in plasma, patients with (n = 8) or without (n = 11) α Ep9 Abs were compared. Interestingly, the comparison uncovered a strong positive sigmoidal association between IL-6 and AST unique to patients with α Ep9 Abs ($R^2 = 0.968$, Spearman's $r = 1.0$, p-value <0.0001, n = 8) (Red line, Figure 2-5A); correlation of IL-6 and AST in patients with α Ep9 Abs remained strong even after removal of the data point at the highest IL-6 concentration. Conversely, a slight negative trend was observed in patients lacking α Ep9 Abs (Spearman's $r = -0.575$, p-value= 0.0612, n = 13). Thus, the presence of α Ep9 Abs could disambiguate the sometimes contradictory association of IL-6 with disease severity.

2.4 Discussion

This study introduces a two-step test as a prognostic for predicting COVID-19 disease severity and its worst outcomes. Specifically, α Ep9 Abs can effectively predict severe disease (specificity 83.6%). However, combining presence of α Ep9 Abs with DRFS ≥ 3 provides much higher specificity (96.7%) for predicting severe disease. Previously, α N IgGs have been recognized as a focal site for an antibody response ^{18,19,21,37} and associated with disease severity and poor outcomes ^{11,37,38}.

The present investigation expands on previous reports that recognize various regions of the RNA binding domain of N protein as focal sites for anti-SARS-CoV-2 antibody response.

For example, the phage display-based VirScan identified an epitope region spanning residues 141-196 and microarrays further isolated peptides including residues 134-171, 155-171, 153-190, and 153-171^{18,19,21}. The above investigations, however, do not find correlations between any these epitopes and disease severity. Our results are confirmed by observations from a patient cohort in Singapore, which identify an epitope (residues 153-170) very similar to Ep9 (residues 152-172) and shows a correlation between antibody response against the epitope and pneumonia and the tissue damage markers (CRP and LDH)²⁰. In our investigation, we examine in-depth patient clinical histories, test results, disease outcomes ranging from asymptomatic to fatal, and longer longitudinal profiling post-symptom onset, to determine the association of a larger subset of markers and risk factors. Such data allows calculation of the DRFS. Together with the presence of α Ep9 Abs, patient DRFS allows early discrimination of severe from non-severe disease outcomes. Additionally, fine epitope mapping demonstrates that α Ep9 Abs strongly and uniquely correlate with COVID-19 disease severity relative to other α N Abs.

We hypothesize that the underlying mechanism relating α Ep9 Abs to increased disease severity involves an overzealous immune response. Specifically, we observe early seroconversion and strong early upregulation of α Ep9 IgGs (Figure 2-3E). Similar IgG observations have been correlated with poor viral neutralization and clearance, resulting in increased COVID-19 severity^{10,38,39}. Also, high levels of IL-6 are observed for α Ep9-positive patients with increased levels of the tissue damage marker AST; this correlation does not exist for patients lacking α Ep9 Abs (Figure 2-5A). The sensitivity to IL-6 concentration before AST-monitored organ damage suggests anti-IL-6 therapeutics could be an effective management in the early and rapidly progressive stages of respiratory distress for α Ep9-

positive patients^{34,40-44}. Since binding to N protein by α Ep9 antibodies is unlikely to enhance uptake of SARS-CoV-2, an antibody-dependent enhancement mechanism could invoke antigen uptake by macrophages. This mechanism could stimulate complement activation and the cytokine storm observed here as elevated IL-6 response. Further investigation is required to determine the basis for increased disease severity in α Ep9 patients.

The data demonstrate that α Ep9 positive patients with DRFS ≥ 3 are 13.42 times (Likelihood Ratio) more likely to have severe COVID-19 disease symptoms within the study cohort (n = 86). The presence of α Ep9 without DRFS is less effective as a prognostic (Likelihood Ratio of 3.17). Despite its high specificity (96.7%), the sensitivity of this two-step test is 44% (n = 86). However, this test could predict a subset of patients with a specific immune response (i.e., early IgG response and IL-6 dependent immune hyperactivity), and could suggest targeted treatment options (e.g., targeting IL-6 and its pathways).

Importantly, α Ep9 Abs appear early in the course of disease. Thus, such a prognostic could outperform traditional markers for the cytokine storm such as IL-6, which appears 6-8 days after symptom onset^{34,42}; all plasma collected from α Ep9 positive patients (n = 7, Figure 2-2D) between 1 to 6 days post-symptoms onset demonstrate detectable levels of α Ep9 IgG (≥ 2 fold over negative control). Early detection of α Ep9 Abs in patients could be used to triage and treat COVID-19 prior to the onset of its most severe symptoms; delayed treatments of IL-6 targeting drugs can decrease their efficacy or be counterproductive^{34,40-45} (Figure 2-5B). The α Ep9 Ab biomarker could identify patients most likely to benefit from anti-IL-6 therapeutics and avoid ineffective treatments.

This study demonstrates the usefulness of fine epitope mapping, but the following limitations should be noted. Short linear epitopes, unlike conformational epitopes in larger domains, might not resemble the tertiary structure of an antigen. Post-translational modifications, such as glycosylation were omitted for the phage-displayed S protein epitopes; the COVAM antigens, however, are produced in baculovirus or HEK-293 cells, which could glycosylate the antigens. Our analysis is largely based upon a population of 86 COVID-19 patients and 5 healthy individuals, with the majority of Hispanic descent. The conclusions could be further strengthened with follow-up investigations in a larger population. Additionally, the population examined here only included three asymptomatic individuals, and additional testing is required to verify absence of α Ep9 Abs in such patients. The sample size of patients with multiple antibody targets was too limited to allow correlation analysis; future investigations could examine associations between α Ep9 and other Abs. Abs recognizing other SARS-CoV-2 structural proteins could also exhibit similar characteristics to α Ep9 Abs.

Existing diagnostic platforms could readily be adapted to test for α Ep9 Abs, and the DRFS calculation is quite simple to implement (e.g., assay with eGFP-Ep9 fusion demonstrated here). As shown here, α Ep9 Abs do not recognize orthologous sequences from closely related coronaviruses, providing good specificity for α Ep9 as a prognostic. Previous studies have shown that the high homology of N protein among related coronaviruses can lead to high false positive rates in serodiagnostics with full-length N antigen ⁴⁶. Thus, the two-step prognostic reported here could mitigate the worst outcomes of COVID-19, particularly for patients at high risk.

2.5 Materials and Methods:

Cloning. For phage display of epitopes, the pm1165a phagemid vector as previously described ⁴⁷ as engineered to encode an N-terminal FLAG-tag and a C-terminal fusion to the P8 coat protein of M13-phage. This template, termed FlagTemplate, was used for subcloning of SARS-CoV-2, SARS, MERS, HKU-1 and NL63 epitopes. A vector map of the FlagTemplate (Figure 2-S6A), cloning procedures, and a list of oligos (Table 2-S3) for Q5 site-directed mutagenesis and Gibson Assembly are provided here.

Short (approximately 30 amino acids) putative epitopes for phage display and *E. coli* expression as eGFP fusion peptides in the pet28 vector were cloned via Q5 site-directed mutagenesis according to the manufacturer's instructions. A vector map of Ep9 fused peptide to eGFP, termed eGFP-Ep9, is shown below (Figure 2-S6B). For large epitopes (>500 bp), such as Ep17, Gibson Assembly (New England Biolabs) was conducted in two PCR steps with the FlagTemplate or pCAGGS containing the SARS-CoV-2 S protein gene (BEI Resources) to generate the vectors and inserts, respectively. The Gibson Assembly (2 μ L) or KLD (Kinase, Ligase, DpnI) mix (5 μ L) was transformed into Nova Blue *E. coli* competent cells, and transformants were plated on a carbenicillin-supplemented (50 μ g/mL) agar plate before incubation at 37 °C overnight. Five single colonies were selected to inoculate 4 mL of SOC media in a 15 mL culture tube supplemented with carbenicillin (50 μ g/mL). The seed cultures were incubated at 37 °C with shaking at 225 rpm for 8-12 h. Phagemid DNA was isolated using the QIAprep spin miniprep kit according to the manufacturer's instructions. The successful subcloning of the ORF encoding each epitope was verified via DNA sequencing (Genewiz). The full-length N protein in a pLVX-EF1 α -IRES-Puro plasmid was a generous gift from Prof. Rachel Martin (UCI).

Purification and preparation of phage. Phage were propagated and purified using procedures previously described ⁴⁸ with the following changes. A single colony was selected to inoculate 15 mL of 2YT and shaken at 37 °C until the OD₆₀₀ reached 0.6. After incubation at 37 °C for 45 min, 8 mL of the primary culture was used to inoculate 300 mL of 2YT supplemented with carbenicillin (50 µg/mL), kanamycin (20 µg/mL), and isopropyl β-D thiogalactopyranoside (IPTG, 30 µM).

To precipitate the phage, the cultures were centrifuged at 10 krpm (15300 x g) for 10 min at 4 °C. The supernatant was decanted into a centrifuge tube containing 60 mL PEG-8000 (20%, w/v) and NaCl (2.5 M). The tube was inverted 10 times and stored on ice for 30 min followed by an additional centrifugation at 10 krpm (15300 x g) for 20 min at 4 °C. The supernatant was decanted, and tubes were centrifuged for an additional 4 min at 4 krpm (2429 x g) at 4 °C. The pellets were resuspended in PBS (10 mM phosphate, 137 mM NaCl, pH 7.2) with TWEEN 20 (0.05%, v/v) and glycerol (10%, v/v), separated into 1 mL aliquots, flash frozen with liquid nitrogen, and stored at -80 °C. For binding assays via ELISA, the purified phage was thawed on ice, precipitated a second time as before. The quality of each phage preparation was routinely checked by quality control ELISA, termed QC ELISA, to a FLAG peptide fused to the N-terminus of each epitope (Figure 2-6); additionally, PCR using Oligo69 and Oligo70 followed by DNA sequencing (Genewiz) was performed for every phage preparation. Such quality control allowed for identification of toxic clones; for example, C8, was apparently toxic to *E. coli*, and three protein epitopes failed to express in *E. coli* for

unknown reasons. The phage concentration was determined by absorbance at 260 nm using a coefficient of molar absorptivity of $0.003 \text{ nM}^{-1} \text{ cm}^{-1}$ and diluted to 40 nM in PBS.

Expression and Purification of eGFP-Ep9 and N protein. A pET28c plasmid containing Ep9 fused to an N-terminal eGFP (Figure 2-S6B) was transformed into BL21 DE3* *E. coli* heat shock, competent cells. A single colony was transferred to LB media (20 mL) supplemented with kanamycin (40 $\mu\text{g}/\text{mL}$) and incubated at 37 °C for 18 h. An aliquot of the starter culture (2.5 mL) was transferred to LB media with 1% glucose (250 mL LB in a 1 L baffled flask). After reaching an OD_{600} between 0.4-0.6, the culture was induced through addition of IPTG (0.5 mM) before incubation at 25 °C for 18 h. The cells were centrifuged (15,300 x g) for 20 min at 4 °C, and the cell pellet was resuspended in lysis buffer (25 mM Tris-HCl and 200 mM NaCl, pH 8.0 and supplemented with protease inhibitor cocktail) followed by sonication. The lysate was subjected to centrifugation (26,892 rcf, 45 min, 4 °C). The supernatant was incubated with charged Ni-IMAC resin overnight on a rotary shaker (150 rpm at 4 °C). The resin was equilibrated in a column, washed with wash buffer (20 mM imidazole in lysis buffer), and the purified protein was eluted using elution buffer (250 mM imidazole in lysis buffer). Elutions containing the purified protein were visualized using 10% or 12% SDS-PAGE (Bio-rad Mini-PROTEAN Tetra electrophoresis system) stained with Coomassie brilliant blue stain (Figure 2-12). The eluted fractions containing the purified eGFP-Ep9 were pooled and buffer exchanged for 3 column volumes (20 mL) with lysis buffer without imidazole using a 10 kDa cutoff microconcentrator (Vivaspin, Fisher Scientific). The protein concentration was determined by a bicinchoninic acid (BCA) assay or Bradford assay using the estimated MW (<http://www.expasy.org>). Similar to eGFP-Ep9, the full-length N protein was expressed in 250 ml LB with 1% glucose and induced with 0.25 mM IPTG at

OD₆₀₀ = 0.8. Protein overexpression cultures were incubated at 16 °C for 22 h. Lysis and purification were conducted as described above, using N protein lysis buffer (20 mM Tris-HCl, 300 mM NaCl, 5 mM MgCl₂, 5 mM BME, 10% glycerol pH 8.0). The purified full-length N protein was analyzed using 10% SDS-PAGE (Figure 2-12).

Patient sample collection. The UC Irvine Experimental Tissue Resource operates under a blanket IRB protocol (UCI #2012-8716) that gives ETR personnel ‘Honest Broker’ status and enables the collection of any fluid or tissue remnant in excess to that needed for clinical diagnosis and distribution to investigators under the conditions of their own IRB approval. Patients undergoing COVID testing in the Emergency Department or on the inpatient service with confirmed COVID (+) pharyngeal swabs, were followed for their blood collections daily. Specimens collected originally for diagnostic purposes were processed and stored by the hospital laboratory in a manner compliant with College of American Pathologists (CAP) standards. EDTA-anticoagulated whole blood was stored for 2 days at 4 °C after clinical diagnosis and released for research purposes. Plasma from heparin-anticoagulated blood was centrifuged immediately after collection and preserved at 4 °C for 3-4 days before being released for research use. All COVID (+) specimens were handled under BSL-2 conditions, aliquoted into screw cap cryovials, and stored at -80 °C long term with constant temperature monitoring. Specimens were coded by the ETR with unique de-identifiers, and accompanying clinical information was stripped of PHI, such that investigators could receive specimens under a Non-Human Subjects Determination exemption from the UC Irvine IRB. All samples from SARS-CoV-2 infected patients were inactivated by incubation in a water bath at 56 °C for 30 min ⁴⁹, aliquoted (40 µL each), and stored at -80 °C.

Phage ELISA with plasma. The phage-displayed SARS-CoV-2 epitopes were used in phage ELISAs with patient plasma samples diluted 100-fold in coating buffer (50 mM Na₂CO₃, pH 9.6). After incubation in a 96-well Nunc MaxiSorp flat-bottom microtiter plate with shaking at 150 rpm at 4 °C for 12-18 h, plasma was aspirated by a plate washer (BioTek). Next, the plate was treated with 100 µL per well of ChonBlock Blocking/Sample Dilution Buffer (Chondrex, Inc.) for 1 h with shaking at 150 rpm at room temperature and washed three times with wash buffer (0.05% v/v Tween-20 in PBS). The epitope displaying phage and controls were diluted to 1 nM in ChonBlock Blocking/Sample Dilution Buffer and 100 µL were added to each well before incubating for 2 h with shaking (150 rpm) at room temperature. The plate was then washed three times with wash buffer. The primary antibody, anti-M13-HRP (Creative Diagnostics), was diluted 1:5000 in ChonBlock Secondary Antibody Buffer and 100 µL was added per well; the plate was incubated for 1 h at 150 rpm and room temperature. Following three washes with wash buffer, 1-Step Ultra TMB-ELISA Substrate Solution (100 µL per well, ThermoScientific) was added. Absorbance of TMB substrate was measured twice at 652 nm by UV-Vis plate reader (BioTek) after 5 and 15 min of incubation.

ELISA of eGFP-Ep9 and full-length N protein with plasma. Varying doses, with a maximum concentration of 1.7 µM, of eGFP-Ep9, eGFP-FLAG or full-length N protein (fl-N) were diluted in PBS pH 8.0, and then immobilized on a 96-well Nunc MaxiSorp flat-bottom microtiter plate before incubation on a shaker (150 rpm) at 4 °C for 12 to 18 h. After incubation, unattached proteins were removed through aspiration using a plate washer (BioTek) and wells were blocked with 100 µL ChonBlock Blocking/Sample Dilution Buffer (Chondrex, Inc.) for 30 min with shaking (150 rpm) at room temperature. The plate was then

washed three times with wash buffer (0.05% v/v Tween-20 in PBS). Pooled plasma from five patients within each experimental group was diluted 100-fold in ChonBlock Blocking/Sample Dilution Buffer and 100 μ L was added to each well before incubating for 1 h with shaking (150 rpm) at room temperature. The plate was then washed three times with wash buffer. The detection antibody, IgG Fc Goat anti-Human, HRP (Invitrogen), was diluted 1:5000 in ChonBlock Secondary Antibody Buffer and 100 μ L was added per well; the plate was incubated for 30 min at 150 rpm and room temperature. Following six washes with wash buffer, 1-Step Ultra TMB-ELISA Substrate Solution (100 μ L per well, ThermoScientific) was added. Absorbance of TMB substrate was measured twice at 652 nm by UV-Vis plate reader (BioTek) after 5 and 15 min of incubation.

Serum coronavirus antigen microarray (COVAM). COVAM included 61 antigens across respiratory virus subtypes including 11 antigens from SARS-CoV-2 expressed in either baculovirus or HEK-293 cells as previously detailed ²². These antigens were provided by Sino Biological U.S. Inc. as either catalog products or custom synthesis service products. The antigens were printed onto microarrays, probed with human sera, and analyzed as previously described ⁵⁰⁻⁵². Briefly, lyophilized antigens were reconstituted with sterile water to a concentration of 0.1 mg/mL protein in PBS, and printing buffer was added. Antigens were then printed onto ONCYTE AVID nitrocellulose-coated slides (Grace Bio-Labs) using an OmniGrid 100 microarray printer (GeneMachines). The microarray slides were probed with human sera diluted 1:100 in 1X Protein Array Blocking Buffer (GVS Life Sciences, Sanford, ME) overnight at 4°C and washed with TTBS buffer (20 mM Tris-HCl, 150 mM NaCl, 0.05% Tween-20 in ddH₂O adjusted to pH 7.5 and filtered) three times for 5 min each. A mixture of human IgG and IgM secondary antibodies conjugated to quantum dot fluorophores Q800 and

Q585 respectively was applied to each of the microarray pads and incubated for 2 h at room temperature, and pads were then washed with TTBS three times for 5 min each and dried. The slides were imaged using an ArrayCam imager (Grace Bio-Labs) to measure background-subtracted median spot fluorescence. Non-specific binding of secondary antibodies was subtracted using a saline control. The mean fluorescence of the 4 replicate spots for each antigen was used for analysis.

Statistical analysis. The ELISA data were analyzed in GraphPad Prism 8. Since the total antibody content differs from person to person, the raw absorbance values for every patient sample were normalized and represented as the ratio as compared to a negative control. Analysis of variance (ANOVA) with Dunnett's multiple comparisons was performed to determine if values were statistically significant. Correlations between COVAM IgG/IgM and ELISA were determined by plotting normalized values on an XY graph and performing a non-parametric correlation analysis using a Spearman's rank correlation coefficient test.

For data visualization of clinical patient data, trends in data were evaluated using Knime Analytics Platform software. GraphPad Prism was used to calculate column statistics including mean, standard deviation, SEM, p-values, Odds Ratios, and Likelihood Ratios defined as sensitivity / (1 - specificity). ANOVA with Tukey's multiple comparisons test was used to evaluate antibody response and disease severity between patients with α Ep9 Abs, non-Ep9, α N Abs, or non α N Abs. Comparisons of patients with α Ep9 Abs and non- α Ep9 Abs were conducted using unpaired, two-tailed, parametric t-tests. Contingency graphs were statistically evaluated using Fisher's exact test, for groups with binary categorization, and Chi-squared test for groups with multiple categories. Different datasets were fitted with

linear or non-linear regression methods, the fit with the higher R^2 value was chosen. Correlations between two clinical parameters (e.g., IL-6 and AST) were evaluated using the Pearson coefficient or Spearman coefficients (r) for linear or non-linear regressions, respectively; r -values between 1.0-0.7 were considered strong correlations, r -values between 0.7 and 0.5 were considered a moderate correlation, and values below 0.5 were considered a weak correlation⁵³. The significance of the correlation was evaluated based on p -value <0.05 .

2.6 Supporting Figures and Tables

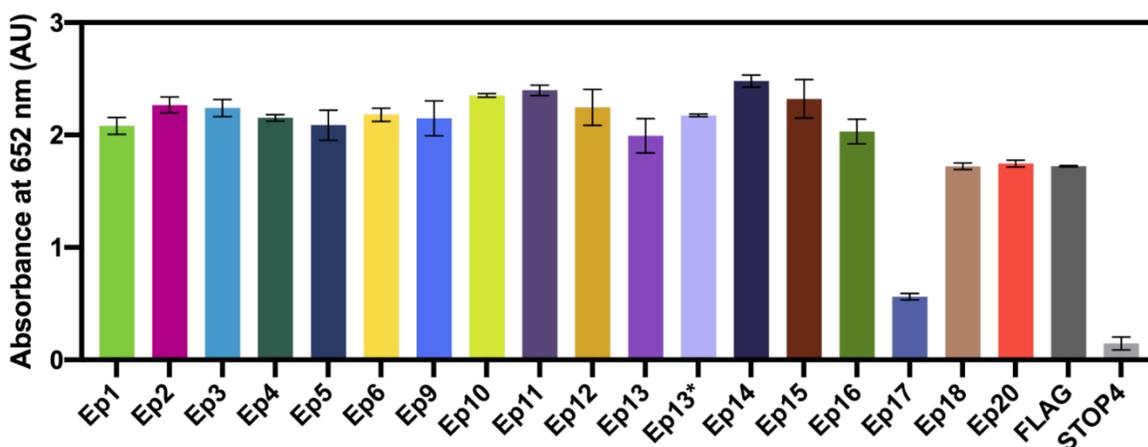


Figure 2-6. Quality control ELISA (QC ELISA) for phage-displayed, epitope candidates. Anti-FLAG antibodies (1:1000 in coating buffer) were immobilized on a microtiter plate. Subsequent steps followed the ELISA protocol provided here. Error bars represent SEM ($n = 3$). Ep8 was apparently toxic to *E. coli*, and Ep7 and Ep19 repeatedly failed sequencing quality controls after phage propagation.

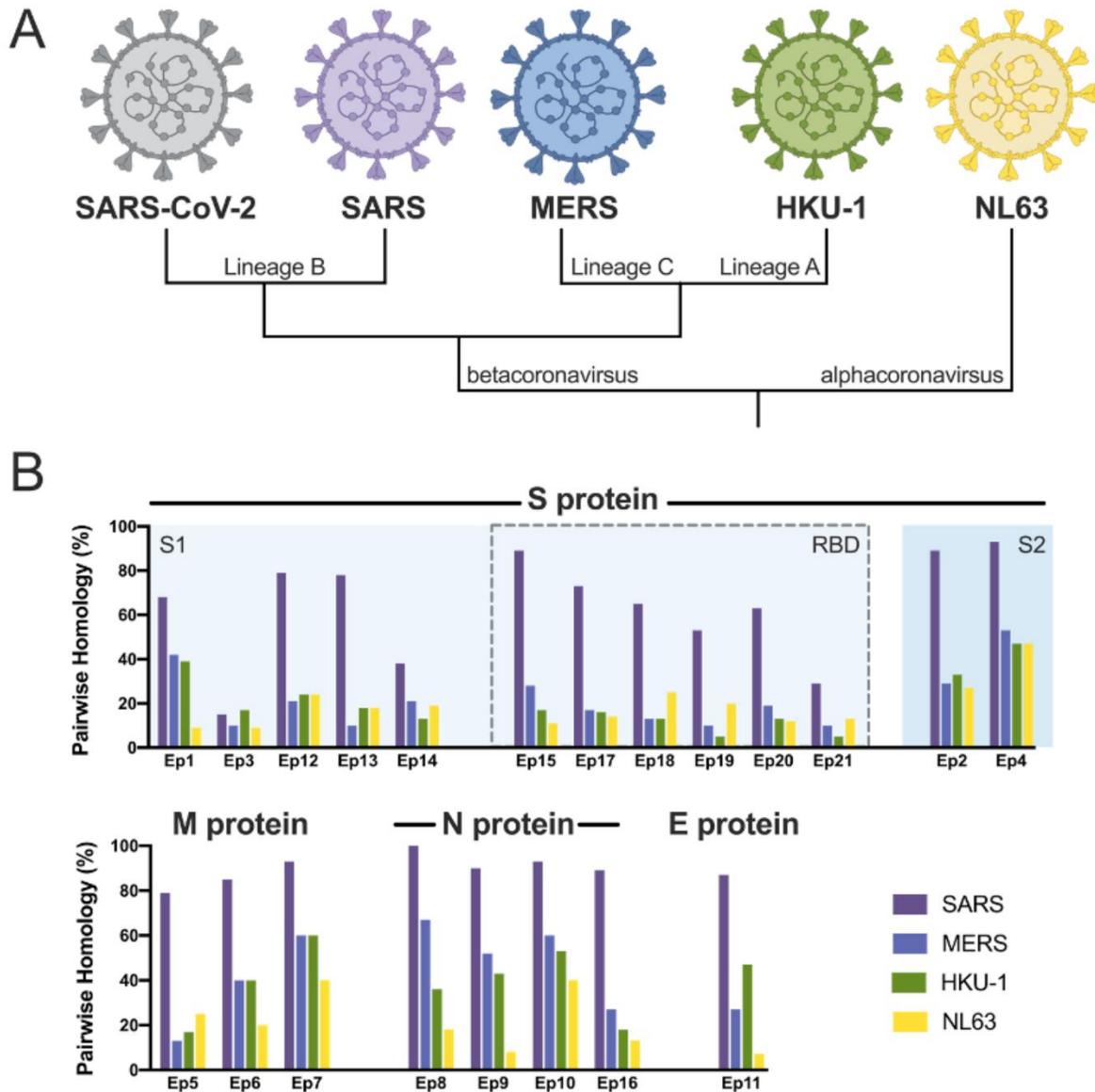


Figure 2-7. Epitope homology of SARS-CoV-2 with four phylogenetically related coronaviruses known to infect humans. **A.** Evolutionary lineages of the human coronaviruses investigated here, including the highly pathogenic (SARS-CoV-2, SARS, and MERS) and the less virulent (HKU-1 and NL63). **B.** The pairwise homology (% amino acid identity) between SARS-CoV-2 and the indicated coronavirus. Labels (top) indicate the proteins and domains (e.g., S1) from which the epitopes are derived.

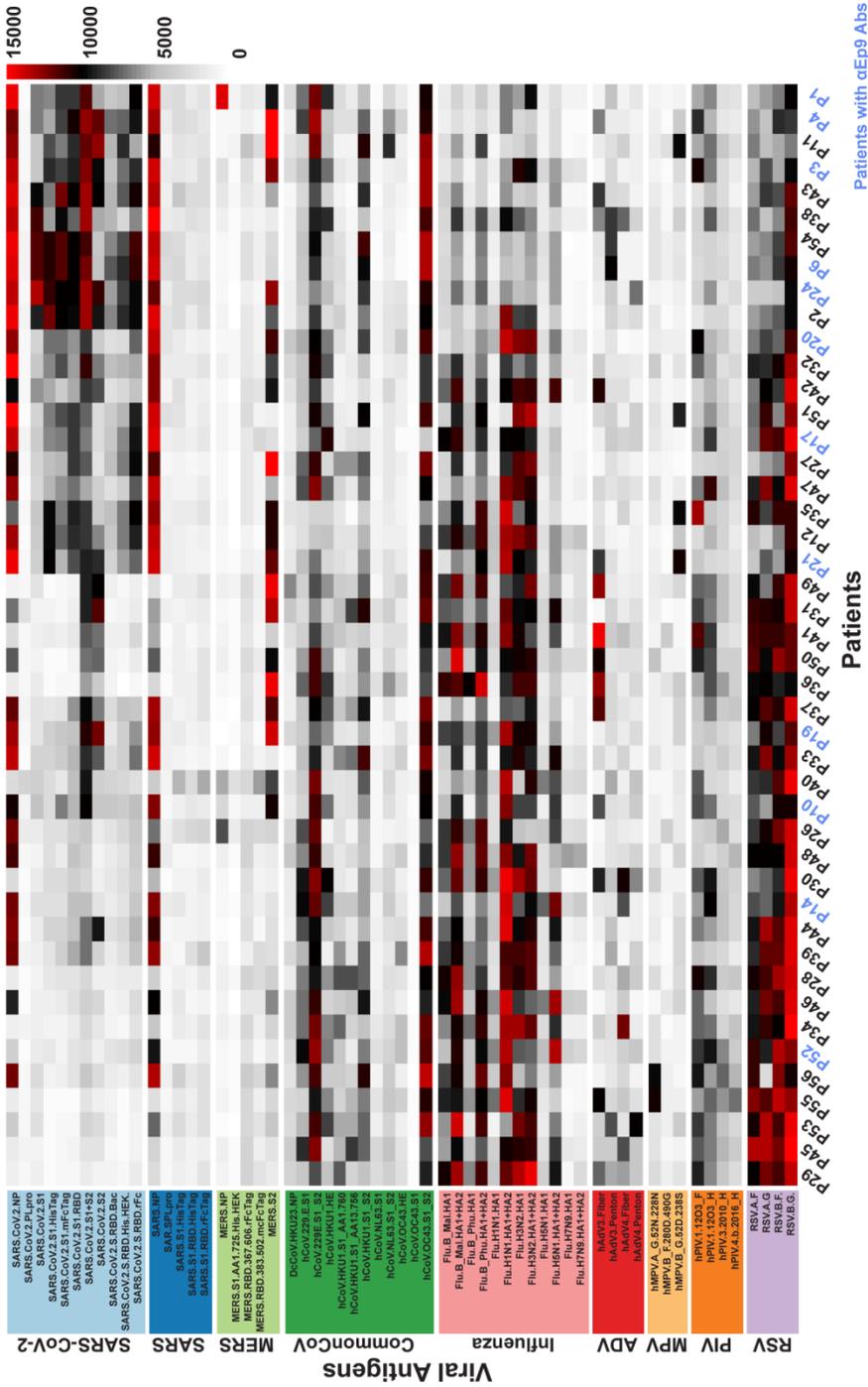


Figure 2-8. COVAM data showing the variation in IgG seroreactivity of patient plasma. The heatmap shows normalized signal intensity from plasma samples (n = 45). Plasma samples are in columns and sorted left to right by increasing average intensity to differentially reactive IgG, and viruses are in rows sorted by decreasing average seroreactivity.

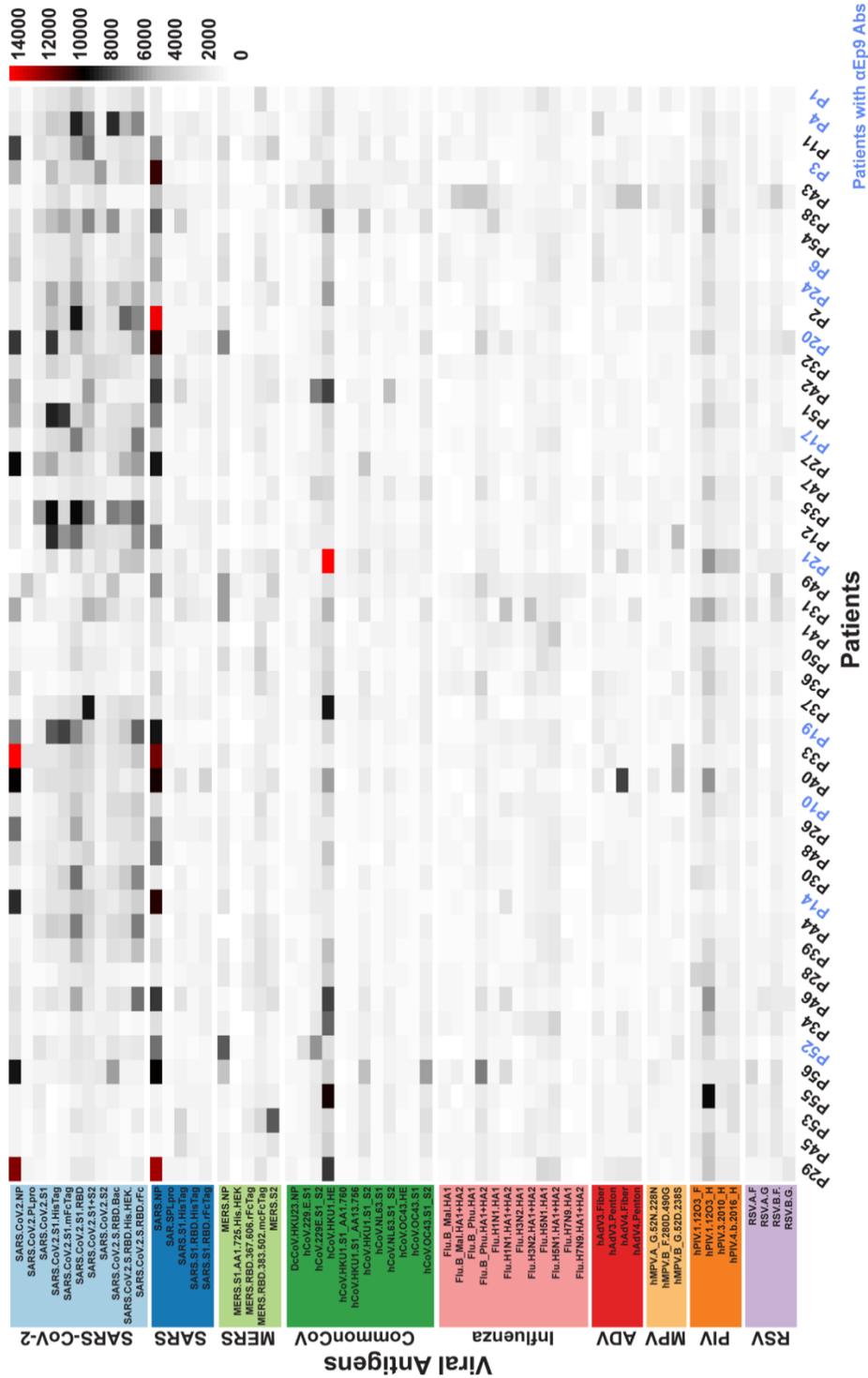


Figure 2-9. Variation in IgM seroreactivity of patient plasma. Heatmap showing normalized signal intensity from plasma samples (n = 45). Plasma samples are in columns and sorted left to right by increasing average intensity to differentially reactive IgM, and viruses are in rows sorted by decreasing average seroreactivity.

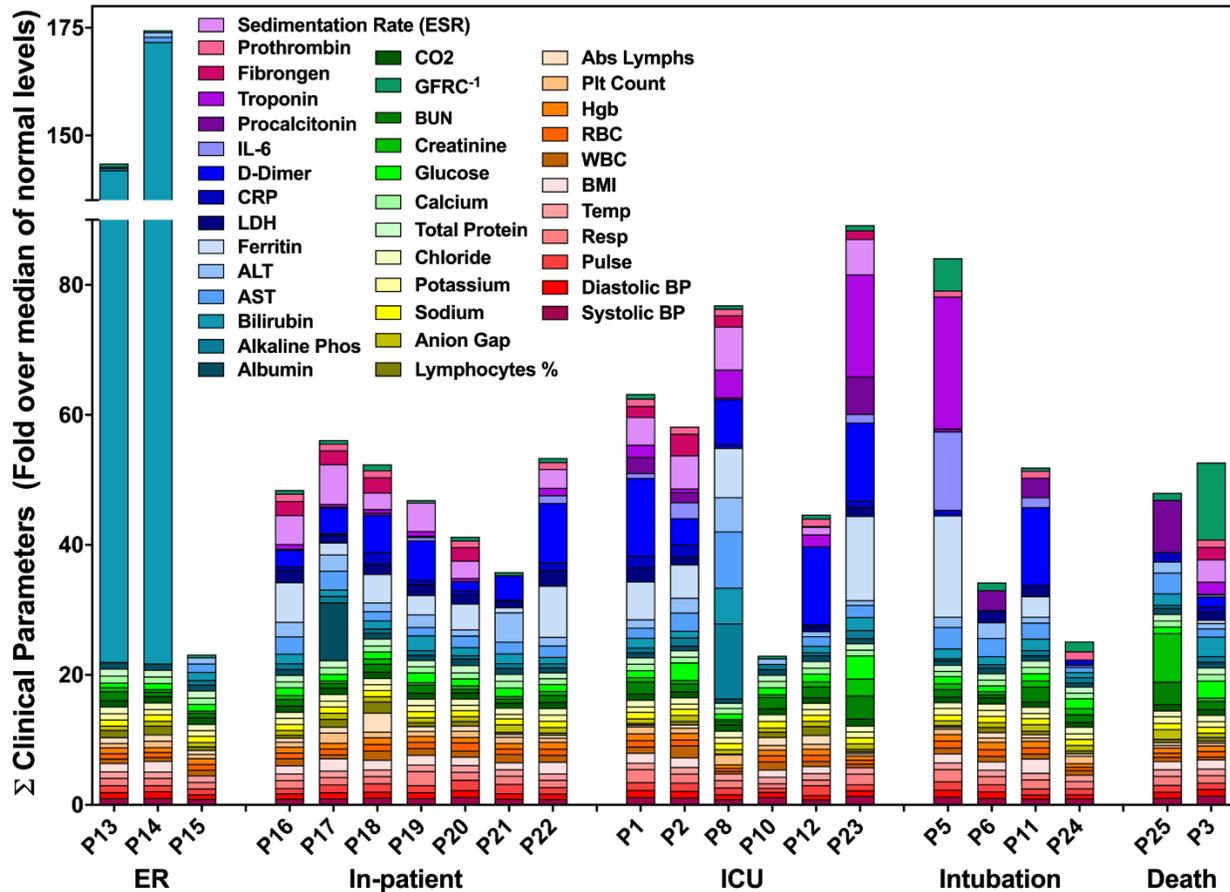


Figure 2-10. Comparison of disease severity and clinical parameters of patients with α Ep9 Abs. The data shown represents the fold change of each clinical parameter over the mean of the normal range. The sum of all the fold changes of the clinical parameters for each Ep9-responsive patient is binned according to COVID-19 disease severity. For facile visualization and comparison of clinical biomarkers between Ep9-responsive patients, the values of each parameter were normalized to fold over the mean of healthy values. No significant trends in clinical parameters (color indicated) were observed with increased disease severity or relative to patients lacking α Ep9 Abs.

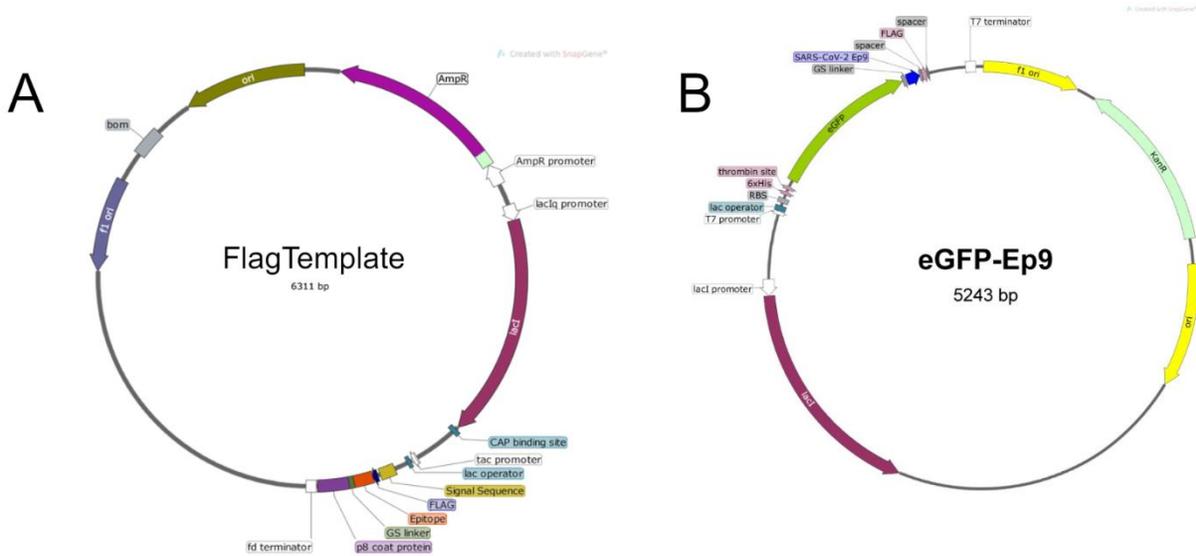


Figure 2-11. Schematic of plasmid map for phage-displayed or recombinant expression epitopes. **A.** Schematic of the FlagTemplate phagemid used for cloning phage-displayed epitopes. The phagemid, termed FlagTemplate, for the subcloning of SARS-CoV-2, SARS, HKU-1 and NL63 epitopes encodes an N-terminal FLAG tag, followed by a GSG linker to the epitope before a C-terminal GGGSGSSS linker to the P8 coat protein of M13-phage. **B.** The plasmid map for recombinant expression of the eGFP-Ep9 fusion. Ep9 was subcloned into a pET28-eGFP fusion vector with a C-terminal FLAG tag. The fusion protein is connected through a linker (GGGSGSS), and two spacers flank the N- and C-termini of the FLAG-tag (SGSG and GSG, respectively). The plasmid backbone lacking the Ep9 sequence was used to express the eGFP-FLAG negative control.

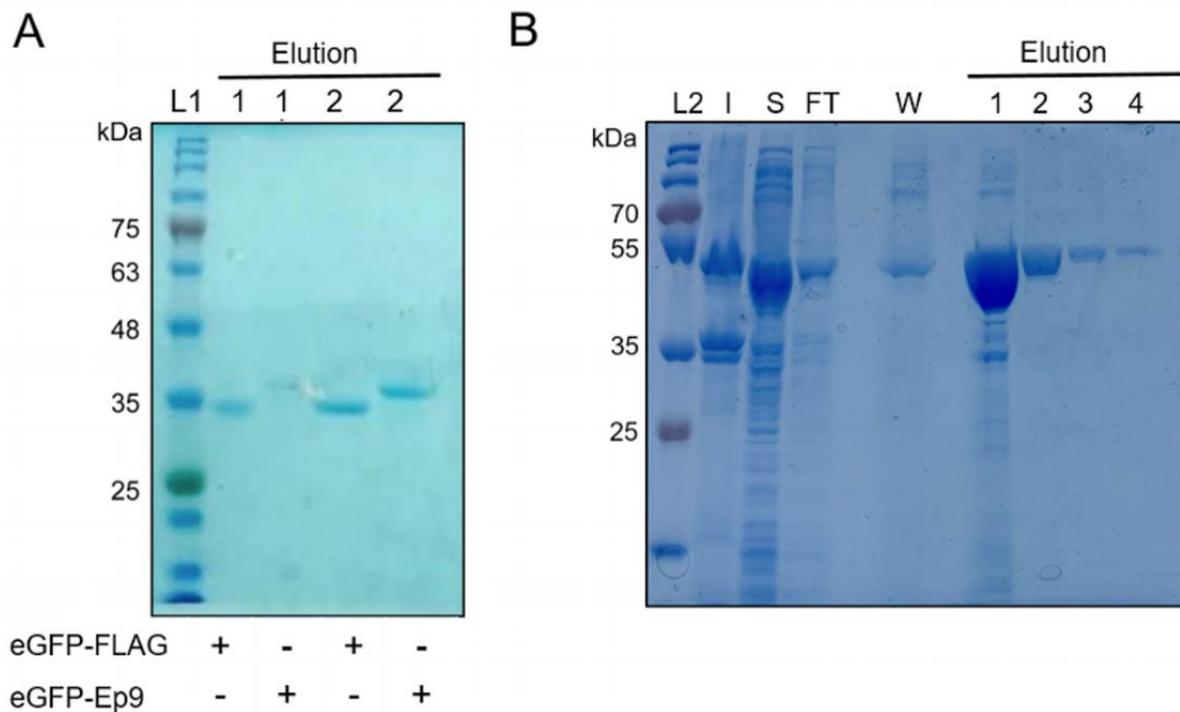


Figure 2-12. eGFP-FLAG, eGFP-Ep9, and full-length N protein purity assessed by 10% SDS-PAGE. Representative SDS-PAGE to visualize **A.** eGFP-FLAG and eGFP-Ep9 or **B.** full-length N protein after immobilized metal (Ni²⁺) affinity chromatography purification. Elution fraction #2 of eGFP-FLAG, eGFP-Ep9, and elution fraction #3 were purified to >95% homogeneity, and migrated at the expected mass of ≈ 32 , ≈ 34 , and ≈ 48 kDa, respectively. Elution fraction #1 of eGFP-FLAG and eGFP-Ep9 was diluted 20-fold. L1, BLUE stain (Goldbio) and L2, Prestain PAGE-Ruler Plus (Thermo Fisher Scientific) protein ladder was used as reference. I = insoluble fraction, S = soluble fractions, FT = flow-through, W = wash.

Table 2-1. Phage-displayed putative epitopes of SARS-CoV-2

| Epitope | Protein | Residues* | Amino Acid Sequence | Ref |
|---------|---------|-----------|--|----------|
| Ep1 | S | 287-317 | DAVDCALDPLSETKCTLKSFTVEKGIYQTSN | 13 |
| Ep2 | S | 802-819 | FSQILPDPSPKPSKRSFIE | 13 |
| Ep3 | S | 15-30 | CVNLTTRTQLPPAYTN | 14 |
| Ep4 | S | 1056-1070 | APHGVVFLHVTVYVPA | 12 |
| Ep5 | M | 1-24 | MADSNGTITVEELKKLLEQWNLVI | 13 |
| Ep6 | M | 132-151 | PLLESELVIGAVILRGHLRI | 13 |
| Ep7 | M | 97-111 | IASFRLFARTRSMWS | 15 |
| Ep8 | N | 41-61 | RPQGLPNNTASWFTALTQHGK | 13 |
| Ep9 | N | 152-172 | ANNAIVLQLPQGTTLPKGFI | 13 |
| Ep10 | N | 264-278 | ATKAYNVTQAFGRRG | 12 |
| Ep11 | E | 52-66 | VKPSFYVYSRVKKNLN | 12 |
| Ep12 | S | 524-598 | VCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSV I | 13 |
| Ep13 | S | 601-640 | GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYS TGS | 13 |
| Ep13* | S | 601-640 | GTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYS TGS | 25,26,52 |
| Ep14 | S | 61-76 | NVTWFHAIHVSQTNGT | 14 |
| Ep15 | S | 373-390 | SFSTFKCYGVSPTKLNDL | 14 |
| Ep16 | N | 354-400 | NKHIDAYKTFPPTEPKKDKKKKADETOALPQRQKKQQ TVTLLPAADL | 13 |
| Ep17 | S | 319-529 | RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNR KRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCF TNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPDDF TGCVIAWNSNNLDSKVGNYNYLYRLFRKSNLKPFER DISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGV GYQPYRVVVLSFELLHAPATVCGPKK | 23 |
| Ep18 | S | 488-507 | CYFPLQSYGFQPTNGVGYQP | 23 |
| Ep19 | S | 429-448 | FTGCVIAWNSNNLDSKVGGN | 23 |
| Ep20 | S | 448-466 | NYNLYRLFRKSNLKPFER | 23 |
| Ep21 | S | 467-487 | DISTEIQAGSTPCNGVEGFN | 23 |
| sEp9 | N | 153-173 | NNNAATVQLPQGTTLPKGFI | |
| mEp9 | N | 141-161 | NNNSAIVTQFAPGTKLPKNFH | |
| hEp9 | N | 166-186 | TTQEAIPTRFPPGTILPQGYI | |
| nEp9 | N | 119-136 | NQKPLEPKFSIALPPELS | |

*Residue numbering from protein sequences deposited in GenBank. Specifically, the accession numbers were as follows: S protein (YP_009724390.1), M protein (YP_009724393.1), N protein (YP_009724397.2), and E protein (YP_009724392.1) from SARS-CoV-2 and N protein from SARS (NP_828855.1), MERS (YP_009047211.1), HKU-1 (YP_173242.1), or NL63 (TP_003771.1).

Table 2-2. Demographics and clinical characteristics of COVID-19 patients categorized by α Ep9 Abs response.

| Characteristics | No α Ep9 Abs (n = 63) | α Ep9 Abs (n = 23) | p-value |
|---|-----------------------------------|--------------------------------|----------|
| Demographics | | | |
| Age (\pm SD) | 49.75 (\pm 18.45) | 47.26 (\pm 18.45) | 0.5668 |
| Sex F:M (%) | 21:42 (44.4/66.7) | 10:13 (43.5/56.5) | 0.4502 |
| Ethnicity n, (%) | 15 (65.2):4 (17.4): 3 (13.0): | 39 (61.9): 8 (12.7): 9 (14.3): | 0.7760 |
| Hispanic: Asian: Caucasian: Black: Other | 1 (4.3): 0 (0) | 3 (4.8): 4 (6.3) | |
| BMI (\pm SD) | 28.9 (\pm 6.4) | 32.0 (\pm 7.9) | 0.0642 |
| Preconditions, n (%) | | | |
| Hypertension | 23 (36.5) | 10 (43.5) | 0.6203 |
| Diabetes | 21 (33.2) | 6 (26.1) | 0.6065 |
| CVD | 6 (9.5) | 2 (8.7) | 1.0000 |
| CAD | 6 (9.5) | 2 (8.7) | 1.0000 |
| CKD/ESRD | 6 (9.5) | 2 (8.7) | 1.0000 |
| Asthma/COPD | 8 (12.7) | 3 (13.0) | 1.0000 |
| Obesity | 24 (38.1) | 13 (56.5) | 0.1461 |
| Cancer | 2 (3.17) | 3 (13.0) | 0.1163 |
| Symptoms, n (%) | | | |
| Total Days of Symptoms | 9.8 (\pm 9.0) | 17 (\pm 10) | 0.0059** |
| Cough | 43 (68.3) | 15 (65.2) | 0.7997 |
| Dyspnea/SOB | 28 (44.4) | 11 (47.8) | 0.8108 |
| Myalgia/Fatigue | 17 (27.0) | 8 (34.8) | 0.5926 |
| Headache | 12 (19.0) | 2 (8.7) | 0.3349 |
| Chest Pain | 7 (11.1) | 3 (13.0) | 1.0000 |
| Anosmia | 4 (6.3) | 2 (8.7) | 0.6561 |
| Stroke-like Symptoms | 0 | 2 (8.7) | 0.0692 |
| Abdominal Pain | 3 (4.8) | 0 | 0.5611 |
| Pulmonary Symptoms [^] (Pneumonia: Other: None) | 16 (25.4): 36 (52.4): 8 (12.7) | 13 (56.5): 7 (30.4): 1 (4.3) | 0.0421* |
| Severity, n (%) | | | |
| Asymptomatic | 3 | 0 | 0.5611 |
| Non-severe: Sever ^{^^} | 51:12 (n, severity 19.0%) | 10:13 (n, severity 56.5%) | 0.0013** |
| Days in Hospital | 5.8 (\pm 8.0) | 11.0 (\pm 10.7) | 0.0183* |
| Days in ICU | 12.6 (\pm 13.2) n = 11 | 12.5 (\pm 6.9), n = 12 | 0.8004 |
| Days on Ventilator | 14.0 (\pm 4.0), n = 6 | 13.0 (\pm 5.4), n = 7 | 0.7934 |

Results are presented as mean \pm SD or n and percentage of population. P-values for continuous variables are calculated using unpaired, two-tailed T-tests. P-values for categorical variables use Fisher's exact test for single value parameters, and Chi-squared test for multi-group variables. *, ** p-values < 0.05, 0.01, respectively. BMI = body mass index, CVD = cardiovascular disease, CAD = coronary artery disease, CKD = chronic kidney disease, ESRD = end-stage renal disease, SOB = shortness of breath, COPD = chronic obstructive pulmonary disease.

[^] Pulmonary symptoms are based on descriptive reports of X-ray and CT scans. "Other" pulmonary symptoms include, but are not limited to, atelectasis, pleural scarring, pleural effusion, pulmonary edema, mild peribronchial thickening.

^{^^} Non-severe include ER and in-patients only, severe includes patients in the ICU, on the ventilator or death.

2.7 References

1. World Health Organization. *Coronavirus Disease (COVID-19)*. (2021) doi:10.1097/jcn.0000000000000710.
2. Llanes, A. *et al.* Betacoronavirus genomes: How genomic information has been used to deal with past outbreaks and the covid-19 pandemic. *Int. J. Mol. Sci.* **21**, 1–28 (2020).
3. Richardson, S. *et al.* Presenting Characteristics, Comorbidities, and Outcomes among 5700 Patients Hospitalized with COVID-19 in the New York City Area. *J. Am. Med. Assoc.* **323**, 2052–2059 (2020).
4. Gallo Marin, B. *et al.* Predictors of COVID-19 severity: A literature review. *Rev. Med. Virol.* **31**, 1–10 (2020).
5. Song, P., Li, W., Xie, J., Hou, Y. & You, C. Cytokine storm induced by SARS-CoV-2. *Clin. Chim. Acta* **509**, 280–287 (2020).
6. Iwasaki, A. & Yang, Y. The potential danger of suboptimal antibody responses in COVID-19. *Nat. Rev. Immunol.* **20**, 339–341 (2020).
7. Arvin, A. M. *et al.* A perspective on potential antibody-dependent enhancement of SARS-CoV-2. *Nature* **584**, 353–363 (2020).
8. Huang, L. R. *et al.* Evaluation of antibody responses against SARS coronaviral nucleocapsid or spike proteins by immunoblotting or ELISA. *J. Med. Virol.* **73**, 338–346 (2004).
9. He, Y., Zhou, Y., Siddiqui, P., Niu, J. & Jiang, S. Identification of immunodominant epitopes on the membrane protein of the severe acute respiratory syndrome-associated coronavirus. *J. Clin. Microbiol.* **43**, 3718–3726 (2005).
10. Long, Q. X. *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat. Med.* (2020) doi:10.1038/s41591-020-0897-1.
11. Batra, M. *et al.* Role of IgG against N-protein of SARS-CoV2 in COVID19 clinical outcomes. *Sci. Rep.* **11**, 3455 (2021).
12. Fast, E., Altman, R. B. & Chen, B. Potential T-cell and B-cell Epitopes of 2019-nCoV. *bioRxiv* (2020).
13. Grifoni, A. *et al.* A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe* **27**, 671–680.e2 (2020).
14. Baruah, V. & Bose, S. Immunoinformatics-aided identification of T cell and B cell epitopes in the surface glycoprotein of 2019-nCoV. *J. Med. Virol.* **92**, 495–500 (2020).
15. Zheng, M. & Song, L. Novel antibody epitopes dominate the antigenicity of spike glycoprotein in SARS-CoV-2 compared to SARS-CoV. *Cell. Mol. Immunol.* **17**, 536–538 (2020).

16. Tilocca, B. *et al.* Comparative computational analysis of SARS-CoV-2 nucleocapsid protein epitopes in taxonomically related coronaviruses. *Microbes Infect.* **22**, 188–194 (2020).
17. Rakib, A. *et al.* Immunoinformatics-guided design of an epitope-based vaccine against severe acute respiratory syndrome coronavirus 2 spike glycoprotein. *Comput. Biol. Med.* **124**, 103967–103983 (2020).
18. Wang, L. *et al.* Serological Responses to Human Virome Define Clinical Outcomes of Italian Patients Infected with SARS-CoV-2. *medRxiv Prepr. Serv. Heal. Sci.* (2020) doi:10.1101/2020.09.04.20187088.
19. Zamecnik, C. R. *et al.* ReScan, a Multiplex Diagnostic Pipeline, Pans Human Sera for SARS-CoV-2 Antigens. *Cell Reports Med.* (2020) doi:10.1016/j.xcrm.2020.100123.
20. Amrun, S. N. *et al.* Linear B-cell epitopes in the spike and nucleocapsid proteins as markers of SARS-CoV-2 exposure and disease severity. *EBioMedicine* **58**, (2020).
21. Musicò, A. *et al.* SARS-CoV-2 epitope mapping on microarrays highlights strong immune-response to n protein region. *Vaccines* **9**, 1–11 (2021).
22. Hedde, P. N. *et al.* A modular microarray imaging system for highly specific COVID-19 antibody testing. *Lab Chip* **20**, 3302–3309 (2020).
23. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281–292 (2020).
24. Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220 (2020).
25. Baric, R. S. Emergence of a Highly Fit SARS-CoV-2 Variant. *N. Engl. J. Med.* **383**, 2684–2686 (2020).
26. Plante, J. A. *et al.* Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **575**, 1–19 (2020).
27. Korber, B. *et al.* Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **182**, 812–827.e19 (2020).
28. Zheng, W. *et al.* Deep-learning contact-map guided protein structure prediction in CASP13. *Proteins Struct. Funct. Bioinforma.* **87**, 1149–1164 (2019).
29. Tian, W. *et al.* Predictors of mortality in hospitalized COVID-19 patients: A systematic review and meta-analysis. *Journal of Medical Virology* vol. 92 1875–1883 (2020).
30. Simonnet, A. *et al.* High Prevalence of Obesity in Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) Requiring Invasive Mechanical Ventilation. *Obesity* **28**, 1195–1199 (2020).
31. Lee, L. Y. W. *et al.* COVID-19 prevalence and mortality in patients with cancer and the effect of primary tumour subtype and patient demographics: a prospective cohort study. *Lancet Oncol.* **21**, 1309–1316 (2020).

32. Zhang, L. *et al.* Clinical Course and Mortality of Stroke Patients With Coronavirus Disease 2019 in Wuhan, China. *Stroke* **51**, 2674–2682 (2020).
33. Charlson, M. E., Pompei, P., Ales, K. L. & MacKenzie, C. R. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J. Chronic Dis.* **40**, 373–383 (1987).
34. Lu, L. *et al.* Preventing Mortality in COVID-19 Patients: Which Cytokine to Target in a Raging Storm? *Front. Cell Dev. Biol.* **8**, 1–8 (2020).
35. Cummings, M. J. *et al.* Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. *Lancet* **395**, 1763–1770 (2020).
36. Yang, Y. *et al.* Plasma IP-10 and MCP-3 levels are highly associated with disease severity and predict the progression of COVID-19. *J. Allergy Clin. Immunol.* **146**, 119–127 (2020).
37. Shrock, E. *et al.* Viral epitope profiling of COVID-19 patients reveals cross-reactivity and correlates of severity. *Science*. eabd4250 (2020).
38. Tan, W. *et al.* Viral Kinetics and Antibody Responses in Patients with COVID-19. *medRxiv* (2020).
39. Jiang, H. wei *et al.* SARS-CoV-2 proteome microarray for global profiling of COVID-19 specific IgG and IgM responses. *Nat. Commun.* **11**, 3581–3592 (2020).
40. Benucci, M. *et al.* COVID-19 pneumonia treated with Sarilumab: A clinical series of eight patients. *Journal of Medical Virology* vol. 92 2368–2370 (2020).
41. Kooistra, E. J. *et al.* Anakinra treatment in critically ill COVID-19 patients: a prospective cohort study. *Crit. Care* **24**, 1–12 (2020).
42. Langer-Gould, A. *et al.* Early Identification of COVID-19 Cytokine Storm and Treatment with Anakinra or Tocilizumab. *Int. J. Infect. Dis.* **99**, 291–297 (2020).
43. Castelnovo, L. *et al.* Anti-IL6 treatment of serious COVID-19 disease. *Medicine (Baltimore)*. **100**, 1–6 (2021).
44. Salama, C. *et al.* Tocilizumab in Patients Hospitalized with Covid-19 Pneumonia. *N. Engl. J. Med.* **384**, 20–30 (2021).
45. Veiga, V. C. *et al.* Effect of tocilizumab on clinical outcomes at 15 days in patients with severe or critical coronavirus disease 2019: randomised controlled trial. *BMJ* **372**, n84 (2021).
46. Yamaoka, Y. *et al.* Whole Nucleocapsid Protein of Severe Acute Respiratory Syndrome Coronavirus 2 May Cause False-Positive Results in Serological Assays. *Clin. Infect. Dis.* **ciaa637**, 1–7 (2020).
47. Murase, K. *et al.* EF-Tu binding peptides identified, dissected, and affinity optimized by phage display. *Chem. Biol.* **10**, 161–168 (2003).

48. Bhasin, A. *et al.* Virus Bioresistor (VBR) for Detection of Bladder Cancer Marker DJ-1 in Urine at 10 pM in One Minute. *Anal. Chem.* **92**, 6654–6666 (2020).
49. Pastorino, B., Touret, F., Gilles, M., de Lamballerie, X. & Charrel, R. N. Heat inactivation of different types of SARS-CoV-2 samples: What protocols for biosafety, molecular detection and serological diagnostics? *Viruses* **12**, (2020).
50. Khan, S. *et al.* Use of an influenza antigen microarray to measure the breadth of serum antibodies across virus subtypes. *J. Vis. Exp.* **2019**, 1–8 (2019).
51. Jain, A. *et al.* Evaluation of Quantum dot immunofluorescence and a digital CMOS imaging system as an alternative to conventional organic fluorescence dyes and laser scanning for quantifying protein microarrays. *Proteomics* **16**, 1271–1279 (2016).
52. Nakajima, R. *et al.* Protein Microarray Analysis of the Specificity and Cross-Reactivity of Influenza Virus Hemagglutinin-Specific Antibodies. *mSphere* **3**, 1–15 (2018).
53. Mukaka, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.* **24**, 69–71 (2012).

CHAPTER 3

Evidence for Deleterious Antigenic Imprinting in SARS-CoV-2 Immune Response

Author contribution: Sanjana R. Sen, and Gregory A. Weiss designed research; Sanjana R. Sen, Emily C. Sanders, Alicia M. Santos, Keertna Bhuvan, Derek Y. Tang, Aidan A. Gelston, and Brian M. Miller performed research; Sanjana R. Sen, Emily C. Sanders and Gregory A. Weiss analyzed data; Joni L. Ricks-Oddie advised on statistical analysis; Sanjana R. Sen and Gregory A. Weiss wrote the manuscript.

3.1 Abstract

A previous report demonstrated the strong association between the presence of antibodies binding to an epitope region from SARS-CoV-2 nucleocapsid, termed Ep9, and COVID-19 disease severity. Patients with anti-Ep9 antibodies (Abs) had hallmarks of antigenic imprinting (AIM), including early IgG upregulation and cytokine-associated injury. Thus, the immunological memory of a previous infection was hypothesized to drive formation of suboptimal anti-Ep9 Abs in severe COVID-19 infections. This study identifies a putative primary antigen capable of stimulating production of cross-reactive, anti-Ep9 Abs. Binding assays with patient blood samples directly show cross-reactivity between Abs binding to Ep9 and only one bioinformatics-derived, homologous potential antigen, a sequence derived from the neuraminidase protein of H3N2 influenza A virus. This cross-reactive binding is highly influenza strain specific and sensitive to even single amino acid changes in epitope sequence. The neuraminidase protein is not present in the influenza vaccine, and the anti-Ep9 Abs likely resulted from the widespread influenza infection in 2014. Therefore, AIM from a previous infection could underlie some cases of COVID-19 disease severity.

3.2 Importance

Infections with SARS-COV-2 result in diverse disease outcomes, ranging from asymptomatic to fatal. The mechanisms underlying different disease outcomes remain largely unexplained.

Previously, our laboratory identified a strong association between the presence of an antibody and increased disease severity in a subset of COVID-19 patients. Here, we report that this severity-associated antibody cross-reacts with viral proteins from an influenza A viral strain from 2014. Therefore, we speculate that antibodies generated against previous infections, like the 2014 influenza A, play a significant role in directing some peoples' immune responses against SARS-COV-2. Such understanding of the sources and drivers of COVID-19 disease severity can help early identification and pre-emptive treatment.

3.3 Introduction

Original antigenic sin or antigenic imprinting (AIM) occurs when the immune response adapted for a primary (or “original”) infection instead targets a similar, but not identical, pathogen¹. Since B-cells undergo affinity maturation post-primary infection, cross-reactive Abs from previous infections can outcompete naïve Abs². AIM ideally accelerates pathogen clearance by targeting highly conserved antigens; however, suboptimal targeting by non-neutralizing, Ab binding can exacerbate disease². The range of outcomes observed in COVID-19, from asymptomatic to fatal, could result from a patient's immunological memory^{1,3}.

Ab cross-reactivity from AIM causes a wide range of disease outcomes. For example, some Abs from healthy individuals previously exposed to other common human coronaviruses (hCoV) could cross-react with SARS-CoV-2 spike protein to neutralize viral pseudotypes⁴. However, other prepandemic Abs with cross-reactivity to SARS-CoV-2 nucleocapsid (NP) and spike proteins did not protect against severe symptoms⁵. Humoral immunity to hCoVs, NL63 and 229E⁶, respiratory syncytial virus, cytomegalovirus and herpes simplex virus-1^{7,8} has been associated with more severe COVID-19 disease.

The presence of Abs with affinity for a 21-mer peptide derived from SARS-CoV-2 NP, an epitope region termed Ep9, have been correlated with severe COVID-19. The patients, termed $\alpha\text{Ep9}(+)$, comprised $\approx 27\%$ of the sampled, SARS-CoV-2-infected population ($n = 186$). The $\alpha\text{Ep9}(+)$ patients ($n = 34$) had high, early levels of αN IgGs, typically within the first week, compared to $\alpha\text{Ep9}(-)$ patients; $\alpha\text{Ep9}(+)$ individuals also experienced cytokine-related, immune hyperactivity⁹. These two observations suggest an AIM-based mechanism for the disease severity observed in $\alpha\text{Ep9}(+)$ patients. Here, we explore the epitope homology landscape and αEp9 Ab cross-reactivity to potentially identify an original antigen driving Ab-based immune response in $\alpha\text{Ep9}(+)$ patients.

3.4 Results and Discussion

Assays measured levels of αEp9 IgGs and IgMs from $\alpha\text{Ep9}(+)$ patients whose plasma was collected at various times post-symptom onset (PSO). Consistent with the hallmarks of AIM tracing a prior infection, αEp9 IgG levels appeared elevated as early as one day PSO in one patients. Similar IgG levels were observed in the patient population over >4 weeks (one-way ANOVA, $p = 0.321$); thus, αEp9 IgG started high and remained high. Levels of αEp9 IgMs amongst patients at various times PSO were also similar (one-way ANOVA, $p = 0.613$). The signals measured for αEp9 IgM levels were significantly lower than the equivalent αEp9 IgG levels (t-test, $p = 0.0181$) (Figure 3-3A); this difference could reflect lower IgM affinity, quantity, or both.

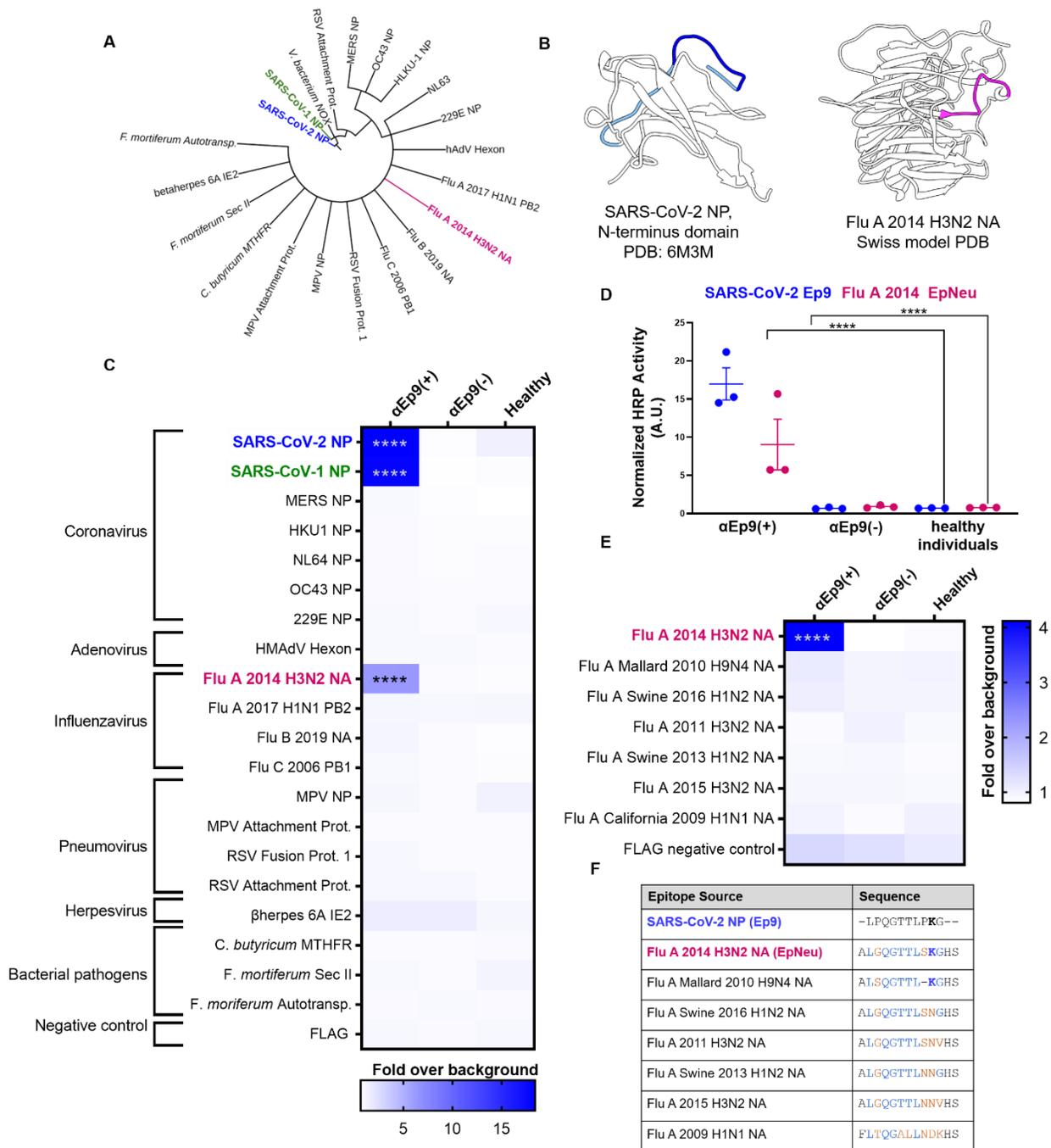


Figure 3-1. Potential OAS epitopes for binding α Ep9 Abs suggested by bioinformatics and tested by phage ELISA. **A.** Cladogram depicting sequence homology of the Ep9 sequence from SARS-CoV-2 to the bioinformatics-identified, closest homologs. Sequence alignments used pBLAST and VAST, and the cladogram was generated by iTOL²⁰. **B.** Structures of SARS-CoV-2 NP RNA binding domain (PDB: 6M3M) and the Flu A 2014 H3N2

NA protein (modeled by SWISS-Model ²¹). SARS-CoV-2 NP highlights Ep9 residues (light and dark blue) and the region homologous region to EpNeu (dark blue). The depicted model of Flu A 2014 H3N2 NA highlights the EpNeu putative antigen (pink). **C.** ELISAs examined binding of phage-displayed potential OAS epitopes to Abs from three sets of pooled plasma from five α Ep9(+) patients, or five α Ep9(-) patients. Pooled plasma from healthy patients was an additional negative control. The colors of the heat map represent the mean binding signal normalized to phage background negative controls (signal from phage without a displayed peptide). **D.** Expansion of data from panel C shows ELISA signals from the independently assayed individual pools shows results from the individual pools (**** $p < 0.0001$ for a two-way ANOVA comparing binding of phage-displayed epitopes listed in panel C to different groups of pooled plasma, *ad hoc* Tukey test). **E.** Using EpNeu as the search template to generate homologous sequences (shown in next panel), ELISAs examined EpNeu homologs' binding to pooled plasma from α Ep9(+), α Ep9(-), or healthy individuals. The data are represented as described in panel C (**** $p < 0.0001$ for two-way ANOVA c phage-displayed epitopes, *ad hoc* Tukey and Dunnett's test as shown). **F.** Amino acid sequence alignment of the closely related Flu A NA homologs of EpNeu from pBLAST¹⁰. Blue and orange residues represent conserved and mismatched amino acids, respectively, relative to Ep9. Bolded residues are important for epitope recognition by α Ep9 Abs. Here, the term Flu refers to influenza.

Searches for sequence and structural homologs of Ep9 using pBLAST¹⁰ and VAST¹¹ databases suggested candidate primary antigens. A structural homolog from betaherpesvirus 6A and 14 other Ep9 sequence homologs were identified. Additionally, Ep9-orthologous regions from six human coronaviruses (SARS-CoV, MERS, OC43, HKU-1, NL63, 229E) were chosen for subsequent assays (Figure 3-1A, and Table 3-1). To expedite the binding measurements, the potential AIM epitope regions were subcloned into phagemids encoding the sequences as fusions to the M13 bacteriophage P8 coat protein. DNA sequencing and ELISA experiments demonstrated successful cloning and consistent phage display, respectively. Two epitopes failed to display on phage and were omitted from subsequent investigation (Table 3-2 and Figure 3-4A).

Phage ELISAs tested binding by Ep9 homologs to α Ep9 Abs. An average response within the patient population was assessed using pooled plasma from three sets of five α Ep9(+) and five α Ep9(-) COVID-19 patients coated onto ELISA plates. Plasma from healthy individuals provided an additional negative control. Confirming previously reported results, SARS-CoV-2 Ep9 and a homologous epitope from SARS-CoV-1 (90% similarity) bound only to plasma from α Ep9(+) patients⁹. The α Ep9 Ab affinity for SARS-CoV-1 is unlikely to drive SARS-CoV-2 AIM due to the former's limited spread in the US¹².

The panel of potential epitopes revealed a candidate epitope from the neuraminidase (NA) protein of an H3N2 influenza A strain, which circulated in 2014 (A/Para/128982-IEC/2014, Accession No. AIX95025.1), termed EpNeu here. The plasma from α Ep9(+), but not α Ep9(-) patients nor healthy individuals, bound EpNeu ($p < 0.0001$, two-way ANOVA *ad hoc* Tukey test) (Figure 3-1C, D). Though Ep9 and EpNeu share 38% amino acid sequence similarity, other candidate epitope regions with significantly higher homology failed to bind to α Ep9(+) plasma (Table 3-1).

Next, the specificity of α Ep9 Abs binding to NA from different viral strains was explored. EpNeu provided a template for further homolog searches in sequence databases. Closely aligned NA sequences isolated from human, avian, and swine hosts in North America were chosen for further analysis (Figure 3-1F, Table 3-1). The sequences were phage-displayed as before. Despite their close similarity to EpNeu (up to 92.3% similarity or only one residue difference), none of the EpNeu homologs bound to Abs from α Ep9(+) patients (Figure 3-1E). A single EpNeu amino acid substitution, K142N (numbering from full-length NA, Accession No. AID57909.1) in an H1N2 swine flu (2016) dramatically decreased binding

affinity to Abs from α Ep9(+) patients ($p < 0.0001$ one-way ANOVA *ad hoc* Tukey). An epitope H4N6 avian influenza A (2010) missing residue S141, but including conserved K142, also greatly reduced binding to Abs from α Ep9(+) patients ($p < 0.0001$ one-way ANOVA *ad hoc* Tukey) (Figure 3-1E, 3-1F). Therefore, S141 and K142 are critical for binding to α Ep9 Abs.

Do Ep9 and EpNeu epitopes bind the same Abs? Data from 34 α Ep9(+) patients demonstrated a strong, highly significant correlation between levels of Abs binding to Ep9 and EpNeu epitopes in patient plasma (Figure 3-2A). Cross-reactivity was confirmed by a sandwich-format assay requiring bivalent, simultaneous binding to both eGFP-fused Ep9 and phage-displayed EpNeu (Figure 3-2B, S3). Cross-reactive Ab binding both Ep9 and EpNeu epitopes in pooled plasma from α Ep9(+) patients, but not in α Ep9(-) patients with other α NP Abs or healthy donors was demonstrated. Thus, we conclude that α Ep9 Abs also recognize the EpNeu epitope.

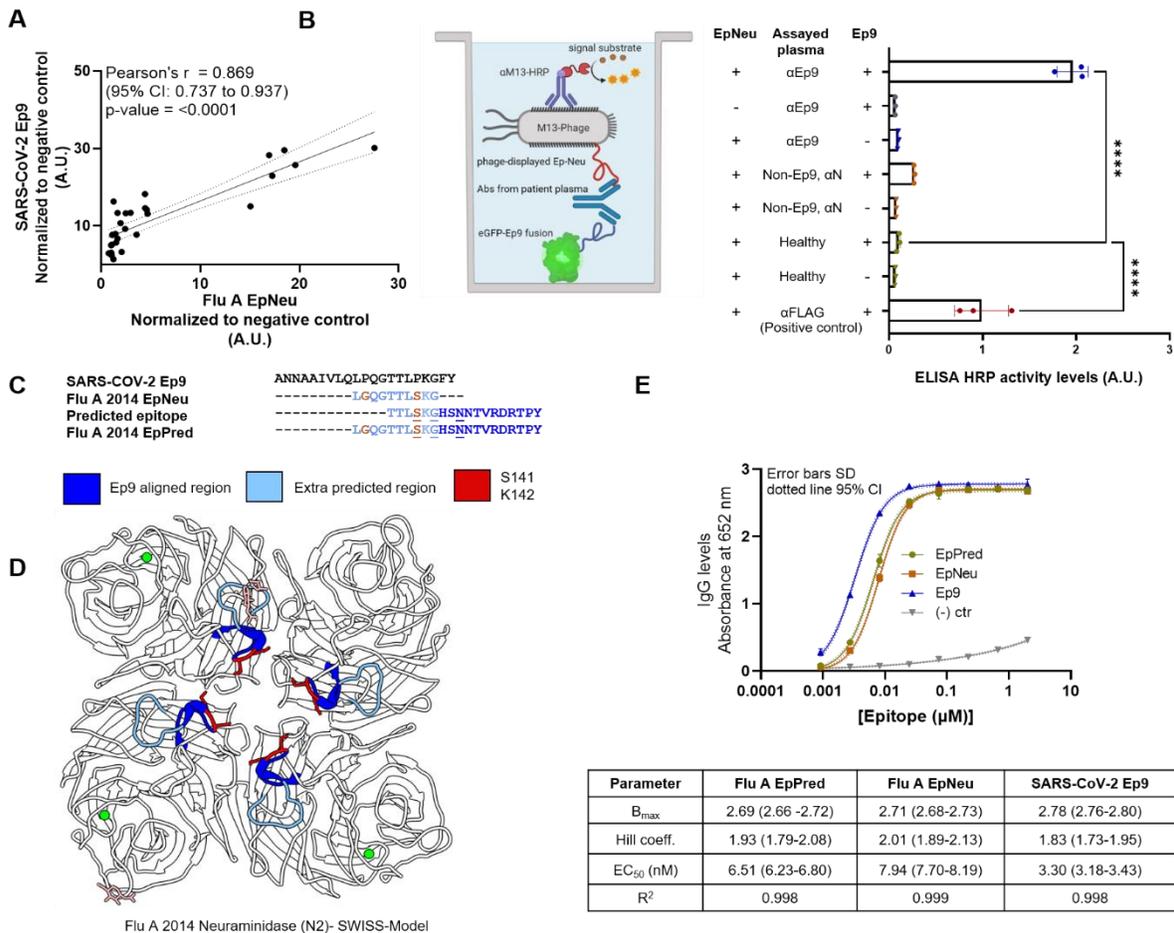


Figure 3-2. Cross-reactive Ab binding to both Ep9 and EpNeu, and EpNeu epitope prediction. **A.** Comparing normalized levels of phage-displayed Ep9 and EpNeu binding to plasma-coated wells from individual α Ep9(+) patients ($n = 34$). A strong correlation is observed, as shown by the depicted statistics. Each point in panels A through C represents data from individual patients. **B.** A schematic diagram of the sandwich ELISA to examine cross-reactivity of α Ep9 Abs. The assay tests for bivalent Ab binding to both Ep9 and EpNeu. Pooled plasma from five α Ep9(+) patients or five α Ep9(-) patients with other α NP Abs was tested for bivalent binding to both eGFP-fused Ep9 and phage-displayed EpNeu. Healthy patient plasma was used as a negative control. For additional negative controls, phage-FLAG and eGFP-FLAG replaced Ep9 and EpNeu, respectively (**** $p < 0.0001$ one-way ANOVA, *ad hoc* Tukey and Dunnett's test shown, with healthy plasma in the presence of EpNeu and Ep9 as negative control). Error bars represent SD. Individual points on bar graph represent technical replicates. **C.** Linear and structural B-cell epitope prediction tools Bepipred 2.0 and Discotope 2.0 suggested an extended, linear epitope region from the influenza A H3N2 2014 NA, including the eight residues of Ep9 Neu (light blue) with an additional ten, C-terminal residues (dark blue). This extended, predicted epitope is termed EpPred. Structural epitope predictions are underlined. Residues on EpNeu that are not aligned with Ep9 are depicted in

orange. **D.** Structural model depicting the influenza A H3N2 2014 NA. The model was generated using SWISS-Model based on the NA structure from influenza A H3N2 Tanzania 2010 (PDB: 4GZS). The NA structure highlights the EpNeu region (light blue), the extended residues in EpPred (dark blue), potential glycosylation sites (light pink), and the residues S141 and K142 (red), which are important for α Ep9 Ab recognition. **E.** Dose-dependent ELISA comparing binding of α Ep9 Abs to Ep9, EpNeu and EpPred. Pooled plasma from five α Ep9(+) patients and five α Ep9(-) patients were tested in triplicates with varying concentrations of eGFP-fused epitopes. The data demonstrates the strongest interactions occurred between α Ep9 Abs and Ep9 with an approximately 2-fold decrease in α Ep9 Abs binding affinity for EpNeu. EpPred bound slightly stronger to α Ep9 Abs than EpNeu; the difference in trend lines of EpNeu and EpPred are statistically significant ($p < 0.0001$, Comparison of Fits). Trendlines represent non-linear regression fit with Hill slope analysis.

We then investigated whether EpNeu could present a viable antigen during infection with 2014 H3N2 (NCBI: txid1566483). Linear epitope analysis of full-length NA protein (Bepipred 2.0)¹³ predicted a candidate antigen with eight residues from EpNeu, including S141 and K142, and ten additional residues (146-155). This predicted epitope region, termed EpPred, includes the conserved catalytic NA residue D151 targeted for viral neutralization by the immune system¹⁴ (Figures 2C, and S4A). A model structure of 2014 H3N2 NA from Swiss-Model¹⁵ and structural epitope prediction (Discotope 2.0)¹⁶ also identified potential epitopes within EpPred (Figures 2C, D and S4B).

eGFP-fused EpPred (Figure 3-4B) was assayed with pooled plasma from five α Ep9(+) patients. Controls included EpNeu and Ep9 (positive) and eGFP FLAG (negative). The α Ep9 Abs bound to Ep9 with \approx 2-fold stronger apparent affinity than for EpNeu (Figure 3-2E). The increased binding strength of Ep9 could result from additional rounds of Ab affinity maturation after the primary infection². The longer length EpPred modestly improved upon the binding of EpNeu to α Ep9 Abs (Figure 3-2E). Thus, α Ep9 Abs likely target a larger epitope of H3N2 2014 NA beyond regions homologous to Ep9; full-length NA's balkiness to

overexpression makes this difficult to test¹⁷. Additionally, the bacterially overexpressed epitopes assayed here do not include post-translational modifications. Taken together, the results support the hypothesis that α Ep9 Abs found in severe COVID-19 disease can result from AIM with H3N2 influenza A. Unfortunately, patient histories typically do not include influenza infections and vaccinations. Isolated from Para, Brazil, the H3N2 2014 strain has unknown spread in North America. However, a severe outbreak of influenza A was recorded in 2014^{18,19}. Since only hemagglutinin was sequenced for strain identification in 2014¹⁹, the candidate AIM strain from the current investigation could not be effectively traced as only its NA sequence was available. Notably, the EpNeu homolog from the 2014 vaccine H3N2 strain (identical to Flu A 2015 H3N2 NA, Accession No. ANM97445.1) does not bind α Ep9 Abs (Figure 3-1E, 3-1F). Therefore, α EpNeu Abs must have been generated against a primary influenza infection, not the vaccine.

This report offers a molecular mechanism for AIM underlying the high-rate of severe COVID-19 in α Ep9(+) patients. Specifically, we demonstrate cross-reactive binding between α Ep9 Abs and a predicted NA epitope from a 2014 influenza A strain. Future studies could examine correlation between a country's rate of the H3N2 2014 influenza and severe COVID-19. Additionally, correlation could be tested using health systems that record influenza infections. Examining epitope conservation and Ab cross-reactivity could predict AIM-based immune responses and disease outcomes in future infections. Identifying detrimental, benign or beneficial AIM pathways could also guide vaccine design.

3.5 Materials and Methods

Sequence and structural alignment analysis. To identify possible sources of primary infection responsible for Ep9 Ab generation, sequence and structural alignment with Ep9 residues and the SARS-CoV-2 NP was conducted. Alignment of Ep9 sequence with the orthologs from other human coronaviruses (hCoVs) such as SARS-CoV, MERS, HKU-1, NL63, 229E and OC43 was conducted using the Benchling sequence alignment tool ²² (<https://benchling.com>). To explore a wider range of human host pathogens pBLAST ¹⁰ (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to search for Ep9 homology in a database of non-redundant protein sequences; common human-host viruses were specified in the organism category. The queries were conducted with the blastp (protein-protein BLAST) program ¹⁰ with search parameters automatically adjusted for short input sequences. Alignments spanning >7 residues were included here. The Vector Alignment Search Tool (VAST) ¹¹ (<https://structure.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml>) was used to find structural alignment between SARS-CoV-2 Ep9 and proteins from other viral and bacterial human host pathogens. Alignment for NP from common hCoV were not further examined, as they had been included in sequence alignment analysis. The aligned sequences were sorted by the number of aligned residues as well as root-mean square deviation (RMDS). The top 50 structurally aligned proteins were then examined for structural homology in the Ep9 epitope region. Regions of proteins that aligned with the Ep9 region were selected for subsequent analysis.

Cloning. Predicted AIM epitopes were subcloned for phage display using the pM1165a phagemid vector ²³ with an N-terminal FLAG-tag and a C-terminal P8 M13-bacteriophage coat protein. AIM constructs were subcloned using the Q5 site-directed

mutagenesis kit (New England Biolabs, Ipswich, MA) as per manufacturer's instructions. After cloning, cells were transformed into XL-1 Blue *E. coli* and spread on carbenicillin-supplemented (50 µg/ml) plates. Individual colonies were then inoculated into 5 ml cultures, and shaken overnight at 37 °C. The phagemid was isolated using the QIAprep spin miniprep kit (Qiagen, Germantown, MD) as per manufacturer's instructions. Cloned sequences were verified by Sanger sequencing (Genewiz, San Diego, CA).

Phage propagation and purification. The Ep9 homologs were expressed as N-terminal fusions to the P8 coat protein of M13 bacteriophage. Plasmids were transformed into SS320 *E. coli* and spread onto carbenicillin-supplemented (50 µg/ml) LB-agar plates before overnight incubation at 37 °C. A single colony was inoculated into a primary culture of 15 ml of 2YT supplemented with 50 µg/ml carbenicillin and 2.5 µg/ml of tetracycline, and incubated at 37 °C with shaking at 225 rpm until an optical density at 600 nm (OD_{600}) of 0.5 to 0.7 was reached. 30 µM IPTG and M13KO7 helper phage at an MOI 4.6 was added to the primary culture, and the culture was incubated for an additional 37 °C with shaking at 225 rpm for 45 min. 8 ml of the primary culture was then transferred to 300 ml of 2YT supplemented with 50 µg/ml of carbenicillin and 20 µg/ml of kanamycin. The cultures were inoculated at 30 °C with shaking at 225 rpm for around 19 h.

The phage propagation culture was centrifuged at 9632 x *g* for 10 min at 4 °C. The supernatant, containing the phage, was transferred into a separate tubes pre-aliquoted with 1/5th volume of phage precipitation buffer (20% w/v PEG-8000 and 2.5 M NaCl), and incubated on ice for 30 min. The solution, containing precipitated phage, was centrifuged for 15 min at 4 °C, and the supernatant was discarded. The precipitated phage was centrifuged a second time at 1,541 x *g* for 4 min at 4 °C, and then dissolved in 20 ml of resuspension

buffer (10 mM phosphate, 137 mM NaCl, pH 7.4 - 8.0 with Tween-20 0.05%, v/v and glycerol 10% v/v). The resuspended pellet solution was divided into 1 ml aliquots, which were flash frozen with liquid nitrogen for storage in -80 °C. Prior to use in ELISA binding assays, the aliquoted phage-displayed constructs were re-precipitated in 0.2 ml of phage precipitation buffer after incubation for 30 min on ice. Aliquots were centrifuged at 12298 x *g* for 20 min at 4 °C and the supernatant was discarded. The phage pellets were re-centrifuged at 1968 x *g* for 4 min at 4 °C, and then resuspended in 1 ml of 10 mM phosphate, 137 mM NaCl, pH 7.4.

Expression and purification of eGFP fusion peptides. pET28c plasmids encoding eGFP fusions to C-terminal Ep9-FLAG, EpNeu-FLAG, EpPred-FLAG and FLAG (negative control) and N-terminal His₆ peptide epitopes, were transformed into BL21DE3 Star *E. coli* chemically competent cells. Transformants were spread on carbenicillin-supplemented (50 µg/ml) LB-agar plates and incubated at 37 °C overnight. Single colonies of each construct were selected to inoculate 25 ml LB media supplemented with carbenicillin (50 µg/ml). After incubation at 37 °C with shaking at 255 rpm overnight, 5 ml of seed cultures were used to inoculate 500 ml of LB media supplemented with carbenicillin (50 µg/ml). Expression cultures were incubated at 37 °C with shaking at 225 rpm until an OD₆₀₀ of ~0.5 was reached. The cultures were induced with 0.5 mM IPTG and incubated at 25 °C for 18 h. The cells were pelleted by centrifugation at 9632 x *g* for 20 min and resuspended in Tris-HCl lysis buffer (20 mM Tris-HCl, 250 mM NaCl, pH 8). Cells were lysed by sonication and the insoluble fractions were pelleted by centrifugation at 24696 x *g*. The supernatant was affinity-purified using Profinity™ IMAC (BioRad, Hercules, CA) resin charged with nickel sulfate. The protein lysate was batch bound overnight to the IMAC resin and purified using gravity columns. Columns were washed with lysis buffer supplemented with 20 mM imidazole, and the elution

fractions were collected from lysis buffer containing 250 mM imidazole. The elution fractions were then buffer-exchanged with lysis buffer lacking imidazole using Vivaspin® 20 Ultrafiltration Units (Sartorius, Goettingen, Germany) with a molecular weight cutoff of 10 kDa. The final buffer imidazole concentrations were calculated to be ~0.1 mM. Purified and buffer-exchanged protein fractions were then visualized using 10% SDS-PAGE with Coomassie dye staining.

Patient sample collection. Samples were collected as previously described ⁹. Briefly, the UC Irvine Experimental Tissue Resource (ETR) operates under a blanket IRB protocol (UCI #2012-8716) which enables sample collection in excess of requirements for clinical diagnosis, and allows distribution to investigators. Plasma was collected from daily blood draws of COVID(+) patients, initially confirmed with pharyngeal swabs. After immediate centrifugation, plasma from heparin-anticoagulated blood was stored for 3-4 days at 4 °C prior to its release for research use. Personal health information was omitted and unique de-identifier codes were assigned to patients to comply with the Non-Human Subjects Determination exemption from the UCI IRB. At the research facility, SARS-CoV-2 virus in plasma samples was inactivated through treatment by incubation in a 56 °C water bath for 30 min²⁴ prior to storage at -80 °C.

Phage ELISAs. As described in previous reports ⁹, pooled plasma from five αEp9(+) patients, five αEp9(-) patients, or healthy patients (Sigma-Aldrich, Saint Louis, MO) were separately prepared in coating buffer (50 mM Na₂CO₃, pH 9.6); the plasma was diluted 100-fold during this step. Plasma samples were then immobilized in 96 well microtiter plates by shaking the plasma solutions at 150 rpm at room temperature (RT) for 30 min. After aspiration and washing by plate washer (BioTek, Winooski, VT), each well was blocked

with 100 μ L of ChonBlock Blocking Buffer (CBB) (Chondrex, Inc., Woodinville, WA) for 30 mins, shaking at 150 rpm at RT. Wells were subsequently washed three times with PBS-T (0.05% v/v Tween-20 in PBS). Next, 1 nM phage-displayed candidate “original” epitopes and controls prepared in CBB was incubated in microtiter wells for 1 h at RT with shaking at 150 rpm. Unbound phage were aspirated and removed using three washes with PBS-T. The peroxidase-conjugated detection antibody, α M13-HRP (Creative Diagnostics, Shirley, NY), was diluted 1000-fold in Chonblock Secondary Antibody Dilution (Chondrex, Inc., Woodinville, WA) buffer; 100 μ l of this solution was added to each well before incubation for 30 min at RT with shaking at 150 rpm. Following aspiration and three washes (100 μ l each), 1-Step Ultra TMB-ELISA Substrate Solution (ThermoScientific, Carlsbad, CA) was added (100 μ l per well). Absorbance of TMB substrate was measured twice at 652 nm by UV-Vis plate reader (BioTek Winooski, VT) after 5 and 15 min of incubation. The experiment was repeated three times using plasma from different α Ep9(+) and α Ep9(-) patients for each experiments, using a total of 15 patients for each group. Each experiment was conducted in technical duplicate.

α Ep9 IgG and IgM ELISA. Plasma from 34 patients, previously tested for the presence of α Ep9 Abs using phage ELISAs⁹, were used to test levels of α Ep9 IgGs and IgMs. 2 μ M eGFP-Ep9 or eGFP-FLAG in PBS pH 8.0 were immobilized onto 96 well microtiter plates via overnight incubation with shaking at 150 rpm at 4 °C. Excess protein was aspirated and removed with three consecutive PBS-T washes. Wells were blocked by adding CBB (100 μ l) before incubation at 30 min at RT with shaking at 150 rpm. Next, α Ep9(+) patient plasma, diluted 1:100 in CBB (100 μ l), was added to duplicate wells before incubation at RT for 1 h with shaking at 150 rpm. The solutions were discarded and sample

wells were washed with PBS-T three times. α Ep9 Abs binding to the potential epitopes was detected using horse radish peroxidase (HRP) conjugated α Human Fc IgG (Thermo Fisher Scientific, Waltham MA) or α IgM μ -chain specific (Millipore Sigma, Temecula, CA) Abs diluted 1:5000 in ChonBlock Sample Antibody Dilution buffer. 100 μ l of detection Abs were added to each sample well, and incubated for 30 min at RT with shaking at 150 rpm. Sample wells were aspirated and washed three times in PBS-T, and the binding signal was detected after addition of TMB substrate (100 μ l per well).

Bivalent Abs binding ELISA. eGFP-Ep9 or eGFP-FLAG was serially diluted (120 nM, 40 nM, 13 nM and 4 nM) in PBS pH 8.0, and added to the appropriate wells in 96 well microtiter plates, followed by shaking overnight at 150 rpm at 4 °C. Excess unbound protein was removed, and the plate was washed three times in PBS-T. Wells were then blocked in CBB and incubated for 30 min at RT. After blocking, pooled plasma (100 μ l per well) from either five α Ep9(+) patients, or five non- α Ep9, α NP(+) patients, or healthy individuals was added to the appropriate wells. Plasma from pooled patients was diluted 100-fold in CBB. As a positive control α FLAG Ab was used as a 1:2000 dilution in CBB. Samples were incubated for 1 h at RT with 150 rpm shaking. The solution was removed by aspiration, and the plate and washed three times with PBS-T. Then 1 nM EpNeu displaying phage or the phage negative control with no epitopes displayed was diluted in CBB. 100 μ l phage solution was added to microtiter wells and incubated for 30 min at RT with shaking at 150 rpm. After aspirating and washing off unbound phage, binding of phage-displayed EpNeu to plasma α Ep9 Abs was visualized using α M13-HRP Ab diluted 1:10,000 in ChonBlock Sample Antibody Dilution buffer. Samples were incubated for 30 min at RT with 150 rpm shaking, and unbound Abs were removed through washing with PBS-T three

times before addition of TMB substrate (100 µl). Experiments were conducted in technical triplicates and repeated three times with different αEp(+) and αEp(-) patient samples.

Dose-dependent ELISA. Wells of microtiter plates were coated with serially diluted concentration of eGFP-Ep9, EpNeu and EpPred or eGFP-FLAG, and incubated overnight at 4 °C before blocking as described above. Next, pooled plasma (100 µl per well) from either five αEp9(+) patients, or five αEp9(-) patients, or healthy individuals at 1:100 total plasma dilution in CBB was added to the appropriate wells. Samples were incubated for 1 h at RT with shaking at 150 rpm. After incubation, unbound solution was removed, and the plates were washed three times with PBS-T. αEp9 IgG levels were detected by adding αFc IgG-HRP diluted 1:5000 in ChonBlock Sample Dilution buffer, followed by incubation for 30 min at RT with shaking at 150 rpm, followed by addition of TMB substrate (100 µl per well). Experiments were conducted in technical triplicates and repeated three times with different αEp(+) and αEp(-) patient samples.

Linear B-cell epitope prediction. Linear epitopes from the Influenza A/Para/128982-IEC/2014(H3N2) neuraminidase protein were predicted using the partial sequence with Accession AIX95025.1 from the National Center for Biotechnology Information's GenBank and the linear B-cell epitope prediction tool, Bepipred 2.0¹³ (<http://www.cbs.dtu.dk/services/BepiPred-2.0/>). The prediction thresholds were set to 0.5. The specificity and sensitivity of epitope prediction at this threshold is 0.572 and 0.586, respectively.

Structure-based B-cell epitope prediction. The structure of Influenza A/Para/128982-IEC/2014(H3N2) neuraminidase protein was modelled using Swiss-Model²¹ (<https://swissmodel.expasy.org/interactive>). Using the ProMod3 3.2.0 tool¹⁵, a

structural model was generated based on the crystal structure (2.35Å, PDB 4GZS 1.A) of a homologous H3N2 neuraminidase with 96.39% sequence identity. Modelling methods and quality assessments are further detailed in the report below.

The structural model of Influenza A/Para/128982-IEC/2014(H3N2) neuraminidase was used to predict structure-based epitopes. Using the *in silico* online platform DiscoTope 2.0¹⁶ (<http://www.cbs.dtu.dk/services/DiscoTope-2.0/>), structure-based epitope propensity scores were calculated to predict likely B-cell epitope residues. The score of -3.7 was set as the threshold for epitope prediction, which estimates a specificity and sensitivity of 0.75 and 0.47, respectively (Figure 3-6)

Statistical analysis. The ELISA data were analyzed in GraphPad Prism 9 (<https://www.graphpad.com>). Since the ELISA assays of 21 potential AIM epitopes were conducted over several microtiter plates for repeated experiments, the raw absorbance values for every patient sample were normalized and represented as the ratio of phage negative control to the signal. For heatmaps, two-way Analysis of variance (ANOVA) with a Tukey adjustment for multiple comparisons tests were conducted for the entire dataset of epitopes. For column comparisons of two groups, for example IgM levels and IgG levels in the α Ep(+) patients, unpaired, two-tailed, parametric t-tests were applied. Additionally, for column comparisons between more than two groups, for example IgM or IgG levels groups by weeks PSO, One-way ANOVA with a Tukey adjustment for multiple comparisons tests were used. Where indicated, an ANOVA with a Dunnett's adjustment were performed to compare results to healthy Abs interactions to α Ep9(+) patient results. Graphs represent SD error bars for technical replicates, defined as replicates of the same conditions in multiple wells of the same plate. Whereas error bars are shown as SEM when an experiment is

repeated with different patient sample sets. Correlations between Ep9 and EpNeu levels in patients were determined by plotting normalized values on an XY graph and performing a linear Pearson's correlation coefficient test, where a r coefficient between 1.0-0.7 were considered strong correlations, values between 0.7 and 0.5 were considered a moderate correlation, and values below 0.5 were considered a weak correlation²⁵. The significance of the correlation was evaluated based on p-value <0.05.

Model Building Report

This document lists the results for the homology modelling project "Untitled Project" submitted to SWISS-MODEL workspace on March 30, 2021, 10:35 a.m.. The submitted primary amino acid sequence is given in Table T1.

If you use any results in your research, please cite the relevant publications:

- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46(W1), W296-W303 (2018). [PM](#) [doi>](#)
- Bienert, S., Waterhouse, A., de Beer, T.A.P., Tauriello, G., Studer, G., Bordoli, L., Schwede, T. The SWISS-MODEL Repository - new features and functionality. *Nucleic Acids Res.* 45, D313-D319 (2017). [PM](#) [doi>](#)
- Studer, G., Tauriello, G., Bienert, S., Biasini, M., Johner, N., Schwede, T. ProMod3 - A versatile homology modelling toolbox. *PLoS Comp. Biol.* 17(1), e1008667 (2021). [PM](#) [doi>](#)
- Studer, G., Rempfer, C., Waterhouse, A.M., Gumienny, G., Haas, J., Schwede, T. QMEANDisCo - distance constraints applied on model quality estimation. *Bioinformatics* 36, 1765-1771 (2020). [PM](#) [doi>](#)
- Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Scientific Reports* 7 (2017). [PM](#) [doi>](#)

Results

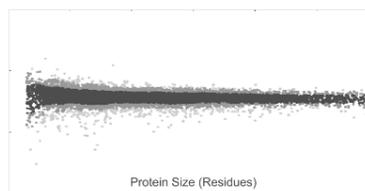
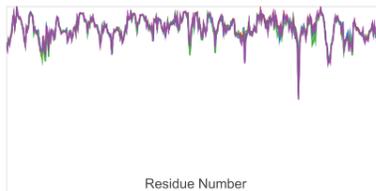
The SWISS-MODEL template library (SMTL version 2021-03-25, PDB release 2021-03-19) was searched with BLAST (Camacho et al.) and HHblits (Steinegger et al.) for evolutionary related structures matching the target sequence in Table T1. For details on the template search, see Materials and Methods. Overall 189 templates were found (Table T2).

Models

The following model was built (see Materials and Methods "Model Building"):

| Model #01 | File | Built with | Oligo-State | Ligands | GMQE | QMEAN |
|---|------|---------------|-------------------------------------|---|------|-------|
|  | PDB | ProMod3 3.2.0 | homo-tetramer (matching prediction) | 4 x CA: CALCIUM ION; 1 x NAG: 2-acetamido-2-deoxy-beta-D-glucopyranose; 1 x NAG-FUC: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose; | 0.95 | -1.40 |

| | | |
|-----------|---|-------|
| QMEAN |  | -1.40 |
| C β |  | -1.37 |
| All Atom |  | -1.65 |
| solvation |  | -1.52 |
| torsion |  | -0.61 |



| Template | Seq Identity | Oligo-state | QSQE | Found by | Method | Resolution | Seq Similarity | Range | Coverage | Description |
|----------|--------------|---------------|------|----------|--------|------------|----------------|---------|----------|---------------|
| 4gzs.1.A | 96.39 | homo-tetramer | 1.00 | HHblits | X-ray | 2.35Å | 0.62 | 1 - 360 | 1.00 | Neuraminidase |

Included Ligands

| Ligand | Description |
|-------------|---|
| 4 x CA | CALCIUM ION |
| 1 x NAG | 2-acetamido-2-deoxy-beta-D-glucopyranose |
| 1 x NAG-FUC | alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose |

Excluded ligands

| Ligand Name.Number | Reason for Exclusion | Description |
|-----------------------|----------------------------|---|
| CA.16 | Not in contact with model. | CALCIUM ION |
| EPE.10 | Not biologically relevant. | 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID |
| EPE.13 | Not biologically relevant. | 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID |
| EPE.18 | Not biologically relevant. | 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID |
| EPE.22 | Not biologically relevant. | 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID |
| NAG.11 | Clashing with protein. | 2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG.14 | Clashing with protein. | 2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG.15 | Clashing with protein. | 2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG.19 | Clashing with protein. | 2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG.20 | Not in contact with model. | 2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG.24 | Clashing with protein. | 2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG-FUC.2 | Clashing with protein. | alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG-FUC.4 | Clashing with protein. | alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG-NAG.8 | Not in contact with model. | 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG-NAG-BMA-MAN-MAN.1 | Clashing with protein. | alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG-NAG-BMA-MAN-MAN.3 | Clashing with protein. | alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG-NAG-BMA-MAN-MAN.5 | Clashing with protein. | alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose |
| NAG-NAG-BMA-MAN-MAN.7 | Not biologically relevant. | alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose |

Target AGGDIWVTRPEYVSCDPDKGNQFALGQGTTLKSGHSNNTVRDRTPYRLLMNELGVPFHLGKQVCIWSSSSSCHDGKAW
 4gzs.1.A AGGDIWVTRPEYVSCDPDKCYQFALGQGTTLNNVHSNNTVRGRTPYRLLMNELGVPFHLGKQVCIWSSSSSCHDGKAW

Target LHCITGDDKNATASFYINGRLVDSVVSWSKEVLRTQESECVCINGTCTVMTDGSASGKADTKILFIEEGKIVHTSTLS
 4gzs.1.A LHCITGDDKNATASFYINGRLVDSVVSWSKEILRTQESECVCINGTCTVMTDGSASGKADTKILFIEEGKIVHTSTLS

Target GSAQHVEECSCYPRYPGVRVCRDNWKGSNRPIDINIKDHSIVSSVYCSGLVGDTPRKNDSSSSHCLNPNNEEGGHV
 4gzs.1.A GSAQHVEECSCYPRYPGVRVCRDNWKGSNRPIDINIKDHSIVSSVYCSGLVGDTPRKNDSSSSHCLDNPNEEGGHV

Target KGWAFDDGNDVWMGRITINETSRLGYETFVKIEGWSNPKSKLQINRQVIVDRGDRSGYSGIFSVVEGKSCINRCFYVELIRG
 4gzs.1.A KGWAFDDGNDVWMGRITINETSRLGYETFVKIEGWSNPKSKLQINRQVIVDRGNRSGYSGIFSVVEGKSCINRCFYVELIRG

Target RKEETEVLWTSNSILLFCGTSYGTGGSWPDAADLNLMPI
 4gzs.1.A RKEETEVLWTSNSIVVFCGTSYGTGGSWPDAADLNLMPI

Target AGGDIWVTREPYVSCDPKGNQFALGQGTTLKSGHSNNTVRDRTPYRLLMNELGVPFHLGKQVCIWSSSSCHDGKAW
 4gzs.1.B AGGDIWVTREPYVSCDPKCYQFALGQGTTLNNVHSNNTVRGRTPYRLLMNELGVPFHLGKQVCIWSSSSCHDGKAW

Target LHCITGDDKNATASFIYNGRLVDSVVSWSKEVLRQTQESECVCTVVMTDGSASGKADTKILFIEEGKIVHTSTLS
 4gzs.1.B LHCITGDDKNATASFIYNGRLVDSVVSWSKEILRTQESECVCTVVMTDGSASGKADTKILFIEEGKIVHTSTLS

Target GSAQHVEECSCYPYRPGVRCVCRDNWKGSNRPVVDINIKDHSIVSSVYVCSGLVGDTPRKNDSSSSGHCLNPNNEEGGHV
 4gzs.1.B GSAQHVEECSCYPYRPGVRCVCRDNWKGSNRPVVDINIKDHSIVSSVYVCSGLVGDTPRKNDSSSSHCLDPNNEEGGHV

Target KGWAFDDGNDVWMGRINETSRLGYETFKVIEGWSNPKSKLQINRQVIVDRGDRSGYSGIFSVEGKSCINRCFYVELIRG
 4gzs.1.B KGWAFDDGNDVWMGRINETSRLGYETFKVIEGWSNPKSKLQINRQVIVDRGNRSGYSGIFSVEGKSCINRCFYVELIRG

Target RKEETEVLWTSNSILLFCGTSPTYGTGSPDAADLNLMPI
 4gzs.1.B RKEETEVLWTSNSIVVFCGTSPTYGTGSPDGADLNLMPI

Target AGGDIWVTREPYVSCDPKGNQFALGQGTTLKSGHSNNTVRDRTPYRLLMNELGVPFHLGKQVCIWSSSSCHDGKAW
 4gzs.1.C AGGDIWVTREPYVSCDPKCYQFALGQGTTLNNVHSNNTVRGRTPYRLLMNELGVPFHLGKQVCIWSSSSCHDGKAW

Target LHCITGDDKNATASFIYNGRLVDSVVSWSKEVLRQTQESECVCTVVMTDGSASGKADTKILFIEEGKIVHTSTLS
 4gzs.1.C LHCITGDDKNATASFIYNGRLVDSVVSWSKEILRTQESECVCTVVMTDGSASGKADTKILFIEEGKIVHTSTLS

Target GSAQHVEECSCYPYRPGVRCVCRDNWKGSNRPVVDINIKDHSIVSSVYVCSGLVGDTPRKNDSSSSGHCLNPNNEEGGHV
 4gzs.1.C GSAQHVEECSCYPYRPGVRCVCRDNWKGSNRPVVDINIKDHSIVSSVYVCSGLVGDTPRKNDSSSSHCLDPNNEEGGHV

Target KGWAFDDGNDVWMGRINETSRLGYETFKVIEGWSNPKSKLQINRQVIVDRGDRSGYSGIFSVEGKSCINRCFYVELIRG
 4gzs.1.C KGWAFDDGNDVWMGRINETSRLGYETFKVIEGWSNPKSKLQINRQVIVDRGNRSGYSGIFSVEGKSCINRCFYVELIRG

Target RKEETEVLWTSNSILLFCGTSPTYGTGSPDAADLNLMPI
 4gzs.1.C RKEETEVLWTSNSIVVFCGTSPTYGTGSPDGADLNLMPI

Target AGGDIWVTREPYVSCDPKGNQFALGQGTTLKSGHSNNTVRDRTPYRLLMNELGVPFHLGKQVCIWSSSSCHDGKAW
 4gzs.1.D AGGDIWVTREPYVSCDPKCYQFALGQGTTLNNVHSNNTVRGRTPYRLLMNELGVPFHLGKQVCIWSSSSCHDGKAW

Target LHCITGDDKNATASFIYNGRLVDSVVSWSKEVLRQTQESECVCTVVMTDGSASGKADTKILFIEEGKIVHTSTLS
 4gzs.1.D LHCITGDDKNATASFIYNGRLVDSVVSWSKEILRTQESECVCTVVMTDGSASGKADTKILFIEEGKIVHTSTLS

Target GSAQHVEECSCYPYRPGVRCVCRDNWKGSNRPVVDINIKDHSIVSSVYVCSGLVGDTPRKNDSSSSGHCLNPNNEEGGHV
 4gzs.1.D GSAQHVEECSCYPYRPGVRCVCRDNWKGSNRPVVDINIKDHSIVSSVYVCSGLVGDTPRKNDSSSSHCLDPNNEEGGHV

Target KGWAFDDGNDVWMGRINETSRLGYETFKVIEGWSNPKSKLQINRQVIVDRGDRSGYSGIFSVEGKSCINRCFYVELIRG
 4gzs.1.D KGWAFDDGNDVWMGRINETSRLGYETFKVIEGWSNPKSKLQINRQVIVDRGNRSGYSGIFSVEGKSCINRCFYVELIRG

Target RKEETEVLWTSNSILLFCGTSPTYGTGSPDAADLNLMPI
 4gzs.1.D RKEETEVLWTSNSIVVFCGTSPTYGTGSPDGADLNLMPI

Materials and Methods

Template Search

Template search with BLAST and HHblits has been performed against the SWISS-MODEL template library (SMTL, last update: 2021-03-25, last included PDB release: 2021-03-19).

The target sequence was searched with BLAST against the primary amino acid sequence contained in the SMTL. A total of 103 templates were found.

An initial HHblits profile has been built using the procedure outlined in (Steinegger et al.), followed by 1 iteration of HHblits against Uniclust30 (Mirdita, von den Driesch et al.). The obtained profile has then be searched against all profiles of the SMTL. A total of 104 templates were found.

Template Selection

For each identified template, the template's quality has been predicted from features of the target-template alignment. The templates with the highest quality have then been selected for model building.

Model Building

Models are built based on the target-template alignment using ProMod3 (Studer et al.). Coordinates which are conserved between the target and the template are copied from the template to the model. Insertions and deletions are remodelled using a fragment library. Side chains are then rebuilt. Finally, the geometry of the resulting model is regularized by using a force field.

Model Quality Estimation

The global and per-residue model quality has been assessed using the QMEAN scoring function (Studer et al.).

Ligand Modelling

Ligands present in the template structure are transferred by homology to the model when the following criteria are met: (a) The ligands are annotated as biologically relevant in the template library, (b) the ligand is in contact with the model, (c) the ligand is not clashing with the protein, (d) the residues in contact with the ligand are conserved between the target and the template. If any of these four criteria is not satisfied, a certain ligand will not be included in the model. The model summary includes information on why and which ligand has not been included.

Oligomeric State Conservation

The quaternary structure annotation of the template is used to model the target sequence in its oligomeric form. The method (Bertoni et al.) is based on a supervised machine learning algorithm, Support Vector Machines (SVM), which combines interface conservation, structural clustering, and other template features to provide a quaternary structure quality estimate (QSQE). The QSQE score is a number between 0 and 1, reflecting the expected accuracy of the interchain contacts for a model built based a given alignment and template. Higher numbers indicate higher reliability. This complements the GMQE score which estimates the accuracy of the tertiary structure of the resulting model.

References

- **BLAST**
Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L. BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421-430 (2009). [doi>](#)
- **HHblits**
Steinegger, M., Meier, M., Mirdita, M., Vöhringer, H., Haunsberger, S. J., Söding, J. HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* 20, 473 (2019). [doi>](#)
- **Uniclust30**
Mirdita, M., von den Driesch, L., Galiez, C., Martin, M.J., Söding, J., Steinegger, M. Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Research* 45, D170–D176 (2016). [doi>](#)

Table T1:

Primary amino acid sequence for which templates were searched and models were built.

```
AGGDIWVTREPYVSCDPDKGNQFALGQGTTLKSGHSNNTVRDRTPYRLLMNLGVPFHLGTRKQVCIAWSSSSCHDGKAWLHVCITGDDKNATASFIYNG
RLVDSVVSWSKEVLRTESECVCTVVMTDGSASGKADTKILFIEEGKIVHTSTLSGSAQHVEECSCYPRYPGVRCVCRDNWKGSNRPPIVDINIKD
HSIVSSYVCSGLVGDTPFRKNDSSSGHCLPNNEEGHGVKGFDDGNDVWMGRRTINETSRLLGYETPKVIEGWSNPKSKLQINRQVIVDRGDRSGYSGI
FSVEGKSCINRCFYVELIRGRKEETEVLWTSNSILLFCGTSYGTGSGWPAADLNLMPFI
```

Table T2:

| Template | Seq Identity | Oligo-state | QSQE | Found by | Method | Resolution | Seq Similarity | Coverage | Description |
|----------|--------------|---------------|------|----------|--------|------------|----------------|----------|---------------|
| 4gzs.1.A | 96.39 | homo-tetramer | 1.00 | HHblits | X-ray | 2.35Å | 0.62 | 1.00 | Neuraminidase |
| 4gzo.1.A | 96.67 | homo-tetramer | 1.00 | HHblits | X-ray | 2.60Å | 0.62 | 1.00 | neuraminidase |

| Template | Seq Identity | Oligo-state | QSQE | Found by | Method | Resolution | Seq Similarity | Coverage | Description |
|----------|--------------|---------------|------|----------|--------|------------|----------------|----------|--------------------------------------|
| 3tia.1.A | 82.78 | homo-tetramer | 1.00 | HHblits | X-ray | 1.80Å | 0.57 | 1.00 | Neuraminidase |
| 4k1j.1.A | 82.78 | homo-tetramer | 1.00 | HHblits | X-ray | 2.20Å | 0.57 | 1.00 | Neuraminidase |
| 4h52.1.A | 82.78 | homo-tetramer | 1.00 | HHblits | X-ray | 1.80Å | 0.57 | 1.00 | Neuraminidase |
| 5huk.1.A | 79.72 | homo-tetramer | 1.00 | HHblits | X-ray | 2.45Å | 0.57 | 1.00 | Neuraminidase |
| 6n4d.1.A | 79.17 | homo-tetramer | 1.00 | HHblits | X-ray | 1.80Å | 0.56 | 1.00 | Neuraminidase |
| 2bat.1.A | 83.06 | homo-tetramer | 1.00 | HHblits | X-ray | 2.00Å | 0.57 | 1.00 | NEURAMINIDASE N2 |
| 1ivg.1.A | 82.78 | homo-tetramer | 1.00 | HHblits | X-ray | 1.90Å | 0.57 | 1.00 | INFLUENZA A SUBTYPE N2 NEURAMINIDASE |
| 1ive.1.A | 82.78 | homo-tetramer | 1.00 | HHblits | X-ray | 2.40Å | 0.57 | 1.00 | INFLUENZA A SUBTYPE N2 NEURAMINIDASE |
| 1inh.1.A | 82.78 | homo-tetramer | 1.00 | HHblits | X-ray | 2.40Å | 0.57 | 1.00 | INFLUENZA A SUBTYPE N2 NEURAMINIDASE |
| 1ing.1.A | 82.78 | homo-tetramer | 1.00 | HHblits | X-ray | 2.40Å | 0.57 | 1.00 | INFLUENZA A SUBTYPE N2 NEURAMINIDASE |
| 6br5.1.A | 95.83 | monomer | - | HHblits | X-ray | 2.04Å | 0.62 | 1.00 | Neuraminidase |
| 2aep.1.A | 90.56 | homo-tetramer | 1.00 | HHblits | X-ray | 2.10Å | 0.60 | 1.00 | neuraminidase |
| 6n6b.1.A | 89.17 | homo-tetramer | 1.00 | HHblits | X-ray | 2.30Å | 0.60 | 1.00 | Neuraminidase |
| 4mwx.1.B | 47.62 | homo-tetramer | 0.86 | HHblits | X-ray | 1.80Å | 0.44 | 0.99 | Neuraminidase |
| 4mwj.1.A | 47.90 | homo-tetramer | 0.85 | HHblits | X-ray | 1.80Å | 0.44 | 0.99 | Neuraminidase |
| 5nz4.1.A | 46.05 | homo-tetramer | 0.88 | HHblits | X-ray | 1.36Å | 0.43 | 0.98 | neuraminidase |
| 2b8h.1.A | 47.62 | homo-tetramer | 0.85 | HHblits | X-ray | 2.20Å | 0.44 | 0.99 | Neuraminidase |
| 4m3m.1.A | 44.94 | homo-tetramer | 0.87 | HHblits | X-ray | 2.10Å | 0.43 | 0.99 | Neuraminidase |
| 5l14.1.A | 47.90 | homo-tetramer | 0.84 | HHblits | X-ray | 1.90Å | 0.44 | 0.99 | Neuraminidase |
| 4mwj.1.A | 49.72 | homo-tetramer | 0.84 | BLAST | X-ray | 1.80Å | 0.45 | 0.98 | Neuraminidase |
| 2b8h.1.A | 49.43 | homo-tetramer | 0.84 | BLAST | X-ray | 2.20Å | 0.45 | 0.98 | Neuraminidase |
| 1nna.1.D | 49.15 | homo-tetramer | 0.84 | BLAST | X-ray | 2.50Å | 0.44 | 0.98 | NEURAMINIDASE |
| 4mju.1.A | 45.51 | homo-tetramer | 0.84 | HHblits | X-ray | 2.35Å | 0.43 | 0.99 | Neuraminidase |
| 4nn9.1.A | 49.15 | homo-tetramer | 0.83 | BLAST | X-ray | 2.30Å | 0.44 | 0.98 | NEURAMINIDASE N9 |
| 1mwe.1.C | 49.15 | homo-tetramer | 0.83 | BLAST | X-ray | 1.70Å | 0.44 | 0.98 | NEURAMINIDASE |
| 6pzw.1.A | 47.90 | homo-tetramer | 0.87 | HHblits | EM | NA | 0.44 | 0.99 | Neuraminidase |
| 6pze.1.A | 47.90 | homo-tetramer | 0.83 | HHblits | X-ray | 2.30Å | 0.44 | 0.99 | Neuraminidase |
| 6pze.1.A | 49.72 | homo-tetramer | 0.83 | BLAST | X-ray | 2.30Å | 0.45 | 0.98 | Neuraminidase |

| Template | Seq Identity | Oligo-state | QSQE | Found by | Method | Resolution | Seq Similarity | Coverage | Description |
|----------|--------------|---------------|------|----------|--------|------------|----------------|----------|--------------------------------------|
| 6pzf.1.A | 47.90 | homo-tetramer | 0.84 | HHblits | X-ray | 2.80Å | 0.44 | 0.99 | Neuraminidase |
| 1nca.2.C | 49.15 | homo-tetramer | 0.84 | BLAST | X-ray | 2.50Å | 0.44 | 0.98 | INFLUENZA A SUBTYPE N9 NEURAMINIDASE |
| 1ncd.1.A | 47.90 | homo-tetramer | 0.84 | HHblits | X-ray | 2.90Å | 0.44 | 0.99 | INFLUENZA A SUBTYPE N9 NEURAMINIDASE |
| 1nmb.1.A | 47.90 | homo-tetramer | 0.81 | HHblits | X-ray | 2.20Å | 0.44 | 0.99 | N9 NEURAMINIDASE |
| 4d8s.1.A | 45.79 | homo-tetramer | 0.88 | HHblits | X-ray | 2.40Å | 0.43 | 0.99 | Neuraminidase |
| 4gb1.1.A | 45.79 | homo-tetramer | 0.85 | HHblits | X-ray | 2.62Å | 0.43 | 0.99 | Neuraminidase |
| 4b7q.1.A | 46.18 | homo-tetramer | 0.85 | HHblits | X-ray | 2.73Å | 0.43 | 0.98 | NEURAMINIDASE |
| 3o9j.1.A | 45.92 | homo-tetramer | 0.84 | HHblits | X-ray | 2.00Å | 0.43 | 0.99 | Neuraminidase |
| 4mjv.1.A | 45.79 | homo-tetramer | 0.83 | HHblits | X-ray | 2.65Å | 0.43 | 0.99 | Neuraminidase |
| 2ht7.1.A | 45.51 | homo-tetramer | 0.81 | HHblits | X-ray | 2.60Å | 0.43 | 0.99 | Neuraminidase |
| 4ks5.1.A | 44.94 | homo-tetramer | 0.77 | HHblits | X-ray | 2.70Å | 0.43 | 0.99 | Neuraminidase |
| 6crd.1.D | 49.15 | homo-tetramer | 0.75 | BLAST | X-ray | 2.57Å | 0.44 | 0.98 | Tetrabrachion,Neuraminidase |
| 6crd.1.B | 49.15 | homo-tetramer | 0.75 | BLAST | X-ray | 2.57Å | 0.44 | 0.98 | Tetrabrachion,Neuraminidase |
| 6crd.1.C | 49.15 | homo-tetramer | 0.75 | BLAST | X-ray | 2.57Å | 0.44 | 0.98 | Tetrabrachion,Neuraminidase |
| 1ncc.1.A | 49.15 | homo-tetramer | 0.86 | BLAST | X-ray | 2.50Å | 0.44 | 0.98 | INFLUENZA A SUBTYPE N9 NEURAMINIDASE |
| 1ncd.1.A | 49.43 | homo-tetramer | 0.83 | BLAST | X-ray | 2.90Å | 0.45 | 0.98 | INFLUENZA A SUBTYPE N9 NEURAMINIDASE |
| 1nmb.1.A | 49.43 | homo-tetramer | 0.81 | BLAST | X-ray | 2.20Å | 0.45 | 0.98 | N9 NEURAMINIDASE |
| 6lxx.1.A | 48.13 | homo-tetramer | 0.84 | BLAST | X-ray | 3.61Å | 0.44 | 0.96 | Neuraminidase |
| 1nma.1.A | 47.62 | homo-tetramer | 0.68 | HHblits | X-ray | 3.00Å | 0.44 | 0.99 | N9 NEURAMINIDASE |
| 1nma.1.A | 49.43 | homo-tetramer | 0.65 | BLAST | X-ray | 3.00Å | 0.45 | 0.98 | N9 NEURAMINIDASE |

The table above shows the top 50 filtered templates. A further 83 templates were found which were considered to be less suitable for modelling than the filtered list.

1a4g.1.A, 1b9v.1.D, 1inf.1.A, 1iny.1.A, 1l7g.1.A, 1mwe.1.C, 1nca.2.C, 1ncb.1.A, 1ncc.1.A, 1nna.1.D, 1vcj.1.A, 2ht7.1.A, 2htv.1.A, 2hu0.1.A, 2qwd.1.A, 3b7e.1.A, 3cl0.1.A, 3cl2.1.A, 3cye.1.A, 3f14.1.A, 3k36.1.A, 3k38.1.A, 3nn9.1.A, 3o9j.1.A, 3sal.1.A, 4b7m.1.C, 4b7q.1.A, 4b7r.1.A, 4cpl.1.A, 4cpo.1.A, 4d8s.1.A, 4fvk.1.A, 4gb1.1.A, 4gdi.1.A, 4gdi.1.B, 4gdj.1.A, 4gez.1.A, 4h53.1.A, 4h53.1.D, 4hzv.1.A, 4hzy.1.A, 4hzz.1.A, 4k3y.1.A, 4k3y.1.C, 4ks5.1.A, 4m3m.1.A, 4mc7.1.A, 4mju.1.A, 4mjv.1.A, 4mwx.1.B, 4nn9.1.A, 4qn3.1.A, 4qn4.1.A, 4qnp.1.A, 4wa5.1.D, 5hug.1.A, 5hum.1.A, 5hun.1.A, 5l14.1.A, 5nn9.1.A, 5nwe.1.A, 5nz4.1.A, 5nze.1.A, 5nzf.1.A, 5nzn.1.A, 6crd.1.A, 6crd.1.B, 6crd.1.C, 6crd.1.D, 6d96.1.A, 6g01.2.D, 6hfy.1.D, 6hp0.1.C, 6lxi.1.A, 6lxx.1.A, 6nn9.1.A, 6pzf.1.A, 6pzw.1.A, 6q20.1.A, 6q23.1.A, 6v4n.1.A, 6v4o.1.A, 7cm1.1.A

Table 3-1. Potential “original” epitopes targeted by αEp9 Abs

| Construct # | Pathogen | Target protein | Accession No. | Residues | Epitope sequences | Similarity % |
|-----------------------------------|---|---|-----------------------|------------------|--|--------------|
| Phage-displayed constructs | | | | | | |
| 1 | SARS-COV-2 (Ep9) | Nucleocapsid | QOX29443.1 | 152-172 | ANNAIVLQLPQGTTLPKGFY | - |
| 2 | SARS-COV-1 | Nucleocapsid | YP_009825061.1 | 153-173 | NNNAATVLLQLPQGTTLPKGFY | 90.5 |
| 3 | MERS | Nucleocapsid | YP_009047211.1 | 141-151 | NNDSAIVTFQFAPGKTLKPNFH | 66.7 |
| 4 | Human coronavirus HKU1 | Nucleocapsid | YP_173242.1 | 166-186 | TTQEAIPTFRFPPTLPGQYY | 57.1 |
| 5 | Human coronavirus NL63 | Nucleocapsid | YP_003771.1 | 119-136 | NQKPLEPKFSIALPPELS | 13.8 |
| 6 | Human coronavirus OC43 | Nucleocapsid | YP_009555245.1 | 167-187 | SSDEAIPTFRFPPTLPGQYY | 71.4 |
| 7 | Human coronavirus 229E | Nucleocapsid | AGW80953.1 | 122-138 | SEPEIPHFNGKLPNGVT | 21.4 |
| 8 | Human adenovirus 61 | Hexon | AQQ81927.1 | 123-164 | ANNAATPQVVFYTEDVNLMPDTHLVFKPAVPNGTIASESL | 17.6 |
| 9 | Human mastadenovirus E | PVIII | YP_068038.1 | 76-114 | AALVYQEIPQPTTVLLPRDAQAEVQLTNSGVQLAGGATL | 31 |
| 10 | Influenza A virus (A/Utah/40/2017) | PB2 polymerase | AVH77902.1 | 225-244 | GSVYIEVLHLTQGTWEQMY | 41.7 |
| 11 | Influenza A virus (EpNeu) (A/Para/128982-IEC/2014(H3N2)) | Neuraminidase, partial | AIX95025.1 | 34-46 | ALGGQTLSKGHS | 38.1 |
| 12 | Influenza B virus (B/California/88/2019) | Neuraminidase | QIA55965.1 | 67-79 | ATKGVLLPEPE | 28.6 |
| 13 | Influenza C virus (C/Singapore/DSO-070193/2006) | Polymerase PB1, partial | AFV68302.1 | 119-145 | AATALQLTVDAIKETGPFKTTMLEY | 34.4 |
| 14 | Human respiratory syncytial virus A | Fusion protein | ASU44644.1 | 87-100 | NNAVTELQLLMQST | 38.1 |
| 15 | Human respiratory syncytial virus A | Attachment glycoprotein | ART28426.1 | 106-116 | GTTPQSTTIPA | 28.6 |
| 16 | Human metapneumovirus | Nucleoprotein, partial | ABO15448.1 | 11-33 | TTTAVTPSSLPQEITLGCILEY | 34.8 |
| 17 | Human metapneumovirus | Attachment glycoprotein, partial | AEW90340.1 | 57-72 | PQQTDKHTALPKSIY | 30.8 |
| 18 | Human betaherpesvirus 6A | Immediate Early protein 2 | AGJ52064.1 | 1396-1422 | AATPIDFVAVKTCNKYAKDNPKIVL | 10 |
| 19 | Verrucomicrobia bacterium | NADH-quinone oxidoreductase (NOX) | PYJ45937.1 | 76-89 | AGVVLQPLQGTTL | 57.1 |
| 20 | Clostridium butyricum | bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase (MTHFD2) | MBE6063617.1 | 94-104 | IILQLPLPKKF | 47.6 |
| 21 | Fusobacterium mortiferum | Type II secretion protein | WP_118233983.1 | 77-99 | VENGAIVLQYDKEIYGLTENFF | 48 |
| 22 | Fusobacterium mortiferum | Autotransporter outer membrane protein | WP_005886362.1 | 449-460 | NGAIVGDLVQGT | 38.1 |
| 23 | Influenza A virus (A/swine/Missouri/A01774733/2016(H1N2) or A/Para/129501-IEC/2014(H3N2)) | Neuraminidase | ANK78229.1 / AIX95013 | 33-145/ 24-36 | ALGGQTLSNGHS | 92.3 |
| 24 | Influenza A (A/swine/Minnesota/A01394278/2013(H3N2)) | Neuraminidase | AHA57095.1 | 134-146 | ALGGQTLLNNGHS | 92.3 |
| 25 | Influenza A virus (A/California/04/2009(H1N1)) | Neuraminidase | AJI76397.1 | 75-89 | TFFLTQGALLNDKHS | 46.7 |
| 26 | Influenza A virus (A/California/111/2015(H3N2)) | Neuraminidase | ANM97445.1 | 133-145 | ALGGQTLLNVHS | 84.6 |
| 27 | Influenza A virus (A/mallard/California/1156/2010(H4N6)) | Neuraminidase | AEK50939.1 | 133-142 | ALSQGTTLKG | 84.6 |
| 28 | Influenza A virus (A/California/33/2011(H3N2)) | Neuraminidase | AGL06761.1 | 133-145 | ALGGQTLLSNVHS | 84.6 |
| 29 | Influenza A virus (EpNeu Pred) (A/Para/128982-IEC/2014(H3N2)) | Neuraminidase, partial | AIX95025.1 | 34-56 | ALGGQTLSKGHSNNTVRDRTPY | - |
| eGFP-fusion constructs | | | | | | |
| 30 | Influenza A virus (EpNeu) (A/Para/128982-IEC/2014(H3N2)) | Neuraminidase | AIX95025.1 | 34-46 | ALGGQTLSKGHS | - |
| 31 | Influenza A virus (EpNeu Pred) (A/Para/128982-IEC/2014(H3N2)) | Neuraminidase | AIX95025.1 | 34-56 | ALGGQTLSKGHSNNTVRDRTPY | - |

Table 3-2. Primers used to subclone potential original epitopes

| Construct # | Primer names: | Oligonucleotide sequence |
|-----------------------------------|---------------|---|
| Phage-displayed constructs | | |
| 1 | primer_1_F | tcaagggactacctgccaaggggttctatGGTGGAGGATCCGGGAGC |
| | primer_1_R | ggtaattgaacacgattgcagcgttattagcTCCACTTCCTTTATCATCGTCATCTTTATAATC |
| 2 | primer_2_F | acttccacaggggaacgacactgccaaggggalltGGTGGAGGATCCGGGAGC |
| | primer_2_R | tgacgtacagtggcagcattgtattccactccTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 3 | primer_3_F | tccgggtacaaagttaccaagaactccacGGTGGAGGATCCGGGAGC |
| | primer_3_R | gcaaatgaglaactatcgctgaatcattgtTCCACTTCCTTTATCATCGTCATCTTTATAATC |
| 4 | primer_4_F | tcccggaacttttaccccaaggatactatGGTGGAGGATCCGGGAGC |
| | primer_4_R | gggaatctagtggaatcgctcctcgagtagiTCCACTTCCTTTATCATCGTCATCTTTATAATC |
| 5 | primer_5_F | agtattgccttgccacctgattatctGGTGGAGGATCCGGGAGC |
| | primer_5_R | aaattcgggtcaagcggctttgattTCCACTTCCTTTATCATCGTCATCTTTATAATC |
| 6 | primer_6_F | tccgggtactgtttgctcagggttactacGGTGGAGGATCCGGGAGC |
| | primer_6_R | ggaaaccgcttgggatggctcatcagagctTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 7 | primer_7_F | ccgaaattccgcaacggcgttactGGTGGAGGATCCGGGAGC |
| | primer_7_R | ttaaagtggggatctctggctcgtTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 8 | primer_8_F | TTCTAAATTCACATCCTCAGTATAGAACACGACTTGTGGGGTGTGCTGCATTGTTGGCTTTATCATCGTCATCTTTATAATCAACCAATGC |
| | primer_8_R | ATGCCAGATACTCATTGGTCTTCAAACCTGCGGTCCCGAATGGCAGCAGATTGCTAGTGAATCTCTCCTTGGTgggagatccgggagc |
| 9 | primer_9_F | ACGCGGGAGCAGCACTGTAGTCGGCTGAGGGATTTCTGGTACACCAGGGCGGCTTTATCATCGTCATCTTTATAATCAACCAATGC |
| | primer_9_R | GACGCGCAAGCCGAGTCCAGTTGACAAACAGCGGTGACAATTGGCTGGCGGTGCCACTCTGggtggagatccgggagc |
| 10 | primer_10_F | acacaagcactctcgggagcaaatgtGGTGGAGGATCCGGGAGC |
| | primer_10_R | gaggtgcagcactcaatagactgacctTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 11 | primer_11_F | gcttctaagggccacagcGGTGGAGGATCCGGGAGC |
| | primer_11_R | gtggtgccttggccaatgcTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 12 | primer_12_F | gctgctccggagccagagGGTGGAGGATCCGGGAGC |
| | primer_12_R | agcacgacgcttctgtagcTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 13 | primer_13_F | gactgaaggtcctttaagggcagcacaaltgtggaalatGGTGGAGGATCCGGGAGC |
| | primer_13_R | tcctaatgcatcaccgtaagttggagagcctgctgcTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 14 | primer_14_F | caactcctcatcagcttacaGGTGGAGGATCCGGGAGC |
| | primer_14_R | aagttctgtgacagcattgtTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 15 | primer_15_F | gacaacctccctgcaGGTGGAGGATCCGGGAGC |
| | primer_15_R | gattgagcgttagtcccTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 16 | primer_16_F | ggaaatcaactctcttgctggcgaactctatGGTGGAGGATCCGGGAGC |
| | primer_16_R | tgccgagcgttagatggcgaactgcccagtagtTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 17 | primer_17_F | acggcactgccgaagtccatctacGGTGGAGGATCCGGGAGC |
| | primer_17_R | atgcttgcagtggtctgttgggTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 18 | primer_18_F | caacaatacgttaaggacaatccgaaagatcgtgctGGTGGAGGATCCGGGAGC |
| | primer_18_R | caggtctgacagctccgacgaagtgcagggcgtagccgcTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 19 | primer_19_F | tccacaagccaccacctGGTGGAGGATCCGGGAGC |
| | primer_19_R | agctgcagcagcagccggcTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 20 | primer_20_F | gttaccaagaatcGGTGGAGGATCCGGGAGC |
| | primer_20_R | ggaagctgcaaaataatTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 21 | primer_21_F | agagattacctggggctgacagagaacttttGGTGGAGGATCCGGGAGC |
| | primer_21_R | ttgtcgtactgtaatacaattgcccgttctccacTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 22 | primer_22_F | gatttagtacaggtaccGGTGGAGGATCCGGGAGC |
| | primer_22_R | gcclacaatggcaccattTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 23 | primer_23_F | CGCTTTCTAAIGGCCACAGCG |
| | primer_23_R | TGGTGCCTTGCCCAATG |
| 24 | primer_24_F | CACCACGCTTaacaatGGCCACAGCG |
| | primer_24_R | CCTTGGCCCAATGCTTTAT |
| 25 | primer_25_F | tccttgaacgacaagcattcaGGTGGAGGATCCGGGAGC |
| | primer_25_R | gctccctgagtcagaagaatgtTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 26 | primer_26_F | attaataacgttcatagtGGTGGAGGATCCGGGAGC |
| | primer_26_R | gtaglacctggccaagcTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 27 | primer_27_F | tgcaaacgtagcattgatcgttcccatcGGTGGAGGATCCGGGAGC |
| | primer_27_R | tgacgaccttcagtgctgcttgagataacgcTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 28 | primer_28_F | ttgtcgaatgtactccGGTGGAGGATCCGGGAGC |
| | primer_28_R | gtggtaccctggccaagagcTTTATCATCGTCATCTTTATAATCAACCAATGC |
| 29 | primer_29_F | ctcgaacaacacgttcgacgacctcctgacGGTGGAGGATCCGGGAGC |
| | primer_29_R | tgccctttagacagagttgcccctggccaagcTTTATCATCGTCATCTTTATAATCAACCAATGC |
| eGFP-fusion constructs | | |
| 30 | primer_30_F | gcttctaagggccacagcAGCGGAAGTGGAGATTATAAAGATGAC |
| | primer_30_R | gtggtgcttggccaatgcGGAGCTCCCGGATCCTCC |
| 31 | primer_31_F | ctcgaacaataaccgtccgcatcggactccgtacAGCGGAAGTGGAGATTATAAAGATGAC |
| | primer_31_R | tgaccttctcaagtagtaccctggccaagcGGAGCTCCCGGATCCTCC |

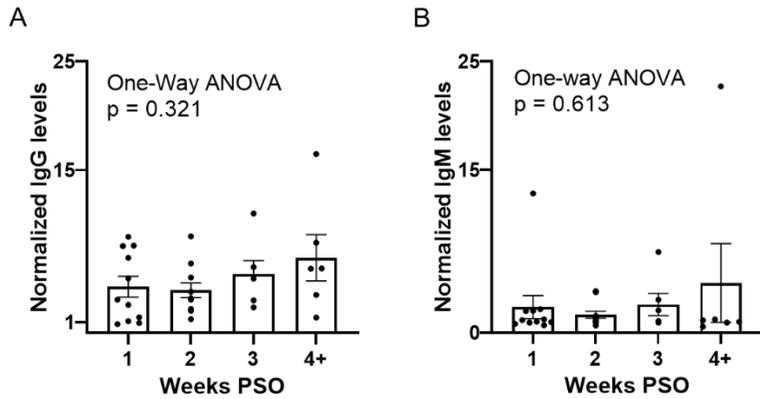


Figure 3-3. Early upregulation of α Ep9 IgGs. ELISA of α Ep9 **A.** IgG and **B.** IgM levels in α Ep9(+) patients (n = 34) from plasma collected at the indicated time periods post-symptom onset (PSO). Statistical analysis was conducted using one-way ANOVA, *ad hoc* Tukey test. Error bars represent SEM.

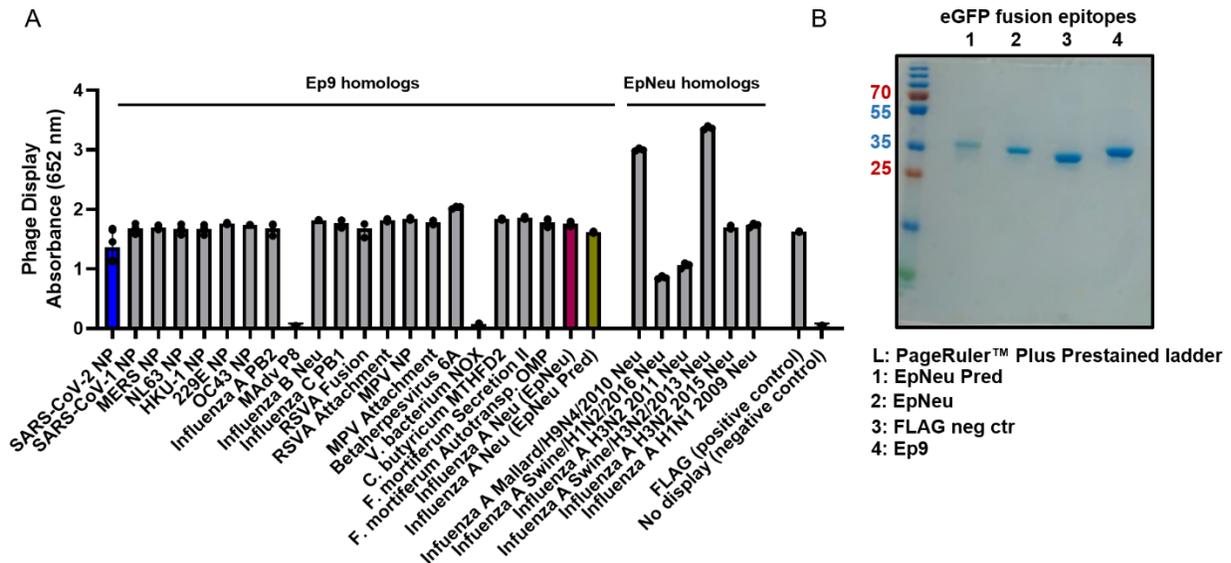


Figure 3-4. Expression of phage-displayed and eGFP-fused potential AIM epitopes. **A.** ELISA demonstrating the display of N-terminal FLAG-tagged potential epitopes fused to the N-terminus of the P8 coat protein. Immobilized α FLAG Abs in microtiter wells bind the displayed FLAG-tag and epitope, and binding is detected with α M-13-HRP Abs as usual. Phage with no epitope displayed provide the negative control. Epitopes for mastadenovirus protein (mAdV) P8 and *V. bacterium* NADH oxidoreductase (NOX) did not display. Error bars represent SD values. **B.** 10% SDS-PAGE gel stained with Coomassie Blue shows His-tag affinity-purified and buffer-exchanged eGFP-fused epitopes, EpPred, EpNeu, FLAG negative control and Ep9.

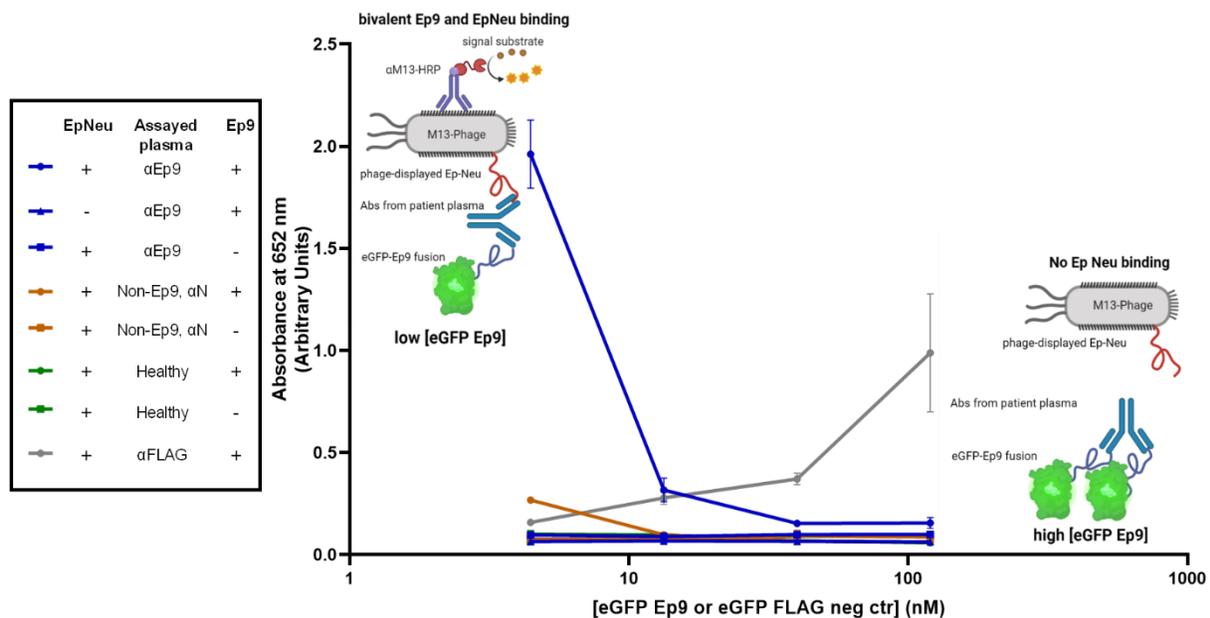


Figure 3-5. Optimization of assay to determine cross-reactivity of α Ep9 Ab to Ep9 and EpNeu. Sandwich ELISA testing the binding of Abs from the pooled plasma of five α Ep9(+) patients, five α Ep9(-) patients with other α NP Abs and healthy individuals. This experiment examines bivalent binding to various doses of immobilized eGFP-fused Ep9 epitope (120, 40, 13 and 4 nM) and phage-displayed EpNeu in solution. The data shows that Abs from α Ep9(+) patients, but not α Ep9(-) or healthy individuals, bivalently bind both EpNeu and Ep9. The positive control (α FLAG 1:2000 fold dilution) at 100 nM eGFP demonstrates concentrations appropriate for bivalent binding to immobilized and in-solution tags. The schematic diagram illustrates the binding observed for bivalence in α Ep9 Abs, where the antibody bridges plate-bound eGFP at its high concentrations. Therefore, Figure 3-2 in the main text uses 4 nM of eGFP Ep9 coated on the plate, and the FLAG positive control uses eGFP at 100 nM. Error bars represent SD.

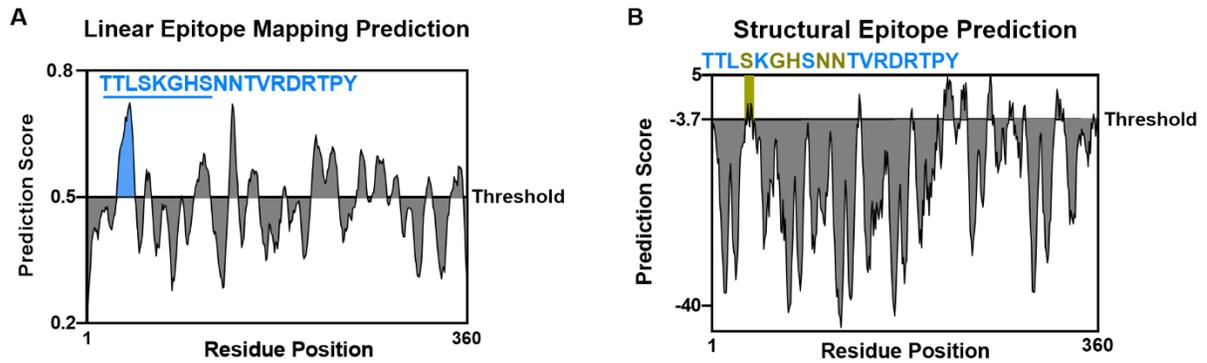


Figure 3-6. Linear and structural epitope mapping prediction of Influenza A H3N2 Neuraminidase.

A. Linear epitope mapping prediction of the Flu A 2014 H3N2 using Bepipred 2.0¹³ demonstrates high prediction scores in a region spanning 18 residues, which includes eight residues from EpNeu (underlined). The additional 10 predicted residues were included as part of an extended epitope termed EpPred. **B.** Structural epitope mapping, using Discotope 2.0¹⁶, of the modelled neuraminidase protein from Flu A 2014 H3N2 (SWISS-model²¹), predicts an epitope of five residues. These were captured by EpPred, including three found in EpNeu.

3.6 References

1. Fierz, W. & Walz, B. Antibody Dependent Enhancement Due to Original Antigenic Sin and the Development of SARS. *Front. Immunol.* **11**, 1120 (2020).
2. Brown, E. L. & Essigmann, H. T. Original Antigenic Sin: the Downside of Immunological Memory and Implications for COVID-19. *mSphere* **6**, (2021).
3. Kohler, H. & Nara, P. A Novel Hypothesis for Original Antigenic Sin in the Severe Disease of SARS-CoV-2 Infection. *Monoclonal Antibodies in Immunodiagnosis and Immunotherapy* vol. 39 107–111 (2020).
4. Ng, K. W. *et al.* Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. *Science (80-.)*. **370**, 1339–1343 (2020).
5. Anderson, E. M. *et al.* Seasonal human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with protection. *medRxiv* vol. 7 2020.11.06.20227215 (2020).
6. Focosi, D. *et al.* Previous humoral immunity to the endemic seasonal alphacoronaviruses NL63 and 229E is associated with worse clinical outcome in COVID-19 and suggests original antigenic sin. *Life* **11**, (2021).
7. Shrock, E. *et al.* Viral epitope profiling of COVID-19 patients reveals cross-reactivity and correlates of severity. *Science*. eabd4250 (2020).
8. Wang, L. *et al.* Serological Responses to Human Virome Define Clinical Outcomes of Italian Patients Infected with SARS-CoV-2. *medRxiv Prepr. Serv. Heal. Sci.* (2020).
9. Sen, S. R. *et al.* Predicting COVID-19 Severity with a Specific Nucleocapsid Antibody plus Disease Risk Factor Score. *mSphere* **6**, e00203-21 (2021).
10. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* vol. 25 3389–3402 (1997).
11. Madej, T. *et al.* MMDB and VAST+: Tracking structural similarities between macromolecular complexes. *Nucleic Acids Res.* **42**, (2014).
12. World Health Organization. WHO | Cumulative Number of Reported Probable Cases of Severe Acute Respiratory Syndrome (SARS). *WHO* 9–11 <http://www.who.int/csr/sars/country/en/> (2013).
13. Jespersen, M. C., Peters, B., Nielsen, M. & Marcatili, P. BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* **45**, W24–W29 (2017).
14. Gentles, L. E., Wan, H., Eichelberger, M. C. & Bloom, J. D. Antibody neutralization of an influenza virus that uses neuraminidase for receptor binding. *Viruses* **12**, 2020.05.08.084954 (2020).
15. Studer, G. *et al.* ProMod3 - A versatile homology modelling toolbox. *PLoS Comput. Biol.* **17**, (2021).
16. Kringelum, J. V., Lundegaard, C., Lund, O. & Nielsen, M. Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking. *PLoS Comput. Biol.* **8**, (2012).
17. Lipničánová, S., Chmelová, D., Godány, A., Ondrejovič, M. & Miertuš, S. Purification of viral neuraminidase from inclusion bodies produced by recombinant *Escherichia coli*. *J. Biotechnol.* **316**, 27–34 (2020).
18. Xie, H. *et al.* H3N2 Mismatch of 2014-15 Northern Hemisphere Influenza Vaccines

- and Head-to-head Comparison between Human and Ferret Antisera derived Antigenic Maps. *Sci. Rep.* **5**, 1–10 (2015).
19. Flannery, B. *et al.* Enhanced Genetic Characterization of Influenza A(H3N2) Viruses and Vaccine Effectiveness by Genetic Group, 2014-2015. in *Journal of Infectious Diseases* vol. 214 1010–1019 (Oxford University Press, 2016).
 20. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).
 21. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
 22. Benchling, I. Benchling [Biology Software]. Retrieved from <https://benchling.com> (2017).
 23. Levin, A. M. Exploring the interaction between the protein kinase A catalytic subunit and caveolin-1 scaffolding domain with shotgun scanning, oligomer complementation, NMR, and docking. *Protein Sci.* **15**, 478–486 (2006).
 24. Pastorino, B., Touret, F., Gilles, M., de Lamballerie, X. & Charrel, R. N. Heat inactivation of different types of SARS-CoV-2 samples: What protocols for biosafety, molecular detection and serological diagnostics? *Viruses* **12**, 735–743 (2020).
 25. Mukaka, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.* **24**, 69–71 (2012).

CHAPTER 4

Development towards a real-time insulin biosensor for Type-1-Diabetes disease management

Author contributions: Sanjana R. Sen, Brian M. Miller, Gaetano Speciale, Madison H. Fletcher, Sudipta Majumdar, Gregory A. Weiss, and Elliot L. Botvinick designed research and analyzed data. Christian A. Totoiu, Katrin Safronyuk, Emily M. Crawley, Bjarne Vincent, Kim Weiss, Sahil Ahuja, Linette Madina, Jessica N. Pham, Brian Miller, Teodora Nedic performed research and analyzed data. Dat Nguyen, Danielle M. Behrens, Toni Wilkinson and Micah M. Lawrence designed experiments and performed FRET and biomedical engineering research and analyzed data.

4.1 Abstract

Therapeutic management of Type 1 Diabetes (T1D) disease can be challenging. The artificial pancreas (AP) automates disease management by delivering insulin with dosage adjusted in real-time in response to blood glucose levels. However, despite automation using advanced control algorithms, patients reliant on the AP risk hyper and hypoglycemia, which can result from discrepancies between the administered and the circulating insulin levels. As such, a direct continuous monitor for circulating insulin could tighten glucose regulation while improving safety and patient response to treatment. Thus, we aim to develop a real-time insulin biosensor for APs modelled upon the human insulin receptor (InR). Using a combination of rationally designed constructs and biopanning of phage display libraries, InR variants were selected for soluble and specific insulin binding. Next, to obtain a dose-dependent optical signal upon insulin binding, selected InR variants were further engineered with genetically-encoded and chemically conjugated FRET pairs. Various assays and constructs were developed to detect FRET-based signaling of insulin binding. These assays include single-protein and two-protein FRET constructs designed in mammalian, phage and *E. coli* platforms.

4.2 Introduction

In the United States, 1.25 million individuals currently live with Type 1 diabetes. Approximately 200,000 T1D individuals are under the age of 20, and a rise in prevalence of 2.7% per annum has been noted among youth (≤ 20 years)¹. T1D occurs as a result of autoimmune destruction of β -pancreatic islet cells²⁻⁴. Insulin, a peptide hormone produced by β -islet cells, regulates glucose uptake and metabolism in the body⁴. In the absence of insulin, T1D patients experience hyperglycemia in which blood glucose far exceeds normal physiological levels. Hyperglycemia can lead to systemic complications such as neurological, vascular, and renal disorders and diabetic ketoacidosis (DKA), which can be fatal^{3,4}. Treatment for T1D involves the administration of exogenous insulin in doses that need to be adjusted relative to blood glucose levels. Excess insulin administration can result in hypoglycemia, which in severe cases can cause coma and death^{3,4}. Therefore, the burdens of T1D disease management stem from consistent efforts to balance levels of blood glucose and insulin in the body. Despite the availability of treatment, less than one third of patients maintain optimal blood glucose levels⁵.

Recent development of the closed-loop, artificial pancreas (AP) provides an option for automated T1D disease management, and thus, may alleviate some of the burdens associated with T1D^{6,7}. The AP is an automated system that couples glucose monitoring to insulin dose regulation in patients with T1D. In the AP, a continuous glucose monitor (CGM) is connected to an insulin pump by a control algorithm to allow automated insulin dose adjustments responding in real-time to blood glucose levels⁶. The effectiveness of the closed-loop system has been evaluated against the Sensor Augmented Pump (SAP) technology, in

which insulin dosage is manually adjusted on an insulin pump based on readings of the CGM^{8,9}.

Since 2011, clinical evaluations of the closed-loop AP systems have transitioned from short term, inpatient trials to the more recent outpatient trials lasting several months in duration⁷. Thorough meta-analysis of 42 clinical trials comprising 1,042 participants, including children, adolescents and adults, report in favor of the AP system compared to the SAP. Their key findings demonstrate that T1D patients on the AP closed-loop system spend an average of 2.5 additional hours in normoglycemia, 2 hours less time in hyperglycemia and 20 minutes less time in hypoglycemia over a 24 h period compared to control groups¹⁰. In light of these improvements, the FDA approved the MiniMed® 670G (Medtronic) in 2016. This hybrid closed-loop system automates basal glucose regulation and administers insulin boluses during meals ^{6-8,10}.

While these technologies have made great strides towards improving glycemic control, T1D patients still struggle with challenges of periodic deviations from normoglycemia, primarily overnight when circulation changes or the infusion set malfunctions due to impeded flow. The resulting patient symptoms can go undetected. T1D patients in closed loop AP systems have been reported to spend an average of 30% of their time in hyperglycemia and about 2.9% in hypoglycemia⁹. The T1D research community thus aims to develop multi-feedback sensors to enhance accuracy and reliability of fully-automated closed-loop T1D management devices ^{7,11}. While the AP directly measures levels of blood glucose, it does not detect levels of insulin in the blood. Thus, discrepancies between the administered and the circulating levels of insulin can expose patients to risks of hyper and hypoglycemia. Precise and direct measurement of blood insulin levels in automated T1D

management devices, in place of indirect estimates of insulin based on glucose levels, could significantly improve blood glucose regulation in AP devices, ultimately leading to improved disease management.

Towards this effort, we aimed to develop a real-time insulin sensor to reduce glycemic deviations resulting from discrepancies in insulin administration and circulation. Insulin detection platforms that are currently available have been developed for *in cellulo* use for diagnostics or for assaying insulin production in β -islet cells for research purposes^{12,13}. Other versions of insulin sensors use antibodies and aptamers to detect and bind insulin with very high affinities and low dissociation rates (k_{off}). The sensors inability to release insulin quickly renders them impractical for use as real-time sensors¹⁴⁻¹⁹.

Real-time biosensing requires (1) fast, reversible binding and (2) a means to detect ligand-binding interaction. Thus, we model our insulin biosensor after the native insulin binder and blood glucose regulator, the insulin receptor (InR). In the body, the wildtype human insulin receptor (InR) binds and releases insulin in real-time, responding to rapid changes in blood glucose levels. The InR's extracellular domain (ECD) undergoes insulin binding-dependent conformational changes, which are well characterized^{20,21}. These conformational changes can be coupled to optical signals using Förster resonance energy transfer (FRET) measurements to detect insulin binding. Due to the InR's natural insulin binding characteristics, our current research objective is to engineer the InR to develop a real-time FRET-based insulin biosensor that is optically responsive to a wide range of insulin concentrations.

The InR is a large (320 kDa) multi-domain, transmembrane, receptor tyrosine kinase (RTK) that forms two symmetric dimers, in which each monomer is composed of two peptide

chains (α and β) (Figure 4-1A) ²⁰⁻²³. Insulin binds to the extracellular domain (ECD) which includes the entire α chain and partial (192 amino acids) beta chain^{20,21,23}. The α and β chains (interchain) as well the dimer (intra- α chain) are held together by disulfide bonds²⁴⁻²⁶. Since insulin binds only to the α chain on the ECD and the expression of the whole transmembrane receptor can prove challenging due to its size and hydrophobicity, we focus our investigations on the α chain of the ECD.

The structure of the InR ECD has been thoroughly investigated and structures of the apo (3.3 Å resolution) and holo (4.3 and 7.4 Å resolution) InR ECD receptor have provided insight into previously unknown interdomain interactions as well as binding-specific conformational changes^{20,21} (Figure 4-1B). The InR α -ECD contains 14 N-glycosylation sites. Mutational scanning has demonstrated that many of these sites can be altered without affecting cell-surface expression, processing, or insulin binding ^{22,27,28}. However, the presence of at least one N-glycosylation at major binding domains has been shown to be necessary for correct folding ^{22,27,28}.

Generally, insulin binding involves three regions of the ECD, the L1 domain of one monomer and the α -helix at the C-terminal (CT) of the second monomer come together to form the high affinity ($K_D \sim 6$ nM) S1 binding site and the FnIII-1 domain of the first monomer forms the lower affinity ($K_D \sim 400$ nM) S2 binding site, which together form a site with a binding affinity of $K_D \sim 200$ pM^{20,29}. As a symmetrical dimer, two binding pockets exist to allow the binding of two molecules of insulin²⁰. Other studies have shown that due to conformational changes upon the binding of a single insulin molecule, negative cooperativity has been reported for insulin binding at physiological concentrations. The affinity of InR for a second insulin is 600x lower than the first ^{20,22,23,29,30}.

In addition to insulin, the InR also binds the insulin-like growth factor (IGF)-I and II at much lower affinities (≤ 350 nM relative to 200 pM for insulin)²². Insulin is most closely related to IGF-I, with approximately 50% amino acid sequence homology³¹. In our investigations, we aimed to design a construct that preferentially binds insulin in affinities higher than to IGF-I, and responds to doses of insulin binding with a corresponding, detectable optical signal.

4.3 Results and Discussion

Mammalian cell expression of soluble insulin binding InR constructs

InR undergoes large conformational changes that make it ideal for FRET-based optical signaling. As observed in the overlaid apo and holo InR structures (Figure 4-1B), the two main changes that occur upon insulin binding involve components of the high affinity insulin binding domains, L1 and the C-terminal insulin binding helix (α CT), which move closer together. Specifically, the N-terminal L1 domain is observed to swing upwards $\sim 55^\circ$ while the α CT is observed to move up 62 Å (6.2 nm), such that upon insulin binding the distance between residues of α CT and L1 is approximately 0.35 nm²⁰. The Förster distance (R_0)- the distance at which is a FRET efficiency is 50% is 4.9 nm for genetically-encoded FRET pairs, cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP), and 6.3 nm for Alexa Fluor 488 and 546 fluorophores^{32,33}. Therefore, the above described ECD holo and apo conformational changes of InR were expected to change FRET efficiency from 50% to its maximum efficiency per molecule. Additionally, previous studies have demonstrated that the InR ECD is able to bind insulin at a wide range of affinities based on deletion and modular arrangement of the domains^{22-24,34,35}.

In this chapter, the ECD and four different constructs, 1-4 (Figure 4-1C and 4-1D), with various combinations of insulin binding sites and their supporting domains were cloned for expression in BL21DE3 cell lines. Unfortunately, the constructs failed to express solubly and were found in the inclusion bodies following cell lysis. Attempts to optimize expression conditions or solubilize the proteins with detergents, such as Sarkosyl and lauryldimethyl amine oxide (LDAO), also did not prove successful.

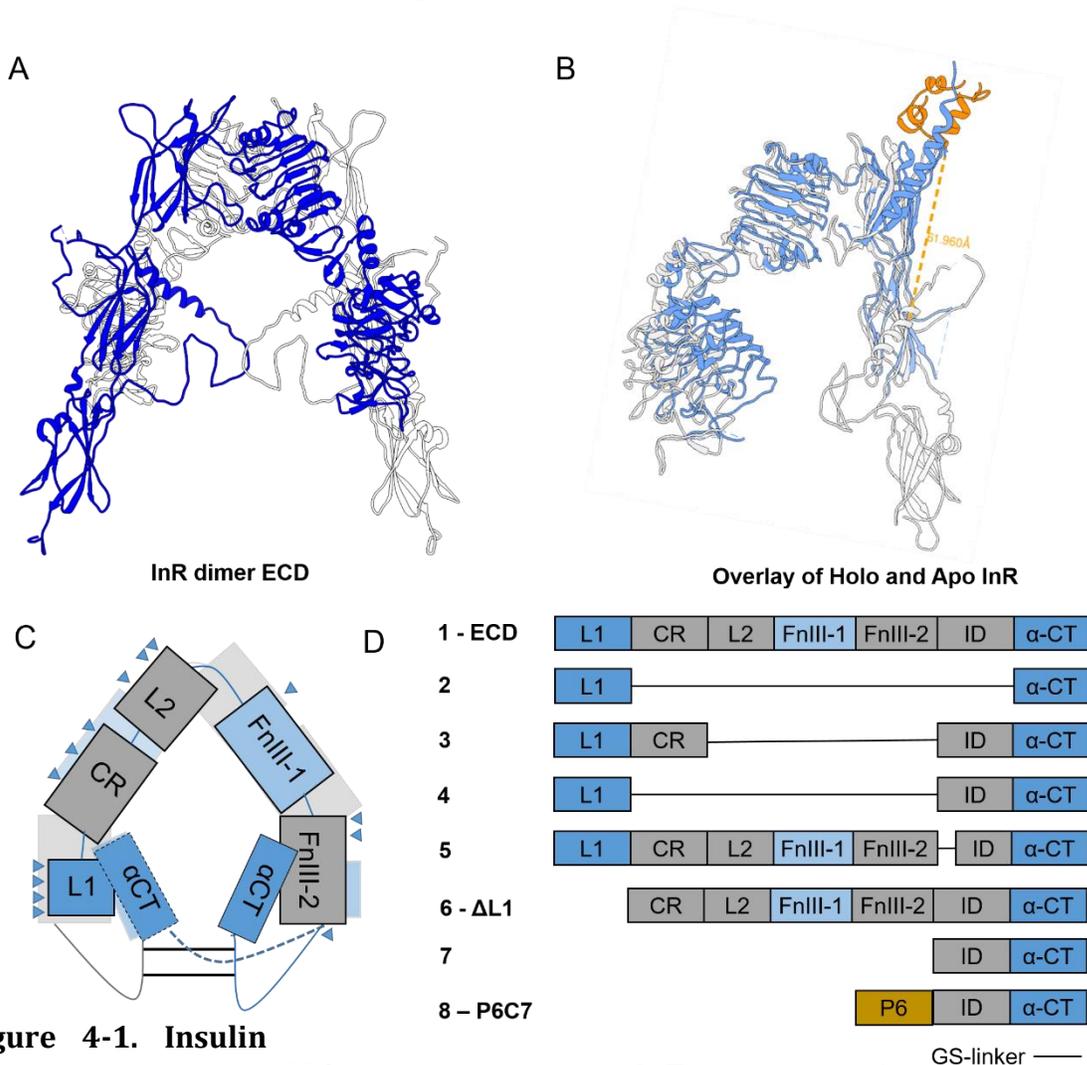


Figure 4-1. Insulin receptor structure and biosensor constructs. **A.** The structure of dimeric InR with monomers depicted in grey and dark blue. **B.** Overlay of apo (blue, PDB: 6CE7) and holo (light grey, PDB: 4XZB) structures of the insulin receptor (InR) ectodomain (ECD). The orange dashed line represents the change in conformation of the αCT domain upon insulin (orange) binding. **C.** Schematic diagram of the dimeric InR depicting the different domains on a single monomer. Darker blue boxes represent components of the high

affinity insulin binding site, the lighter blue represents the low affinity binding site. The dashed line and boxes represent the binding of α CT from the same monomer interacting with the L1. Due to lower structural resolution in flexible ID domain (dashed line), it is unknown whether the α CT and L1 from the same monomer is likely to interact to form the high affinity insulin binding site. **D.** InR-derived constructs used to design an insulin biosensor using various combinations of insulin binding sites and their supporting domains. Orange box represents peptide 6 (P6) that was discovered during phage selections and used to design a chimeric InR construct.

Expression of functional InR and its minireceptor constructs in literature are conducted using mammalian (Chinese Hamster Ovary (CHO)) or insect (baculovirus) cell lines, which allow post-translational modifications (PTMs)³⁶⁻³⁸. To determine whether PTMs would increase soluble and functional expression of the rationally designed InR constructs, stable expression of the InR constructs was conducted in Flp-In™ CHO cells (Thermofisher)³⁹. Constructs were subcloned with an N-terminal albumin secretion tag such that completely folded proteins could be collected in the cellular media. Its C-terminal His₆-tag was subsequently used for protein purification using nickel-affinity chromatography. The protein was detected by a binding assay after its N-terminal FLAG-tag and fluorophores were site-specifically conjugated to an N-terminal SNAP-tag. Using this platform, the full-length ECD, the second largest construct C5, and the smallest C4 were expressed and assayed (Figure 4-2A).

The InR C5 does not contain the α - α dimerizing cysteine triplet (Cys⁶⁸², Cys⁶⁸³, Cys⁶⁸⁵) but does contain Cys⁵²⁴, the second α - α dimerization site^{24,40,41}. As such this construct was expected to form weak dimers. In contrast, the smaller C4 contained only the high affinity binding domains L1 and α CT linked by the flexible insert domain (ID) comprising the dimerizing triplet cysteines. In western blot analyses of InR constructs purified from cell

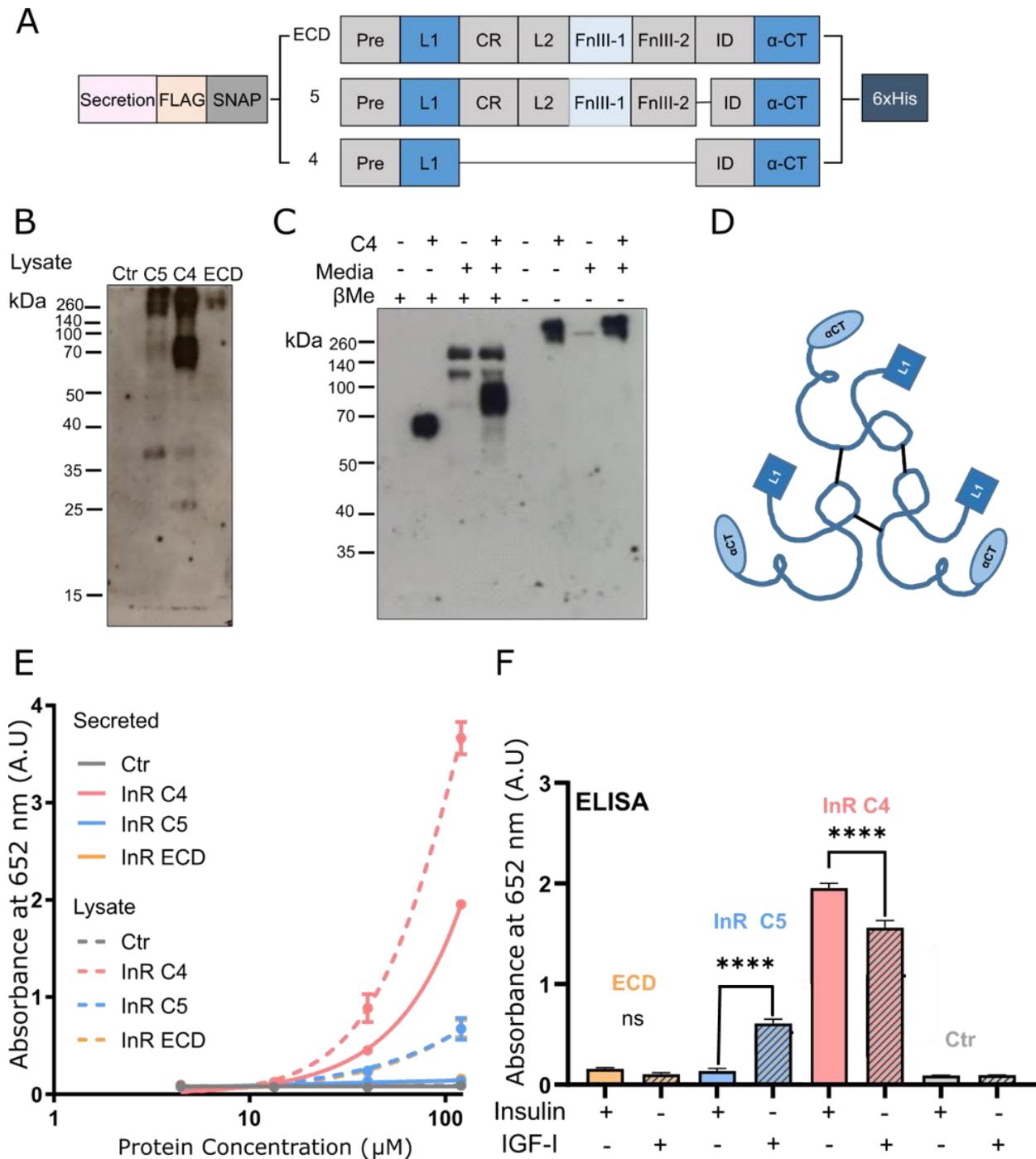


Figure 4-2. Expression and characterization of InR constructs in CHO cells. **A.** Schematic representation of InR constructs ECD, C5 and C4. Constructs contain an N-terminal albumin secretion tag for secretion in to the cellular media, a FLAG-tag for protein detection, a SNAP-tag for site-specific fluorophore conjugation and a C-terminal His₆ tag for affinity purification. **B.** Western blot analysis, using αFLAG-HRP to detect protein, demonstrates expression of InR constructs C5, C4 ECD in cellular lysate. The untransfected cell lysate is used as a negative control to determine any non-specific binding signal. The in the absence

of glycosylations the expected monomeric sizes for ECD, C5 and C4 are 106 kDa, 97 kDa and 57 kDa respectively. In the presence of glycosylations the ECD dimer is known to be approximate 350 kDa⁴². Oligomerization of all constructs are observed in the SDS-PAGE gel; specifically ECD forms dimers, C5 forms dimers and several higher molecular weight oligomers, C4 forms monomers and several higher molecular weight oligomers. **C.** InR C4 collected in media and lysate and run on SDS-PAGE gels in the presence and absence of the reducing agent β -mercaptoethanol (10 mM). Disulfides are found to be responsible for oligomerization. **D.** Schematic representation of InR C4 oligomers (trimer). Black lines represent disulfide bonds. **E.** ELISA showing insulin (2 μ M) binding of InR ECD, C5 and C4 collected from cell lysate (dashed lines) and cell media (solid lines). **F.** ELISA showing binding of indicated InR constructs to 2 μ M insulin (solid bars) or 2 μ M IGF-I (bars with diagonal patterns). InR C4 demonstrates specificity towards insulin whereas, InR C4 appears to prefer binding to IGF-I.

lysate in reducing SDS-PAGE conditions, the InR ECD and C5 still appeared to form dimers > 260 kDa. This suggested strong disulfide bonding or the formation of disulfide-independent aggregates (Figure 4-2B). The fully glycosylated dimeric ECD has been estimated to be 350 kDa⁴², as such the glycosylated monomeric ECD with its SNAP-tag was estimated to be 195 kDa. In contrast, the SNAP-InR C4 variant, which was expected to be 57 kDa, demonstrated higher molecular weight likely due to additional PTMs when secreted into the media and disulfide dependent oligomeric states that were converted to monomers in the presence of reducing agents, respectively (Figure 4-2C and 4-2D).

The constructs were then used to determine insulin binding using ELISAs with insulin-immobilized onto microtiter plates. Results demonstrated that all three constructs from the lysate, ECD, C5 and C4, bound insulin in a dose-dependence with C4 having the strongest binding (Figure 4-2E). Secreted InR C4 also demonstrated insulin binding, but secreted InR C5 and ECD did not show such binding. Next, specific binding to insulin was examined, where InR C4 purified from cell lysate demonstrated slightly better binding to insulin than its homolog IGF-I. In contrast, C5 demonstrated preferred binding to IGF-I over

insulin (Figure 4-2F). Shortening the flexible insert domain likely resulted in a wider binding pocket formed by L1' and α CT, allowing for better binding interactions with the larger IGF-I protein (7.6 kDa) and weaker binding to insulin (5.8 kDa). Surprisingly the ECD did not show strong binding to either protein in this assay. Since insulin was immobilized on the plate it is also possible that immobilization may have hindered access to critical binding surfaces.

Previously, Scholler *et al.* demonstrated whole InR FRET on the cell surface using N-terminal SNAP-tags on each monomer. In their investigations, the group used optimized stoichiometry of fluorescent FRET pairs with benzylguanine (BG) moieties to differentially conjugate InR monomers to generate dimeric FRET constructs and detect FRET-based insulin binding using time resolved FRET (trFRET). The study demonstrated dynamic range between 10 nM to 1 μ M insulin, which was the lowest range observed amongst nine other receptor tyrosine kinase sensors tested by similar methods¹². In our investigation, a similar approach was applied to detect FRET in the InR ECD, C5 and C4, where the constructs were subcloned with N-terminal SNAP-tags and conjugated using cost-effective and readily available Alexa Fluor FRET pairs, BG-Alexa Fluor 488 and BG-Alexa Fluor 546 (ThermoFisher) (Figure 4-3A). Media from untransfected cells subjected to identical conditions but omitting the fluorophores were used as negative controls. FRET assays were conducted using varied doses of the fast-acting insulin derivative, Humalog® (6 - 29 μ M). Humalog® contains mutations P28K and L29P in which the positions of lysine and proline at positions 28 and 29 in the B-chain are inverted⁴³. Only the InR C4 FRET pair was observed to be conjugated by both fluorophores, although the Alexa Fluor 546 conjugation appeared less efficient than the Alexa Fluor 488 (Figure 4-3B).

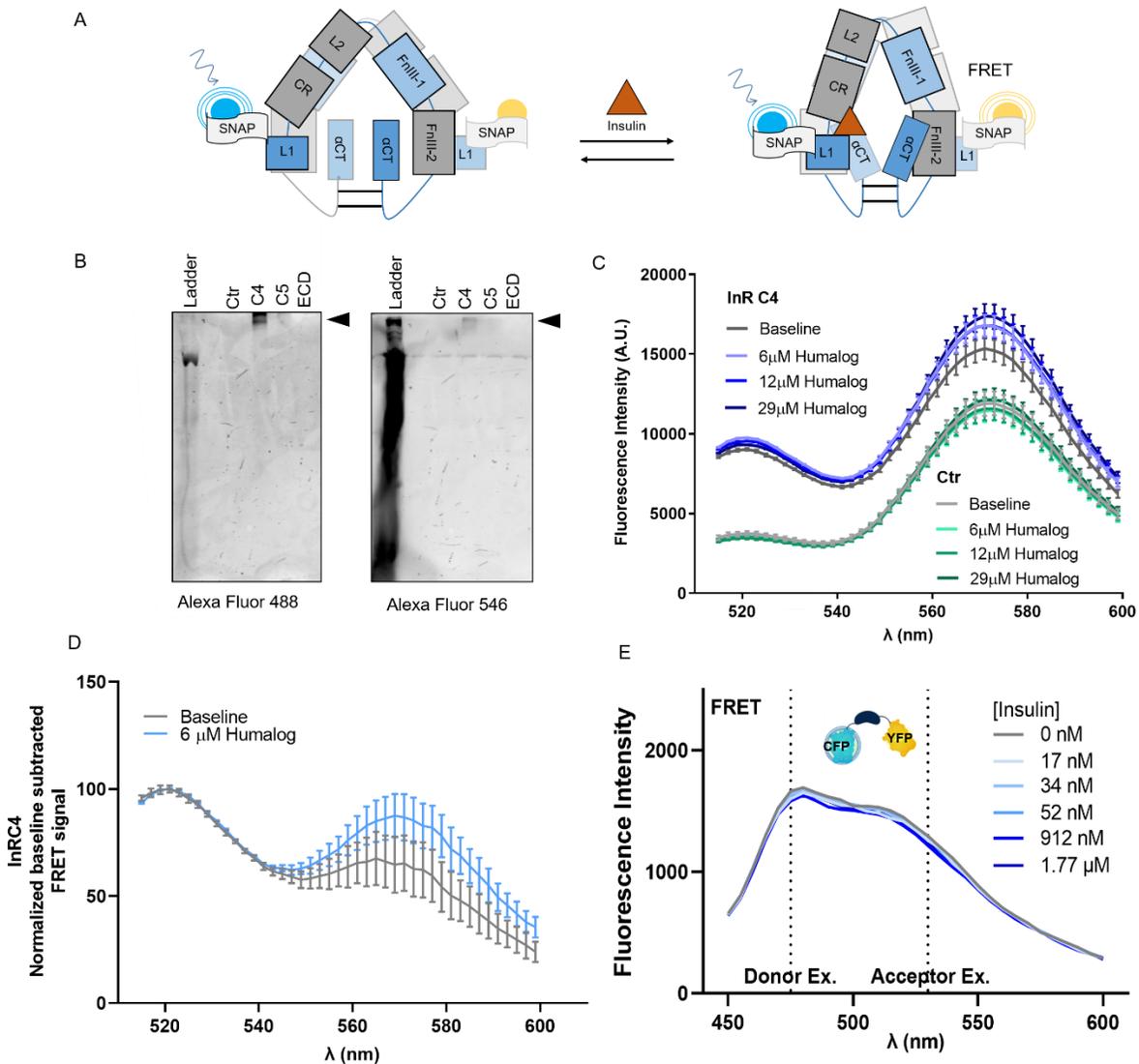


Figure 4-3. Humalog®-dependent change in FRET signal of CHO cell expressed InR C4. **A.** Schematic representation of SNAP-tag based InR FRET construct. Stoichiometrically adjusted (1:6) fluorophore BG-tagged Alexa Fluor 488 and 546 FRET pairs were simultaneously conjugated to SNAP-tags on the N-termini of monomers of purified InR constructs from CHO cell media. **B.** SDS-PAGE gel imaged using appropriate filters for Alexa Fluor 488 and Alexa Fluor 546 on the Typhoon Fluorescence Gel imaging system. Only InR C4 showed conjugation of both fluorophores, although Alexa Fluor 546 conjugation appeared to be less efficient. **C.** The raw spectral data for InR C4 and the untransfected control are represented at baseline and after administration various concentrations of Humalog®. **D.** Graph represents the normalized FRET change in InRC4 upon Humalog administration, after signal subtraction from untransfected negative control and normalization of signal on a percentage scale. **E.** Graph represents FRET signal from monomerized InR C4 construct with N-terminal CFP and C-terminal YFP. No change in FRET is observed with increasing doses of Humalog.

FRET analysis of the SNAP-tagged InR C4 and the untransfected negative control revealed that the control also demonstrated an unexpected FRET peak at the 571 nm, indicating some direct Alexa Fluor 546 acceptor excitation from free-floating fluorophores (Figure 4-3C). Thus, normalization by negative control subtraction was used to determine insulin-dependent conformational changes by InR C4. Results demonstrated 6 μ M Humalog increased FRET signal relative to the “no insulin” baseline (Figure 4-3D). This concentration appeared to saturate insulin binding to FRET InR C4 constructs as no further increases were observed at higher concentrations of insulin. Notably, the well-to-well FRET signal deviation was large, possibly due to pipetting error, as observed by the slightly overlapping SEM error bars between the Humalog-induced signal and baseline. The use of trFRET, as described in Scholler *et al.*¹², could significantly help increase sensitivity of this assay by reducing background signal. However, since the inefficient tagging of both FRET pairs and the inherent inefficiency and inconsistencies of FRET pair dimer formation would still pose problems towards signal optimization, this path was not chosen for future investigation.

To optimize conjugation and signal consistency, a FRET construct using a monomeric InR C4 lacking the disulfide forming cysteines and comprising genetically encoded FRET pairs CFP and YFP, at the N- and C-termini, respectively, was designed from the backbone of the FRET construct, Twitch2B (described in Chapter 5). A similar InR construct adhered to the cell surface was previously shown to have bioluminescent resonance energy transfer (BRET) activity¹³. Surprisingly, the InR C4 construct did not demonstrate any change in FRET relative to varying doses of Humalog. Therefore, it was concluded that the oligomeric form of the InR C4 was likely responsible for any observed FRET change in response to Humalog (Figure 4-3E). The long growth times, low concentrations of protein production and

inefficient purification methods, limited the capacity for high-throughput construct production and screening hindered efforts to further optimize this system. These numerous obstacles of the mammalian protein expression platform indicated that it was not well-suited for further optimization of the insulin biosensor. Thus, in the next steps we address the issue of developing InR constructs on a robust protein expression platform, which also allows high-throughput selections for specific insulin binding and screening for optical signal.

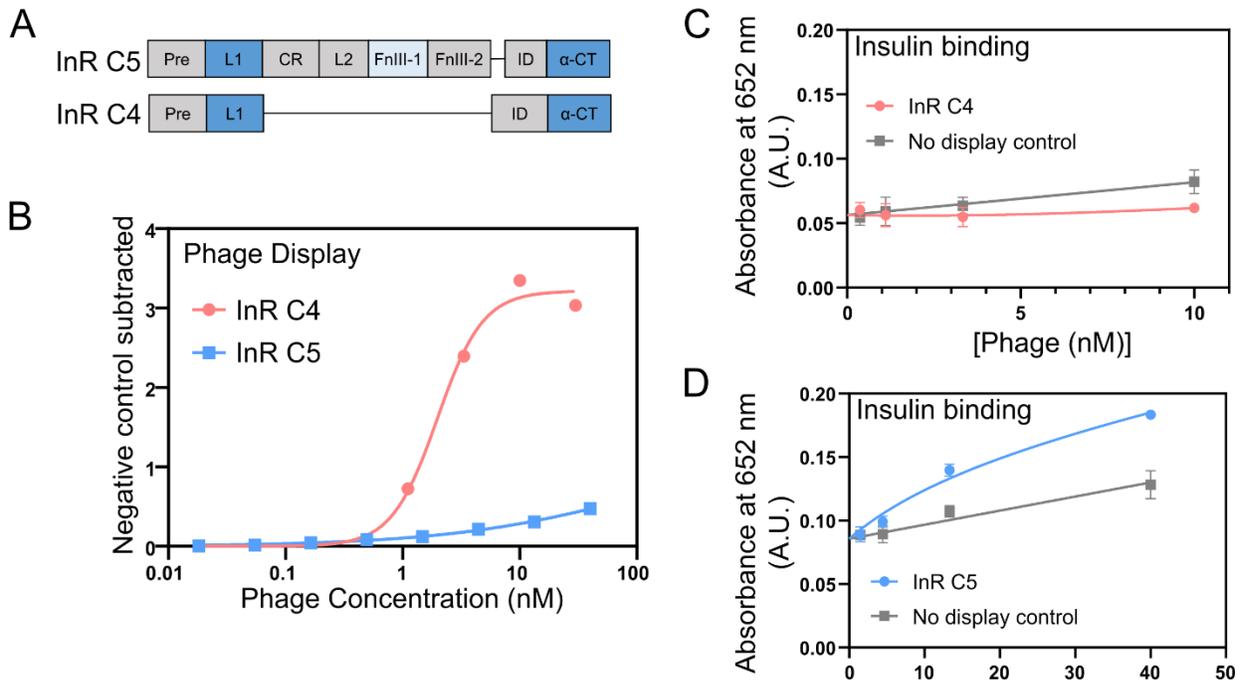


Figure 4-4. Phage display and insulin binding of InR constructs. **A and C.** ELISA represents dose-dependent phage display of InR constructs InR 4 and 5. InR C4 demonstrated maximal phage display. **B.** ELISA represents insulin binding to InR constructs, with 2 μ M insulin immobilized onto microtiter plates and phage binding is detected using α M13-HRP. No insulin binding is observed in these constructs. **D.** Insulin binding to SNAP-tagged InR C5 using immobilized biotinylated insulin (Nanocs Inc.) demonstrates low levels of insulin binding.

chosen (inset, left) for site-directed library generation, to increase variant solubility. The theoretical diversity of the library is 2×10^9 . Using Gibson Assembly cloning techniques a library with approximately 10^4 diversity was generated. **B.** Libraries were selected for two rounds for insulin binding. Library variants between 4 and 5 fold binding over background demonstrated variants that bound insulin better than InR C5 but lacked the L1 domain. Library variants ≥ 5 fold over background insulin binding were very short fragments that only included the library region of the protein. **C.** InR construct schematic showing the common residues found in insulin binding variants selected from the library. Cysteines and prolines were unaltered in the library design, however, more charged residues are observed. Additionally, the L1 domain is found to be lacking in these variants. **D.** ELISA showing the insulin binding of the specific InR variant (highlighted as a red square in panel C), where $2 \mu\text{M}$ insulin is immobilized on microtiter plates and the phage-displayed variant binding is quantified using $\alpha\text{M13-HRP}$.

simultaneously allowing high-throughput selections and screens. As such, C4 and C5 (Figure 4-4A) were successfully displayed on phage (Figure 4-4B). However, only C5 demonstrated low levels of binding to insulin (Figure 4-4C, 4-4D).

To improve insulin binding, solubility, and specificity, SNAP-tagged C5 was chosen as the template for further selections. The cysteine rich (CR) domain, which lies adjacent to the insulin binding L1 domain, contains 25 disulfide forming cysteines and was thought to be a possible region where proper folding for binding and solubility might be especially problematic. Using surface hydrophobicity analysis on UCSF's Chimera visualization tool⁴⁴, the surface exposed hydrophobic regions of the CR domain were identified and targeted for library generation. Libraries were designed such that these hydrophobic regions would be replaced with polar or charged residues - a strategy previously used by the Weiss lab to solubilize the hydrophobic intramembrane domain of calveolin-1⁴⁵. Prolines and cysteines that are usually important for overall domain conformation were not changed (Figure 4-5A). Selectants were screened after two rounds of selection with insulin binding and one round

with IGF-I negative selection, (see further details in Materials and Methods section). Peculiarly, while InR C5 variant selectants contained more hydrophilic CR residues as encoded by the library, they also omitted the L1 domain entirely (Figure 4-5B, 4-5D). The results suggest that the L1 domain, and not the CR, was the likely culprit for improper folding of InR in *E. coli* and very low insulin binding.

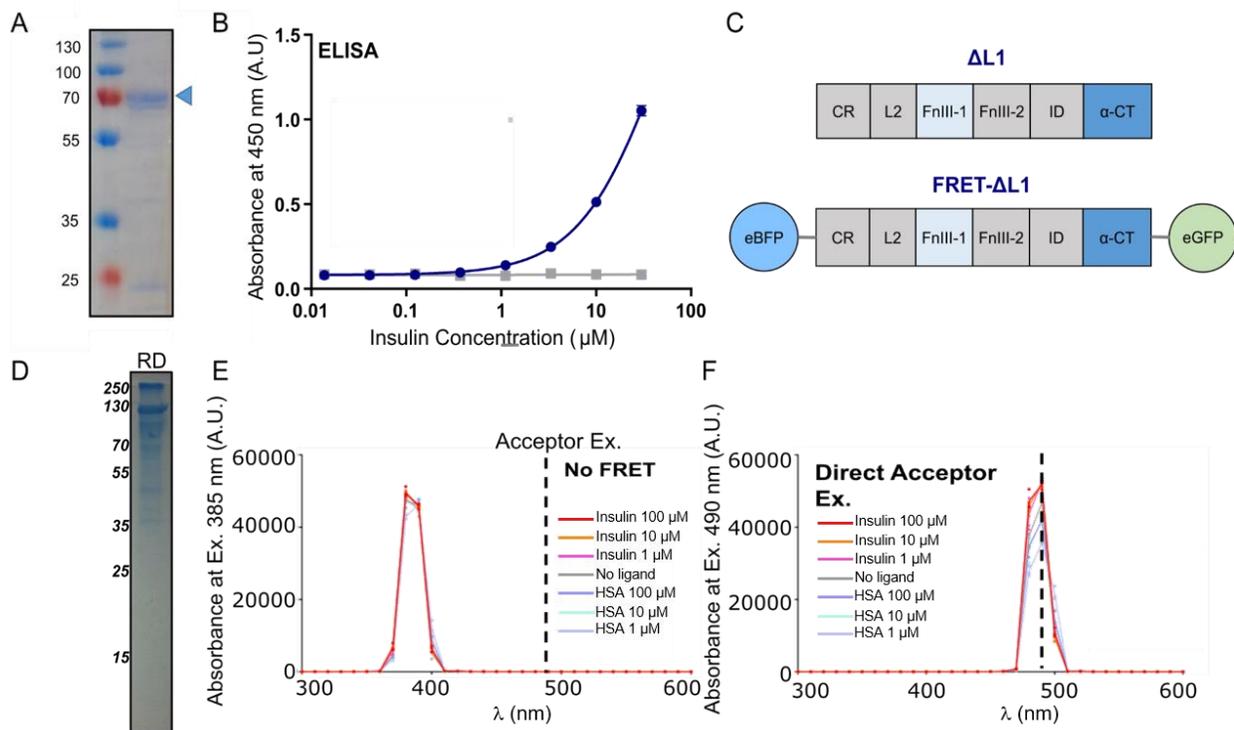


Figure 4-6. *E. coli*-expressed L1 domain disrupts InR solubility and insulin binding. **A** SDS-PAGE gel shows the solubilized and refolded ΔL1 construct at its expected size of 70 kDa. **B**. ELISA represents insulin binding of ΔL1 , where 143 nM of ΔL1 is immobilized on the microtiter plates and biotinylated insulin (Nanocs Inc.) at various concentrations is used to determine binding using streptavidin-HRP. **C**. Schematic representation of ΔL1 and FRET construct of ΔL1 . **D**. SDS-PAGE representing soluble FRET ΔL1 as a monomer and a dimer after refolding using rapid dilution. **E**. FRET assay showing donor peak after direct donor excitation but FRET signal from the acceptor. **F**. Spectral graph represents acceptor peak of FRET ΔL1 (eGFP) after direct excitation.

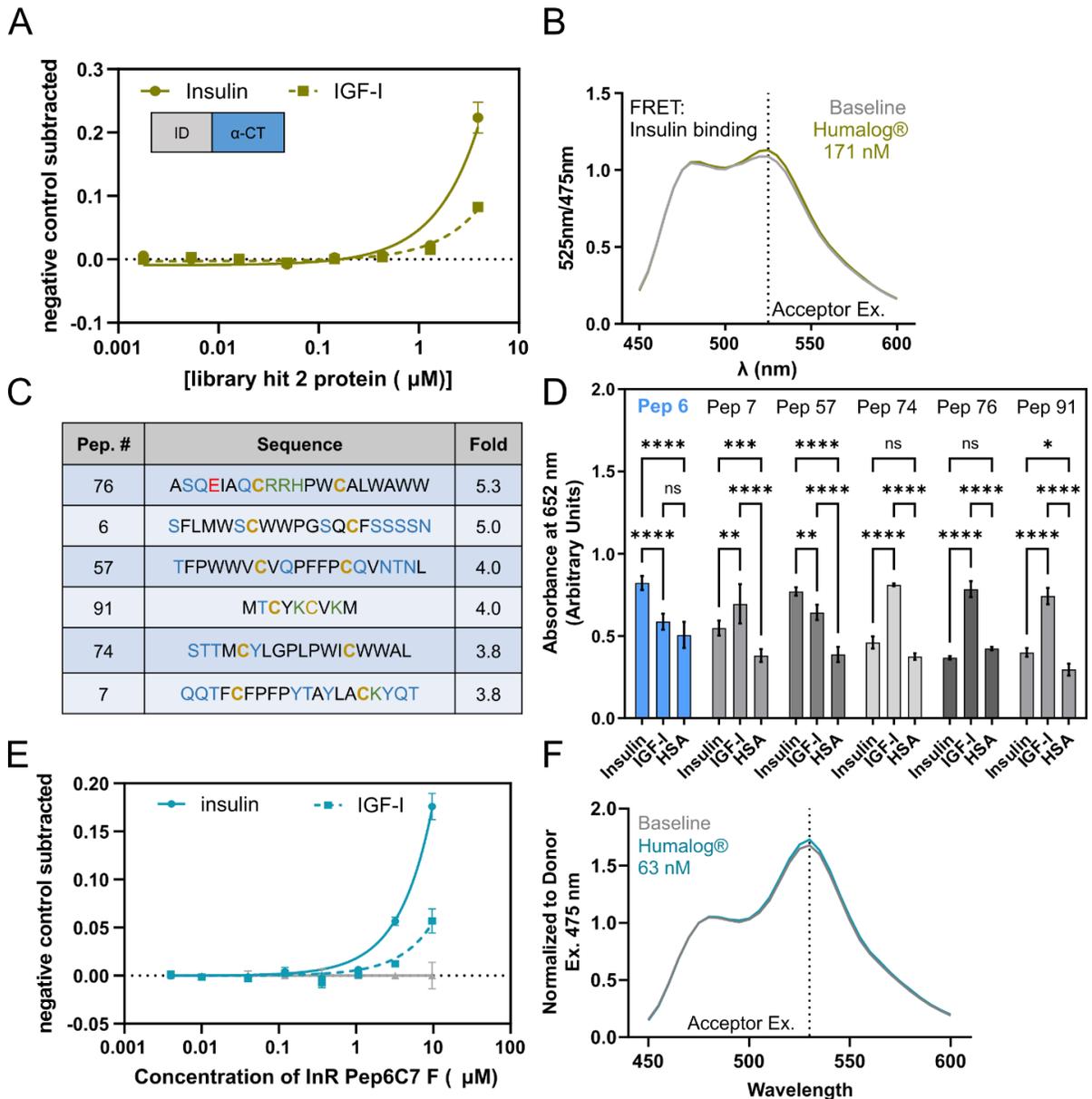
The L1 domain is not functionally expressed in simpler organisms

To test the hypothesis that the absence of the L1 domain could result in improved InR solubility in *E. coli* and binding to insulin, the domain was omitted from the new construct, Δ L1 (Figure 4-6C). Δ L1 was more soluble than other InR constructs (Figure 4-6A), though a significant amount of protein was still found in the inclusion bodies. ELISAs for insulin binding were conducted such that 143 nM of soluble Δ L1 was immobilized onto microtiter plates and biotinylated insulin at varied concentrations was added in solution and detected through the addition of streptavidin-HRP, followed by a colorimetric HRP substrate. Hydrolysed N-hydroxysuccinamide-biotin in the absence of insulin was added at equimolar concentrations as a negative control. Results demonstrated that insulin binding was observed at insulin concentrations starting at 185 fold over the highest physiological concentrations (370 nM – 30 μ M, Figure 4-6B). Thus, the dynamic range of insulin binding to Δ L1 was still too low for physiological insulin detection.

To determine if conformational change of Δ L1 upon insulin binding could be used to generate an insulin dependent FRET signal, the construct was subcloned in a pET28c plasmid with the FRET pair fluorescent proteins, eBFP and eGFP, flanking its N- and C-termini, respectively. The monomeric construct was 130 kDa and mostly insoluble. Denaturation and refolding of the FRET- Δ L1 construct using rapid dilution resulted in visibly fluorescent soluble protein, which was observed to be at the correct monomeric and dimeric sizes in an SDS-PAGE gel (Figure 4-6D). Then, 1 μ M of refolded FRET- Δ L1 was assayed for a FRET signal at varying doses of insulin (1, 10 and 100 μ M). Equimolar concentrations of human serum albumin (HSA) was used as a negative control. While both FRET fluorophores could be directly excited, unfortunately, no detectable levels of FRET was observed for this

protein (Figure 4-6E, 4-6F). Therefore, it was concluded that the L1 domain may contribute to InR misfolding when expressed in simpler organisms, where Δ L1 InR slightly improved solubility and insulin binding. The omission of a properly folded L1 domain, however, also resulted in the loss of the expected conformational change upon insulin binding.

Figure 4-7. Designing compact, soluble, and specific insulin binding constructs. A.



Graph represents insulin binding ELISA where a compact construct, termed C7 (expressed solubly in *E. coli*) binds at various doses to immobilized 2 μ M insulin and IGF-I. **B.** FRET assay with InR C7 flanked with CFP and YFP does not demonstrate sufficient FRET in response to varying doses on Humalog. **C.** Amino acid sequences of peptide selectants with the highest insulin binding after four rounds of selections for specific insulin binding are depicted. **D.** Peptide 6 (P6) is subcloned on to the N-terminus of compact construct 1, termed P6C7. ELISA represents insulin and IGF-I binding to P6C7, as previously described, and binds specifically to insulin. **F.** FRET construct P6C7, with flanking CFP and YFP, does not show any FRET in response to Humalog binding.

The Peptide-InR Chimeric construct, P6C7

Since the L1 domain of the InR high affinity binding site could not be functionally expressed in *E. coli* or phage display platforms, we subsequently investigated whether the α CT alone could bind insulin. In previous literature the synthesized α CT helix could bind insulin, although binding was shown in the presence of a separately expressed L1 domain expressed in BHK cells³⁴. We therefore examined whether the α CT helix bound to the flexible ID domain, with CFP and YFP at its N- and C- terminal flanking ends, could also specifically bind insulin and result in the desired FRET signal. The new construct was termed InR C7. ELISAs assessing binding of C7 to insulin or IGF-I demonstrated that the highest concentration of C7 tested (4 μ M) bound insulin 1.6 fold over IGF-I (Figure 4-7A). However, FRET analysis of the C7 construct did not demonstrate any discernible signal upon Humalog binding at concentrations ranging from 170 pM to 6 μ M (Figure 4-7B). The data thus reveals that while the C7 alone is capable of binding insulin, it does not change the FRET signal. This is expected since, with only a single binding site, dramatic conformational changes are unlikely to occur. Additionally, the entire C7 construct is approximately 58 Å, which is very close to the R_0 value of the CFP-YFP FRET pair (49 Å). Therefore, a significant amount of

FRET was already present at baseline levels and slight changes in FRET upon insulin binding were unlikely to be useful for measurements.

To find other simple insulin binding components that can be used for real-time insulin detection, selections for specific insulin binding were conducted using the previously generated Mega Random Peptide Library (MRPL)⁴⁶. Selections were conducted for four rounds, including positive selections for insulin binding and negative selections against IGF-I binding. Spot assays of 192 total colonies resulted in 21 unique peptides that bound insulin >3 fold over non-specific binding to the bovine serum albumin (BSA) negative control. Of these 21 selectants, the six phages with highest binding over background were further tested for specificity over IGF-I. Results demonstrated that despite selections against IGF-I, 4 out of 6 peptides preferred binding to IGF-I over insulin. Two 20 residue peptides, P6 and P57, showed significant binding to insulin relative to HSA negative controls ($p < 0.0001$, two-way ANOVA, *ad hoc* Tukey). P6 additionally showed no significant binding to IGF-I relative to HSA (Figure 4-7C, 4-7D). Therefore, Pep 6 was chosen for subsequent analyses.

To generate a longer construct ($\sim 128 \text{ \AA}$, excluding fluorophores) that binds to insulin and undergoes a possible conformational change, the P6 sequence was subcloned to the N-terminus of the InR C7 FRET construct. Based on structural measurements and the average size of a 20 residue peptide, approximating 3.5 \AA per residue, it was estimated that the P6C7 construct would have a FRET efficiency of 0.3% compared to the estimated 27% for C7. ELISAs assessing dose-dependent insulin and IGF-I binding demonstrated approximately 1.5 fold preferred insulin binding over IGF-I at P6C7 concentrations of 3.22 and 9.67 μM (Figure 4-7E). Notably, the overall background subtracted signal of P6C7 was lower than that of C7 alone. Unfortunately, FRET assays for this construct also did not result in an insulin

dependent signal, while a significant FRET signal was still observed at baseline (Figure 4-7F). The FRET observed over baseline was greater than that observed for C7 alone. Therefore, it is likely that the P6 and C7 interact with each other. While the peptide and InR-derived compact constructs can bind insulin with some specificity, insulin-dependent FRET signals, which likely require conformational changes over larger distances, could not be observed.

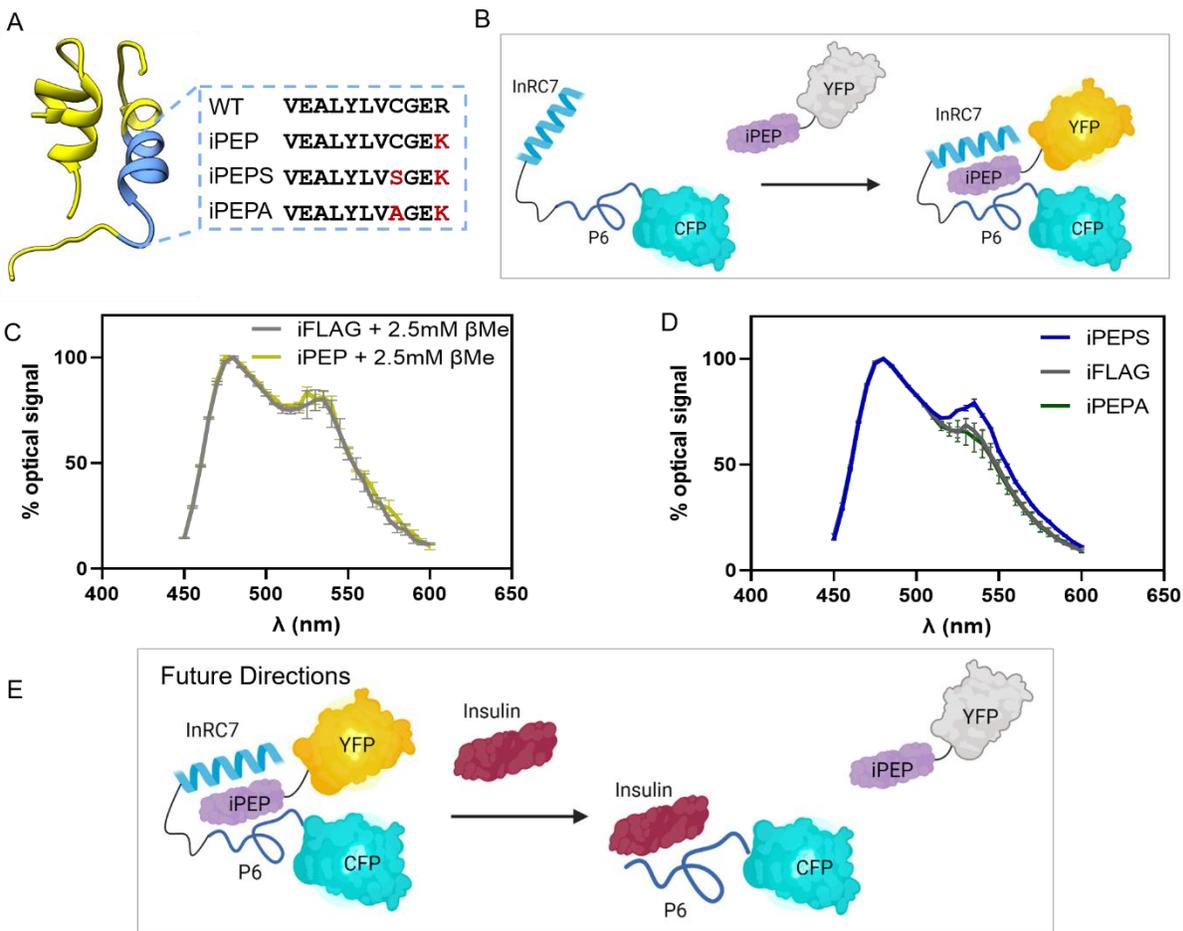


Figure 4-8. A two-protein FRET Assay design. **A.** Schematic representing a two protein FRET assay where an insulin B-chain derivative iPEP is tagged with a C-terminal YFP and the P6C7 is tagged with an N terminal CFP. The saturated binding of the two InR and iPEP is expected to generate a maximal FRET signal. Varying concentrations of unlabeled insulin are then expected to outcompete the iPEP and dose-dependently reduce FRET. **B.** iPEP alone did not demonstrate a FRET signal **C.** An SDS-PAGE of iPEP indicates that oligomerization resulting from a free cysteine in the iPEP sequence is likely. **D.** iPEP variants with its single

cysteine replaced with alanine (iPEPA) and serine (iPEPS) were generated and tested for FRET. iPEPS and not iPEPA generated a FRET signal relative to the iPEPS baselines.

A two-protein FRET Assay

Next, we devised a two-protein FRET system that detects changes in FRET based on binding of two different proteins each with a single FRET pair fluorophore attached. Two-protein FRET systems have several advantages over FRET constructs that depend on conformational changes of a single protein. Firstly, larger differences in binding distance between bound and unbound states allows for very low signal-to-noise. Secondly, stoichiometric control over each FRET component allows tunable sensitivity and dynamic range. Thirdly, the signal for the donor and acceptor molecules can be normalized with more accurate controls, allowing for more accurate calculations of FRET efficiency⁴⁷.

This new FRET system design for insulin binding includes the P6C7 with an N-terminal CFP protein and an insulin B-chain derivative, here termed iPEP, with a C-terminal YFP tag (Figure 4-8A). We propose a biosensor that, in the absence of insulin, will emit a maximum FRET signal under saturated binding conditions between iPEP and P6C7 (Figure 4-8B, schematic). As insulin is administered to the system, the insulin will outcompete the iPEP and result in dose-dependent reduction in FRET (Figure 4-8E, schematic). The iPEP construct has been previously used for structural analyses to determine binding interactions between insulin and the insulin-degrading enzyme (IDE)⁴⁸. Structural analysis of InR α CT and insulin interactions report that residue F714 from InR α CT interacts with a hydrophobic crevice within insulin, which contains B chain sites V12 and L15²⁹. Therefore, it was hypothesized that the iPEP should weakly bind P6C7.

To test P6C7 and iPEP binding, a FRET-based assay was developed such that non-specific binding, acceptor excitation, and molecular crowding could be normalized using a YFP and FLAG tagged negative control without the iPEP. This control protein was termed iFLAG. To avoid self-aggregation due to disulfide formation with the single cysteine residue C8 on iPEP, binding assays were first conducted under reducing conditions and no change in FRET was observed (Figure 4-8C). The reducing conditions may disrupt disulfide bonds formed within P6C7 and thus disrupt its insulin binding capabilities. Therefore, the two iPEP-derived constructs with the mutations, C8A or C8S, termed iPEPA or iPEPS, were designed to test FRET-based binding to P6C7. Results demonstrate an observable FRET signal in response to 10 μ M iPEPS relative to the negative control iFLAG, but not with iPEPA (Figure 4-8D). Testing higher concentrations resulted in FRET within the negative control demonstrating a possible crowding effect. Therefore, further increases in specific binding signal could not be observed. Future work towards the development of this assay should include outcompeting the signal using varying doses of insulin and developing FRET constructs and conjugation methods that allow for more sensitive detection of FRET with lower concentrations of protein (Figure 4-8E).

4.4 Conclusions and Future Directions

This study establishes that the InR expression and function is highly dependent on PTMs. Literature demonstrates that InR-based optical insulin sensors can be designed in limited (mammalian and insect) cell lines and under surface-tethered conditions. Our investigation demonstrates that CHO cell expressed, untethered, oligomeric InR C4, containing the L1 and the α CT together, can bind insulin and undergo conformational

changes that are visualized as FRET. The difference in signal, however, has very low signal-to-noise ratios. Furthermore, dose-dependent increases in FRET were not observed. Further testing with linkers between L1 and α CT could be examined. However, the mammalian expression system does not allow high-throughput library expression and screening.

The study also examines whether human InR or rationally designed deletion variants could be functionally expressed in simpler expression systems such as phage or *E. coli* that are more permissive to high-throughput selections and screens. The current investigation identifies that with the exception of the L1 domain, the InR can be functionally expressed in simpler organisms, such as phage or *E. coli*. Notably, smaller InR variants, such as InR C4 comprising the L1 domain could be solubly expressed in *E. coli* and displayed on phage, but had lost the ability to bind insulin in the absence PTMs. Conversely, constructs missing just the L1 domain are only slightly soluble, but importantly retain the ability to bind insulin. This Δ L1 construct, however, does not undergo the large conformational changes necessary for FRET signal generation do not occur.

In this study, we also establish that the α CT domain alone is capable of binding insulin in the absence of all other domains of the InR. For future experiments, the InR C7 construct containing only the α CT flanked by the ID flexible linker can be used in conjugation with the insulin B-chain-derived iPEPS to test its binding and suitability for the two-protein FRET platform developed in this study. Additionally, other native insulin binding proteins such as IGFBP-7⁴⁹ and insulin-degrading enzyme⁴⁸, that have limited binding dependent conformational change, can be used as candidates for the development of two protein assays. The relative binding affinities to iPEPS and insulin can allow selection of protein pairs with the most suitable insulin detection range and kinetics for real-time sensing. To reduce steric

hindrance during binding and increase signal-to-noise of the two-FRET protein assay, the large genetically-encoded fluorophores could be replaced with bioconjugated small molecule FRET pairs⁵⁰. Such FRET pairs can be tested for reduced bleed-through and direct acceptor excitation than is observed between CFP and YFP.

This investigation lays the foundations for a FRET-based insulin biosensor design and assay development for furthering automated T1D disease management. The study highlights the utility of native proteins as biosensing tools for essential biomarkers in chronic diseases and their treatment. The use of multiple platforms for protein expression, including mammalian cell lines, M13 bacteriophage display, and *E. coli*, highlight both the advantages and caveats of each and how each platform can be critical to various stages of biosensor development. The protein-based approaches and assays described here can be further expanded to incorporate other important metabolites of T1D, like β -hydroxybutyrate for detecting diabetic ketoacidosis⁵¹, or glucagon for its direct detection in patients treated with dual hormone automated devices. The real-time detection of biomarkers for progressive diseases can also result in effective, monitored early treatment⁵². As such, on-going and future projects in our laboratory can build on the assays and sensor engineering approaches from this investigation towards the development of a series of protein-based biosensors for accurate, real-time monitoring of disease conditions.

4.5 Materials and methods

Generating stable cell lines expressing InR constructs in using the Flp-In™ CHO System. The Flp-In™ CHO cell line comprises an integrated FRT site that can be cleaved by Flp recombinase and a LacZ-Zeocin™ fusion gene downstream an ATG start site and

controlled by an SV60 promoter. InR constructs were subcloned into a pcDNA5/FRT expression vector using the Gibson Assembly 2x Hifi kit, as per manufacturer's instructions. Gene expression in the pcDNA5/FRT vector is controlled by a human CMV promoter. The vector also contains a hygromycin resistance gene with an FRT site inserted at the 5' coding region, but lacking an ATG start site.

The subcloned pcDNA5/FRT plasmid and a pOG44 plasmid that constitutively expresses Flp recombinase were co-transfected into the Flp-In™ CHO cell line (ThermoFisher Scientific). Specifically, 20 uL Lipofectamine® 2000, 7 µg of POG44, and 7 µg construct (pcDNA5/FRT plasmid excluded for control cells) were prepared in 300 uL Gibco Dulbecco's Modified Eagle Medium (DMEM) and incubated for 5 min at room temperature (RT). The DNA-lipid mixture was then added to Flp-In™ CHO cells media when cells were 60% confluent. Cells were incubated for 48 hours after transfection and passaged using trypsin in DMEM with 10% fetal bovine serum (FBS) at a density of about 8×10^5 cells/ml. Cells were then treated with 2 mg/mL hygromycin. Insertion of the cleavage of the FRT site in the cell line and the insertion FRT cassette in the pcDNA5/FRT vector alters resistance from Zeocin to hygromycin. Thus, selection and growth of stably transfected cells with hygromycin was conducted. Untransfected control cells were treated with 1:1000 zeocin antibiotic. After 24 hours, when significant cell death of untransfected cells is observed, culture media was replaced with DMEM containing 10% FBS, 1% glutamine, 1% penicillin/streptomycin, and 2 mg/mL hygromycin. Remaining cells were grown to 100% confluence and re-passaged to about 2×10^6 cells/mL and maintained in multiple 100 mm cell culture plates.

Harvesting and purification of InR constructs expressed in Flp-In™ CHO cells:

Once cells reached 100% confluence (~ 3 days), the media was collected and stored at 4 °C. The cells were then washed with 1xPBS, lysed using lysis buffer (2 mM EDTA, 50 mM TRIS-HCl, 25 mM NaCl) and scraped off plates. Cells were further lysed by passing its contents through a 26-gauge syringe and needle. Lysate was centrifuged for 10 min at 9632 x *g* and the supernatant was collected.

His-tagged constructs were purified by batch binding to nickel sulfate charged IMAC resin (BioRad) overnight. Batch bound resin was loaded to gravity columns and washed with lysis buffer (2 mM EDTA, 50 mM TRIS-HCl, 25 mM NaCl,) with 5 mM imidazole. Constructs were eluted from column using lysis buffer with 250 mM imidazole.

Detecting protein expression using Western Blots. Western blot was conducted as previously described⁵³. InR constructs were diluted in 6x Laemmli loading buffer (9% SDS, 50% glycerol, 0.03% Bromophenol Blue, 375 mM Tris-HCl, 30% β-mercaptoethanol) and denatured at 95 °C for 5 minutes. SDS-PAGE using 10% gel was conducted and transferred paper onto nitrocellulose membrane (0.2 μm, BioRad) using chilled transfer buffer (25 mM Tris base, 190 mM glycine, 20% methanol, at pH 8.3) and a voltage of 100 V for 1 hour. The membrane was then blocked using sterile filtered (0.22 μm pore size) 2% bovine serum albumin (BSA) in TBS-T (20mM M Tris, 150 mM NaCl, 0.1% Tween-20) for 1 hour at room temperature on a shaker, and incubated in α-FLAG antibody diluted 1:1000 in blocking buffer overnight at 4 °C. The following day, the membrane was washed three times with TBS-T for 5 minutes each and incubated with 1:2000 α-mouse-HRP for 1 hour at room temperature. After a final four washes in TBST for 10 minutes each, membrane was imaged

using SuperSignal enhanced chemiluminescence substrate (ThermoFisher Scientific). Western blots were imaged on Amersham Hyperfilm™ ECL high performance film (GE Healthcare).

Protein expression and purification. pET28c plasmids with kanamycin resistance gene encoding Δ L1, Δ L1 FRET InR, and SNAP InR C5 constructs were subcloned using ligation independent cloning methods as previously described⁵⁴. InR C7 and P6C7 with N-terminal CFP and C-terminal YFP were cloned into a pRSET vector with carbenicillin resistance gene using Gibson Assembly 2x Hifi kit (New England BioLabs) as per manufactures instructions. InR constructs were transformed into BL21DE3 (Thermo Scientific™) chemically competent *E. coli* cells, plated into antibiotic-supplemented LB-agar plates, and incubated overnight at 37 °C. A single colony was then inoculated into a 50 ml LB seed culture supplemented with antibiotics (40 µg/ml carbenicillin and 50 µg/ml of kanamycin) and grown overnight at 37 °C on a shaker (225 rpm). 10 mL of seed culture was then added to a 500 mL antibiotic-supplemented LB enrichment culture, which was grown to optical density at 600 nm (OD_{600}) between 0.5-0.7. IPTG (1 mM for Δ L1 and Δ L1 FRET or 0.5 mM for all other InR constructs) was added to induce protein expression and incubated at 25 °C for 18 hours on a shaker (225 rpm). Cells were then pelleted by centrifugation at 9632 x *g* for 20 minutes and resuspended in lysis buffer. For Δ L1, Δ L1 FRET InR, and SNAP InR C5, the lysis buffer used was 50 mM NaH_2PO_4 , 300 mM NaCl at pH 8.0, supplemented with 10 mM β -mercaptoethanol and 1x HALT protease inhibitor. For compact InR FRET constructs, C7 and P6C7, the lysis buffer used was 20 mM Tris-HCl, 250 mM NaCl, pH 8, with 1x HALT protease inhibitor). Cells were lysed by sonication, the insoluble fractions were

pelleted by centrifugation at 24696 x *g* and stored at 4 °C for further refolding experiments where needed.

For refolding experiments, the pellets were further dissolved in 30 mL 8M urea for five days in a shaker at 4 °C. The suspension was then supplemented with 10 mM imidazole before purification. 10 mL of the mixture was used for IMAC purification as detailed below and the remaining 20 mL was used for VFD refolding and purification.

For soluble proteins, the supernatant after lysis and purification was affinity-purified using Profinity™ IMAC (BioRad, Hercules, CA) resin charged with nickel sulfate. For unfolded proteins, the denaturing protein solution was used for batch binding. The protein lysates were batch bound overnight to the charged IMAC resin and purified using gravity columns. Columns were washed with lysis buffer supplemented with 20 mM imidazole and the elution fractions were collected with lysis buffer containing 250 mM imidazole. For FRET experiments, the elution fractions were then buffer-exchanged and concentrated with lysis buffer lacking imidazole using Vivaspin® 20 Ultrafiltration Units (Sartorius) with a molecular weight cutoff of 10 kDa. The final buffer imidazole concentrations were calculated to be ~0.1 mM. Purified and buffer-exchanged protein fractions were then visualized using 10-12% SDS-PAGE with Coomassie dye staining.

Protein refolding using rapid dilution. Unfolded and purified InR constructs were concentrated to a final volume of 500 µl using a 30k molecular weight cut-off Vivaspin 20 ultrafiltration units as per manufacturer's instructions. 5 mL of refolding buffer (25 mM HEPES, 150 mM NaCl and 0.05% Triton X-100 at pH 7.8) was prepared in a microcentrifuge tube. A small stir bar was added to the bottom of the tube and was held up by ice in 50 mL

beaker on top of a stir plate. Small volumes of the unfolded sample was drawn up in a 26-gauge syringe. A single drop was added every 5 minutes to the refolding buffer until the whole volume of unfolded protein had been added and no apparent precipitate was observed. The 5 mL of refolded protein was then re-concentrated as described above and used for further analysis on an SDS-PAGE gel as well as in binding and FRET assays.

SNAP-tag bioconjugation. Stocks of benzylguanine (BG) tagged fluorophores, termed SNAP-Surface Alexa Fluor 546 and SNAP-Alexa Fluor 488 (New England Biolabs), were dissolved in 50 μ L DMSO to make 1 mM stock solutions. SNAP-Alexa Fluor 488 was further diluted 1:4 in DMSO to make a working solution of 250 μ M, as per manufacturer's instructions. InR constructs were buffer exchanged (25 mM HEPES, 300 mM NaCl, pH 8.0) to remove imidazole after purification. Approximately 1 μ M of purified InR constructs and the equivalent volume of purified untransfected cell lysate (negative control) was used in the conjugation reaction. 1mM DTT and 10 μ M of SNAP-Alexa Fluor 488 and 60 μ M of SNAP-Surface Alexa Fluor 546 (1:6 ratio) was added to the volume containing the InR construct. The total reaction volume was 150 μ L. The reaction tubes were protected from light and incubated at 4 °C on a shaker (150 rpm). The components were then dialyzed overnight at 4 °C in 4 L of reaction buffer to remove any unconjugated fluorophores.

Fluorescence Imaging of SDS-PAGE Gels. SDS-PAGE gels (0.75 mm thickness) were used to visualize InR constructs tagged with SNAP- with Alexa Fluor 488 and Alexa Fluor 546. Gels were imaged using a Typhoon Trio+ Scanner (GE Healthcare) in "fluorescence acquisition" mode. For Alexa Fluor 488 signal, the Green (532 nm) laser and the 526 SP filter at 500 PMT voltage was used. Alexa Fluor 546 fluorescent signal was detected using the

Green (532 nm) laser and 580 BP filter at 400 PMT voltage. Images were recorded at 25-micron pixel. The imaging of fluorescent signals total protein was visualized using Coomassie dye.

FRET measurements. For FRET measurements, 96-well black polystyrene plates (Corning) were used. InR constructs were diluted 1/10th (approximately 100 nM) for a final volume of 100 μ l. The protein dilution, z-position and gain were optimized for detection of acceptor and donor fluorophores. Insulin or Humalog® (Eli Lilly) at various concentrations was added in volumes of 0.5 μ l. Volume control negative controls using equal volumes of buffer were used. Each experimental group was repeated in triplicate. For Alexa Fluor 488 and 546 FRET pair, excitation of the donor fluorophore at 485 nm (20 nm bandwidth) with emission scans from 510 nm to 600 nm (10 nm bandwidth) were conducted. To determine emission from direct acceptor excitation 530 nm (10 nm bandwidth), emission scans from 555 nm to 600 nm (10 nm bandwidth) were conducted. For CFP and YFP FRET pairs, excitation of the donor fluorophore at 405 nm (20 nm bandwidth) with emission scans from 450 nm to 600 nm (10 nm bandwidth) were conducted. To determine emission from direct acceptor excitation 575 nm (10 nm bandwidth), emission scans from 515 nm to 600 nm (10 nm bandwidth) were conducted. Prior to adding insulin baseline measurements were recorded.

FRET measurements for refolded Δ L1 FRET. FRET assays for the Δ L1 FRET InR was conducted using an FL600 fluorescence plate reader with reader function controlled by KC4 data reduction software. Fluorescence was determined using a 485 nm bandpass emission filter and 530 nm, 25 nm bandpass emission filter with an instrument sensitivity setting of

135. Standard curves of eGFP were conducted to determine which concentration of Δ L1 FRET should be used in the assay. 1 μ M of Δ L1 FRET was prepared in buffer (25 mM HEPES, 150 mM NaCl, 0.05% Triton X-100, pH 7.8). Then, final concentrations of 0, 1, 10, 100 μ M of insulin or the negative control human serum albumin (HSA) was added in a total volume of 20 μ L to make the final well volumes 120 μ L. Emission spectra from 300 – 600 nm was collected from eBFP donor excitation at 385 nm, as well as direct eGFP acceptor excitation at 490 nm.

Two-protein FRET Assay. The two protein FRET assay involves the binding of the compact InR construct, P6C7 with an N-terminal CFP tag, and the insulin B-chain derivatives iPEP, iPEPA and iPEPS with a C-terminal YFP and FLAG tag. A negative control iFLAG, is composed of just the YFP with a C-terminal FLAG tag. For the FRET assay, control wells consisting of 1) 500 nM P6C7 alone, with PBS added as volume control, 2) 10 μ M iPEP, iPEPA, iPEPS and iFLAG alone without P6C7 and 3) PBS only well were used to determine background and signal bleed-through when donor (CFP) and acceptor (YFP) were individually excited. The negative control for inner filter effects and other non-FRET signals included 500 nM of P6C7 with 10 μ M of iFLAG. Finally, the experimental groups consisted of 500 nM of P6C7 with 10 μ M of either iPEP, iPEPA or iPEPS. FRET assays were conducted using a Tecan Spark plate reader (Tecan), optimal z-position for each plate was calculated, and the optimal gain was found to be 100 for CFP excitation and 70 for YFP excitation. The donor was excited at 405 nm with an excitation bandwidth of 7.5 nm, and emission spectra were collected from 450 nm to 600 nm. The acceptor was also directly excited at 475 nm with an excitation bandwidth of 7.5 nm, as emission spectra were collected from 515 nm to 600 nm. The final FRET signal was calculated such that the background of the iPEP and

iFLAG negative controls were subtracted from wells containing P6C7 and the signals are normalized to a percentage scale where the lowest mean value on the spectra has a value of 0 and the largest a value of 100.

Phage propagation and precipitation. InR constructs were subcloned into the pM1165a phagemid, which encodes an N-terminal FLAG-tag and C-terminal fusion to the P8 coat protein. Subcloned phagemids were transformed into F pilus positive, amber suppressing *E. coli* cell lines, ER2738 or XL-1 Blue. For constructs expected to dimerize, the amber stop codon, TAG, was added in place of a stop codon to allow expression on P8 unfused monomers that can dimerize with phage displayed constructs. Transformed plasmids were plated onto carbenicillin (50 µg/mL)-supplemented LB-agar plates and incubated at 37 °C overnight. A single colony from the plate was inoculated in a 4 mL primary culture 2YT supplemented with carbenicillin (50 µg/mL, phagemid selection) and tetracycline (2.5 µg/mL, cell line selection) and grown overnight at 37 °C in a shaker at 225 rpm. The following day, 100 µL of the primary culture was used to inoculate a second 4 mL of 2YT supplemented with carbenicillin and tetracycline. The secondary culture was grown to an OD₆₀₀ of 0.5 - 0.7. M13K07 helper phage, at multiplicity of infection (MOI) 4.6, was added to the culture and incubated and returned to the shaker for 1 hour. Next, the entire 4 mL culture was transferred into a secondary culture containing 150 mL 2YT supplemented with carbenicillin and kanamycin (40 µg/mL, M13K07 selection). The cultures were grown for 18 hours at 25 °C and 225 rpm.

In centrifuge tubes, the secondary culture was centrifuged at 9632 x *g* for 10 minutes at 4 °C. The supernatant was transferred into a separate tube of pre-filled with 60 mL of

chilled precipitation buffer (PEG-8000 (20%, w/v) and 2.5 M NaCl). The resulting solution was mixed by inverting the tube 10 times and left to incubate on ice for 30 minutes. The tubes were centrifuged again at $9632 \times g$ for 15 minutes at 4 °C. The phage was pelleted as a white streak and the supernatant was discarded. The tubes were re-centrifuged at $1,541 \times g$ for 4 minutes at 4 °C to condense the pellet. The supernatant was carefully removed and the pellet was resuspended in 3 mL PBS pH 8.0 with 10% glycerol and transferred into microcentrifuge tubes, where aliquots flash frozen using liquid N₂.

A second precipitation was conducted prior to use in binding assays. Aliquots from the first precipitation were thawed on ice and 0.2 mL of chilled precipitation buffer was added directly to the aliquot and gently mixed by inverting. The microcentrifuge tubes were then incubated for 30 minutes on ice before centrifugation at $9632 \times g$ for 20 minutes at 4 °C. The supernatant was discarded, and the tubes were re-centrifuged at $1,541 \times g$ for 4 minutes at 4 °C. After the removal of any remaining supernatant, the pellet was resuspended in 1 mL of 1xPBS pH 8.0. Concentration of the phage was calculated using the Jasco V-730 Spectrophotometer and a quartz cuvette (Starna Cells, Inc) to measure absorbance values at 268 nm and 320 nm (for baseline), where $(A_{268} - A_{320}) = 1$ is known to have a phage concentration of 8.31 nM or 5×10^9 phage/mL⁵⁵.

ELISAs for display and insulin binding. Phage display of ELISAs were conducted by detection of the N-terminal FLAG tag as previously described ⁵⁶. Insulin binding ELISAs were conducted using the configurations described below.

Immobilized insulin: 2 μM insulin was immobilized directly onto 96-well NUNC microtiter plates (Thermofisher) and blocked with 400 μL 0.02% casein blocking buffer in 1xPBS at room temperature with shaking at 150 rpm. Phage displayed InR constructs or His₆

affinity purified InR constructs were then added (100 μ L) at the indicated concentrations and incubated for 1 hour at room temperature with shaking at 150 rpm. Wells were then washed with 300 μ L 1xPBS-T (0.05% Tween-20) three times. Next, and then 100 μ L detection antibodies, α -M13-HRP (for phage constructs) or α -FLAG-HRP (purified protein), diluted 2000-fold in 1xPBS-T was added to each well and incubated at room temperature for 30 minutes with shaking at 150 rpm. Wells were washed thrice in 1xPBS-T and twice in 1xPBS. Binding signal was detected with 100 μ L TMB colorimetric HRP substrate (BioRad Inc.) at 652 nm or at 450 nm if the reaction was stopped using 2M H₂SO₄. Absorbance was measured using a UV-Vis plate reader (BioTek).

Biotinylated insulin in solution: Purified InR constructs were immobilized onto 96-well NUNC microtiter plates (Thermofisher) and blocked with 400 μ L 0.02% casein blocking buffer in 1xPBS at room temperature with shaking at 150 rpm. Indicated concentrations of biotinylated insulin (Nanocs Inc.) were added to each well (100 μ L) and incubated for 1 hour at room temperature with shaking at 150 rpm. Wells were washed three times in 1xPBS-T (300 μ L). Streptavidin-HRP (100 μ L), diluted 10,000-fold in 1xPBS-T, was used to detect insulin binding as described above.

Site-directed library cloning and preparation. The degenerate sequence (Table 1) was synthesized (Genewiz) as an oligomer with 5' and 3' non-degenerate flanking sequences and used to create the InR CR library using Gibson assembly. The synthesized dsDNA sequence was then amplified in a PCR reaction similarly to other fragments of the SNAP-C5 gene. The N- and C- terminal fragments of the SNAP-C5 gene were amplified using a previously cloned phagemid encoding a SNAP-tagged InR C5 construct as the template. The

SNAP-InRC5-CR library was also cloned into a phagemid vector used for p8-display. The PCR reaction was conducted using 5 U of Herculase II fusion DNA polymerase and buffer (Agilent), 25 mM dNTP (Thermofisher), 30 ng of DNA template, 12.5 nM forward (F) and reverse (R) primers. The PCR reaction was conducted by initially denaturing dsDNA at 95 °C for 2 minutes, then 30 cycles of denaturation at 95 °C for 20 seconds, annealing for 20 seconds at the temperatures shown in Table 1, and extension at 72 °C for 45 seconds, followed by a final extension at 72 °C for 3 minutes. 2uL of the PCR reaction was then visualized on a 1% agarose gel, the remaining was purified using the Zymo Clean and Concentrator® kit (Zymo Research). Purified PCR products were subcloned using 2x Hifi Gibson Assembly kit (New England Biosciences) as per manufacturer’s instructions.

Table 4-1. Gibson Assembly primers and template used for CR library Cloning

| Primer Name | Sequence | Annealing T (°C) |
|----------------------|--|------------------|
| CR library template | 5'-CAGGGCTGCCACCAG-BRTRWARWTSRTRAWRAWVAWTGTVWTCCGRAWTGCCGVRTRSCBRTRMARWG-AATTCCAGCAACTTGCTGTG-3' | |
| N-terminus SNAP-C5-F | 5'-GATAAGATTTTCAGGGCGCCACTGGACAAAGACTGCGAAATGAAG-3' | 53 |
| N-terminus SNAP-C5-R | 5'-GACGTAAGTGGTGGCAGCCCTGCCTCCGCGAGTTCTTGCATTTG-3' | 53 |
| Library CR- F | 5'-AATTCCAGCAACTTGCTGTGCACCCCATGCCTGGGTCCCT-3' | 49 |
| Library CR- R | 5'-GACCCGACGCGGTCCGGCTATGGCCTAGGGTCCCTCGGCAC-3' | 49 |
| C-terminus SNAP-C5-F | 5'-CACAAATGCAAGAAGCTCGCGGAGGCAGGGCTGCCACCAG-3' | 57 |
| C-terminus SNAP-C5-R | 5'-GACAGGGACCCAGGCATGGGGTGCACAGCAAGTTGCTGGAATT-3' | 57 |
| Phagemid vector- F | 5'-GTGCCGAGGACCCTAGGCCATAGCCGACCGCGTCCGGTC-3' | 58 |
| Phagemid vector- R | 5'-CATTTCCGAGTCTTTGTCCATTGGCGCCCTGAAAATCTTATCGTCG-3' | 58 |

The Gibson Assembly reaction (2 µL out of 20 µL) was transformed into Nova Blue cells. The transformation solution was then used to inoculate a 40 mL culture of carbenicillin (50µg/ml) - supplemented LB. The culture was grown overnight at 37 °C in a shaker at 225 rpm. The library pool was sent for sequencing and showed diversity in the library region. The miniprep kit (Qiagen) was then used to extract the plasmids. 20 µg of plasmid DNA was

then electroporated into electrocompetent XL-1 Blue cells using 2.5 kV field strength, 125 Ohms resistance, 50 μ F capacitance. This resulted in field peak voltages between 2380-2460 V and a time constant of 5.2 ms. Electroporated cells were then used to inoculate a 23 mL SOC media culture and shaken (225 rpm) in incubator for 20 minutes. A 50 μ L aliquot was then removed to determine titers, and 10^9 pfu of M13K07 helper phage was added to the culture and incubated in the same shaker for 1 hour. After the incubation, the helper phage-infected culture was added to a larger 500 mL of carbenicillin (50 μ g/ml) and kanamycin (40 μ g/ml) – supplemented 2YT culture and incubated overnight at 37 °C, shaking at 225 rpm for 14 hours. Electroporation of samples library samples directly after Gibson Assembly generated titers of 10^4 cfu/mL; therefore, this is the diversity of the library. After growth amplification and electroporation, the library was amplified 10^4 fold. Phage precipitation of the library was conducted as described above.

Phage selections. Selections for insulin-specific insulin binding were conducted using positive selection for insulin binding and negative selections for the insulin homolog, IGF-I. For round 1 of selections, 15 wells from two separate 96-well Nunc microtiter plate were coated with either 1.7 μ M insulin or 1.7 μ M IGF-I in 1xPBS and incubated overnight at 4 °C on a 150 rpm shaker. Wells were blocked at room temperature with 0.2% casein in 1xPBS for 30 minutes and then 100 μ L of 60 nM InR CR phage library was incubated with the IGF-I-coated plate for 1 hour at room temperature on a shaker at 150 rpm. Then, unbound phage library solution was collected, while the variants on the bound plate were discarded. The IGF-I unbound phage was then added to the insulin-coated plates (100 μ L) and incubated for 1.5 hours at room temperature, shaking at 150 rpm. The unbound solution was then discarded and wells were washed eight times in 1xPBS with 0.05% Tween-20. To

release bound phage from the plates, wells were incubated with 100 μ L of 0.1 M HCl and sonicated in a water bath for 10 minutes. The phage elute was collected from the wells and neutralized with 1M Tris, pH 8.0. Phage was then used to inoculate 22 mL growth culture of XL-1 Blue cells at O.D.= 0.5, in tetracyclin (5 μ g/mL)-supplemented 2YT and incubated for 45 minutes at 37 °C in a 225 rpm shaker. After incubation, 50 μ L of culture was removed for titration and M13K07 helper phage was added to the culture at a multiplicity of infection (MOI) value of 4.6. The culture was returned to the incubator for an additional 45 minutes. Then, the 22 mL culture of infected cells were used to inoculate 500 mL 2YT culture supplemented with carbenicillin (50 μ g/ml) and kanamycin (40 μ g/ml), which was incubated for 18 hours at 37 °C, shaking at 225 rpm. The following day, the phage was precipitated as described above and used for a second round of selections, which only included positive selection for insulin binding.

The second round was conducted as described above, except 0.2% BSA in 1xPBS was used for blocking, wells were incubated with insulin for 1 hour and washed 12 times after the insulin binding step. Individual colonies from the second round of selections were grown in 96-well deep well plates and infected with M13K07. The phage solutions were assayed on insulin or BSA (negative control)-coated Nunc plates for specific insulin binding in an ELISA format, as described above. The fold increase in binding signal of insulin over BSA was calculated and variants with signal > 4 fold was considered a hit and sequenced.

4.6 References

1. Wherrett, D. K. *et al.* Defining Pathways for Development of Disease-Modifying Therapies in Children With Type 1 Diabetes: A Consensus Report. *Diabetes Care* **38**, 1975–1985 (2015).
2. Todd, J. A. Etiology of Type 1 Diabetes. *Immunity* **32**, 457–467 (2010).
3. Bluestone, J. A., Herold, K. & Eisenbarth, G. Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature* **464**, 1293–1300 (2010).
4. Atkinson, M. a., Eisenbarth, G. S. & Michels, A. W. 3R: Type 1 diabetes. *Lancet* **383**, 69–82 (2014).
5. García-Pérez, L. E., Álvarez, M., Dilla, T., Gil-Guillén, V. & Orozco-Beltrán, D. Adherence to therapies in patients with type 2 diabetes. *Diabetes Therapy* **4**, 175–194 (2013).
6. X., D., Z.-C., L., L., Z., W.-P., Z. & F., H. Artificial Pancreas as an Effective and Safe Alternative in Patients with Type 1 Diabetes Mellitus: A Systematic Review and Meta-Analysis. *Diabetes Ther.* **9**, 1269–1277 (2018).
7. Boughton, C. K. & Hovorka, R. Is an artificial pancreas (closed-loop system) for Type 1 diabetes effective? *Diabet. Med.* **36**, 279–286 (2019).
8. Bekiari, E. *et al.* Artificial pancreas treatment for outpatients with type 1 diabetes: Systematic review and meta-Analysis. *BMJ (Online)* **361**, (2018).
9. Thabit, H. *et al.* Home Use of an Artificial Beta Cell in Type 1 Diabetes. *N. Engl. J. Med.* **373**, 2129–2140 (2015).
10. Bekiari, E., Karagiannis, T., Haidich, A. B. & Tsapas, A. Meta-analysis of artificial pancreas trials: methodological considerations. *The Lancet Diabetes and Endocrinology* **5**, 685 (2017).
11. Georga, E. I., Protopappas, V. C., Bellos, C. V. & Fotiadis, D. I. Wearable systems and mobile applications for diabetes disease management. *Health and Technology* **4**, 101–112 (2014).
12. Scholler, P. *et al.* HTS-compatible FRET-based conformational sensors clarify membrane receptor activation. *Nat. Chem. Biol.* **13**, 372–380 (2017).
13. Shigeto, H. *et al.* Insulin sensor cells for the analysis of insulin secretion responses in single living pancreatic β cells. *Analyst* **144**, 3765–3772 (2019).
14. Malkoc, A. *et al.* Enhancing Glycemic Control via Detection of Insulin Using Electrochemical Impedance Spectroscopy. *J. Diabetes Sci. Technol.* **11**, 930–935 (2017).
15. Xu, M., Luo, X. & Davis, J. J. The label free picomolar detection of insulin in blood serum. *Biosens. Bioelectron.* **39**, 21–25 (2012).
16. Luo, X., Xu, M., Freeman, C., James, T. & Davis, J. J. Ultrasensitive label free electrical detection of insulin in neat blood serum. *Anal. Chem.* **85**, 4129–4134 (2013).
17. Yu, Y., Guo, M., Yuan, M., Liu, W. & Hu, J. Nickel nanoparticle-modified electrode for ultra-sensitive electrochemical detection of insulin. *Biosens. Bioelectron.* **77**, 215–219 (2016).
18. Chhasatia, R. *et al.* Non-invasive, in vitro analysis of islet insulin production enabled by an

- optical porous silicon biosensor. *Biosens. Bioelectron.* **91**, 515–522 (2017).
19. Wang, Y. *et al.* A near infrared fluorescence resonance energy transfer based aptamer biosensor for insulin detection in human plasma. *Chem. Commun.* **50**, 811–813 (2014).
 20. Scapin, G. *et al.* Structure of the insulin receptor-insulin complex by single-particle cryo-EM analysis. *Nature* **556**, 122–125 (2018).
 21. Croll, T. I. *et al.* Higher-Resolution Structure of the Human Insulin Receptor Ectodomain: Multi-Modal Inclusion of the Insert Domain. *Structure* **24**, 469–476 (2016).
 22. De Meyts, P. *et al.* Insulin and IGF-I receptor structure and binding mechanism. in *Mechanisms of Insulin Action: Medical Intelligence Unit* (2007). doi:10.1007/978-0-387-72204-7_1
 23. De Meyts, P. & Whittaker, J. Structural biology of insulin and IGF1 receptors: Implications for drug design. *Nature Reviews Drug Discovery* **1**, 769–783 (2002).
 24. Brandt, J., Andersen, A. S. & Kristensen, C. Dimeric Fragment of the Insulin Receptor α -Subunit Binds Insulin with Full Holoreceptor Affinity. *J. Biol. Chem.* **276**, 12378–12384 (2001).
 25. Surinya, K. H. *et al.* Role of insulin receptor dimerization domains in ligand binding, cooperativity, and modulation by anti-receptor antibodies. *J. Biol. Chem.* **277**, 16718–16725 (2002).
 26. Ward, C. W., Menting, J. G. & Lawrence, M. C. The insulin receptor changes conformation in unforeseen ways on ligand binding: Sharpening the picture of insulin receptor activation. *BioEssays* **35**, 945–954 (2013).
 27. Elleman, T. C. *et al.* Mutational analysis of the N-linked glycosylation sites of the human insulin receptor. *Biochem.J.* **347 Pt 3**, 771–779 (2000).
 28. Heleen, L., Caro, P., Ohali, A., Gorden, P. & Collier, E. Mutational analysis of the NH₂-terminal glycosylation sites of the insulin receptor α -subunit. *Diabetes* **43**, 240–246 (1994).
 29. Menting, J. G. *et al.* How insulin engages its primary binding site on the insulin receptor. *Nature* **493**, 241–245 (2013).
 30. Menting, J. G., Ward, C. W., Margetts, M. B. & Lawrence, M. C. A thermodynamic study of ligand binding to the first three domains of the human insulin receptor: Relationship between the receptor α -chain C-terminal peptide and the Site 1 insulin mimetic peptides. *Biochemistry* **48**, 5492–5500 (2009).
 31. Puche, J. E. & Castilla-Cortázar, I. Human conditions of insulin-like growth factor-I (IGF-I) deficiency. *Journal of Translational Medicine* **10**, (2012).
 32. Bajar, B. T., Wang, E. S., Zhang, S., Lin, M. Z. & Chu, J. A guide to fluorescent protein FRET pairs. *Sensors (Switzerland)* **16**, (2016).
 33. Berney, C. & Danuser, G. FRET or no FRET: A quantitative comparison. *Biophys. J.* **84**, 3992–4010 (2003).
 34. Kristensen, C., Wiberg, F. C., Schäffer, L. & Andersen, A. S. Expression and characterization of a 70-kDa fragment of the insulin receptor that binds insulin. Minimizing ligand binding domain of the insulin receptor. *J. Biol. Chem.* **273**, 17780–17786 (1998).

35. Schaefer, E. M., Siddle, K. & Ellis, L. Deletion analysis of the human insulin receptor ectodomain reveals independently folded soluble subdomains and insulin binding by a monomeric α -subunit. *J. Biol. Chem.* **265**, 13248–13253 (1990).
36. Paul, J. I., Tavare, J., Denton, R. M. & Steiner, D. F. Baculovirus-directed expression of the human insulin receptor and an insulin-binding ectodomain. *J. Biol. Chem.* **265**, 13074–13083 (1990).
37. Lou, M. *et al.* The first three domains of the insulin receptor differ structurally from the insulin-like growth factor 1 receptor in the regions governing ligand specificity. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12429–12434 (2006).
38. Maggi, D., Laurino, C., Andraghetti, G. & Cordera, R. The overexpression of insulin receptor makes CHO cells resistant to the action of IGF-1: Role of IRS-1. *Biochem. Biophys. Res. Commun.* **205**, 693–699 (1994).
39. Kim, M. S. & Lee, G. M. Use of Flp-mediated cassette exchange in the development of a CHO cell line stably producing erythropoietin. *J. Microbiol. Biotechnol.* (2008).
40. Sparrow, L. G. *et al.* The disulfide bonds in the C-terminal domains of the human insulin receptor ectodomain. *J. Biol. Chem.* **272**, 29460–29467 (1997).
41. Schäffer, L. & Ljungqvist, L. Identification of a disulfide bridge connecting the α -subunits of the extracellular domain of the insulin receptor. *Biochem. Biophys. Res. Commun.* **189**, 650–653 (1992).
42. Gutmann, T. *et al.* Cryo-EM structure of the complete and ligand-saturated insulin receptor ectodomain. *J. Cell Biol.* **219**, (2020).
43. Holleman, F. & Hoekstra, J. B. L. Insulin Lispro. *N. Engl. J. Med.* **337**, 176–183 (1997).
44. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
45. Hajduczki, A., Majumdar, S., Fricke, M., Brown, I. A. M. & Weiss, G. A. Solubilization of a membrane protein by combinatorial supercharging. *ACS Chem. Biol.* **6**, 301–307 (2011).
46. Ogata, A. F. *et al.* Virus-enabled biosensor for human serum albumin. *Anal. Chem.* **89**, 1373–1381 (2017).
47. Algar, W. R., Hildebrandt, N., Vogel, S. S. & Medintz, I. L. FRET as a biomolecular research tool — understanding its potential while avoiding pitfalls. *Nat. Methods* **16**, 815–829 (2019).
48. Stefanidis, L., Fusco, N. D., Cooper, S. E., Smith-Carpenter, J. E. & Alper, B. J. Molecular Determinants of Substrate Specificity in Human Insulin-Degrading Enzyme. *Biochemistry* **57**, 4903–4914 (2018).
49. Ruan, W. *et al.* Interaction between IGFBP7 and insulin: A theoretical and experimental study. *Sci. Rep.* **6**, 1–9 (2016).
50. Medintz, I. L., Goldman, E. R., Lassman, M. E. & Mauro, J. M. A fluorescence resonance energy transfer sensor based on maltose binding protein. *Bioconjug. Chem.* **14**, 909–918 (2003).
51. Dhatariya, K. Blood ketones: Measurement, interpretation, limitations, and utility in the management of diabetic ketoacidosis. *Review of Diabetic Studies* **13**, 217–225 (2016).

52. Kim, D. H. *et al.* Real-time Monitoring of Biomarkers in Serum for Early Diagnosis of Target Disease. *Biochip Journal* **14**, 2–17 (2020).
53. Sen, S. *et al.* Retinoblastoma protein modulates the inverse relationship between cellular proliferation and elastogenesis. *J. Biol. Chem.* **286**, 36580–36591 (2011).
54. Aslanidis, C. & de Jong, P. J. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res.* **18**, 6069–6074 (1990).
55. Mohan, K., Penner, R. M. & Weiss, G. A. Biosensing with Virus Electrode Hybrids. *Curr. Protoc. Chem. Biol.* **7**, 53–72 (2015).
56. Sen, S. R. *et al.* Predicting COVID-19 Severity with a Specific Nucleocapsid Antibody plus Disease Risk Factor Score. *mSphere* **6**, e00203-21 (2021).

CHAPTER 5

Photostable and Proteolysis-Resistant Förster Resonance Energy Transfer-Based Calcium Biosensor

Reprinted (adapted) with permission from Nguyen, D. *et al.* Photostable and Proteolysis-Resistant Förster Resonance Energy Transfer-Based Calcium Biosensor. *Anal. Chem.* **92**, 7683–7689 (2020). Copyright 2020 American Chemical Society.

Author contributions: Dat Nguyen and Elliot L. Botvinick designed research, analyzed data, and wrote the manuscript. Dat Nguyen, Danielle M. Behrens, Avid Najdahmadi, and Micah M. Lawrence performed FRET and biomedical engineering research. Dat Nguyen, Sanjana Sen, Jessica N. Pham, and Gaetano Speciale, Micah M. Lawrence performed protein expression, purification and bioconjugation research and analyzed data. Sudipta Majumdar and Gregory A. Weiss designed research and analyzed data. Sanjana Sen, Jessica N. Pham, Sudipta Majumdar and Gregory A. Weiss partially wrote manuscript.

5.1 Abstract

Molecular sensors from protein engineering offer new methods to sensitively bind to and detect target analytes for a wide range of applications. For example, these sensors can be integrated into probes for implantation, and then yield new and valuable physiological information. Here, a new Förster Resonance Energy Transfer (FRET)-based sensor is integrated with an optical fiber to yield a device measuring free Ca^{2+} . This membrane-encapsulated optical fiber (MEOF) device is composed of a sensor-matrix that fills polytetrafluoroethylene (PTFE) with an engineered Troponin C (TnC) protein fused to a pair of FRET fluorophores. The FRET efficiency is modulated upon Ca^{2+} ion binding. The probe further comprises a second, size-excluding filter-membrane that is synthesized by filling the pores of an PTFE matrix with a polyethylene glycol dimethacrylate (PEGDMA) hydrogel; this design ensures protection from circulating proteases and the foreign body response. The two membranes are stacked and placed on a thin, silica optical fiber for optical excitation and detection. Results show the biosensor responds to changes in Ca^{2+} concentration within minutes with a sensitivity ranging from 0.01 to 10 mM Ca^{2+} , allowing discrimination of hyper

and hypocalcemia. Furthermore, the system reversibly binds Ca^{2+} to allow continuous monitoring. This work paves the way for the use of engineered structure-switching proteins for continuous optical monitoring in a large number of applications

5.2 Introduction

The development of continuous biomedical sensors provides clinicians and researchers real-time data on clinically relevant and new physiological signals.¹ Currently, the catalog of continuous sensors is vastly outweighed by the number of clinically relevant analytes, which are largely analyzed with point-of-care (POC) devices or at clinical laboratories. For example, ionized calcium (Ca^{2+}), an essential mineral for muscle contraction, bone development, nerve impulses, blood clotting, and regulating heart beat propagation, is assayed by a calcium blood test; this test requires a healthcare professional to draw blood from a patient's median cubital vein and send it to a clinical laboratory for a complete metabolic panel analysis.^{2,3} The time between depositing a sample and receiving results may be several hours, or approximately one hour in emergency cases. To eliminate the latency caused by hospital lab delay, POC devices such as Abbott I-STAT can perform on-site assays, including Ca^{2+} , providing results within a few minutes. However, the frequency of assay is still dependent upon typically infrequent, professional blood draws. Though laboratory assays of Ca^{2+} are precise and accurate, the measurements are intermittent as compared to physiological Ca^{2+} dynamics. For example, in clinical cases, such as rapid blood transfusion during liver transplantations, Ca^{2+} concentrations can exhibit rapid transients at very low concentrations (e.g., drops by 0.1 mM Ca^{2+} in 5 min), underlying the need for a continuous Ca^{2+} sensor.^{2,3}

Advances in protein engineering have yielded new classes of binding macromolecules that display exquisite ligand binding specificity and yield quantifiable signals upon such ligand or target binding.^{4,5} For example, Maeshime, et al. developed a Förster Resonance Energy Transfer (FRET)-based molecular Mg^{2+} sensor to monitor Mg^{2+} dynamics during the cell cycle. This sensor comprises the structure-switching *Escherichia coli* (*E. coli*) Mg^{2+} transporter CorA (CorA-CD) positioned between the fluorescent proteins (FPs) enhanced cyan fluorescent protein (eCFP) and Venus (a variant of yellow fluorescent protein, YFP).⁶ Furthermore, a genetically encoded calcium indicator, termed Cameleon, has been reported; this sensor joins structure-switching calmodulin (CaM) and CaM-binding peptides of Myosin Light-Chain Kinase positioned between eCFP and eYFP to monitor intracellular calcium.⁷ These FRET sensors demonstrate the specificity and sensitivity of genetically encoded protein biosensors.

Here, the calcium FRET sensor Twitch-2B is considered for use in a continuous Ca^{2+} monitor. Twitch-2B was developed by Thestrup et al. for intracellular monitoring of Ca^{2+} transients during T-lymphocyte activation and action potential propagation in mouse cortical neurons.⁸ Twitch-2B is derived from toadfish *Opsanus tau* Troponin C (TnC), a muscular actin-associated protein that undergoes structure-switching upon Ca^{2+} binding. Twitch-2B comprises a modified TnC (equilibrium dissociation constant for Ca^{2+} , $K_D = 200$ nM) space with linkers, each fused at their free ends to the FPs mCerulean3 (cyan FP variant) and cpVenus^{cd} (yellow FP variant), at the N- and C-termini, respectively.⁸ Twitch-2B was determined to be a candidate sensing molecule for a continuous physiological Ca^{2+} probe because of its reversible binding kinetics, stability in vivo, and sensitivity to varying Ca^{2+} concentrations.

A number of calcium sensing modalities have been developed to monitor calcium. Asif et al. developed an electro-chemical sensor to Ca^{2+} comprising functionalized biocompatible ZnO nanorods. In vitro testing shows a log-linear relationship between sensor voltage and Ca^{2+} ranging from 100 nM to 10 mM.⁹ Shortreed et al. functionalized the distal end of an optical fiber with the calcium sensitive dye Calcium Green and reported a unique emission spectrum for Ca^{2+} concentrations ranging from 37.6 nM to 39.8 μM .¹⁰

These reported strategies lack a method to prevent interactions with physiological macromolecules, including antibodies, proteases, and other soluble proteins, upon device implantation. Proteins from the foreign body response (FBR) can foul the surface and adversely affect sensing for in vivo applications.¹¹ In consideration of an implantable Ca^{2+} sensor, a new type of Ca^{2+} sensor is presented that combines a FRET-based sensing molecule with a new membrane to provide the requisite protection for in vivo applications.

An optical fiber device was developed and referred to here as the membrane-encapsulated optical fiber (MEOF, Figure 5-1). MEOF first comprises a hydrophilic polytetrafluoroethylene (PTFE) sheet impregnated with a solution of a protein FRET sensor and placed in contact with the end of an optical fiber. MEOF also includes an outer membrane composed of a size-selective polyethylene glycol dimethacrylate (PEGDMA) 2000 hydrogel interpenetrating a second PTFE sheet. Results are presented for Ca^{2+} -sensing MEOF wherein the FRET sensor is a new photo-resilient variant of Twitch-2B, SNAP-Twitch-2B-CLIP (STC), which replaces the readily photobleached FPs with photostable Alexa Fluor 488 and Alex Fluor 546 as the donor and acceptor fluorophore, respectively.¹³ This Ca^{2+} probe shows a reversible response curve for concentrations ranging between 0.01 mM to 10 mM, and is

photostable over prolonged excitation scans, a necessary capability for future implantation. The probe also operates in real-time, which is critical for a continuous Ca^{2+} biosensor.

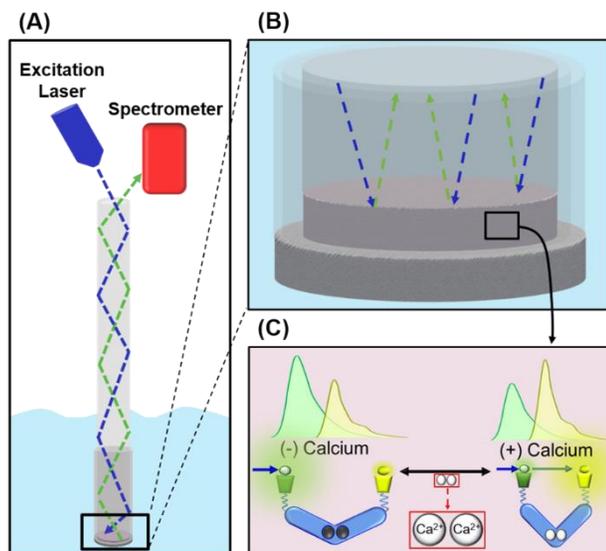


Figure 5-1. Schematic of the FRET calcium ion probe. **A.** Excitation light results in a FRET signal analyzed by a spectrometer. **B.** A size-excluding filter-membrane and sensor-matrix encapsulate the optical fiber. **C.** The FRET signal is modulated by calcium ion concentration.

5.3 Experimental Section

Filter-Membrane Fabrication: Monomeric PEGDMA 2000 precursor (Sigma Aldrich) and Milli-Q water were mixed with a flea-sized micro stirring bar (Grainger) in a 30%/70% (w/w) ratio in a dark room until a homogenous solution was obtained. Photoinitiator 2-Hydroxy-4'-(2-hydroxyethoxy)-2-methylpropiophenone (Irgacure 2959, Sigma Aldrich) was then mixed at 1% (w/w) and illuminated by a UV (365 nm), 8 W VWR Dual UV Transilluminator for 10 min to form a hydrogel. Subsequently, filter-membranes were fabricated by polymerizing the hydrogel precursor solution while in PTFE (FHLC01300, 80% porosity, 0.45 μm pore size, 145 μm thick, Sigma Aldrich).

Permeability Testing: Permeability of hydrogels was tested using 3 mm diameter disks of PEGDMA 2000. The hydrogels were first allowed to swell overnight in Phosphate Buffered Saline Solution (PBS, pH 7.4, Thermo Fischer Scientific). Swelling reaches a steady state in 5 h (Supporting Information, Figure 5-10). The swollen hydrogels were then placed in a custom-made rubber chamber (frame size: 20 mm x 20 mm, mat opening: 15 mm x 15 mm, height: 1 mm McMaster-Carr), and incubated in PBS solutions containing Alexa Fluor 488 - labeled insulin (5.8 kDa, Supporting Information) or Precision Plus Protein Kaleidoscope Prestained Protein Standards (Bio-Rad; 10 kDa to 250 kDa) for 30 min.¹² Confocal microscopy was performed using either a Leica Sp8 or Zeiss LSM700 inverted microscope using 488 and 633 nm laser lines and a 10x or 20x objective lens. Emission was collected from 520 to 700 nm and 650 to 700 nm under 488 and 633 nm laser line excitation, respectively. Images were acquired every 30 s.

The permeability of the filter-membranes was tested by first creating 1 mm PTFE disks with a 1 mm biopsy punch (World Precision Instruments). The disks were then glued (Loctite 4981, Henkel) to the face of MRE-025 microrenathane tubing (ID: 0.30 mm, OD: 0.63 mm, Braintree Scientific) and the hydrogel precursor solution was polymerized within the pores of the PTFE. A bare optical fiber (FG200LEA, 0.22 NA, 200 μm core, 220 μm cladding, Thor Labs) was guided down the tubing and placed in contact with the membrane. Loctite 4981 glue was used to seal the backend of the silica optical fiber to the microrenathane tubing. The distal end of each optical fiber terminated in an SMA thread and was coupled to a custom benchtop spectrometer apparatus (Figure 5-11). A set of membrane-fiber units were fabricated for each test solution described below. After incubation in PBS, the tip of each optical fiber construct was immersed in a solution containing fluorescein (FITC, 389.4

Da, Sigma Aldrich), green fluorescent protein (GFP, 26.9 kDa), Alexa Fluor 488 - labeled insulin (5.8 kDa) or fluorescein isothiocyanate-dextran (3-5 kDa or 150 kDa, Sigma Aldrich).^{13,14}

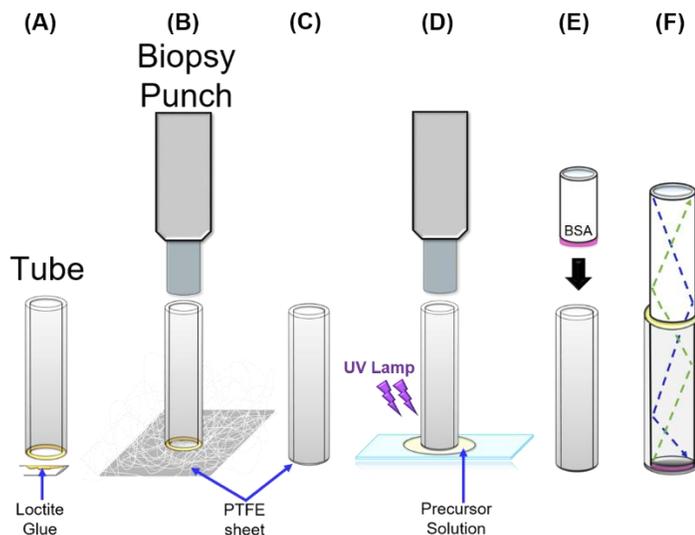


Figure 5-2. Fabrication of the membrane-encapsulated optical fiber (MEOF). **A.** Microrenathane tubing is cut, and Loctite 4981glue is applied by a small needle to its tip. The tubing is placed onto a sheet of PTFE cut by biopsy punch **B.** to produce a capped tube **C.** PEGDMA 2000 hydrogel precursor solution is allowed to impregnate the PTFE sheet **D.** such that after polymerization, a hybrid PEGDMA2000/PTFE membrane is produced. An optical fiber guides an PTFE sheet, wetted with FRET protein solution (shown in pink), to the end of the tubing **E.** and then sealed on the back end with Loctite 2981 glue **F.**

5.4 Fabrication of the Membrane-Encapsulated Optical Fiber (MEOF)

MEOF probes were fabricated as illustrated in Figure 5-2. Microrenathane tubing was cut into 5 mm segments (Figure 5-2A), and the glue was then applied to the bottom annulus of the tubing (Figure 5-2A) to adhere the tubing to a PTFE sheet (Figure 5-2B). A 1 mm biopsy punch was slipped over the tubing to cut the PTFE sheet into a circle of 1 mm diameter. (Figure 5-2C). Next, the hydrogel precursor solution was pipetted on a glass slide and the PTFE-capped tubing was pressed into the solution. The hydrogel precursor solution

permeates throughout the PTFE due to their compatible hydrophilicity and was polymerized with a 365 nm 8 W VWR Dual UV Transilluminator for 10 min, producing a membrane-capped tube (Figures 5-2D). A second hydrophilic PTFE sheet was wetted with 2 μ L of FRET sensor solution (4 mg/mL) by first pipetting the FRET protein solution onto the PTFE material and then compressing this sheet between two microscope slides to form the sensor-matrix. In parallel, a silica optical fiber was incubated overnight in a 2% bovine serum albumin (BSA, Chromatopur) /water solution to block the optical fiber's surface and prevent surface-induced aggregation of the FRET sensor. Next, the sensor-matrix was placed onto the tip of the BSA-coated optical fiber (Figure 5-2E) and the assembly was pushed through the tubing until it contacted the filter-membrane (Figure 5-2F). Glue was used to seal the backend of the silica optical fiber to the microrenathane tubing, and the resulting MEOF was incubated in protein buffer media at 4 °C for at least 5 h. The assembly quality was assessed by MEOFs encapsulating GFP to monitor for leakage. These MEOFs were incubated in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer for 5 h (Figure 5-12), and the solution was tested for leaking GFP using a fluorescent plate reader (Spark, TECAN). No signal intensity was observed, confirming encapsulation of the protein within the sensor-matrix.

FRET Calcium Sensing: Each spectral measurement was acquired in triplicate with exposure times of 100 ms for the Twitch-2B sensor or 2 s for the STC sensor. Back-ground (dark signal)-subtracted spectra were processed in MATLAB. Custom MATLAB code was used to process the digitized signals and calculate the FRET ratiometric signal, which here is defined as the peak acceptor intensity signal divided by the peak donor intensity signal ($I_{\text{acceptor}}/I_{\text{donor}}$), a metric commonly used to quantify FRET efficiency.¹⁵

Calcium solutions ($\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, Sigma Aldrich) at the indicated concentration were placed in microcentrifuge tubes capped with parafilm to prevent evaporation. MEOF probes were tested by dipping the devices into these calcium solutions. GraphPad Prism was used for to fit data points onto nonlinear regression models and for statistical analysis.

5.5 Characterization of the Filter-Membrane

A well-known event, the FBR activates upon device implantation.¹¹ Of particular importance to the MEOF Ca^{2+} probe, the FBR recruits immune cells that secrete proteases, which could cleave the Twitch-2B and render the device inoperable. Thus, the device must exclude protease access to the molecular sensor.

Here, PEG hydrogels are formulated to size-exclude proteases secreted by immune cells. Such proteases are considerably smaller than immunoglobulins (~ 150 kDa), and much larger than calcium ions.¹⁶ The smallest of these known proteases is Chymase, which has a MW of approximately 30 kDa.¹⁷ PEG-derived hydrogels were selected for protein exclusion due to the tunability of their mesh size, swelling, and tensile strength.^{18,19} Additionally, PEG is considered to be generally nonfouling and biocompatible.²⁰ Screening of various PEG formulations and derivatives lead to the selection of PEGDMA 2000 hydrogels, formulated as described in the Experimental Section, and were shown to exclude proteins as small as 10 kDa, but permit transport of insulin (5.8 kDa). Fluorescence signals from the fluorescently labeled insulin at the hydrogel/fluid interface show very rapid kinetics (Figure 5-3A,5-3C); in comparison, the 10 kDa and larger proteins show no penetration of the hydrogel as measured by fluorescence (Figure 5-3B, 5-3C).

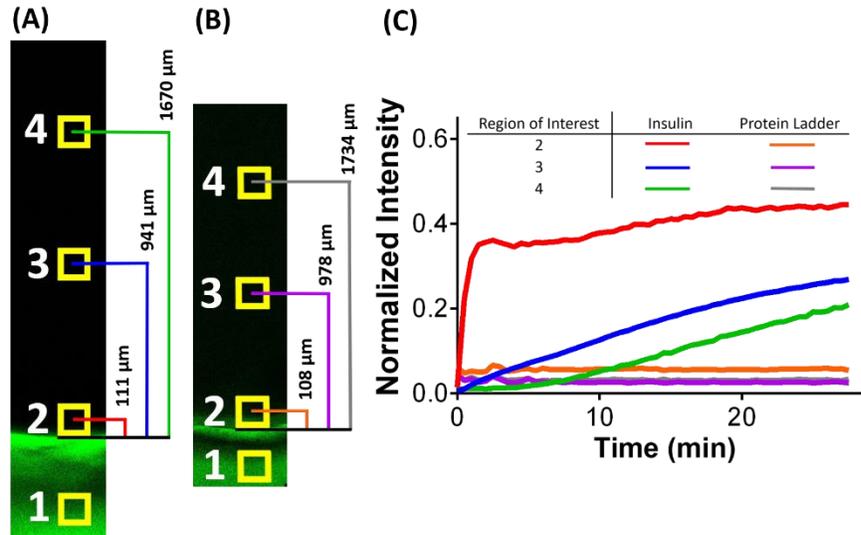


Figure 5-3. Examining the permeability of the PEGDMA 2000 hydrogel. **A.** Confocal micrographs of the PEGDMA 2000 hydrogel surrounded by Alexa Fluor 488-labeled insulin (20X image) or **B.** a protein ladder (Bio-Rad Protein Kaleidoscope; 10X image) were captured at $t = 0$. **C.** Intensity profiles highlight permeabilization by proteins into the indicated regions of interests (ROIs) over time with normalization for fluorescence in ROI 1. ImageJ software was used to analyze ROI intensity.

The PEGDMA 2000 hydrogels are very soft when hydrated, however and too fragile to place on the tip of a probe for implantation. Thus, a protective PTFE layer was devised. The hydrogel precursor solution was polymerized within the non-solid phase of commercially available hydrophilic PTFE sheets. Scanning electron microscopy (FEI Magellan 400 XHR, Electron Microscopy Solutions) after iridium sputter coating (10nm, EMS 150T, Electron Microscopy Solutions) reveal the natural porosity of PTFE sheets (Figure 5-4A) and their impregnation by the PEGDMA 2000 hydrogel (Figure 5-4B). This composite, termed the filter-membrane, combines the size-exclusion properties of the PEGDMA 2000 hydrogel with the durability of PTFE sheets, and offers a robust composite material for the tip of the Ca^{2+} sensing device.

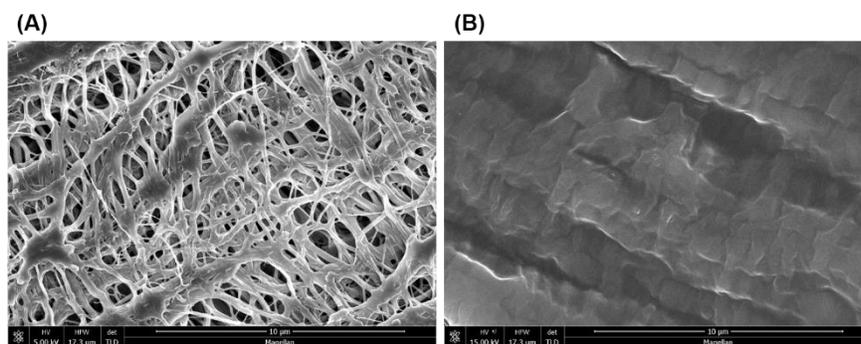


Figure 5-4. SEM images of PTFE and the filter-membrane. Scale Bar = 10 µm. A. The PTFE sheets contain 0.45 µm pores that are infiltrated with PEGDMA 2000 hydrogels B. after UV polymerization. Filter-membranes were tested to ensure they retain the size selectivity of the hydrogel alone. For each test, a filter-membrane was coupled to the end of an optical fiber as described in the Experimental Section. Figure 5-5 reports that FITC, 3-5 kDa fluorescein isothiocyanate-dextran and Alexa Fluor 488-labeled insulin can diffuse through the filter-membrane, while large macromolecules GFP and 150 kDa fluorescein isothiocyanate-dextran cannot. Figures S4A and S4B display the time-dependent emission spectra of Alexa Fluor 488 – labeled insulin and GFP. These results show the filter-membrane retains the size-exclusion properties of the hydrogel.

| | Fluorescein | Dextran | Insulin | GFP | Dextran |
|------------------------|-------------|---------|---------|------|---------|
| Molecular Weight (kDa) | 0.38 | 3-5 | 5.8 | 26.9 | 150 |
| Membrane Permeable? | Yes | Yes | Yes | No | No |

Figure 5-5. Properties of the filter-membrane. Tabulated permeability results indicate the observed penetration by fluorescently labeled analytes.

5.6 Twitch-2B Ca²⁺ Probe

Calcium sensing membrane-encapsulated optical fiber (MEOF) biosensors were fabricated as detailed in the Experimental Section, where the FRET molecular sensor was Twitch-2B. Important to the FRET sensor, the sensor-matrix allows Twitch-2B to remain in solution, preserving Twitch-2B structure-switching upon Ca²⁺ binding and unbinding. The resulting MEOF probe has the following key properties. 1) enabling the passage of small target analytes (e.g., ions and small peptides) yet exclusion of larger proteins (such as serine

proteases), 2) allowing FRET sensors to undergo conformational changes upon binding and releasing target analytes, and 3) guiding excitation light to the sensor-matrix and fluorescently emitted light to the spectrometer.

In Vitro Testing of MEOF Ca²⁺ Probes

To determine the efficacy of the Twitch-2B MEOF Ca²⁺ probe, probes were incubated in test buffer (25 mM 3-(N-morpholino) propanesulfonic acid (MOPS), 100 mM KCl, pH 7.6) supplemented with increasing concentrations of Ca²⁺. For each measurement, 0.7 μW of 405 nm laser light was used to excite the protein sensor for 100 ms. Results show the FRET ratio ($I_{\text{acceptor}} / I_{\text{donor}}$) being sensitive to increasing Ca²⁺ concentration from 0 to 14 mM, above which the sensor response saturates (Figure 5-6A). Importantly, the MEOF sensor responds in the physiological concentration range of Ca²⁺; normal non-protein bound Ca²⁺ concentration in human blood plasma ranges from 1.1 to 1.3 mM with pathological values ranging from 0.33 mM (hypocalcemia) to 2.5 mM (hypercalcemia).^{11,21,22} The probe-response curve was fit to a Hill equation, which determined the K_D and Hill constant to be 2.3 mM and 0.78, respectively (R^2 (Coefficient of Determination) = 0.99, $Sy.x$ (Standard Error of the Estimate) = 0.07) consistent with binding affinity in a useful range. These results demonstrate that the intracellular FRET sensor Twitch-2B can be adapted to function in a hand-held optical device.

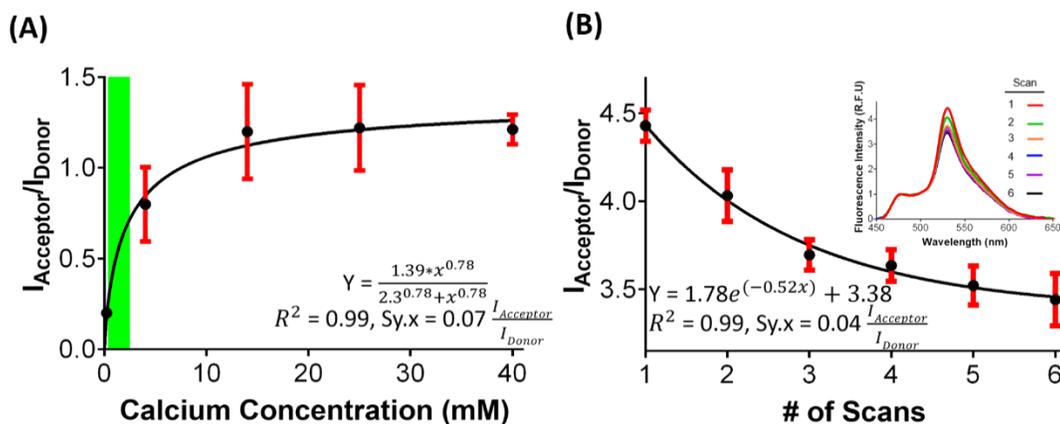


Figure 5-6. Efficacy of Twitch2B to sense calcium ions. A. The dose-dependent response of the MEOF and its data fit to a Hill equation demonstrates the MEOF's usefulness for sensing physiological concentrations of Ca^{2+} (green). $I_{\text{acceptor}}/I_{\text{donor}}$ is the emission ratio after subtracting the spectra for calcium ion-free MOPS buffer. **B.** Photobleaching of mCerulean3 and cpVenus^{cd} in calcium-free MOPS buffer is significant, and renders Twitch2B inappropriate for a MEOF calcium ion monitor; the inset depicts the emission spectra normalized by mCerulean3 peak intensity.

However, the Twitch-2B acceptor fluorophore, cpVenus^{cd} (a YFP derivative) is susceptible to photobleaching, particularly as compared to the donor fluorophore mCerulean3 (a CFP derivative).²³ The photostability of the Twitch-2B MEOF device was therefore investigated by repeated exposure to 405 nm while in calcium-free test buffer. The FRET ratio was measured after each exposure and showed significant photobleaching over time (Figure 5-6B), following an exponential decay in FRET ratio ($R^2 = 0.99$, $Sy.x = 0.04$), as expected for exponential photobleaching.^{24,25} Further, it has been shown that the CFP-YFP FRET pair may suffer from phototoxicity when excited with violet wavelengths, resulting in inactivation of the FRET pair.²³ To account for possible effects specific to the MEOF, photobleaching experiments were conducted using solutions of Twitch-2B. Twitch-2B solutions were exposed to 45 mW, 405 nm laser light for three min; a non-illuminated

solution served as a negative control. Solutions were analyzed with the spectroscopic fluorescent plate reader under three conditions: donor excitation in test buffer, acceptor excitation in test buffer, and donor excitation in test buffer supplemented with CaCl₂. Spectroscopic data (Figure 5-14) confirm photobleaching and consequential loss of sensitivity to Ca²⁺.

5.7 The Development of a Photostable Twitch-2B FRET Construct

To overcome cpVenus^{cd} photobleaching during MEOF measurements, photostable fluorescent molecules were sought to replace the mCerulean3-cpVenus^{cd} FRET pair. Small molecule Alexa Fluor dyes are highly suitable candidates for the MEOF FRET sensor as they provide greater photostability and brightness compared to genetically encoded fluorescent proteins and other fluorescent dyes like Cy3 and fluorescein.^{26,27} To conjugate Alexa Fluor dyes, genetically-encoded SNAP- and CLIP-tags were integrated into Twitch-2B. SNAP and CLIP-tags, both derived from the human O⁶-alkylguanine transferase; catalyze the covalent attachment of an O⁶-benzylguanine (O⁶-BG) and O⁶-benzylcytosine (O⁶-BC), respectively.²⁸

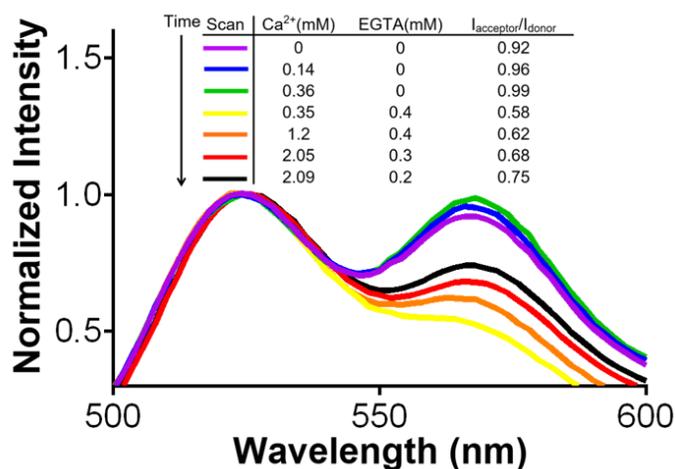


Figure 5-7. STC Ca²⁺ response in a fluorescence plate reader normalized by Alexa Fluor 488 peak intensity. The FRET ratio response is sensitive to increasing concentrations of

Ca²⁺. The Twitch-2B sensor was reengineered with SNAP and CLIP tags flanking its N- and C termini, resulting in, SNAP-Twitch-2B-CLIP, termed STC. The plasmid encoding STC and plasmid map are shown in Figure 5-15 and 5-16. The new construct was expressed at high yields and purified using affinity and size exclusion chromatography (Figure 5-17A). Commercially available O⁶-BG Alexa Fluor 488 and O⁶-BC-Alexa Fluor 546-BC (New England Biolabs) were covalently attached onto the SNAP and CLIP tags of STC (Figure 5-17B, 5-17C), and the FRET construct retained its dose-dependent response to Ca²⁺ in a 96-well fluorescent plate assays (Figure 5-7). The STC construct was, henceforth, used in further investigations of the MEOF probes.

5.8 Characterization of MEOF Probes with the STC Construct

Here, probe photostability was assessed using a 450 nm laser light to excite the donor fluorophore Alexa Fluor 488. Each MEOF was incubated in Reference Buffer (RB, 50 mM HEPES, 50 mM NaCl, 10% glycerol, 10 mM β -mercaptoethanol (β ME), pH 7.6) and exposed to 0.71 mW laser for 2 s per reading. Laser light energy per measurement is 20E3 times greater for STC as compared to Twitch-2B. No significant FRET ratio change was observed during the experiment (Figure 5-8A, Relative Standard Deviation, RSD = 0.57%). Thus, integration of the new donor and acceptor fluorophores provide sufficient photostability for continuous Ca²⁺ monitoring.

The sensor rise time was assessed by first incubating MEOF probes in RB and then in RB supplemented with 1.25 mM CaCl₂. The FRET ratio was recorded every 3 min for 30 min after media exchange. The data follow an exponential plateau function ($R^2 = 0.98$, $S_{y,x} = 0.004$) with a rise time of approximately 6 min, without detectable photobleaching within the plateau (RSD = 0.47%) (Figure 5-8B). The 6 min rise time is similar to that of commercial analytical devices, for example, continuous glucose monitors, known to be effective in guiding treatment of conditions having rapid and slow transients such as seen in type 1 diabetes.²⁹

Next, MEOF probes were incubated in RB solutions having Ca^{2+} concentration ranging from 0 to 20 mM. The data shows sensitivity to Ca^{2+} across the measured range (0.01 mM to 10 mM), and a functional form consistent with the Hill equation (Figure 5-8C; $R^2 = 0.99$, $S_{y,x} = 0.01$) having a Hill constant and K_D values of 0.4 and 0.1 mM, respectively. As indicated by the green region in Figure 5-8C, the MEOF is sensitive across the physiological range of ionized Ca^{2+} concentration in the blood, and extends the dynamic range of the Ca^{2+} sensing modalities previously mentioned.¹¹⁻¹³ This extended sensitivity range correlates with the Hill constant being less than one, affording a stretched 'linear region' of sensitivity.³⁰

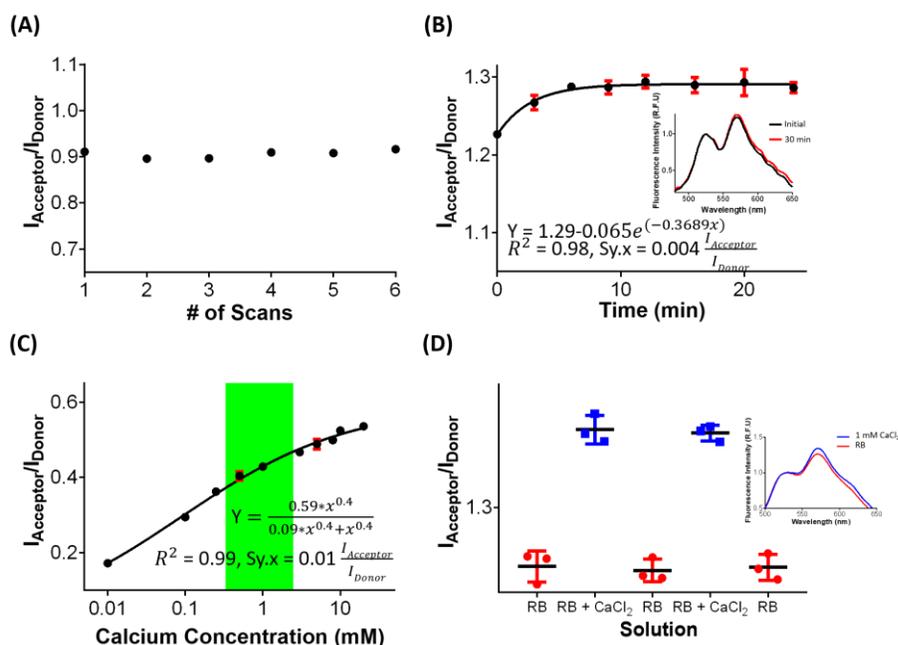


Figure 5-8. Ca^{2+} sensing with the MEOF probe. **A.** The STC exhibits excellent stability in reference buffer. **B.** Steady state measurements are observed 6 min after the addition of calcium ions; the inset depicts the emission spectra normalized by donor peak intensity. **C.** The STC exhibits a dose-dependent response, and the data fits to a Hill equation. $I_{\text{acceptor}}/I_{\text{donor}}$ is emission ratio after subtracting spectra in calcium ion-free RB. **D.** The MEOF device has excellent reproducibility after serial exposures to either buffer supplemented with Ca^{2+} (1 mM) or RB.

To test for sensor reversibility, MEOF probes were serially placed in RB solutions with or without 1 mM CaCl₂ (Figure 5-8D;). A one-way ANOVA was conducted to compare effects of Ca²⁺ on FRET ratio. There was a significant effect of Ca²⁺ across all groups ($P < 0.0001$). The Tukey post-hoc comparison with adjusted P values shows no significant differences between the three RB solution groups or the two RB + 1 mM CaCl₂ groups ($p \varepsilon 0.99$ for each comparison), but significant differences between each pairing of RB or RB + 1 mM CaCl₂ groups ($p < 0.0001$ for each comparison). Notably, sensor dynamics are considerably slower when comparing Ca²⁺ unbinding (Figure 5-18) to binding (Figure 5-8B). Although the cytoplasmic Twitch-2B has previously demonstrated rapid unbinding kinetics in cellulo, cells contain orders of magnitude lower calcium concentration as compared to the blood and contain unique pumps and transporters that rapidly reduce cytoplasmic Ca²⁺ concentration.³¹ And so, it is unsurprising that MEOF fall-times are relatively long given the passive diffusion of Ca²⁺ ions at concentrations mimicking the blood.

5.9 Protection from Proteolytic Enzymes

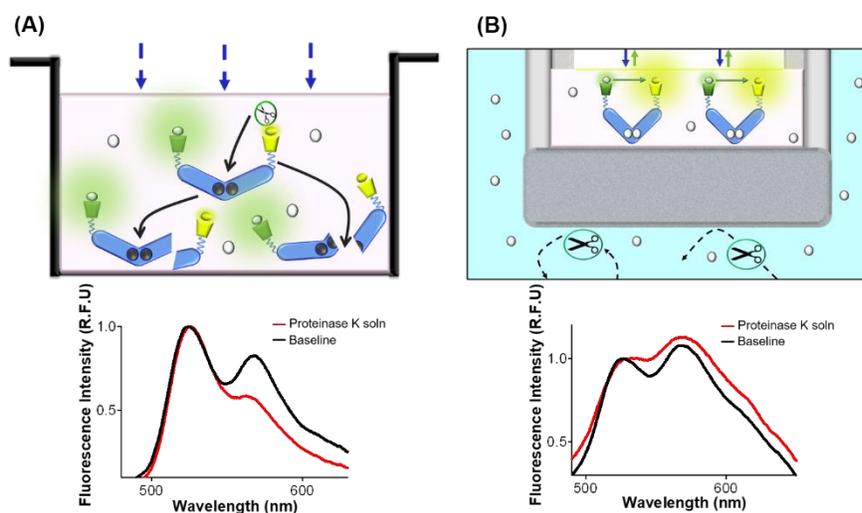


Figure 5-9. Sensor Isolation from Proteinase K. **A.** Proteinase K cleaves STC resulting in a diminished acceptor intensity signal. **B.** However, the MEOF successfully protects the protein sensor as demonstrated by the observed increase in FRET efficiency.

To model immune system proteolytic enzymes secretion during the FBR, the STC MEOF was incubated with the endopeptidase serine protease Proteinase K (28.9 kDa, New England Biolabs).³² The solution comprises 692 μM proteinase K in a solution of 20 mM Tris-HCl, 1 mM CaCl_2 and 50% glycerol by volume. Emission spectra confirm that proteinase K in the Ca^{2+} solution rapidly cleaves the STC and results in a 26% loss in acceptor fluorescence (Figure 5-9A). Conversely, uncleaved STC should show an increase in acceptor fluorescence over time due to FRET. This increase is observed for the MEOF incubated in the proteinase K solution for 30 min, demonstrating that the filter-membrane successfully excludes the proteinase K (Figure 5-9B). These findings highlight the need for and the key role of the size-excluding filter-membrane to circumvent the FBR.

5.10 Conclusion

This effort has produced an optical fiber-based probe, MEOF, designed for prolonged measurements of Ca^{2+} concentrations in vivo. Such capabilities result from (1) encapsulating protein FRET sensors within the pores of a mechanically robust PTFE sheet, such that FRET sensors are free to undergo conformation changes, (2) excluding host proteases by a hybrid PTFE-PEGDMA membrane, and (3) elimination of deleterious photobleaching of FPs in the Twitch-2B FRET sensor by installing small molecules for the donor and acceptor FPs, respectively. This approach is generalizable to the larger class of protein FRET sensors (e.g. ATP and Zn^{2+}).^{33,34} Further, the molecular weight cut off properties of the hybrid filter-membrane can be tuned for other applications or FRET sensors by new hydrogel formulation

comprising PEG derivatives, or other commonly used hydrogels such as alginate and hyaluronic acid. In summary, the MEOF probe can be easily adopted to include new FRET sensors to generate a new class of analytical probes.

5.11 Materials and Methods

Expression and Purification of SNAP-Twitch-2B-CLIP (STC). Plasmids encoding STC were transformed into chemically competent BL21(DE3)* *E. coli* cells. After the cells were transferred to LB plates supplemented with carbenicillin (0.050 mg/mL), the cells were incubated at 37 °C for 12-18 h. An individual colony from the transformation plate was inoculated with LB (70 mL) supplemented with carbenicillin (0.050 mg/mL). The flasks were incubated at 37 °C with shaking at 220 rpm for 12-18 h. This seed culture (15 mL) was transferred into LB (1 L) supplemented with carbenicillin (0.050 mg/mL) before incubation at 37 °C with shaking at 220 rpm. Upon reaching a cell density of OD₆₀₀ between 0.50 and 0.56, the culture was induced through addition of IPTG (0.5 mM) before incubation at 25 °C with shaking at 225 rpm for 6 h. The cultures were centrifuged at 6084.1 rcf at 4 °C for 20 min. The supernatant was discarded, and the cell pellet was resuspended in lysis buffer (20 mL of 50 mM HEPES buffer, 300 mM NaCl, 10 mM β-mercaptoethanol (βME), pH 7.6). A protease inhibitor cocktail (HALT) was added to the suspended cells at the manufacturer's recommended concentration, and the cells were lysed by sonication at a 50% amplitude pulsation (1 s on and 1 s off) for a 1-min cycle and a 2-min rest between each cycle, for a total of 5 cycles. The lysate was centrifuged at 26891.1 rcf at 4 °C for 45 min. The supernatant was loaded onto a nickel-charged microporous polymethacrylate resin (MIDA, Purolite). The supernatant underwent batch binding with the charged resin overnight in 4 °C with a slow

tilt rotation. After batch binding, the mixture was homogenized by inversion and applied to a gravity column. Flow through, wash, and elution fractions were collected using the lysis buffer, lysis buffer with imidazole (40 mM), and lysis buffer with a higher concentration of imidazole (250 mM), respectively. The fractions containing the desired protein was further purified using fast protein liquid chromatography (FPLC) using a Superdex™ 75 10/300 GL column (GE Healthcare). The column was washed with 1 volume of 20% ethanol and 1 volume of double-distilled water and pre-equilibrated with 1 column volume of a size exclusion reference buffer (RB, 50 mM HEPES, 50 mM NaCl, 10% glycerol, 10 mM β ME, pH 7.6) at a flow rate of 0.2 mL/min. The protein was eluted with the RB at 0.2 mL/min. A 12% polyacrylamide SDS-PAGE gel confirmed the homogeneity of the protein.

Bioconjugation of Fluorophores onto STC. At 37 °C for 1 h in the dark, the purified protein was incubated with BG-Alexa Fluor 488 and BC-Alexa Fluor 546 dyes (New England Biolabs) in a 1:2:2 molar ratio (protein:BG-dye:BC-dye) in DMSO with the addition of dithiothreitol (DTT, 1 mM). Unbound dye was removed by microfiltration of the solution with a 10 kDa molecular weight cut off (MWCO) microfilter (Amicon Ultra-0.5 Centrifugal Filter Unit) at 17136 rcf at 4 °C for 5 min.

Expression and Purification of Twitch-2B. Using BL21(DE3) (Thermo Scientific™) chemically competent *E. coli* cells, the miniprepmed CFP-Twitch2B-YFP plasmids were heat shock transformed. The cells were plated on warm LB agar plates supplemented with carbenicillin (50 mg/mL) and placed in a 37 °C incubator to grow for 12-18 h. Seed cultures were made by picking single colonies from the transformation plates and inoculating into a culture tube with 5 mL LB and 5 μ L of carbenicillin (50 mg/mL). The tubes were left to grow

at 37 °C, shaking at 220 rpm, for 6 h. An expression culture was made by adding the seed culture in 1 L of LB and 1 mL of carbenicillin (50 mg/mL). The expression culture was left to grow at 37 °C, shaking at 220 rpm, until the OD₆₀₀ is at 0.6. Once it has reached this OD₆₀₀, the culture was induced with isopropyl-β-D-thio-galactoside (IPTG, 1 mM). The expression culture was moved to a pre-cooled 25 °C incubator for induction at 225 rpm for 12 to 18 h.

The culture was removed from the incubator and centrifuged at 6084.1 rcf at 4 °C for 20 min to harvest the cells. The cells were resuspended in 20 mL of lysis buffer (25 mM MOPS buffer, 100 mM KCl, pH 7.6). A protease inhibitor cocktail (HALT) was added to the suspended cells at the manufacturer's recommended concentration. The suspended cells were lysed by sonication at a 50 % amplitude pulsation (1 second on and 1 second off) for a 1-min cycle and a 2-min rest between each cycle, with a total of 3 cycles. The lysate was centrifuged at 26891.1 rcf at 4 °C for 45 min. The supernatant was loaded onto a nickel charged immobilized metal affinity chromatography (IMAC) resin (1.5 mL of IMAC charged with NiSO₄ and equilibrated with the lysis buffer) The supernatant underwent batch binding to the charged IMAC resin overnight in 4 °C with constant shaking.

After batch binding, the IMAC/supernatant mixture was homogenized by inversion and loaded into a gravity column. The following fractions were collected: a flow through fraction (FT), a wash fraction using a solution of lysis buffer with 25 mM imidazole (W), and elution fractions using a solution of lysis buffer with 250 mM imidazole (E1-5). SDS-PAGE confirmed the expression of the protein and its purity. Ethylenediaminetetraacetic acid (EDTA, 0.5 M) was added to the eluted fractions containing the protein of interest and the

mixture was dialyzed in 4 L of lysis buffer overnight and concentrated to approximately 1 mL.

Tagging Insulin with N-Hydroxysuccinimide Ester (NHS)-Alexa Fluor 488. In order to label insulin with Alexa Fluor 488, Alexa Fluor 488-NHS ester (Thermo Fisher Scientific) and recombinant human insulin were combined at final concentrations of 20 μ M and 84 μ M, respectively, in a total volume of approximately 180 μ L. The volume was split into 30 small reactions and incubated in a thermocycler for 1 h at 4 °C. The total volume of the reaction was serially dialyzed for 3 h and then overnight in 2 L of PBS pH 8.0, such that unbound, hydrolyzed NHS-Alexa Fluor 488 could be removed from the labelled insulin solution.

Fluorescence Imaging of SDS-PAGE Gels. Tagging of SNAP- and CLIP-tags with Alexa Fluor 488 and Alexa Fluor 546 was visualized by imaging the fluorescently tagged proteins on a 0.75 mm thickness SDS-PAGE using a Typhoon Trio+ Scanner (GE Healthcare) in “fluorescence acquisition” mode. Fluorescent signals from Alexa Fluor 488 was imaged using the Green (532 nm) laser and the 526 SP filter at 500 PMT voltage, whereas for detecting Alexa Fluor 546, the Green (532 nm) laser and 580 BP filter at 400 PMT voltage was used. Images were recorded at 25-micron pixel size for high resolution scanning. Once fluorescent signals were imaged, the total protein was visualized using Coomassie dye.

Fluorescent Intensity Measurements by TECAN Spark Plate Reader. In a Corning black polystyrene 96 well plate, 100 μ L of buffer (either TB or RB) was used to fill the wells unless otherwise stated. 4 μ L of protein sensor solution was added to each well, and 405 nm and 450 nm excitation light were used excite the donor fluorophores on Twitch-2B and STC,

respectively. Emission spectra from 450 to 600 nm and 500 to 600 nm were collected for Twitch-2B and STC. Excitation and emission bandwidths were set to 5 nm.

Determining Protein Concentration using the Bradford Assay. Protein concentration was determined with the Bradford protein assay. BSA standards and the protein samples were introduced to Coomassie Brilliant Blue G-250 dye (Bio-Rad) for protein binding in 96 well Costar plate (SKU#3595), which undergoes a colorimetric change. Absorbance is measured at 595 nm, and the protein concentration is determined through interpolation of the standard curve.

Supporting Figures

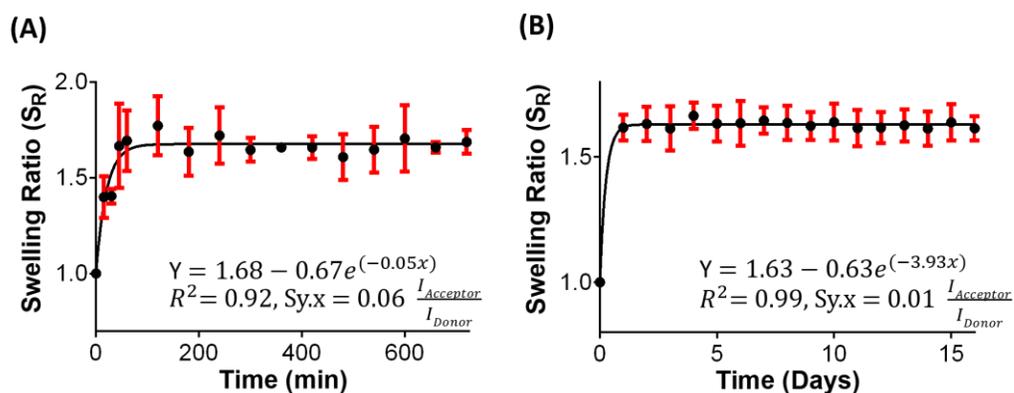


Figure 5-10. Swelling studies of PEGDMA 2000.

PEGDMA 2000 hydrogels were allowed to swell in PBS, pH 7.4 at room temperature to reach a steady state mass. To quantify the degree of swelling, the swelling ratio of the hydrogel was calculated. The formula is shown in equation 1

$$S_R = \frac{W_s}{W_i} * 100 \quad \text{Equation 1}$$

where S_R is the hydrogel swelling ratio, W_s is the hydrogel swell mass after incubation in PBS, pH 7.4 at room temperature and W_i is the initial hydrogel mass right after polymerization.³⁵ The results indicate that **A.** the hydrogel reaches a steady state mass after 5 h of incubation after which **B.** there is no change in S_R for up to two weeks (relative standard deviation (RSD) = 0.87%).

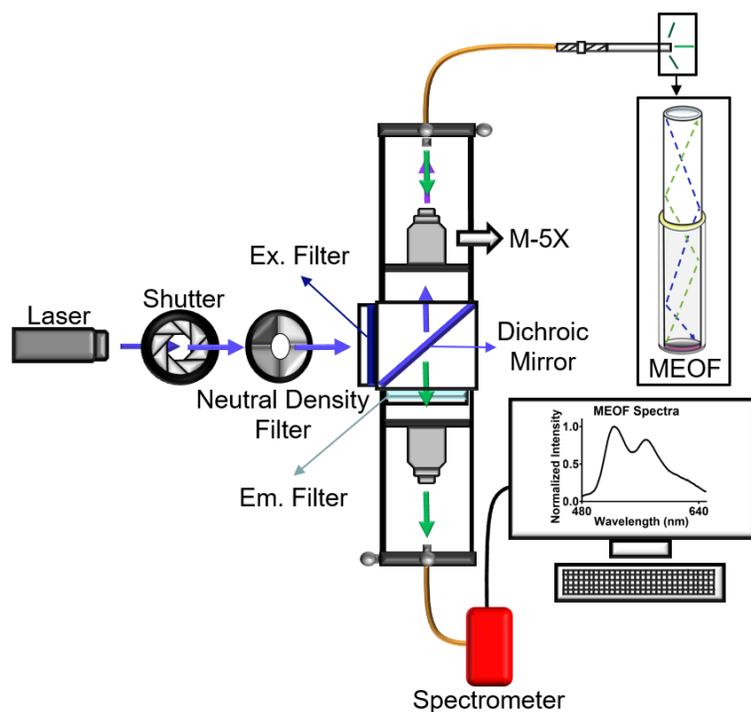


Figure 5-11. Spectroscopic Benchtop System. A benchtop optical system was constructed for MEOF excitation and emission collection. The system was integrated into a cage system to maintain fiber alignment. In the case of STC, excitation light is provided by a 450 nm continuous wave laser (4.5 mW, CPS450 ThorLabs) that is mechanically shuttered (CX2450B, NM Laser Product). The beam passes through a variable neutral density filter wheel (NDC-25C-4, Thorlabs) to reduce power in the beam. The beam then passes through a filter cube block (CM1-DCH with attached SM30 circular filter mounts tubes, Thorlabs), comprising a 450 nm bandpass filter (ET448/19x, Chroma) tuned to the laser wavelength, a dichroic beam splitting mirror with center wavelength at 458 nm (ZT458rdc, Chroma), and a 465 nm long pass filter (AT465lp, Chroma) in the emission pass to filter out any back-reflected laser light. Laser light reflected off the dichroic mirror is focused by a M-5x microscope objective lens (0.10 NA, Newport) into a SMA-SMA optical fiber patch cable (M92L01, Thorlabs). This optical fiber is attached to a SMA female plate mounted within a XY translation stage (ST1XY-S Translation Mount, Thorlabs) for fine X-Y alignment of the optical fiber relative to the focused laser beam. The MEOF is coupled to the distal end of the patch cable by a SMA to SMA mating sleeve (ADASMA, Thorlabs). Fluorescence emission is collected by the MEOF and collimated into our cage system, passing through the dichroic mirror and the 465 nm longpass filter, where it is coupled into an optical fiber. This coupling uses the same strategy as for the laser light into the MEOF system. The distal end of the optical fiber directs light into a spectrometer (CCS200, Thorlabs). ThorLabs software collects spectrometer data that is subsequently analyzed in MATLAB (MathWorks). In the case of the Twitch-2B construct, the laser, excitation, dichroic, and emission filters were replaced with a 405 nm continuous wave laser (Newport LQA305-40P laser, 45 mW maximum power), FBH405-10, DMLP425R, and FELH0450 filters (Thorlabs) respectively.

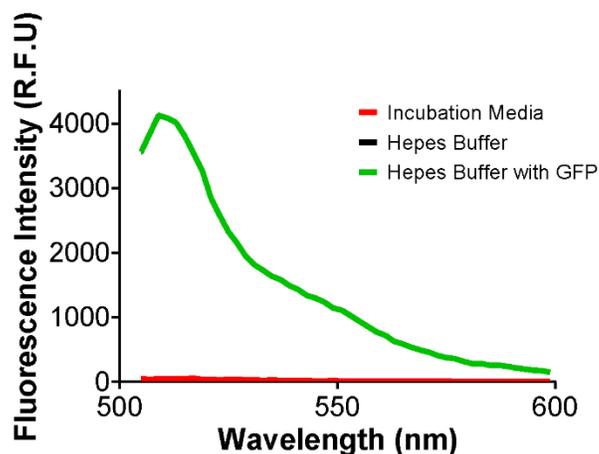


Figure 5-12. GFP encapsulation within the MEOF. GFP was loaded into the sensor-matrix of the MEOF probe and incubated in HEPES buffer for 5 h. 150 μ L of the HEPES buffer was then analyzed with a fluorescence plate reader (red). For the experiment, negative and positive controls include 150 μ L of HEPES buffer without GFP (black) and 150 μ L of HEPES buffer spiked with 0.1 μ L of GFP at 4 mg/mL (green). The spectral curves demonstrate GFP encapsulation within the MEOF.

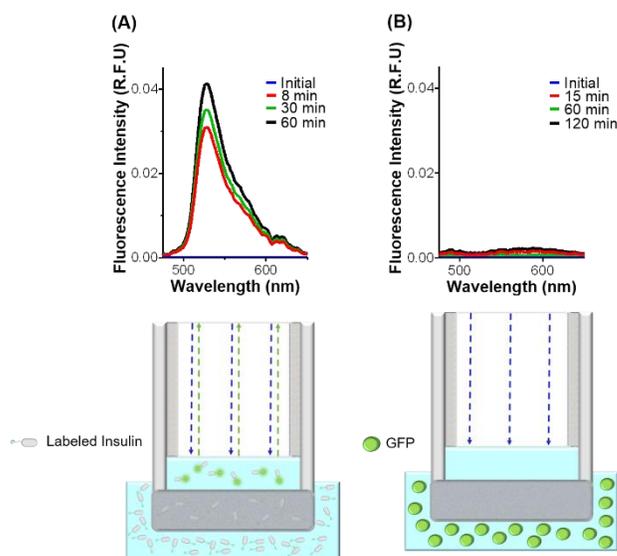


Figure 5-13. Selective permeability of the filter-membrane. The spectral curves demonstrate GFP encapsulation within the MEOF. Diffusion spectra (top) and schematic representation (bottom) of fluorescently labeled (A) insulin and (B) GFP demonstrate selective permeability by the MEOF.

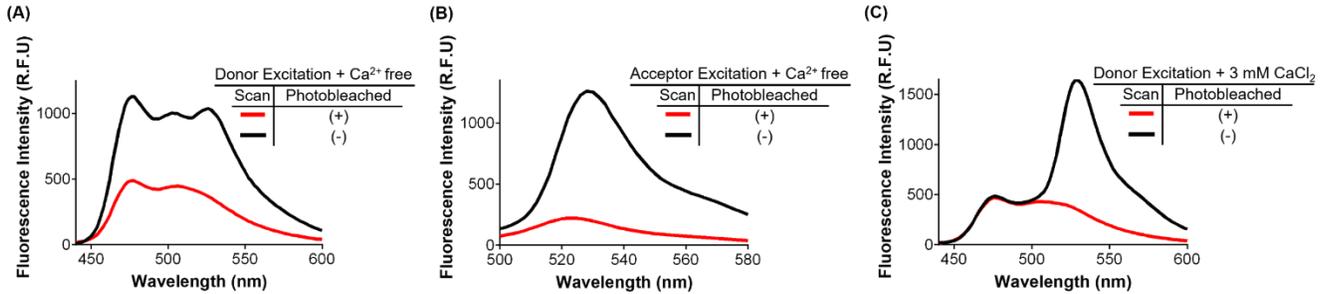


Figure 5-14. FRET properties of non-photobleached and photobleached Twitch2B samples. Both solutions contained 150 μ L of Twitch2B protein (4 mg/mL). For the photobleached samples, both **A.** donor excitation at 405 nm and **B.** direct acceptor excitation at 475 nm have an absent acceptor fluorescence signal. **C.** Moreover, donor excitation with the introduction of Ca²⁺ does not increase the FRET efficiency in the photobleached sample.

ATGCGGGTCTCATCATCATCATCATGGTATGGACAAAGACTGCGAAATGAAGCGCACCAC
 CCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCGAACAGGGCCTGCACCGTATCATCT
 TCCTGGGCAAAGGAACATCTGCCGCCGACGCCGTGGAAGTGCCTGCCCCAGCCGCCGTGCTGGGC
 GGACCAGAGCCACTGATGCAGGCCACCGCCTGGCTCAACGCCTACTTTCACCAGCCTGAGGCCAT
 CGAGGAGTTCCCTGTGCCAGCCCTGCACCACCCAGTGTTCAGCAGGAGAGCTTTACCCGCCAGG
 TGCTGTGAAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCTACAGCCACCTGGCCGCC
 CTGGCCGGCAATCCCGCCGCCACCGCCGCCGTGAAAACCGCCCTGAGCGGAAATCCCGTGCCCAT
 CTGATCCCTGCCACCGGTGGTGCAGGGCGACCTGGACGTGGGGGGCTACGAGGGCGGGCTCGC
 CGTGAAAGAGTGGCTGCTGGCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGGTGCGATGC
 AAGTTGCCGACGCCAGCGAAGAGGAACTGAGCGAGTGCTTCAGAATCTTCGACTTCGACGGCAAC
 GGCTTCATCGACAGAGAGGAGTTTGGCGACATCATCAGACTGACCGGCGAGCAGCTGACCGACGA
 GGACGTGGACGAGATCTTCGGGCACTCCGACACCGACAAGAACGGCAGAATCGATTTGACGAGT
 TCCTGAAGATGGTGGAAAACGTGCAGCCATCTACCCGAGCTCGACAAAGACTGCGAAATGAAG
 CGCACCACCTGGATAGCCCTCTGGGCAAGCTGGAAGTGTCTGGGTGCGAACAGGGCCTGCACCG
 TATCATCTTCCTGGGCAAAGGAACATCTGCCGCCGACGCCGTGGAAGTGCCTGCCCCAGCCGCC
 TGCTGGGCGGACCAGAGCCACTGATCCAGGCCACCGCCTGGCTCAACGCCTACTTTCACCAGCCT
 GAGGCCATCGAGGAGTTCCCTGTGCCAGCCCTGCACCACCCAGTGTTCAGCAGGAGAGCTTTAC
 CCGCCAGGTGCTGTGAAACTGCTGAAAGTGGTGAAGTTCGGAGAGGTCATCAGCGAGAGCCACC
 TGCCGCCCTGGTGGGCAATCCCGCCGCCACCGCCGCCGTGAACACCGCCCTGGACGGAAATCCCG
 TGCCCATCTGATCCCTGCCACCGGTGGTGCAGGGCGACAGCGACGTGGGGCCCTACCTGGGC

GGGCTCGCCGTGAAAGAGTGGCTGCTGGCCACGAGGGCCACAGACTGGGCAAGCCTGGGCTGGG
T

Legend: HisTag Snap Tag Linkers Twitch-2B CLIP tag

Figure 5-15. Nucleotide sequence encoding the STC protein.

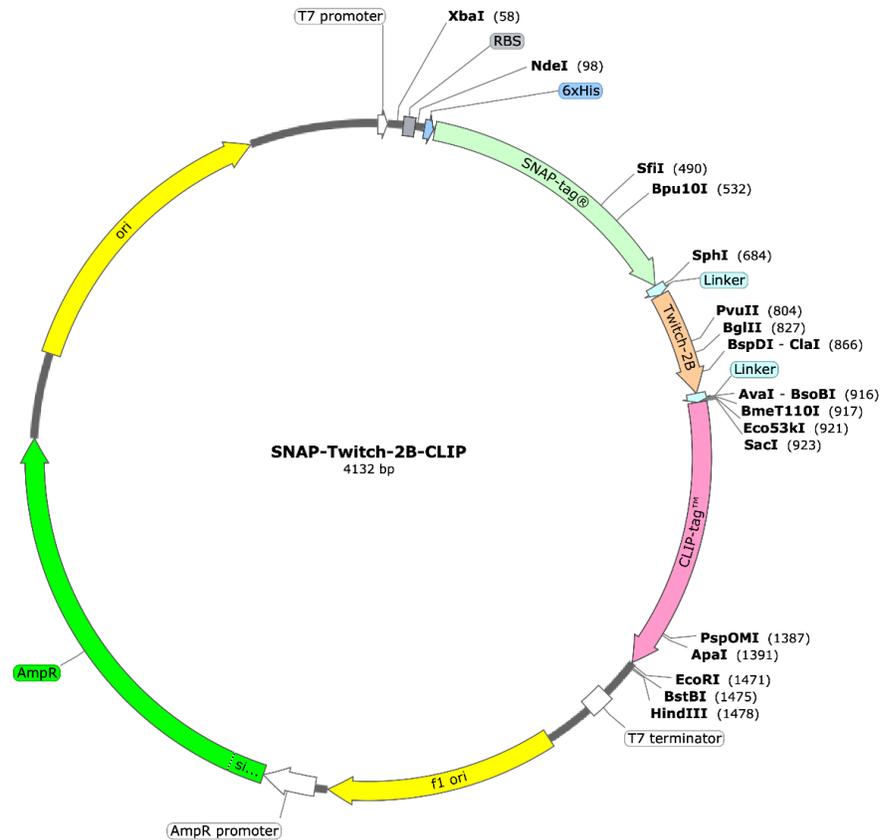


Figure 5-16. Plasmid Map of the gene encoding the STC protein in the pRSETB vector.

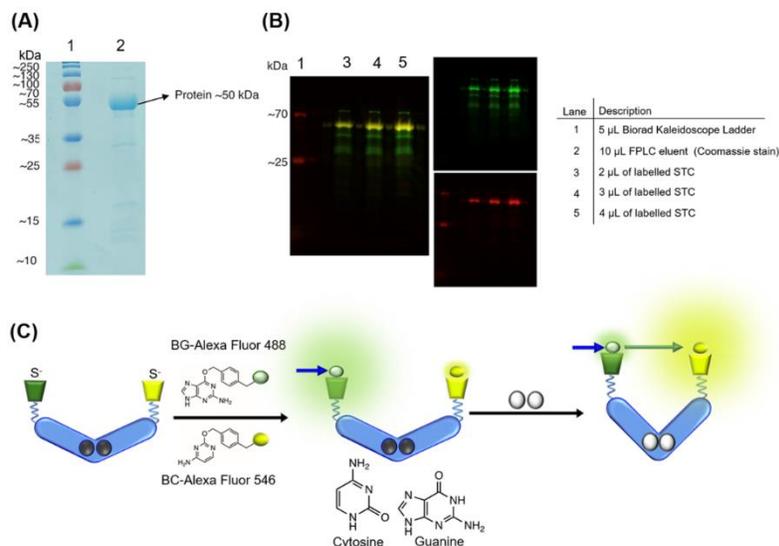


Figure 5-17. Production of STC. (A) SDS-PAGE was used to analyze the FPLC fractions to show the high protein yield and purity of STC. A. After STC protein was fluorescently labelled with Alexa Fluor 488 and 546 dyes using orthogonal SNAP and CLIP tags, which was confirmed with a B. Typhoon fluorescence scanner.³⁶⁻³⁸ The covalent linkage of the Alexa Fluor dyes onto SNAP and CLIP tag are shown in C. where cytosine and guanine are displaced in the reaction.

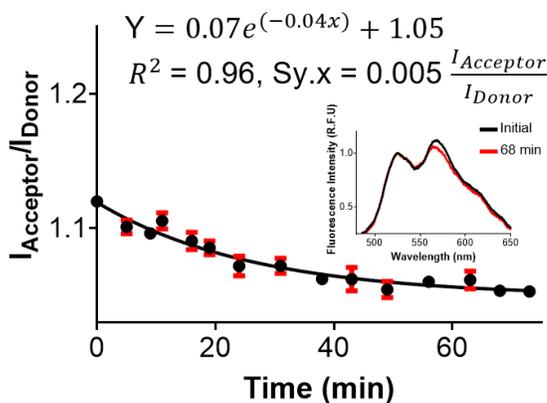


Figure 5-18. Unbinding of Ca^{2+} . The $I_{Acceptor}/I_{Donor}$ ratio falls after media exchange from RB with $CaCl_2$ (1.25 mM) solution to calcium-free RB. Inset: Emission spectra at times 0 and 68 min.

5.12 References

1. Seshadri, D. R. *et al.* Wearable sensors for monitoring the physiological and biochemical profile of the athlete. *npj Digit. Med.* **2**, (2019).
2. Díaz, J. *et al.* Correlation among ionized calcium, citrate, and total calcium levels during hepatic transplantation. *Clin. Biochem.* **28**, 315–317 (1995).
3. Chung, H. S., Cho, S. J. & Park, C. S. Effects of liver function on ionized hypocalcaemia following rapid blood transfusion. *J. Int. Med. Res.* **40**, 572–582 (2012).
4. Chockalingam, K. & Zhao, H. Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. in *AIChE Annual Meeting, Conference Proceedings* 9272 (2005). doi:10.1073/pnas.0409206102
5. Di Roberto, R. B., Chang, B. & Peisajovich, S. G. The directed evolution of ligand specificity in a GPCR and the unequal contributions of efficacy and affinity. *Sci. Rep.* **7**, (2017).
6. Maeshima, K. *et al.* A Transient Rise in Free Mg²⁺ Ions Released from ATP-Mg Hydrolysis Contributes to Mitotic Chromosome Condensation. *Curr. Biol.* **28**, 444-451.e6 (2018).
7. Lindenbarg, L. & Merckx, M. Engineering genetically encoded FRET sensors. *Sensors (Switzerland)* **14**, 11691–11713 (2014).
8. Thestrup, T. *et al.* optimized ratiometric calcium sensors for functional in vivo imaging of neurons and t lymphocytes. *Artic. Nat. methods* / **11**, 175 (2014).
9. Asif, M. H. *et al.* Functionalized zinc oxide nanorod with ionophore-membrane coating as an intracellular Ca²⁺ selective sensor. *Appl. Phys. Lett.* **95**, (2009).
10. Shortreed, M., Kopelman, R., Kuhn, M. & Hoyland, B. Fluorescent Fiber-Optic Calcium Sensor for Physiological Measurements. *Anal. Chem.* **68**, 1414–1418 (1996).
11. Kastellorizios, M., Tipnis, N. & Burgess, D. J. Foreign body reaction to subcutaneous implants. *Adv. Exp. Med. Biol.* **865**, 93–108 (2015).
12. Fu, Z., R. Gilbert, E. & Liu, D. Regulation of Insulin Synthesis and Secretion and Pancreatic Beta-Cell Dysfunction in Diabetes. *Curr. Diabetes Rev.* **9**, 25–53 (2012).
13. Hink, M. A. *et al.* Structural dynamics of green fluorescent protein alone and fused with a single chain Fv protein. *J. Biol. Chem.* **275**, 17556–17560 (2000).
14. Woting, A. & Blaut, M. Small intestinal permeability and gut-transit time determined with low and high molecular weight fluorescein isothiocyanate-dextran in C3H mice. *Nutrients* **10**, (2018).
15. Algar, W. R., Hildebrandt, N., Vogel, S. S. & Medintz, I. L. FRET as a biomolecular research tool — understanding its potential while avoiding pitfalls. *Nat. Methods* **16**, 815–829 (2019).
16. Lipman, N. S., Jackson, L. R., Trudel, L. J. & Weis-Garcia, F. Monoclonal Versus Polyclonal Antibodies: Distinguishing Characteristics, Applications, and Information Resources. *ILAR J.* **46**, 258–268 (2005).
17. Dell'italia, L. J., Collawn, J. F. & Ferrario, C. M. Multifunctional role of chymase in acute and chronic tissue injury and remodeling. *Circulation Research* **122**, 319–336 (2018).

18. Lin, C. C. & Anseth, K. S. PEG hydrogels for the controlled release of biomolecules in regenerative medicine. *Pharmaceutical Research* **26**, 631–643 (2009).
19. Lanasa, S. M., Hoffecker, I. T. & Bryant, S. J. Presence of pores and hydrogel composition influence tensile properties of scaffolds fabricated from well-defined sphere templates. *J. Biomed. Mater. Res. - Part B Appl. Biomater.* **96 B**, 294–302 (2011).
20. Jang, H. J., Shin, C. Y. & Kim, K. B. Safety evaluation of polyethylene glycol (PEG) compounds for cosmetic use. *Toxicol. Res.* **31**, 105–136 (2015).
21. Renaghan, A. D. & Rosner, M. H. Hypercalcemia: etiology and management. *Nephrol. Dial. Transplant.* **33**, 549–551 (2018).
22. Anastasopoulos, D., Kefaliakos, A. & Michalopoulos, A. Is plasma calcium concentration implicated in the development of critical illness polyneuropathy and myopathy? *Crit. Care* **15**, R247 (2011).
23. Bajar, B. T., Wang, E. S., Zhang, S., Lin, M. Z. & Chu, J. A guide to fluorescent protein FRET pairs. *Sensors (Switzerland)* **16**, (2016).
24. Van Munster, E. B., Kremers, G. J., Adjobo-Hermans, M. J. W. & Gadella, T. W. J. Fluorescence resonance energy transfer (FRET) measurement by gradual acceptor photobleaching. *J. Microsc.* **218**, 253–262 (2005).
25. Derieppe, M., Bos, C., De Greef, M., Moonen, C. & Denis De Senneville, B. Correction for photobleaching in dynamic fluorescence microscopy: Application in the assessment of pharmacokinetic parameters in ultrasound-mediated drug delivery. *Phys. Med. Biol.* **61**, 588–600 (2015).
26. Panchuk-Voloshina, N. *et al.* Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. *J. Histochem. Cytochem.* **47**, 1179–1188 (1999).
27. Regoes, A. & Hehl, A. B. SNAP-tagTM mediated live cell labeling as an alternative to GFP in anaerobic organisms. *Biotechniques* **39**, 809–812 (2005).
28. Gautier, A. *et al.* An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chem. Biol.* **15**, 128–136 (2008).
29. Zaharieva, D. P. *et al.* Lag Time Remains with Newer Real-Time Continuous Glucose Monitoring Technology During Aerobic Exercise in Adults Living with Type 1 Diabetes. *Diabetes Technol. Ther.* **21**, 313–321 (2019).
30. Kurganov, B. I., Lobanov, A. V., Borisov, I. A. & Reshetilov, A. N. Criterion for Hill equation validity for description of biosensor calibration curves. *Anal. Chim. Acta* **427**, 11–19 (2001).
31. Bagur, R. & Hajnóczky, G. Intracellular Ca²⁺ Sensing: Its Role in Calcium Homeostasis and Signaling. *Molecular Cell* **66**, 780–788 (2017).
32. Mótyán, J., Tóth, F. & Tózsér, J. Research Applications of Proteolytic Enzymes in Molecular Biology. *Biomolecules* **3**, 923–942 (2013).
33. Sanford, L. & Palmer, A. Recent Advances in Development of Genetically Encoded Fluorescent Sensors. in *Methods in Enzymology* **589**, 1–49 (Academic Press Inc., 2017).
34. Vinkenborg, J. L. *et al.* Genetically encoded FRET sensors to monitor intracellular Zn²⁺ homeostasis. *Nat. Methods* **6**, 737–740 (2009).

35. Leach, S. P. Z. and J. B. Hydrolytically degradable poly(ethylene glycol) hydrogel Biomacromolecules. *Biomacromolecules* **11**, 1348–1357 (2011).
36. Keppler, A. *et al.* A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **21**, 86–89 (2003).
37. Keppler, A. *et al.* Labeling of fusion proteins of O6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro. *Methods* **32**, 437–444 (2004).
38. Keppler, A., Pick, H., Arrivoli, C., Vogel, H. & Johnsson, K. Labeling of fusion proteins with synthetic fluorophores in live cells. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9955–9959 (2004).

CHAPTER 6

Under 5 Minute Immunoblot Assays by Vortex Fluidic Device Acceleration

Author contributions: Emily C. Sanders and Gregory A. Weiss designed research; Emily C. Sanders Sanjana Sen, Aidan A. Gelston, Alicia M. Santos, Xuan Luo, Keertna Bhuvan, Derek Y. Tang performed research. Emily C. Sanders, Gregory A. Weiss and Colin L. Raston wrote the manuscript.

6.1 Abstract

Unlocking the potential of personalized medicine in point-of-care settings requires a new generation of biomarker and proteomic assays. Ideally, assays could inexpensively perform hundreds of quantitative protein measurements in parallel at the bedsides of patients. This goal greatly exceeds current capabilities. Furthermore, biomarker assays are often challenging to translate from benchtop to clinic due to difficulties achieving the necessary selectivity, sensitivity, and reproducibility. To address these challenges, we developed an efficient (<5 min), robust (low CoVs), and inexpensive (decreasing reagent use and cost by >70%) immunoassay method. Specifically, the immunoblot membrane is dotted with the sample and then developed in a vortex fluidic device (VFD) reactor. All assay steps – blocking, binding, and washing – leverage the unique thin-film microfluidics of the VFD. The approach can accelerate direct, indirect, and sandwich immunoblot assays. The applications demonstrated include assays relevant to both the laboratory and the clinic.

6.2 Introduction

The promises of personalized medicine require efficient, inexpensive testing for the presence and concentration of biomarkers. Low-cost diagnostics for broad deployment of precision medicine also represent a health justice issue, as high-tech medical devices often neglect resource-limited areas.¹ The extreme disparity between technologically lagging and advanced settings directly impacts disease mortality and morbidity, particularly for

infectious diseases.² Thus, a clear need exists for a simple, cost-effective platform technology to advance precision medicine worldwide.

Point-of-care (PoC) tests have revolutionized diagnostics and patient care. For example, the rapid Strep A test has reduced unnecessary antibiotic treatments with clear benefits to public health.³ Similarly, PoC influenza tests can allow early antiviral intervention, if conducted within 72 hours post-symptom onset.^{4,5} The pregnancy test for chorionic gonadotropin has changed women's reproductive health and has emerged as both the most common at-home and PoC diagnostic test.⁶ Despite these successful examples, a clear gap exists between the thousands of evidence-based biomarkers reported and their validation in the clinic.⁷ Thus, technology allowing PoC biomarker validation and widespread deployment is required to close this gap.

Already used extensively for biomarker-based tests, immunoblot assays (IAs) offer a low-tech, but highly effective disease diagnostic. For example, an IA was developed as a cost-effective tool for detection of Dengue, a rampant ailment in countries lacking medical infrastructure for more complicated testing.⁸ Similarly, IAs have been developed for the diagnosis of myofibrillar myopathies.⁹ Most prominently, an IA is used in concert with an enzyme-linked immunosorbent assay (ELISA) to diagnose human immunodeficiency virus.^{10,11}

However, IAs typically incur high costs and have complex protocols and low sensitivity.^{12,13} Despite these limitations, IAs in laboratories are often used to optimize conditions for the more experimentally demanding and time-consuming Western blot. A conventional IA typically requires >2 hr and consumes significant amounts of expensive

reagents (e.g., primary and secondary antibodies).¹⁴ Kurien *et al.* previously described a shortened, >40 min IA protocol reliant on processing the blot with reagents pre-warmed to 37 °C, which suggests thermally driving equilibration is one approach to accelerating IAs.¹⁵ An alternative, especially for temperature sensitive applications, the Vitrozm Zoom Blot Plate, a single-use apparatus, can perform a multiplexed IA in 60 min.¹⁶

Eliminating the IA's background is key to improving its sensitivity. Wu *et al.* determined that inefficient washing is the main contributor to high background in IAs.¹⁷ We envisioned applying the mechanical energy of a vortex fluidic device (VFD) to provide stringent washes, accelerated equilibration, and decrease IA background. Previously, the VFD has been used to drive protein purification and tethering¹⁸, recovery of DNA from formalin-preserved tissue¹⁹, protein folding²⁰, and embedding active enzyme in xerogels.²¹ Here, we report using the VFD to improve IA sensitivity and reduce processing time to <5 min.

6.3 Results and Discussion

To access the unique microfluidics of the VFD, IAs were performed entirely inside a VFD quartz reactor. Briefly, antigens were dotted on a nitrocellulose membrane that was sandwiched between two sheets of filter paper. The paper assembly was rolled into a

cylinder and placed concentrically within the VFD reactor. All subsequent blocking, binding, and washing steps then took place inside the spinning VFD.

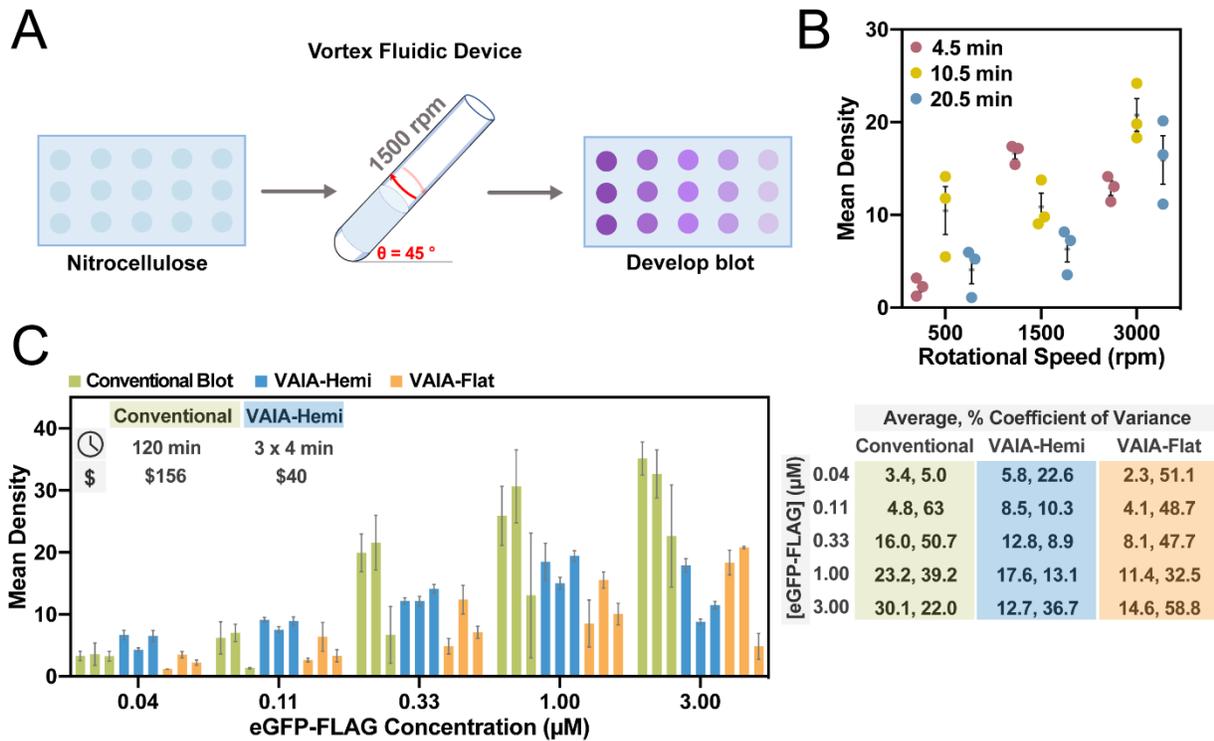


Figure 6-1. Optimization of the VAIA method. **A.** Nitrocellulose membranes dotted with target antigens are concentrically placed in a VFD reactor. The VFD reactor is tilted at 45° and rotated at 1500 rpm throughout the assay steps. Membranes are removed from the VFD reactor and the assay dots are visualized with a colorimetric reagent. **B.** An indirect phage IA was used to optimize VAIA rotational speed and assay time. Systematic screening of 9 combinations of these parameters revealed 1500 rpm and a 4.5 min assay time yielded the highest signal-to-noise ratios and lowest error. Error bars represent SEM (n = 3). **C.** A direct eGFP-FLAG detection IA was used to compare quantitation with VAIA to the conventional IA method. VAIA was faster, more robust, and more sensitive than the conventional method. The fluid flow was examined by performing VAIA with either hemispherical-bottom VFD reactors or flat-bottom VFD reactors. Error bars represent SEM for each group of dots on each immunoblot (n = 3). Coefficients of variance indicate variation between separate immunoblots (n = 3).

The initial optimization of the VFD-accelerated IA (VAIA) was performed with a previously described, robust assay format.²² The dotted HSA antigen was captured by addition of a small quantity (2 mL of 1 nM phage) of a phage-displayed HSA binding ligand and visualized by colorimetric activity of a peroxidase-conjugated, phage-specific antibody (2 mL of 1:10,000 diluted antibody in PBS) (Figure 6-1A). The VFD conditions, including

rotational speed and time, were subject to optimization. Ultimately, 1500 rpm and 4.5 min total of blocking (1 min), phage binding (2 min), antibody binding (1 min), and washing (2 x 15 s) steps yielded the greatest signal-to-noise ratios and the lowest levels of non-specific background binding (Figure 6-1B).

Next, a direct assay further demonstrated the generality of the technique and revealed the effects of fluid flow on VAIA (Figure 6-1C). The model protein, enhanced green fluorescent protein with a C-terminal FLAG-tag (eGFP-FLAG)²³, was detected with an anti-eGFP peroxidase-conjugated antibody. The assay was performed in either a hemispherical quartz VFD tube (VAIA-Hemi) or a flat-bottom quartz VFD tube (VAIA-Flat). At the optimal speed, the curved hemispherical base of the tube and the curved wall of the tube are expected to create a Coriolis fluid flow impacting the inner surface of the tube. Overall, the fluid flow afforded by the VAIA-Hemi resulted in more sensitivity and lower variability, where the rapid processing arises from the Coriolis fluid flow inducing high mass transfer into and out of the membrane.

The multifunctional eGFP-FLAG fusion protein provided a useful model to demonstrate a multitude of classic IA methods. The FLAG-tag was used as the antigen in both indirect (4.5 min) and direct (3.25 min) IA formats (Figure 6-2A). Both formats delivered a dose-dependent response and the expected binding patterns (i.e., one format has higher sensitivity due to stronger binding affinity). Specifically, the direct FLAG immunoblot has drastically reduced signal below 0.33 μM . However, the indirect FLAG immunoblot maintains a detectable signal to concentrations of eGFP-FLAG as low as 0.04 μM . These data are

consistent with the finding that indirect immunoassays are generally more sensitive due to signal amplification by the secondary antibody.^[26]

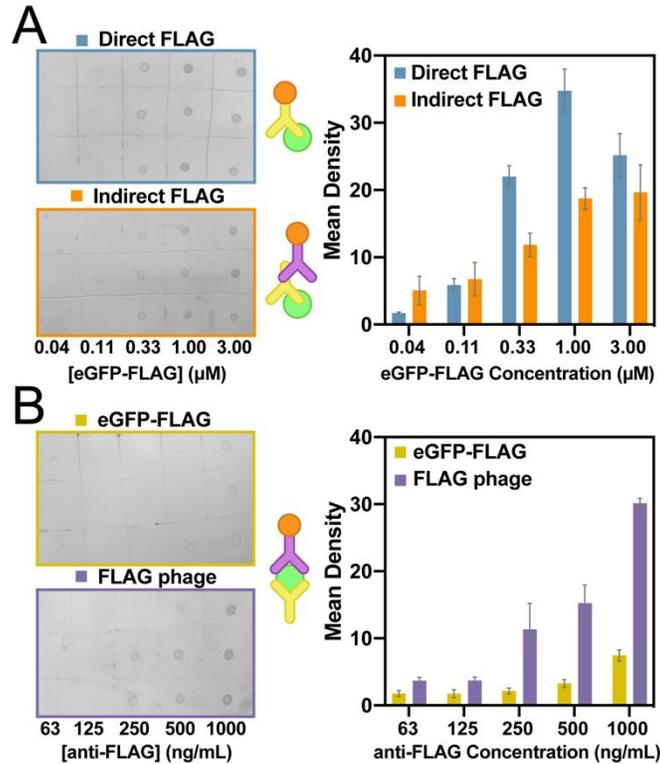


Figure 6-2. Generalization of VAIA to three common immunoassay formats. **A.** Both an indirect and direct anti-FLAG VAIA allowed dose-dependent quantification of eGFP-FLAG. The different binding modes cause the data to vary in both sensitivity and saturation limit. **B.** Anti-FLAG antibodies captured either FLAG phage or eGFP-FLAG for sandwich VAIAs. Detection of FLAG phage was more sensitive, but both methods resulted in useful signal. A schematic for each assay format is provided (middle). Error bars represent SEM ($n = 3$).

Another common IA, the sandwich-format, features an antigen entrapped between two noncompetitive antibodies. VAIA enabled two different sandwich-format assays to be performed with each requiring <5 min (Figure 6-2B). An anti-FLAG antibody was dotted on the membrane and captured either eGFP-FLAG or FLAG-tagged M13 bacteriophage (FLAG phage). The eGFP-FLAG fusion was sandwiched with anti-eGFP-HRP; the FLAG phage was sandwiched with anti-M13-HRP. Both immunoblots demonstrated dose-dependent binding;

however, the FLAG phage signal was significantly more intense and sensitive. Taken as a whole, the wide variety of IA formats demonstrate the generality and robustness of VAIA.

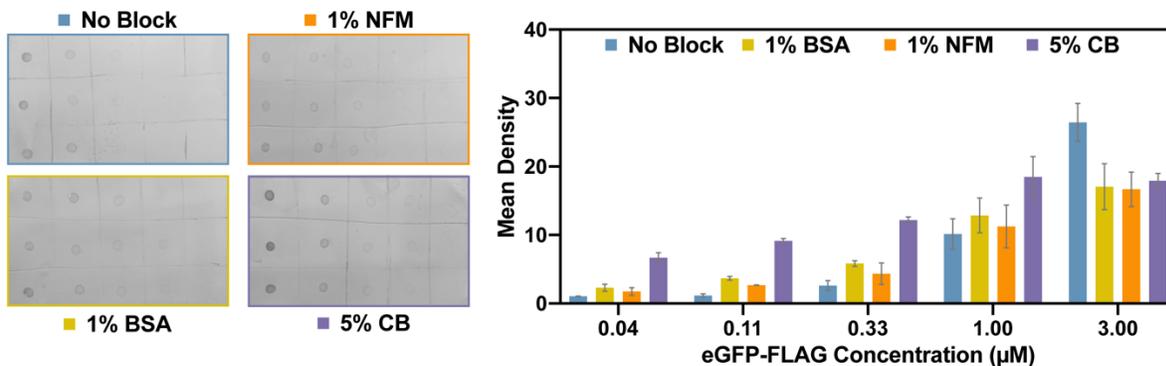


Figure 6-3. Generalizability of VAIA with different blocking agents. A direct detection of eGFP-FLAG with anti-eGFP-HRP was performed with no blocking (PBS-T), 1% bovine serum albumin (BSA), 1% non-fat milk (NFM), and 5% ChonBlock (5% CB) to demonstrate the utility of VAIA with blocking systems currently used in research and clinical settings. Error is represented as SEM (n = 3).

The previously described assays were all performed with a commercial blocking agent, ChonBlock (CB). However, researchers typically block with solutions of bovine serum albumin (BSA), non-fat milk (NFM), or high concentrations of the non-ionic detergent Tween-20. Therefore, the eGFP-FLAG direct detection with anti-eGFP-HRP was repeated with these more common blocking conditions (Figure 6-3). All four blocking conditions resulted in a robust signal; however, 5% CB had the best signal-to-noise and therefore the best sensitivity. The experiments illustrate the adaptability of the VAIA platform for application to a variety of IA conditions and reagents.

VAIA also works well for potential clinical applications. In clinical samples, endogenous proteins can be challenging to detect due to the complex composition of the biofluid, including interfering substances. Human serum albumin (HSA) levels in the body are a biomarker for malnutrition, cirrhosis, and kidney disease.^{24,25} Here, detection of endogenous HSA in diluted plasma, sera, urine, and blood from human patients validated

testing of biofluids with VAIA (Figure 6-4A). HSA levels were the highest in blood, followed by plasma, then sera. In urine, the measured HSA levels were undetectable (i.e., comparable to the negative control). This result is predictable, as urine from healthy donors should be relatively HSA-free.²⁶

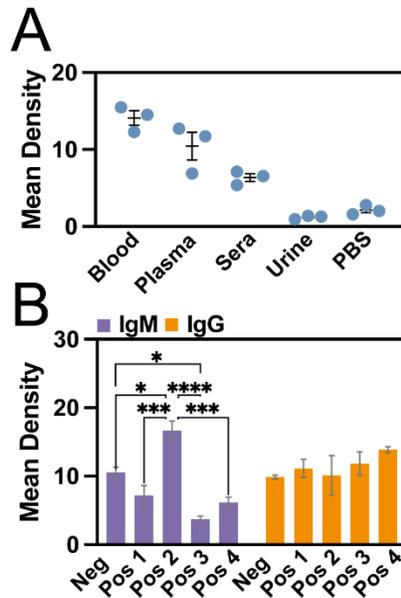


Figure 6-4. Clinical potential of VAIA with biofluids. **A.** An indirect HSA immunoassay in plasma, sera, urine, blood, and PBS demonstrated VAIA applicability to a variety of biofluids. **B.** IgM and IgG levels in plasma from healthy and COVID-19 patients were assayed with VAIA. Error bars represent SEM (n = 3). ANOVA with Tukey’s multiple comparisons yields p-values of *<0.05, ***<0.001, ****<0.0001.

Several diseases can be diagnosed through the assessment of immunoglobulin levels in biofluids.²⁷⁻²⁹ However, the current state-of-the-art immunoglobulin assays are lengthy and complicated. Here, pooled plasma from healthy donors and COVID-19 patients further characterized the clinical potential of VAIA (Figure 6-4B). Interestingly, immunoglobulin G (IgG) levels were consistent amongst the pooled plasma (no significance by ANOVA), and immunoglobulin M (IgM) levels varied drastically from one another. This data agrees with previous measurements from our laboratory.²³ In summary, the strong wash conditions of VAIA can overcome interfering substances during IAs with clinical samples.

The approach described here could find use in many chemical processes requiring the interactions of solid and liquid states. In diagnostics, for example, molecular recognition often requires molecules in liquid-phase to bind to a target affixed to solid support. The VFD-driven rapid equilibration could accelerate equilibration of otherwise slow binding events. Furthermore, the >10-fold acceleration, combined with decreased cost, and simplified execution suggests the work reported here could advance IAs in academic, industrial, and clinical spaces.

We conclude by noting that VAIA satisfies the requirements for bringing proteomic assays to PoC settings. VAIA is robust, rapid, and technically simple to execute, unlike conventional IAs. This conclusion was verified by an independent operator who replicated the Materials and Methods. Additionally, the large number of assay formats in this report demonstrates VAIA's adaptability to a variety of established immunoassay formats. The approach can be readily scaled to examine hundreds to potentially thousands of proteins in one assay using a <5 min, inexpensive sandwich format assay. Most importantly, the data are unambiguous, digitizable with a cell phone camera, and robust. Therefore, VAIA could address the gap between development and implementation of biomarker-based precision medicine.

6.4 Materials and Methods

Propagation and Purification of M13 Phage-Displayed Peptide Ligands. The HSA-binding- or FLAG-phagemid was transformed into SS320 Escherichia coli (*E. coli*) competent cells and heat-shocked at 42 °C for 40 s. The cells were plated onto a pre-warmed LB-carbenicillin plate at 37 °C overnight. Near a flame, a 20 mL primary culture of 2YT (16 g

tryptone, 10 g yeast extract, 5 g NaCl in 1 L autoclaved, nanopure water) with carbenicillin (50 mg/mL) and tetracycline (5 mg/mL) was inoculated with a single colony and incubated with shaking at 225 rpm and 37 °C until an OD₆₀₀ of 0.5 to 0.7 was reached. Next, the primary culture was induced by addition of isopropyl β-D thiogalactopyranoside (IPTG, 30 μM) and M13K07 helper phage (8 nM, 5 μL). After incubation for 45 min at 37 °C with shaking at 225 rpm, 8 mL of the primary culture was transferred to 150 mL of 2YT supplemented with carbenicillin (50 μg/mL), kanamycin (20 μg/mL), and IPTG (30 μM) and incubated for 18 h at 30 °C with shaking at 225 rpm.

To harvest the phage, the expression culture was centrifuged at 10 krpm (15300 x g) for 10 min at 4 °C. The supernatant was transferred to a 250 mL centrifuge bottle with 30 mL of 20% (w/v) PEG-8000/2.5 M NaCl and inverted three times. After incubation on ice for 45 min, an additional centrifugation was done at 10 krpm (15300 x g) for 30 min at 4 °C. Without disturbing the pellets, the supernatant was decanted, and the centrifuge bottles were centrifuged with pellets facing away from the central axis of the rotor at 4 krpm (2,500 x g) for 4 min at 4 °C. The pellets were resuspended in 1X PBS (10 mM phosphate, 137 mM NaCl, pH 7.2) supplemented with TWEEN-20 (0.05%, v/v) and glycerol (10%, v/v), aliquoted into 1.5 mL Eppendorf tubes, flash frozen with liquid nitrogen, and stored at -80 °C. As required, the phage were thawed on ice. One-fifth of the total volume of 20% (w/v) PEG-8000/2.5 M NaCl was added, and the phage was incubated on ice for 45 min. Next, the phage were centrifuged at 10 krpm (15300 x g) for 30 min at 4 °C. The supernatant was decanted, and the pellets were resuspended in 1X PBS. The phage were centrifuged a final time at 10 krpm (15300 x g) for 10 min and decanted into a clean tube. The phage concentration was

determined by measuring the absorbance at 268 nm, multiplying the A_{268} value by a factor of 8.14, and correcting for the dilution factor.

Expression and IMAC Purification of eGFP-FLAG. A pET28c plasmid containing eGFP-FLAG fused to an N-terminal His-tag was transformed into BL21 DE3* E. coli heat-shock, competent cells. Cells were plated on a carbenicillin-supplemented (50 $\mu\text{g}/\text{mL}$) LB-agar plate and incubated at 37 °C overnight. A single colony was transferred to an overnight culture of 25 mL LB media (10 g tryptone, 10 g NaCl, 5 g yeast extract in 1 L autoclaved nanopure water) supplemented with carbenicillin (50 $\mu\text{g}/\text{mL}$) and incubated at 37 °C for 18 h. An expression culture of 500 mL LB supplemented with carbenicillin (50 $\mu\text{g}/\text{mL}$) was inoculated with 5 mL of the seed culture before incubation at 37 °C with shaking at 225 rpm until an $\text{OD}_{600} \sim 0.5$ was reached. The cultures were induced by addition of 0.5 mM IPTG and incubated for 18 h at 25 °C. To harvest the protein, the culture was centrifuged (9632 x g) for 20 min, resuspended in lysis buffer (20 mM Tris-HCl, 250 mM NaCl, pH 8), and sonicated. The lysate was centrifuged (26,892 rcf, 45 min, 4 °C), and the supernatant was incubated with charged Ni^{2+} -charge IMAC (Profinity™) resin overnight on a rotary shaker (150 rpm at 4 °C). The resin-bound protein was purified by gravity column, first by washing with lysis buffer containing 20 mM imidazole and then eluted with lysis buffer containing 250 mM imidazole. The eluted fractions were visualized using a 12% SDS-PAGE (Bio-rad Mini-PROTEAN Tetra electrophoresis system) stained with Coomassie brilliant blue dye (Figure 6-5). The eluted fractions containing the purified eGFP-FLAG were pooled, and the buffer exchanged for 3 column volumes (20 mL) with lysis buffer without imidazole using a 10 kDa

cutoff microconcentrator (Vivaspin, Fisher Scientific). The protein concentration was determined by Bradford assay using the estimated MW.

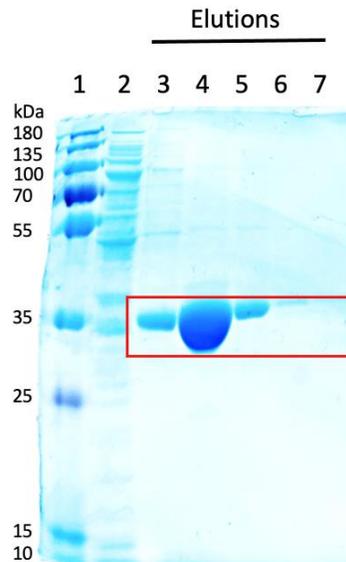


Figure 6-5. SDS-PAGE analysis of eGFP-FLAG expression and purification. The fusion protein had an expected molecular weight of ~32 kDa, as indicated with a red box. From left to right, the lanes are the following: 1) PageRuler Plus Prestained Protein Ladder, 2) Flowthrough, 3) Wash, 4) Elution 1, 5) Elution 2, 6) Elution 3, 7) Elution 4.

Optimization of VAIA Rotational Speed and Assay Time:

The general process of arraying samples for testing is as follows. The nitrocellulose membrane was trimmed to fit the inner circumference of the VFD reactor (5 cm) and marked with a pencil to create a 1 x 1 cm square for each sample. The negative control (1 μ L of PBS) or experimental antigen (1 μ L of 50 mg/mL HSA in PBS) was applied as a dot directly to the membrane, which was allowed to dry as usual. Next, the membrane was sandwiched between two pieces of filter paper cut to the size of the membrane; this assembly prevents damage to the membrane during processing. The filter paper membrane sandwich was shaped into a cylinder and placed concentrically within the glass VFD reactor for assay.

VAIA speed and assay time were optimized through an indirect phage assay. These assays consisted of three distinct steps performed with different times in 1:2:1 ratios

(blocking : phage binding : antibody binding) (Figure 6-6). Different rotational speeds were also explored for each ratio. First, each membrane was blocked with 3 mL 5X Chonblock (Chondrex Inc.) in PBS (blocking buffer) at the indicated speeds and times, and the solution was discarded. Next, 3 mL of HSA-binding phage (HSA-L3) diluted in blocking buffer was added to the VFD reactor for processing at the indicated speeds and times. The solution was discarded, and 3 mL of 1X PBS with 0.05% (v/v) Tween-20 (PBS-T) was added to the reactor to wash the membrane for 15 s before discarding the solution. Next, the HSA/HSA-L3 binding interaction was detected through addition of 3 mL anti-M13-HRP (Creative Diagnostics) diluted in blocking buffer (1:10,000) before VFD processing at the indicated speeds and times. Finally, this solution was discarded, and the membrane sandwich was washed once more as previously described.

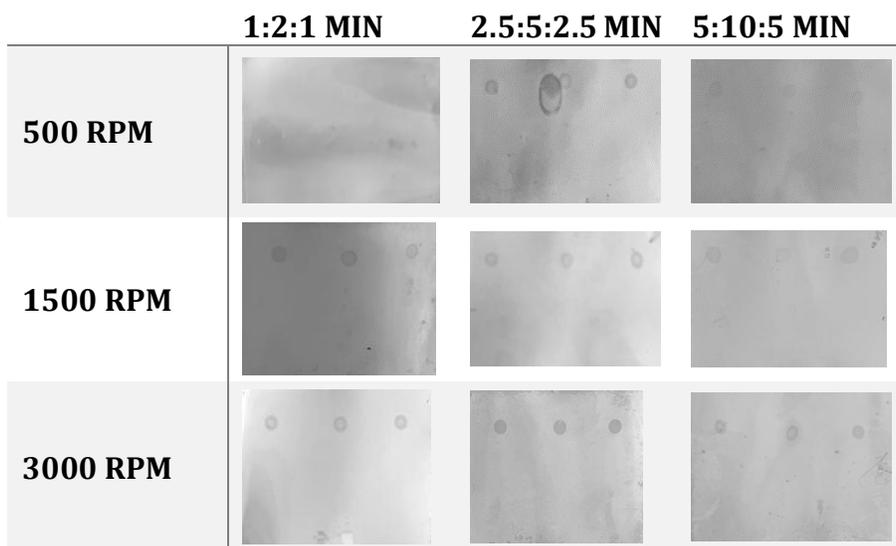


Figure 6-6. Images of HSA-phage VAIs with varying speeds and assay times. For each blot, the top row is dotted with HSA, the bottom row is dotted with PBS. Images were taken under a light table with an iPhone 12 mini camera and imported into ImageJ for further analysis. For images shown in this manuscript, an additional adjustment of contrast was performed.

After each assay was complete, the membrane was removed from the VFD reactor through gentle tapping and placed on a clean paper towel. Sufficient TMB solution was added to cover the membrane surface. After sufficient time passed for color to develop (1 to 5 min), each membrane was rinsed with water and imaged after drying. All membranes were imaged using top-down lighting by a light table with an iPhone 12 Mini camera (Figure 6-7). The images were transferred into Image-J and converted to 8-bit grayscale format, as directed by the software. Next, the background was subtracted with a 50.0 rolling ball radius and light-dark inverted. The mean density of each assay dot was measured and imported into Prism 9.0 for further analysis.

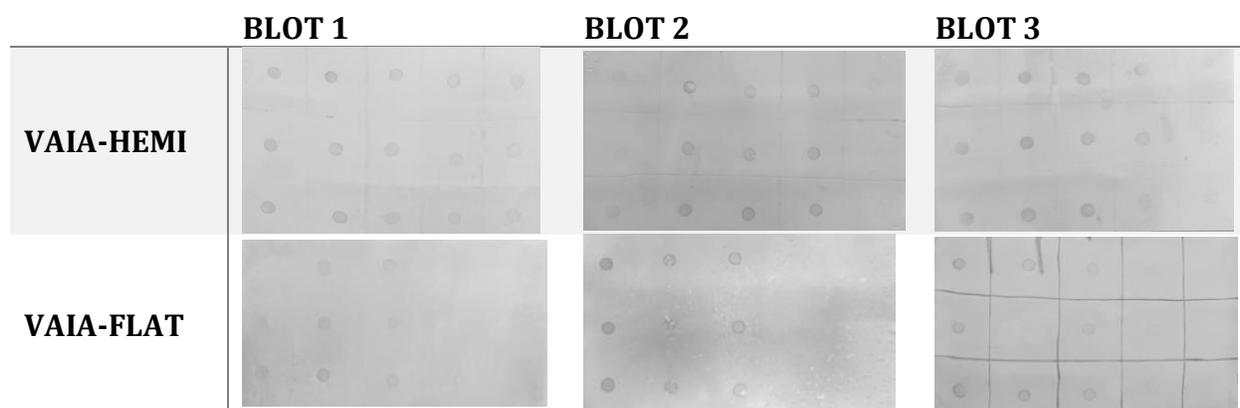


Figure 6-7. Images of eGFP-FLAG VAIAs. From left to right on each blot, the eGFP-FLAG concentrations are: 3.00 μM , 1.00 μM , 0.33 μM , 0.11 μM , 0.04 μM . Images were acquired under a light table with an iPhone 12 mini camera and imported into ImageJ for further analysis. For images shown in this manuscript, an additional adjustment of contrast was performed.



Figure 6-8. Images of conventional eGFP-FLAG IAs. From left to right for each blot, the eGFP-FLAG concentrations are: 3.00 μM , 1.00 μM , 0.33 μM , 0.11 μM , 0.04 μM . Images were acquired under a light table with an iPhone 12 mini camera and imported into ImageJ for further analysis. For images shown in this manuscript, an additional adjustment of contrast was performed after data analysis.

Direct Detection of eGFP-FLAG with Conventional Dot Blot. The membranes were prepared as described above. These assays were performed in separate plastic containers. To prevent non-specific binding, 10 mL of blocking buffer were added to each container, shaken at 150 rpm for 30 min, then discarded. Next, 10 mL of 1:1,000 anti-eGFP-HRP in blocking buffer were added to each container, shaken at 150 rpm for 60 min, then discarded. Each membrane was washed with 10 mL PBS-T with shaking at 150 rpm for 10 min three times. The membranes were developed, and the data were processed as described above (Figure 6-8).

Direct VAIA Detection of eGFP-FLAG with anti-FLAG-HRP. The membranes were prepared and blocked as previously described. Next, 2 mL of 1:1,000 anti-FLAG-HRP (Sigma-Aldrich) in blocking buffer were added to the tube and spun at 1500 rpm for 2 min. The solution was then discarded, and the membranes were washed and processed as described above.

Indirect VAIA Detection of eGFP-HRP with anti-FLAG and anti-mouse-HRP. The membranes were prepared and blocked as previously described. The solution was then

discarded. Next, 2 mL of 1:1,000 anti-FLAG (Sigma-Aldrich) in blocking buffer were added to the tube and spun at 1500 rpm for 2 min. The solution was discarded, and the membrane was washed by adding 2 mL of PBS-T and spinning at 1500 rpm for 15 s. After discarding the wash solution, 2 mL of 1:10,000 anti-mouse-HRP in blocking buffer were added to the tube, spun at 1500 rpm for 2 min, then discarded. The solution was discarded, and the membranes were washed and processed as previously described.

Sandwich VAIA Detection of eGFP-FLAG with anti-FLAG and anti-eGFP-HRP.

Membranes were prepared as previously described. Five serial dilutions of anti-FLAG (1:100, 1:1000, 1:10,000, 1:100,000, 1:1,000,000) were prepared and 1 μ L of each was added in triplicate on each of the five columns within the grid. The membrane was sandwiched in filter paper and blocked as previously described. Next, 2 mL of 3 μ M eGFP-FLAG in blocking buffer were added, and the tube was spun at 1500 rpm for 2 min. The solution was then discarded, and the membrane was washed as previously described. Next, 2 mL of 1:1000 anti-eGFP-HRP in blocking buffer were added to the tube and spun at 1500 rpm for 2 min, then discarded. The solution was discarded, and the membrane was washed, and the data were processed as previously described.

Sandwich VAIA Detection of eGFP-FLAG with anti-FLAG and FLAG-Binding Phage.

The membrane was prepared and blocked as previously described. Next, 2 mL of 1 nM FLAG-binding phage in blocking buffer were added to the tube and spun at 1500 rpm for 2 min. The solution was then discarded, and the membrane washed as described above. Next, 2 mL of 1:1000 anti-M13-HRP in blocking buffer were added to the tube and spun at 1500 rpm for 2 min, then discarded. The solution was discarded, and the membrane was washed, and data processed as previously described.

Direct VAIA Detection of HSA in Biofluids. The membrane was prepared as previously described. Plasma, sera, urine, blood were diluted (1:100) in PBS and dotted on the membrane in triplicate. Additionally, PBS was dotted on the membrane as a negative control. An image of the membrane was acquired as before. The membrane was sandwiched in filter paper and blocked as previously described. Next, 2 mL of anti-HSA-HRP (Bethyl Laboratories, 1:1000) in blocking buffer were added to the tube and spun at 1500 rpm for 2 min. The solution was discarded, and the membrane was washed and imaged as described.

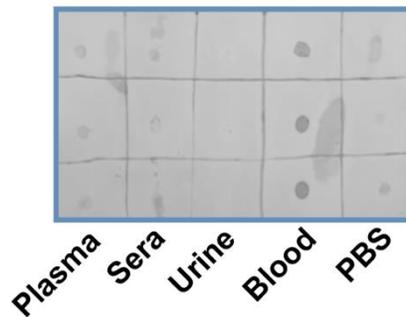


Figure 6-9. Image of VAIA assay completed in a variety of biofluids. Images were acquired under a light table with an iPhone 12 mini camera and imported into ImageJ for further analysis. For images shown in this manuscript, an additional adjustment of contrast was performed after data analysis.

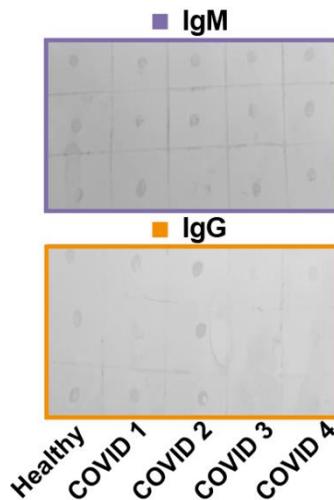


Figure 6-10. Images of patient plasma VAIA. IgMs and IgGs were detected with anti-IgM-HRP or anti-IgG-HRP, respectively. Images were acquired under a light table with an iPhone 12 mini camera and imported into ImageJ for further analysis. For images shown in this manuscript, an additional adjustment of contrast was performed after data analysis.

To account for the initial dark color of the diluted blood, the mean density values were calculated by subtracting the values before and after the assay (Figure 6-9).

Antibody Detection in COVID-19 Patient Plasma. The membrane was prepared as previously described. Plasma from healthy patients and COVID-19 patients were diluted (1:100) in PBS and dotted on the membrane in triplicate. The membrane was sandwiched in filter paper and blocked as previously described. Next, 2 mL of anti-IgG-HRP (Sigma-Aldrich, 1:1000) or anti-IgM-HRP (Invitrogen, 1:1000) in blocking buffer were added to the tube and spun at 1500 rpm for 2 min. The solution was discarded, and the membrane was washed and imaged as described above (Figure 6-10).

6.5 References

1. Howitt, P. *et al.* Technologies for global health. *Lancet* **380**, 507–535 (2012).
2. Gavazzi, G., Herrmann, F. & Krause, K. H. Aging and infectious diseases in the developing world. *Clin. Infect. Dis.* **39**, 83–91 (2004).
3. Hendi, S. Bin, Malik, Z. A., Khamis, A. H. & Al-Najjar, F. Y. A. High diagnostic accuracy of automated rapid Strep A test reduces antibiotic prescriptions for children in the United Arab Emirates. *BMC Pediatr.* **21**, 1–6 (2021).
4. Balish, A., Garten, R., Klimov, A. & Villanueva, J. Analytical detection of influenza A(H3N2)v and other A variant viruses from the USA by rapid influenza diagnostic tests. *Influenza Other Respi. Viruses* **7**, 491–496 (2013).
5. Uyeki, T. M. Influenza diagnosis and treatment in children: A review of studies on clinically useful tests and antiviral treatment for influenza. *Pediatr. Infect. Dis. J.* **22**, 164–177 (2003).
6. Gnoth, C. & Johnson, S. Strips of hope: Accuracy of home pregnancy tests and new developments. *Geburtshilfe Frauenheilkd.* **74**, 661–669 (2014).
7. Selleck, M. J., Senthil, M. & Wall, N. R. Making Meaningful Clinical Use of Biomarkers. *Biomark. Insights* **12**, 1–7 (2017).
8. Falconar, A. K. I. & Romero-Vivas, C. M. E. A simple, inexpensive, robust and sensitive dot-blot assay for equal detection of the nonstructural-1 glycoprotein of all dengue virus serotypes. *Viol. J.* **10**, 1–13 (2013).
9. Marini, M. *et al.* Immunoblot as a potential diagnostic tool for myofibrillar myopathies. *Electrophoresis* **36**, 3097–3100 (2015).
10. Esteban, J. I. *et al.* Importance of Western Blot Analysis in Predicting Infectivity of Anti-Htlv-iii/Lav Positive Blood. *Lancet* **326**, 1083–1086 (1985).
11. CDC Laboratory Procedure Manual. *HIV Western Blot Confirmatory Test.* CDC (2013).
12. Montagnese, F., Babačić, H., Eichhorn, P. & Schoser, B. Evaluating the diagnostic utility of new line immunoassays for myositis antibodies in clinical practice: a retrospective study. *J. Neurol.* **266**, 1358–1366 (2019).
13. Harrell, J., Rubio, X. B., Nielson, C., Hsu, S. & Motaparathi, K. Advances in the diagnosis of autoimmune bullous dermatoses. *Clin. Dermatol.* **37**, 692–712 (2019).
14. Stott, D. I. Immunoblotting and dot blotting. *Journal of Immunological Methods* vol. 119 153–187 (1989).
15. Kurien, B. T., Danda, D., Bachmann, M. & Scofield, R. H. SDS PAGE to immunoblot in one hour. *Methods Mol. Biol.* 449–454 (2015) doi:10.1007/978-1-4939-2694-7.
16. Mai, J. & Xie, R. Zoom Plate™ Dot Blot. *Nat. Methods Appl. Note* 1–3 (2015).

17. Wu, M., Stockley, P. G. & Martin, W. J. An improved western blotting technique effectively reduces background. *Electrophoresis* **23**, 2373–2376 (2002).
18. Britton, J., Dyer, R. P., Majumdar, S., Raston, C. L. & Weiss, G. A. Ten-Minute Protein Purification and Surface Tethering for Continuous-Flow Biocatalysis. *Angew. Chemie - Int. Ed.* **56**, 2296–2301 (2017).
19. Totoiu, C. A. *et al.* Vortex fluidics-mediated DNA rescue from formalin-fixed museum specimens. *PLoS One* **15**, 1–13 (2020).
20. Britton, J., Smith, J. N., Raston, C. L. & Weiss, G. A. Protein folding using a vortex fluidic device. *Methods Mol. Biol.* **1586**, 211–220 (2017).
21. Luo, X. *et al.* Vortex Fluidic-Mediated Fabrication of Fast Gelled Silica Hydrogels with Embedded Laccase Nanoflowers for Real-Time Biosensing under Flow. *ACS Appl. Mater. Interfaces* **12**, 51999–52007 (2020).
22. Ogata, A. F. *et al.* Virus-enabled biosensor for human serum albumin. *Anal. Chem.* (2017) doi:10.1021/acs.analchem.6b04840.
23. Sen, S. *et al.* Predicting COVID-19 Severity with a Specific Nucleocapsid Antibody plus Disease Risk Factor Score Authors. *mSphere* 2020.10.15.341743 (2021).
24. Lee, S. *et al.* Development of human serum albumin selective fluorescent probe using thieno[3,2-b]pyridine-5(4h)one fluorophore derivatives. *Sensors* **19**, (2019).
25. Doumas, B. T. & Peters, T. Serum and urine albumin: A progress report on their measurement and clinical significance. *Clin. Chim. Acta* **258**, 3–20 (1997).
26. Choi, S. *et al.* A rapid, simple measurement of human albumin in whole blood using a fluorescence immunoassay (I). *Clin. Chim. Acta* **339**, 147–156 (2004).
27. Strassburg, C. P. Autoimmune hepatitis. *Best Pract. Res. Clin. Gastroenterol.* **24**, 667–682 (2010).
28. Phillips, A. C., Carroll, D., Drayson, M. T. & Batty, G. D. Raised levels of immunoglobulin G, A and M are associated with an increased risk of total and cause-specific mortality: The Vietnam Experience Study. *J. Epidemiol. Community Health* **69**, 129–135 (2015).
29. Zhang, H. *et al.* Serum IgG subclasses in autoimmune diseases. *Med. (United States)* **94**, e387 (2015).

CHAPTER 7

Summary and Conclusions

The past few decades have seen a remarkable shift in the medical field towards precision medicine. Large-scale data collection through omics platforms and the availability of global health databases have provided a broad understanding of the factors that contribute to disease and the highly variable disease outcomes that can result^{1,2}. Factors such as genetics, medical history, environment, and lifestyle all play important roles in an individual's disease response¹. Precision medicine approaches help categorize patients based on their responses to allow early, preventative treatment, or tailor clinical interventions to accurately manage disease with minimal side effects. Here, we develop multiple tools towards furthering the goals of precision medicine in both preventing and managing diseases.

Precision medicine is particularly important for diseases that result in highly variable outcomes. The pandemic has seen patient outcomes running the gamut from asymptomatic to fatal. Chapter 1 describes instances of similarly severe symptoms being driven by different immunological pathways, making treatment even more complicated³. Therefore, identifying and characterizing a patient's predisposition for specific disease trajectories can be vitally important to effectively determine pre-emptive routes of treatment. With the advent of quick Ab testing for COVID-19 exposure^{4,5} and growing evidence of the role of the humoral immune response^{3,6,7} in driving disease outcomes, antibodies (Ab) are promising as biomarkers for rapid, accessible prognostic testing for COVID-19 disease severity.

Towards this goal, we identified an Ab biomarker that is associated with increased disease severity (Chapter 2). This Ab can be used as an early predictor for severe COVID-19, being detectable as early as one day after symptom onset. Since the presence of this Ab only identifies patients that are especially prone to severity caused by the pro-inflammatory cytokines, IL-6, early interventions with IL-6-targeting treatments⁸ may be particularly beneficial to this patient subset. Our study thus identifies a biomarker for early triaging and possible pre-emptive treatment against severe COVID-19 for a specific subset of patients.

The association between COVID-19 disease severity and metabolic disease further shows that disease outcomes can be highly dependent on an individual's overall health. In Chapter 2, we establish that certain comorbidities, such as diabetes, hypertension, and age >50, exacerbate disease specifically in patients positive for the above-mentioned severity-associated Ab. A set of co-morbidities were thus incorporated into our Ab disease prediction test to increase the specificity of predicting disease to 97%⁹. In addition to comorbidities, we determined that a history of viral infections may also influence disease (Chapter 3). Cross-reactivity of the severity-associated SARS-CoV-2 Ab with a neuraminidase epitope from an influenza A viral strain suggested that boosting of Abs from a previous influenza A strain infection may influence COVID-19 disease severity. The role of immunological memory in driving disease outcomes has been previously reported in COVID-19 patients α -spike Abs^{7,10,11}. Together these studies show that immunological context greatly impacts disease outcomes of SARS-CoV-2 infection, and further highlights the crucial need for personalized medical approaches to treat and manage the disease.

Rapid, high-throughput, multi-analyte testing can improve patient categorization and triage. While traditional ELISA-based assays require 3-5 h¹² and rapid high-throughput multi-antigen microarray testing systems can take between 2-4 h¹³, new advancements in testing platforms have recently led to the development of rapid COVID-19 Ab and antigen testing platforms by Abbott that take approximately 15 mins^{4,5,14}. In Chapter 6, we describe an efficient and low-cost immunoblot-based testing method using the Vortex Fluidic Device (VFD) that can take under 5 mins from blocking to detection. The method can be used as a qualitative diagnostic platform for indirect and direct immunoblotting strategies. This platform can be used for serological testing directly from patient serum and blood and can be used for assaying multiple patient samples at a time. Such rapid, high-throughput testing methods are crucial to urgent triaging and treatment.

Maintaining a precise molecular balance with medical interventions under disease conditions can be challenging, as an individual's response to treatment and disease states themselves can be dynamic. Frequent monitoring of biomarkers can help increase dosing accuracy. However, this can be cumbersome, and further non-compliance can result in misdosing and serious side effects. Solutions to these problems in precision medicine have focused on automated, real-time monitoring and dosage adjusting. Automated disease management for type-1-diabetes (T1D) has been at the forefront of such innovations with the advent of the artificial pancreas¹⁵. Despite advances, current technologies lack complete precision and can result in suboptimal glucose levels in patients, which can lead to serious conditions¹⁶. Monitoring the presence of multiple analytes involved in balancing glucose regulation can allow for more accurate dosing calculations. In Chapter 4, we describe the utility of a real-time insulin sensor, to supplement the continuous glucose monitor in the

artificial pancreas. We engineer the native human insulin-binding insulin receptor (InR) towards developing an optical biosensor. Using modular rearrangement of the InR binding sites, we develop multiple soluble, insulin-binding variants on three different protein expression platforms. We further characterize which domains are conducive to soluble, functional expression in robust, high-throughput expression platforms in the absence of its post-translational modifications. Finally, we design optical-assaying systems involving conformational changes of a single compact variant in mammalian cells and a two-protein assay using a chimeric InR variant and an insulin variant. These platforms lay the foundations for further development of a biosensor for insulin, or other analytes, that can enhance dosing accuracy for automated disease-management systems.

Bioconjugation tools used in developing insulin sensor variants were also used to modify existing calcium sensors for improved use in implantable devices. Chapter 5, describes the development of a photostable version of the Twitch 2B calcium sensor¹⁷. By replacing the N- and C-terminal photobleaching-prone, genetically-encoded FRET-pair fluorophores with SNAP- and CLIP-tags, photostable fluorophores were site-specifically and orthogonally conjugated to each tag (Chapter 5 and Appendix 3). Our biomedical engineering collaborators from Elliot Botvinick's laboratory further demonstrated the stable and real-time sensing capacity of this new calcium sensor within a device platform. Ca²⁺ is responsible for various physiological processes, and dramatic fluctuations can occur during medical procedures such as blood transfusions. Thus, real-time monitoring can allow for quick detection and intervention to prevent serious muscular, nervous, or cardiac outcomes¹⁸⁻²⁰.

By the next decade, systemic integration of health information and technologies is expected to vastly improve the adoption and expanse of precision medicine practices in health care. The chapters in this thesis discuss using protein engineering techniques to develop assays and tools towards the improvement of precision medicine. These tools promote rapid testing, early triaging, and pre-emptive treatment for severe diseases. Furthermore, assaying platforms and biosensors developed here can inform the design of multi-analyte sensors that allow precise dosage-adjusted delivery of treatments based on the unique and dynamic needs of individual patients.

7.1 References

1. Van Eyk, J. E. & Snyder, M. P. Precision Medicine: Role of Proteomics in Changing Clinical Management and Care. (2018). doi:10.1021/acs.jproteome.8b00504
2. Jain, K. K. Role of Proteomics in the Development of Personalized Medicine. in *Advances in Protein Chemistry and Structural Biology* **102**, 41–52 (Academic Press Inc., 2016).
3. Zohar, T. *et al.* Compromised Humoral Functional Evolution Tracks with SARS-CoV-2 Mortality. *Cell* **183**, 1508-1519.e12 (2020).
4. Shah, M. M. *et al.* Performance of Repeat BinaxNOW SARS-CoV-2 Antigen Testing in a Community Setting, Wisconsin, November-December 2020. *Clin. Infect. Dis.* (2021).
5. Serei, V. D. *et al.* Comparison of Abbott ID NOW COVID-19 rapid molecular assay to Cepheid Xpert Xpress SARS-CoV-2 assay in dry nasal swabs. *Diagnostic Microbiology and Infectious Disease* **99**, 115208 (2021).
6. Chakraborty, S. *et al.* Proinflammatory IgG Fc structures in patients with severe COVID-19. *Nat. Immunol.* **22**, 67–73 (2021).
7. Dugan, H. L. *et al.* Profiling B cell immunodominance after SARS-CoV-2 infection reveals antibody evolution to non-neutralizing viral targets. *Immunity* **54**, (2021).
8. Langer-Gould, A. *et al.* Early Identification of COVID-19 Cytokine Storm and Treatment with Anakinra or Tocilizumab. *Int. J. Infect. Dis.* **99**, 291–297 (2020).
9. Sen, S. R. *et al.* Predicting COVID-19 Severity with a Specific Nucleocapsid Antibody plus Disease Risk Factor Score. *mSphere* **6**, e00203-21 (2021).
10. Anderson, E. M. *et al.* Seasonal human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with protection. *medRxiv* **7**, 2020.11.06.20227215 (2020).
11. Wang, L. *et al.* Serological Responses to Human Virome Define Clinical Outcomes of Italian Patients Infected with SARS-CoV-2. *medRxiv Prepr. Serv. Heal. Sci.* (2020).
12. Alhajj, M. & Farhana, A. *Enzyme Linked Immunosorbent Assay*. *StatPearls* (StatPearls Publishing, 2021).
13. Hedde, P. N. *et al.* A modular microarray imaging system for highly specific COVID-19 antibody testing. *Lab Chip* **20**, (2020).
14. Valdivia, H. O. *et al.* Field validation of a magneto-optical detection device (Gazelle) for portable point-of-care Plasmodium vivax diagnosis. *PLoS One* **16**, e0253232 (2021).
15. Bekiari, E., Karagiannis, T., Haidich, A. B. & Tsapas, A. Meta-analysis of artificial pancreas trials: methodological considerations. *The Lancet Diabetes and Endocrinology* **5**, 685 (2017).
16. Thabit, H. *et al.* Home Use of an Artificial Beta Cell in Type 1 Diabetes. *N. Engl. J. Med.*

- 373, 2129–2140 (2015).
17. Nguyen, D. *et al.* A Photo-stable and Proteolysis-Resistant FRET-Based Calcium Biosensor. (2020).
 18. Ogunlesi, T. A., Lesi, F. E. A. & Oduwole, O. Prophylactic intravenous calcium therapy for exchange blood transfusion in the newborn. *Cochrane Database of Systematic Reviews* **2017**, (2017).
 19. Vasudeva, M. *et al.* Hypocalcemia in trauma patients: A systematic review. *Journal of Trauma and Acute Care Surgery* **90**, 396–402 (2021).
 20. Jankowski, S. & Vincent, J. L. Calcium Administration for Cardiovascular Support in Critically Ill Patients: When is it Indicated? *J. Intensive Care Med.* **10**, 91–100 (1995).