UNIVERSITY OF CALIFORNIA SAN DIEGO

New Purpose for Old *Streptococcus pyogenes* Antigens: Modified Group A Carbohydrate as a Vaccine Antigen and Cas9 as a Regulator of Bacterial Virulence

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

in

Biomedical Sciences

by

Nina J. Gao

Committee in charge:

Professor Victor Nizet, Chair
Professor Partho Ghosh
Professor David Pride
Professor Manuela Raffatellu
Professor Karsten Zengler

2021

The dissertation of Nina J. Gao is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

To everyone personally victimized by *Streptococcus pyogenes*.

EPIGRAPH

You can't change the world alone - you will need some help -

and to truly get from your starting point to your destination

takes friends, colleagues, the good will of strangers and a strong coxswain to guide them.

—William H. McRaven

TABLE OF CONTENTS

Dissertation Approval Page iii

Dedication iv

Epigraph v

Table of Contents vi

List of Abbreviations ix

List of Figures xii

List of Tables xiii

List of Supplemental Files xiv

Acknowledgements xv

Vita xix

Abstract of the Dissertation xx

Chapter 1 Immunobiology of the Classical Lancefield Group A Streptococcal Carbohydrate Antigen 1
1.1 Abstract 2
1.2 Preface to Chapter 1 3
1.3 Introduction: Group A *Streptococcus* and its species defining antigen 4
1.4 Discovery of GAC and its role in diagnostics 5
1.5 GAC chemical structure and genetic basis of its biosynthesis 6
1.6 New insight into GAC function 9
1.7 Natural immunogenicity of GAC 11
1.8 Suspected role of GAC in GAS-induced autoimmunity 13
1.9 GAC as a candidate GAS vaccine antigen 15
1.10 Conclusions 18
1.11 Acknowledgements 19

Chapter 2 Envisioning a Safe, Universal, Multi-Valent Group A Streptococcal Vaccine 20
2.1 Abstract 21
2.2 Preface to Chapter 2 22
2.3 Introduction 23
2.4 Results 26
2.4.1 Expression and purification of GAS protein antigens 26

2.4.2 Generation of SpyAD-GAC^{PR} conjugate 28
2.4.3 Evaluation of GAS vaccine antigens for immunogenicity . . 30
2.4.4 Evaluation of rabbit antisera in opsonophagocytic killing, blocking SLO activity, and passive protection in murine challenge model 32
2.4.5 Multivalent immunization in mice generates antibody response with improved capacity to bind to native surface antigens and direct opsonophagocytic killing 36
2.4.6 Immunization with a multivalent glycoconjugate vaccine provides significant protection against systemic and intradermal GAS challenge 38
2.4.7 Immunization with SpyAD-GAC^{PR} does not induce antibodies cross-reactive to human heart or brain tissue epitopes . . 39
2.4.8 Multivalent conjugate combination vaccine yields protection against GBS challenge 41
2.5 Discussion 42
2.6 Methods 45
2.6.1 Bacterial strains and generation of SpyAD knockout strain . 45
2.6.2 GAC purification 46
2.6.3 Cloning, Expression and Purification of SLO, C5a peptidase, and SpyAD[4pAMF] 47
2.6.4 Multi-angle light scattering (MALS) analysis 48
2.6.5 Dibenzocyclooctyne (DBCO)-derivatization of GAC^{PR} and conjugation to SpyAD[4pAMF] 49
2.6.6 Anthrone assay for total polysaccharide concentration . . . 49
2.6.7 Generation of immunized rabbit serum 50
2.6.8 Antibody titer of rabbit serum by ELISA 50
2.6.9 IgG binding to GAS strains 51
2.6.10 Primary human neutrophil opsonophagocytic killing (OPK) assay 51
2.6.11 Primary human neutrophil oxidative burst assay 52
2.6.12 *In vivo* mouse immunization studies 53
2.6.13 Western blot analysis for heart / brain lysate cross reactivity 54
2.6.14 Statistical analysis 54
2.7 Acknowledgements 54

Chapter 3 Functional and Proteomic Analysis of *Streptococcus pyogenes* Virulence Upon Loss of Its Native Cas9 Nuclease 56
3.1 Abstract 57
3.2 Preface to Chapter 3 58
3.3 Introduction 59
3.4 Results 61
3.4.1 Generation of a GAS cas9 Mutant (Δ*cas9*) 61

3.4.2 Cas9 Deficiency Is Associated With Reduced Abundance of Key GAS Virulence Determinants and Regulatory Factors . 63
3.4.3 Loss of Cas9 Is Associated With GAS Virulence Attenuation 67
3.5 Discussion 71
3.6 Materials and Methods 75
3.6.1 Bacterial Strains and Culture Conditions 75
3.6.2 Genetic Manipulation of GAS (Construction of Δ*cas9* Strain and Δ*cas9* Complementation) 75
3.6.3 RNA Extraction and qPCR Assays 76
3.6.4 Western Immunoblot Assays 77
3.6.5 Proteomics Sample Preparation 77
3.6.6 Quantitative Proteomics, Protein Identification, and Analysis 78
3.6.7 Hyaluronic Acid Capsule Assays 79
3.6.8 Quantification of Cell Wall-Attached M Protein by Flow Cytometry 79
3.6.9 Cysteine Protease Activity Assays 80
3.6.10 Hemolysis Assays 80
3.6.11 Whole Blood Assays 81
3.6.12 Adherence Assays 81
3.6.13 Animal Experiments 81
3.6.14 Statistical Analysis 82
3.7 Acknowledgements 82

Chapter 4 Conclusions 84
4.1 Conclusions 85
4.2 Summary of results 85
4.3 The importance of revisiting basic science 87
4.4 Potential impact of my dissertation work 88

Appendix A Supplemental Information for Chapter 3 90
A.1 Supplementary Figures 90
A.2 Supplementary Tables 93

Bibliography 95

LIST OF ABBREVIATIONS

aa: amino acid

ANOVA: analysis of variance

ARF: acute rheumatic fever

BRC: baby rabbit complement

CFU: colony forming unit

Cas: CRISPR-associated genes

cat: chloramphenicol acetyltransferase gene

CDAP: 1-Cyano-4-dimethyl aminopyridinium tetrafluoroborate

CFPS: cell-free protein synthesis

CIES: carrier-induced epitopic suppression

CRISPR: clustered regularly interspaced short palindromic repeats

Da: dalton

DBCO: dibenzocyclooctyne

DCFH-DA: 2,7-dichlorofluorescein diacetate

DMEM: Dulbecco's modified Eagle's medium

ELISA: enzyme-linked immunosorbent assay

Erm: erythromycin

FBS: fetal bovine serum

FL: full length

GAC: group A carbohydrate

GAS: group A *Streptococcus*, *Streptococcus pyogenes*

GBS: group B *Streptococcus*, *Streptococcus agalactiae*

GlcNAc: *N*-acetylglucosamine

GRAB: protein G-related α2-macroglobulin-binding protein

HA: hyaluronic acid

HBSS: Hank's balanced salt solution

HiB: *Haemophilus influenzae* type B

HRP: horseradish peroxidase

IgG: immunoglobulin G

IM: intramuscular

IP or i.p.: intraperitoneal

IRB: Institutional Review Board

kDa: kilodalton

MOI: multiplicity of infection

MRSA: methicillin-resistant *Staphylococcus aureus*

nnAA: non-natural amino acid

NIAID: National Institute of Allergy and Infectious Disease

NIH: National Institutes of Health

OD: optical density

OPK: opsonophagocytic killing

pAMF: *p*-azidomethyl phenylalanine, a non-natural amino acid

PBS: phosphate-buffered saline

PBS-T: PBS + 0.05% Tween 20

PCR: polymerase chain reaction

PMA: phorbol myristate acetate

PMSF: phenylmethylsulfonyl fluoride

PSRU: polysaccharide repeating unit

PVDF: polyvinylidene fluoride

qRT-PCR: quantitative reverse transcription PCR

RAST: Rapid Annotations Subsystems using Technology server

RBC: red blood cell, erythrocytes

RHD: rheumatic heart disease

ScpA: GAS C5a peptidase

SDS: sodium dodecyl sulfate

SEC: size exclusion chromatography

SEM: standard error of the mean

SIC: streptococcal inhibitor of complement

SLO: streptolysin O

SLS: streptolysin S

TCS: two-component (regulatory) system

TAMRA: tetramethylrhodamine fluorescent dye

TBS: Tris-buffered saline

THA: Todd-Hewitt broth with 15g/L agar

THB: Todd-Hewitt broth

TM: transmembrane

TMT: tandem mass tag

LIST OF FIGURES

Figure 1.1: Biosynthesis of the group A carbohydrate (GAC), the most abundant component of the group A streptococcal cell wall 7
Figure 1.2: Roles of the group A carbohydrate (GAC) vs. its component polyrhamnose backbone (GAC^{PR}) or GlcNAc sidechain in the bacteriology, pathobiology, and immunobiology of group A *Streptococcus*. 10

Figure 2.1: Expression, purification and analysis of GAS protein antigens. 29
Figure 2.2: GAC^{PR} derivatization schematic and generation of conjugate vaccine antigen. 31
Figure 2.3: Immunogenicity of GAS vaccine antigens in rabbits. 33
Figure 2.4: Evaluation of rabbit vaccine antisera in GAS opsonophagocytic killing, blocking SLO activity, and passive protection in murine challenge model. 35
Figure 2.5: A multivalent vaccine SpyAD-GAC^{PR} + SLO(ΔC101) + C5a peptidase vaccine elicits IgG responses in mice that recognize the surface of diverse GAS strains and promote opsonophagocytic killing of M1 GAS. 37
Figure 2.6: Mice immunized by multivalent conjugate combination are protected when challenged by lethal GAS systemic infection or intradermal skin infection. 38
Figure 2.7: Assessment of antigen-specific antisera for cross-reactivity to human heart lysate. 40
Figure 2.8: Mice immunized by multi-valent conjugate combination are protected when challenged by lethal GBS systemic infection. 42

Figure 3.1: Deletion of the cas9 gene does not affect GAS growth and morphology. . . 62
Figure 3.2: Cas9-deficiency is associated with RAST functional enrichment and differential abundance of GAS virulence determinants and regulators. 64
Figure 3.3: Lack of Cas9 is impaired with significant changes in key virulence factors and pathogenic functionalities of GAS. 68
Figure 3.4: Lack of Cas9 is impaired with significant changes in key virulence factors and pathogenic functionalities of GAS. 70
Figure 3.5: Schematic representation of the network of GAS regulatory proteins and virulence factors affected upon loss of Cas9. 74

Figure A.1: Genetic complementation of Δ*cas9* GAS M1T1 5448 strain leads to marked growth defect. 91
Figure A.2: Abundance of proteins detected by tandem mass tag proteomics in wild type (WT) and Δ*cas9* GAS M1T1 5448 strains. 92

LIST OF TABLES

Table 1.1: GAS immunization strategies utilizing GAC. 17

Table A.1: Minimum inhibitory concentration (MIC) of cell wall synthesis inhibitory antibiotics against wild type (WT) or Δ*cas9* GAS M1T1 5448 strains. 94

Table A.2: Normalized abundance of all proteins detected from wild-type and Δ*cas9* GAS 5448 strains by tandem mass tag proteomics. 94

Table A.3: Normalized abundance of proteins that were identified to be significantly enriched either in wild type (WT) or Δ*cas9* GAS 5448 strains. 94

LIST OF SUPPLEMENTAL FILES

A2_TMTproteomics.xlsx

A3_RASTsignificant.xlsx

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge Dr. Victor Nizet as my mentor and support throughout my PhD. Thank you for your guidance over these past 6 years. I'm grateful for the opportunity to not only expand my scientific repertoire, but also to work with a biotechnology company and potentially make a difference in the world.

I would next like to thank my thesis committee: without your guidance and support, I would not have been motivated to finish my PhD studies. I am grateful for you taking the time to be a part of my scientific career. Thank you to Dr. Partho Ghosh for making me feel welcome when I collaborated in my first rotation here at UC San Diego. Thank you to Dr. Karsten Zengler who frequently caught me eating lunch at odd hours and showed genuine concern for my well-being. Thank you to Dr. David Pride who not only provided guidance with my research, but also in being a patient teacher when I attended clinical microbiology rounds to learn more about medical microbiology. Thank you to Dr. Manuela Raffatellu for encouraging me to pursue further studies; I hope you find that I've grown since my undergraduate days researching in your lab.

I'd like to acknowledge and thank Professor Bruce Hamilton for acting as a mentor for me through the genetics training program. You always asked the tough questions and encouraged us to think critically and grow as people.

I am grateful for the clinical microbiologists who I have come to know. I'm grateful the learning opportunities I have had at clinical microbiology rounds with UC San Diego Health at the Center for Advanced Laboratory Medicine (CALM). Thank you to the microbiology lab directors Drs. Susan Realegeno, David Pride, and Sharon Reed for making me feel welcome. I'd also like to thank biomedical sciences PhD alumni Omai Garner and Morgan Pence for meeting/speaking with me and giving advice as I explored clinical microbiology as a career. I look forward to working in the realm of medical microbiology in the future, and hope to see you all again.

I'd like to give special thanks to my former PI: Dr. Daniel Salomon. Thank you for believing in me (even when I felt I was struggling), and supporting me as I transitioned from a

research technician to a PhD graduate student. I wish you could see me and how far I've come since I started working in your lab at the Scripps Research Institute.

Another special thank you to Ervin Rodas Lima, an undergraduate student whom I had the opportunity and honor to mentor. Your curiosity and drive reinvigorated my drive for research and knowledge. I won't forget your excitement and eagerness to learn.

I'd like to thank all my current and former labmates for making my lab experience fun and entertaining. I apologize, I will likely unintentionally leave people out, but it does not mean I was not grateful. I'd like to acknowledge my baymates over time: thank you to Janet Liu, Angelica Riestra, Chih-Ming Tsai, Bryan Hancock, Jwa-Kyung Kim, and Alison Vrbanac for all the (reassuring, supportive, fun, and on occasion scientific) conversations. Thank you to Alison Coady for supporting me scientifically and inspiring me to stay fit and even run a half-marathon, something I never expected I could do. Thank you to Ross Corriden for the conversations and advice for life in lab and out. And of course, a shout out to everyone who's made the PhD lab experience better (in no particular order): Elisabet Bjånes, Chris LaRock, Doris LaRock, Leo Lin, Nick Dillon, Stephanie Brandt (Heberlein), Katy Patras, Andrés Valderrama, Sarah Burgdorf, Raymond Zurich, Anne Phan, Lucy Catteau, Nitasha Menon, Emilia Hoste, Anel Lizcano, Ingrid Cornax, Yohei Kohno, Anabel Flores, Marvic Carillo Terrazas, Sudeshna Shah, Daniel Sandoval, and Joshua Olson.

Thank you to my labmates in Professor Mark Walker's lab at University of Queensland in St. Lucia, Queensland, Australia. Tania Rivera-Hernandez, Andrew Turner, Amelia Soderholm, Mira Syahira, Hanna Richter, Mark Walker, and more: you all made me feel very welcome during my 3 month stay, despite my "Americanisms".

Science does not happen in a vacuum, and I could not have accomplished what I have without support from my friends in San Diego and/or pre-graduate school life. Special thanks for the orange crew: Hannah Tsunemoto, Roland Liu, and Samuel Lin - I'm happy that we met through graduate courses and immediately got along. I'm ever grateful for our lifelong

friendship. Thank you to my friends in the biomedical sciences PhD program; there's really no appropriate way for me to thank you for the fun and scientific adventures we have had - thank you Alison Vrbanac, Lisa Marotz, Gregory Golden, Angeline Chen, Ryan Marina, Robert Mills, Anastasia Gromova, Anthony Vu, Christina Alarcon, Saugat Poudel, Dana Steffen (Ramms), Elie Farah, Sarah Shires, Samantha Jones, Gabriela Goldberg, Samantha Jones, Cayce Dorrier, Natalie Hollingsworth, Doug Zuill, Gabriel Ascui, and many more. Thank you to all the non-BMS graduate friends I've come to know: Antonia Darragh, Chris Piggott, Brooke Anderson, Gina Powers, JR Caldera, and anyone else I may have forgotten. Shout out to our bar trivia team The Durangos (Chase Marotz, Jake Banfield-Weir, Geoffrey Berz, Adrian Harris, Alison Vrbanac, Natalie Hollingsworth, Andy Ostericher, and Anastasia Gromova) for giving me something to look forward to on Wednesday nights, and for your patience despite my lack of pop culture knowledge. And of course, thank you to my friends from before graduate school, for their patience and support for me (even when they didn't quite understand what I as doing). Thank you to Ivan Check, Pyrrha Arrastia, Grace Kim, Alina Shiotsu, Angela Gao, Sabrina Tran, Grant Taylor, Jonathan Nguyen, John Nguyen, Franny Acosta, Dorian Kim, Thar Soe, Beck Levin, Anthony Aguilar, and Erica Carry for all your kindness and empathy; you all accommodated my strange work schedule and also kept me going. I could go on forever thanking you all.

Last but not least, I'd like to acknowledge my family, particularly my siblings (Diana Gao and Frank Gao) and my grandmother. You all have supported my scientific career from the beginning, even when none of you really understood what I wanted to study or do. Also, thank you to my aunt Kathy, uncle Steve, and my parents. We didn't always agree on my goals and pursuits, but I'd like to think you always wanted what was best for me.

Chapter 1, in full, is submitted for publication of material as it may appear in *Infection and Immunity*. "**Nina J. Gao**, Ervin Rodas Lima, and Victor Nizet. *Immunobiology of the Classical Lancefield Group A Streptococcal Carbohydrate Antigen.*" The dissertation author was a primary investigator and author of this work.

Chapter 2 contains some material being prepared for submission for publication and some material is, in part, a reprint of the material as it appears in *Infectious Microbes and Disease*. "**Nina J. Gao**, Satoshi Uchiyama, Lucy Pill, Samira Dahesh, Joshua Olson, Leslie Bautista, Shilpa Maroju, Aym Berges, Janet Z. Liu, Raymond H. Zurich, Nina van Sorge, Jeff Fairman, Neeraj Kapoor, Victor Nizet. *Site-Specific Conjugation of Cell Wall Polyrhamnose to Protein SpyAD Envisioning a Safe Universal Group A Streptococcal Vaccine*. Infect Microb Dis 2020;00(00):00–00. doi: 10.1097/IM9.0000000000000044". The dissertation author was a primary investigator and author of this work.

Chapter 3, in full, is a reprint of material as it appears in *Frontiers in Microbiology* 2019 by **Nina J. Gao**, Mahmoud M. Al-Bassam, Saugat Poudel, Jacob M. Wozniak, David J. Gonzalez, Joshua Olson, Karsten Zengler, Victor Nizet, and J. Andrés Valderrama. *Functional and Proteomic Analysis of Streptococcus pyogenes Virulence Upon Loss of Its Native Cas9 Nuclease*. DOI: 10.3389/fmicb.2019.01967. The dissertation author was the primary investigator and author of this paper.

VITA

2013	Bachelor of Science in Microbiology & Immunology *cum laude*, University of California Irvine
2013-2015	Research Technician, Laboratory of Daniel R. Solomon, The Scripps Research Institute, La Jolla
2016-2017	Graduate Research Grant from UC San Diego Global Health Institute
2017-2019	Training Grant in Genetics, University of California San Diego
2021	Doctor of Philosophy in Biomedical Sciences, University of California San Diego

PUBLICATIONS

Nina J. Gao, Satoshi Uchiyama, Lucy Pill, Joshua Olson, Leslie Bautista, Shilpa Maroju, Aym Berges, Janet Z. Liu, Raymond H. Zurich, Nina van Sorge, Jeff Fairman, Neeraj Kapoor, Victor Nizet. "Site-Specific Conjugation of Cell Wall Polyrhamnose to Protein SpyAD Envisioning a Safe Universal Group A Streptococcal Vaccine." *Infect Microb Dis*. 2020 Dec 30;00:00000. (Early online.)

Nina J. Gao, Mahmoud M. Al-Bassam, Saugat Poudel, Jacob M. Wozniak, David J. Gonzalez, Joshua Olson, Karsten Zengler, Victor Nizet, J. Andrés Valderrama. "Functional and Proteomics Analysis of Streptococcus pyogenes Virulence Upon Loss of Its Native Cas9 Nuclease." *Front Microbiol*. 2019 Aug 22;10:1967.

J. Andrés Valderrama, Angelica M. Riestra, **Nina J. Gao**, Christopher N. LaRock, Naveen Gupta, S. Raza Ali, Hal M. Hoffman, Partho Ghosh, Victor Nizet. "Group A streptococcal M protein activates the NLRP3 inflammasome." *Nat Microbiol*. 2017 Oct; 2(10):1425-1434.

Ling-Juan Zhang, George L. Sen, Nicole L. Ward, Andrew Johnston, Kimberly Chun, Yifang Chen, Christopher Adase, James A. Sanford, **Nina Gao**, Melanie Chensee, Emi Sato, Yi Fritz, Jaymie Baliwag, Michael R. Williams, Tissa Hata, Richard L. Gallo. "Antimicrobial Peptide LL37 and MAVS Signaling Drive Interferon-beta Production by Epidermal Keratinocytes during Skin Injury." *Immunity*. 2016 Jul 19;45(1):110-30.

Stefan Jellbauer, Araceli Perez Lopez, Judith Behnsen, **Nina Gao**, Thao Nguyen, Clodagh Murphy, Robert A. Edwards, Manuela Raffatellu. "Beneficial Effects of Sodium Phenylbutyrate Administration during Infection with *Salmonella enterica* Serovar *Typhimurium*." *Infect Immun*. 2016 Aug 19;84(9):2639-52.

Thomas C. Whisenant, Eigen R. Peralta, Lauren D. Aarreberg, **Nina J. Gao**, Steven R. Head, Philip Ordoukhanian, James R. Williamson, Daniel R. Salomon. "The Activation-Induced Assembly of an RNA/Protein Interactome Centered on the Splicing Factor U2AF2 Regulates Gene Expression in Human CD4 T Cells." PLoS One. 2015 Dec 7; 10(12):e0144409.

ABSTRACT OF THE DISSERTATION

New Purpose for Old *Streptococcus pyogenes* Antigens: Modified Group A Carbohydrate as a Vaccine Antigen and Cas9 as a Regulator of Bacterial Virulence

by

Nina J. Gao

Doctor of Philosophy in Biomedical Sciences

University of California San Diego, 2021

Professor Victor Nizet, Chair

Group A *Streptococcus* (GAS, *Streptococcus pyogenes*) is one of the most important human bacterial pathogens, causing hundreds of millions of infections around the globe annual in a wide spectrum of disease presentations. In this dissertation work, I will examine two different molecules of GAS: the group A carbohydrate (GAC) and Cas9 nuclease. Both molecules were initially discovered decades ago, but regained popularity and interest in recent years with re-examination of potential roles in pathogenesis and virulence.

GAC has since been explored beyond its structural function in the streptococcal cell, with putative roles in evasion of host immune factors and development of lasting host immunity. The

highly conserved and essential nature of the GAC posits the molecule as a candidate antigen for a universal GAS vaccine. However, due to the prescence of *N*-acetylglucosamine (GlcNAc) sidechain on GAC, which is a hypothesized driver of post-streptococcal autoimmune sequelae, modification of the GAC to remove the autoreactive epitope is necessary to make a safe antigen (denoted as GAC^{PR} for the polyrhamnose alone).

To develop our safe and immunogenic GAC-conjugate, we collaborated with Vaxcyte Inc. to use cell-free protein synthesis technology to incorporate non-native amino acids (nnAA) as site-specific attachment sites for the polysaccharide. This technology allows for the generation of precise glyco-conjugates with GAC^{PR} covalently linked to streptococcal virulence factor SpyAD. Finally, I test our SpyAD-GAC^{PR} conjugate for efficacy in a multi-valent vaccine in a primary neutrophil opsonophagocytic assay and murine infection challenges.

Cas9, now famous for the precise gene-editing technology, was originally discovered in the GAS type IIA CRISPR system. Notably, this CRISPR system was frequently found in bacterial pathogens and evaluated for potential roles in virulence regulation and control. We examine the role of Cas9 with regards to GAS pathogenesis. Using proteomics, we identify key protein regulators of GAS virulence networks significantly reduced virulence due to the absence of endogenous Cas9. Follow-up studies confirm Cas9-deficient GAS has reduction of specific virulence factors and significantly reduced virulence in whole human blood and mouse skin infection model, providing strong evidence that Cas9 plays a regulatory role in virulence, though the precise molecular mechanism is unknown.

Chapter 1

Immunobiology of the Classical Lancefield Group A Streptococcal Carbohydrate Antigen

1.1 Abstract

Streptococcus pyogenes, also known as group A Streptococcus (GAS), is one of the preeminent human bacterial pathogens causing hundreds of millions of infections each year worldwide. The bacterium is easily identified and diagnosed by a rapid antigen test against the group A carbohydrate (GAC), the species-defining polysaccharide found in the bacterial cell wall. The GAC is comprised of a polyrhamnose backbone with N-acetylglucosamine (GlcNAc) sidechains and has been found on all clinical isolates of GAS; the high degree of conservation of the biosynthetic operon and corresponding carbohydrate suggests it plays a critical role for bacterial survival. Despite the discovery of the GAC in the 1930s and the widespread prevalence of the rapid diagnostic test reliant on the GAC, few studies explored the GAC beyond chemical structure and essential structure within the cell wall until renewed interest in GAS vaccine development sparked studies in GAS pathogenesis and GAC immunogenicity. This review summarizes the current literature on GAC structure, biosynthesis, natural immunogenicity in human infections, its role in GAS pathogenesis and virulence, and implications for vaccine development.

1.2 Preface to Chapter 1

This chapter serves to summarize the current literature on the group A carbohydrate, the Lancefield antigen that serves as a molecular identifier of *Streptococcus pyogenes*. To date, this is the first literature review focusing on this antigen and its roles for GAS. Though much progress has been made recently, there are still many knowledge gaps that are crucial to understand the role of GAC in pathogenesis of GAS infections or autoimmune disease. However, due to the urgent need for an effective and safe GAS vaccine, these weaknesses should not deter the potential for GAC as an universal vaccine antigen.

1.3 Introduction: Group A *Streptococcus* and its species defining antigen

Streptococcus pyogenes, commonly known as group A Streptococcus (GAS), is a preeminent human pathogen causing hundreds of millions of infections each year worldwide. The tremendous global disease burden of this gram-positive bacterial pathogen is skewed significantly toward resource-limited parts of the world [161, 180]. The most common GAS disease manifestations are superficial mucosal infections, in particular pharyngitis ("strep throat"), and skin infections, notably impetigo, which can be self-limiting or effectively managed with oral or topical antibiotics. However, GAS also has significant invasive disease potential, and can disseminate through deep tissues or the bloodstream to cause sepsis, necrotizing fasciitis or streptococcal toxic shock syndrome, potentially life-threatening conditions without medical or surgical intervention. And unique among human bacterial pathogens in terms of scale, GAS can trigger hallmark post-infectious, immunologically-mediated pathologies, in particular rheumatic heart disease (RHD), that represent a major source of morbidity and mortality throughout many parts of the developing world. New strategies for effective treatment and prevention of GAS infection and its complications remain a major public health priority, with vaccines targeting the pathogen still in early stages of development or evaluation.

The global success of GAS as a human pathogen and its wide array of disease manifestations reflect a complex host-pathogen interaction, with numerous bacterial virulence factors and toxins that act in concert to promote epithelial attachment and barrier disruption, resist innate clearance mechanisms, and provoke host immune and inflammatory responses [180]. Among these bacterial factors is a unique and highly abundant cell wall component, the group A carbohydrate (GAC), which is universally conserved and indeed defines the species. GAC is the basis for accurate and rapid clinical detection of GAS infection, essential for prescribing prompt and effective treatment to reduce the risk of invasive infection and post-infectious

immune-mediated complications with potential lifelong consequences [6, 32]. However, despite its utility as an immunodiagnostic and a potential role in triggering autoimmune cross-reactivity post-GAS infection (see later section), little else was understood about GAC until recent years. New genetic and chemical structure discoveries have shepherded great progress in understanding the GAC biosynthetic process, its contribution to virulence, and safely harnessing its natural immunogenicity in vaccine design. This review serves to summarize the current literature on GAC, the eponymous antigen that both defines and provides structural integrity to GAS.

1.4 Discovery of GAC and its role in diagnostics

In the early days of microbial diagnostics, the differing species of β-hemolytic streptococci were virtually indistinguishable, and became grouped together in a single rubric as "*Streptococcus hemolyticus*". However, during the 1920s and 1930s, Dr. Rebecca Lancefield discovered a series of bacterial cell wall and surface antigens as a means to identify and classify the β-hemolytic streptococci. Originally categorized as "C-substance", these "group antigens" were eventually identified to be carbohydrate in nature through her experiments using enzymatic digests [95]. By the 1930s, a molecular basis for streptococci group classification was established and carried out by antibody-mediated latex bead agglutination and antiserum protection tests [96]. Antibody recognition of group antigens quickly found application for identification of clinically relevant pathogenic streptococci [104], and remains to this day the gold standard for rapid antigen detection and diagnostics for GAS.

The original classification of β-hemolytic streptococci into groups A through M has undergone significant reorganization as the bacterial species were more precisely characterized and genomic comparisons became possible. Though many β-hemolytic Streptococcus group classifications were not specific to a single species, the group A streptococci initially comprised only one species: *Streptococcus pyogenes*, producing GAC as a species-unique identifier. Despite

the long history of reliable pathogen identification dependent on the GAC, it is only in the last decade that the GAC-encoding operon and biosynthesis process have been characterized largely due to advancements in streptococcal genetics and glycobiology techniques, as well as renewed interest in the potential of GAC as an effective vaccine antigen. Interestingly, anti-GAC antibody cross-reactivity is now recognized from certain rare group C-variant streptococci strains (*S. dysgalactiae* subsp. *equisimilis*) [160, 117] and *Streptococcus castoreus*, a recently identified group A streptococcal bacterium isolated from beavers [100]; both bacteria have genetic loci homologous to the operon of GAS genes essential for the biosynthetic pathway.

1.5 GAC chemical structure and genetic basis of its biosynthesis

The thick gram-positive GAS cell wall contains peptidoglycan, lipoteichoic acids, and various structural proteins, but GAC itself makes up an remarkable 30-50% of cell wall composition by weight [116]. Antibody binding experiments have found that the GAC is localized primarily to the outermost surface of the cell wall [116], but also intercalated within a mesh-like structure to the peptidoglycan [121], to which GAC is anchored via phosphodiester bonds [74]. Chemical composition analyses indicate that GAC is composed of a linear polyrhamnose chain decorated with N-acetylglucosamine (GlcNAc) sidechains [116]. Polyrhamnose comprising the GAC backbone are connected by alternating $\alpha - L - (1 \rightarrow 3)$ and $\alpha - L - (1 \rightarrow 2)$ glyosidic linkages, with the β-D-GlcNAc attached to every other residue on the rhamnose backbone at the 3-position [91]. A recent study also detected glycerol phosphates (GroP) present on the C6-hydroxyl group of approximately 25% of GlcNAc [46], a significant modification previously undetected due to harsh extraction methods classically used for GAC purification.

GAC biosynthesis is encoded by a 12 gene cluster (gac operon) that is highly conserved across GAS genomes. A recent analysis found that 97% of GAS genomes (2,017 of 2,083) had

high sequence conservation of $> 70\%$ sequence homology for the entire 12 gene cluster [38], further supporting conclusions from a smaller dataset [174]. The first 7 genes of the *gac* operon (*gacABCDEFG*) encode for synthesis of the core polyrhamnose structure and are conserved across groups A, B, C, and G streptococci [190]. In GAS, some genomes exhibited frameshift mutations within several gac genes, suggesting that not all genes are essential for survival and suggesting the potential existence of compensatory genes [38, 71]. To date, not all the genes or their products have been unambiguously delineated to specific roles, but recent work has expanded our knowledge of individual gene functions and the GAC biosynthesis process (summarized in Figure 1).

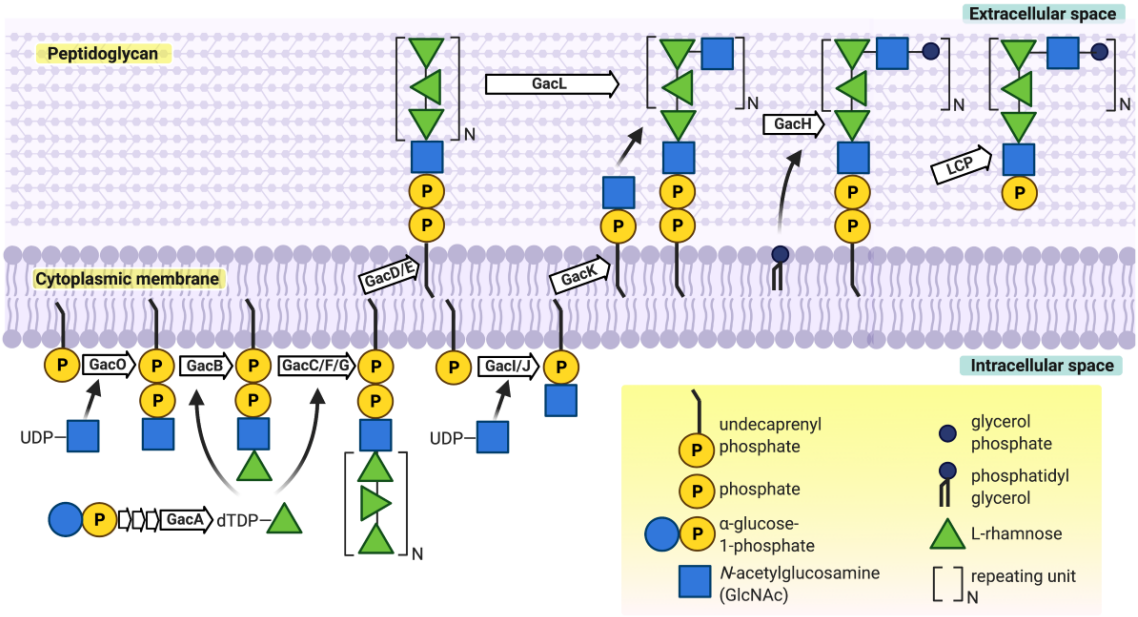

Figure 1.1: **Biosynthesis of the group A carbohydrate (GAC), the most abundant component of the group A streptococcal cell wall** A schematic diagram that summarizes the current literature on GAC biosynthesis. Enzymatic processes are noted by gene products in white arrows. The polyrhaman backbone is assembled on the inner leaflet of the cytoplasmic membrane by GacBCFG before flipped to the outer leaflet by GacDE complex. The GlcNAc sidechain and glycerol phosphates are added onto the GAC by GacL and GacH, respectively, before a Lytr-CpsA-Psr (LCP) family enzyme transfers and attaches the completed GAC to the peptidoglycan via a phosphodiester bond. Created with BioRender.com

Similar to many polysaccharides exposed on the bacterial outer surface, including capsular polysaccharides and wall teichoic acids, GAC synthesis commences with GacO affixing GlcNAc to undecaprenyl phosphate (UDP) on the intracellular side of the cell membrane [121]. The gacB gene encodes a rhamnosyltransferase that synthesizes the committed step in the GAC biosynthesis: translocation of the first rhamnose residue onto the membrane-bound UDP-GlcNAc acceptor in the inner leaflet of the bacterial cell membrane [190]. Free dTDP-L-rhamnose is produced from α-glucose-1-phosphate by the enzymes encoded by gacA and an operon located distally from GAC operon but well-studied for rhamnose synthesis, the *rmlABC* operon [171]. GacA is a metal-independent dTDP-4-dehydrorhamnose reductase enzyme, and unlike homologues in the RmlD family, GacA uniquely functions as a monomer instead of a homodimer [171]. Sequential elongation of the polyrhamnose chain is mediated by glycosyltransferases GacC, GacF, and GacG [190]. GacD and GacE encode heterodimers of an ATP-dependent ABC transporter, actively translocating the completed polyrhamnose chain to the extracellular side of the cell membrane.

The remaining five genes of the gac operon are predicted to encode functions extending off the rhamnan chain. GacJ, a small protein associated with the bacterial membrane, complexes with GacI for improved catalytic efficiency to produce free UDP-GlcNAc [190]. Though UDP-GlcNAc can freely diffuse across the cytoplasmic membrane on its own, it is also transported by the Wzx family flippase enzyme GacK to the extracellular side of the membrane [145]. GacL, a putative glycosyltransferase, uses the UDP-GlcNAc substrate to link GlcNAc to the polyrhamnose [145]. GacH cleaves phosphatidylglycerol to release and attach glycerol phosphate to C6 on approximately 25% of GlcNAc, potentially to reinforce attachment and stability of GAC on the cell wall, though the precise function is unclear [46, 74]. Finally, the GAC is transferred and covalently linked to the peptidoglycan layer via phosphodiester bond by an enzyme of the Lytr-CpsA-Psr (LCP) family. The final glycan has a reported mass that correlates to an estimated 18 trisaccharide repeating units [84], though different purification methods result in varying

average polysaccharide sizes [70, 84].

1.6 New insight into GAC function

For decades, GAC was assumed to function solely in the structural integrity of the streptococcal cell wall. The polysaccharide is undeniably a crucial cell wall component, comprising up to 50% of its mass [116]. Depletion of L-rhamnose [171] or deletion of genes required for synthesis GAC polyrhamnose backbone [174] are lethal to the bacterium, demonstrating the essential role of GAC for cell wall viability. Monte Carlo simulations on NMR spectroscopy data predict a conformationally restricted polysaccharide [165], consistent with GAC acting as a rigid structural support framework for the bacterial cell wall. Functional role(s) of GAC beyond cell wall structure are less well delineated, in part due to the interconnected network that GAC biosynthesis shares with synthesis of other key cell wall glycopolymers including peptidoglycan [121]. However, high conservation of polyrhamnose cell wall polysaccharides (including GAC) in the order Lactobacillales provides supporting evidence that GAC may functionally replace wall teichoic or teichuronic acid structures found in other gram-positive bacteria but absent in β-hemolytic streptococci [17, 121]. This notion implies GAC performs key roles in pathogenesis, cell shape, regulation of cell division, and other aspects of bacterial cell wall physiology [14].

Beyond the polyrhamnose backbone that is considered essential, the entire GAS *gac* operon is highly conserved and all human clinical isolates express native GAC with its intact GlcNAc sidechain. However, examples in which the polysaccharide will lose GlcNAc after serial passage in mice and rabbits have been described [119], suggesting a role for the sidechain in human-specific pathogenicity or immune evasion. Specific epitopes of various rhamnose cell wall polysaccharides, including other streptococcal group polysaccharides, promote bacterial resistance to immunological clearance, though molecular mechanisms are not yet defined [121].

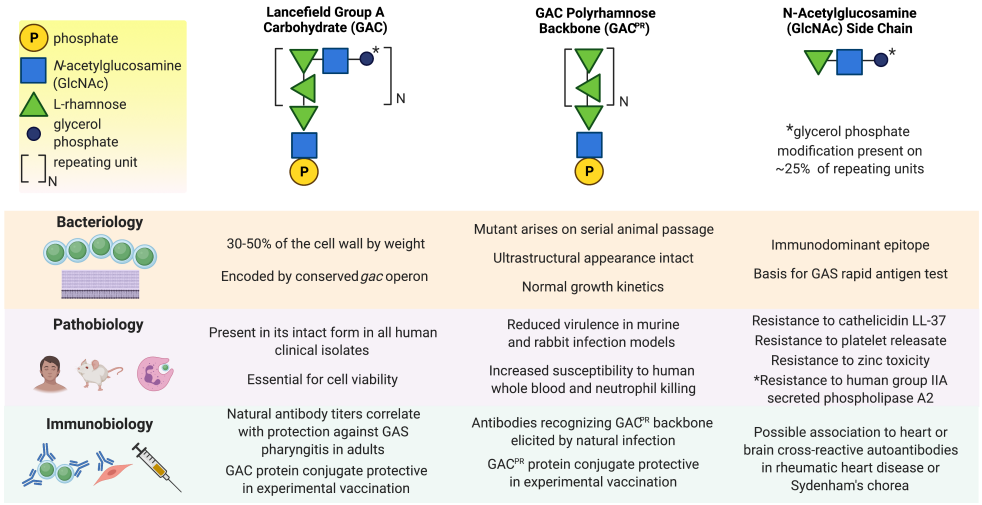

Figure 1.2: **Roles of the group A carbohydrate (GAC) vs. its component polyrhamnose backbone (GAC^{PR}) or GlcNAc sidechain in the bacteriology, pathobiology, and immunobiology of group A *Streptococcus*.**

The discovery of the *gac* operon [174] for the first time allowed the generation of an isogenic GAS mutant lacking the GlcNAc sidechain but retaining the polyrhamnose backbone, akin to the "A variant" strains isolated following animal passage [119]. The GlcNAc-deficient (ΔGAC) mutant had similar ultrastructural appearance to the parental strain on transmission electron microscopy, and did not exhibit a general defect in cell wall integrity, as susceptibility to autolysis, reactive oxygen species, lysozyme, nafcillin or vancomycin were equivalent to wild-type [174]. GAS growth in various media was unaffected by loss of the GlcNAc sidechain, but the average chain length of the mutant strains was increased [174].

Despite these many phenotypic similarities, a ΔGAC mutant in the hypervirulent, globally-disseminated M1T1 GAS background was markedly attenuated for virulence compared to the parent strain in both murine systemic infection and rabbit pneumonia models [174]. Moreover, the isogenic ΔGAC mutant was more susceptible than the wild-type GAS strain to killing in human whole blood and by purified human neutrophils, corroborating a function of the GlcNAc sidechain in resistance to immune clearance. Mechanistic studies showed increased killing of GAS to human

platelet releasate, which is rich in cationic antimicrobial peptides, and to the cationic human defense peptide cathelicidin (LL-37); surface plasmon resonance confirmed greater binding of LL-37 to the ΔGlcNAc mutant GAC compard to wild-type GAC polysaccharide [174]. Later, it was found the GroP modification present on some GAC GlcNac sidechains is required for cationic bactericidal enzyme human Group IIA secreted phospholipase A2 (hGIIA) to efficiently act against GAS; hGIIA killed GAS strains lacking GlcNAc at one-tenth the concentration that is lethal for wild-type GAS [46, 172]. The GacH required for GroP decoration also conveyed resistance to zinc toxicity [46], though the mechanism of protection is not yet understood.

GlcNAc is a common sugar present in glycan structures on mammalian cell surfaces and extracellular matrix, and GAC sidechain is hypothesized to play a role in mimicking human epitopes, perhaps helping GAS avoid immune detection by masking polyrhamnose, a non-human glycan motif found in bacteria [121], fungi [111], and protozoan parasites [162]. Though human lectins with a precise specificity for rhamnose or polyrhamnose have yet to been identified [121], another hypothesis suggests that the GlcNAc may directly bind host receptors to skew the immune response in favor of bacterial survival within the host, as documented with the GlcNAc present in the lipooligosaccharide of *Neisseria gonorrhoeae* [175]. It is important to note that the GlcNAc sidechain on GAC is not an universal pre-requisite for virulence; the degree to which loss of GlcNAc affects GAS susceptibility to innate immune clearance is strain-dependent [71].

1.7 Natural immunogenicity of GAC

Following GAS infection, humans naturally develop antibodies against various GAS antigens, including GAC. Among these, antibody titers to streptolysin O (SLO), a cholesterol-dependent cytolysin produced by GAS commonly serves as a clinical standard to determine recent GAS infections. When humans develop antibodies against GAC, serum binding assays indicate the anti-GAC antibody peak is slower and more modest compared to that of SLO. For example,

one study found the anti-SLO titer increased significantly post-GAS infection (228%) compared to a modest increase in anti-GAC titer (22%) [189]. Despite a strong antibody response to GAS infections, the anti-SLO titer level peaks in children and declines with age [189], suggesting that while anti-SLO antibodies protect in acute or active infections, they may not provide lasting immunity or protection. In contrast, mean antibody titers to the GAC in humans increase slowly with age, peaking between 14 and 17 years of age before decreasing slightly in the twenties age group [189]; these serum titers strongly correlate with a reduced rate of GAS infection after age 17 [151]. Therefore, antibodies recognizing GAC may be important in the development of lasting immunity against GAS infections, despite weak natural immunogenicity compared to other GAS antigens.

The critical factors for the development of immunity against GAC are not well understood. As is the case for other polysaccharides, the molecular weight of GAC fragments plays a role in immunogenicity and the accompanying inflammatory response. Immunization of rabbits with synthetic GAC polysaccharides of varying lengths showed a significant protective response with a hexasaccharide, equivalent to two repeat units [84]. The hexasaccharide GAC may represent the minimum sufficient antigen as it contains all possible structural motives of the repeating subunits [84]. In a rat arthritis model, rats injected with larger polysaccharide chains experienced acute edema and arthritis, while an equivalent mass of polysaccharides composed of smaller fragments induced arthritis but no edema [22]. The inflammatory response to GAC thus varies depending on the length of the chain, likely due to the differences in molecular recognition and processing of the polysaccharide antigen. The trisaccharide unit branch point and the size of the total polysaccharide determine the epitopes to which anti-GAC antibodies bind [165].

Most antibodies directed against the GAC are specific to the GlcNAc moiety, the immunodominant epitope of the GAC and the basis for the Lancefield group assignment [151]. Early immunization experiments with A variant (animal-passaged) GAS strains lacking GlcNAc confirmed that these strains induced antisera against the polyrhamnose on GAS [118]. Antibodies to

purified native GAC are protective in various infection models and opsonizing across different M protein serotypes in phagocytosis assays [151]. Likewise, anti-rhamnan antibodies raised against ΔGAC enhance neutrophil opsonophagocytic killing of multiple *emm* serotypes GAS in vitro and protect against lethal challenge through passive immunization in a murine infection model [174]. As the terminal, β-linked GlcNAc sugars play such a dominant role in the serological specificity of the polysaccharide, concerns for immunological cross-reactivity with mammalian connective tissues containing this sugar motif have been raised [118], and are discussed in the subsequent section.

1.8 Suspected role of GAC in GAS-induced autoimmunity

GAS infections can trigger significant post-infectious, immune-mediated disease sequelae, in particular acute rheumatic fever (ARF) which if recurrent can lead to rheumatic heart disease (RHD), a major cause of cardiovascular morbidity and mortality in several resource poor regions of the world. ARF may arise following 3-6% of GAS pharyngitis infections that are not promptly treated, with the autoimmune reaction affecting the heart, joints, skin and/or central nervous system [48], the latter including the syndrome of Sydenham's chorea, reflective of basal ganglia dysfunction [137]. Uncommon in adult subjects, ARF is most frequently seen in children and adolescents aged 5-15 years of age [32, 33], coinciding with peak incidence of GAS pharyngitis cases. ARF and progression to RHD can be prevented with the use of primary or secondary antibiotic prophylaxis, respectively, to target GAS [12]. Current diagnostic criteria for ARF, originally published in 1944 as the Jones criteria and updated in 1992, feature a list of major and minor clinical manifestations, but also require documentation of recent GAS infection via positive throat culture, positive rapid antigen test, or elevated/rising streptococcal antibody titer (typically anti-streptolysin O and sometimes anti-DNaseB) [15].

Original studies by Goldstein isolated glycopeptides from human heart valves that share

immunological properties with the GlcNac sidechain of GAC [62]. Anti-streptococcal monoclonal antibodies derived from human ARF patients that reacted with heart cells were then noted to bind the glycoprotein laminin found in the basement membrane underlying human heart valve endothelium [56]. Further studies found that responses against GlcNAc were strongly linked to antibody responses against cardiac myosin and other α-helical coiled-coil proteins, including streptococcal M proteins [31, 159]; these α-helical peptide motifs are hypothesized to underpin cross-reactivity between GAC epitope and the myocardium or heart valves [1, 56].

Individual monoclonal antibodies against the GlcNAc epitope have been isolated from patients with rheumatic carditis [1] and Sydenham's chorea, one of the Jones criteria for ARF that affects neuronal cell signaling [89], suggesting that the same epitope may be responsible for multiple ARF manifestations. Of note, one such mAb from an ARF patient that recognized native GAC on the wild-type GAS surface did not recognize an isogenic GAS knockout mutant that lacked the GlcNAC sidechain [174]. Overall, ARF patients had 2- to 3-fold higher concentrations of anti-GAC antibodies at the initial time point of illness compared to GAS pharyngitis patients [112], and elevated levels of serum antibodies recognizing GAC persisted for longer periods of time in patients with RHD compared to the normal rate of decline in ARF patients without carditis [43].

However, the presence of cross-reactive antibodies in and of itself does not establish a central role in promoting ARF/RHD pathogenesis; a sustained immune response must also be driven to break tolerance. Models have been proposed in which anti-GlcNAc antibodies attach to the valvular endothelium and trigger upregulation of CXCL9/Mig, which attracts CD8+ and CD4+ T cells, and vascular cell adhesion molecule-1 (VCAM-1), which facilitates the extravasation of these T cells into the heart valve tissue (reviewed in [42]). However, a comparison of sera from normal rabbits and GAC-immunized rabbits demonstrated no significant difference in antibody binding to fixed human tissues [147] or human cardiac tissue lysates via ELISAs [174], suggesting that the GAC may not be sufficient or solely responsible for the development of ARF.

A clinical trial for a M-protein based GAS vaccine in the 1960's sounded alarm when 3 of the 21 participants developed suspected or definite ARF [114, 115], leading the US Food and Drug Administration to issue a ban on GAS vaccine development that was not lifted until after an expert panel brought together in 2005 by the US National Institute of Allergy and Infectious Disease demonstrated the therapeutic potential and need for a GAS vaccine [11]; research was permitted to proceed, however the panel clearly established the need for safety and testing for autoreactivity as a special requirement specific for GAS vaccines. The incompletely understood mechanism(s) of ARF/RHD immunopathology remain critical knowledge gaps, and biomarkers to detect early signs of these diseases are urgently needed to help steer the development of a safe and effective GAS vaccine [110].

1.9 GAC as a candidate GAS vaccine antigen

GAS differs from other prominent human streptococcal pathogens, notably group B *Streptococcus* (GBS) and *Streptococcus pneumoniae*, which express immunodominant exopolysaccharide capsules that confer serotype specificity, promote virulence and are validated as vaccine antigens [149]. Rather, the polysaccharide capsule expressed by the vast majority of GAS strains is composed of high molecular weight hyaluronic acid, a ubiquitous component of host tissues, and therefore immunologically inert [183]. Rather, GAS serotype specificity is conferred by the hypervariable N-terminal domains of the surface M protein, encoded by more than 200 *emm* gene variants, among which cross-protection is not ensured [153, 163]. Against these complexities, the species defining GAC antigen has several hallmarks of an ideal GAS vaccine candidate. GAC is universally conserved, essential for GAS survival, abundant on the bacterial cell surface, and accessible to antibody binding regardless of the degree of encapsulation [84]. Whereas human sera rarely contain antibodies to multiple M protein serotypes, IgG antibody against GAC, present in sera from children from diverse geographic areas, promoted opsonophagocytosis of several M

type-specific GAS strains in a titer-dependent manner [151]. That said, native GAC has two key limitations as a vaccine antigen: (i) large polysaccharides alone are not sufficiently immunogenic to generate a robust vaccine immune response, and (ii) there are potential safety concerns due to potential auto-reactivity related to ARF/RHD discussed above.

Polysaccharides covalently linked to a carrier, including immunogenic proteins, lipopeptides, or gold nanoparticles, can become effective vaccine antigens, such as the polysaccharide capsule protein conjugates found in approved vaccines for *S. pneumoniae*, *Haemophilus influenzae* type B, and *Neisseria meningitidis*. To boost immunogenicity of the native GAC, several studies have employed conjugation methods that generated protective antibody responses in animals (summarized in Table 1). Classic protein carriers employed include modified bacterial toxins such as tetanus toxoid [147, 148] or CRM_{197}, a non-toxic mutant diphtheria toxin [84]. Previous murine challenge studies using two different M serotype GAS strains showed that immunization with tetanus toxoid-GAC conjugates protected against intranasal colonization and intraperitoneal lethal infection [147]. Subsequently, synthetic GAC molecules of varying lengths were conjugated to CRM_{197} and used to determine size-dependent immunogenicity—a synthetic dodecasaccharide (equivalent to four trisaccharide repeating units of native GAC) was sufficient to generate antibody titers equivalent to purified native GAC and elicited a protective response in infectious challenge [84]. GAC sequences of varying sizes were also conjugated to GAS surface protein C5a peptidase and induced IgG1 antibodies that promoted opsonophagocytic killing; immunized mice showed reduced lung injury and mortality following GAS challenge in a pneumonia model [181].

The discovery of the GAS *gac* operon allowed targeted mutagenesis of the bacterium to eliminate the GlcNAc sidechain on GAC and subsequent purification of the polyrhamnose backbone (GAC^{PR}) from the resulting mutant [174]. GAC^{PR} has been explored as a potential universal GAS vaccine antigen devoid of the autoimmunity concerns ascribed to the GlcNAc moiety. Using an immunogenic pneumococcal protein as a carrier for wild-type GAC and

Table 1.1: **GAS immunization strategies utilizing GAC.**

Vaccine	GAC	Formulation	References(s)
GAS CHO-TT	native wild-type	Conjugated to tetanus toxoid (TT)	Sabharwal *et al.* 2006 [147, 148]
CRM_{197} conjugates	synthetic wild-type of varying isomers (6mer and 12mer)	Conjugated to CRM_{197}	Kabanova *et al.* 2010 [84]
GAC-SP0435 conjugate	native wild-type and mutant	Conjugated to recombinant pneumococcal protein SP_0435 (elongation factor)	van Sorge *et al.* 2010 [174]
ΔGAC-ADI	native mutant	Conjugated GAS arginine deiminase (ADI)	Rivera-Hernandez *et al.* 2016 [143]
C5a peptidase-GAC	synthetic wild-type in various oligomers (3mer, 6mer, 9mer)	Conjugated to C5a peptidase	Zhao *et al.* 2019 [188] Wang *et al.* 2020 [181]
Random GAC glycoconjugates	native wild-type	Random conjugation to CRM_{197}, streptolysin O, SpyCEP, or SpyAD	Di Benedetto *et al.* 2020 [39]
Gold nanoparticle	synthetic poly-rhamnose	Conjugation to gold nanoparticles	Pitirollo *et al.* 2020 [135]
SpyAD-GAC^{PR}	native mutant	Site-directed conjugation to GAS SpyAD	Gao *et al.* 2021 [58]
Trirhamnosyl-lipopeptide	synthetic tri-rhamnose	Conjugation to self-adjuvanting Ac-PADRE - lipid core	Khatun *et al.* 2021 [88]

GAC^{PR}, it was found upon conjugation that GAC^{PR} was sufficient to generate anti-GAC antibody responses comparable to native GAC [174]; the GAC^{PR} conjugated onto arginine deiminase (ADI), a streptococcal protein and virulence factor, also induced strong anti-GAC antibody titers in BALB/c mice and afforded partial protection mouse bacteremia and skin infections models [143]. For enhanced immunogenicity and broad coverage of GAS strains, our own recent study examined the use of a multivalent vaccine formulation composed of C5a peptidase, SLO, and GAC^{PR} conjugated to GAS surface protein SpyAD. This use of multiple, highly conserved GAS proteins in addition to the GAC^{PR}-induced broad immunity to multiple GAS strains of different serotypes as demonstrated by opsonophagocytic killing assays and two different murine models of GAS infection without evidence of cross-reactivity to human heart or brain tissue lysates [58]. Finally, a new study successfully mounted an anti-GAC immune response by conjugating synthetic rhamnose-GlcNAc trisaccharide conjugated to lipopeptide engineered to self-assemble into immunogenic, self-adjuvanting lipid-core complexes [88]. Opsonic activity resulting from immunized mice showed 75-97% protection against four different clinically relevant strains of GAS, and showed that the GlcNAc was not required for the vaccine antigen to induce protective immunity [88]. Collectively, these studies suggest that the GAC (wild-type or modified to lack GlcNAc) may potentially be sufficient as an universal vaccine antigen for protective immunity against GAS, but requires boosted immunogenicity achieved by a variety of different adjuvants or conjugation strategies.

1.10 Conclusions

The Lancefield GAC antigen has played a prominent historical role in medical microbiology and a crucial clinical role in the laboratory diagnosis of GAS. For nearly a century, the abundance, conservation and universal presence of GAC in human GAS isolates implied an essential function in cell wall structure and bacterial survival. However, it has only been with the

recent discovery of the biosynthetic *gac* operon that mechanistic research into GAC assembly and other potential functional roles in host immune evasion and resistance to bactericidal proteins or metals became possible. Development of natural immunity against GAC spurred its potential as a universal GAS vaccine antigen, but this has met with caution due to correlations between its GlcNAc sidechain and important autoimmune sequelae of RHD or Sydenham's chorea. Identification and functional understanding of the genetically encoded biosynthetic process and chemical approaches to GAC component synthesis have enabled immunogenicity studies of modified GAC (e.g. GAC^{PR}) that might allay these safety concerns. Should further elucidation of the immunobiology of the eponymous species-defining GAC contribute to GAS disease prevention as well as diagnostics, it would indeed be a fitting tribute to the outsized legacy of Rebecca Lancefield.

1.11 Acknowledgements

Chapter 1, in full, is submitted for publication of material as it may appear in *Infection and Immunity*. "**Nina J. Gao**, Ervin Rodas Lima, and Victor Nizet. *Immunobiology of the Classical Lancefield Group A Streptococcal Carbohydrate Antigen.*" This chapter is dedicated to the memory of co-author Ervin Rodas-Lima (1999-2021). The authors' work on GAS pathogenesis and immunity has been supported by NIH/NIAID grant and the CARB-X accelerator. Figures created on Biorender.com. The dissertation author was a primary investigator and author of this work.

Chapter 2

Envisioning a Safe, Universal, Multi-Valent Group A Streptococcal Vaccine

2.1 Abstract

Development of an effective vaccine against the leading human bacterial pathogen group A Streptococcus (GAS) is a public health priority. The species defining group A cell wall carbohydrate (GAC, Lancefield antigen) can be engineered to remove its immunodominant N-acetylglucosamine (GlcNAc) sidechain, implicated in provoking autoimmune cross-reactivity in rheumatic heart disease, leaving its polyrhamnose core (GAC^{PR}). Here we generate a novel protein conjugate of the GAC^{PR} and test the utility of this conjugate antigen in active immunization. Instead of conjugation to a standard carrier protein, we selected SpyAD, a highly conserved GAS surface protein containing both B-cell and T-cell epitopes relevant to the bacterium that itself shows promise as a vaccine antigen. SpyAD was synthesized using the Xpress CF™ cell-free protein expression system, incorporating a non-natural amino acid to which GAC^{PR} was conjugated by site-specific click chemistry to yield high molecular mass SpyAD-GAC^{PR} conjugates and avoid disruption of important T-cell and B-cell immunological epitopes. The conjugated SpyAD-GAC^{PR} elicited antibodies that bound the surface of multiple GAS strains of diverse M types and promoted opsonophagocytic killing by human neutrophils. Active immunization of mice with a multivalent vaccine consisting of SpyAD-GAC^{PR}, together with candidate vaccine antigens streptolysin O and C5a peptidase, protected against GAS challenge in a systemic infection model and localized skin infection model, without evidence of cross reactivity to human heart or brain tissue epitopes. This general approach may allow GAC to be safely and effectively included in future GAS subunit vaccine formulations with the goal of broad protection without autoreactivity.

2.2 Preface to Chapter 2

Previously in chapter 1, I summarized the group A carbohydrate and its potential as an universal vaccine antigen, though with caveats with potential safety concerns over the GlcNAc motif inducing auto-reactive immune response (and potentially fatal heart disease). This chapter demonstrates the advantage of generating a glyco-conjugate for vaccination potential by combining the modified group A carbohydrate (GAC^{PR}) with an immunogenic protein carrier SpyAD to induce protective immune responses to the native GAC. Furthermore, as a part of a multi-valent vaccine formulation, this SpyAD-GAC^{PR} is capable of inducing a protective immune response in both human *ex vivo* models and murine infectious challenge models.

This study is a first definitive step to make the universal group A streptococcal vaccine a reality, while taking vaccine safety (in terms of removing the auto-reactivity potential by removing the GAC GlcNAc) into account at the development stage.

2.3 Introduction

Group A *Streptococcus* (GAS, *S. pyogenes*) is one of the most important human bacterial pathogens, estimated to cause more than 600 million cases of pharyngitis ('strep throat') and 100 million cases of impetigo annually across the globe.[18, 161] In particular, pharyngitis is highly prevalent in school-aged children and a major source of antibiotic prescriptions worldwide, driving selective pressure for antibiotic resistance throughout the human microflora[54, 60, 178]. In recent decades, GAS has been increasingly associated with severe invasive forms of infection, sometimes in previously healthy individuals, including sepsis, necrotizing fasciitis, and toxic shock syndrome. Finally, GAS is the trigger for important post-infectious immune-mediated diseases, in particular post-streptococcal glomerulonephritis [176], acute rheumatic fever (ARF), and rheumatic heart disease (RHD). Deaths associated with RHD ranks GAS as a leading cause of cardiovascular mortality in the developing world[134]. At least 33 million people are currently affected by RHD, with approximately 275,000 deaths annually (60% age $<$70) and 9 million disability-adjusted life years lost[19]. Almost all RHD deaths occur in low-income and middle-income countries.[18]

As a result of its disease manifestations, GAS ranks among the top 10 infectious causes of human mortality[138]. Despite the high disease burden and global demand, there is to date no safe and efficacious commercial vaccine against GAS. A number of phenotypic features of the pathogen pose particular challenges to vaccine development. First, in contrast to the diverse capsular polysaccharides of *Streptococcus pneumoniae* that are the basis of multivalent glycoconjugate vaccines in clinical use worldwide[59, 113], GAS possesses an invariant capsule of high molecular weight hyaluronic acid (HA). The capsule is thus structurally identical to the ubiquitous HA present in human connective tissues, and therefore immunologically inert[87, 164, 170, 184]. Instead, GAS serology is classically assigned by its most abundant and highly immunogenic surface-anchored protein, the M protein, which poses a second vaccine challenge—M proteins are

highly polymorphic. There are over 220 described sequence variants of the encoding *emm* gene, classified into 48 *emm*-clusters of closely related M proteins sharing functional and structural properties[153]. Heterologous M proteins do not typically elicit cross-protective immunity to M-types from other clusters[53]. Finally, an unique challenge to GAS vaccine development is its link to the serious post-infectious, immune-mediated ARF and RHD, whose precise molecular pathogenesis remains poorly understood[12, 19, 66]. A prevailing theory holds that GAS molecular mimicry of host cell epitopes mediates B and/or T cell cross-reactions with human tissue antigens in the heart, brain, or other tissues[67].

Beginning in the late 1970's, the U.S. Food and Drug Administration (FDA) made the rare decision to suspend GAS vaccine development (21 Code of Federal Regulations 610.19) for nearly three decades. This decision followed an immunization study in which three participants developed suspected or definite ARF[115]. In 2004, the US National Institutes of Allergy and Infectious Diseases (NIAID) convened experts to review epidemiologic, microbiological and immunologic factors involved in preclinical and clinical development of a safe and effective GAS vaccine that facilitated lifting of the FDA ban[11]. Their summary report concluded that molecular mimicry represented a major obstacle to vaccine development, and that GAS antigens including M proteins and group A carbohydrate possess epitopes linked to B and/or T cell reactivity to human tissue antigens. The panel recommended: "Because the precise role of molecular mimicry in the pathogenesis of ARF has not been established, every effort should be made to exclude tissue cross-reactive epitopes during vaccine development."[11] Recently, the World Health Organization, International Vaccine Institute, Wellcome Trust, Bill and Melinda Gates Foundation and other stakeholders convened in advisories on GAS vaccine development to address scientific challenges for this paramount global health need[131, 178].

Ideal candidate antigens for inclusion in a GAS vaccine would be (a) highly immunogenic and elicit antibodies that promote opsonophagocytosis or inhibit virulence, (b) exhibit broad conservation across strains contributing to global disease epidemiology, and (c) be chosen to

avoid autoimmune cross-reactivity with human tissue epitopes[38, 72]. All GAS strains possess the species-defining Lancefield group A carbohydrate (GAC), composed of a polyrhamnose backbone with an immunodominant N-acetylglucosamine (GlcNAc) sidechain[116, 117], of which ˜25% is decorated with glycerolphosphate[46]. Representing 40-50% of the GAS cell wall by weight[116], GAC serves as the basis of current rapid antigen testing for GAS infection[6]. Earlier mouse immunization studies with protein-conjugated native or synthetic GAC vaccines show clear efficacy against multiple GAS M types[84, 147]. Serum anti-GAC antibodies are likewise present in healthy individuals and peak around 17 years of age, strongly correlating with decreased GAS infection risk[189].

For the above reasons of immunogenicity and conservation, GAC has garnered considerable interest as a universal GAS vaccine antigen. However, this has also elicited concern, since experimental evidence for autoreactivity of antibodies that recognize the native GAC GlcNAc sidechain against human tissues has been communicated by different research groups. For example, glycoproteins from human heart valves elicit antibodies that bind GAC in a manner blocked by GlcNAc (but not rhamnose or other glycans)[62], and persistence of anti-GlcNAc/GAC antibodies (up to 20 years) are a marker of poor prognosis of RHD valvular disease, whereas antibodies against streptolysin O (SLO) and the polyrhamnose core of GAC fade independently of valvular complications[43]. Also, anti-GAC GlcNAc antibodies that cross-react with heart or brain tissue are present in sera of ARF patients with cardiac or neurological complications[56, 89].

An approach to modify the GAC to eliminate the potential cross-reactive GlcNAc epitope was achieved by Van Sorge *et al.* upon discovery of the 12-gene GAS *gac* gene cluster encoding the biosynthetic machinery for GAC production[174]. This work generated an isogenic mutant (Δ*gacI*) that expressed only the polyrhamnose backbone of GAC without the GlcNAc sidechain[145, 174]. This Δ*gacI* mutant was attenuated for virulence in mouse and rabbit infection models, and showed increased sensitivity to killing by human whole blood, neutrophils, and platelet-derived antimicrobials in serum[174]. A biotin conjugate of the modified GAC

structure containing only the non-mammalian carbohydrate rhamnose (GAC^{PR}), generated high antibody titers in rabbits that promoted opsonophagocytic killing of GAS strains of multiple M types by human neutrophils and protected against systemic GAS challenge in mice upon passive immunization[174]. No cross-reactivity of anti-GAC^{PR} antisera was observed against human heart tissue lysates[174].

In the present work, we sought to generate a protein conjugate of the GAC^{PR} for use in safe, universal subunit vaccines against this important pathogen, and to test the utility of this conjugate antigen in active immunization. As a novel approach, instead of conjugation to a standard carrier protein to engender T-cell-mediated immunity, such as tetanus toxoid or CRM_{197}, we selected SpyAD, a highly conserved GAS surface protein contains both B-cell and T-cell epitopes relevant to the bacterium and itself shows promise as a vaccine antigen[10, 52, 55, 141]. The conjugates SpyAD-GAC^{PR} elicited antibodies that bound the surface of multiple GAS strains of diverse M types and promoted opsonophagocytic killing by human neutrophils. Active immunization of mice with a multivalent vaccine consisting of SpyAD-GAC^{PR} in combination with candidate vaccine antigens streptolysin O (SLO) and C5a peptidase provided significant protection against GAS challenge in a systemic infection model and localized skin infection model, without evidence of cross reactivity to human heart epitopes.

2.4 Results

2.4.1 Expression and purification of GAS protein antigens

GAS expresses several secreted and membrane-anchored virulence factors that are important for disease pathogenesis *in vivo*[27, 180]. For the present vaccine study, we selected three protein antigens that are strongly conserved with high genomic carriage rate across a globally representative and clinically diverse collection of 2,083 GAS genomes, coupled to low amino acid sequence variation[38] and evidence of natural immunogenicity—streptolysin O (SLO), C5a

peptidase, and SpyAD. SLO is a secreted, pore-forming, cholesterol-dependent cytolysin; not only is SLO a virulence factor unique to GAS, but anti-SLO serum titer is indicative of recent GAS infection [158]. Immunization of mice with inactivated SLO toxoid induces SLO-neutralizing antibodies and protects against experimental infection with the pathogen[23, 167]. C5a peptidase is a surface-expressed GAS peptidase that cleaves C5a complement to inactivate the chemoattractant, delaying recruitment of phagocytes to the site of GAS infection.C5a peptidase alters clearance and trafficking of group A streptococci by infected mice[81]. C5a peptidase-specific antibody responses strongly correlate with anti-SLO titers in children with pharyngitis[25, 158], and the protein also elicited protective vaccine responses in murine models[25, 80]. Finally, "*S. pyogenes* adhesion and division protein" (SpyAD), which plays dual roles in GAS adherence to host cells and bacterial cell division[55], was identified as a potential GAS vaccine candidate in several throughput screens of the GAS surface proteome[9, 52, 144] and was part of a 7-valent GAS vaccine that proved successful in murine immunization studies[141].

To express the above proteins for use as vaccine antigens, we applied Vaxcyte's proprietary XpressCF+™ cell-free protein synthesis (CFPS) expression platform, which is based on extraction of the *E. coli* cellular machinery required for transcription, translation, and energy production into a cell-free mixture capable of continuous oxidative phosphorylation. With this CFPS platform, we successfully expressed the immunogenic cores of SLO, C5a peptidase, and SpyAD. For native FL-SLO, we expressed both a N-terminal truncation fragment spanning aa 79-571 (designated SLO) as well as a N- and C-terminal truncation fragment comprising aa 79-470 (designated SLO(ΔC101)) (Figure 2.1A). Truncated SLO were designed to optimize product yield in our CFPS (further improved with truncation of 101 amino acids off the C-terminus) while preserving key immunogenic epitopes and detoxifying the cytolysin by removing the tryptophan-rich loop that mediates insertion into cellular membranes. Additionally, we expressed C5a peptidase and SpyAD with precise mutations for enzymatic inactivation (Figure 2.1A). Next, using amber stop codon suppression and the addition of orthogonal tRNA and tRNA synthetase pair during

CFPS, we successfully expressed a variant of SpyAD that contains p-azidomethyl phenylalanine (pAMF) non-native amino acids (nnAA) to replace 4 native lysines (Figure 2.1A, marked by black triangles). The quantitative expression yield of each antigen was $>$ 200 mg/L as estimated by incorporation of 14C-leucine into the translating polypeptide (Fig. 2.1B). SDS-PAGE analysis of synthesized proteins confirmed high purity fractions of each of the expected peptide antigens ($>$95% as shown in Figure 2.1C and D). Incorporation of pAMF into SpyAD was confirmed by conjugation to DBCO-TAMRA dye and fluorescence readout. Unlike native SpyAD, the variant with 4 pAMF sites was successfully labeled with the dye to confirm pAMF incorporation (Figure 2.1D).

2.4.2 Generation of SpyAD-GAC^{PR} conjugate

Using the in-house isolation and purification protocol, highly pure preparations of mutant GAS cell wall carbohydrate containing only polyrhamnose (GAC^{PR}) were generated from a GAS mutant strain that lacked the GlcNac sidechain of GAC and the surface hyaluronan capsule (Δ*gacI*Δ*hasA*). Using 1-Cyano-4-dimethyl aminopyridinium tetrafluoroborate (CDAP) chemistry, GAC^{PR} was dibenzocyclooctyne (DBCO)-derivatized for use in a conjugation reaction with SpyAD[4pAMF], shown schematically in Figure 2.2A. Analysis of SEC-MALS data estimated an average molar mass of 7.2 kDa for GAC^{PR} and 87.3 kDa for SpyAD[4pAMF], the latter in close agreement with the theoretical molar mass of a SpyAD monomer (Figure 2.2B and C). The molar mass of the purified DBCO-derivatized GAC^{PR} was similar to the molar mass of the native polysaccharide (data not shown). For conjugation, DBCO-derivatized GAC^{PR} was incubated with SpyAD[4pAMF] at a 1:1 ratio to facilitate strain-promoted Cu^{2+}-free click chemistry reaction to generate conjugates. After the conjugation reaction and dialysis to remove excess GAC^{PR}, the resulting conjugates were analyzed using SEC-MALS. As shown in Figure 2.2D, the conjugates elute as two separate polydisperse distributions (Peak1 and Peak2) with average molar masses ranging from 2 MDa to 135 kDa, respectively, yielding a combined average molar mass of 153.4

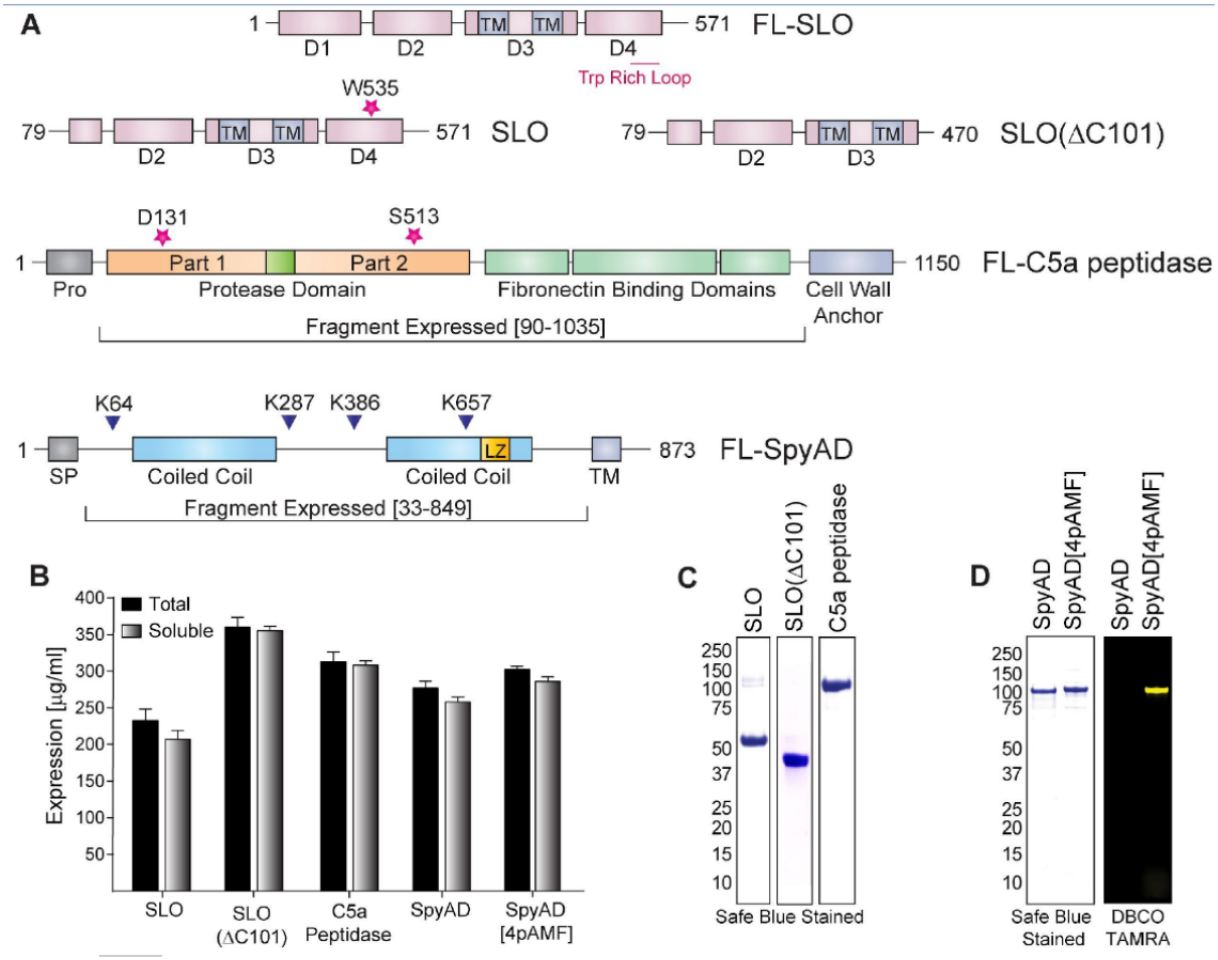

Figure 2.1: **Expression, purification and analysis of GAS protein antigens.** **(A)** Modular architecture of GAS protein antigens is depicted schematically while highlighting amino acid changes (red stars) / truncations in SLO or C5a peptidase and pAMF incorporation sites on SpyAD (black triangles) along with the immunogenic core of the proteins expressed in cell free protein synthesis (CFPS). Domains: D# = protein domain, TM = transmembrane, Pro = pro-sequence, SP = signal peptide. **(B)** Protein expression yield in CFPS estimated using incorporation of 14C-leucine into the translating polypeptide. Safe blue stained SDS-PAGE of SLO and C5a peptidase **(C)** and SpyAD **(D, left)** confirm purity of expressed proteins. pAMF incorporation into SpyAD[4pAMF] was confirmed through selective labeling with DBCO-TAMRA fluorescent dye **(D, right)**.

kDa for the conjugates, which is significantly higher than the average molar mass of either GAC^{PR} (˜7.2 kDa) or SpyAD[4pAMF] (87.3 kDa) alone, thereby confirming successful conjugation into the final product, referred to as SpyAD-GAC^{PR} from here onwards.

2.4.3 Evaluation of GAS vaccine antigens for immunogenicity

Immunization of New Zealand White rabbits was performed to generate antisera for functional evaluation of the candidate GAS vaccine antigens. Rabbits were immunized with 5 µg of individual protein antigen (SLO, C5a peptidase, SpyAD[4pAMF]) or SpyAD-GAC^{PR} conjugate (equivalent to 5 µg of polysaccharide). To determine immunoglobulin G (IgG) titers elicited by vaccination, terminal bleed (day 35) rabbit antisera were evaluated by ELISA using plates coated with synthesized proteins. For all three protein antigens (SpyAD, C5a peptidase, and SLO), the group of rabbits immunized with the recombinant protein showed significantly increased (3- to 4-$\log_{10}$ fold) antibody titers against the target antigen compared to either pooled serum from the rabbits before immunization ("pre-immune (pooled)") or the other immunized rabbit groups (Figure 2.3A), confirming specificity of the IgG response to each respective antigen. The anti-SpyAD protein titer generated by the SpyAD-GAC^{PR} conjugate was not inferior to that of the SpyAD recombinant protein alone, suggesting that key B cell epitopes were not impacted by pAMF sites. Flow cytometry was used to evaluate the binding affinity of rabbit-derived IgG to the surface of eight live wild-type GAS strains of different M protein serotypes (M1-6, M12, M28). For the great preponderance of strains, the respective immunized serum yielded an increase in binding of 100 - 400% over the baseline IgG binding of the pre-immune serum (Figure 2.3B). For six of the eight strains, the GAS surface binding of the antisera raised against the SpyAD-GAC^{PR} conjugate roughly doubled the level of IgG binding seen with antisera raised against SpyAD alone. To examine the post-immunization antibody response to native GAC antigen, we generated a genetic knockout of SpyAD in M1 5448 strain (ΔSpyAD) and tested rabbit serum IgG binding to the cell surface of the live bacterium (Figure 2.3C). Fluorescent signals detected similar IgG

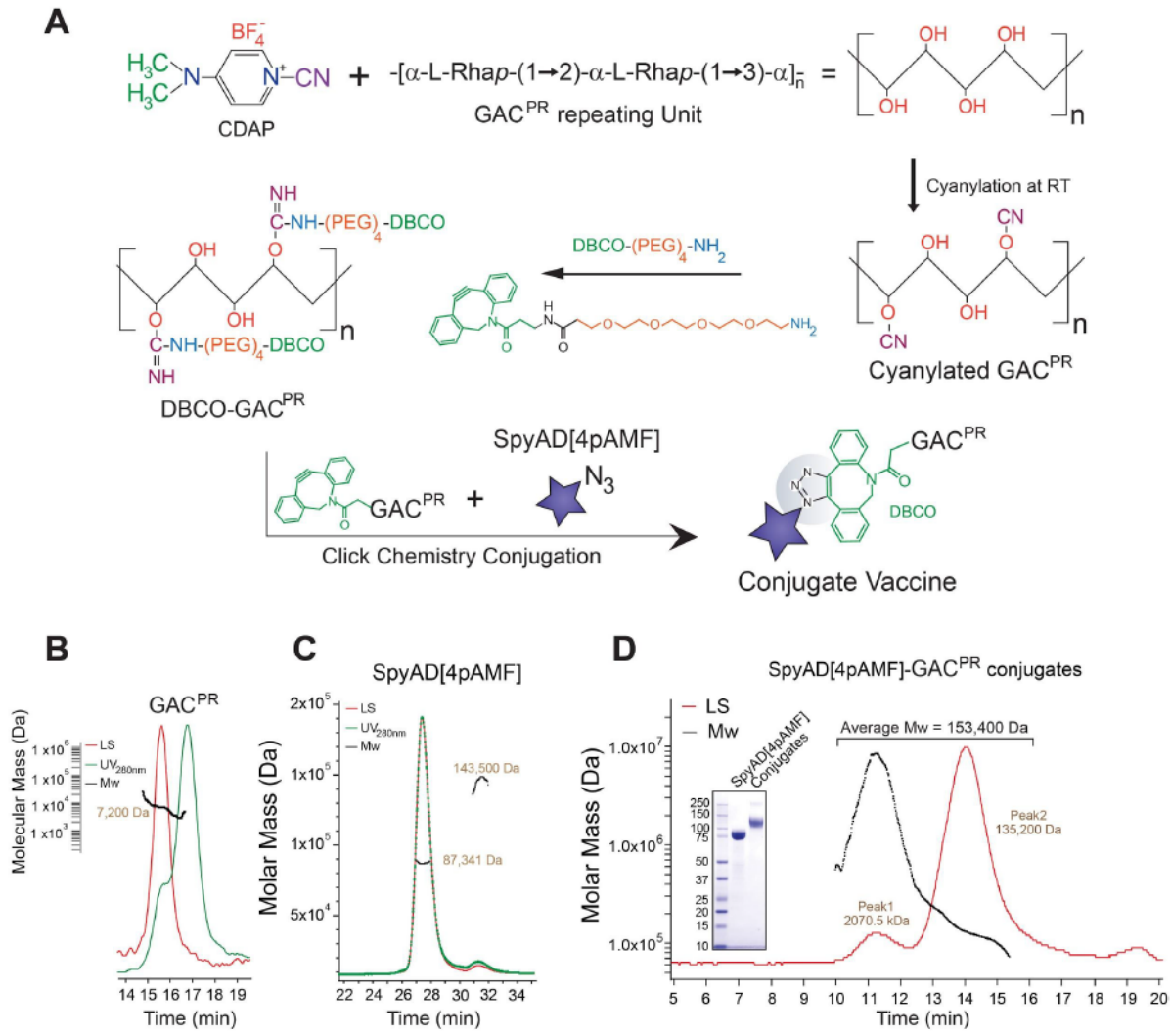

Figure 2.2: **GAC^{PR} derivatization schematic and generation of conjugate vaccine antigen.** **(A)** Schematic outlines chemical reactions utilizing CDAP chemistry for DBCO derivatization of GAC^{PR} and theoretical depiction of Cu^{2+}-free click chemistry reaction for generating conjugate vaccine. **(B)** SEC-MALS analysis of purified native GAC^{PR} estimates an average molecular mass of 7.2 kDa. **(C)** SEC-MALS analysis of purified native SpyAD[4pAMF] estimates an average molecular mass of 87.3 kDa. **(D)** Post-click chemistry SEC-MALS analysis of purified conjugates estimates an average combined molecular mass of 153.4 kDa. Inset shows the Safe Blue stained SDS-PAGE analysis of SpyAD[4pAMF] pre- (left) and post-conjugation (right).

binding of SpyAD or SpyAD-GAC^{PR} antisera to the surface of wild-type M1 GAS. As expected, the IgG fluorescent signal was for SpyAD antiserum against the ΔSpyAD, but a clear signal was still present for the SpyAD-GAC^{PR} antiserum, demonstrating that it contains native GAC-binding IgG.

2.4.4 Evaluation of rabbit antisera in opsonophagocytic killing, blocking SLO activity, and passive protection in murine challenge model

The ability of the rabbit antisera raised against the GAS vaccine antigens (SLO, C5a peptidase, SpyAD and SpyAD-GAC^{PR}) to promote human neutrophil opsonophagocytic killing (OPK) was evaluated using GAS strains of five different M protein serotypes (M1-M5). This assay was performed with 30 minutes pre-opsonization with the respective heat-inactivated antisera or pre-immune sera control, then 30 min exposure to freshly isolated human neutrophils at multiplicity of infection (MOI) of 0.1 bacteria per neutrophil in the presence of 2% baby rabbit complement. To test potential protection from a combination of antigens, we tested an antiserum mixture composed of one-third of each antiserum raised against SLO, C5a peptidase, and SpyAD-GAC^{PR} to maintain a consistent total serum concentration (designated as "Combo" in Figure 2.4A). As a positive control for OPK, we used anti-M1 rabbit serum against M1 GAS and recovered a low percentage of recovered CFU due to effective killing (Figure 2.4A, gray bar of first graph). All antisera induced statistically significant increases by one-way ANOVA compared to pre-immune sera in short-term OPK of GAS (Figure 2.4A). Of note, our assay uses a 100 to 1,000-fold greater GAS inoculum than other published OPK assays employing the human promyelocytic leukemia HL-60 cell line at MOI of 0.001[83] or 0.0001[150]; a higher bacterial inoculum in our assay more accurately recapitulates the immunization target of supporting primary human innate immune cell function. In our *ex vivo* human neutrophil OPK assay, antisera raised against the SpyAD protein and SpyAD-GAC^{PR} performed similarly. Moreover, the antisera combination "Combo" performed similarly or better than the individual

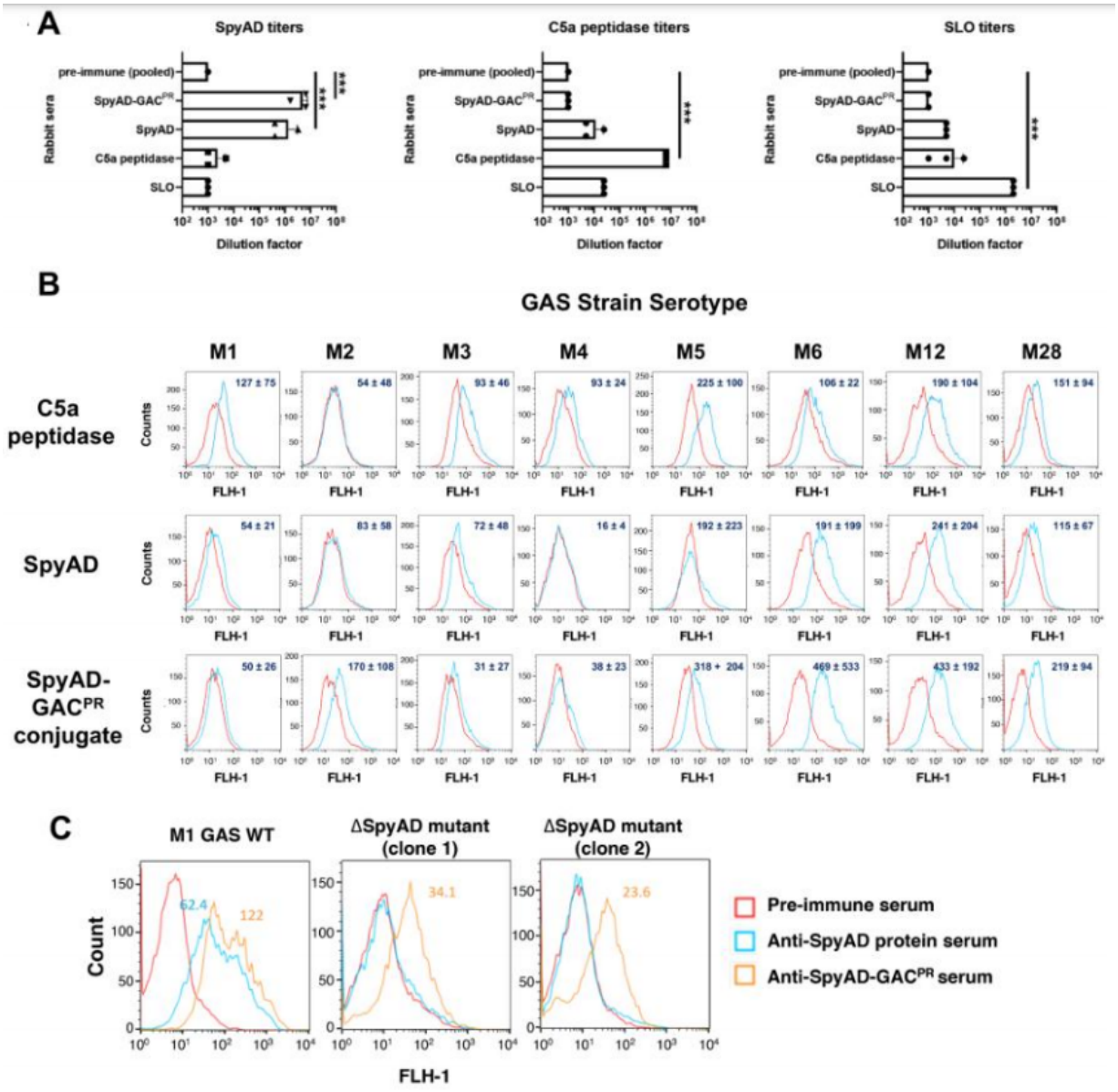

Figure 2.3: **Immunogenicity of GAS vaccine antigens in rabbits.** **(A)** ELISA was performed on rabbit serum to quantify IgG titers raised against each protein antigen. Each point represents serum derived from one animal, with 2-3 rabbits per group. **(B)** Eight GAS strains of different M-types were used to determine rabbit serum IgG binding to native antigens by flow cytometry. Histograms show representative fluorescent signals from IgG binding; red represents pre-immune sera and blue immune sera. Numbers in dark blue the top right corner of each histogram shows the mean percentage of increased signal of immunized rabbit serum over pre-immunized serum signal. **(C)** Rabbit IgG binding to ΔSpyAD mutant GAS confirms the presence of IgG recognizing native GAC in the SpyAD-GAC^{PR} antiserum.

components, indicating no cross-interference in their OPK functions. As SLO is a secreted cytotoxin not anchored to the GAS surface, its contribution to enhanced neutrophil killing is accrued not from increased phagocytic uptake, but rather reducing SLO-mediated membrane damage and impairment of neutrophil antimicrobial functions.[167] We found that the anti-SLO rabbit immune serum, whether elicited by the SLO or SLO(ΔC101) antigens, significantly inhibited hemolytic activity of purified SLO against human red blood cells up to a dilution of 1:2,048 (Figure 2.4B). Furthermore, anti-SLO or anti-SLO(ΔC101) immune serum equally preserved neutrophil oxidative burst function (superoxide generation) against GAS supernatant (SLO)-mediated suppression (Figure 2.4C).

The potential of vaccination using SpyAD-GAC^{PR} alone or in combination with other GAS protein antigens, namely SLO or SLO(ΔC101) plus C5a peptidase, to provide protection against systemic GAS infection was evaluated in murine models of passive immunization. As a first proof-of-principle, rabbit antisera were transferred into adult CD-1 mice intravenously by retro-orbital injection before challenging each animal with 1 x 10^7 colony forming units (CFU) of the virulent serotype M1 GAS strain 89155 injected into the peritoneal cavity (Figure 2.4D). Control mice receiving pre-immune rabbit serum showed <10% survival within 24 hours of infection whereas mice immunized with either the SpyAD-GAC^{PR} antiserum or a combination of SLO + C5a peptidase antisera (volume divided in a 1:1 ratio) showed modest protection (20% and 26.7% survival through day 5, respectively), but most importantly the mice immunized with the multivalent SpyAD-GAC^{PR} + SLO + C5a peptidase (volume divided in a 1:1:1 ratio) had significant protection against mortality with 53.3% survival through day 5 (P = 0.0048 vs. pre-immune serum group, Figure 2.4D)

In summary, due to the variation in IgG binding to GAS strains of different M types (Fig. 2.3B), likely due to the variation in antigen surface expression levels, and the superiority of the combination antisera in OPK assays (Fig. 2.4A) and passive immunization (Fig. 2.4B), we concluded that a multivalent vaccine formulation was required to broadly cover GAS strains and

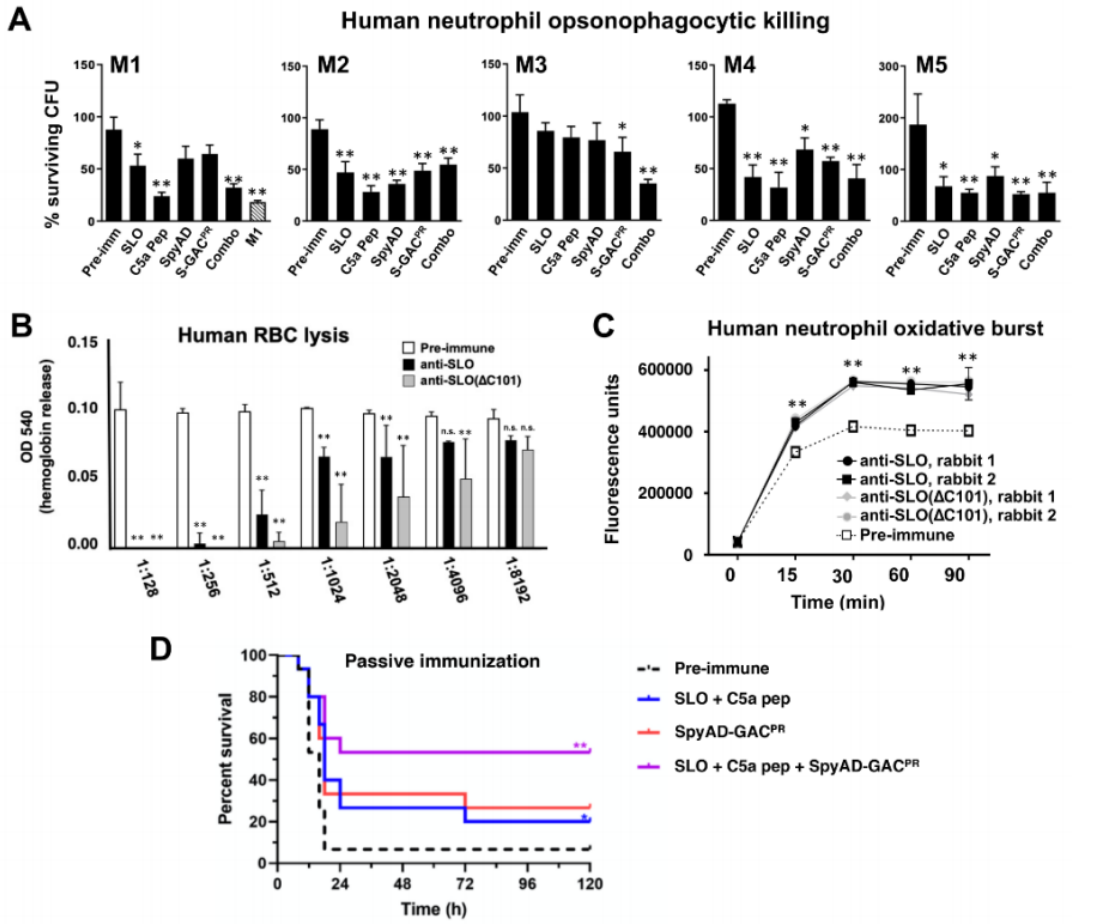

Figure 2.4: **Evaluation of rabbit vaccine antisera in GAS opsonophagocytic killing, blocking SLO activity, and passive protection in murine challenge model. (A)** Opsonophagocytic killing of GAS of different M serotypes by human neutrophils in the presence of immune serum from immunized rabbits compared to serum from pre-immunized rabbit serum. **(B)** Anti-SLO or anti-SLO(ΔC101) antisera block lysis of human red blood cells by purified SLO. **(C)** Anti-SLO or anti-SLO(ΔC101) antisera enhance oxidative burst capacity of human neutrophils exposed to GAS supernatants containing SLO as quantified by 2,7-dichlorofluoroscein diacetate fluorescence. Statistical analyses were performed by one-way ANOVA compared to pre-immune serum; $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***). **(D)** CD-1 female mice receive an intravenous dose of rabbit antisera followed by intraperitoneal lethal challenge with $1x10^7$ M1 89155 strain GAS. Data shown are from two independent experiments with N = 15 (combined) mice per group; statistics are calculated from log-rank Mantel Cox test for Kaplan-Meier plots.

proceeded to active immunization experiments with antigen combinations.

2.4.5 Multivalent immunization in mice generates antibody response with improved capacity to bind to native surface antigens and direct opsonophagocytic killing

Mouse sera were collected from 9 immunized (and 10 mock immunized) animals on day 42 following a three dose immunization (intramuscular injections of antigens on days 0, 14 and 28) with SpyAD-GAC^{PR} + SLO(ΔC101) + C5a peptidase or mock immunization control, and the efficacy of these antisera tested for surface IgG binding to 20 GAS strains via flow cytometry. For all 20 GAS strains of differing *emm* types tested, the multivalent combination antisera from all mice bound clearly more IgG to the bacterial surface than the mock immunization animal sera (Figure 2.5A). Murine IgG binding to methicillin-resistant *Staphylococcus aureus* (MRSA) strain USA300 was comparable between antisera groups, confirming the specificity of bound murine IgG from the SpyAD-GAC^{PR} + SLO(ΔC101) + C5a peptidase for GAS.

The mock and combination immunized mouse serum was subsequently tested for promoting *ex vivo* opsonophagocytic killing of M1 89155 GAS by freshly isolated human neutrophils (Figure 2.5B). When GAS were opsonized with post-immune SpyAD-GAC^{PR} + SLO(ΔC101) + C5a peptidase mouse antisera, only ˜20% of the original inoculum of M1 GAS was recovered following neutrophil exposure, compared to ˜60% CFU recovery in control studies in which GAS were opsonized with mock immune mouse sera (Figure 2.5B, $P < 0.01$).

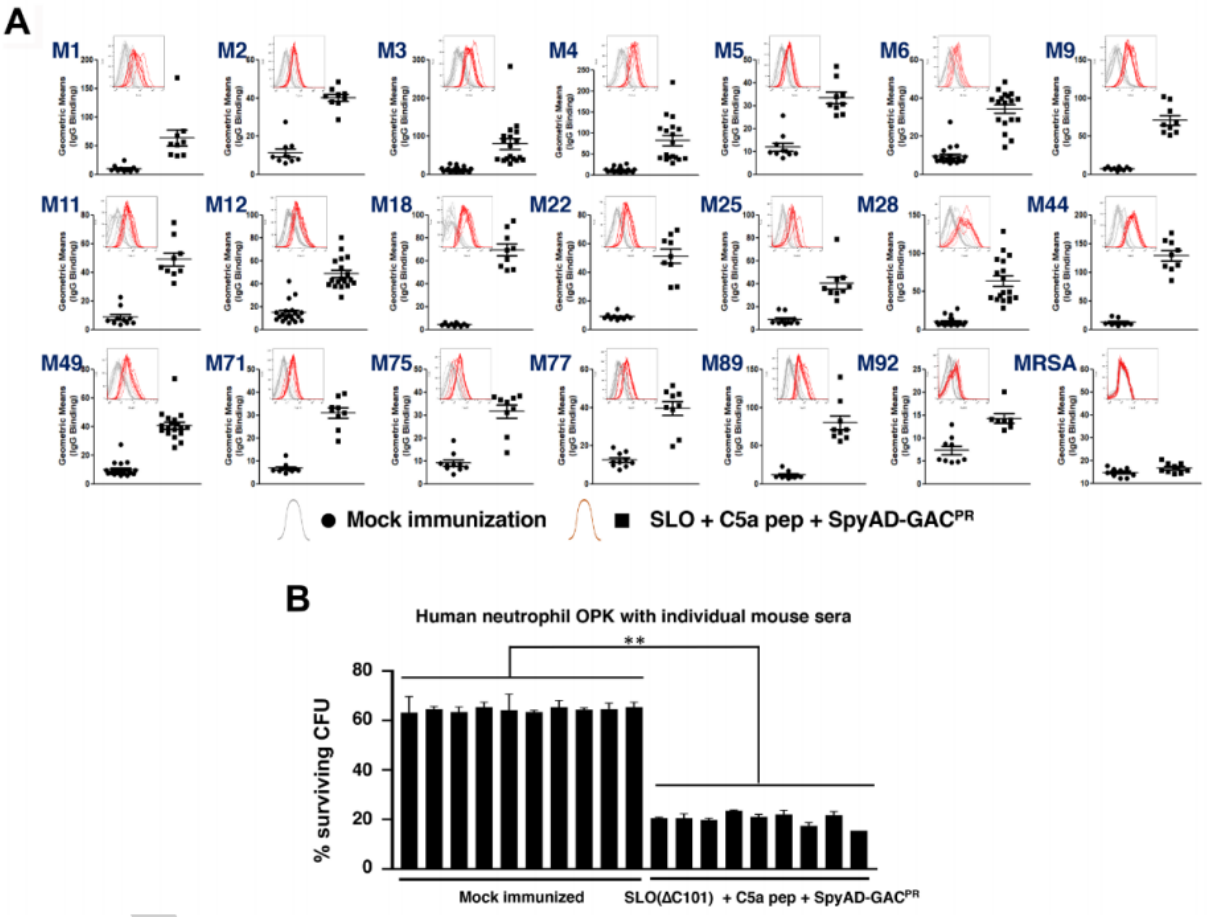

Figure 2.5: **A multivalent vaccine SpyAD-GAC^{PR} + SLO(ΔC101) + C5a peptidase vaccine elicits IgG responses in mice that recognize the surface of diverse GAS strains and promote opsonophagocytic killing of M1 GAS.** Serum was collected from female CD-1 mice immunized with 3 doses of combination SpyAD-GAC^{PR} + SLO(ΔC101) + C5a peptidase vaccine or mock immunization control. **(A)** Fluorescent intensities of murine IgG from mock antisera (gray on inset histogram) or multivalent combination antisera (red on inset histograms) bound to GAS surface antigens of multiple serotypes are quantified via flow cytometry. Mean fluorescence intensities are summarized in column scatter plots, with each point representing an individual mouse serum. Methicillin-resistant *Staphylococcus aureus* USA300 served as a control to show specificity of IgG binding to GAS. **(B)** M1 GAS bacteria was pre-opsonized with murine post-immune serum and tested in a human neutrophil opsonophagocytic killing (OPK) assay. Each bar shown is the result from an individual mouse serum, with error bars representing technical replicates. $p < 0.01$ (**) for each mouse by one-way ANOVA vs. mock immunized.

2.4.6 Immunization with a multivalent glycoconjugate vaccine provides significant protection against systemic and intradermal GAS challenge

Active immunization of mice was performed with intramuscular injections of antigens at days 0, 14 and 28 followed by intraperitoneal M1 GAS challenge on day 42 (Figure 2.6A). In this model, the SpyAD-GAC^{PR} + SLO + C5a peptidase triple combination vaccine yielded striking 100% protection against the lethal challenge (P = 0.0004), whereas SpyAD-GAC^{PR} alone (20% protection) or two protein SLO + C5a peptidase formulation (40% protection) showed modest increases in mouse survival that did not achieve statistical significance (Figure 2.6A).

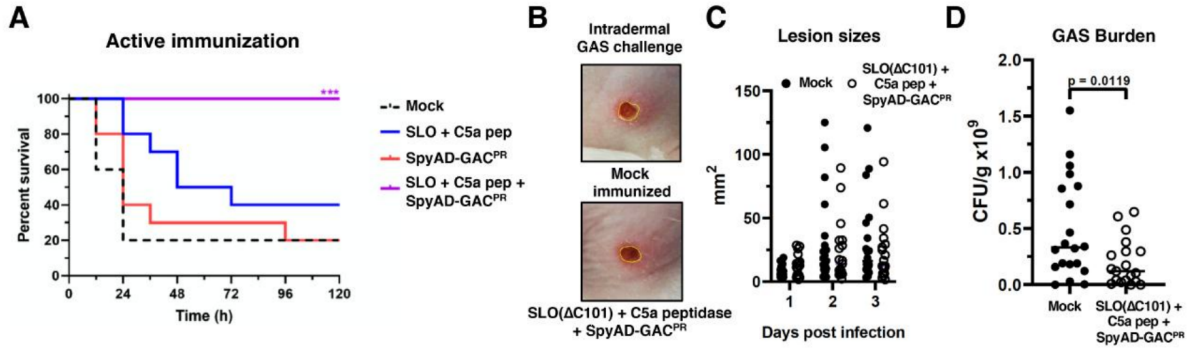

Figure 2.6: **Mice immunized by multivalent conjugate combination are protected when challenged by lethal GAS systemic infection or intradermal skin infection.** Female, wild-type CD-1 mice received three intramuscular doses of alum-adjuvanted vaccine formulations before infectious challenge with M1 GAS. **(A)** Immunized mice were subjected to lethal intraperitoneal challenge with 1x10^7 CFU M1 89155 strain GAS and tracked for survival. N = 10 mice per group for lethal challenges, with statistics calculated from log-rank Mantel Cox test for Kaplan-Meier plots. Immunized mice were intradermally challenged with 1x10^6 CFU M1 GAS and tracked for skin lesion development. Representative photos of lesions at day 3 are shown in **(B)**, with the total affected area for each mouse quantified in **(C)**. **(D)** Skin lesions were harvested by day 3 post-infection and homogenized to enumerate bacterial burden. For intradermal infections, data shown reflects two independent experiments with groups of 10; statistics are calculated with combined data by unpaired t-test with Welch's correction **(C and D)**.

Mice immunized with SpyAD-GAC^{PR} + SLO(ΔC101) + C5a peptidase or mock (PBS and adjuvant alone) control were challenged using intradermal M1 GAS infection. The size of

necrotic lesions generated by the resulting acute inflammatory response did not differ significantly between the two groups (Figure 2.6B and C), but the recovered bacterial CFU per gram was reduced more than two-fold in the debrided tissue at the site of infection in the SpyAD-GAC^{PR} + SLO(ΔC101) + C5a peptidase immunized group (Figure 2.6D, P = 0.0012).

2.4.7 Immunization with SpyAD-GAC^{PR} does not induce antibodies cross-reactive to human heart or brain tissue epitopes

GAS vaccine development programs have the unique challenge of ensuring the formulation does elicit cross-reactive immune responses capable of recognizing self-antigens in heart and brain tissue implicated in the pathogenesis of RHD. We performed western blot analysis of the SLO, C5a peptidase, and SpyAD-GAC^{PR} rabbit immune sera on normal human heart or brain lysates separated by polyacrylamide gel electrophoresis. For each antisera tested, immunizing antigen alone was run in the leftmost lane of each group to conform antibody recognition. Antisera raised against the GAS M1 protein using the same rabbit immunization protocol served as a control. As shown in Figure 2.7, while anti-M1 antisera reacted to very high molecular weight components of the lysate, no cross-reactivity to the human heart tissue was seen for the SLO, C5a peptidase and SpyAD-GAC^{PR} antisera, although they recognized the respective cognate GAS protein (or protein carrier in the case of SpyAD-GAC^{PR}). This result is consistent with (a) the lack of sequence homology of SLO, C5a peptidase or SpyAD to human proteins, and (b) knowledge that GAC^{PR} is comprised solely of repeating rhamnose, a sugar absent in humans, following the genetic deletion of its GlcNAc sidechain, which represented a common mammalian sugar epitope. We acknowledge that though our data reveals a lack of cross-reactivity of rabbit serum antibodies to human tissues, this does not exclude the possibility of cross-reactivity in humans nor does it alone guarantee vaccine safety.

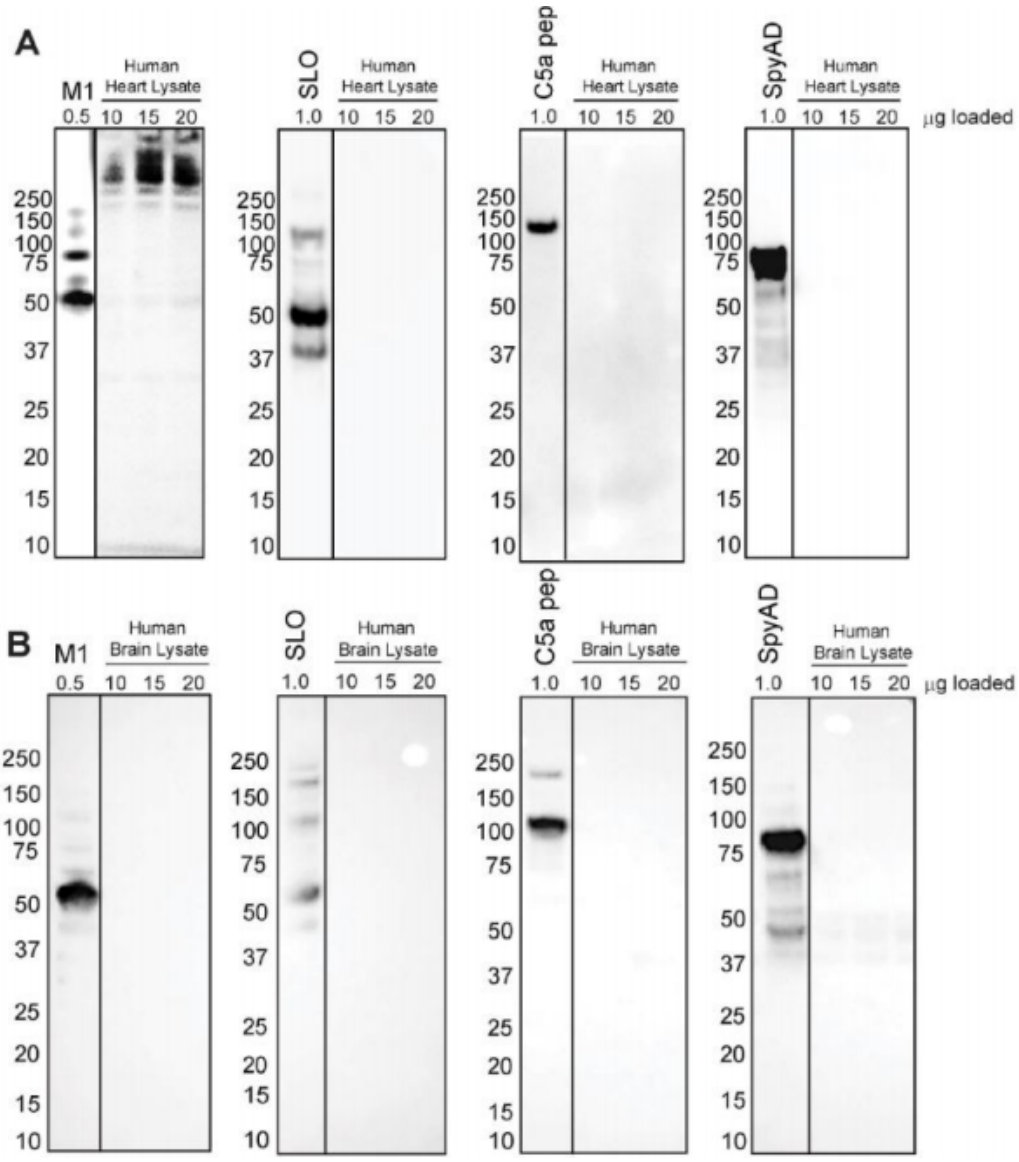

Figure 2.7: **Assessment of antigen-specific antisera for cross-reactivity to human heart lysate.** Antisera from rabbits immunized with M1 protein, full length SLO, C5a peptidase or the SpyAD-GAC^{PR} conjugates were used in western blot analysis against human heart **(A)** or brain **(B)** lysate. Unlike antisera raised against purified M1, antisera generated against each of the GAS vaccine antigens only react to the corresponding recombinant protein immunogen (lane 1) but do not possess detectable cross-reactivity to normal human organ lysate (lanes 2-4) for each antisera tested.

2.4.8 Multivalent conjugate combination vaccine yields protection against GBS challenge

With the exception of streptolysin O, which is a virulence factor unique to GAS, the protein antigens C5a peptidase [21, 24] and SpyAD used in our multivalent vaccine share some sequence conservation with *Streptococcus agalactiae*, also known as the group B *Streptococcus* (GBS). GBS can be found as a a colonizer of human gastrointestinal or genitounrinary tracts, but it is also the most common infectious cause of morbidity and mortality in neonates via bacterial transfer from the mother to drive GBS pneumonia, bacteremia, and menigitis [146, 173]. Since the 1980s, GBS has also been increasing trend with invasive disease, particularly in the elderly [51]. Though there are many GBS vaccine strategies targeting variations of the capsular polysaccharides, GBS capsule suffers the same flaws as GAS M protein: though these antigens define serology and are known to be immunogenic, vaccines would need to cover all variants to avoid serotype global distribution and prevalence which would render the vaccine formulation useless [101].

In GBS, C5a peptidase (encoded by *scpB*) is a surface peptidase highly conserved with GAS C5a peptidase, with previously published cross-protection between the two homologs [24, 25]. Current literature on SpyAD and its homologues are scarce; however, SpyAD (also known by its former locus tag *spy0269* in GAS) shares up to 81% amino acid homology of predicted protein with GBS "SEC10/PgrA surface exclusion-containing protein", primarily at the surface exclusion PgrA domain (NCBI reference sequence WP_100244355.1). There are no publications focused on the biology or structure of GBS SEC10 protein at this time. Stretches of polyrhamnose exist within the complex, highly branched group B carbohydrate (GBC) of GBS [121], allowing for potential cross-protection yielded by immunization with GAC^{PR}.

To test the potential protection from the triple combination against GBS infection, we challenged immunized female CD-1 with two different GBS strains: (1) COH1 strain representa-

tive of serotype III GBS and (2) A909 strain representative of serotype Ia (Figure 2.8). Serotypes Ia-V are the most commonly reported for GBS strains isolated from asymptomatic colonization and disease manifestations for neonates and adults, with serotype Ia and III accounting for 12.6% and 29.4%, respectively, of GBS colonized in healthy, non-pregnant people [173]. For both GBS lethal intraperitoneal infections, mice immunized with the multi-valent combination of GAS vaccine antigens were protected from mortality compared to the mock immunized groups.

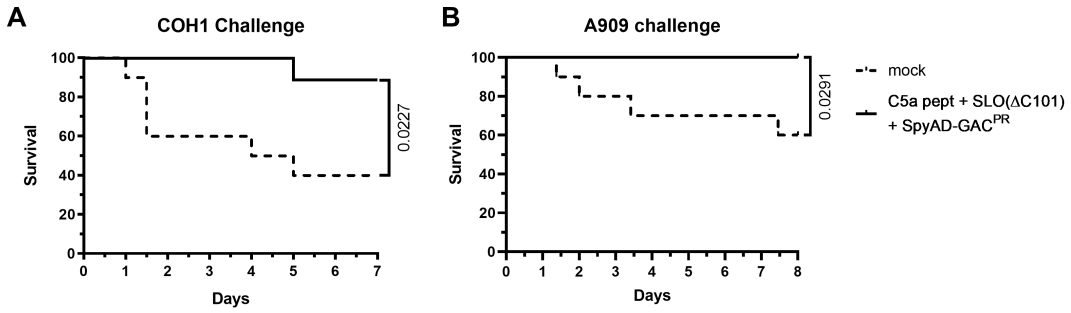

Figure 2.8: **Mice immunized by multi-valent conjugate combination are protected when challenged by lethal GBS systemic infection.** Female, wild-type CD-1 mice received three intramuscular doses of alum-adjuvanted vaccine formulations before infectious challenge with GBS strains. Immunized mice were subjected to lethal intraperitoneal challenge with either **(A)** $2.5x10^8$ CFU COH1 GBS strain (serotype III) or **(B)** $2x10^8$ CFU A909 GBS strain (serotype Ia) and tracked for survival. N = 9-10 mice per group, with statistics calculated from log-rank Mantel Cox test for Kaplan-Meier plots.

2.5 Discussion

In the present work, we deployed a novel conjugation technology to an universally conserved, highly abundant GAS cell wall carbohydrate molecule (GAC^{PR}) that has been engineered to eliminate a suspected cross-reactive epitope (GlcNAc). To functionalize SpyAD as a carrier, the protein was synthesized using the Xpress CF™ cell-free system, incorporating a non-natural amino acid (nnAA) to which GAC^{PR} could be conjugated in a site-specific manner using click chemistry, generating high molecular mass SpyAD-GAC^{PR} conjugates. The sites of nnAA incorporation were chosen so as to avoid disruption of important immunological epitopes, including

both T-cell and B-cell moieties. The successful outcome expands the proven utility of this conjugation platform, which has previously been applied in a 32-valent preclinical pneumococcal conjugate vaccine program[49] and in studies of an experimental malaria conjugate vaccine.[86] The use of bio-orthogonal attachment chemistry incorporated into the non-natural amino acids theoretically allows for more efficient and potent antigen presentation to the immune system, simplified purification, and more well-defined structure of these semi-synthetic immunogens.

We also applied a cell-free *in vitro* transcription-translation system for GAS protein antigen production that has high potential for speed and linear scalability. An advantage of this cell-free system is that viability of a microbial expression strain is not required; removal of this constraint creates unique avenues for optimizing protein production, especially for antigens that might confer cytotoxicity in a cell-based system. Indeed, expression yields for the GAS protein antigens studied herein using the cell-free system were much higher than previously reported in the literature, in particular for the cytotoxin SLO (˜200-fold improvement). Additionally, the lack of a cellular membrane also allows for the exogenous addition of components to manipulate transcription, translation, and folding, and for site-specific incorporation of non-native amino acids into the polypeptide. The protein expression process can be modulated and easily sampled over time with standard industrial microbial fermentation and process equipment.

Another innovative aspect of this work was our aim to devise a GAC^{PR} protein conjugate that avoided using the common pneumococcal vaccine carrier protein antigen CRM_{197}[44], a non-toxic mutant of diphtheria toxin[120]. As the ultimate goal of a GAS vaccine development program would be to enter the pediatric immunization schedule, simultaneous vaccination with different glyco-antigens on the same carrier protein is potentially associated with dampened immune responses. First, pre-existing immunity to a carrier protein may diminish the ensuing immune response to a new antigen conjugated to the same carrier, i.e. carrier-induced epitopic suppression (CIES)[50]. This was seen in children immunized simultaneously with tetanus toxoid-conjugated pneumococcal vaccine and either (a) *Haemophilus influenzae* type B (Hib)

polysaccharide conjugated likewise to tetanus toxoid or, (b) Hib conjugated to diphtheria toxin, and the anti-Hib immune response was demonstrably superior in the latter group[34]. Second, combining different vaccines with CRM_{197} as carrier can induce "bystander interference" extending even to unrelated antigens present in a multivalent subunit vaccine, speculated to derive from competition for limited molecular and cellular resources for antigen presentation within the lymph node[35, 50, 78]. In our study, GAC^{PR} was conjugated to the surface-anchored GAS protein (SpyAD), itself a highly conserved vaccine antigen candidate for the pathogen[10, 52, 141]. SpyAD retained its immunogenicity following rational targeted introduction of non-natural amino acids for site-specific conjugation that left its critical human B and T cell epitopes exposed.

In the present proof-of-principle studies, the novel SpyAD-GAC^{PR} conjugate was combined with two additional universally conserved GAS virulence factors, the surface anchored C5a peptidase and the potent secreted cytotoxin SLO, yielding a 3-component, 4-valent combination vaccine rationally designed to avoid cross-reactivity with human heart muscle epitopes, here corroborated by western blot analysis. This vaccine elicited antibodies that bound the surface of intact GAS of different serotypes, promoted human neutrophil OPK, and showed protection in a murine model of systemic and localized skin M1 GAS infection.

Other multi-subunit formulations are currently under preclinical exploration for GAS vaccination. Among the protein antigens we selected, C5a peptidase and SpyAD were two of the seven proteins selected for inclusion in a multicomponent vaccine (Spy7) that showed efficacy in reducing GAS dissemination in a murine intramuscular infection model[141], SpyAD and SLO were included in a five protein component (5CP) vaccine that reduced murine GAS skin lesion development and accelerated lesion recovery[10], and SLO and C5a peptidase were included in a formulation (Combo5) that produced a reduction in pharyngitis and tonsillitis in a GAS nonhuman primate mucosal infection model[143]. The use of GAS proteins, including SLO, SpyCEP or SpyAD, as both antigen and carrier protein to conjugate GAC, has also recently been reported[40], but this approach (i) use non-specific conjugation methods that can disrupt the

polysaccharide backbone through periodate mediated oxidation, and (ii) were applied to native GAC, containing the potentially cross-reactive GlcNAc sidechain epitope. Our approach instead utilizes CDAP to only derivatize polysaccharide backbone hydroxyls, which allows a site-specific conjugation strategy that can preserve critical protective immune epitopes of the carrier protein when known.

In sum, our general approach may allow the signature, species defining GAC antigen to be safely and effectively included in future GAS subunit vaccine formulations with the goal of broad protection without autoreactivity.

2.6 Methods

2.6.1 Bacterial strains and generation of SpyAD knockout strain

M1 (*emm1*) strain 89155, an invasive disease isolate of U.S. origin from the WHO Collaborating Center for Reference and Research on Streptococci at the University of Minnesota, was the GAS strain used for IgG-binding flow cytometry, neutrophil opsonophagocytic killing assay, neutrophil oxidative burst assay, and murine infectious challenges; this strain shows the common small fragment chromosomal restriction enzyme analysis (REA) pattern 1c64. Additional U.S. Centers for Disease Control and Prevention reference strains used were 3752-05 (*emm2*), 4041-05 (*emm3*), 3979-05 (*emm4*), 4623-05 (*emm5*), 4045-05 (*emm6*), 3749-05 (*emm9*), 3979-05 (*emm11*), 4523-05 (*emm12*), 4626-05 (*emm22*), 4039-05 (*emm28*), 3756-05 (*emm44*), 3487-05 (*emm49*), 4044-05 (*emm77*), and 4264-05 (*emm89*) were kindly provided by B.W. Beall at the CDC Streptococcal Reference Laboratory. M1 5448 is an *emm1* GAS isolate from a patient with severe invasive GAS disease[85]; wild-type 5448 and the SpyAD genetic knock out were only used for IgG-binding flow cytometry (Figure 2.3). A909 (serotype Ia)(ATCC #BAA-1138) and COH1 (serotype III, ATCC #BAA-1176) were selected as representative GBS strains. All GAS and GBS strains were propagated in liquid Todd-Hewitt broth (THB) and either THB agar

plates (THA) or tryptic soy agar plates with 5% sheep's blood (Hardy Diagnostics A10) overnight at 37°C in ambient air without shaking. Used as a control, MRSA USA300 strain TCH1516 was originally isolated from an adolescent patient with severe sepsis syndrome at the Texas Children's Hospital in Houston [63] and propogated in liquid lysogeny broth overnight at 37°C in ambient air with shaking.

To generate the genetic knockout of SpyAD (5448ΔSpyAD) via plasmid integration into the chromosome, an intragenic fragment (300 bp) of *spyAD* (original locus tag Spy0269) was amplified from M1T1 GAS 5448 chromosome using forward primer Spy0629_For_EcoRI (5'-GAATTCAGCAGATCGTAATCGC-3') and reverse primer Spy0629_Rev_EcoRI (5'-GAATTCC CACGTTTAATACC-3'). The PCR product was recovered by TA cloning into pCR2.1-TOPO (Invitrogen), and then subsequently cloned into the temperature-sensitive erythromycin (Erm) resistant plasmid pHY304. The resultant plasmid was transformed into wild-type M1 GAS strain 5448 by electroporation and transformants were plated on THA-Erm 2 μg/ml. Single-crossover chromosomal insertions were identified by shifting to the non-permissive temperature (37°C) while maintaining Erm selection. Integrational knockouts were confirmed by PCR.

2.6.2 GAC purification

GAC^{PR} was purified from GAS 5448Δ*hasA*Δ*gacI*, a M1 serotype strain genetically engineered to lack both the hyaluronic acid capsule and the GlcNAc sidechain on GAC. A bacterial cell pellet from 6.25L growth culture was resuspended and sonicated in ice cold 48% aqueous hydrogen fluoride, stirred at 4°C for 48h, then dialyzed against ice cold deionized H_2O, centrifuged to remove cellular debris, and supernatants lyophilized. Cell wall material was treated with proteinase K overnight, dialyzed in deionized water to remove salts and proteins before supernatants were once again lyophilized. GAC^{PR} was purified from these lyophilized supernatants by size-exclusion chromatography, with positive fractions pooled and re-lyophilized. GAC^{PR} identification and purity were confirmed by gas chromatography/mass spectrometry of

alditol acetate derivatives and linkage analysis performed on partially methylated alditol acetate derivatives as a service of the UC San Diego Glycotechnology Core.

2.6.3 Cloning, Expression and Purification of SLO, C5a peptidase, and SpyAD[4pAMF]

Genes were designed using Biomax ProteoExpert (https://ssl.biomax.de/ ProteoExpert/index.jsp) or DNA 2.0 GeneDesigner (https://www.dna20.com/genedesigner2/) algorithms [182] and re-synthesized (DNA 2.0, Menlo Park, CA). The codon-optimized gene for expression of native SLO (aa79-571), SLO(ΔC101) (aa79-470) and C5a peptidase (aa90-1035) or SpyAD(aa33-849) without or with non-native amino acid namely p-azidomethyl phenylalanine (pAMF) incorporation sites [K64,287,386,657] to generate SpyAD[4pAMF] variant was synthesized at ATUM (Menlo Park, CA) and subcloned with an N-terminal methionine into a proprietary vector. Each of the genes contained an N-terminal his6-tag followed by a TEV protease site [ENLYFQG] for purification of untagged protein. *In vitro* protein expression using Xpress cell free protein synthesis (CFPS) or XpressCF+TM platform was performed as described elsewhere. [86] For titer estimates, expression of SLO (79-571), SLO(C101) (79-470), C5a peptidase(90-1035) or SpyAD[4pAMF] was quantitated through incorporation of 14C-leucine (GE Life Sciences, Piscataway, NJ) into the translating polypeptide during CFPS at 25°C. 4 µl of either the complete CFPS reaction or the reaction supernatant were blotted onto an anion exchange filter membrane, extensively washed to remove unbound material, and heat dried for 30 min. Finally, the filter membrane was evenly coated with scintillation fluid, air dried and the counts recorded to estimate the total and soluble yield of the expressed proteins.

Controlled large-scale antigen expression utilized a DASbox mini bioreactor system for 10 h at 25°C with constant 650 rpm stirring, pH 7.2, dissolved oxygen 30%. After 10 h, reactions were harvested and spun down at 15,000 x g at 4°C for 30 min, passed through a 0.45 µm filter, filtrate loaded on a 5 ml HisTrap excel column (Cytiva) equilibrated and extensively washed

with Buffer A [50 mM Tris, 150 mM M NaCl, 10 mM imidazole] until absorbance returned to baseline. Proteins were eluted using a 50% Step gradient of Buffer B [50 mM Tris, 150 mM M NaCl, 500 mM Imidazole], and elution fractions pooled, concentrated and incubated with excess purified his6-tagged TEV protease overnight while dialyzing against Buffer A. The dialyzed cleavage reaction was loaded back onto a pre-equilibrated HisTrap excel 5 ml column (Cytiva) and untagged proteins collected in the flow-through fractions. Thereafter, the flow-through was concentrated and loaded onto a Superdex 200 26/60 size exclusion chromatography (SEC) column pre-equilibrated with Buffer S [50 mM Tris pH 8.0, 150 mM NaCl] and purity of eluted fractions assessed by SDS-PAGE gel analysis. Purified SpyAD[4pAMF] was incubated with excess Dibenzocyclooctyne-PEG4-tetramethylrhodamine (DBCO-TAMRA) dye for 1h at room temperature to confirm pAMF sites by SDS-PAGE gel and fluorescence readout was recorded using a Syngene G-box gel imager.

2.6.4 Multi-angle light scattering (MALS) analysis

SEC MALS-UV-RI was performed with an Agilent HPLC 1100 degasser, temperature-controlled auto-sampler (4°C), column compartment (25°C) and UV-VIS diode array detector (Agilent, Santa Clara, CA) in line with a DAWN-HELEOS multi-angle laser light scattering detector and Optilab T-rEX differential refractive interferometer (Wyatt Technology, Santa Barbara, CA) coupled to three TOSOH columns in series: TSKgel Guard PWXL 6.0 mm ID x 4.0 cm long, 12 µm particle; TOSOH TSKgel 6000 PWXL 7.8 mm ID x 30 cm long, 13 µm particle; and a TSKgel 3000 PWXL 7.8 mm ID x 30 cm long, 7 µm particle. A mobile phase consisting of 0.2 µm filtered 1x PBS with 5% (v/v) acetonitrile was used at a 0.5 mL/min flow rate and 50-100 µg sample was injected for analysis. Agilent Open Lab software was used to control the HPLC, and Wyatt Astra 7 software was used for data collection and molecular weight analysis.

2.6.5 Dibenzocyclooctyne (DBCO)-derivatization of GAC^{PR} and conjugation to SpyAD[4pAMF]

To a 6 mM solution of GAC^{PR} in 100 mM Borate Buffer pH 8.5, three equivalents (to the polysaccharide repeating unit) of 1-cyano-4-dimethylaminopyridinium tetrafluoroborate (CDAP; from 100 mg/mL solution in acetonitrile) were added with vigorous stirring to facilitate cyanylation at reactive hydroxyl groups. 5 min after addition of CDAP, 2 molar equivalents of dibenzocyclooctyne-amine linker stock in DMSO was added such that the final DMSO concentration was 5% (v/v). After DBCO-derivatization, 200 mM glycine was added to the reaction to quench unreacted cyanate esters. The DBCO-derivatized polysaccharide was purified via zeba spin desalting column and the purity of the recovered material was assessed by reverse phase. A single peak in HPLC when absorbance was monitored at 309 nm confirmed complete removal of excess DBCO linker and other reaction byproducts. Finally, the polysaccharide concentration was measured using anthrone assay and dibenzocyclooctyne concentration was measured using absorbance at 309 nm. These two values were combined to give an estimate of the percentage of polysaccharide derivatized with a dibenzocyclooctyne functional group. For conjugation, %DBCO derivatization of the GAC^{PR} was kept between 5-10%. Thereafter, SpyAD[4pAMF] was mixed with DBCO-derivatized GAC^{PR} at a 1:1 ratio [0.5 mg/ml each] to facilitate conjugation via click chemistry. Post-conjugation, the reaction mixture was dialyzed against a 50 kDa cutoff membrane to remove excess unreacted free polysaccharide. The recovered conjugates were analyzed by SEC MALS and the concentration was estimated using an anthrone assay.

2.6.6 Anthrone assay for total polysaccharide concentration

A stock of 2 mg/ml of the anthrone reagent (Sigma-Aldrich, CAS#90-44-8) was prepared in cold sulfuric acid while a 1 mM stock of polysaccharide repeating unit (PSRU) comprising 2x

rhamnose was prepared in water as a standard. In triplicate wells, 100 µl of PSRU stock (serially diluted into reference standards) or the unknown samples (diluted 1:3) were plated (96-well plate) followed by addition of 200 µl/well of the anthrone reagent stock. All reactions were thoroughly mixed and sealed with a plate cover for incubation at 95°C for 10 min. The plate was briefly placed on ice to cool to ambient temperature before absorbance is measured at 620 nm using a UV/VIS plate reader. To determine concentration of unknown samples, PSRU standard concentrations and absorbances were used to generate a least-square fit regression.

2.6.7 Generation of immunized rabbit serum

Age-matched New Zealand white (NZW) rabbits were intramuscularly (IM) injected with 5 µg of protein antigens (SLO, C5a peptidase, SpyAD, M1) or SpyAD-GAC^{PR} conjugate (equivalent of 5 µg of polysaccharide) in succinate buffer adjuvanted with Adju-phos® (Invivogen). Immunizations (250 µl per injection per dose) were performed on day 0, 14 and 28 with terminal bleed for serum performed on day 35.

2.6.8 Antibody titer of rabbit serum by ELISA

Anti-protein antibody titers of rabbit antisera were determined by ELISA. 3 µg recombinant protein in PBS per well was incubated overnight in a high-binding flat bottom 96-well plate (Corning #3361) at room temperature, antigen-coated plates washed three times in PBS with 0.05% Tween-20 (PBST), blotted dry, and blocked for 2 h in PBS + 1% bovine serum albumin (R&D Systems cat#DY995). Plates were washed x3 with PBST and blotted dry before addition of rabbit antisera, serially diluted in PBS with 1% BSA starting from 1:1000 dilution. Antisera were incubated for 2 h at room temperature, plates washed three times with PBST, blotted dry, then incubated with horseradish peroxidase (HRP)-conjugated goat-anti rabbit IgG secondary antibody (Southern Biotech cat#4050-05), diluted 1:4000 in PBS with 1% BSA, for 2 h at room

temperature. Plates were washed three times with PBST and blotted dry before addition of TMB substrate (BD cat#555214) per manufacturer's instruction and incubated in the dark for 5 min. Horseradish peroxidase (HRP) reaction was stopped by the addition of 2N sulfuric acid (Sigma) before optical density was read at 450 nm. Antibody titer calculated as highest serum dilution where the signal exceeded the signal of blank wells plus three standard deviations, and all samples were run in at least duplicate for 2-3 rabbits per immunization group. For subsequent experiments, serum from the rabbits with the highest anti-protein antigen titers were used: 1:2,050,000 for SLO, 1:6,550,000 for C5a peptidase, 1:3,280,000 for SpyAD, and 1:6,550,000 for SpyAD-GAC^{PR}. Statistics shown use one-way ANOVA with Dunnett's multiple comparisons test to compare each immunized group with the dilution factor of the pooled pre-immune system.

2.6.9 IgG binding to GAS strains

All GAS strains were grown to mid-logarithmic growth phase (OD600nm = 0.4) and washed in PBS before blocking incubation in 10% donkey serum for 1h at room temperature. Murine or rabbit antisera was added to bacteria to 2% final volume and incubated for 1 h at room temperature to allow antibodies to bind to bacteria surfaces. Samples were washed with PBS and incubated with 1:200 donkey anti-rabbit IgG conjugated with AlexaFluor 488 fluorophore (Thermo Fisher #21206), protected from light for 30 min at room temperature. Samples were washed in PBS once and run on a BD FACSCalibur. Signal intensity was analyzed using FlowJo software (Tree Star) and reported as a percentage of increased mean fluorescence intensity signal in individual rabbit antiserum compared to pooled pre-immunized rabbit sera.

2.6.10 Primary human neutrophil opsonophagocytic killing (OPK) assay

Neutrophils were isolated from blood drawn from healthy human donors with consent, as approved by UC San Diego institutional review board. Neutrophils were pre-incubated with

baby rabbit complement (BRC, PelFreez cat#31061) and heat-inactivated fetal bovine serum (FBS) for 10 min. All GAS strains were grown to mid-logarithmic growth phase (OD600nm = 0.4), washed in PBS, and incubated with heat-inactivated murine or rabbit antisera for 30 min at 37°C. For combination of rabbit serum, total anti-serum volume was kept the same but consisted of an even pool of serum from multiple rabbits. Neutrophils were added to bacteria at a multiplicity of infection (MOI) = 0.1 bacteria per neutrophil, briefly centrifuged to ensure contact, and allowed to incubate for 30 min at 37°C with 5% CO2. Final concentrations of components were % murine or rabbit serum, 2% FBS, and 2% BRC, and the remaining volume comprised bacteria and neutrophils in PBS. At experiment termination, samples were serially diluted in PBS and plated onto THB agar plates for CFU enumeration. Sera from pre-immunized rabbits were pooled and used as control to measure non-specific, baseline bacterial killing by neutrophils for the rabbit antisera samples. Student's t-test was used to compare the mouse serum OPK assay, and one-way analysis of variance (ANOVA) with Bonferroni's multiple comparisons post-test was used to compare the rabbit pre-immune serum killing with each immune serum group. At minimum, each serum or serum combination was tested in triplicate assays using neutrophils from two different donors to ensure statistical confidence.

2.6.11 Primary human neutrophil oxidative burst assay

Primary human neutrophils were isolated as described for the OPK assay. 2×10^6/mL human neutrophils were loaded with 20 μM 2,7-dichlorofluorescein diacetate (DCFH-DA; Fisher) in Hank's balanced salt solution (HBSS, Cellgro) without Ca^{2+} and Mg^{2+} and incubated with rotation at 37°C for 20 min. Neutrophils were resuspended in HBSS with Ca^{2+} and Mg^{2+} to a density of 1×10^6 cells/well before treatment with rabbit serum (1:64 final concentration) followed by supernatant from mid-logarithmic growth phase M1 89155 bacterial cultures (diluted 1:10). Lastly, 50 ng/ml phorbol myristate acetate (PMA) was added to wells, and incubated at 37°C and 5% CO_2, and fluorescence intensity at 485 nm excitation/520 nm emission quantified

on an Enspire plate reader (Perkin Elmer) at the indicated time points.

2.6.12 *In vivo* mouse immunization studies

All mouse experiments were approved by the UC San Diego Institutional Animal Care and Use Committee and conducted per accepted veterinary standards. For passive immunization experiments, 200 µl of rabbit antisera was intravenously delivered via retro-orbital injection in anesthetized wild-type female CD-1 mice (Charles River), 8-10 weeks of age. After 5 min, mice were challenged with $1x10^7$ CFU M1 strain 89155 GAS by intraperitoneal (i.p.) injection and tracked for survival. For active immunization, wild-type female CD-1 mice (Charles River) were immunized every 14 days for a total of 3 doses starting at 5-7 weeks of age. Intramuscular immunizations delivered consisted of 100 µl total volume per mouse per dose, including 50 µl of Alhydrogel 2% aluminum hydroxide adjuvant (Invivogen), prepared per manufacturer's instructions. 14 days after final immunization, mice were infected with $1x10^7$ CFU M1 89155 GAS by i.p. injection and tracked for survival. Statistics of Kaplan-Meier survival curves calculated using log-rank Mantel-Cox test.

For intradermal infection, female CD-1 mice were immunized as previously mentioned. Mice were shaved and chemically depilated under isoflurane anesthesia and allowed to recover for a day prior to infection. Isoflurane-anesthetized mice were injected with 20 µl prepared M1 89155 culture (1×10^6 CFU) intradermally using 26 ½ gauge needles on Hamilton 1000 µl glass syringe (cat# 81320) with PD600 repeating dispenser (cat# 83700) for reproducible precision. Developing lesions was tracked and photographed daily. On day 3 post-infection, mice were euthanized, lesions excised and homogenized using MagNaLyser (Roche) and serially diluted in technical triplicate onto tryptic soy agar plates with 5% sheep's blood (Hardy Diagnostics A10) for CFU enumeration. Lesion sizes (area) were quantified using FIJI[155]. Statistics of both lesion sizes and recovered CFU were calculated using unpaired t-test with Welch's correction.

2.6.13 Western blot analysis for heart / brain lysate cross reactivity

Varying amounts of normal adult human heart tissue lysate (Novus Biologicals cat# NB820-59217) or brain lysate (Novus Biologicals cat. # NB820-59177) incubated with SDS-containing denaturing loading dye were separated by SDS-PAGE using 4-12% Bis-Tris gels before transfer onto a PVDF membrane using the manufacturer's protocol on iBlot (Thermo Fisher). The blot was blocked at ambient temperature for 1 h, followed by probing with rabbit antisera generated against each of the GAS antigens (diluted 1:1000). After three 30 min washes, HRP conjugated anti-rabbit (Jackson ImmunoResearch Laboratories, Inc., Cat # 211-035-109) secondary antibody (diluted 1:10,000) was added and chemiluminescence recorded on a Syngene G-Box F3 image scanner after incubation of the blot with the pico substrate (Thermo Fisher Scientific Cat # 34580). All blot blocking, washes, and antibody / serum dilutions were performed in TBS + 5% BSA.

2.6.14 Statistical analysis

All statistical analyses were performed using GraphPad Prism version 8.0.0 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com. P values were summarized for respective analyses as: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***).

2.7 Acknowledgements

This chapter contains some material being prepared for submission for publication and some material is, in part, a reprint of the material as it appears in *Infectious Microbes and Disease*. "**Nina J. Gao**, Satoshi Uchiyama, Lucy Pill, Samira Dahesh, Joshua Olson, Leslie Bautista, Shilpa Maroju, Aym Berges, Janet Z. Liu, Raymond H. Zurich, Nina van Sorge, Jeff Fairman, Neeraj Kapoor, Victor Nizet. *Site-Specific Conjugation of Cell Wall Polyrhamnose to Protein SpyAD Envisioning a Safe Universal Group A Streptococcal Vaccine*. Infect Microb

Dis 2020;00(00):00–00. doi: 10.1097/IM9.0000000000000044". The dissertation author was a primary investigator and author of this work.

Research was funded by CARB-X, Vaxcyte, Inc., and NIH Grant AI077780. NJG was supported in part by a Ruth L. Kirschstein Institutional National Research Award from the National Institute for General Medical Sciences, T32 GM008666.

Chapter 3

Functional and Proteomic Analysis of *Streptococcus pyogenes* Virulence Upon Loss of Its Native Cas9 Nuclease

3.1 Abstract

The public health impact of *Streptococcus pyogenes* (group A *Streptococcus*, GAS) as a top ten cause of infection-related mortality in humans contrasts with its benefit to biotechnology as the main natural source of Cas9 nuclease, the key component of the revolutionary CRISPR-Cas9 gene editing platform. Despite widespread knowledge acquired in the last decade on the molecular mechanisms by which GAS Cas9 achieves precise DNA targeting, the functions of Cas9 in the biology and pathogenesis of its native organism remain unknown. In this study, we generated an isogenic serotype M1 GAS mutant deficient in Cas9 protein and compared its behavior and phenotypes to the wild-type parent strain. Absence of Cas9 was linked to reduced GAS epithelial cell adherence, reduced growth in human whole blood *ex vivo*, and attenuation of virulence in a murine necrotizing skin infection model. Virulence defects of the GAS Δ*cas9* strain were explored through quantitative proteomic analysis, revealing a significant reduction in the abundance of key GAS virulence determinants. Similarly, deletion of *cas9* affected the expression of several known virulence regulatory proteins, indicating that Cas9 impacts the global architecture of GAS gene regulation.

3.2 Preface to Chapter 3

The application of Cas9 as a gene editing tool has led to a revolution in scientific research and potential therapeutic applications. This re-discovery casts light once again on the CRISPR-Cas system, first identified in GAS. Recent publications have identified that Cas9 plays a role in regulation of virulence in other pathogenic bacteria. In collaboration with the expertise of the Zengler and Gonzalez labs, we sought to explore the role of Cas9 beyond defense against bacteriophages.

3.3 Introduction

Clustered regularly interspaced short palindromic repeats (CRISPR) genes and CRISPR-associated (Cas) genes are recognized as an adaptive immune system that allows prokaryotic organisms to defend against plasmids, bacteriophages and transposons [8]. CRISPR-Cas systems are widely distributed in many bacterial and archaeal genomes [16, 108], and are evolutionarily classified in two main classes, with class II as the most representative and uniquely driven by the nuclease Cas9 [108]. Type II CRISPR-Cas systems occur only in bacteria, and not in archaea [69].

A variety of important human pathogens possess a type II CRISPR-Cas system, including bacterial species that cause acute or chronic infections [103]. Several lines of investigation support the notion that endogenous bacterial factors involved in stress responses and virulence gene regulation might interact to modulate the expression of CRISPR-Cas genes. For example, mutants in stress adaptation regulatory proteins RelAQ down-regulate cas genes in *Enterococcus faecalis* [186], deletion of the osmotic regulator OmpR represses cas gene expression in *Yersinia pestis* [57], and *Escherichia coli* two-component regulatory system (TCS) BaeSR modulates cas genes expression in response to cell envelope stress [133].

Genomic analyses of virulence features in diverse pathogenic bacteria suggest roles of CRISPR-Cas beyond defense against foreign DNA and viruses, including potential involvement in regulation of endogenous gene expression [123], including those encoding virulence factors [92]. These hypotheses have been supported experimentally in a number of cases. For example, using Cas9 and tracrRNA as regulators, *Francisella novicida* represses a key surface-expressed lipoprotein (BLP), avoiding recognition of the pathogen by host cellular receptors [152]. In addition, CRISPR-Cas modulates swarming and biofilm formation in *Pseudomonas aeruginosa* [187], CRISPR-associated Cas2 enhances intracellular infection by *Legionella pneumophila* [68], a CRISPR type II system contributes to *Campylobacter jejuni* attachment to and invasion of human intestinal epithelium [102], and cas9 deletion reduces *Neisseria meningitidis* epithelial cell adher-

ence and invasion [152]. Recently, inactivation of *cas9* in *Streptococcus agalactiae* was shown to impair epithelial cell adherence and macrophage intracellular survival, which is translated to decreased virulence of the Δ*cas9* mutant strain in zebrafish and murine infection models [106].

Although Cas9 nuclease is found in many bacterial genomes, the native source of the Cas9 used in genome engineering is *Streptococcus pyogenes* (group A *Streptococcus*, GAS). Dubbed the most significant genetic tool of the 21st century [132], GAS Cas9 has been utilized for precise and efficient gene editing in species ranging from bacteria [82], to yeast [41], to cynomolgus monkeys [126], and human cell lines [30]. While the GAS CRISPR-Cas9 system is one of the best understood biochemically [109], its influence on the physiology and the pathogenesis of its native organism remain unknown. This is striking since GAS remains a top ten cause of infection-associated mortality worldwide, producing a wide spectrum of diseases with multiple clinical manifestations, ranging from mild impetigo and pharyngitis, to severe invasive toxic shock syndrome and necrotizing fasciitis [18, 31].

Group A *Streptococcus* possesses a multitude of surface-bound and secreted virulence factors that subvert innate defenses and allow the pathogen to survive and replicate in the human host [168, 180]. Control of virulence gene expression in GAS involves a complex, interconnected network of TCS and specific and/or global transcriptional regulators. Together, these virulence regulators integrate environmental host cues with the pathogen's own metabolic state, as well as feedback signals from the expressed genome, into a coordinated response [177].

In this study, we present evidence that endogenous Cas9 impacts GAS pathogenesis. Specifically, Cas9 is required for efficient GAS adherence to epithelial cells, growth in human blood, and full virulence in a murine skin infection model. Unbiased proteomic analysis shows how Cas9 influences the protein abundance of several key GAS virulence factors and regulators of virulence gene expression.

3.4 Results

3.4.1 Generation of a GAS cas9 Mutant (Δ*cas9*)

To explore the functional role of Cas9 in GAS, we generated a precise in-frame allelic exchange mutant in the background of the well characterized globally disseminated GAS serotype M1T1 strain 5448 [85], wherein the *cas9* gene in the type IIA CRISPR operon was replaced with chloramphenicol acetyltransferase (*cat*) gene (Figure 3.1A) as validated by PCR and sequence analysis. Functional confirmation of gene deletion in the GAS Δ*cas9* strain was achieved in both mid-exponential and stationary growth phase cultures by real-time qPCR (Figure 3.1B) and western immunoblot using specific antibodies raised against the GAS nuclease (Figure 3.1C). Additionally, we confirmed that the replacement of cas9 with the cat gene did not exert polar effects on the expression of downstream cas genes, as the transcriptional levels of *cas1*, *cas2*, and *csn2* did not differ significantly between WT and Δ*cas9* strains (Figure 3.1B).

Loss of Cas9 did not affect growth in bacteriological media (Figure 3.1D), and wild-type (WT) and Δ*cas9* GAS strains had similar morphology and chain length distribution in brightfield microscopy imaging (Figure 3.1E), and similar susceptibility to cell wall-active antibiotics vancomycin and penicillin (Supplementary Table A.1). Genetic complementation of the Δ*cas9* strain with a plasmid expressing full-length Cas9 under a constitutive promoter restored production of the Cas9 protein product above WT levels (A.1A); however, the complemented strain showed a significant defect in growth compared to both WT and Δ*cas9* strains (A.1B). This result was consistent with Cas9 overexpression-mediated toxicity, limiting our attempts to complement specific Δ*cas9* phenotypes (A.1C and D), while at the same time suggesting a non-canonical function of Cas9 in GAS.

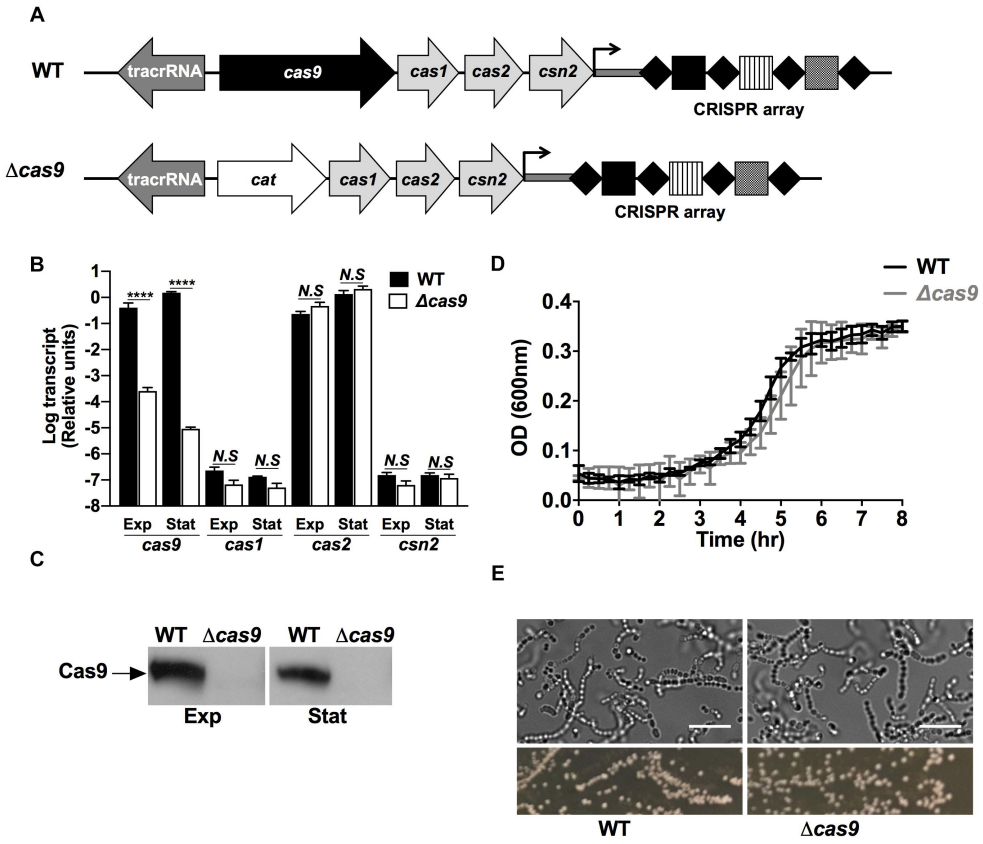

Figure 3.1: **Deletion of the cas9 gene does not affect GAS growth and morphology.** **(A)** Schematic of the genomic organization of type II-A CRISPR-Cas loci in GAS 5448 wild type (top) and Δ*cas9*) (bottom) strains. The cas genes are represented in light gray with *cas9* highlighted in black. The tracrRNA is shown in dark gray. Substitution of cas9 by the *cat* gene in the Δ *cas9*) strain is represented in white. The CRISPR array of GAS 5448 is constituted by the leader sequence (dark gray bar), four repeats (black diamonds) and three spacers (squares). **(B,C)** Quantification of *cas9*, *cas1*, *cas2*, and *csn2* mRNA transcripts by RT-qPCR **(B)** and expression of Cas9 protein by western blot **(C)** from wild type (WT) and *cas9* mutant (Δ*cas9*) GAS strains grown to mid-exponential (Exp) or stationary (Stat) cell growth phases. **(D)** Cellular growth curves of WT Δ*cas9* GAS strains grown at 37°C in THB media. **(E)** Microscopic visualization (top panels) and colonies morphology (bottom panels) of WT and Δ*cas9*) GAS strains grown to stationary growth phase in THB media or on THA agar plates, respectively. Scale bar (10 μm) is indicated. For each experiment, samples were assayed at least in triplicate. Data in (**(B)** and **(D)** are plotted as the mean ± SEM and are pooled and representative of three independent experiments, respectively. Data in **(B)** was analyzed by two-way ANOVA multiple comparisons.

3.4.2 Cas9 Deficiency Is Associated With Reduced Abundance of Key GAS Virulence Determinants and Regulatory Factors

A number of studies in pathogenic bacteria have suggested Cas proteins play important roles in biological processes beyond the well-studied adaptive immune system that protects against foreign DNA [123]. To explore the overall functional impact of Cas9 on GAS pathophysiology, we compared the proteomic profiles of GAS WT and Δ*cas9* strains from cells grown to mid-logarithmic phase using quantitative, multiplexed proteomics [98]. We found that deletion of *cas9* resulted in drastic remodeling of the GAS proteome. From a total of 1,224 proteins identified, the abundance of 340 proteins was significantly decreased, and 405 proteins significantly increased, in the Δ*cas9* mutant compared to the WT parent GAS strain (Supplemental Figure A.2 and Supplemental Table A.2).

To understand functional changes attributable to Cas9 deficiency, we undertook gene annotation by Rapid Annotations Subsystems using Technology (RAST server) [4] and classified all identified proteins into subcategories by predicted gene function (Supplementary Table A.3). Among the identified proteins, RAST analysis yielded 17 functional subcategories significantly affected by loss of Cas9, with 11 subcategories enriched in the WT strain, and 6 subcategories enriched in the Δ*cas9* strain (Figure 3.2A). The most enhanced subcategories in the Δ*cas9* mutant strain were "Sugar alcohols" and "Membrane Transport" (Figure 3.2A), including glycerophosphoryl diester phosphodiesterase (encoded by M5005_Spy0647) and co-factor transporters (encoded by *cbiQ* and *cbiO1*) (Supplementary Table A.3). In contrast, the most pronounced subcategories enriched in the WT strain were "Electron Accepting Reactions" and "Respiration" (Figure 3.2A), including important bacterial metabolic proteins such as arsenate reductase (encoded by *arcA*), glycerol dehydrogenase (encoded by *gdlA*), and ferrodoxin (encoded by M5005_Spy0616) (Supplementary Table A.3). Together these observations suggest that loss of Cas9 may indirectly lead to imbalances in the metabolic status of GAS.

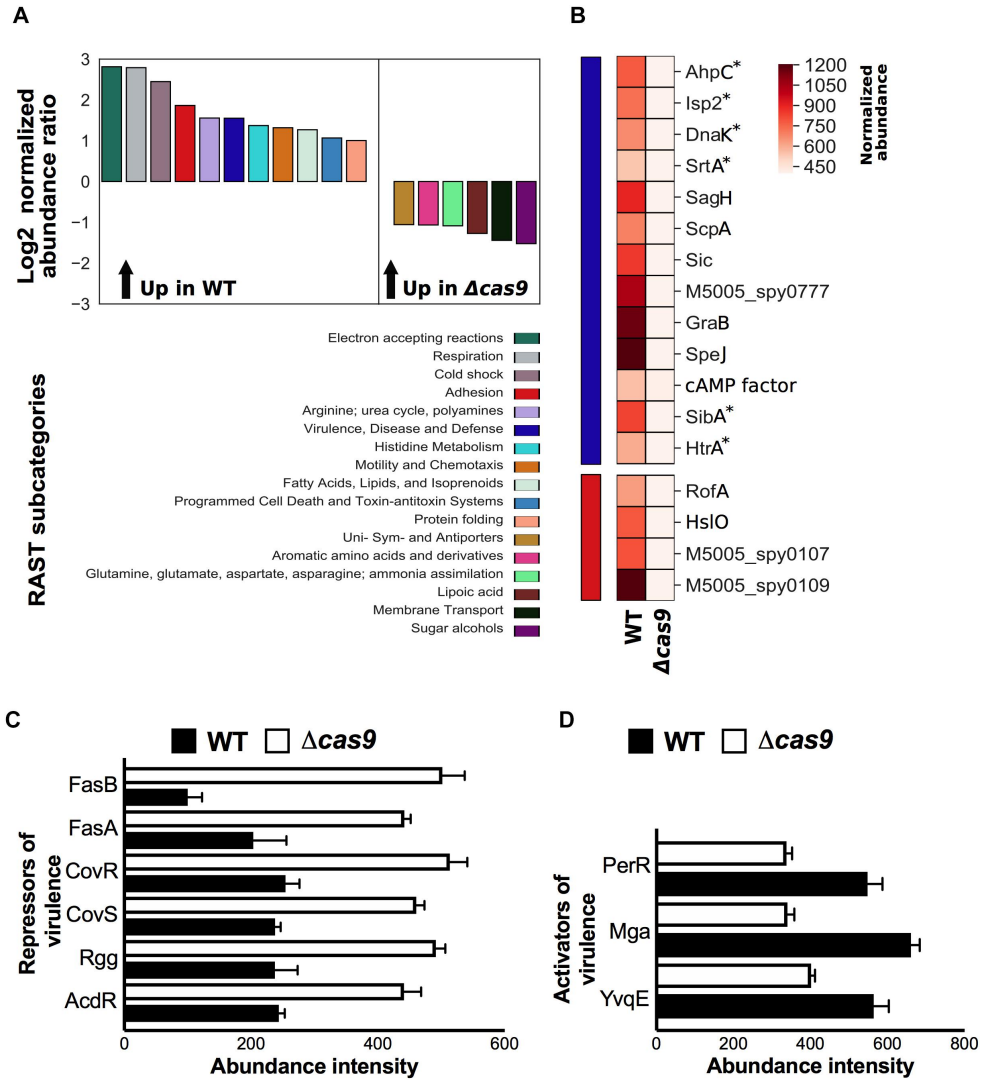

Figure 3.2: **Cas9-deficiency is associated with RAST functional enrichment and differential abundance of GAS virulence determinants and regulators.** **(A)** Comparison of significant differential protein abundance at the RAST subcategory level (Log2 normalized abundance ratio) between wild type (WT) and isogenic *cas9* mutant (Δ*cas9*) GAS strains. Significant enrichment in protein abundance in the wild-type (up in WT) or in the *cas9*-deficient mutant strain (up in Δ*cas9*) is indicated (top panel), with a matching color code of the corresponding RAST subcategory (bottom panel). **(B)** Statistical representation of the differential abundance of proteins identified in the WT or Δ*cas9* GAS strains and grouped under "virulence, disease, and defense" (blue bar) and "adhesion" (red bar) RAST subcategories. Color pattern of the normalized protein abundance is indicated **(C,D)** Comparison of normalized abundance of transcriptional repressors **(C)** and activators **(D)** of GAS virulence identified by proteomics in the WT and Δ*cas9* GAS strains. Proteins annotated with asterisks indicate that were manually added (curated) into the relevant RAST subcategory based on published evidence. All differential proteins listed in **(B,D)** were significant by t-test ($p < 0.05$).

Of particular interest, virulence-related RAST subcategories were markedly reduced in the Δ*cas9* mutant compared to the WT GAS strain. To that end, we focused our proteomic analysis on the normalized abundance levels of individual gene products grouped within the "Adhesion" (red) and "Virulence, Disease, and Defense" (blue) RAST subcategories (Figure 3.2B). At least four proteins associated with GAS adhesion, including the chaperone Hslo, the transcriptional regulator of adhesins RofA, pilin (M5005_spy0109), and pilus ancillary protein 1 (M5005_spy01070) were significantly reduced in the Δ*cas9* strain, potentially impacting GAS host cell adherence and colonization. At least 13 additional GAS virulence-associated proteins had significantly reduced abundances in the Δ*cas9* strain (Figure 3.1), including immunogenic secreted products Isp2 and SibA, and proteins involved in folding and maturation of other secreted GAS virulence determinants. Among these is HtrA, a serine protease directly associated with maturation of two key pathogenic factors: cysteine protease SpeB and pore-forming toxin streptolysin S (SLS) [105].

Group A Streptococcus genes encoding proteins induced during human neutrophil phagocytosis [179] were also significantly reduced upon Cas9 deficiency, including detoxifiers of cell-damaging reactive oxygen species (ROS), such as AhpC and DnaK, and sortase A (SrtA), the transpeptidase required for cell wall anchoring of surface virulence factors such as M protein, protein G-related α2-macroglobulin-binding protein (GRAB), and protein F [7, 140]. Proteins involved in GAS evasion of the complement system, which has a central role in innate immunity, were also diminished in the Δ*cas9* strain, These included (a) serine protease ScpA that specifically cleaves C5a, a key chemoattractant factor that also helps coordinate activation of the classical, alternative and lectin-binding complement pathways [26] and (b) SIC (streptococcal inhibitor of complement), which further inactivates antimicrobial factors including cathelicidin defense peptides, α-defensins, secretory leukocyte protease inhibitors and lysozyme [27].

Other virulence-related proteins found to be less abundant in the Δ*cas9* strain include protein G-related α2-macroglobulin-binding protein (GRAB), which protects important GAS virulence determinants from proteolytic degradation [139]; CAMP factor, linked with GAS

epithelial cell adherence and resistance to macrophage phagocytosis [94, 93], SagH (an SLS export transmembrane protein), and superantigen SpeJ. The reduced abundance of virulence-related proteins was a first clue to the possibility of virulence attenuation in the Δ*cas9* mutant.

A complex network of two-component regulatory systems (TCS), global and specific transcriptional regulators exert efficient and rapid control over the expression of all the aforementioned GAS virulence-related proteins and other factors relevant to pathogenicity. Since Cas9 controls key transcriptional regulatory elements in other pathogens [106], we next analyzed the differential abundance of virulence regulatory proteins between GAS WT and Δ*cas9* strains. Remarkably, the abundance of several transcriptional repressors of virulence were increased in the Δ*cas9* strain (Figure 3.2C). This included the master TCS CovR/CovS, which influences transcription of up to 15% of all GAS chromosomal genes, including repression of hyaluronic acid capsule, SLS precursor SagA, streptokinase (SkA), cysteine protease SpeB and other secreted GAS factors [64]. Also increased in the Δ*cas9* mutant was the TCS FasA/FasB, which down-regulates transcription of genes encoding GAS adhesins in a growth phase-dependent fashion (e.g., *fbp54*, *mrp*) [90]. Finally, transcriptional regulatory proteins Rgg, which down-regulate several genes associated with GAS virulence [20] and AdcR, involved in the repression for adaptive responses to zinc limitation [154] were also more abundant in the Cas9-deficient strain.

In contrast, transcriptional activators of virulence determinants were diminished in the Δ*cas9* strain (Figure 3.2D). One such example is Mga, the best-characterized stand-alone virulence regulator of GAS, which induces a core set of virulence genes, including M protein, the most abundant GAS surface protein. Similarly, protein abundance levels of the transcriptional regulator PerR and the histidine kinase YvqE were reduced in the absence of Cas9; these proteins are known to directly up-regulate GAS responses to oxidative stress and thereby enhancing resistance and virulence in the host [65], and signaling-mediated control of biofilm formation and pilus expression [79], respectively.

In summary, loss of Cas9 is associated with changes of several GAS virulence-related

regulatory elements, generally fitting a pattern of reduced activators and enhanced repressors, suggesting an important role of the nuclease on the overall virulence of the bacterium.

3.4.3 Loss of Cas9 Is Associated With GAS Virulence Attenuation

To functionally validate our proteomic observations pointing toward an altered virulence phenotype in the Δ*cas9* strain, we first compared the expression and/or activity of well-known GAS virulence determinants. The GAS hyaluronic acid (HA) capsule varies in thickness across different strains [2]. High level HA capsule expression can produce a mucoid colony morphology and plays a critical role in resistance to opsonophagocytosis and evasion of the host innate immune response [184, 37]. Visual comparisons between the WT and Δ*cas9* strains did not reveal differences in mucoid morphology of the bacterial colonies (Figure 3.1E, bottom panels), and the amount of capsular HA extracted from mid-exponential growth phase bacteria was similar in the two strains by hyaluronan specific ELISA (Figure 3.3A). These findings were consistent with the proteomics results showing similar expression of hyaluronan synthase (HasA) in both strains (Supplementary Table A.2).

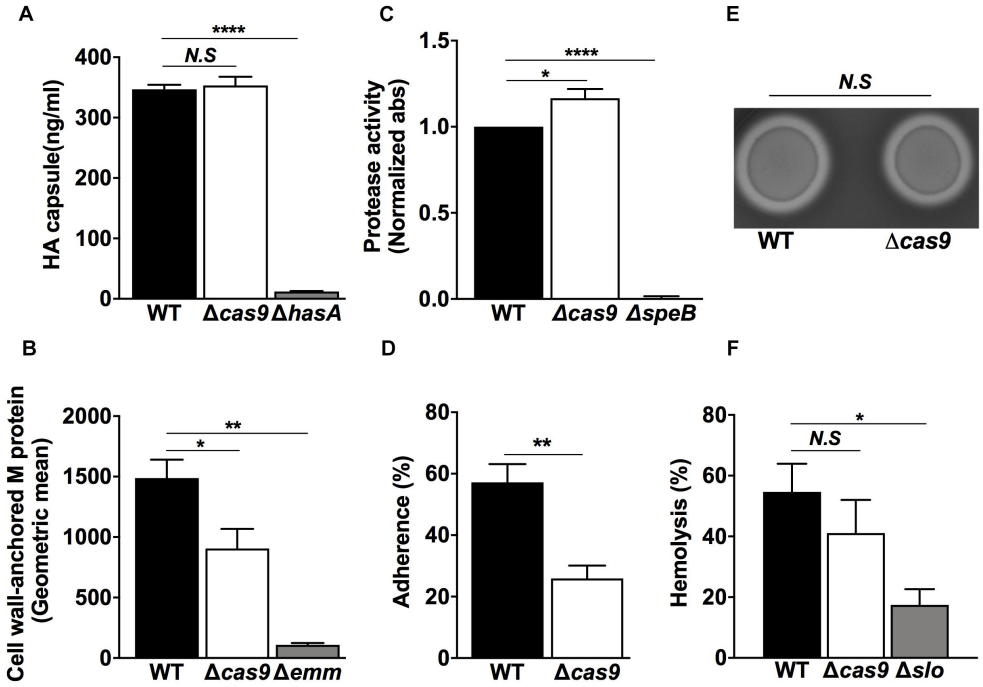

Figure 3.3: **Lack of Cas9 is impaired with significant changes in key virulence factors and pathogenic functionalities of GAS.** GAS WT and Δ*cas9* strains were assessed for **(A)** capsule expression by ELISA, **(B)** quantification of M protein-anchored to the cell wall by flow cytometry, **(C)** SpeB protease activity by cleavage of azocasein, **(D)** capacity to adhere to HaCaT human skin keratinocytes, **(E)** β-hemolysis on blood-agar media, and **(F)** ability to lyse human red-blood cells. Isogenic GAS mutant strains in capsule (Δ*hasA*), M protein (Δ*emm*), SpeB (Δ*speB*), and SLO (Δ*slo*) were used as negative controls in **(A–C, F)**, respectively. For each experiment, samples were assayed at least in triplicate. Data in **(A–D, F)** are plotted as the mean ± SEM, pooled from three independent experiments and analyzed by Student's t-test.

The surface-anchored M protein forms the basis for the serological differentiation of GAS strains, and influences several pathogenic properties of the bacterium such as epithelial cell adherence[130] and resistance to opsonophagocytosis. M protein can also bind several host components including fibrinogen and immunoglobulin G [61], and block membrane-lytic activities by sequestering antimicrobial peptides [99] and histones [45]. M protein has pro-inflammatory properties that drive the pathogenesis of streptococcal sepsis[73] and activate host IL-1β signaling through NLRP3 inflammasome activation [169]. Using flow cytometry, we measured a significant reduction of cell wall-associated M protein in the Δ*cas9* strain compared to the WT strain (Figure 3.3B), consistent with a trend toward lower M protein detected by our proteomic experiments (Supplementary Table A.2), though this fell short of statistical significance (p-value = 0.06). Reduction of cell wall-associated M protein could be also attributed to the reduced amounts of the specific M protein transcriptional activator (Mga) and/or its surface anchor sortase (SrtA), found in the Δ*cas9* mutant by proteomics (Figures 3.2B and D, respectively).

M protein is also one of the multiple GAS virulence factors recognized and cleaved by cysteine protease SpeB, the most predominant secreted protein produced by the pathogen[5, 125]. SpeB contributes to the establishment of localized skin infections [28] and enhances GAS persistence and dissemination by degrading multiple host proteins [47, 129, 157]. Our whole cell proteomic analysis did not show differences in the abundance of SpeB between WT and Δ*cas9* strains (Supplementary Table A.2) but did not capture the secreted protein fraction of SpeB. Thus, we directly studied extracellular protease activity of SpeB from bacterial supernatants of both strains using the azocasein assay [75] and found a significant increase in protease activity in the Δ*cas9* strain compared to the WT counterpart (Figure 3.3C). This increased cysteine protease activity could contribute to the reduced surface-attached M protein observed in the Δ*cas9* strain (Figure 3.3B).

A primary step in GAS colonization of the host is adhesion to host epithelial cells [13, 128]. Adhesion-related proteins were highlighted in our proteomic analysis, with a significant reduction

in the abundance of proteins, such as pilin and pilus components, in the Δ*cas9* mutant strain (Figures 3.2B and C). We compared adherence of WT versus Δ*cas9* to human HaCaT keratinocytes and found a significant three-fold reduction in Δ*cas9* adhesion (Figure 3.3D), suggesting a Cas9-dependent effect on GAS ability to bind to human host cell.

β-hemolysis is a hallmark phenotypic feature of GAS [127] and the oxygen-stable streptolysin S (SLS) is the main factor responsible for red cell lysis on blood agar media. SLS forms hydrophilic pores in a broad array of epithelial and immune cell types [122, 124]. SagA, the secreted structural propeptide for the SLS toxin [36, 127] was not detected in our proteomic analysis, nor were there differences in abundance of the oxygen-labile pore-forming, cholesterol-dependent streptolysin O (SLO) another important secreted toxin [107], between the WT and Δ*cas9* strains (Supplementary Table A.2). Consistent with these findings, the WT and the Δ*cas9* mutant strains did show significant differences in the zone of β-hemolysis surrounding the GAS colonies (Figure 3.3E) or hemolysis in a liquid phase red blood cell lysis assay (Figure 3.3F).

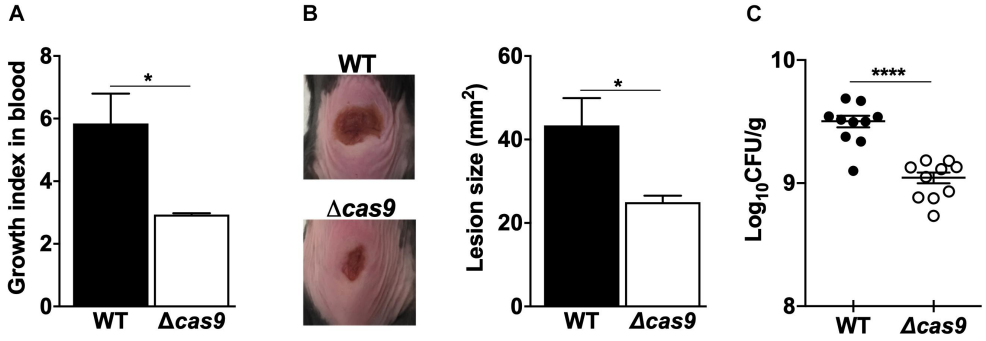

Figure 3.4: **Lack of Cas9 is impaired with significant changes in key virulence factors and pathogenic functionalities of GAS.** GAS WT and Δ*cas9* strains were assessed for **(A)** Capsule expression by ELISA, **(B)** Quantification of M protein-anchored to the cell wall by flow cytometry, **(C)** SpeB protease activity by azocasein assay, **(D)** Capacity to adhere to HaCaT human skin keratinocytes, **(E)** β-hemolysis on blood-agar media, and **(F)** ability to lyse human red-blood cells. Isogenic GAS mutant strains in capsule (Δ*hasA*), M protein (Δ*emm*) SpeB (Δ*speB*), and SLO (Δ*slo*) were used as negative controls in **(A–C,F)**, respectively. For each experiment, samples were assayed at least in triplicate. Data in **(A–D,F)** are plotted as the mean ± SEM, pooled from three independent experiments and analyzed by Student's t test.

Our proteomic analysis and phenotypic assays suggested a Cas9-associated control over

some of the key virulence determinants of GAS. To explore the cumulative effect of these changes, we first compared the capacity of WT vs. the Δ*cas9* GAS strains to proliferate in human whole blood *ex vivo* and found attenuated growth in the Δ*cas9* mutant (Figure 3.4A). Moving further to an *in vivo* murine infection model, we followed the development of necrotic skin ulcers following subcutaneous challenge of mice with WT vs. Δ*cas9* GAS strains. Twenty-four hours post-infection, lesions were significantly larger in WT GAS-infected mice compared to those challenged with the Δ*cas9* mutant (Figure 3.4B). When the mice were euthanized and lesions harvested for colony forming units (CFUs) enumeration 48 h after infection, a significantly higher amount of WT GAS bacteria were recovered compared to Δ*cas9* mutant bacteria (Figure 3.4C). In summary, these data suggest that Cas9 plays an important role during GAS infection in a mammal host, and this effect might reflect influences of the nuclease on several different virulence phenotypes and virulence-related regulatory factors.

3.5 Discussion

The discovery and molecular characterization of RNA-programmable Cas9 nuclease emerged from basic research on the type II CRISPR-Cas system from GAS and has provided a revolutionary biotechnological tool for genome engineering, with promising potential to develop novel strategies to fight and cure many diseases[142]. Despite the attention that GAS Cas9 has received and the major health problem that GAS infections continue to exert on the public health, the native biological role of Cas9 and its contribution for GAS pathogenesis has yet to be reported. In this study, we provide initial experimental evidence that Cas9 has a significant effect on GAS virulence associated phenotypes *in vitro* and *in vivo*. These effects expand the biological significance of GAS Cas9 beyond its well-known role as the key component of the adaptive immune system that can precisely recognize and target foreign DNA.

Deletion of *cas9* in GAS did not affect growth kinetics nor gross morphology of the

bacterium, consistent with observations upon loss of Cas9 orthologs in other organisms such as *F. novicidia*, *N. meningitidis* and GBS, where cell viability of Cas9 deficient strains was also not impacted [152, 106]. To date, the evidence suggests that Cas9 is not involved in the control of essential gene products.

Cas9 RNA transcripts and protein levels were independent of GAS growth phase in bacteriological media, consistent with constitutive Cas9 expression; however, our experiments cannot exclude the possibility that changes in Cas9 expression occur *in vivo* during encounters of the pathogen with host factors. Our *in vitro* experiments demonstrate that cas9 deletion is associated with a significant reduction in the GAS capacity to adhere to epithelial cells. Moreover, our *ex vivo* experiments in human whole blood infected with live WT or Cas9-deficient GAS bacterial strains show a significant contribution of endogenous Cas9 expression to bacterial growth, a key feature of the pathogen for dissemination within the host. During murine necrotic skin infection, absence of Cas9 was linked to diminished size of necrotic skin ulcers and reduced bacterial load within the harvested wounds.

Our proteomic studies suggest that the virulence phenotypes displayed upon loss of Cas9 are mediated in part by control over the abundance of at least four proteins related with GAS adhesion and other thirteen proteins associated with GAS virulence and defense, including pilus structural components, adhesins, key proteins that mediate resistance to reactive oxygen species, immunogenic secreted products, complement inhibitory factors and toxins. Additionally, our functional experiments show a reduction in M protein on the bacterial surface, which may be reflected by enhanced SpeB proteolytic activity in the Δ *cas9* mutant. Reduced SpeB activity is observed in the GAS transition to systemic infection [5, 27], and normal Cas9 function may be required for this functional shift.

Group A Streptococcus Cas9-mediated control of key virulence determinants finds a parallel in Cas9-mediated regulation of BLP inẽxtitF. novicida, which enables the pathogen to dampen TLR2-dependent inflammatory response and to survive within host cells [152]. Also,

the observed broad effect of Cas9 on GAS virulence regulation is reminiscent of the multiple virulence pathways regulated by the CRISPR-Cas9 system of *C. jejuni*, including those encoding lipoproteins, flagella, and chemotaxis-related factors [156].

Absence of GAS Cas9 was associated with differential abundance in several virulence-related transcriptional regulatory factors, including enhanced levels of well-known transcriptional repressors of virulence. Conversely, abundance of important activators of virulence was diminished as a consequence of Cas9 deficiency, suggesting that Cas9 mediates a coordinated balance for the expression of the virulence machinery of GAS, including two of the most important and best studied GAS global regulators (e.g., the TCS CovR/CovS and the transcriptional regulator Mga). Similar evidence of Cas9 regulatory effect over virulence-related transcriptional regulators have been seen in GBS, where the nuclease influenced transcriptional regulator RegR, the modulator of hyaluronidase activity, a key virulence factor involved in GBS blood-brain barrier invasion during meningitis [106]. Based on our experimental evidences, we present a schematic for Cas9-mediated virulence control (Figure 3.5), in which Cas9 could control the expression of several GAS virulence determinants, both directly or indirectly through its regulatory effect on the expression of key transcriptional regulators of virulence.

Since deletion of Cas9 is associated with significant changes in the abundance of more than 40% (n = 745) of GAS products encoded by genes dispersed throughout the genome, including those involved in diverse cellular processes such as stress response, protein metabolism, gene regulation and pathogenesis, among other functions, our studies suggest that Cas9 is a global regulator of GAS virulence and physiology.

Considering the highly specific endonuclease activity of Cas9, one potential mechanism underlying Cas9 effects on virulence regulation is that the nuclease may complex with the tracrRNA encoded immediately downstream of Cas9 in the GAS genome (Figure 3.2A). In that manner, Cas9 could interact with operator regions of genes encoding the virulence determinants observed to be affected by loss of Cas9, leading to degradation or alteration of the corresponding

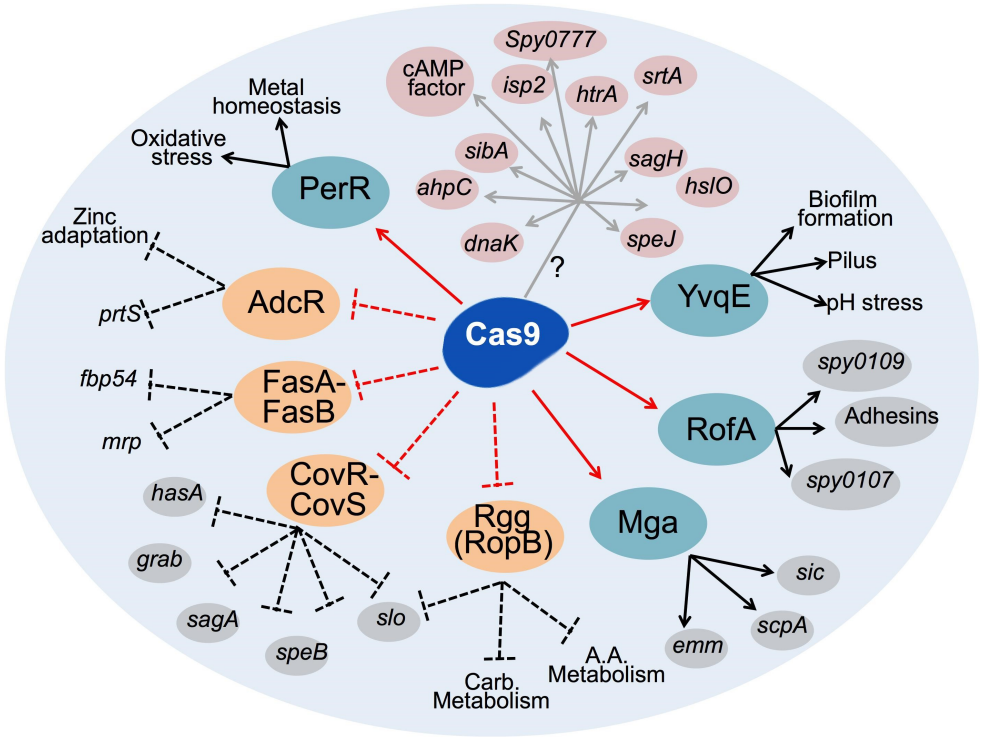

Figure 3.5: **Schematic representation of the network of GAS regulatory proteins and virulence factors affected upon loss of Cas9.** Cas9 may directly (gray arrows) activate expression of several GAS virulence determinants (highlighted in pink) through an unknown molecular mechanism (marked with "?"). Cas9 also upregulates (solid red arrows) some transcriptional activators of virulence (highlighted in teal), further augmenting the expression of numerous virulence factors (black arrows). Conversely, Cas9 negatively controls (dashed red lines) the expression of several transcriptional repressors of virulence (highlighted in orange). Consequently, Cas9 blocks the repressor role (dashed black lines) of these regulators on the expression of key GAS virulence factors. Virulence determinants downstream of transcription factors experimentally confirmed in this study are highlighted in gray. Other virulence factors or functions known from previous studies to be regulated by the highlighted repressors and activators of GAS virulence are depicted with black text only. "Carb. Metabolism": carbohydrates metabolism, "A.A. metabolism": amino acids metabolism.

transcripts. Further studies are required to address this hypothesis or other potential molecular mechanisms of Cas9-mediated regulation on GAS pathogenesis. Elucidating these mechanisms will help in understanding whether these new findings are associated with canonical functions of GAS CRISPR-Cas9 system, such as adaptive immunity mediated through spacer acquisition, and whether they can have direct impact on horizontal gene transfer with consequences on GAS evolution and ecology.

3.6 Materials and Methods

3.6.1 Bacterial Strains and Culture Conditions

GAS M1T1 5448 was originally isolated from a patient with necrotizing fasciitis and streptococcal toxic shock syndrome [85]. All GAS strains were routinely propagated at 37°C on Todd-Hewitt agar (THA, Difco) or in static liquid Todd-Hewitt broth (THB).

3.6.2 Genetic Manipulation of GAS (Construction of Δ*cas9* Strain and Δ*cas9* Complementation)

Precise in-frame allelic replacement of the cas9 gene was performed using established methodology [136]. We first generated PCR products immediately up and downstream of the cas9 gene. 1000 bp upstream was amplified with primers cas9upFw (5′-ccgctcgagtcctgtggagcttagtaggttt agcaagatggcagc-3′) and cas9upRv (5′-tatccagtgatttttttctccatttttgcctcctaaaataaaaagtttaaattaaatcca-3′). Subsequently, 1060 bp of sequence downstream of cas9 was amplified with primers cas9downFw (5′-tactgcgatgagtggcagggcggggcgtaatggctggttggcgtactgttgtggt-3′) and cas9downRv (5′-cccaagcttgacctgcattgattggatgctccaaatctcttgag-3′). Primers cas9upFw and cas9downRv were designed with 25 bp 5′ extensions corresponding to the 5′ and 3′ ends of the chloramphenicol acetyltransferase (*cat*) gene, respectively. The 660 bp fragment corresponding to the *cat* gene

was PCR amplified from a previous GAS allelic replacement strain with primers catFw (5′-atggagaaaaaaatcactggatatacc-3′) and catRv (5′-ttacgccccgccctgccactcatcgca-3′). The upstream, downstream and full *cat* fragments were assembled in a second round of PCR using primers cas9upFw and cas9downRv. The resultant PCR amplicon was subcloned into temperature-sensitive vector pHY304, and allelic exchange mutagenesis in GAS 5448 was performed following double crossover as described previously to generate the stable mutant 5448Δ*cas9* strain.

Attempt of Δ*cas9* complementation was carried out by amplification of *cas9* gene with primers cas9Fw (5′-tccccccgggtgaaggaggcaaaaatggataagaaatactcaataggcttaga-3′) and cas9Rv (5′-gctctagatcagtcacctcctagctgactcaaatcaatgc-3′). The resulting PCR product was cloned into TOPO-XL vector (Invitrogen), subcloned into the multicopy plasmid pDCerm and transformed into the Δ*cas9* strain. The Δ*cas9* strain complemented showed a significant defect in cell growth (Supplementary Figure A.1), confounding results and therefore excluded from study.

3.6.3 RNA Extraction and qPCR Assays

Total RNA was isolated from GAS bacteria pellets using RNEasy isolation kit (Qiagen), with an additional bead-beating step with 1.0 mm glass beads (Sigma). Synthesis of total cDNA was performed using iScript cDNA Synthesis kit (Bio-Rad). For each sample to be retrotranscribed, an exact amount of 1 µg of RNA was used as template. Real-time PCR assays were conducted in a CFX96 Real-Time System (Bio-Rad). Three biological and three technical replicates were analyzed for each sample. 20 µl reactions contained 1 µl cDNA, 10 µl SYBR fast qPCR Master Mix (KAPA Biosystems) and 0.25 µM of each target-specific primer. Primer pairs cas9qPCRFw (5′-aaatacagaccgccacagtatc-3′) and cas9qPCRRv (5′-tcttccgacgtgtataccttcta-3′), cas1qPCRFw (5′-acgccaattggttgaaactc-3′) and cas1qPCRRv (5′-acgacggcatttagatacgc-3′), cas2q PCRFw (5′-ttgatatgccgacggacac-3′) and cas2qPCRRv (5′-aaaagcctccccaagaaatac-3′), csn2qPCR Fw (5′-ggcggtacaattcttgtgct-3′) and csn2qPCRRv (5′-cgatttcacttcgggtttct-3′) and gyrAqPCR-Fw (5′-gaagtgatccctggacctga-3′) and gyrAqPCR-Rv (5′-cccgacctgtttgagttgtt-3′) were used to amplify

transcripts from the *cas9*, *cas1*, *cas2*, *csn2*, and *gyrA* (encoding the DNA gyrase subunit A and used as an internal control to normalize the sample data) genes, respectively. Amplifications were carried out with 1 denaturation cycle (95°C for 5 min), followed by 45 cycles of amplification (95°C for 10 s; 60°C for 10 s; 72°C for 10 s). After amplification, melting curves were generated to confirm amplification of a single product. Relative *cas9* mRNA transcript levels were determined using ΔCt method and normalized with the mRNA transcript levels of *gyrA* housekeeping gene.

3.6.4 Western Immunoblot Assays

Cells from GAS cultures were pelleted and lysed following enzymatic digestion with mutanolysin and lysozyme as previously described[77]. Cell lysates were quantified for protein content by Pierce BCA Assay, and equal amounts of each sample analyzed were separated by SDS-PAGE, immunoblotted and visualized with Supersignal® WestPico Chemiluminescent Substrate (Thermo Fisher Scientific) and CL-XposureTM Film (Thermo Fisher Scientific). The following antibodies were used for immunoblotting: anti-CRISPR-Cas9 (ab204448; abcam), ECLTM anti-Rabbit IgG peroxidase-conjugated (NA934V; GE Healthcare).

3.6.5 Proteomics Sample Preparation

Three independent cultures of mid-exponential growth phase GAS cells were pelleted and resuspended in lysis buffer containing 50 mM HEPES, 3% sodium dodecyl sulfate (SDS, Fisher), 75 mM NaCl (Sigma), 1 mM NaF (Sigma), 1 mM beta-glycerophosphate (Sigma), 1 mM sodium orthovanadate (Sigma), 10 mM sodium pyrophosphate (Sigma), 1 mM phenylmethylsulfonyl fluoride (PMSF, Sigma) and 1X cOmplete mini EDTA-free protease inhibitor cocktail tablet (Roche). Bacterial homogenates were sonicated to ensure complete lysis. Subsequent sample preparation, including 10-plex tandem mass tag (TMT) labeling, was performed as previously described[98].

3.6.6 Quantitative Proteomics, Protein Identification, and Analysis

Resulting ".raw" MS data files were processed using Proteome Discoverer 2.1 (Thermo Fisher). MS2 spectra were searched against a protein database derived from GAS strain MGAS5005 genome (GenBank: CP000017.2). Mass tolerances of 50 ppm and 0.6 Da were used for MS1 and MS2 spectra, respectively. Search parameters included full digest by trypsin with a maximum of two missed cleavages per peptide, static modifications of TMT 10-plex reagents on lysines and peptide N-termini (+ 229.162932 Da) and carbamidomethylation of cysteines (+ 57.02146 Da), and variable oxidation of methionine (+ 15.99492 Da). Results were filtered to a 1% false discovery rate using a target-decoy strategy at both the peptide and protein level.

For quantitative analysis, reporter ion intensities for the TMT reagents were extracted from MS3 spectra. Only peptide spectral matches exceeding an average signal:noise greater than 10 and an isolation interference less than 25% were retained for downstream analysis. Data were normalized as previously described[97]. Briefly, the reporter ion value for each peptide was summed to the protein level. The summed values were first normalized to the bridge channel value for each protein then to median of the entire bridge channel. To account for differences in amount of peptide labeled, the quantitative information was then normalized to the median of the entire dataset and reported as the normalized, summed signal:noise ratios per protein, per sample. Datasets and corresponding annotated spectra are available through ProteomeXchange (PXD012568).

To determine proteins of significantly difference abundances, a F-test was first used to compare the variances of each protein in each condition. If the variances were equal, a standard Student's t-test was performed, but if the variances were unequal, Welch's correction was included. The GAS MGAS5005 genome was annotated in the RAST database in order to systematically organize genes into categories, subcategories, and subsystems. Differentially abundant proteins with P values < 0.05 were first identified. Normalized intensities for significant genes were summed per each RAST subcategory. Subcategories with greater than two-fold

change were plotted as the log2 of the WT to Δ*cas9* ratio ($Log_2 > 1$ or $Log_2 < -1$). Significantly changing proteins were also ranked using pi score as previously described[185], considering all differential proteins with level $\alpha < 0.001$ as significant. The "Virulence, disease and defense" and "Regulation of virulence" RAST subcategories were manually curated in order to ensure a comprehensive account of known virulence determinants and virulence regulatory proteins reported in the literature. All manually annotated proteins were highlighted as ($*$) in Figures 3.2B, D and Supplementary Table A.3. Python 2.7 was used to plot and analyze data. The data and relevant code are available upon request.

3.6.7 Hyaluronic Acid Capsule Assays

Hyaluronic acid extraction was performed as previously described[76]. Briefly, pellets from 5 mL GAS cells grown to mid-logarithmic were resuspended in 500 µl deionized water. In a 2 mL screw-cap tube, 400 µl of cells suspension and 1mL chloroform (Sigma-Aldrich) were combined and vortexed at maximum speed for 10 min. Samples were then centrifuged at 13,000 $\times$ g for 10min. The resulting aqueous phase was collected and diluted 1:40 for Quantikine hyaluronan ELISA (DHYAL0; R&D Systems) for quantification according to manufacturer's instructions.

3.6.8 Quantification of Cell Wall-Attached M Protein by Flow Cytometry

Mid-exponential phase GAS cells were probed for surface-attached M protein as previously described [174]. Briefly, bacteria pellets were washed with PBS before incubation with mouse anti-serum raised against GAS M1 protein. Samples were subjected to additional PBS wash steps to remove excess of antibodies, following incubation with goat anti-mouse IgG (H+L) AlexaFluor 488 secondary antibodies (Thermo Scientific). Samples were run on BD FACS Canto II without fixation and analyzed on FlowJo X7 (TreeStar).

3.6.9 Cysteine Protease Activity Assays

SpeB protease activity was determined as previously described[29]. Briefly, overnight GAS cultures were diluted 1:50 into fresh THB media and cultured for 17 h at 37°C to early stationary phase. Cultures were centrifuged at 3,200 × g, and supernatants were filter sterilized. Equal volumes of filtered supernatant and activation buffer (1 mM EDTA, 100 mM sodium acetate, and 20 mM freshly prepared DTT) were mixed and incubated at 40°C for 30 min. 2% azocasein (Sigma) was dissolved in activation buffer and added to the activated supernatant in a 1:1 (v/v) ratio. The mixture was then incubated 1 additional h at 40°C. Excess azocasein was precipitated with the addition of trichloroacetic acid (Sigma) to a final concentration of 15% (w/v) and centrifuged for removal. Supernatants were transferred into a 96-well plate. Absorbance was measured at 366nm and normalized to wild-type levels to determine relative protease activity.

3.6.10 Hemolysis Assays

Red blood cells (RBC) were prepared for hemolysis from whole blood drawn from healthy volunteer donors. Hirudin tubes-containing whole blood were left to settle naturally at room temperature on the bench for 1 h before washing the RBC pellet with PBS. The fraction of RBC was resuspended in PBS to 2% (v/v). Equal volumes of mid-logarithmic growth phase bacteria and 2% RBC suspension were mixed in V-bottom 96 well plates. After incubation for 1 h at 37°C, plates were spun down to pellet intact RBC. Supernatant was transferred to a new plate for OD measure at 450 nm. Percent of lysis was calculated using PBS as a negative control (0% lysis) and 0.025% Triton X100 as a positive control (100% lysis). For blood agar hemolysis, 10 µl of mid-logarithmic phase bacterial cultures were spotted on to 5% sheep blood in tryptic soy agar base (Hardy Diagnostics A10). Plates were incubated overnight at 37°C, imaged with ruler using Gel Doc XR+ gel documentation system (Bio-Rad) and resulting zones of hemolysis were quantified for average radius using FIJI[155].

3.6.11 Whole Blood Assays

1×10^6 colony forming units (CFU) from mid-exponential growth phase GAS were resuspended in 20 µl PBS and mixed with 80 µl of whole blood. Samples were incubated at 37°C with rotation. After 2 h of incubation, 10 µl aliquots were diluted, plated on THA plates and incubated overnight for CFU enumeration. Blood was drawn from healthy volunteer donors into hirudin tubes by trained phlebotomists following a protocol for simple phlebotomy approved by the UCSD Institutional Review Board/Human Research Protection Program. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

3.6.12 Adherence Assays

Adherence assays were performed as described previously[166] but using the HaCaT human skin keratinocyte cell line. HaCaT cells were obtained from ATCC and propagated as monolayer in RPMI 1640 medium + 10% fetal bovine serum (FBS). For assays, cells were plated at 5×10^5 cells/well in 24 well plates. Immediately prior to assay, the culture media on the HaCaT cells was replaced with fresh RPMI 1640 + 2% FBS. GAS strains grown to the mid-exponential phase were resuspended in RPMI 1640 + 2% FBS and added to the HaCaT cells at a multiplicity of infection (MOI) of 10. Plates were centrifuged at 800 g $\times$ 5 min to ensure GAS-HaCaT contact. Infected cells were incubated at 37°C with 5% CO_2 for 30 min then lysed with trypsin and 0.025% Triton X100, serially diluted, and plated onto THA plates for CFU enumeration.

3.6.13 Animal Experiments

The UCSD Institutional Animal Care and Use Committee approved all animal use and procedures. In compliance with ethical guidelines, to minimize the number of animals, we used a minimum of five mice for each experimental group (except where indicated in the figure legends) to ensure statistical power. All mice were randomly distributed into the different groups as

indicated in the corresponding figure legend. 8- to 10-week-old C57BL/6 mice were infected subcutaneously with 1×10^8 CFUs of either GAS wild type or Δ*cas9* strains resuspended in 100 µl of PBS. Lesions were imaged daily and surface area quantified using ImageJ software. At 48 h post-infection, lesions were excised, homogenized, and plated as dilutions onto THA plates for enumeration of bacterial CFU.

3.6.14 Statistical Analysis

The data were collected from three independent experiments in triplicate, unless otherwise indicated. Data were combined and represented as mean ± standard error of the mean. Results were either analyzed by unpaired Student's t-test or by two-way ANOVA using GraphPad Prism version 7. P values were summarized for respective analyses as: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***).

3.7 Acknowledgements

This chapter, in full, is a reprint of material as it appears in *Frontiers in Microbiology* 2019 by **Nina J. Gao**, Mahmoud M. Al-Bassam, Saugat Poudel, Jacob M. Wozniak, David J. Gonzalez, Joshua Olson, Karsten Zengler, Victor Nizet, and J. Andrés Valderrama. *Functional and Proteomic Analysis of Streptococcus pyogenes Virulence Upon Loss of Its Native Cas9 Nuclease*. DOI: 10.3389/fmicb.2019.01967. The dissertation was the primary investigator and author of this paper.

The authors thank Samira Dahesh and Satoshi Uchiyama from the Nizet laboratory for providing protocols and advice on GAS genetic manipulation. They also thank Stephanie Brandt for her expertise and assistance with the mouse infection experiments.

Work published in this research article was funded by grants from the National Institutes of Health (NIH) / National Institute of Allergy and Infectious Disease (NIAID), National Institute

for General Medical Sciences (NIGMS), and National Institute Arthritis and Musculoskeletal and Skin Diseases (NIAMS).

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Chapter 4

Conclusions

4.1 Conclusions

4.2 Summary of results

My dissertation opens with Chapter 1 as a review of current literature on the group A carbohydrate. Bacterial whole genome sequencing technology and the search for GAS vaccine candidates led Nina van Sorge *et al.* to identify the *gac* operon encoding the biosynthetic enzymes required for GAC [174]. This was a hallmark paper that led to renewed interest in the GAC. Genetic manipulations of the GAC allowed testing of functional roles beyond cell wall rigidity and garned increased potential as a candidate vaccine antigen. The disastrous early clinical trials for a GAS vaccine in the 1960's immediately deterred GAC research, emphasizing the etiological associations with rheumatic heart disease and Syndenham's chorea. However, with increasing incidence of severe, invasive GAS disease and increasing global burden of GAS disease (especially in at-risk populations such as the Australian Aborigines), GAS vaccination research resumed with special attention to safety and lack of reactivity to human self antigens. This chapter is, to my knowledge, the first to compile GAC-relevant research into a single review article.

Chapter 2 uses the currently knowledge on GAC to devise a site-specific glyo-conjugate strategy for GAS vaccine development. Here, we use a modified form of the GAC as the polyrhamnose core without the human autoreactive epitope of N-acetylglucosamine (referred to as GAC^{PR}) conjugated to a carrier protein at precise, site-specific attachment sites to retain native protein immunogenic epitopes. This strategy avoids the potentially-overused conjugate proteins such as tetanus toxoid and CRM197 and instead allows the use of yet another GAS antigen (in this study, the GAS adherence and division protein SpyAD) as a multivalent combination with reduced component complexity. We use antigen ELISA and bacterial surface IgG-binding flow cytometry to confirm induction of antigenicity of antigens with no loss of antibody titer in the conjugate compared to the carrier protein alone. Furthermore, the surface binding assay

confirms that vaccination-induced antibodies recognize native surface antigens across many different serotypes of GAS. Primary human neutrophil opsonophagocytosis assays using the same rabbit antisera demonstrate the functional advantage of the conjugate with comparable, if not superior, opsonophagocytic killing of GAS strains of different serotypes when compared to the SpyAD[4pAMF] carrier protein antisera. This will allow for the addition of the highly conserved GAC antigen to be used in conjunction with conserved protein antigens as a multi-valent GAS vaccine combination with reduced complexity: two GAS-derived antigens may be combined into one component while retaining full immunogenic potential of both antigens.

To create a broadly protective multi-valent GAS vaccine, we employ the SpyAD[4pMF]-GAC^{PR} glyco-conjugate with 2 more highly conserved protein GAS antigens: streptolysin O (SLO) and C5a peptidase. All individual components show great vaccine potential due to broad protection, but combined they show superior anti-streptococcal results against multiple GAS serotypes in human neutrophil opsonophagocytic killing assay, murine lethal systemic challenge, and murine skin infection challenge models of disease. As an exciting bonus, we find that this vaccine formulation is also capable of inducing protection against live group B *Streptococcus* (GBS) infections, yet another major cause of morbidity and mortality around the world, for mothers, neonates, and the elderly who are particularly at risk.

Finally in chapter 3, we examined the potential role of Cas9 virulence. We generated an isogenic knockout mutation of the *cas9* gene and examined changes in the proteome profile changes. Loss of endogenous Cas9 led to broad down-regulation of known virulence proteins. To confirm functional effects of the loss of Cas9, we quantified several known virulence factors include the hyaluronan capsule, SpeB protease, surface bound M protein, and hemolysis of the wild-type and Δ*cas9* GAS strains to confirm reduced virulence. Finally, we used the GAS strains in in human blood and a mouse skin infection model to confirm that mutant Δ*cas9* GAS strain is less capable than the isogenic wild-type GAS strain at evading host immune responses as a successful pathogen. This data corroborates other studies examining the role of Cas9 in bacterial

pathogenesis and regulation of virulence.

4.3 The importance of revisiting basic science

Most of my doctoral work was based on scientific findings from a century ago that has recently been brought into the spotlight and re-examined with greater purpose. For both the GAC and Cas9, these GAS components were assumed to play a basic role of bacterial wall stability and viral defense, respectively, for nearly a century after the initial finding. For the group A carbohydrate, the growing need for a safe, effective GAS vaccine drives research for an universal antigen. The research is only possible now due to major advances in scientific knowledge and molecular tools that allow us to test the questions in ways previously thought impossible. In the Cas9 project, though we were unable to definitely confirm the role nor molecular mechanism, we found that without endogenous Cas9, the GAS bacterium has impaired virulence with reduced expression of protein virulence factors, reduced survival in whole human blood, and reduced lesion size when subcutaneously injected into mice. This study corroborates findings of many existing studies of bacterial pathogens and the role played by Cas9 in regulation of virulence, expanding the long-standing assumption that the Cas9 solely functioned as a means of anti-viral defense in the type II CRISPR/Cas system.

My dissertation work exemplifies the importance of revisiting basic science; though both the GAC and Cas9 have been found in GAS and accepted for a single purpose. However, advances in knowledge and scientific tools and revisitation of these molecules have expanded their purpose beyond what their original functions. GAC now presents as a strong contender as an universal vaccine antigen, which Cas9 has sparked a gene-editing revolution with the diagnostic and therapeutic applications.

4.4 Potential impact of my dissertation work

When I began my doctoral research, I knew from my brief rotation with Dr. Victor Nizet's laboratory that group A *Streptococcus* would be the pathogen I wanted to study. A bacterial pathogen, best known as the cause of a familiar childhood pharyngitis, actually holds a great potential for human morbidity and mortality on a global scale. The implications for public health and clinical disease for GAS were a strong motivators for me to pursue research projects on GAS virulence and immune responses. I am grateful to Global Health Institute's graduate researcher award for allowing me to pursue research with a global perspective, including support for a 3-month research trip abroad with the laboratory of Professor Mark Walker at the University of Queensland in Australia. It's important to recognize and remember that though GAS infections are usually seen as "just strep throat" here, many populations around the world are at higher risk for severe invasive disease or greater mortality rates due to post-streptococcal autoimmune sequelae.

I am grateful for the opportunity to work on development of a safe, universal GAS vaccine. Our results are promising that our multi-valent formulation provides broad protection across different serotypes and disease presentations with no signs of inducing cross-reactive antibodies that may drive rheumatic heart disease. I am excited with our results, but there is still much to do before the project yields a GAS vaccine. Our work focused on examining the humoral response as a metric for vaccine immunogenicity, though I am curious about the cellular immunity induced by our glyco-conjugate. Recent studies propose the ability of glyco-conjugates to induce T cells responsive to carbohydrate antigens [3]. Further research into the contributions of induced cellular-mediated immunity, as well as contributions of individual components within the vaccine, would aid in justification of this formulation as an effective GAS vaccine. Furthermore, understanding the protective response (humoral and cellular) to GAS infections, even with the caveat of imperfect infection models in mice, may allow for optimal choice of vaccine adjuvant

to improve vaccine efficacy, as demonstrated by Rivera-Hernandez *et al.* [143].

A critical gap in knowledge remains for the molecular pathogenesis of rheumatic heart disease (RHD). GAS vaccine safety hinges on the evaluating the potential of the vaccine to drive an auto-immune response. Though our vaccine formulation selects antigens with consideration for potential autoreactive epitopes such as the GlcNAc, research into key RHD epitopes based on convalescent serum from RHD disease patients compared to healthy controls, and in comparison to GAS antigens, may yield insight into dangerous pathogenic epitopes to avoid in vaccine design. Identification of RHD disease epitopes will ensure safe vaccine design and hopefully propel an universal GAS vaccine safely through clinical trials and to market.

Appendix A

Supplemental Information for Chapter 3

A.1 Supplementary Figures

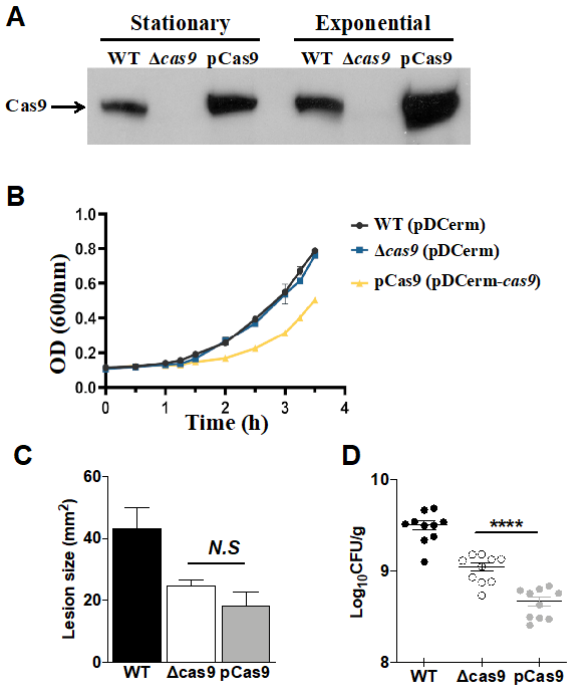

Figure A.1: **Genetic complementation of Δ*cas9* GAS M1T1 5448 strain leads to marked growth defect. (A)** Western blot analysis of Cas9 protein expression in wild type (WT), Cas9-deficient (Δ*cas9*), and *cas9*-plasmid complemented (pCas9) GAS strains from cells frown at stationary or exponential growth phases. **(B)** Cellular growth curves of WT, Δ*cas9*, and pCas9 GAS strains grown in THB media at 37°C. GAS strains in **(B)** carried either empty plasmid (pDCerm) or the *cas9*-expressing plasmid (pDCerm-*cas9*). **(C and D)** Subcutaneous infection of C57BL/6 mice with WT, Δ*cas9*, or Cas9-complemented (pCas9) strains of GAS. Average lesion sizes **(C)** and enumeration of CFU recovered from excised lesions **(D)** at 48 hours post-infection. Data in **(C and D)** are plotted as the mean ± SEM and represent data combined from two independent experiments, analyzed by Student's t-test. *N.S.* = not significant ($p > 0.05$), **** $P < 0.001$.

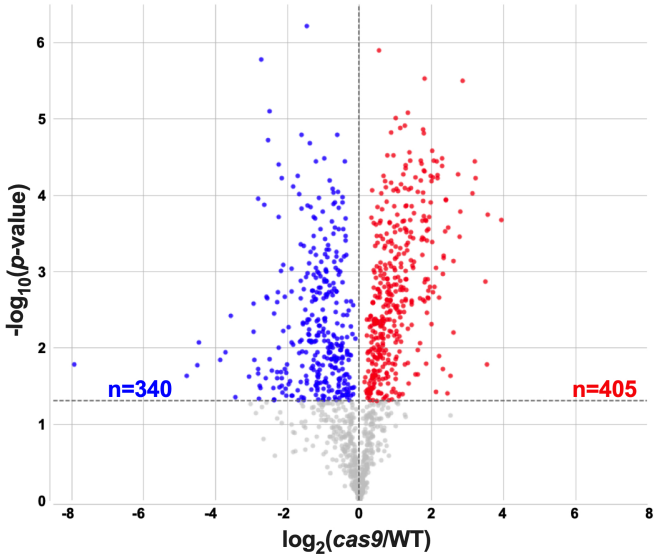

Figure A.2: **Abundance of proteins detected by tandem mass tag proteomics in wild type (WT) and Δ*cas9* GAS M1T1 5448 strains.** Volcano plot shows proteins significantly more abundant in the wild type (blue dots, n = 340) or in the Δ*cas9* (red dots, n = 405) GAS M1T1 5448 strains. Significant proteins were initially identified by Student's t-test with $p < 0.05$.

A.2 Supplementary Tables

Table A.1: **Minimum inhibitory concentration (MIC) of cell wall synthesis inhibitory antibiotics against wild type (WT) or Δ*cas9* GAS M1T1 5448 strains.** MIC determined by microbroth dilution in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Todd-Hewitt broth.

Antibiotic	wild type (WT)	Δ*cas9*
penicillin G, μg/ml	0.015-0.031	0.015-0.031
vancomycin, μg/ml	1.0-2.0	1.0-2.0

Table A.2: **A2_TMTproteomics.xlsx Normalized abundance of all proteins detected from wild-type and Δ*cas9* GAS 5448 strains by tandem mass tag proteomics.** Peptide reads were searched against database derived from the MGAS5005 M1 serotype *S. pyogenes* genome (GenBank: CP000017.2) Columns include: GenBank protein ID, protein description, and full-raw data from three independent biological samples corresponding to each protein identified. F test, *P* value and pi score statistical analyses are also listed in the corresponding column. #, $p <$ 0.05 (protein abundance deemed statistically significant by Student's t-test). *, alpha level $<$ 0.05 (protein abundance deemed statistically significant by pi score).

Table A.3: **A3_RAST_significant.xlsx Normalized abundance of proteins that were identified to be significantly enriched either in wild type (WT) or Δ*cas9* GAS 5448 strains.** Significant proteins were grouped into subcategories using RAST. out of the 745 significantly differentially quantified proteins between both strains, 474 proteins had annotations by RAST server (RAST annotation +). Some proteins annotated by RAST in different categories (RAST annotation +/-) or not annotated at all (RAST annotation -) were manually added into the corresponding functional category based on reported literature. Columns include: GenBank protein ID, corresponding gene name, protein description, RAST category and sub-category, as well as full-raw data from three independent biological samples corresponding to each protein identified as significant. F test, *P* value and pi score statistical analyses are also listed in the corresponding column. #, $p < 0.05$ (protein abundance deemed statistically significant by Student's t-test). *, alpha level < 0.05 (protein abundance deemed statistically significant by pi score).

Bibliography

[1] Elisabeth E. Adderson, Alexander R. Shikhman, Kent E. Ward, and Madeleine W. Cunningham. Molecular Analysis of Polyreactive Monoclonal Antibodies from Rheumatic Carditis: Human Anti-N-Acetylglucosamine/Anti-Myosin Antibody V Region Genes. *The Journal of Immunology*, 161(4):2020–2031, 1998.

[2] C. D. Ashbaugh, H. B. Warren, V. J. Carey, and M. R. Wessels. Molecular analysis of the role of the group A streptococcal cysteine protease, hyaluronic acid capsule, and M protein in a murine model of human invasive soft-tissue infection. *J Clin Invest*, 102(3):550–60, 1998.

[3] Fikri Avci, Francesco Berti, Peter Dull, John Hennessey, Viliam Pavliak, A. Krishna Prasad, Willie Vann, Michael Wacker, and Olivier Marcq. Glycoconjugates: What it would take to master these well-known yet little-understood immunogens for vaccine development. *mSphere*, 4(5):e00520–19, 2019.

[4] R. K. Aziz, D. Bartels, A. A. Best, M. DeJongh, T. Disz, R. A. Edwards, K. Formsma, S. Gerdes, E. M. Glass, M. Kubal, F. Meyer, G. J. Olsen, R. Olson, A. L. Osterman, R. A. Overbeek, L. K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G. D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke, and O. Zagnitko. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, 9:75, 2008.

[5] R. K. Aziz, M. J. Pabst, A. Jeng, R. Kansal, D. E. Low, V. Nizet, and M. Kotb. Invasive M1T1 group A Streptococcus undergoes a phase-shift *in vivo* to prevent proteolytic degradation of multiple virulence factors by SpeB. *Molecular Microbiology*, 51(1):123–134, 2003.

[6] S. Banerjee and C. Ford. *CADTH Rapid Response Reports*. Canadian Agency for Drugs and Technologies in Health., Ottawa (ON), 2018.

[7] Timothy C. Barnett and June R. Scott. Differential Recognition of Surface Proteins in Streptococcus pyogenes by Two Sortase Gene Homologs. *Journal of Bacteriology*, 184(8):2181–2191, 2002.

[8] R. Barrangou, C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, 315(5819):1709–1712, 2007.

[9] Giuliano Bensi, Marirosa Mora, Giovanna Tuscano, Massimiliano Biagini, Emiliano Chiarot, Mauro Bombaci, Sabrina Capo, Fabiana Falugi, Andrea G. O. Manetti, Paolo Donato, Erwin Swennen, Marilena Gallotta, Manuela Garibaldi, Vittoria Pinto, Nico Chiappini, James M. Musser, Robert Janulczyk, Massimo Mariani, Maria Scarselli, John L. Telford, Renata Grifantini, Nathalie Norais, Immaculada Margarit, and Guido Grandi. Multi high-throughput approach for highly selective identification of vaccine candidates: the group a streptococcus case. *Molecular & Cellular Proteomics*, 11(6), 2012.

[10] Shuai Bi, Meiyi Xu, Ya Zhou, Xinxin Xing, Adong Shen, and Beinan Wang. A Multi-component Vaccine Provides Immunity against Local and Systemic Infections by Group A Streptococcus across Serotypes. *mBio*, 10(6), 2019.

[11] Alan L. Bisno, Fran A. Rubin, P. Patrick Cleary, and James B. Dale. Prospects for a Group A Streptococcal Vaccine: Rationale, Feasibility, and Obstacles—Report of a National Institute of Allergy and Infectious Diseases Workshop. *Clinical Infectious Diseases*, 41(8):1150–1156, 2005.

[12] P. D. Bright, B. M. Mayosi, and W. J. Martin. An immunological perspective on rheumatic heart disease pathogenesis: more questions than answers. *Heart*, 102(19):1527–32, 2016.

[13] S. Brouwer, T. C. Barnett, T. Rivera-Hernandez, M. Rohde, and M. J. Walker. *Streptococcus pyogenes* adhesion and colonization. *FEBS Lett*, 590(21):3739–3757, 2016.

[14] Stephanie Brown, Jr. Santa Maria, John P., and Suzanne Walker. Wall teichoic acids of gram-positive bacteria. *Annual review of microbiology*, 67:313–336, 2013.

[15] R. J. Burke and C. Chang. Diagnostic criteria of acute rheumatic fever. *Autoimmun Rev*, 13(4-5):503–7, 2014.

[16] D. Burstein, C. L. Sun, C. T. Brown, I. Sharon, K. Anantharaman, A. J. Probst, B. C. Thomas, and J. F. Banfield. Major bacterial lineages are essentially devoid of CRISPR-Cas viral defence systems. *Nature Communications*, 7(1):10613, 2016.

[17] Élise Caliot, Shaynoor Dramsi, Marie-Pierre Chapot-Chartier, Pascal Courtin, Saulius Kulakauskas, Christine Péchoux, Patrick Trieu-Cuot, and Michel-Yves Mistou. Role of the Group B Antigen of Streptococcus agalactiae: A Peptidoglycan-Anchored Polysaccharide Involved in Cell Wall Biogenesis. *PLOS Pathogens*, 8(6):e1002756, 2012.

[18] J. R. Carapetis, A. C. Steer, E. K. Mulholland, and M. Weber. The global burden of group A streptococcal diseases. *The Lancet Infectious Diseases*, 5(11):685–694, 2005.

[19] Jonathan R. Carapetis, Andrea Beaton, Madeleine W. Cunningham, Luiza Guilherme, Ganesan Karthikeyan, Bongani M. Mayosi, Craig Sable, Andrew Steer, Nigel Wilson, Rosemary Wyber, and Liesl Zühlke. Acute rheumatic fever and rheumatic heart disease. *Nature Reviews Disease Primers*, 2(1):15084, 2016.

[20] M. S. Chaussee. Rgg Influences the Expression of Multiple Regulatory Loci To Coregulate Virulence Factor Expression in *Streptococcus pyogenes*. *Infection and Immunity*, 70(2):762–770, 2002.

[21] Q. Cheng, D. Stafslien, S. S. Purushothaman, and P. Cleary. The group b streptococcal c5a peptidase is both a specific protease and an invasin. *Infect Immun*, 70(5):2408–13, 2002.

[22] C. Chetty, D.G. Klapper, and J.H. Schwab. Soluble Peptidoglycan-Polysaccharide Fragments of the Bacterial Cell Wall Induce Acute Inflammation. *Infect Immun*, 38(2):1010–1019, 1982.

[23] Emiliano Chiarot, Cristina Faralla, Nico Chiappini, Giovanna Tuscano, Fabiana Falugi, Gabriella Gambellini, Annarita Taddei, Sabrina Capo, Elena Cartocci, Daniele Veggi, Alessia Corrado, Simona Mangiavacchi, Simona Tavarini, Maria Scarselli, Robert Janulczyk, Guido Grandi, Immaculada Margarit, and Giuliano Bensi. Targeted Amino Acid Substitutions Impair Streptolysin O Toxicity and Group A Streptococcus Virulence. *mBio*, 4(1), 2013.

[24] I. Chmouryguina, A. Suvorov, P. Ferrieri, and P. P. Cleary. Conservation of the c5a peptidase genes in group a and b streptococci. *Infection and immunity*, 64(7):2387–2390, 1996.

[25] P. P. Cleary, Y. V. Matsuka, T. Huynh, H. Lam, and S. B. Olmsted. Immunization with C5a peptidase from either group A or B streptococci enhances clearance of group A streptococci from intranasally infected mice. *Vaccine*, 22(31-32):4332–41, 2004.

[26] P. Patrick Cleary, Usha Prahbu, James B. Dale, Daniel E. Wexler, and Jeffrey Handley. Streptococcal c5a peptidase is a highly specific endopeptidase. *Infection and Immunity*, 60(12):5219–5223, 1992.

[27] J. N. Cole, T. C. Barnett, V. Nizet, and M. J. Walker. Molecular insight into invasive group a streptococcal disease. *Nature Reviews Microbiology*, 9(10):724–736, 2011.

[28] J. N. Cole, J. D. McArthur, F. C. McKay, M. L. Sanderson-Smith, A. J. Cork, M. Ranson, M. Rohde, A. Itzek, H. Sun, D. Ginsburg, M. Kotb, V. Nizet, G. S. Chhatwal, and M. J. Walker. Trigger for group A streptococcal M1T1 invasive disease. *The FASEB Journal*, 20(10):1745–1747, 2006.

[29] M. Collin and A. Olsén. Generation of a mature streptococcal cysteine proteinase is dependent on cell wall-anchored m1 protein. *Mol Microbiol*, 36(6):1306–18, 2000.

[30] L. Cong, F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini, and F. Zhang. Multiplex genome engineering using crispr/cas systems. *Science*, 339(6121):819–823, 2013.

[31] M. W. Cunningham. Pathogenesis of group a streptococcal infections. *Clin Microbiol Rev*, 13(3):470–511, 2000.

[32] M. W. Cunningham. Rheumatic fever, autoimmunity, and molecular mimicry: the streptococcal connection. *Int Rev Immunol*, 33(4):314–29, 2014.

[33] M. W. Cunningham. *Post-Streptococcal Autoimmune Sequelae: Rheumatic Fever and Beyond*. University of Oklahoma Health Sciences Center, Oklahoma City, 2016.

[34] Ron Dagan, Juhani Eskola, Claude Leclerc, and Odile Leroy. Reduced response to multiple vaccines sharing common protein epitopes that are administered simultaneously to infants. *Infection and Immunity*, 66(5):2093–2098, 1998.

[35] Ron Dagan, Jan Poolman, and Claire-Anne Siegrist. Glycoconjugate vaccines and immune interference: A review. *Vaccine*, 28(34):5513–5523, 2010.

[36] J. B. Dale, E. Y. Chiang, D. L. Hasty, and H. S. Courtney. Antibodies against a Synthetic Peptide of SagA Neutralize the Cytolytic Activity of Streptolysin S from Group A Streptococci. *Infection and Immunity*, 70(4):2166–2170, 2002.

[37] J. B. Dale, R. G. Washburn, M. B. Marques, and M. R. Wessels. Hyaluronate capsule and surface M protein in resistance to opsonization of group A streptococci. *Infect Immun*, 64(5):1495–501, 1996.

[38] Mark R. Davies, Liam McIntyre, Ankur Mutreja, Jake A. Lacey, John A. Lees, Rebecca J. Towers, Sebastián Duchêne, Pierre R. Smeesters, Hannah R. Frost, David J. Price, Matthew T. G. Holden, Sophia David, Philip M. Giffard, Kate A. Worthing, Anna C. Seale, James A. Berkley, Simon R. Harris, Tania Rivera-Hernandez, Olga Berking, Amanda J. Cork, Rosângela S. L. A. Torres, Trevor Lithgow, Richard A. Strugnell, Rene Bergmann, Patric Nitsche-Schmitz, Gusharan S. Chhatwal, Stephen D. Bentley, John D. Fraser, Nicole J. Moreland, Jonathan R. Carapetis, Andrew C. Steer, Julian Parkhill, Allan Saul, Deborah A. Williamson, Bart J. Currie, Steven Y. C. Tong, Gordon Dougan, and Mark J. Walker. Atlas of group a streptococcal vaccine candidates compiled using large-scale comparative genomics. *Nature Genetics*, 51(6):1035–1043, 2019.

[39] R. Di Benedetto, F. Mancini, M. Carducci, G. Gasperini, D. G. Moriel, A. Saul, F. Necchi, R. Rappuoli, and F. Micoli. Rational Design of a Glycoconjugate Vaccine against Group A Streptococcus. *Int J Mol Sci*, 21(22), 2020.

[40] Roberta Di Benedetto, Francesca Mancini, Martina Carducci, Gianmarco Gasperini, Danilo Gomes Moriel, Allan Saul, Francesca Necchi, Rino Rappuoli, and Francesca Micoli. Rational design of a glycoconjugate vaccine against group a streptococcus. *International Journal of Molecular Sciences*, 21(22), 2020.

[41] James E. DiCarlo, Julie E. Norville, Prashant Mali, Xavier Rios, John Aach, and George M. Church. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Research*, 41(7):4336–4343, 03 2013.

[42] Leanne M. Dooley, Tarek B. Ahmad, Manisha Pandey, Michael F. Good, and Michael Kotiw. Rheumatic heart disease: A review of the current status of global research activity. *Autoimmunity Reviews*, 20(2):102740, 2021.

[43] B. A. Dudding and E. M. Ayoub. Persistence of streptococcal group a antibody in patients with rheumatic valvular disease. *The Journal of experimental medicine*, 128(5):1081–1098, 1968.

[44] P. Durando, S. N. Faust, M. Fletcher, P. Krizova, A. Torres, and T. Welte. Experience with pneumococcal polysaccharide conjugate vaccine (conjugated to crm197 carrier protein) in children and adults *. *Clinical Microbiology and Infection*, 19:1–9, 2013.

[45] S. Döhrmann, C. N. LaRock, E. L. Anderson, J. N. Cole, B. Ryali, C. Stewart, P. Nonejuie, J. Pogliano, R. Corriden, P. Ghosh, and V. Nizet. Group A Streptococcal M1 Protein Provides Resistance against the Antimicrobial Activity of Histones. *Scientific Reports*, 7:43039, 2017.

[46] Rebecca J. Edgar, Vincent P. van Hensbergen, Alessandro Ruda, Andrew G. Turner, Pan Deng, Yoann Le Breton, Najib M. El-Sayed, Ashton T. Belew, Kevin S. McIver, Alastair G. McEwan, Andrew J. Morris, Gérard Lambeau, Mark J. Walker, Jeffrey S. Rush, Konstantin V. Korotkov, Göran Widmalm, Nina M. van Sorge, and Natalia Korotkova. Discovery of glycerol phosphate modification on streptococcal rhamnose polysaccharides. *Nature Chemical Biology*, 15(5):463–471, 2019.

[47] A. Eriksson and M. Norgren. Cleavage of antigen-bound immunoglobulin G by SpeB contributes to streptococcal persistence in opsonizing blood. *Infection and immunity*, 71(1):211–217, 2003.

[48] S. Esposito, S. Bianchini, M. Fastiggi, M. Fumagalli, L. Andreozzi, and D. Rigante. Geoepidemiological hints about *Streptococcus pyogenes* strains in relationship with acute rheumatic fever. *Autoimmun Rev*, 14(7):616–21, 2015.

[49] J Fairman, J Heinrichs, W Chan, and N Kapoor. Self-adjuvanted immunogenic conjugates (international patent publication wo/2020/010016), January 2020.

[50] H Findlow and R Borrow. Interactions of conjugate vaccines and co-administered vaccines. *Human Vaccines & Immunotherapeutics*, 12(1):226–230, 2016. PMID: 26619353.

[51] Louise K. Francois Watkins, Lesley McGee, Stephanie J. Schrag, Bernard Beall, Jennifer Hudson Jain, Tracy Pondo, Monica M. Farley, Lee H. Harrison, Shelley M. Zansky, Joan Baumbach, Ruth Lynfield, Paula Snippes Vagnone, Lisa A. Miller, William Schaffner, Ann R. Thomas, James P. Watt, Susan Petit, and Gayle E. Langley. Epidemiology of

Invasive Group B Streptococcal Infections Among Nonpregnant Adults in the United States, 2008-2016. *JAMA Internal Medicine*, 179(4):479–488, 04 2019.

[52] Andrea Fritzer, Beatrice M. Senn, Duc Bui Minh, Markus Hanner, Dieter Gelbmann, Birgit Noiges, Tamás Henics, Kai Schulze, Carlos A. Guzman, John Goodacre, Alexander von Gabain, Eszter Nagy, and Andreas L. Meinke. Novel conserved group a streptococcal proteins identified by the antigenome technology as vaccine candidates for a non-m protein-based vaccine. *Infection and Immunity*, 78(9):4051–4067, 2010.

[53] Hannah R Frost, Delphine Laho, Martina L Sanderson-Smith, Paul Licciardi, Susan Donath, Nigel Curtis, Joseph Kado, James B Dale, Andrew C Steer, and Pierre R Smeesters. Immune Cross-Opsonization Within emm Clusters Following Group A Streptococcus Skin Infection: Broadening the Scope of Type-Specific Immunity. *Clinical Infectious Diseases*, 65(9):1523–1531, 07 2017.

[54] C. Gagliotti, R. Buttazzi, S. Di Mario, F. Morsillo, and M. L. Moro. A regionwide intervention to promote appropriate antibiotic use in children reversed trends in erythromycin resistance to streptococcus pyogenes. *Acta Paediatr*, 104(9):e422–4, 2015.

[55] Marilena Gallotta, Giovanni Gancitano, Giampiero Pietrocola, Marirosa Mora, Alfredo Pezzicoli, Giovanna Tuscano, Emiliano Chiarot, Vincenzo Nardi-Dei, Anna Rita Taddei, Simonetta Rindi, Pietro Speziale, Marco Soriani, Guido Grandi, Immaculada Margarit, and Giuliano Bensi. Spyad, a moonlighting protein of group a streptococcus contributing to bacterial division and host cell adhesion. *Infection and Immunity*, 82(7):2890–2901, 2014.

[56] J. E. Galvin, M. E. Hemric, K. Ward, and M. W. Cunningham. Cytotoxic mAb from rheumatic carditis recognizes heart valves and laminin. *J Clin Invest*, 106(2):217–24, 2000.

[57] H. Gao, Y. Zhang, Y. Han, L. Yang, X. Liu, Z. Guo, Y. Tan, X. Huang, D. Zhou, and R. Yang. Phenotypic and transcriptional analysis of the osmotic regulator ompr in yersinia pestis. *BMC Microbiology*, 11(1):39, 2011.

[58] N.J. Gao, S. Uchiyama, L. Pill, S. Dahesh, J. Olson, L. Bautista, S. Maroju, A. Berges, J.Z. Liu, R.H. Zurich, N. van Sorge, J. Fairman, N. Kapoor, and V. Nizet. Site-Specific Conjugation of Call Wall Polyrhamnose to Protein SpyAD Envisioning a Safe Universal Group A Streptococcal Vaccine. *Infectious Microbes & Diseases*, 2021.

[59] K. Aaron Geno, Gwendolyn L. Gilbert, Joon Young Song, Ian C. Skovsted, Keith P. Klugman, Christopher Jones, Helle B. Konradsen, and Moon H. Nahm. Pneumococcal capsules and their types: Past, present, and future. *Clinical Microbiology Reviews*, 28(3):871–899, 2015.

[60] Jeffrey S. Gerber, Priya A. Prasad, Alexander G. Fiks, A. Russell Localio, Robert W. Grundmeier, Louis M. Bell, Richard C. Wasserman, Ron Keren, and Theoklis E. Zaoutis. Effect of an Outpatient Antimicrobial Stewardship Intervention on Broad-Spectrum Antibiotic Prescribing by Primary Care Pediatricians: A Randomized Trial. *JAMA*, 309(22):2345–2352, 06 2013.

[61] P. Ghosh. Variation, indispensability, and masking in the m protein. *Trends in Microbiology*, 26(2):132–144, 2018.

[62] I. Goldstein, P. Rebeyrotte, J. Parlebas, and B. Halpern. Isolation from heart valves of glycopeptides which share immunological properties with *Streptococcus haemolyticus* group A polysaccharides. *Nature*, 219(5156):866–8, 1968.

[63] Blanca E. Gonzalez, Gerardo Martinez-Aguilar, Kristina G. Hulten, Wendy A. Hammerman, Jorge Coss-Bu, Anna Avalos-Mishaan, Edward O. Mason, and Sheldon L. Kaplan. Severe Staphylococcal Sepsis in Adolescents in the Era of Community-Acquired Methicillin-Resistant Staphylococcus aureus. *Pediatrics*, 115(3):642–648, 2005.

[64] Morag R. Graham, Laura M. Smoot, Cristi A. Lux Migliaccio, Kimmo Virtaneva, Daniel E. Sturdevant, Stephen F. Porcella, Michael J. Federle, Gerald J. Adams, June R. Scott, and James M. Musser. Virulence control in group A Streptococcus by a two-component gene regulatory system: Global expression profiling and in vivo infection modeling. *Proceedings of the National Academy of Sciences*, 99(21):13855–13860, 2002.

[65] R. Grifantini, C. Toukoki, A. Colaprico, and I. Gryllos. Peroxide stimulon and role of perr in group a streptococcus. *Journal of Bacteriology*, 193(23):6539–6551, 2011.

[66] L. Guilherme, S. Freschi de Barros, K.F. Kohler, S.R. Santos, F. Morais Ferreira, W.R. Silva, R. Alencar, E. Postol, and J. Kalil. Rheumatic Heart Disease: Pathogenesis and Vaccine. *Current Protein & Peptide Science*, 19(9):900–908, 2018.

[67] Luiza Guilherme and Jorge Kalil. Rheumatic fever and rheumatic heart disease: Cellular mechanisms leading autoimmune reactivity and disease. *Journal of Clinical Immunology*, 30(1):17–23, 2010.

[68] F. F. Gunderson and N. P. Cianciotto. The crispr-associated gene cas2 of legionella pneumophila is required for intracellular infection of amoebae. *MBio*, 4(2):e00074–13, 2013.

[69] D. H. Haft, J. Selengut, E. F. Mongodin, and K. E. Nelson. A guild of 45 crispr-associated (cas) protein families and multiple crispr/cas subtypes exist in prokaryotic genomes. *PLOS Computational Biology*, 1(6):1–10, 11 2005.

[70] Shigeyuki Hamada, Nobuo Okahashi, Toshiko Yamamoto, Ichijiro Morisaki, Suzanne M. Michalek, and Jerry R. McGhee. Isolation and immunochemical characterization of carbohydrate antigens prepared from group A *Streptococcus pyogenes*. *Zentralblatt für Bakteriologie, Mikrobiologie und Hygiene. 1. Abt. Originale. A, Medizinische Mikrobiologie, Infektionskrankheiten und Parasitologie*, 256(1):37–48, 1983.

[71] A. Henningham, M. R. Davies, S. Uchiyama, N. M. van Sorge, S. Lund, K. T. Chen, M. J. Walker, J. N. Cole, and V. Nizet. Virulence Role of the GlcNAc Side Chain of the Lancefield Cell Wall Carbohydrate Antigen in Non-M1-Serotype Group A Streptococcus. *MBio*, 9(1), 2018.

[72] Anna Henningham, Christine M. Gillen, and Mark J. Walker. *Group A Streptococcal Vaccine Candidates: Potential for the Development of a Human Vaccine*, pages 207–242. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[73] H. Herwald, H. Cramer, M. Mörgelin, W. Russell, U. Sollenberg, A. Norrby-Teglund, H. Flodgaard, L. Lindbom, and L. Björck. M protein, a classical bacterial virulence determinant, forms complexes with fibrinogen that induce vascular leakage. *Cell*, 116(3):367–379, 2004.

[74] H. Heymann, J. M. Manniello, and S. S. Barkulis. Structure of streptococcal cell walls. v phosphate esters in the walls of group a: Streptococcus pyogenes. *Biochemical and Biophysical Research Communications*, 26(4):486–491, 1967.

[75] A. Hollands, R. K. Aziz, R. Kansal, M. Kotb, V. Nizet, and M. J. Walker. A naturally occurring mutation in ropb suppresses speb expression and reduces m1t1 group a streptococcal systemic virulence. *PLoS One*, 3(12):e4102, 2008.

[76] A. Hollands, M. A. Pence, A. M. Timmer, S. R. Osvath, L. Turnbull, C. B. Whitchurch, M. J. Walker, and V. Nizet. Genetic switch to hypervirulence reduces colonization phenotypes of the globally disseminated group a streptococcus m1t1 clone. *J Infect Dis*, 202(1):11–9, 2010.

[77] T T Huang, H Malke, and J J Ferretti. Heterogeneity of the streptokinase gene in group a streptococci. *Infection and Immunity*, 57(2):502–506, 1989.

[78] RICHARD A. INSEL. Potential alterations in immunogenicity by combining or simultaneously administering vaccine componentsa. *Annals of the New York Academy of Sciences*, 754(1):35–48, 1995.

[79] M. Isaka, I. Tatsuno, J. Maeyama, H. Matsui, Y. Zhang, and T. Hasegawa. The yvqe two-component system controls biofilm formation and acid production in streptococcus pyogenes. *APMIS*, 124(7):574–585, 2016.

[80] Y. Ji, B. Carlson, A. Kondagunta, and P. P. Cleary. Intranasal immunization with c5a peptidase prevents nasopharyngeal colonization of mice by the group a streptococcus. *Infect Immun*, 65(6):2080–7, 1997.

[81] Y. Ji, L. McLandsborough, A. Kondagunta, and P. P. Cleary. C5a peptidase alters clearance and trafficking of group a streptococci by infected mice. *Infect Immun*, 64(2):503–10, 1996.

[82] W. Jiang, D. Bikard, D. Cox, F. Zhang, and L. A. Marraffini. Rna-guided editing of bacterial genomes using crispr-cas systems. *Nature Biotechnology*, 31(3):233–239, 2013.

[83] Scott Jones, Nicole J. Moreland, Marta Zancolli, Jeremy Raynes, Jacelyn M.S. Loh, Pierre R. Smeesters, Shiranee Sriskandan, Jonathan R. Carapetis, John D. Fraser, and David

Goldblatt. Development of an opsonophagocytic killing assay for group a streptococcus. *Vaccine*, 36(26):3756–3763, 2018.

[84] A. Kabanova, I. Margarit, F. Berti, M. R. Romano, G. Grandi, G. Bensi, E. Chiarot, D. Proietti, E. Swennen, E. Cappelletti, P. Fontani, D. Casini, R. Adamo, V. Pinto, D. Skibinski, S. Capo, G. Buffi, M. Gallotta, W. J. Christ, A. S. Campbell, J. Pena, P. H. Seeberger, R. Rappuoli, and P. Costantino. Evaluation of a group a streptococcus synthetic oligosaccharide as vaccine candidate. *Vaccine*, 29(1):104–14, 2010.

[85] R. G. Kansal, A. McGeer, D. E. Low, A. Norrby-Teglund, and M. Kotb. Inverse relation between disease severity and expression of the streptococcal cysteine protease, speb, among clonal m1t1 isolates recovered from invasive group a streptococcal infection cases. *Infection and Immunity*, 68(11):6362–6369, 2000.

[86] Neeraj Kapoor, Ivana Vanjak, James Rozzelle, Aym Berges, Wei Chan, Gang Yin, Cuong Tran, Aaron K. Sato, Alexander R. Steiner, Thao P. Pham, Ashley J. Birkett, Carole A. Long, Jeff Fairman, and Kazutoyo Miura. Malaria derived glycosylphosphatidylinositol anchor enhances anti-pfs25 functional antibodies that block malaria transmission. *Biochemistry*, 57(5):516–519, 2018.

[87] Forrest E. Kendall, Michael Heidelberger, and Martin H. Dawson. A serologically inactive polysaccharide elaborated by mucoid strains of group a hemolytic streptococcus. *Journal of Biological Chemistry*, 118(1):61–69, 1937.

[88] F. Khatun, C. C. Dai, T. Rivera-Hernandez, W. M. Hussein, Z. G. Khalil, R. J. Capon, I. Toth, and R. J. Stephenson. Immunogenicity Assessment of Cell Wall Carbohydrates of Group A Streptococcus via Self-Adjuvanted Glyco-lipopeptides. *ACS Infect Dis*, 7(2):390–405, 2021.

[89] C.A. Kirvan, S.E. Swedo, J.S. Jeuser, and M. W. Cunningham. Mimicry and autoantibody-mediated neuronal cell signaling in Syndenham chorea. *Nat Med*, 9(7), 2003.

[90] B. Kreikemeyer, M. D. P. Boyle, B. A. Buttaro, M. Heinemann, and A. Podbielski. Group a streptococcal growth phase-associated virulence factor regulation by a novel operon (fas) with homologies to two-component-type regulators requires a small rna molecule. *Molecular Microbiology*, 39(2):392–406, 2001.

[91] Uwe C. Kreis, Vikram Varma, and B. Mario Pinto. Application of two-dimensional NMR spectroscopy and molecular dynamics simulations to the conformational analysis of oligosaccharides corresponding to the cell-wall polysaccharide of Streptococcus group A. *International Journal of Biological Macromolecules*, 17(3-4):117–130, 1995.

[92] C. Kuenne, A. Billion, M. A. Mraheil, A. Strittmatter, R. Daniel, A. Goesmann, S. Barbuddhe, T. Hain, and T. Chakraborty. Reassessment of the listeria monocytogenes pan-genome reveals dynamic integration hotspots and mobile genetic elements as major components of the accessory genome. *BMC Genomics*, 14(1):47, 2013.

[93] Mie Kurosawa, Masataka Oda, Hisanori Domon, Toshihito Isono, Yuki Nakamura, Issei Saitoh, Haruaki Hayasaki, Masaya Yamaguchi, Shigetada Kawabata, and Yutaka Terao. Streptococcus pyogenes camp factor promotes bacterial adhesion and invasion in pharyngeal epithelial cells without serum via pi3k/akt signaling pathway. *Microbes and Infection*, 20(1):9–18, 2018.

[94] Mie Kurosawa, Masataka Oda, Hisanori Domon, Issei Saitoh, Haruaki Hayasaki, and Yutaka Terao. Streptococcus pyogenes camp factor attenuates phagocytic activity of raw 264.7 cells. *Microbes and Infection*, 18(2):118–127, 2016.

[95] R. C. Lancefield. The Antigenic Complex of *Streptococcus Haemolyticus* : I. Demonstration of a Type-Specific Substance in Extracts of *Streptococcus Haemolyticus*. *J Exp Med*, 47(1):91–103, 1928.

[96] R. C. Lancefield. A serological differentiation of human and other groups of hemolytic streptococci. *J Exp Med*, 57(4):571–595, 1933.

[97] J. D. Lapek, P. Greninger, R. Morris, A. Amzallag, I. Pruteanu-Malinici, C. H. Benes, and W. Haas. Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities. *Nature Biotechnology*, 35(10):983–989, 2017.

[98] Jr. Lapek, J. D., R. H. Mills, J. M. Wozniak, A. Campeau, R. H. Fang, X. Wei, K. van de Groep, A. Perez-Lopez, N. M. van Sorge, M. Raffatellu, R. Knight, L. Zhang, and D. J. Gonzalez. Defining host responses during systemic bacterial infection through construction of a murine organ proteome atlas. *Cell Systems*, 2018.

[99] C. N. LaRock, S. Döhrmann, J. Todd, R. Corriden, J. Olson, T. Johannssen, B. Lepenies, R. L. Gallo, P. Ghosh, and V. Nizet. Group a streptococcal m1 protein sequesters cathelicidin to evade innate immune killing. *Cell Host & Microbe*, 18(4):471–477, 2015.

[100] P. A. Lawson, G. Foster, E. Falsen, S. J. Markopoulos, and M. D. Collins. *streptococcus castoreus* sp. *nov.*, isolated from a beaver (castor fiber). *Int J Syst Evol Microbiol*, 55(Pt 2):843–6, 2005.

[101] Shun Mei Lin, Yong Zhi, Ki Bum Ahn, Sangyong Lim, and Ho Seong Seo. Status of group b streptococcal vaccine development. *Clinical and experimental vaccine research*, 7(1):76–81, 2018.

[102] R. Louwen, D. Horst-Kreft, A.G. de Boer, L. van der Graaf, G. de Knegt, M. Hamersma, A.P. Heikema, A.R. Timms, B.C. Jacobs, J.A. Wagenaar, H. P. Endtz, J. Van der Oost, J.M. Wells, E.E. Nieuwenhuis, A.H. van Vliet, P.T. Willemsen, P. Van Baarlen, and A. van Belkum. A novel link between *Campylobacter jejuni* bacteriophage defence, virulence and Guillain Barré syndrome. *Eur J Clin Microbiol Infect Dis*, 32(2):207–26, 2013.

[103] R. Louwen, R. H. Staals, H. P. Endtz, P. van Baarlen, and J. van der Oost. The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol Mol Biol Rev*, 78(1):74–88, 2014.

[104] Y. A. Lue, I. P. Howit, and P. D. Ellner. Rapid grouping of beta-hemolytic streptococci by latex agglutination. *J Clin Micro*, 8(3):326–328, 1978.

[105] William R. Lyon and Michael G. Caparon. Role for Serine Protease HtrA (DegP) of Streptococcus pyogenes in the Biogenesis of Virulence Factors SpeB and the Hemolysin Streptolysin S. *Infection and Immunity*, 72(3):1618–1625, 2004.

[106] K. Ma, Q. Cao, S. Luo, Z. Wang, G. Liu, C. Lu, and Y. Liu. cas9 Enhances Bacterial Virulence by Repressing the regR Transcriptional Regulator in *Streptococcus agalactiae*. *Infect Immun*, 86(3), 2018.

[107] J. C. Madden, N. Ruiz, and M. Caparon. Cytolysin-Mediated Translocation (CMT): A Functional Equivalent of Type III Secretion in Gram-Positive Bacteria. *Cell*, 104(1):143–152, 2001.

[108] K. S. Makarova, Y. I. Wolf, O. S. Alkhnbashi, F. Costa, S. A. Shah, S. J. Saunders, R. Barrangou, S. J. J. Brouns, E. Charpentier, D. H. Haft, P. Horvath, S. Moineau, F. J. M. Mojica, R. M. Terns, M. P. Terns, M. F. White, A. F. Yakunin, R. A. Garrett, J. van der Oost, R. Backofen, and E. V. Koonin. An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology*, 13(11):722–736, 2015.

[109] L. A. Marraffini. *The CRISPR-Cas system of Streptococcus pyogenes: function and applications*. University of Oklahoma Health Sciences Center, 2016.

[110] W. J. Martin, A. C. Steer, P. R. Smeesters, J. Keeble, M. Inouye, J. Carapetis, and I. P. Wicks. Post-infectious group a streptococcal autoimmune syndromes and the heart. *Autoimmun Rev*, 14(8):710–25, 2015.

[111] Viviana Martinez, Miles Ingwers, James Smith, John Glushka, Ting Yang, and Maor Bar-Peled. Biosynthesis of udp-4-keto-6-deoxyglucose and udp-rhamnose in pathogenic fungi ¡em¿magnaporthe grisea¡/em¿ and ¡em¿botryotinia fuckeliana¡/em¿*¡sup¿¡/sup¿. *Journal of Biological Chemistry*, 287(2):879–892, 2012.

[112] T. B. Martins, J. L. Hoffman, N. H. Augustine, A. R. Phansalkar, V. A. Fischetti, J. B. Zabriskie, P. P. Cleary, J. M. Musser, L. G. Veasy, and H. R. Hill. Comprehensive analysis of antibody responses to streptococcal and tissue antigens in patients with acute rheumatic fever. *Int Immunol*, 20(3):445–52, 2008.

[113] Malihe Masomian, Zuleeza Ahmad, Lai Ti Gew, and Chit Laa Poh. Development of next generation streptococcus pneumoniae vaccines conferring broad protection. *Vaccines*, 8(1), 2020.

[114] B. F. Massell, J. G. Michael, J. Amezcua, and M. Siner. Secondary and apparent primary antibody responses after group a streptococcal vaccination of 21 children. *Applied microbiology*, 16(3):509–518, 1968.

[115] Benedict F. Massell, Larry H. Honikman, and Jacqueline Amezcua. Rheumatic Fever Following Streptococcal Vaccination: Report of Three Cases. *JAMA*, 207(6):1115–1119, 02 1969.

[116] M. McCarty. The lysis of group A hemolytic streptococci by extracellular enzymes of Streptomyces albus. II. Nature of the cellular substrate attacked by the lytic enzymes. *J Exp Med*, 96(6):569–80, 1952.

[117] M. McCarty. Variation in the group-specific carbohydrate of group a streptococci. ii. studies on the chemical basis for serological specificity of the carbohydrates. *The Journal of experimental medicine*, 104(5):629–643, 1956.

[118] M. McCarty. The streptococcal cell wall, 18 December 1969 1969.

[119] M. McCarty and R. C. Lancefield. Variation in the group-specific carbohydrate of group a streptococci. i. immunochemical studies on the carbohydrates of variant strains. *The Journal of experimental medicine*, 102(1):11–28, 1955.

[120] Ravi P.N. Mishra, Ravi S.P. Yadav, Christopher Jones, Salvatore Nocadello, George Minasov, Ludmilla A. Shuvalova, Wayne F. Anderson, and Akshay Goel. Structural and immunological characterization of E. coli derived recombinant CRM197 protein used as carrier in conjugate vaccines. *Bioscience Reports*, 38(5), 09 2018.

[121] M. Y. Mistou, I. C. Sutcliffe, and N. M. van Sorge. Bacterial glycobiology: rhamnose-containing cell wall polysaccharides in gram-positive bacteria. *FEMS Microbiol Rev*, 40(4):464–79, 2016.

[122] T. Miyoshi-Akiyama, D. Takamatsu, M. Koyanagi, J. Zhao, K. Imanishi, and T. Uchiyama. Cytocidal Effect of Streptococcus pyogenes on Mouse Neutrophils In Vivo and the Critical Role of Streptolysin S. *The Journal of Infectious Diseases*, 192(1):107–116, 07 2005.

[123] F. J. M. Mojica, C. Díez-Villaseñor, J. García-Martínez, and E. Soria. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution*, 60(2):174–182, 2005.

[124] E. M. Molloy, P. D. Cotter, C. Hill, D. A. Mitchell, and R. P. Ross. Streptolysin S-like virulence factors: the continuing sagA. *Nature Reviews Microbiology*, 9:670, 2011.

[125] D. C. Nelson, J. Garbe, and M. Collin. Cysteine proteinase SpeB from Streptococcus pyogenes - a potent modifier of immunologically important host and bacterial proteins. *Biol Chem*, 392(12):1077–88, 2011.

[126] Y. Niu, B. Shen, Y. Cui, Y. Chen, J. Wang, L. Wang, Y. Kang, X. Zhao, W. Si, W. Li, A. P. Xiang, J. Zhou, X. Guo, Y. Bi, C. Si, B. Hu, G. Dong, H. Wang, Z. Zhou, T. Li, T. Tan, X. Pu, F. Wang, S. Ji, Q. Zhou, X. Huang, W. Ji, and J. Sha. Generation of Gene-Modified Cynomolgus Monkey via Cas9/RNA-Mediated Gene Targeting in One-Cell Embryos. *Cell*, 156(4):836–843, 2014.

[127] V. Nizet. Streptococcal β-hemolysins: genetics and role in disease pathogenesis. *Trends in Microbiology*, 10(12):575–580, 2002.

[128] A. H. Nobbs, R. J. Lamont, and H. F. Jenkinson. *Streptococcus* Adherence and Colonization. *Microbiology and Molecular Biology Reviews*, 73(3):407–450, 2009.

[129] P. Nyberg, M. Rasmussen, and L. Björck. α2-Macroglobulin-Proteinase Complexes Protect Streptococcus pyogenes from Killing by the Antimicrobial Peptide LL-37. *Journal of Biological Chemistry*, 279(51):52820–52823, 2004.

[130] N. Okada, M. K. Liszewski, J. P. Atkinson, and M. Caparon. Membrane cofactor protein (cd46) is a keratinocyte receptor for the m protein of the group a streptococcus. *Proceedings of the National Academy of Sciences of the United States of America*, 92(7):2489–2493, 1995.

[131] Joshua Osowicki, Johan Vekemans, David C. Kaslow, Martin H. Friede, Jerome H. Kim, and Andrew C. Steer. Who/ivi global stakeholder consultation on group a streptococcus vaccine development: Report from a meeting held on 12–13 december 2016. *Vaccine*, 36(24):3397–3405, 2018.

[132] Elizabeth Pennisi. The CRISPR Craze. *Science*, 341(6148):833–836, 2013.

[133] R. Perez-Rodriguez, C. Haitjema, Q. Huang, K. H. Nam, S. Bernardis, A. Ke, and M. P. DeLisa. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol Microbiol*, 79(3):584–99, 2011.

[134] F. Peters, G. Karthikeyan, J. Abrams, L. Muhwava, and L. Zühlke. Rheumatic heart disease: current status of diagnosis and therapy. *Cardiovasc Diagn Ther*, 10(2):305–315, 2020.

[135] O. Pitirollo, F. Micoli, F. Necchi, F. Mancini, M. Carducci, R. Adamo, C. Evangelisti, L. Morelli, L. Polito, and L. Lay. Gold nanoparticles morphology does not affect the multivalent presentation and antibody recognition of Group A Streptococcus synthetic oligorhamnans. *Bioorg Chem*, 99:103815, 2020.

[136] Craig A. Pritzlaff, Jennifer C. W. Chang, Shrin P. Kuo, Glen S. Tamura, Craig E. Rubens, and Victor Nizet. Genetic basis for the -haemolytic/cytolytic activity of group b streptococcus. *Molecular Microbiology*, 39(2):236–248, 2001.

[137] M. Punukollu, N. Mushet, M. Linney, C. Hennessy, and M. Morton. Neuropsychiatric manifestations of Sydenham's chorea: a systematic review. *Dev Med Child Neurol*, 58(1):16–28, 2016.

[138] Anna P. Ralph and Jonathan R. Carapetis. *Group A Streptococcal Diseases and Their Global Burden*, pages 1–27. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[139] Magnus Rasmussen, Hans-Peter Müller, and Lars Björck. Protein GRAB of *Streptococcus pyogenes* Regulates Proteolysis at the Bacterial Surface by Binding α2-Macroglobulin. *Journal of Biological Chemistry*, 274(22):15336–15344, 1999.

[140] Assaf Raz and Vincent A. Fischetti. Sortase A localizes to distinct foci on the *Streptococcus pyogenes* membrane. *Proceedings of the National Academy of Sciences*, 2008.

[141] Mark Reglinski, Nicola N. Lynskey, Yoon Jung Choi, Robert J. Edwards, and Shiranee Sriskandan. Development of a multicomponent vaccine for *Streptococcus pyogenes* based on the antigenic targets of IVIG. *Journal of Infection*, 72(4):450–459, 2016.

[142] Anaïs Le Rhun, Andrés Escalera-Maurer, Majda Bratovič, and Emmanuelle Charpentier. CRISPR-Cas in *Streptococcus pyogenes*. *RNA Biology*, 16(4):380–389, 2019. PMID: 30856357.

[143] T. Rivera-Hernandez, M. Pandey, A. Henningham, J. Cole, B. Choudhury, A. J. Cork, C. M. Gillen, K. A. Ghaffar, N. P. West, G. Silvestri, M. F. Good, P. M. Moyle, I. Toth, V. Nizet, M. R. Batzloff, and M. J. Walker. Differing Efficacies of Lead Group A Streptococcal Vaccine Candidates and Full-Length M Protein in Cutaneous and Invasive Disease Models. *MBio*, 7(3), 2016.

[144] Manuel J. Rodríguez-Ortega, Nathalie Norais, Giuliano Bensi, Sabrina Liberatori, Sabrina Capo, Marirosa Mora, Maria Scarselli, Francesco Doro, Germano Ferrari, Ignazio Garaguso, Tiziana Maggi, Anita Neumann, Alessia Covre, John L. Telford, and Guido Grandi. Characterization and identification of vaccine candidate proteins through analysis of the group A *Streptococcus* surface proteome. *Nature Biotechnology*, 24(2):191–197, 2006.

[145] Jeffrey S. Rush, Rebecca J. Edgar, Pan Deng, Jing Chen, Haining Zhu, Nina M. van Sorge, Andrew J. Morris, Konstantin V. Korotkov, and Natalia Korotkova. The molecular mechanism of *N*-acetylglucosamine side-chain attachment to the Lancefield group A carbohydrate in *Streptococcus pyogenes*. *Journal of Biological Chemistry*, 292(47):19441–19457, 2017.

[146] Neal J Russell, Anna C Seale, Megan O'Driscoll, Catherine O'Sullivan, Fiorella Bianchi-Jassir, Juan Gonzalez-Guarin, Joy E Lawn, Carol J Baker, Linda Bartlett, Clare Cutland, Michael G Gravett, Paul T Heath, Kirsty Le Doare, Shabir A Madhi, Craig E Rubens, Stephanie Schrag, Ajoke Sobanjo-ter Meulen, Johan Vekemans, Samir K Saha, Margaret Ip, and for the GBS Maternal Colonization Investigator Group. Maternal Colonization With Group B Streptococcus and Serotype Distribution Worldwide: Systematic Review and Meta-analyses. *Clinical Infectious Diseases*, $65(\text{suppl}_2) : S100 - -S111, 112017$.

[147] H. Sabharwal, F. Michon, D. Nelson, W. Dong, K. Fuchs, R. C. Manjarrez, A. Sarkar, C. Uitz, A. Viteri-Jackson, R. S. Suarez, M. Blake, and J. B. Zabriskie. Group A streptococcus (GAS) carbohydrate as an immunogen for protection against GAS infection. *J Infect Dis*, 193(1):129–35, 2006.

[148] Hemant Sabharwal, Daniel Nelson, Francis Michon, and John Zabriskie. The streptococcal group A carbohydrate protects against nasal colonization with group A streptococci in mice. *International Congress Series*, 1289:329–331, 2006.

[149] M. Sadarangani. Protection Against Invasive Infections in Children Caused by Encapsulated Bacteria. *Front Immunol*, 9:2674, 2018.

[150] Sanaz Salehi, Claudia M. Hohn, Thomas A. Penfound, and James B. Dale. Development of an opsonophagocytic killing assay using hl-60 cells for detection of functional antibodies against streptococcus pyogenes. *mSphere*, 3(6), 2018.

[151] L. G. Salvadori, M. S. Blake, M. McCarty, J. Y. Tai, and J. B. Zabriskie. Group A streptococcus-liposome ELISA antibody titers to group A polysaccharide and opsonophagocytic capabilities of the antibodies. *J Infect Dis*, 171(3):593–600, 1995.

[152] T. R. Sampson, S. D. Saroj, A. C. Llewellyn, Y. L. Tzeng, and D. S. Weiss. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature*, 497(7448):254–7, 2013.

[153] M. Sanderson-Smith, D. M. De Oliveira, J. Guglielmini, D. J. McMillan, T. Vu, J. K. Holien, A. Henningham, A. C. Steer, D. E. Bessen, J. B. Dale, N. Curtis, B. W. Beall, M. J. Walker, M. W. Parker, J. R. Carapetis, L. Van Melderen, K. S. Sriprakash, and P. R. Smeesters. A systematic and functional classification of *Streptococcus pyogenes* that serves as a new tool for molecular typing and vaccine development. *J Infect Dis*, 210(8):1325–38, 2014.

[154] Misu Sanson, Nishanth Makthal, Anthony R. Flores, Randall J. Olsen, James M. Musser, and Muthiah Kumaraswami. Adhesin competence repressor (AdcR) from Streptococcus pyogenes controls adaptive responses to zinc limitation and contributes to virulence. *Nucleic Acids Research*, 43(1):418–432, 12 2014.

[155] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. Fiji: an open-source platform for biological-image analysis. *Nat Methods*, 9(7):676–82, 2012.

[156] Muhammad A. B. Shabbir, Yanping Tang, Zihui Xu, Mingyue Lin, Guyue Cheng, Menghong Dai, Xu Wang, Zhengli Liu, Zonghui Yuan, and Haihong Hao. The Involvement of the Cas9 Gene in Virulence of *Campylobacter jejuni*. *Frontiers in Cellular and Infection Microbiology*, 8:285, 2018.

[157] Samuel A. Shelburne, Chanel Granville, Maria Tokuyama, Izabela Sitkiewicz, Payal Patel, and James M. Musser. Growth Characteristics of and Virulence Factor Production by Group A *Streptococcus* during Cultivation in Human Saliva. *Infection and Immunity*, 73(8):4723, 2005.

[158] Anita Shet, Edward L. Kaplan, Dwight R. Johnson, and P. Patrick Cleary. Immune Response to Group A Streptococcal C5a Peptidase in Children: Implications for Vaccine Development. *The Journal of Infectious Diseases*, 188(6):809–817, 09 2003.

[159] A. R. Shikhman, N. S. Greenspan, and M. W. Cunningham. Cytokeratin peptide SFGSGFGGGY mimics N-acetyl-beta-D-glucosamine in reaction with antibodies and lectins, and induces *in vivo* anti-carbohydrate antibody response. *The Journal of Immunology*, 153(12):5593, 1994.

[160] S. T. Shulman and E. M. Ayoub. Serologic cross-reactions among streptococcal group A, A-variant, and C polysaccharides. *Clin Immunol Immunopathol*, 28(2):229–42, 1983.

[161] A Sims Sanyahumbi, S Colquhoun, R Wyber, and JR Carapetis. *Global Disease Burden of Group A Streptococcus*, book section Global Disease Burden of Group A Streptococcus. University of Oklahoma Health Sciences Center, Oklahoma City, 2016.

[162] Bibhuti N. Singh, Gary R. Hayes, John J. Lucas, Ulf Sommer, Nelly Viseux, Ekaterina Mirgorodskaya, Radiana T. Trifonova, Rosaria Rita S. Sassi, Catherine E. Costello, and Raina N. Fichorova. Structural details and composition of trichomonas vaginalis lipophosphoglycan in relevance to the epithelial immune function. *Glycoconjugate Journal*, 26(1):3, 2008.

[163] A. C. Steer, I. Law, L. Matatolu, B. W. Beall, and J. R. Carapetis. Global emm type distribution of group A *streptococci*: systematic review and implications for vaccine development. *Lancet Infect Dis*, 9(10):611–6, 2009.

[164] Gene H. Stollerman and James B. Dale. The Importance of the Group A Streptococcus Capsule in the Pathogenesis of Human Infections: A Historical Perspective. *Clinical Infectious Diseases*, 46(7):1038–1045, 04 2008.

[165] R. Stuike-Prill and B. M. Pinto. Conformational analysis of oligosaccharides corresponding to the cell-wall polysaccharide of the *Streptococcus* group A by Metropolis Monte Carlo simulations. *Carbohydrate research*, 279:59–73, 1995.

[166] Anjuli M. Timmer, Sascha A. Kristian, Vivekanand Datta, Arthur Jeng, Christine M. Gillen, Mark J. Walker, Bernard Beall, and Victor Nizet. Serum opacity factor promotes group A streptococcal epithelial cell invasion and virulence. *Molecular Microbiology*, 62(1):15–25, 2006.

[167] Satoshi Uchiyama, Simon Döhrmann, Anjuli M. Timmer, Neha Dixit, Mariam Ghochani, Tamara Bhandari, John C. Timmer, Kimberly Sprague, Juliane Bubeck-Wardenburg, Scott I. Simon, and Victor Nizet. Streptolysin o rapidly impairs neutrophil oxidative burst and antibacterial responses to group a streptococcus. *Frontiers in Immunology*, 6:581, 2016.

[168] J. A. Valderrama and V. Nizet. Group A Streptococcus encounters with host macrophages. *Future Microbiology*, 13(1):119–134, 2017.

[169] J. A. Valderrama, A. M. Riestra, N. J. Gao, C. N. LaRock, N. Gupta, S. R. Ali, H. M. Hoffman, P. Ghosh, and V. Nizet. Group A streptococcal M protein activates the NLRP3 inflammasome. *Nature Microbiology*, 2(10):1425–1434, 2017.

[170] I. van de Rijn, B. Bernish, and D. L. Crater. Analysis of hyaluronic acid capsule expression in group a streptococci. *Adv Exp Med Biol*, 418:965–9, 1997.

[171] S. L. van der Beek, Y. Le Breton, A. T. Ferenbach, R. N. Chapman, D. M. van Aalten, I. Navratilova, G. J. Boons, K. S. McIver, N. M. van Sorge, and H. C. Dorfmueller. GacA is essential for Group A Streptococcus and defines a new class of monomeric dTDP-4-dehydrorhamnose reductases (RmlD). *Mol Microbiol*, 98(5):946–62, 2015.

[172] V. P. van Hensbergen, E. Movert, V. de Maat, C. Luchtenborg, Y. Le Breton, G. Lambeau, C. Payre, A. Henningham, V. Nizet, J. A. G. van Strijp, B. Brugger, F. Carlsson, K. S. McIver, and N. M. van Sorge. Streptococcal Lancefield polysaccharides are critical cell wall determinants for human Group IIA secreted phospholipase A2 to exert its bactericidal effects. *PLoS Pathog*, 14(10):e1007348, 2018.

[173] Merel N. van Kassel, Sanne W.C.M. Janssen, Sanne Kofman, Matthijs C. Brouwer, Diederik van de Beek, and Merijn W. Bijlsma. Prevalence of group b streptococcal colonization in the healthy non-pregnant population: a systematic review and meta-analysis. *Clinical Microbiology and Infection*, 2021.

[174] N. M. van Sorge, J. N. Cole, K. Kuipers, A. Henningham, R. K. Aziz, A. Kasirer-Friede, L. Lin, E. T. M. Berends, M. R. Davies, G. Dougan, F. Zhang, S. Dahesh, L. Shaw, J. Gin, M. Cunningham, J. A. Merriman, J. Hutter, B. Lepenies, S. H. M. Rooijakkers, R. Malley, M. J. Walker, S. J. Shattil, P. M. Schlievert, B. Choudhury, and V. Nizet. The classical lancefield antigen of group a streptococcus is a virulence determinant with implications for vaccine design. *Cell Host Microbe*, 15(6):729–740, 2014.

[175] Sandra J. van Vliet, Liana Steeghs, Sven C. M. Bruijns, Medi M. Vaezirad, Christian Snijders Blok, Jésus A. Arenas Busto, Marcel Deken, Jos P. M. van Putten, and Yvette van Kooyk. Variation of *Neisseria gonorrhoeae* lipooligosaccharide directs dendritic cell-induced T helper responses. *PLoS pathogens*, 5(10):e1000625–e1000625, 2009.

[176] René G. VanDeVoorde. Acute poststreptococcal glomerulonephritis: The most common acute glomerulonephritis. *Pediatrics in Review*, 36(1):3–13, 2015.

[177] Luis A. Vega, Horst Malke, and Kevin McIver. *Virulence-Related Transcriptional Regulators of Streptococcus pyogenes*. The University of Oklahoma Health Sciences Center., Oklahoma City OK, 2016.

[178] Johan Vekemans, Fernando Gouvea-Reis, Jerome H Kim, Jean-Louis Excler, Pierre R Smeesters, Katherine L O'Brien, Chris A Van Beneden, Andrew C Steer, Jonathan R Carapetis, and David C Kaslow. The Path to Group A Streptococcus Vaccines: World Health Organization Research and Development Technology Roadmap and Preferred Product Characteristics. *Clinical Infectious Diseases*, 69(5):877–883, 01 2019.

[179] J. M. Voyich, D. E. Sturdevant, K. R. Braughton, S. D. Kobayashi, B. Lei, K. Virtaneva, D. W. Dorward, J. M. Musser, and F. R. DeLeo. Genome-wide protective response used by group A Streptococcus to evade destruction by human polymorphonuclear leukocytes. *Proceedings of the National Academy of Sciences*, 100(4):1996–2001, 2003.

[180] M. J. Walker, T. C. Barnett, J. D. McArthur, J. N. Cole, C. M. Gillen, A. Henningham, K. S. Sriprakash, M. L. Sanderson-Smith, and V. Nizet. Disease manifestations and pathogenic mechanisms of Group A Streptococcus. *Clin Microbiol Rev*, 27(2):264–301, 2014.

[181] S. Wang, Y. Zhao, G. Wang, S. Feng, Z. Guo, and G. Gu. Group A Streptococcus Cell Wall Oligosaccharide-Streptococcal C5a Peptidase Conjugates as Effective Antibacterial Vaccines. *ACS Infect Dis*, 6(2):281–290, 2020.

[182] Mark Welch, Sridhar Govindarajan, Jon E. Ness, Alan Villalobos, Austin Gurney, Jeremy Minshull, and Claes Gustafsson. Design parameters to control synthetic gene expression in escherichia coli. *PLOS ONE*, 4(9):1–10, 09 2009.

[183] M. R. Wessels. *Cell Wall and Surface Molecules: Capsule*. University of Oklahoma Health Sciences Center © The University of Oklahoma Health Sciences Center., Oklahoma City (OK), 2016.

[184] M. R. Wessels, A. E. Moses, J. B. Goldberg, and T. J. DiCesare. Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci. *Proc Natl Acad Sci USA*, 88(19):8317–8321, 1991.

[185] Y. Xiao, T. H. Hsiao, U. Suresh, H. I. Chen, X. Wu, S. E. Wolf, and Y. Chen. A novel significance score for gene selection and ranking. *Bioinformatics*, 30(6):801–7, 2014.

[186] X. Yan, C. Zhao, A. Budin-Verneuil, A. Hartke, A. Rincé, M. S. Gilmore, Y. Auffray, and V. Pichereau. The (p)ppGpp synthetase RelA contributes to stress adaptation and virulence in Enterococcus faecalis V583. *Microbiology*, 155(10):3226–3237, 2009.

[187] M. E. Zegans, J. C. Wagner, K. C. Cady, D. M. Murphy, J. H. Hammond, and G. A. O'Toole. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of Pseudomonas aeruginosa. *J Bacteriol*, 191(1):210–9, 2009.

[188] Yisheng Zhao, Subo Wang, Guirong Wang, Hui Li, Zhongwu Guo, and Guofeng Gu. Synthesis and immunological studies of group A Streptococcus cell-wall oligosaccharide–streptococcal C5a peptidase conjugates as bivalent vaccines. *Org Chem Front*, 6:3589–3596, 2019.

[189] Robert A. Zimmerman, Arthur H. Auernheimer, and Angelo Taranta. Precipitating Antibody to Group a Streptococcal Polysaccharide in Humans. *The Journal of Immunology*, 107(3):832–841, 1971.

[190] Azul Zorzoli, Benjamin H. Meyer, Elaine Adair, Vladimir I. Torgov, Vladimir V. Veselovsky, Leonid L. Danilov, Dusan Uhrin, and Helge C. Dorfmueller. Group A, B, C, and G Streptococcus Lancefield antigen biosynthesis is initiated by a conserved $\alpha-d-$GlcNAc$-\beta-1,4-l-$rhamnosyltransferase. *The Journal of Biological Chemistry*, 294(42):15237–15256, 2019.