

LONDON  
SCHOOL of  
HYGIENE  
& TROPICAL  
MEDICINE



**Development of a Glycoconjugate Vaccine Against  
Group A Streptococcus (GAS)**

**Keira Burns**

**Thesis submitted in accordance with the requirements for  
the degree of  
Doctor of Philosophy  
of the  
University of London  
September 2022**

**Department of Infection Biology**

**Faculty of Infectious and Tropical Diseases**

**LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE**

Funded by National Institute for Biological Standards and Control  
(NIBSC)

Research group affiliation(s): Dr. Alexandra Shaw (NIBSC)  
Dr. Fatme Mawas (NIBSC)  
Prof. Brendan Wren (LSHTM)

## Declaration

I, Keira Burns, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.



Keira Burns

Countersigned by supervisors



Dr. Alexandra Shaw



Prof. Brendan Wren



Dr. Fatme Mawas

## Abstract

Group A *Streptococcus* (GAS) is responsible for superficial infections, systemic disease, and autoimmune complications globally. Despite extensive research no commercial GAS vaccine currently exists. The highly abundant, conserved surface exposed Lancefield Group A Carbohydrate (GAC) is of interest as a vaccine candidate due to evidence of protective properties of anti-GAC antibodies. However, for effective vaccines it is necessary to conjugate polysaccharides to protein carriers to improve immune memory responses. This thesis investigates glycoconjugate production as an approach for the development of GAS vaccines, comparing chemical linkage and an *E. coli* cell based (“biological”) glycoengineering system as distinct manufacturing methods.

For chemical conjugation reactions, either wildtype GAC or GAS\_Rha (GAC extracted from mutant  $\Delta gacI$  NCTC-8198 strain devoid of autoimmunogenic GlcNAc epitopes) was enzymatically extracted from GAS cells and conjugated to proteins using carbodiimide crosslinker chemistry. Biological conjugates were built in *E. coli* strain CLM24 (W3110  $\Delta lpxM$ ,  $\Delta waaL$ ) through recombinant polyrhamnose and protein expression, conjugated by chromosomally encoded *Campylobacter jejuni* PglB oligosaccharyltransferase. Three chemical glycoconjugate vaccines were successfully generated; GAS protein antigen SpyAD conjugated to GAC as a ‘double hit’ GAS vaccine, and classical carrier protein Tetanus Toxoid (TT) conjugated to either GAC or GAS\_Rha. One biological glycoconjugate vaccine was also generated with an optimised *Streptococcal* carrier protein conjugated to recombinant polyrhamnose.

GAC and protein IgG antibody production was investigated by ELISA after three subcutaneous immunisations of BALB/c mice with the produced glycoconjugate vaccines. Anti-GAC antibody responses were significantly elevated in mice immunised with TT-GAC, but not with SpyAD-GAC, in comparison with the adjuvant only group. In contrast, the TT-GAS\_Rha chemical conjugate and *Streptococcus* specific biological conjugate, both containing rhamnose epitopes only, failed to induce anti-GAC antibody production above baseline. Meanwhile, SpyAD-GAC produced GAS protein specific opsonic antibodies, inducing significant opsonophagocytosis of GAS cells demonstrating that SpyAD remains immunogenic following conjugation. Although both methods successfully generated GAS glycoconjugates, in this study chemical linkage of wildtype GAC produced a more robust immune response than rhamnose variants conjugated either chemically or

using PgIB glycoengineering technology. Observations and technological advances developed in this thesis contribute to optimal GAS specific glycoconjugate vaccine design and production.

## Acknowledgments

First and foremost, I wish to express my wholehearted appreciation and gratitude to my supervisors Dr. Alexandra Shaw, Dr. Fatme Mawas, and Prof. Brendan Wren.

Thank you for your constant support, guidance, and encouragement throughout this 4-year rollercoaster journey. I am so appreciative for you giving me such an opportunity and believing in my capabilities. I am also so grateful Allie for your unwavering scientific expertise, random ideas and experimental plans, and general enthusiasm for science! Thank you all also for your valuable comments on drafts of this thesis.

I would also like to extend my gratitude to members of Bacteriology (NIBSC), both past and present, who have been wonderful colleagues, helping me grow both professionally and personally. Particularly I wish to thank members of B23 and B25, specifically Arif Felek for being the best coffee companion and physiochemical analysis wizard, as well as Martina Kristof who has been an absolute pillar of support, of which I don't think I would have survived without (and the copious number of blueberry muffins that we consumed!). I am so grateful for sharing my PhD journey with you.

Thank you also to everyone who has worked with me and collaborated on this project with a special acknowledgment to Dr. Helge Dorfmueller who kindly provided plasmid constructs for testing within the bioconjugation system. On this note, I would also like to thank the Wren group members who were so generous with their time, knowledge, and advice for which I am so grateful. Thank you also to NIBSC for the funding to undertake this project, as well as providing the best environment to develop into the scientist that I am today.

Finally, I would like to thank my friends for being around for a moan and always a glass of wine. They have all been a never-ending source of fun times, cheering me up and helping me see past the PhD in those more challenging times. Also thank you to my family and loved ones who have always shown me the love and faith I needed to believe in my capabilities, especially to my late nannan who always told me not to worry as 'you can do it, thas a Burnsey!'. Without all these people in my life this would have been near on impossible.

# Contents

Declaration.....	2
Abstract.....	3
Acknowledgments.....	5
List of Figures.....	10
List of Tables.....	12
List of Abbreviations.....	13
<b>CHAPTER 1</b> .....	16
Introduction.....	16
1.1. An introduction to Group A <i>Streptococcus</i> (GAS).....	16
1.2. GAS Infections and Disease.....	16
1.3. GAS Transmission.....	19
1.4. Epidemiology of GAS Infections.....	19
1.5. GAS Pathogenesis.....	22
1.6. Autoimmunity - GAS and Host Interplay.....	32
1.7. Host Response to GAS Infection.....	34
1.8. Control, Treatment and Prevention of GAS Infections.....	35
1.9. GAS Vaccine History.....	37
1.10. Glycoconjugate Vaccines.....	45
1.11. Research Aims, Objectives, and Hypotheses.....	49
<b>CHAPTER 2</b> .....	50
Materials and Methods.....	50
2.1. Materials.....	50
2.2. DNA Manipulation.....	52
2.3. Protein Techniques.....	56
2.4. Immunoassays.....	58
2.5. Quantification Techniques.....	59
2.6. Data Analysis and Presentation.....	60
<b>CHAPTER 3</b> .....	61
Use of Chemical Conjugation to Manufacture GAS Glycoconjugate Vaccines.....	61
3.1. Introduction.....	61
3.1.1. Group A Carbohydrate (GAC) Properties Related to Glycoconjugate Vaccine Production.....	61
3.1.2. Glycoconjugate Protein Carriers.....	62
3.1.3. Chemical Conjugation Approaches.....	63
3.2. Aim, Hypothesis and Objectives.....	68
3.3. Materials and Methods.....	69

3.3.1.	Bacterial Strains .....	69
3.3.2.	Bacterial Growth Conditions .....	69
3.3.3.	GAS Carrier Protein Selection, Cloning, Expression and Purification	69
3.3.4.	Preparation of Tetanus Toxoid (TT) Monomer .....	71
3.3.5.	$\Delta gacI$ Mutant Generation for Polyrhamnose Polysaccharide (GAS_Rha) Production .....	72
3.3.6.	Polysaccharide Extraction and Purification .....	72
3.3.7.	Polysaccharide Characterisation .....	75
3.3.8.	EDC Carbodiimide Conjugation .....	77
3.3.9.	Glycoconjugate Analytical Methods .....	78
3.3.10.	Immunogenicity Analysis.....	81
3.4.	Results .....	84
3.4.1.	Design and Analysis of GAS Specific Carrier Proteins.....	84
3.4.2.	Proposed Vaccine Components Are Immunogenic.....	90
3.4.3.	Extraction and Characterisation of Polysaccharide from GAS cells for Use in Chemical Conjugation Reactions .....	96
3.4.4.	Selection of EDC Conjugation Chemistry Approach .....	104
3.4.5.	Chemical Conjugation of TT Classical Protein Carrier and Wildtype GAC	107
3.4.6.	Chemical Conjugation of GAS Protein Antigens and Wildtype GAC	113
3.4.7.	Chemical Conjugation of TT Classical Protein Carrier and GAS_Rha	117
3.5.	Discussion .....	122
3.5.1.	Glycoconjugate Antigen Selection .....	122
3.5.2.	EDC Carbodiimide Conjugation Chemistry Suitability .....	126
3.6.	Summary .....	133
<b>CHAPTER 4</b>	.....	<b>134</b>
Use of Bioconjugation to Manufacture GAS Glycoconjugate Vaccines .....		134
4.1.	Introduction.....	134
4.1.1.	Discovery of Protein Glycan Coupling Technology (PGCT) .....	134
4.1.2.	<i>C. jejuni pgl</i> Operon Biosynthesis Mechanism and Expression.....	135
4.1.3.	Glycoconjugate Production using <i>C. jejuni</i> PglB Bioconjugation Inside <i>E. coli</i> Cells.....	137
4.1.4.	Recombinant Polyrhamnose Expression in <i>E. coli</i> Based on Group A Carbohydrate (GAC) and Homologs in <i>Streptococcus mutans</i> .....	138
4.2.	Aim, Hypothesis and Objectives .....	140
4.3.	Methods.....	141
4.3.1.	Bacterial Strains .....	141
4.3.2.	Bacterial Growth Conditions .....	143

4.3.3.	Cloning of GAS antigens in Commercial <i>E. coli</i> Strains .....	143
4.3.4.	Production of Electrocompetent Glycoengineering <i>E. coli</i> Strains for Polysaccharide Expression Testing and 'One Pot' <i>In Vitro</i> Testing .....	144
4.3.5.	Production of Multi-plasmid PGCT Systems via Electroporation Transformation of CLM24 <i>cedA::pglB E. coli</i> Cells.....	144
4.3.6.	Optimisation of Recombinant Rhamnose Expression .....	145
4.3.7.	Optimisation of GAS Antigen Protein Expression.....	147
4.3.8.	Preparation of Cell Free Lysate Extracts for Inclusion in 'One Pot' <i>In Vitro</i> Glycosylation Reactions .....	147
4.3.9.	'One Pot' <i>In Vitro</i> Glycosylation Assay.....	149
4.3.10.	Cell-Based Glycoprotein Production Inside <i>E. coli</i> Cells.....	149
4.3.11.	Glycoprotein Analysis.....	150
4.4.	Results .....	152
4.4.1.	Protein Carriers Used for Bioconjugation Approaches .....	152
4.4.2.	Polysaccharide Extracted From GAS Cells as Compatible Substrates Recognised and Transferred by PglB to Model Carrier Protein AcrA .....	154
4.4.3.	<i>E. coli</i> Expressed Recombinant Rhamnose for Use in Bioconjugation 156	
4.4.4.	'One Pot' <i>In Vitro</i> Reactions to Assemble Bioconjugates .....	166
4.4.5.	Cell-Based Glycosylation to Assemble Bioconjugates .....	168
4.5.	Discussion .....	175
4.5.1.	Successful Conjugation Testing of Proteins with Polysaccharide Variants using PglB OST .....	175
4.5.2.	Optimisation and Characterisation of Recombinant Rhamnose Polymers .....	177
4.5.3.	Suitability of Proteins as Bioconjugate Carriers Using Both Glycosylation Methods.....	178
4.5.4.	Characterising and Troubleshooting Bioconjugate Vaccine Production 182	
4.6.	Summary .....	183
<b>CHAPTER 5</b>	.....	185
Immunological Analysis of Chemical Glycoconjugate and Bioconjugate Vaccines		185
5.1.	Introduction.....	185
5.1.1.	Polysaccharide Antigens Relating to Glycoconjugate Immunogenicity 185	
5.1.2.	Protein Antigens Relating to Glycoconjugate Immunogenicity .....	186
5.1.3.	Manufacturing Methods Relating to Glycoconjugate Immunogenicity 187	
5.2.	Aim, Hypothesis and Objectives .....	189
5.3.	Methods.....	190



5.3.1.	Animals and Immunisations .....	190
5.3.2.	Immunoassays .....	190
5.4.	Results .....	192
5.4.1.	Antibody Titres Generated from Glycoconjugate Vaccine Immunisation 192	
5.4.2.	Function of Antibodies Generated from Glycoconjugate Vaccine Immunisation .....	203
5.5.	Discussion .....	207
5.5.1.	Immune Response to Chemical Glycoconjugate Vaccines – TT-GAC and SpyAD-GAC .....	207
5.5.2.	Immune Response to Chemical Glycoconjugate Vaccines – TT-GAC and TT-GAS_Rha .....	210
5.5.3.	Immune Response to Bioconjugate Vaccine – NanA-rpORha .....	212
5.5.4.	Additional Functional Immunological Analysis of Glycoconjugate Vaccines.....	214
5.5.5.	Vaccine Safety Testing .....	216
5.6.	Summary .....	217
<b>CHAPTER 6</b>	.....	<b>218</b>
Overall Discussion	.....	218
6.1.	Broad Themes and Thesis Scope.....	218
6.1.1.	Identifying Novel Immunogenic Antigens for Glycoconjugate Vaccine Inclusion.....	218
6.1.2.	Overall Glycoconjugate Structure and Immunogenicity from Different Manufacturing Methods .....	219
6.1.3.	Limitations of Glycoconjugate Production .....	221
6.2.	A Preferred Glycoconjugate Manufacturing Method.....	221
6.3.	Future Perspectives and Remaining Questions .....	223
6.3.1.	Current Limitations to Glycoconjugate Vaccine Analysis.....	223
6.3.2.	Glycoconjugate Antigen Inclusion Towards Clinical Efficacy Targets. 224	
6.3.3.	Alternative Approaches to Glycoconjugate Vaccines .....	225
6.4.	Concluding Remarks .....	227
References	.....	228

## List of Figures

- **Figure 1.1:** The *gac* operon in GAS produces Group A Carbohydrate (GAC) containing a polyrhamnose backbone decorated with alternate GlcNAc side chains.
- **Figure 1.2:** GAS cell surface virulence factor repertoire and M protein structure.
- **Figure 1.3:** Schematic representation of immunological processing and presentation of glycoconjugate vaccines by B cells to initiate T cell help effector functions.
- **Figure 1.4:** Summary of glycoconjugate vaccine antigen selection.
- **Figure 3.1:** Simplified schematic representation of chemical conjugation.
- **Figure 3.2:** Enzymatic extraction method of GAC and GAS\_Rha from wildtype M1 NCTC-8198 strain and associated mutant  $\Delta$ *gacl* GAS cells.
- **Figure 3.3:** Classical protein carrier Tetanus Toxoid (TT) and GAS protein carriers used for chemical conjugation reactions.
- **Figure 3.4:** Antibody response to GAS protein antigens in pooled human IgG (IVIG) and following subcutaneous immunisation of BALB/c mice.
- **Figure 3.5:** Surface exposure and recognition of GAS antigens by protein specific mouse antisera.
- **Figure 3.6:** Neutralising capability of anti-SpyCEP immune sera for IL-8 cleavage activity during incubation of GAS culture supernatants with recombinant IL-8 cytokine.
- **Figure 3.7:** Generation and phenotypic characterisation of *gacl* gene deletion in GAS M1 NCTC-8198.
- **Figure 3.8:** Analysis of extracted and purified wildtype GAC and mutant GAS\_Rha polysaccharides.
- **Figure 3.9:** Purified polysaccharide wildtype GAC and GAS\_Rha signals at UV<sub>214nm</sub> wavelength.
- **Figure 3.10:** Chemical conjugation strategy and validation of polysaccharide carboxylic group availability.
- **Figure 3.11:** Analysis of TT-GAC conjugation reaction.
- **Figure 3.12:** H-PAD polysaccharide quantification of purified TT-GAC conjugation reaction.
- **Figure 3.13:** Reverse Phase HPLC chromatograms of chemical glycoconjugate TT-GAC.
- **Figure 3.14:** Analysis of SpyAD-GAC conjugation reaction.
- **Figure 3.15:** HPLC-SEC chromatogram of chemical glycoconjugate SpyAD-GAC.
- **Figure 3.16:** Analysis of TT-GAS\_Rha conjugation reaction.
- **Figure 3.17:** HPLC-SEC chromatogram of chemical glycoconjugate TT-GAS\_Rha.
- **Figure 4.1:** The *pgl* locus in *Campylobacter jejuni* (NCTC 11168) produces a heptasaccharide for bioconjugation onto carrier proteins by PglB OST enzyme.
- **Figure 4.2:** Glycoconjugate production using *C. jejuni* PglB bioconjugation inside *E. coli* host cells.
- **Figure 4.3:** Expression, localisation, and *C. jejuni* heptasaccharide glycosylation of recombinant GAS carrier proteins using the 'one pot' *in vitro* bioconjugation system.
- **Figure 4.4:** 'One pot' *in vitro* glycosylation of wildtype (LLO-GAC) and mutant (LLO-GAS\_Rha) polysaccharides as PglB OST substrates for conjugation to AcrA carrier protein.

- **Figure 4.5:** Schematic representation of genetic loci and recombinant rhamnose structures encoded on plasmids pHD0249, pHD677 + pHD0689, and pHD0677 + pHD0256.
- **Figure 4.6:** CLM24 *E. coli* cells express recombinant rhamnose variants rpWTRha and rpMRha in different media and induction time points.
- **Figure 4.7:** Superior recombinant rhamnose polymer expression in CS2775 *E. coli* cells compared to CLM24 *E. coli* cells in different media.
- **Figure 4.8:** Purification of rpORha polymers isolated from CS2775 *E. coli* cell surfaces.
- **Figure 4.9:** 'One pot' *in vitro* glycosylation testing of rpORha for conjugation to AcrA carrier protein.
- **Figure 4.10:** 'One pot' *in vitro* glycosylation testing of rpORha for conjugation to GAS protein SpyAD.
- **Figure 4.11:** Expression, localisation, and *C. jejuni* heptasaccharide glycosylation of recombinant GAS carrier proteins using the cell-based bioconjugation system.
- **Figure 4.12:** Conjugation of *S. pneumoniae* NanA carrier protein with rpORha in chromosomal PglB *E. coli* strain, CLM24*cedA::pglB*.
- **Figure 5.1:** Subcutaneous immunisation schedule of BALB/c mice for GAS specific glycoconjugate immunogenicity study.
- **Figure 5.2:** ELISA analysis of anti-GAC IgG antibody titres in day 13, 27 and 42 antisera of mice immunised with wildtype GAC containing glycoconjugates.
- **Figure 5.3:** ELISA analysis of anti-GAC IgG antibody titres in day 13, 27 and 42 antisera of mice immunised with modified GAC containing glycoconjugates.
- **Figure 5.4:** ELISA analysis of anti-GAS\_Rha IgG antibody titres in day 13, 27 and 42 antisera of mice immunised with modified GAC containing glycoconjugates.
- **Figure 5.5:** ELISA analysis of anti-TT antibody titres in day 13, 27 and 42 antisera of mice immunised with TT containing glycoconjugates.
- **Figure 5.6:** ELISA analysis of anti-SpyAD antibody titres in day 13, 27 and 42 antisera of mice immunised with SpyAD containing glycoconjugate.
- **Figure 5.7:** Flow cytometry analysis of GAS M1 NCTC 8189 cells stained with mouse antisera against glycoconjugate vaccines.
- **Figure 5.8:** Glycoconjugate and protein antigen immune sera can induce neutrophil-mediated GAS killing assessed by a HL60-based opsonophagocytosis assay (OPA).

## List of Tables

- **Table 1.1:** Summary of GAS vaccine candidates currently in development.
- **Table 1.2:** GAS glycoconjugate immunisation strategies using GAC as a vaccine component.
- **Table 2.1:** List of antibodies and lectins used in this study.
- **Table 2.2:** Plasmids used in this study.
- **Table 2.3:** Oligonucleotides used in this study.
- **Table 3.1:** Choice of chemical conjugation strategies in licensed commercially available glycoconjugate vaccines.
- **Table 3.2:** List of *E. coli* and GAS strains used in chemical conjugation work.
- **Table 3.3:** H-PAD elution conditions.
- **Table 3.4:** RP-HPLC elution conditions.
- **Table 3.5:** GAS protein antigens selected for use as carriers for chemical conjugation.
- **Table 3.6:** Predicted immunological properties of selected GAS antigens.
- **Table 3.7:** Predicted physicochemical properties of selected GAS antigens.
- **Table 3.8:** Antibody responses to GAS protein antigens in pooled human IgG (IVIG) and IgG titres following subcutaneous immunisation of BALB/c mice.
- **Table 3.9:** SEC-MALS analysis of purified polysaccharide wildtype GAC and GAS\_Rha.
- **Table 3.10:** L-Lysine (K) availability as a measure of amine groups on GAS protein carriers and monomeric TT for targeted EDC chemical conjugation.
- **Table 3.11:** Quantification of chemical glycoconjugate vaccine components.
- **Table 3.12:** Total and conjugate ratios of chemical glycoconjugate vaccines.
- **Table 4.1:** List of *E. coli* and GAS strains used in bioconjugation work.
- **Table 4.2:** List of antibiotics used in transformed *E. coli* strains.
- **Table 4.3:** Genes encoding rpWTRha (pHD0249).
- **Table 4.4:** Genes encoding the modified acceptor stem (pHD0677), rpMRha (pHD0689) and rpORha (pHD0256).
- **Table 4.5:** Characteristics of NanA-rpORha bioconjugate.
- **Table 5.1:** GAS specific glycoconjugate vaccine groups and antigen components.

## List of Abbreviations

ACN	Acetonitrile
AD	Assay Diluent
ADI	Arginine Deiminase
Alum	Aluminium Hydroxide
APC	Antigen Presenting Cell
APSGN	Acute Poststreptococcal Glomerulonephritis
Ara	Arabinose
ARF	Acute Rheumatic Fever
Bac	Bacillosamine
BCA	Bicinchoninic Acid
BCR	B Cell Receptor
BHI	Brain Heart Infusion
BSA	Bovine Serum Albumin
CFU	Colony Forming Unit
CHIMS	Controlled Human Infection Models
CovR/S	Control Of Virulence
CTD	Carboxylic Terminal Domain
DALYs	Disability-Adjusted Life Years
DDM	N-Dodecyl- $\beta$ -D-Maltoside
DMF	Dimethylformamide
dNTP	Deoxyribonucleotide Triphosphate
DoA	Design of Analysis
DT	Diphtheria Toxoid
DTT	Dithiothreitol
ECM	Extracellular Matrix
EDC	1-Ethyl-3-(3-dimethylaminopropyl) Carbodiimide
EDTA	Ethylenediaminetetraacetic Acid
ELISA/spot	Enzyme-Linked Immunosorbent Assay / Spot
ExPEC	Extra-Intestinal Pathogenic <i>Escherichia coli</i>
FACs	Fluorescence-Activated Cell Sorting
FBS	Fetal Bovine Serum
FDA	Food and Drug Administration
GAC	Group A Carbohydrate
Gal	Galactose
GalNac	N-Acetylgalactosamine
GAS	Group A Streptococcus
GAS_Rha	Group A Streptococcus Rhamnose (Polysaccharide extracted from $\Delta$ <i>gacI</i> mutant cells)
GBS	Group B <i>Streptococcus</i>
Glc	Glucose
GlcNAc	N-acetylglucosamine
GMMA	Generalised Modules for Membrane Antigen
GMT	Geometric Mean Titre
GSK	GlaxoSmithKline
HA	Hyaluronic Acid
Hib	<i>Haemophilus influenzae</i> type B
HIC	High Income Country
HLA	Human Leukocyte Antigen
H-PAD	High-Performance Anion-Exchange Chromatography With Pulsed Amperometric Detection
HPLC-SEC	High Pressure Liquid Chromatography coupled with Size Exclusion Chromatography

HRP	Horseradish Peroxidase
HVR	Hyper Variable Region
IFN-	Interferon -
iGAS	Invasive GAS
IL-	Interleukin -
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
IVIG	Intravenous Immune Globulin
KDO	Ketodeoxyoctonic Acid
LAL	Limulus Amebocyte Lysate
LB	Lysogeny Broth
LCP	Lytr-CpsA-Psr Enzyme
LDS	Lithium Dodecyl Sulfate
LLO	Lipid Linked Oligosaccharide
LMIC	Low / Middle Income Country
LPS	Lipopolysaccharide
LTA	Lipoteichoic Acid
mAb	Monoclonal Antibody
MALS	Multi Angle Light Scattering
MES	2-(N-morpholino) Ethanesulfonic Acid
Mga	Multiple Gene Regulator
MHC	Major Histocompatibility Complex
<i>m/m</i>	Molar Ratio
$M_n$	Number Average Molecular Weight
MOPS	3-(N-morpholino) Propanesulfonic acid
MurNAc	N-acetylmuramic Acid
$M_w$	Molecular Weight
MWCO	Molecular Weight Cut Off
NETs	Neutrophil Extracellular Traps
NMR	Nucleic Magnetic Resonance
NTD	Amine (N) Terminal Domain
OPA	Opsonophagocytosis Assay
OST	Oligosaccaryltransferase Enzyme
PADRE	Pan HLA-DR-binding Epitope Peptide
PANDAS	Polyclonal Antibody Paediatric Autoimmune Neuropsychiatric Disorders
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline with Tween
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PGCT	Protein Glycan Coupling Technology
PHE	Public Health England
PBMC	Peripheral Blood Mononuclear Cell
PML	Polymorphonuclear Leucocytes
RALP	RofA-like Proteins
RGP	<i>S. mutans</i> Rhamnose-Containing Glucose Polymers
RHD	Rheumatic Heart Disease
rpMRha	<i>E. coli</i> Produced Recombinant Poly Modified Rhamnose (pHD0677 + pHD0689)
rpORha	<i>E. coli</i> Produced Recombinant Poly Optimised Rhamnose (pHD0677 + pHD0256)
rpWTRha	<i>E. coli</i> Produced Recombinant Poly Wildtype Rhamnose (pHD0249)
SBA	Soybean Agglutinin
SDS-PAGE	Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography

SIC	Streptococcal Inhibitor of Complement
SLO	Streptolysin O
SLS	Streptolysin S
SMEZ	Streptococcal Mitogenic Exotoxin Z
SNP	Single Nucleotide Polymorphisms
SOC	Super Optimal broth with Catabolite Repression
SOF	Streptococcal Serum Opacity Factor
Spe	Streptococcal Pyrogenic Exotoxins
SpyAD	<i>Streptococcus pyogenes</i> Adhesion and Division protein
SpyCEP	<i>Streptococcus pyogenes</i> Cell-Envelope Proteinase
SSA	Streptococcal Superantigen
STSS	Streptococcal Toxic Shock Syndrome
Sulfo-NHS	N-hydroxysulfosuccinimide
TBE	Tris-borate-EDTA
TCR	T Cell Receptor
TF	Trigger Factor
TFA	Trifluoroacetic Acid
TLR	Toll Like Receptor
TMB	3,3',5,5'-Tetramethylbenzidine
TNBSA	2,4,6-Trinitrobenzenesulfonic Acid
TNF $\alpha$	Tumour Necrosis Factor alpha
TT	Tetanus Toxoid
Und (P) P	Undecaprenol (Pyro) Phosphate
WCL	Whole Cell Lysates
WHO	The World Health Organisation
w/v	weight/volume
w/w	weight/weight
2-ME	2-Mercaptoethanol

# CHAPTER 1

## Introduction

### 1.1. An introduction to Group A *Streptococcus* (GAS)

Streptococci is a heterogenous group composed of over 50 bacterial species, some part of the normal human microbiota, and others classed as major human pathogens. From this genus, *Streptococcus pyogenes* also known as Group A *Streptococcus* (GAS) is one of the most important and common pathogens responsible for significant morbidity and mortality rates globally. GAS are beta-haemolytic Gram-positive chain forming cocci, adapted for transmission within humans. GAS is distinguished from other beta-haemolytic streptococci species by Lancefield serotyping, a classification system based on the unique cell wall anchored carbohydrate expression<sup>1</sup>. There are twenty Lancefield groups, A-U, excluding I and J groupings, with GAS named due to the presence of Group A Carbohydrate (GAC) on it's surface<sup>1</sup>. This differs from other relevant human pathogenic streptococci belonging to B (*S. agalactiae*), C (*S. equisimilis*, *S. equi*, *S. zooepidemicus*, *S. dysgalactiae*), and G (*S. anginosus*) groupings due to differences in monosaccharide composition and chain structure.

### 1.2. GAS Infections and Disease

GAS infections are responsible for a wide spectrum of disease, typically colonising throat and skin epithelial surfaces, with the ability to also invade cells resulting in deep tissue and bloodstream invasive infections. GAS colonisation can lead to a range of outcomes, from asymptomatic carriage through to severe disease mediated by not only colonisation and cellular invasion, but also toxin secretion and host immunological effects. The World Health Organisation (WHO) global disease burden figures rank GAS as the ninth leading cause of human mortality<sup>2</sup>, with a more recent report suggesting GAS is the fifth most lethal pathogen based on annual mortality<sup>3</sup>.

#### 1.2.1. *Mild Superficial GAS Infections*

GAS infections are often mild causing superficial throat and skin infections such as pharyngitis or impetigo respectively<sup>4,5</sup>. Pharyngitis, also known as Strep Throat, is often cleared by the host's immune system within a couple of weeks, or by antibiotic intervention<sup>4</sup>. GAS is the most common bacterial cause of pharyngitis, especially in children aged between 5 and 14<sup>6</sup> with around 600 million predicted cases per year<sup>7</sup>. Although prevalent globally, childhood pharyngitis incidence rates are 5 to 10 times



higher in low and middle income countries (LMICs) than in high income countries (HICs)<sup>7</sup>. Similarly, impetigo is also most prevalent in LMIC children<sup>8,9</sup> as well as in certain communities in HICs such as Māori and Pacific children in New Zealand<sup>10</sup>. Although not deadly themselves, pharyngitis and impetigo contribute hugely to morbidity, antibiotic usage, healthcare expenditure, as well as global economic and productivity losses<sup>11</sup>.

### *1.2.2. Autoimmune GAS Infections*

Repeated superficial infections can lead to the development of more serious autoimmune mediated GAS complications such as Acute Rheumatic Fever (ARF)<sup>12,13</sup>. ARF, characterised by heart and joint inflammation, is a major driver of morbidity and mortality worldwide, particularly when ARF progresses to causes permanent long-term heart damage through Rheumatic Heart Disease (RHD). RHD is the most common cause of paediatric heart disease worldwide<sup>14</sup>, as well as a major cause of maternal mortality in LMICs<sup>15–17</sup>. RHD is often described as a systemic disorder of economic disadvantage and social inequality<sup>18</sup>, disproportionately affecting LMIC children particularly within South Asia, central sub-Saharan Africa<sup>7,19</sup>, as well as indigenous populations in HICs such as Australia<sup>20</sup>. Although preventable and almost eliminated from some HICs through effective antibiotic intervention, RHD remains to cause on average 320,000 deaths annually, and is responsible for 10.5 million estimated disability-adjusted life years (DALYs)<sup>19</sup>.

Repeated GAS infections can also cause Acute Poststreptococcal Glomerulonephritis (APSGN) through immune complexes depositing on kidney glomerular membranes leading to oedema, urinary sediment abnormalities, and hypertension<sup>21</sup>. Similar to ARF/RHD, APSGN is most prevalent in LMICs, but outbreaks are often associated with “nephritogenic” GAS isolates<sup>6,22,23</sup> causing severe morbidity burdens<sup>24,25</sup>. Additionally, a less well characterised group of autoimmune complications following repeated superficial infections are neurological symptoms such as paediatric autoimmune neuropsychiatric disorders (PANDAS). Although clinical definitions are disputed, development is thought to be associated with anti-GAS antibodies recognising various host basal ganglia epitopes<sup>26</sup>.

### *1.2.3. Toxin Mediated Disease*

GAS infections are also associated with toxin mediated disease, common when infecting strains secrete high levels of bacteriophage-encoded streptococcal pyrogenic exotoxins (Spe)<sup>22,27,28</sup>. Toxin mediated disease affects people of all age groups globally, however, babies less than 1 year of age, people aged over 65, and

pregnant women are most susceptible<sup>29</sup>. Scarlet fever, the most common GAS toxin-mediated disease, in the 19th and early 20th centuries caused a significant burden of childhood morbidity and mortality<sup>30–32</sup>. Although rates have been steadily declining over the last 200 years, the United Kingdom, United States and Canada have all reported increased scarlet fever incidence<sup>33–35</sup>, highlighting that toxin mediated disease remains a significant health problem, even within HICs.

#### *1.2.4. Systemic Infections*

Although less common, GAS can also become invasive with a fifth of invasive GAS (iGAS) infections presenting as bloodstream infections<sup>29,36</sup>. Systemic disease is serious with poor clinical outcomes especially without intervention<sup>37–39</sup>. iGAS most commonly presents as bacteraemia and cellulitis, but can also cause necrotising fasciitis or Streptococcal Toxic Shock Syndrome (STSS) in response to superantigen production causing high fever, hypotension, and sometimes multiorgan failure<sup>40</sup>. There has been a global increase in iGAS infections since the late 1980s<sup>41–49</sup>, and as with all GAS infections, rates of iGAS disease are higher in LMICs<sup>50–53</sup> as well as indigenous Australian populations<sup>54,55</sup>.

Although some GAS related disease is falling or almost eradicated in HICs, infections persist globally, particularly in LMICs in both communicable and non-communicable disease forms. High burden and broad disease spectrums, coupled with increasing incidence rates, demonstrates the global need for an effective GAS preventative measure such as a vaccine<sup>56</sup>.

#### *1.2.5. Asymptomatic Nasopharyngeal Colonisation*

Given the broad disease spectrum GAS is not considered a commensal organism. However, GAS asymptomatic carriage was first described in the human nasopharynx during disease outbreaks in 1945<sup>57</sup>, and has more recently been observed in the throats of asymptomatic school children<sup>58–60</sup>.

Carriage is thought to occur when an individual carries GAS without contributing to infection<sup>61</sup>. The exact rate of carriage in a given population, and the role it plays in GAS pathogenesis is currently poorly defined, with a number of working hypotheses often based on observations of different pharyngeal flora between carriers and non-carriers<sup>62</sup>. These include poor GAS clearance from the nasopharynx due to  $\beta$ -lactamase producing pharyngeal flora<sup>62–66</sup>, or interactions between organisms which may have an inhibitory effect on GAS pharyngeal adhesion mechanisms<sup>67</sup>. There are also alternative theories relating to GAS bacteria themselves such as infection models and tonsil biopsies showing GAS internalisation into epithelial cells surviving

intracellularly<sup>68-71</sup>. GAS may also form biofilms increasing the antibiotic concentration required for successful killing<sup>72,73</sup>.

Distinguishing between carriage and infection is important for not only assessing transmission rates, but also recording GAS related disease, and implementing preventative measures. This is challenging with one study suggesting 66 – 75% of ARF individuals do not recall prior pharyngeal infection and may have in fact been asymptomatic carriers<sup>74,75</sup>. This suggests that GAS nasopharynx carriage may still act as a source and driver of host transmission as well as eventual disease.

### 1.3. GAS Transmission

GAS is transmitted by direct contact with saliva or nasal secretions from a colonised person, as well as skin contact following impetigo infection<sup>76,77</sup>. GAS maintenance and transmission is primarily mediated by two tissue sites, with GAS remaining viable for long periods of time in saliva<sup>78</sup> and on the skin<sup>79</sup>. GAS carriage makes the spread of infection and ultimately disease onset difficult to control, especially in overcrowded communities with poor sanitation<sup>6</sup>, and close living arrangements such as military camps<sup>18,80,81</sup>. In these instances, person to person transmission is more likely meaning outbreaks are common, with GAS often described as a community pathogen.

Larger GAS outbreaks can also occur due to the emergence and dominance of successful clones with increased transmissibility, and / or acquisition of toxins and antibiotic resistance determinants improving survival<sup>34,45,82,83</sup>. An example was a large outbreak of scarlet fever in Hong Kong and mainland China<sup>84-86</sup> where the dominant circulating strains had acquired tetracycline and macrolide resistance, as well as virulence factors such as streptococcal superantigen (SSA), and Streptococcal Pyrogenic Exotoxin C (SpeC) improving the isolates ability to induce toxin mediated disease<sup>34</sup>.

### 1.4. Epidemiology of GAS Infections

Despite clinical, epidemiological and laboratory research, no single strain or serotype has been attributed solely to specific disease. The exact mechanisms and selective pressures underpinning the broad disease phenotypes are not fully understood. There are however some general observations, with GAS skin infections most common in tropical LMICs and throat infections in industrialised HICs<sup>87</sup>. However, host genetic susceptibility, population immunity, environmental factors, and antibiotic access all influence disease progression and severity. This is

in addition to bacterial factors such as the circulating strains virulence factor repertoire as discussed in the context of increased scarlet fever incidence<sup>84–86</sup>.

#### 1.4.1. GAS Typing

Successful epidemiological studies rely on the accurate characterisation and typing of GAS strains. In 1928 Dr. Rebecca Lancefield proposed a system to identify and type GAS isolates using serotype-specific antiserum against the M protein<sup>88</sup>, as the immunodominant cell surface antigen and key virulence determinant<sup>89,90</sup>. Anti-M protein antibodies confer host protection and were later shown to be directed primarily against the hypervariable amine (N) terminal domain (NTD) of the mature protein. However, in the 1980s following increased iGAS incidence, the immunological serotype detection method could not successfully type all of the circulating isolates<sup>91–93</sup>. It was discovered that sequencing 10 – 15% of the 5' region of the *emm* gene encoding the M protein's NTD could reliably predict M serotypes<sup>94–97</sup>. Since adopting sequence typing, over 220 variants have been identified globally<sup>98</sup>, stated as "*emm* types", distinguished from serological "M types".

With such a wide variety of circulating *emm* types, more recently, strains have been functionally classified into 48 *emm* clusters based on the 5' *emm* sequence<sup>99</sup>. Clusters contain strains possessing closely related M proteins which share similarities in binding and structural properties. For example, one study showed that out of 175 different *emm* sequences, 143 *emm* types had closely related sequences and function, belonging to just 16 clusters<sup>99</sup>.

Additionally, to complement and enhance immunological and sequence M typing, additional GAS characterisation methods have been adopted, specifically T protein serotyping, and detection of streptococcal serum opacity factor (SOF)<sup>100</sup>. Currently there are ~ 20 recognised T types identified based on genes encoding the pili structure. With only a few exceptions, GAS isolates belonging to the same *emm* type often belong to the same T type<sup>101,102</sup>, however, this approach is yet to demonstrate any correlation between T type and specific GAS infection sites<sup>100</sup>. Typing can also be performed based on the heterogenous surface exposed and secreted virulence factor SOF. SOF typing was initially serological based on neutralising enzymatic activity<sup>103</sup>, but more recently, sequencing of the 5' end of the *sof* gene has also been used. Despite this, *emm* sequencing remains the gold standard of GAS typing used in epidemiological studies and surveillance of circulating strains.

#### 1.4.2. GAS Molecular Epidemiology

Large scale GAS typing has shown epidemiological differences in the geographical and socioeconomical distribution of *emm* types<sup>104–106</sup>. Although major infecting *emm* types have changed over time<sup>107</sup>, LMICs continue to display higher, more heterogenous *emm* type diversity compared to HICs. In HICs most circulating isolates belong to few *emm* types, notably 1, 3, 12 and 28, accounting for 40% of disease<sup>98,108–113</sup> and dominating asymptomatic carriage<sup>109</sup>. On the contrary, these *emm* types are rarely found in LMICs<sup>108,114–116</sup>. Interestingly, one study suggested that GAS isolates can be loosely divided into 3 country groupings; group 1 containing Canada, USA, Mexico, Western Europe, Korea and Japan, group 2 containing Eastern Europe and China, and group 3 containing predominantly LMICs: Brazil, Ethiopia, Israel, India, Nepal, Fiji in addition to HIC Australia. Group 3 compared to groups 1 and 2 have a higher diversity of *emm* types, and shared no predominant *emm* type between each of the countries within the group<sup>105</sup>. However, in some instances, epidemiological information is unavailable or incomplete meaning full GAS disease burden and association with different geographical regions is not always fully understood. Additionally, to complicate matters, *emm* strain distribution differences can be observed not only globally but within a single country, such as overcrowded impoverished slums showing higher strain diversity than wealthier areas in Brazil<sup>117</sup>.

Epidemiological patterns can also be observed through the association with certain M-protein types and particular GAS disease presentations. For example, increased ARF incidence in the 1980s in the USA was strongly associated with circulating M18 strain<sup>118</sup>, and M1 and M3 strains have been responsible for severe iGAS over the past two decades<sup>39,93</sup>. Moreover, M1 has disseminated globally, persisting as one of the most prevalent and frequently isolated serotypes from both iGAS and pharyngitis infections<sup>83,108</sup>. Additionally, 5 major chromosomal *emm* gene patterns have been associated with skin and throat infections. Gene patterns are determined by the number and arrangement of the *emm* subfamily genes situated in tandem near the Multiple Gene Regulator region (Mga)<sup>119</sup>. Generally, A – C chromosomal patterns relate to GAS isolates which cause pharyngitis, D causing impetigo, and E causing both pharyngitis and impetigo, with APSGN found among all chromosomal patterns. Typical M serotypes are also associated with these groups with A/D/E chromosomal patterns containing the most M types and C containing the least, with just one M type<sup>119–121</sup>.

## 1.5. GAS Pathogenesis

GAS as a well-adapted versatile pathogen possesses an array of colonisation factors, escape mechanisms, and survival strategies. GAS pathogenesis understanding has improved through the publication of a number of GAS genomes<sup>118,122–129</sup>. These have identified more than 40 putative virulence-associated genes, controlled by 13 independent regulatory mechanisms<sup>130</sup>. Virulence factor identification has accelerated understanding of the molecular mechanisms underpinning pathogenesis, divided into distinct pathogen-host events namely adhesion and colonisation, immune evasion, and invasive spread.

### 1.5.1. *Adhesion and Colonisation*

Prior to disease, GAS must attach to mucosal and / or cutaneous surfaces. GAS has a number of adhesins to facilitate attachment to the nasopharynx in the tonsil epithelium and dermal epithelial cells<sup>124</sup>. At least 17 adhesins have been identified, of which lipoteichoic acid (LTA), M protein, and fibronectin binding proteins are the most characterised<sup>131</sup>. Cell attachment is a complex event, simplified into a two-step process, specifically bacterial adhesion followed by colonisation, mediated by cell surface virulence components.

#### 1.5.1.1. *The Role of Capsule and Surface Polysaccharide in Host Cell Adhesion and Colonisation*

GAS cell surfaces are multifunctional, being protective by forming a cell barrier, as well as playing a role in virulence enabling initial attachment to colonise hosts. The GAS cell surface is composed of a hyaluronic acid (HA) capsule, the Group A Carbohydrate (GAC), surface exposed proteins covalently linked to the peptidoglycan layer and LTA (Figure 1.2a).

The HA capsule is a high molecular weight polymer made from N-Acetyl-Glucosamine (GlcNAc) and glucuronic acid surrounding the peptidoglycan layer. Unlike other prominent human streptococcal species, for example, Group B *Streptococcus* (GBS) and *Streptococcus pneumoniae*, GAS does not have an immunogenic exopolysaccharide, with the capsule being immunologically inert. The HA capsule rather enables host survival<sup>132</sup> and immune evasion<sup>133,134</sup>, forming similar structures to those found in humans<sup>135</sup>. GAS strains vary in HA expression, with encapsulated strains better than non-encapsulated strains at host colonisation<sup>132</sup>. High HA expression, however, may not be required for long-term GAS colonisation, with some strains isolated from asymptomatic individuals possessing a frameshift mutation resulting in reduced or abolished capsule production<sup>136</sup>.

Unlike the HA capsule, the expression of GAC is conserved across all GAS strain cell surfaces<sup>1,137–140</sup> and is essential to GAS bacterial survival as demonstrated by plasmid integrational mutagenesis studies<sup>141</sup>. It provides structural support as an environmental barrier, as well as playing a role in cell morphology, division<sup>142</sup>, and pathogenesis<sup>139,141,143</sup>. GAC is a conformationally restricted polysaccharide<sup>144</sup> with an average polymer molecular mass of 8.9 +/- 1.0 kDa, corresponding to 18 repeating units<sup>139</sup> made up of a linear polyrhamnose backbone with alternating GlcNAc sidechains with  $\alpha$ -L-(1→2) and  $\alpha$ -L-(1→3) glycosidic linkages at position 3, with a trisaccharide repeating unit of [3 $\alpha$ -L-Rhap(1→2)[ $\beta$ -D-GlcpNAc(1→3)] $\alpha$ -L-Rhap(1→3)]<sub>n</sub> (Figure 1.1c)<sup>139,145</sup>. GlcNAc is attached to every other rhamnose residue extending out to the periphery from the rhamnose helix core<sup>146</sup>.

All human GAS clinical isolates have GAC containing GlcNAc sidechains, but following serial passage in mice and rabbits, the GlcNAc sidechain can be lost, resulting in an “A variant” strain<sup>147</sup>. GlcNAc residues are immunodominant with their experimental removal resulting in reduced virulence of M1T1 strain in mouse and rabbit pneumonia infection models<sup>141</sup>. Additionally, the variant strain was more susceptible to killing by human whole blood and purified human neutrophils due to greater binding of cationic human defence peptides, specifically cathelicidin LL-37, to the mutant polysaccharide<sup>141</sup>.

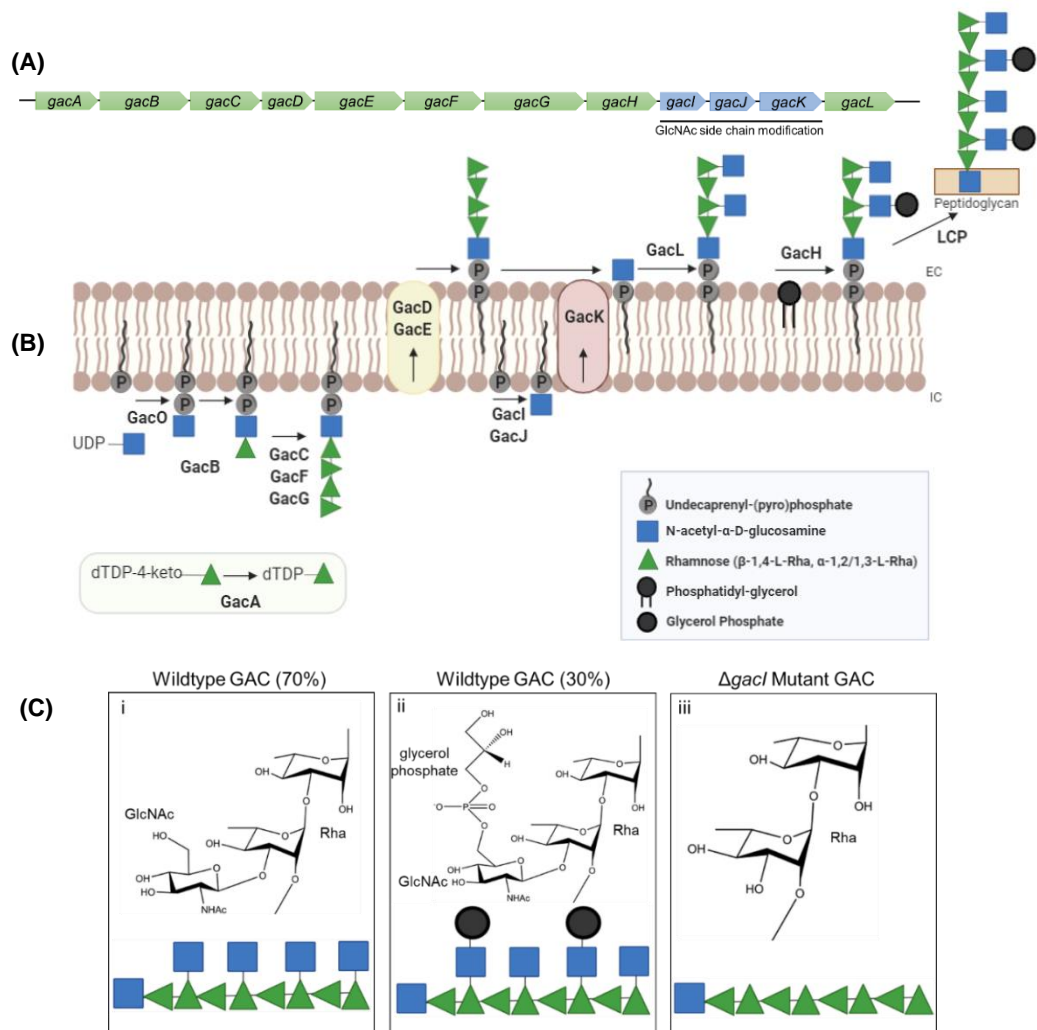
Approximately 25 - 30% of GAC GlcNAc residues contain glycerol phosphate (GroP) modifications, specifically on the C6-hydroxyl group. This polysaccharide modification has remained elusive for some time, most likely due to harsh acidic extraction methods used to purify GAC, which can affect polymer composition<sup>148</sup>. The function of GroP has not been fully characterised, however, it may enable immune evasion, specifically through resistance to zinc toxicity, a killing mechanism deployed by neutrophils<sup>149</sup>.

GAC polymers are encoded by a conserved 12 gene cluster termed *gacA-L* (*gac* operon)<sup>141,150</sup> (Figure 1.1a). The operon is highly conserved with one study finding that 2,017 of 2,083 tested GAS genomes had > 70% DNA sequence identity for the entire 12 gene cluster<sup>150</sup> supporting observations from a smaller dataset<sup>141</sup>. The latter study, performing a plasmid integrational mutagenesis scan across cluster showed that the first seven genes (*gacA-G*) enabling polyrhamnose biosynthesis, are conserved across other streptococci groups, specifically A, B, C and G<sup>151</sup>. These genes and therefore L-Rhamnose substrate availability are essential to GAS survival<sup>141,142</sup>, whereas *gacI-K* genes, implicated in GlcNAc sidechain attachment

and modification<sup>141</sup>, are believed to not be required for GAS when propagated in Todd-Hewitt liquid culture, but results in the loss of GAS latex agglutination reactivity<sup>143,150</sup>.

Gene identification and functional enzymatic characterisation has enabled a working model explaining GAC precursor biosynthesis, assembly, and transport to the extracellular space (Figure 1.1b). Firstly, GlcNAc is attached to undecaprenyl phosphate (Und P) by GacO, initiating GAC biosynthesis on the intracellular side of the cell membrane<sup>152</sup>. Translocation of the first rhamnose residue to the membrane-bound UndP-GlcNAc in the cell membrane inner leaflet then follows, as a committed step to GAC biosynthesis, through GacB, a rhamnosyltransferase<sup>151</sup>. From  $\alpha$ -glucose-1-phosphate, free dTDP-L-rhamnose is synthesised by GacA, as a metal-independent dTDP-4-dehydrorhamnose reductase enzyme, as well as the *rmIABC* operon located distally from the GAC operon<sup>142</sup>. GacC, GacF, and GacG glycosyltransferases then sequentially elongate the polyrhamnose chain<sup>151</sup> which once completed is translocated across the membrane to the extracellular side, by an ATP-dependent ABC transporter encoded by GacD and GacE functioning as a heterodimer. The remaining genes are predicted to encode rhamnose chain modifications on GAS cell surfaces. GacJ, a small membrane associated protein interacts with GacI to synthesise free UndP-GlcNAc<sup>151</sup>, which can either diffuse across the cytoplasmic membrane or alternatively be transported by GacK encoding a Wzx family flippase enzyme to the extracellular side of the membrane<sup>145</sup>. GacL as a putative glycosyltransferase then attaches UndP-GlcNAc to the polyrhamnose backbone<sup>145</sup>. This GlcNAc may be subsequently modified by GacH cleaving phosphatidylglycerol to release and attach glycerol phosphate<sup>149</sup>. Once fully synthesised the final step is for the complete GAC polymers to be transferred and covalently linked to the peptidoglycan layer via a phosphodiester bond by Lyt-CpsA-Psr (LCP) enzyme.





**Figure 1.1: The *gac* operon in GAS produces Group A Carbohydrate (GAC) containing a polyrhamnose backbone decorated with alternate GlcNAc side chains.**

**(A)** Schematic representation of the *gac* operon (*gacA-L*) in GAS. Horizontal arrows represent each gene designation with colour denoting predicted gene function. Green, polyrhamnose biosynthesis; blue, GlcNAc biosynthesis. Adapted from Sorge *et al*<sup>141</sup>.

**(B)** Schematic diagram of GAC biosynthesis. GAC biosynthesis is initiated on lipid linked GlcNAc attached to the inner leaflet of the periplasmic membrane for polyrhamnose synthesis catalysed by rhamnosyltransferase enzymes (GacBCFG). After polymerisation, the polyrhamnose backbone is flipped to the outer leaflet by an ABC transporter (GacDE complex) before GlcNAc (GacL) and glycerol phosphates (GacH) are transferred to the polyrhamnose backbone as a sidechain modification. A LytR-CpsS-Psr (LCP) phosphotransferase protein attaches GAC to peptidoglycan via a phosphodiester bond. Adapted from Rush *et al*<sup>145</sup>.

**(C)** GAC trisaccharide repeating unit [3]α-L-Rhap(1→2)[β-D-GlcpNAc(1→3)]α-L-Rhap constituting 70% wildtype (WT) GAC (i), and 30% glycerol phosphate modified GAC (ii) [3]-α-L-Rhap-(1 → 2)[β-d-GlcpNAc6P(S)Gro-(1 → 3)]-α-L-Rhap-(1 → 3)].  $\Delta gacI$  mutants and certain "A variant" GAS strains have the repeating structure [3)-α-L-Rhap-( 1 → 2)α-L-Rhap-( 1 → 3)-α] deficient in GlcNAc sidechains (iii). Original figure created with BioRender.com and Chemdraw (RRID:SCR\_016768).

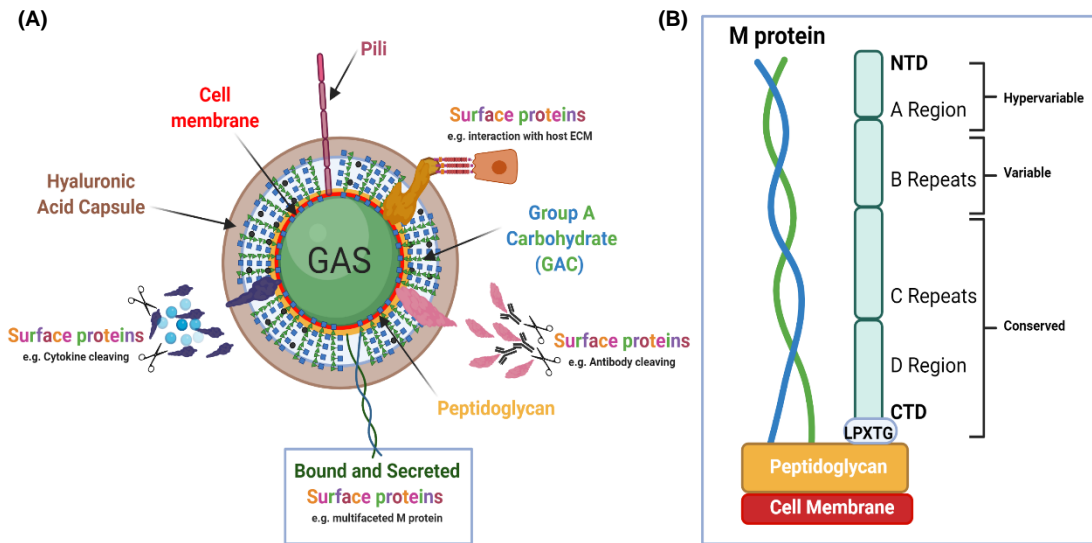
#### 1.5.1.2. *The Role of Surface Exposed Proteins in Host Cell Adhesion and Colonisation*

GAS cell surfaces are decorated with many proteins facilitating different functions. Most proteins are anchored to the cell wall peptidoglycan by sortase enzymes<sup>153–155</sup>, categorised into housekeeping sortase SrtA, attaching proteins via their LPXTG motif<sup>153</sup> or the pilus-associated sortases<sup>155</sup>. There are also a small number of lipid-anchored putative surface proteins<sup>156</sup>. The first step of GAS adhesion is a weak long-range attachment to pharyngeal or dermal epithelial cells via hydrophobic interactions. This is primarily mediated by LTA attaching to host cell extracellular matrix (ECM) component fibronectin to bring GAS in close contact<sup>157</sup> before high affinity binding occurs via lectin-carbohydrate and protein-protein interactions<sup>158</sup>. There are multiple ECM targets including fibronectin, collagen, laminin and vitronectin, with multiple GAS surface exposed proteins displaying functional redundancy, explaining adhesin repertoire and expression strain diversity.

Pili enable cell adhesion as long flexible rods which extend up to 3 µm from GAS cell surfaces<sup>159</sup>. They are made up of pilin subunits, which polymerise via transpeptidation reactions and are decorated with one or more ancillary protein subunits (Ap1 / Cpa and AP2), forming mature pilus structures<sup>160–162</sup>. T type 1 and 2 pili are most characterised, showing adhesion to a variety of epithelial cells using fibronectin via the pilus tip<sup>163–165</sup>. T type 3 pili show differences binding to type I collagen<sup>166,167</sup> using the Cpa ancillary protein subunit at the base of the structure<sup>168</sup>.

The most abundant and characterised adhesin are M proteins, covering the entire GAS cell surface through LPXTG attachment to the peptidoglycan layer<sup>89</sup>(Figure 1.2b). M proteins are dimeric alpha helical, coiled coil, fibrillar proteins which can be divided into three regions, the amine (N) terminal domain (NTD) distal to the cell wall extending into the environment<sup>4</sup>, the conserved carboxylic terminal domain (CTD) anchored to the cell wall, and the central domain<sup>89</sup>. The central region can be further sub-divided into A, B, C and D regions (Figure 1.2b), differing in size and number in different GAS strains<sup>169</sup>. Region A from the central part along with approximately 11 amino acids from the NTD constitutes as the hypervariable region (HVR) forming the basis of the Lancefield serological and *emm* sequence typing classification systems as described previously.

M proteins are multifaceted, aiding bacterial colonisation in the early stages of infection through high affinity adhesion and persistence<sup>170</sup>, as well as inhibiting phagocytosis as an immune evasion strategy. Proteins which structurally resemble M proteins are also present on GAS cell surfaces, encoded by genes belonging to the *emm* superfamily. Such M-like proteins are similar to M proteins, enabling host cell adhesion through LTA interactions, maintaining stability when colonising the nasopharynx<sup>158,171</sup>.



**Figure 1.2: GAS cell surface virulence factor repertoire and M protein structure.**

**(A)** Schematic of a GAS cell surface. Hyaluronic acid (HA) capsule in brown, peptidoglycan (cell wall) in yellow, and cell membrane in red. GAC polysaccharide polymers (blue / green) and surface exposed proteins (multicoloured) are attached to the cell wall surrounding the cell. Surface bound and secreted protein function inferred such as cytokine and immunoglobulin cleavage and adhesion to host epithelial cells via ECM components. Boxed structure shows M protein attached to the cell wall.

**(B)** Model of a M protein alpha helical dimer. One M protein dimer is in green and the other dimer is in blue. The NTD and CTD domains and central region (A, B, C and D) are shown.

Original figure created with BioRender.com.

The next most abundant class of surface exposed adhesins recognising ECM components are those that bind to fibronectin. Fibronectin is the most common and key host extracellular target due to the affinity between GAS and host integrin receptors on epithelial cells. GAS possesses at least 11 fibronectin binding adhesins<sup>172</sup>. PrtF1<sup>173</sup>, also known as SfbI (Streptococcal Fibronectin Binding Protein)<sup>174</sup>, is the most characterised adhering to epithelial cells<sup>175</sup> as well as dermal

cells<sup>176</sup>. PrtF2 is another common fibronectin binding protein displaying high affinity to fibronectin<sup>177</sup>, as well as fibrinogen<sup>177,178</sup>. Other fibronectin binding adhesins include PFBP<sup>179</sup>, Fbp54<sup>180</sup>, and SfbX<sup>181</sup> belonging to one of two families depending on shared structural motifs<sup>182,183</sup>.

Additionally, there is a different protein family known as collagen-like proteins, specifically Scl1 and Scl2 containing Gly-X-X motifs within a triple-helical elongated structure<sup>184</sup>. Such proteins contain a globular V domain and are homotrimeric, extending out from the cell surface using a rod-shaped CL domain to allow host cell interactions<sup>184</sup>. Scl1 proteins are most characterised, binding to a variety of host components such as  $\alpha_2\beta_1$  and  $\alpha_{11}\beta_1$  integrins on fibroblasts, endothelial, and epithelial cells to promote cell internalisation<sup>185,186</sup>, increasing intracellular survival and re-emergence<sup>187,188</sup>. Binding to host cell fibronectin and laminin is also possible with these proteins to facilitate biofilm formation as an immune evasion strategy<sup>189,190</sup>.

### 1.5.2. Immune Evasion

GAS dissemination and invasive disease is attributable to avoidance of host immune clearance mechanisms. GAS possesses a number of virulence factors to prevent phagocytosis, namely but not exclusive to M proteins, M-like proteins, and the HA capsule. GAS immune evasion proteins often prevent host complement mediated attack by binding factor H, a complement control protein involved in regulation<sup>191</sup>. For example, the M protein induces antiphagocytic effects through factor H binding, interfering with opsonisation via the alternative complement pathway, through impairing C3b labelling onto GAS surfaces, reducing polymorphonuclear leucocytes (PML) recognition and bacterial clearance<sup>192</sup>. Excessive HA capsule in highly mucoid strains (e.g. M18 and M24) also enables effective phagocytosis resistance<sup>132</sup>, through HA anionic properties reducing C3b opsonisation<sup>4</sup>. Secreted cysteine proteases, such as IdeS (Mac-1 / MspA), also reduces phagocytosis through binding to PML surfaces via CD16<sup>193</sup>.

GAS also avoids host immune responses through proteins which degrade immunoglobulins to reduce direct and complement mediated opsonisation (Figure 1.2a). M proteins can bind to the Fc part of IgG and IgA immunoglobulins as well as to C4b directly in a *emm* type specific fashion via the hypervariable region<sup>194-196</sup>. M-like proteins, such as Mrp and Enn, also carry out antiphagocytic effects by binding IgG or IgA<sup>197</sup> cooperating with M protein function to interfere with immunoglobulin- and complement- mediated opsonisation<sup>198</sup>. This is in addition to IdeS which can

degrade IgG antibodies<sup>199</sup>, and other virulence factors such streptococcal inhibitor of complement (SIC) protein, binding to C5b complexes to prevent terminal complement membrane attack complex formation<sup>200</sup>.

Other surface bound or secreted virulence factors also contribute to immune evasion strategies by interacting with chemokines and cytokines (Figure 1.2a). Conserved surface exposed C5a peptidase, a large multi-domain protein, inactivates C5a complement from PML surfaces, preventing phagocytosis and complement-mediated chemotaxis<sup>201,202</sup>. GAS also secretes *Streptococcus pyogenes* Cell-Envelope Proteinase (SpyCEP, ScpC), which cleaves CXC chemokines such as interleukin (IL) -8 reducing neutrophil recruitment<sup>203</sup>, as well as secreted cysteine proteases, such as streptococcal pyrogenic exotoxin B (SpeB) which converts IL-1 $\beta$  into an active molecule<sup>204</sup>, generating active peptides such as kinins<sup>205</sup>, and histamine<sup>206</sup>, as well as degrading bacterial surface bound IgG<sup>207</sup>. Some secreted virulence factors, such as SpyCEP, are also bound to the GAS cell surface<sup>208</sup>.

GAS can also induce cytolysin activity to destroy host cells and neutrophil extracellular traps (NETs), and form biofilms to reduce effective clearance. Streptolysin O (SLO) pore-forming toxin promotes phagocytosis resistance, suppressing neutrophil oxidative burst, and impairing neutrophil migration to enable bloodstream survival<sup>209</sup>. A switch in bacterial lifestyle to form biofilms is an additional method, impeding not only host clearance mechanisms but also antibiotic effectiveness. In addition to mediating adhesion, T type 2 pili play a role in bacterial auto aggregation<sup>210</sup> to increase host survival<sup>211,212</sup>. Multifaceted M proteins may also be important in long-term GAS colonisation and survival, enabling aggregate formation on the tonsillar epithelia, and promoting biofilm formation<sup>59,213–215</sup>.

### 1.5.3. Systemic Toxicity and GAS Bacterial Dissemination

GAS can cause systemic toxicity mediated by virulence factors directly interacting and activating immune cells. Spes and SSA are important virulence factors associated with iGAS disease. One key mechanism in STSS onset is the non-specific activation of T cells by SSA leading to Th1 proinflammatory responses<sup>216,217</sup>. There are also many Spe's used by circulating GAS strains<sup>84</sup> such as SpeA<sup>218</sup>, SpeC<sup>219</sup>, SpeH<sup>220</sup> and Streptococcal Mitogenic Exotoxin Z (SMEZ)<sup>221</sup> which enhance pro-inflammatory host responses. Secreted extracellular toxin superantigens, SpeA and SpeC, are historically known as scarlatinal toxins, responsible for increased scarlet fever incidence<sup>222–224</sup>.

SSA structures possess both distinct T-cell receptor (TCR) and Major Histocompatibility Complex (MHC) class II binding sites<sup>225</sup>, allowing SSAs to act as immunostimulatory molecules, binding outside of the MHC class II peptide-binding site, leading to overstimulation and activation of a large number of T cells<sup>226,227</sup>. This has important downstream effects through the release of large amounts of inflammatory cytokines such as Tumour Necrosis Factor alpha (TNF- $\alpha$ ), IL-1 $\beta$ , and T cell mediators, IL-2 and Interferon (INF) gamma (INF- $\gamma$ ), causing complement, coagulation, and fibrinolytic cascades, which can lead to hypotension and multiorgan failure characteristic of iGAS infections.

GAS can also cause invasive disease through other secreted virulence factors, such as SLO, a cholesterol dependent oxygen-labile cytolysin, prevalent in iGAS associated isolates<sup>228</sup>. SLO's main function is to interact with cholesterol in target cells forming multi-subunit pores, lysing erythrocytes, leukocytes, macrophages, and platelets<sup>229,230</sup>. Cell lysis provides a bacterial survival advantage through binding to released haemoglobin which can then be used as an iron source, and mediating tissue destruction important in necrotising fasciitis and bacterial dissemination<sup>231</sup>.

Streptolysin S (SLS) is an additional cytolysin also implicated in iGAS. SLS expression gives GAS the characteristic  $\beta$ -haemolytic phenotype when cultured on blood agar. Similar to SLO, SLS exhibits cytotoxic effects against epithelial cells, neutrophils, lymphocytes, and platelets in intracellular, cell-surface-bound, and extracellular forms<sup>232</sup>. This specifically occurs by activation of cellular calpain, a host cysteine protease which enables degradation of E-cadherin leading to mucosal epithelium disruption<sup>232</sup>. SLS therefore contributes to GAS translocation across epithelial cell monolayers, which can lead to lesion formation and tissue destruction<sup>233</sup>.

Systemic toxicity is associated with widespread infection when GAS cells reach local and distant lymph nodes as an extracellular pathogen<sup>234,235</sup>. GAS is able to access the bloodstream through the lymphatic system, remaining extracellular in transit, important in bacterial dissemination characteristic of iGAS<sup>234</sup>. In addition to travelling to sites distant from initial colonisation, GAS can also become adapted to intracellular survival within non-phagocytic epithelial cells<sup>236-238</sup>. Intracellular GAS has been observed within surgically excised tonsils from asymptomatic patients<sup>68,69</sup> suggesting intracellular niches to be an important source for recurrent infections and carriage<sup>69,239</sup>. GAS cell internalisation provides a survival advantage, impeding host clearance and enabling bacterial persistence. This is through damage to the

epithelial integrity resulting in cell death and exposure of underlying tissues, providing an intracellular route to deeper tissues. Similar to bacterial adherence, internalisation likely requires multiple steps including interaction between adhesion proteins and host integrins<sup>237,240</sup>, followed by host cytoskeleton rearrangements<sup>172</sup>.

A key GAS virulence factor believed to be involved in internalisation is the HA capsule<sup>241</sup>, binding CD44 glycoproteins on epithelial cell surfaces to induce cytoskeletal rearrangements<sup>242,243</sup>. Cell signalling disrupts intercellular junctions, causing membrane ruffling and epithelium penetration<sup>241</sup>. M proteins<sup>237,244</sup>, fibronectin binding proteins such as Sfb<sup>245</sup> and Fba<sup>246</sup>, in addition to toxins and proteases, such as SLS<sup>247</sup>, Streptokinase, and SpeB have also all been implicated in GAS internalisation. This is through fibronectin binding triggering plasma membrane invaginations and aggregation aiding GAS internalisation. In addition to epithelial cells, GAS survival has also been demonstrated within phagocytic cells, such as macrophages<sup>248</sup>. GAS can escape phagolysosomes enabling multiplication in the hosts cytoplasm<sup>249</sup>, thought to be mediated by SLO and its co-toxin expression<sup>250</sup>. This aids intracellular survival as well as resistance to macrophage killing as an innate immune evasion strategy.

#### *1.5.4. Control of Virulence*

The ability to control virulence factor expression during GAS infection is important for bacterial survival. GAS systems respond to environmental signals to control gene products involved in adhesion, colonisation and invasion<sup>251,252</sup>. Some GAS genes encoding virulence factors are under control by two-component regulatory systems composed of a membrane bound sensor and a cytoplasmic response regulator. One of the most characterised and key regulatory systems is the CovR/S system (Control Of Virulence; also known as csrRS, Capsule Synthesis Regulator)<sup>253</sup>, with CovS as the sensor protein and CovR as the repressor molecule. Microarray expression technologies predict that CovR/S controls the transcription of 15% of all chromosomal GAS genes<sup>254</sup>, encoding proteins involved in host-pathogen interactions, enabling differential expression during GAS growth, and response to environmental stress<sup>255</sup>. CovR/S has also been shown to repress the expression of a number of virulence factor genes including HA capsule (*hasABC*), DNase, *sda*, and SpyCEP<sup>203,254,256,257</sup>.

Another well studied transcriptional regulator is the Multiple Gene Regulator (Mga) responding to host cellular environments and making up 10 % of the GAS genome<sup>258</sup>. Mga controls genes important in adhesion and homeostasis<sup>259</sup> as a DNA

binding protein<sup>260</sup> controlling its own expression<sup>261</sup> in addition to the expression of the M protein family (*emm*, *mrp*, *enn* and *arp*), C5a peptidase (*scpA*), SOF (*sof*), SIC (*sic*), and Scl1 (*sclA*)<sup>262</sup>.

Additional gene expression regulators include RofA-like Proteins (RALPs), specifically RofA<sup>263</sup> controlling PrtF1 fibronectin binding adhesin expression in response to oxygen concentration<sup>264</sup>. Other regulators of note include Rgg, known as RofB, as a global regulator of several genes and other regulators such as the Mga and CovR/S systems<sup>265</sup>, as well as FasBCA three-component system controlling expression of certain fibrinogen and fibronectin binding proteins<sup>266</sup>.

## 1.6. Autoimmunity - GAS and Host Interplay

Since the 1970s a link has been observed between GAS cell components generating anti-GAS antibodies leading to ARF development<sup>267,268</sup>. Initial experiments demonstrated that whole GAS cells, cell walls and membranes could absorb human antibodies which recognise the heart<sup>269-271</sup>. Additionally, rabbit sera from GAS immunisation and ARF patient sera both react with heart and skeletal myosin<sup>272</sup>. Studies in the 1980s determined shared epitopes between GAS directed monoclonal antibodies (mAbs) and murine and human tissues<sup>273-275</sup>.

Immunofluorescent staining of heart tissue showed cross reactive mAbs recognised the heavy chain of skeletal and cardiac host myosin<sup>276-278</sup>, in addition to GAS M proteins<sup>275,279-281</sup> and the GAS cell surface carbohydrate, specifically GAC GlcNAc residues<sup>273-275,282</sup>. Additional host structures have also been identified including keratin in skin, tropomyosin in cardiac muscle<sup>283,284</sup>, and laminin and vimentin in the ECM of heart valves<sup>285,286</sup>.

### 1.6.1. *GAS Autoimmunity – M proteins*

One of the main epitopes recognised by GAS specific antibodies are  $\alpha$ -helical coiled-coil molecules, present in host cardiac and skeletal myosin, tropomyosin, vimentin, laminin and keratin<sup>287</sup>. GAS M proteins are also  $\alpha$ -helical coiled-coil molecules<sup>288,289</sup>, and are the best studied GAS virulence factor implicated in autoimmune sequelae. There are seven residues shared between M proteins and host ECM components forming the structural molecular mimicry, in addition to some shared amino acid sequence homology to human skeletal and cardiac myosin<sup>290</sup>. Primarily, M proteins induce autoimmune effects through the generation of antibodies, however cytotoxic T cell populations which recognise myosin have also been reported in ARF sera<sup>281,291</sup> and patient RHD heart tissue<sup>292</sup>. M5 protein is most



characterised for ARF association<sup>293</sup>, with mapped autoimmune epitopes<sup>294,295</sup>, and generating T cell stimulation in animal models<sup>296</sup>.

### *1.6.2. GAS Autoimmunity – GlcNAc Residues on GAC*

A link has also been demonstrated between ARF patient anti-myosin and laminin mAbs and GlcNAc residues on GAC<sup>297,298</sup>. Clinically high anti-GlcNAc mAb titres and host myosin correlate with chronic rheumatic valvulitis, an ARF complication<sup>299</sup>, as well as longer lived anti-GAC antibody populations in RHD patients<sup>299</sup>.

Specifically, anti-GlcNAc antibodies have been found deposited on the basement membrane underlying the heart valve endothelium<sup>283</sup> and the valve surface<sup>292</sup> within RHD patients<sup>270</sup>. This is mediated by increased vascular cell adhesion molecule-1 (VCAM-1) expression<sup>300</sup> and CXC-9 CXC chemokine upregulation<sup>301</sup>, leading to CD8<sup>+</sup> and CD4<sup>+</sup> T cell and macrophage extravasation to valvular tissues. This results in imbalanced Th2 type immune responses often observed in chronic RHD<sup>292,301–304</sup>.

### *1.6.3. GAS Autoimmunity – Other Antigens*

Additional GAS cell surface epitopes have also been identified using affinity purified anti-myosin antibodies present in ARF sera. These include a 60 kDa streptococcal actin-like protein<sup>305</sup>, a 60 kDa wall membrane antigen<sup>282</sup> and a 67 kDa protein antigen<sup>306</sup>. Cross reactive antibodies also recognise heart myosin and actin, due to shared conformational properties with M proteins, or MHC class II molecules in the case of the 67 kDa protein<sup>306</sup>.

### *1.6.4. Mechanism of Heart Tissue Damage*

GAS bacterial factors, host molecular mimicry, and exaggerated and unregulated host immune responses are fundamentally responsible for autoimmune sequelae. GAS antigen processing and presentation through B and T cell interaction leads to stimulation of cross reactive cytotoxic mAbs in the presence of complement<sup>297,307</sup> causing inflammation and tissue damage resulting in carditis<sup>308,309</sup>. The main driver is excessive and unbalanced cytokine release by peripheral blood mononuclear cells (PBMCs) and antigen presenting cells (APCs) such as macrophages at infection sites<sup>310–312</sup> as well as presence of heart targeting T cells<sup>313,314</sup>. Cytokine elevation of TNF- $\alpha$ <sup>313,314</sup>, IL-10<sup>303,311,315</sup>, IL-6<sup>303,311</sup>, IL-17<sup>313</sup>, IL-23 and IFN- $\gamma$ <sup>316</sup>, are implicated in ARF / RHD pathogenesis. Specifically, IL-10 and IL-6 are potent B cell activators leading to increased survival, differentiation, and proliferation, which exacerbates autoantibody production. Skewed cytokine profiles also lead to more PBMC cell infiltration, activation of endothelial cells, increased vascular

permeability, and increased adhesion molecule expression, resulting in tissue damage providing an additional trigger for further T cell and macrophage activity<sup>303,316</sup>.

## 1.7. Host Response to GAS Infection

### 1.7.1. *Host Response to Repeated Infection Inducing Autoimmunity*

Relatively few repeated GAS infections lead to autoimmunological disease<sup>317</sup> therefore host susceptibility likely influences autoimmune sequelae development. For example, age and gender influences ARF susceptibility, with young children and females more likely to develop RHD from ARF than males<sup>318,319</sup>. However, main ARF/RHD risk factors relate to differential expression of host genes involved in cytotoxicity, chemotaxis, apoptosis and immune mechanisms<sup>310,320,321</sup>. Specifically, innate immune polymorphisms include mutations in mannan binding lectin genes<sup>4,322–324</sup>, the TLR-2 gene<sup>325</sup>, the IL-10 gene promoter region<sup>326,327</sup>, and IL-6 genetic regions<sup>312,328</sup>. Adaptive immune polymorphisms are also responsible, such as variations in the DR or DQ regions of class II human leukocyte antigens (HLA), particularly HLA-DR7 and HLA-DQA1-DQB1<sup>317,329–331</sup>. Such mutations reduce effective GAS clearance<sup>332</sup>, as well as cytokine production levels<sup>315,326,333</sup>.

### 1.7.2. *Host Response to Natural Infection*

Distinct from counterproductive autoimmune responses, protective responses after natural infection can take up to 20 years to fully develop<sup>334</sup>. A single GAS infection does not lead to general streptococcal immunity and people generate different anti-GAS antibody levels throughout their lives<sup>335,336</sup>. For example, children more commonly have non-invasive superficial infections compared to adults, and iGAS disease is most common in the elderly due to immunosenescence and complicating co-morbidities<sup>37,337,338</sup>. Adults have the highest anti-GAS antibody levels of all age groups<sup>334,339–341</sup>, suggesting that natural exposure through repeated infections may generate some level of immunity over time<sup>334</sup>.

Slow immunity acquisition may be related to GAS antigenic diversity, particularly the immunodominant M protein<sup>342,343</sup>. Acquired immunity is achieved through type specific protection against homologous strains<sup>335</sup> specifically generating opsonic antibodies against the infecting serotype<sup>344,345</sup>. This has been observed in antibodies isolated from people who have recovered from GAS upper respiratory infections only binding the infecting *emm* type heat-killed strain<sup>346</sup>. Therefore, unless reinfected with the same GAS strain within a short period of time, immunity is short lived, and acquisition of long-lasting type-specific immunity is a slow process.

Humoral responses directed against a range of cell wall associated and secreted antigens are pivotal in GAS neutralisation and clearance. Mucosal antibodies such as IgA and IgM prevent GAS colonisation and adherence in murine models<sup>347</sup>, and type specific IgG antibodies promote GAS clearance by complement mediated phagocytosis<sup>4</sup>. Some of the dominant antigens recognised by these antibodies include the M protein, C5a peptidase, LTA, and fibronectin binding proteins<sup>12,348</sup>.

## 1.8. Control, Treatment and Prevention of GAS Infections

### 1.8.1. *Antibiotics Targeting GAS*

GAS retains susceptibility to  $\beta$ -lactam antibiotics<sup>349</sup>, used as the main method to prevent infection and associated complications<sup>350</sup>. Widespread use of benzathine penicillin G (BPG) has been pivotal in successful treatment of GAS pharyngitis, and has played an important role in ARF reduction observed in HICs<sup>351</sup>. GAS has generally remained susceptible to other  $\beta$ -lactam antibiotics such as cephalosporin and amoxicillin<sup>352</sup>, with macrolides and clindamycin used clinically to treat infections in penicillin allergy cases<sup>353–355</sup>.

Recently, however, some isolates have gained resistance to macrolides, clindamycin, and lincosamide antibiotics<sup>33,34,354,356</sup>. GAS strains have also displayed rare mutations in the penicillin-binding protein 2B (PBP2) reducing  $\beta$ -lactam susceptibility, including amoxicillin<sup>357</sup>, confirmed in a recent study observing *emm43* and *emm4* iGAS associated strains in the USA<sup>358</sup>. Such mutations in PBP2 have led to penicillin resistance in *S. pneumoniae*<sup>359</sup>. In addition to antimicrobial resistance risk, penicillin alone has limited efficacy in GAS carriage eradication<sup>360</sup>, with a combination of penicillin with other antibiotics such as rifampicin and azithromycin necessary to obtain carriage clearance<sup>361,362</sup>. Such treatment routes are associated with high costs and are often not implemented or available in LMICs where pharyngitis and carriage rates are highest. Additionally, there can be difficulty distinguishing strep throat from viral pharyngitis<sup>363</sup>, meaning infections are important drivers of broad spectrum antibiotic usage, which in some instances may be unnecessary<sup>364</sup>.

### 1.8.2. *Intravenous Immunoglobulin G (IVIG) Treatment*

In iGAS cases with serious life-threatening sepsis, necrotising fasciitis or STSS, antibiotics may not be completely effective and intravenous immunoglobulin G (IVIG) treatment is required. This entails injection with pooled plasma donor neutralising IgG antibodies to dampen proinflammatory responses and reduce cytotoxic T cell proliferation. Though effective, IVIG is expensive and without an

exhaustive source, used only in exceptional circumstances. Nevertheless, IVIG has been used to identify potential vaccine antigenic targets with an aim to skew responses towards neutralising and opsonic antibody protection<sup>365</sup>.

### *1.8.3. Vaccines Targeting GAS*

Vaccination is considered the most successful health intervention, reducing infectious disease incidence<sup>366</sup>. Antibiotics and IVIG treatment are not suitable to control GAS infections at the population level or stop transmission within communities<sup>56</sup>. Safe and efficacious vaccines are therefore required to control GAS morbidity and mortality, as well as potentially transmission towards pathogen eradication, possible with humans as the only natural host.

Vaccine development is a complex and lengthy process, however, GAS vaccine development has had arguably a more complicated history compared to most with an official impeded status recognised by the WHO, loosely related to bacterial, host and scientific effects. First and foremost, GAS is a complex pathogen with high genomic heterogeneity, different virulence factor expression profiles between strains, and complicated diverse global epidemiology. A genomic study showed that no single protein investigated during vaccine development has been 100% conserved between all the analysed GAS isolates<sup>150</sup>. Protein sequence variation within vaccine candidates, such as within the most progress NTD M protein-based vaccines, leads to the requirement of multicomponent inclusion to obtain acceptable cross protection levels. Additionally, the broad disease spectrum makes finding an effective long-term vaccine strategy challenging, as does disproportionate GAS disease burdens and diversity within LMICs.

Host specific effects have also impeded vaccine development through autoimmunity safety concerns as discussed previously. Early crude GAS vaccines reporting apparent autoimmune sequelae, led to uncertain and insufficient GAS vaccine market incentives. However, recent concerted efforts globally have improved awareness, aiming to support commercialisation and eventual product licensure. Key players include the Coalition to Accelerate New Vaccines for Group A Streptococcus (CANVAS) aiming to combat ARF/RHD and iGAS infections<sup>367</sup>, as well as the Wellcome Trust supported SAVAC – Strep A Vaccine Consortium, which is developing economic and business cases to assess safety, correlates of protection and vaccine targets. Such collaborations and investments have undoubtedly led to improved preclinical antigen discovery, vaccine formulation, and use of vaccine efficacy infection models. A vaccine summary is stated in Table 1.1.

## 1.9. GAS Vaccine History

GAS vaccine research started in the early 20<sup>th</sup> century aiming to prevent scarlet fever. Interest quickly waned as outbreaks reduced and ARF incidence declined in the USA and Europe due to penicillin discovery and implementation<sup>368</sup>. However, disease burden persisted globally particularly in LMICs with poor antibiotic access and higher morbidity and mortality associated with autoimmune complications. Recently the WHO called for better control and prevention of acute and chronic infections, as well as RHD autoimmune complication<sup>56</sup> through accelerated vaccine development and regulatory and policy implementation. Towards this aim, the WHO sponsored the publication of two documents, a framework towards Preferred Product Characteristics, and Vaccine Research and Development Roadmap<sup>369</sup>. The roadmap highlights vaccine efficacy against pharyngitis and skin infections as a near-term strategic goal with a hope to prevent more severe infections<sup>56</sup>.

### 1.9.1. *Whole Cell and Crude M protein Vaccines*

Early vaccines consisted of whole heated killed GAS developing into crude M protein preparations. The highly reactogenic vaccine was used to immunise 4,000 young adults but failed to show successful disease prevention<sup>370</sup>. This was succeeded by GAS cell wall fractions and partially purified M proteins as antigens in the 1960s which were shown to be reactogenic but also produced inconsistent immune responses<sup>371,372</sup>.

This was followed by a study which repeatedly immunised children with partially purified M proteins<sup>373</sup>, which lead to concerns that GAS vaccines may induce autoimmune symptoms. Specifically, the group published controversial findings linking two partially purified M protein candidates to vaccine induced autoimmune sequelae, reporting an apparent ARF increase in 3 out of 21 vaccinated volunteers<sup>374</sup>. Whether the apparent ARF increase was caused by the vaccine is still debated, however, such findings resulted in a US Food and Drug Administration (FDA) ban on the use of GAS organisms and derivatives in bacterial vaccines in 1978<sup>375</sup>. The ban was lifted in 2006 following evidence of non-vaccine confounding factors influencing ARF susceptibility in the vaccinated subjects<sup>376</sup>. Since resuming vaccine research, no similar adverse safety signals have been observed, however, the ban significantly slowed vaccine development, with no reported trials documented during a 25-year period<sup>56</sup>. To date, pre-clinically 28 vaccine candidates have been shown to be protective against GAS infections using a number of animal models<sup>377,378</sup>. However, currently there are no licenced vaccines, and only a small number of candidates have been tested in human clinical trials<sup>377,378</sup>.

**Table 1.1: Summary of GAS vaccine candidates currently in development.**

NTD = amine (N) terminal domain, CTD = carboxylic terminal domain, HVR = hyper variable region, GMMA = generalised modules for membrane antigen.

Type	Name / Identity	Antigen Details	Stage of development	References
Purified Whole M protein		Purified M protein	Preclinical Proof of concept in human challenge Phase I	379-381
Multivalent NTD M protein	StreptAnova	HVR 4, 6 and 8 valent	Phase I	382-384
		HVR 26 valent	Phase I and Phase II	385,386
		HVR 30 valent	Phase I	377,387
		M peptides displayed on surface of commensal <i>L. lactis</i>	Pre-clinical	388,389
CTD M protein	StreptInCor	C-repeat epitope	Pre-clinical	390,391
	J8	Minimal Epitope	Pre-clinical Phase I	378,392
	J14/p145	Minimal Epitope	Pre-clinical	393
	J14 Combo	10 valent HVR	Pre-clinical	394
Purified protein antigens (Multicomponent)	Combo3	SpyCEP, SpyAD and SLO	Pre-clinical	395
	Combo5	Arginine Deiminase (ADI), Trigger Factor (TF), C5a peptidase, SpyCEP, SLO	Pre-clinical Proof of concept in non-human primate model	396,397
	Spy7	SpyAD, C5a peptidase, Spy0762, Spy0651, oligopeptide binding, nucleoside-binding protein, pullulanase, nucleoside-binding protein	Pre-clinical	398,399
	5CP	SpyCEP, SLO, SpyAD, C5a peptidase, sortase A	Pre-clinical	394
Purified protein antigens (Single Component)		C5a peptidase	Pre-clinical	400
		Fibronectin binding protein	Pre-clinical	401
		Streptococcal protective antigen	Pre-clinical	385
		Serum opacity factor	Pre-clinical	402
		Streptococcal pyrogenic exotoxin A/B/C	Pre-clinical	403-405
		Streptococcal pili T antigen	Pre-clinical	159,406
		SpyCEP	Pre-clinical	208
		Other common antigens	Pre-clinical	407
		Identified but untested antigens; G-related alpha2-macroglobulin binding (GRAB), Protein, metal transporter of streptococcus (MtsA), superoxidase dismutase, lipoproteins	Pre-clinical	408156
	Wildtype GAC		Native GAC conjugated to Tetanus toxoid (TT)	Pre-clinical
		Native GAC with CRM <sub>197</sub>	Pre-clinical	410
		Native GAC with GAS antigens SpyCEP, SpyAD and SLO	Pre-clinical	411
Mutant GAC (Rhamnose Polymer)		Rhamnose defective of GlcNAc side chains with <i>S. pneumoniae</i> carrier protein	Pre-clinical	141
		Rhamnose defective of GlcNAc side chains with GAS Antigen SpyAD	Pre-clinical	412
Synthetic GAC Oligosaccharides		Synthetic carbohydrate with C5a peptidase	Pre-clinical	413
		Synthetic carbohydrate with CRM <sub>197</sub>	Pre-clinical	139
Novel Antigen Display		GAC displayed in gold nanoparticles	Pre-clinical	414
		GAC displayed in GMMA	Pre-clinical	415
		Synthetic GAC repeating unit presented in lipid-core complex	Pre-clinical	416

### 1.9.2. Purified M Protein Vaccines

The most progressed GAS vaccines are multivalent subunit vaccines in phase I and II clinical trials containing highly purified recombinant M proteins<sup>377,385,387</sup>. Such vaccines are built on immunisation data dating back to the 1970s, demonstrating protection from infection with homologous M types<sup>379,380</sup>, and presence of protective serotype specific bactericidal antibodies inducing phagocytic GAS killing<sup>381</sup>.

#### 1.9.2.1. Multivalent Amino (N) Terminal Domain (NTD) M Protein Vaccines

From the M protein structure (Figure 1.2b), the hypervariable N terminal domain (NTD) contains epitopes which evoke the most bactericidal antibodies.

Recombinant multivalent NTD M peptide vaccines in pre-clinical and clinical trials show good safety and immunogenicity profiles<sup>384–386</sup>. Over the years, M protein-based vaccines have become more complex in their formulation and can be divided into 6-valent, 26-valent, and 30-valent vaccines, from most basic to most progressed. Initially the leading multivalent vaccine consisted of NTD portions of 6 strains, predominantly found in the US<sup>384</sup>. However, due to regional variation between strains, the vaccine failed to show cross strain protection<sup>384,417</sup>. This vaccine was subsequently redesigned with increased valency, containing 26 M protein NTD peptides (StreptAvax), commonly found in Europe and North America<sup>385</sup>. The vaccine was well tolerated, generating elevated protein specific opsonic antibody responses in humans<sup>385</sup>, without autoreactive antibody generation<sup>385</sup>.

Nevertheless the 26-valent StreptAvax vaccine still failed to protect against a broad coverage of GAS strains, particularly against isolates found in Asia and the Pacific continents<sup>108</sup>. It was therefore further developed to contain 30 M protein NTD peptides (StreptAnova™)<sup>387</sup>. The increased valency demonstrated greater protection and efficacy, specifically through bactericidal opsonic antibodies directed against all included vaccine serotypes, as well as 24 additional non-vaccine serotypes<sup>387</sup>. This indicated broad coverage and neutralising ability beyond vaccine serotypes due to similarities between M proteins in the same *emm* clusters<sup>99</sup>. Recently in 2020, StreptAnova™ phase I trials began and have demonstrated that the vaccine is safe, immunogenic, and well tolerated without inducing any autoimmune antibodies<sup>377</sup>.

However, a concern with widespread multicomponent NTD M protein vaccine implementation is the emergence and dominance of non-vaccine serotypes<sup>384</sup>. However, given that the StreptAnova™ vaccine was able to provide cross protection

to M proteins not included in the formulation, these concerns may be negated. Nevertheless, vaccine-induced escape serotypes have been observed for other bacterial vaccines<sup>418,419</sup>, and can be theoretically applied to GAS vaccines, with one study reporting only 685 out of 2,083 tested genomes contain a vaccine NTD *emm* subtype<sup>150</sup>. This equates to 33% antigenic coverage within the 30-Valent StreptAnova™ vaccine formulation<sup>150</sup>, threatening to reduce protection in areas where people are most at risk of serious GAS complications<sup>108,420</sup>. For this reason, the conserved CTD and middle region of M proteins have been investigated to circumvent sequence variation<sup>378,421</sup>.

#### 1.9.2.2. Multivalent Carboxylic Terminal Domain (CTD) M Protein Vaccines

A group of alternative M protein subunit vaccines consist of minimal B cell epitopes from the CTD region, namely J8, J14 and J145. These vaccines have generated antibody responses protecting mice from mucosal, soft tissue and systemic infection, independent of the challenging *emm* type<sup>392,422</sup>. The most progressed CTD M protein-based vaccine is StreptInCor currently undergoing phase I/IIa clinical trials. This vaccine consists of a conserved 55 amino acid polypeptide, containing linked B and T cell epitopes similar to the CTD of M5 protein sequences. The vaccine demonstrated protection in mice<sup>390</sup>, and good preclinical toxicity and tolerability<sup>391</sup>.

However, CTD M protein vaccines may not protect against systemic deep tissue infections, and are predicted to present less opsonic and bactericidal epitopes than NTD vaccines<sup>396,423,424</sup>. This concern was addressed through the development of MJ8VAX (J8-DT) vaccine which consists of a 29 amino acid peptide conjugated to Diphtheria Toxoid (DT)<sup>392,393</sup>. In the tested animal model, MJ8VAX stimulated opsonic antibody production, inducing protection against intranasal and intraperitoneal infection<sup>425,426</sup>. Following promising safety and immunogenicity studies<sup>392,426–428</sup> phase I clinical trials showed that MJ8VAX was immunogenic and well tolerated<sup>378</sup>. However, over time antibody levels decreased and the overall vaccine efficacy was disputed<sup>378</sup>.

Since J8 there has been alternative variations trialled, for example the conjugation of the CTD peptide to the universal T cell epitope, T-helper Pan HLA-DR-binding epitope peptide (PADRE), as well as J14 and J145 vaccines, engineered to have amino acid replacements to generate more robust and potent immune responses compared to J8<sup>429</sup>. A number of J14 and J145 formulations and routes of administration have also been tested including sophisticated adjuvants c-di-AMP



and BPPCysMPEG showing superior immunogenicity when delivered via the mucosal route within lipopeptide nanocarriers<sup>430</sup>. Another novel approach to improve immunogenicity was to fuse three J14 repeats to different NTD M peptides containing a C terminal Gram positive sortase motif. The aim was to encourage conjugation of TLR-2 agnostic FSL-1, and was shown to induce high antibody titres, protecting mice from lethal GAS challenge<sup>394</sup>.

### 1.9.3. *Non-M Protein Vaccines*

Alternative 'next generation' GAS antigens are also under investigation, aiming to find abundant and cross-protective antigenic epitopes effective at different infection stages<sup>428</sup>. Ideal candidates would be suitable for use globally through broad strain coverage but many non-M protein vaccines remain to be tested in human clinical trials<sup>431,432</sup>. Alternative antigens explored include pili<sup>159,406</sup>, ECM and fibronectin binding proteins<sup>401</sup>, C5a peptidase<sup>400</sup>, streptococcal pyrogenic exotoxins (SpeA/B/C)<sup>403–405</sup>, among others<sup>385,407,408</sup>. Important in this work, *S. pyogenes* cell envelope protease – SpyCEP (Spy0416), and *S. pyogenes* Adhesion and Division protein – SpyAD (Spy0269) have also been investigated as antigens both alone<sup>208</sup>, in combination<sup>395,433,434</sup> as well as recently as protein carriers in glycoconjugate vaccines<sup>411,412</sup>.

### 1.9.4. *Multicomponent Vaccines*

Due to the observed variation in circulating GAS serotypes, vaccines containing multiple antigens are often more desirable by offering better cross-protection. Multicomponent vaccines have the potential to overcome antigenic variation and provide a multi-hit approach to immunogenicity. Many multicomponent vaccines have focussed on combined individual candidates in a single formulation<sup>396,433</sup>. These have often been identified by reverse vaccinology methods, combining antigen localisation, sequence conservation, and immunogenicity<sup>395</sup>, or isolated from pooled human IgG IVIG<sup>399</sup>.

#### 1.9.4.1. *Protein Subunit GSK Combo3 Vaccine (SpyCEP, SLO, SpyAD,)*

GlaxoSmithKline (GSK) have developed a combination vaccine containing three conserved antigens, namely SpyCEP, SpyAD, and SLO (Streptolysin O - Spy0167)<sup>395</sup>. These prevalent and conserved surface exposed or secreted antigens were selected based on high-throughput proteomic approaches in combination with immune-array and flow cytometry techniques<sup>395</sup>. Preclinical investigation demonstrated consistent antigen specific bactericidal antibody production effective

in clearance of four M serotypes when challenged via intraperitoneal and intranasal routes within a murine model<sup>395</sup>.

#### 1.9.4.2. *Protein Subunit Combo5 Vaccine (SpyCEP, SLO ADI, TF, and C5a peptidase)*

Based on combo3 vaccine's encouraging immunogenicity<sup>395</sup>, SpyCEP and SLO were taken forward for inclusion in a five GAS protein antigen multicomponent vaccine. The additional antigens included C5a peptidase (ScpA)<sup>435</sup>, Trigger Factor (TF)<sup>436</sup> and an inactive version of Arginine Deiminase (ADI)<sup>437</sup>. TF and ADI enzymes are anchorless conserved surface proteins<sup>436</sup>, however there may be coverage concerns due to protein sequence variation and heterogeneity of SLO and C5a peptidase<sup>150</sup>. When combined and formulated with Aluminium Hydroxide (Alum) adjuvant, immunised mice displayed protection from cutaneous GAS challenge, but not from systemic infection<sup>396</sup>. Interestingly, however, when using the same invasive murine model, and formulating Combo5 with different experimental adjuvants, such as saponin QS21, increased protective efficacies were observed, through a more balanced Th1/Th2-type immune response important for iGAS protection<sup>397</sup>. Combo5 immune responses have also been characterised in non-human primate models showing a successful reduction in pharyngitis but not colonisation<sup>434</sup>.

#### 1.9.4.3. *Protein Subunit 5CP Vaccine*

A similar multicomponent vaccine containing five protein antigens known as 5CP has also been developed, consisting of C5a peptidase, SpyAD, SpyCEP, SLO in addition to sortase A, a membrane associated cysteine protease<sup>394</sup>. The vaccine, containing antigens from a number of pathogenic mechanisms such as adhesion, phagocytosis, cytolysis, and chemotaxis, was adjuvanted with CpG-oligodeoxynucleotides and immunised via the intranasal route. The vaccine showed protection against intranasal, skin and systemic challenge, stimulating a controlled Th17 responses via local and systemic antibodies for at least 6 months following the last immunisation<sup>394</sup>.

#### 1.9.4.4. *Protein Subunit Spy7 Vaccine*

Further increased antigen inclusion has also been investigated to broaden coverage. Spy7 vaccine contains several surface antigens recognised by pooled IVIG, specifically C5a peptidase, oligopeptide-binding protein, pullulanase, nucleoside-binding protein, hypothetical membrane associated protein Spy0762, cell surface protein Spy0651, and SpyAD<sup>399</sup>. The seven antigens were recombinantly expressed in *E. coli* and formulated with Freund's adjuvant before

immunisation, demonstrating reduced GAS M1 and M3 dissemination in a murine model<sup>365</sup>. However, Spy7 immunogenicity when formulated with an human approved adjuvant such as Alum remains to be characterised<sup>438</sup>.

#### 1.9.5. GAS Polysaccharide Vaccines

Carbohydrate capsules incorporated into vaccine preparations have protected against a number of streptococcal species, most notably *S.pneumoniae*, however, the GAS HA capsule is an unsuitable vaccine target<sup>135,439</sup>. For this reason, the Group A Carbohydrate (GAC), has been suggested as an alternative polysaccharide vaccine candidate. GAC branched trisaccharide repeating units have been shown to be immunogenic<sup>440,441</sup>, through studies using rabbit and human antiserum<sup>442,443</sup>, computer simulations<sup>144,146,444</sup> and Nucleic Magnetic Resonance (NMR) structural techniques<sup>441</sup>.

Anti-GAC antibody titres are thought to correlate with reduced GAS infection incidence in adults, suggesting that although weakly immunogenic compared to other GAS antigens, anti-GAC antibodies are important in long lasting immunity<sup>409</sup>. Affinity purified anti-GAC antibodies have also been shown to have opsonic properties specifically against M3, M6, M14 and M28 serotypes<sup>445</sup>. Sabharwal *et al*<sup>409</sup> also demonstrated an inverse relationship between anti-GAC antibody levels and throat colonisation, as well as protection against systemic and intranasal GAS challenge in mice vaccinated with GAC conjugated to a protein carrier.

However, as discussed GlcNAc side chains are believed to be responsible for post infection autoimmune sequelae<sup>410</sup>. To address potential vaccine safety concerns, Van Sorge *et al*<sup>441</sup> used modified GAC devoid of GlcNAc and conjugated it to *S. pneumoniae* protein SP0435. They showed that antibodies raised against rhamnose polymers only promoted phagocytic killing of multiple GAS serotypes and could protect mice systemically following passive immunisation with rabbit anti-GAC antisera<sup>141</sup>. Both studies, along with findings by Kabanova *et al*<sup>139</sup> justifies further investigation into both wildtype and modified GAC as protective conserved vaccine antigens for inclusion in glycoconjugate vaccine design (Table 1.2).

**Table 1.2: GAS glycoconjugate immunisation strategies using GAC as a vaccine component.**

- = information not available.

Group	Protein	Saccharide	Conjugation Chemistry	Clean up	Physicochemical Analysis	Immunological Analysis
Sabharwal <i>et al.</i> , 2006 <sup>409</sup>	TT	Wildtype GAC extracted from GAS cells	Reductive amination with sodium cyanoborohydride	TFF MWCO	-	IgG titre (ELISA) Passive murine protection model Murine challenge model – subcutaneous and intra nasal against colonisation and survival Autoimmunity cross reactivity
Kabanova <i>et al.</i> , 2010 <sup>139</sup>	CRM <sub>197</sub> CRM <sub>197</sub>	Synthetic wildtype GAC fragments (6mer and 12mer) Wildtype GAC extracted from GAS cells	Activation via disuccinimidyl adipate and reductive amination with sodium cyanoborohydride	Ultrafiltration	<sup>1</sup> H NMR spectrum of GAC Mass Spec	IgG titre (ELISA) Antibody function (opsonophagocytic killing) Antibody binding (flow cytometry - 2 GAS strains) Murine protection model
Auzanneau <i>et al.</i> , 2013 <sup>446</sup>	TT	Synthetic GAC fragments	Cysteamine GAC activation and diethyl squarate amine reaction	Ultrafiltration	Epitope mapping Saturation transfer difference NMR Colourmetric assays	IgG and IgM titre (ELISA) Passive murine protection model with rabbit antisera - systemic infection
Sorge <i>et al.</i> , 2014 <sup>141</sup>	Recombinant pneumococcal protein SP_0435	Mutant GAC (polyrhamnose) extracted from $\Delta gacI$ GAS cells	Affinity interactions - streptavidin / biotin	Size exclusion HPLC	-	Antibody binding (flow cytometry - 8 GAS strains) Antibody function (opsonophagocytic killing) Autoimmunity cross reactivity
Rivera-Hernandez <i>et al.</i> , 2016 <sup>396</sup>	Arginine deiminase protein	Native mutant extracted from GAS cells	Cyanylation using 1-cyano-4-dimethylaminopyridinium tetrafluoroborate	Size exclusion HPLC	-	IgG titre (ELISA) Antibody binding (flow cytometry - 2 GAS strains) Indirect bactericidal assay Skin and subcutaneous challenge model
Zhao <i>et al.</i> , 2019 <sup>447</sup> Wang <i>et al.</i> , 2020 <sup>413</sup>	C5a peptidase	Synthetic wildtype oligomers (3mer, 6mer, 9mer)	Acylation reactions	Size exclusion HPLC	MALDI-TOF	IgG titre (ELISA) Antibody function (opsonophagocytic killing) Antibody binding (flow cytometry - 1 GAS strain) Murine challenge model – subcutaneous CD4+ T cell activation and proliferation
Di Benedetto <i>et al.</i> , 2020 <sup>411</sup>	CRM <sub>197</sub> SpyCEP SpyAD SLO	Wildtype GAC extracted from GAS cells	Oxidation reduction amination (with and without oxidised GAC)	Ultrafiltration	Molecular mass (HPLC-SEC - refractive index and fluorescence detection)	IgG titre (ELISA) Antibody binding (flow cytometry – 1 GAS strain) Antibody function (functional cleavage assays)
Pitirollo <i>et al.</i> , 2020 <sup>414</sup>	Conjugation to gold nanoparticles	Synthetic polyrhamnose	Thiol-ending linker	Ultrafiltration	UV-Vis H-PAD Dynamic light scattering Electron Microscopy Nanoparticle tracking analysis	Competitive ELISA
Gao <i>et al.</i> , 2021 <sup>412</sup>	SpyAD, SLO and C5a peptidase	Mutant GAC (polyrhamnose) extracted from $\Delta gacI$ GAS cells	Site-direct CLICK chemistry	Dialysis	SEC-MALS	IgG titre (ELISA) Antibody binding (flow cytometry - 8 GAS strains) Antibody function (opsonophagocytic killing, functional blocking) Passive protection murine challenge model – systemic and skin Autoimmunity cross reactivity
Khatun <i>et al.</i> , 2021 <sup>416</sup>	Conjugation Ac-PADRE–lipid core	Synthetic tri-rhamnose	Alkyne–azide cycloaddition	Semipreparative RP-HPLC	HPLC-SEC Dynamic light scattering	IgG titre (ELISA) Antibody function (opsonophagocytic killing)

## 1.10. Glycoconjugate Vaccines

Identification of protective antibodies induced by pneumococcal polysaccharide enabled the development of the first vaccine composed of purified polysaccharide<sup>448,449</sup>. During the 1970s and 1980s polysaccharide vaccines against meningococcus serogroup ACWY, *S. pneumoniae* and *Haemophilus influenzae type b* (Hib) were licensed. However, bacterial polysaccharides alone were found to be poorly immunogenic especially in children due to poor antibody boosting and limited isotype switching from IgM<sup>450</sup>.

### 1.10.1. *Progression Towards Glycoconjugate Vaccines*

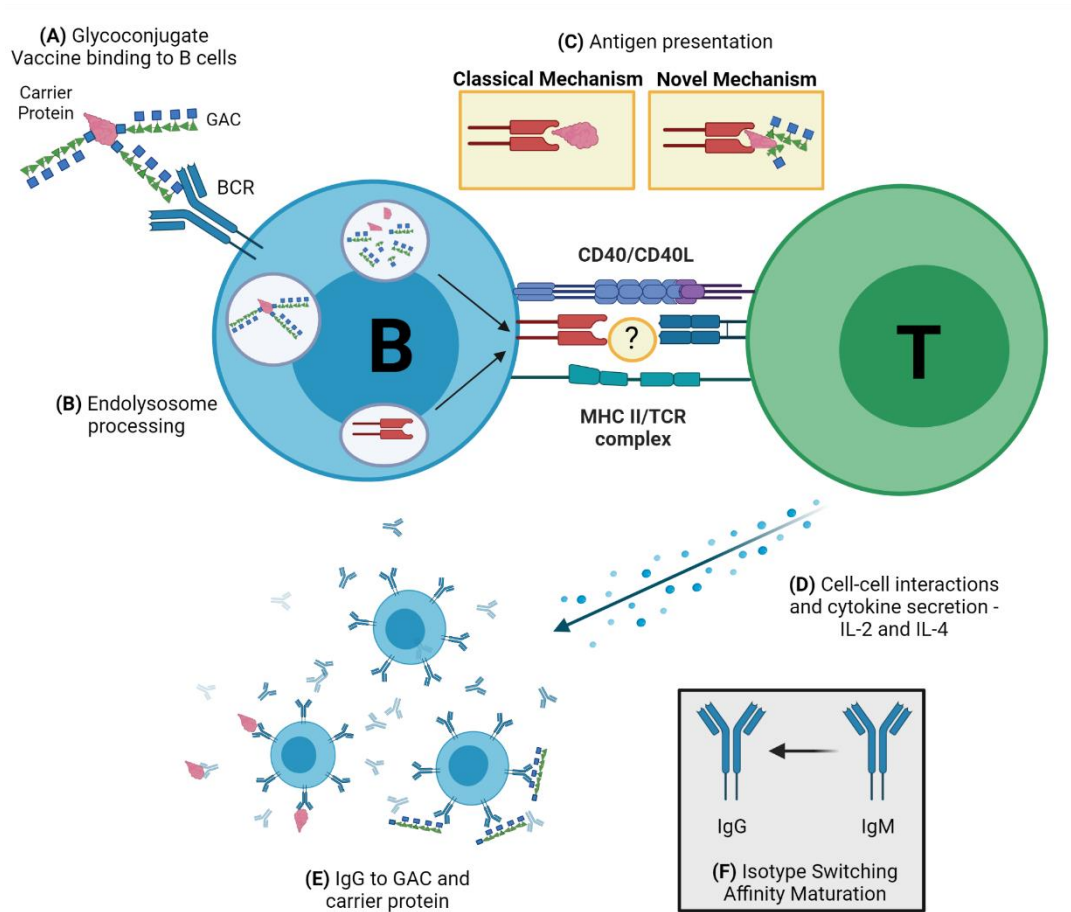
Generally the disease burden of encapsulated bacteria is highest during the first years of life, a complication for the low efficacy of polysaccharide vaccines at this age<sup>451–453</sup>. Polysaccharide antigens alone are T cell independent antigens, containing repetitive structures which although recognised and cross-linked by B cell receptors (BCR) on B cells, cannot provide T cell epitopes. This leads to B cell differentiation into plasma cells which secrete antibodies directed against polysaccharide epitopes. However, there is no T cell helper function as the mechanism of B to T cell cooperation is immature in infants<sup>454,455</sup>, reducing immunological memory<sup>456,457</sup>. Conjugation of polysaccharide to a carrier protein enables the protein moiety to provide such T cell epitopes strengthening and improving immune response longevity<sup>458</sup>.

This was demonstrated in the early 20<sup>th</sup> century through the discovery of altered immune responses to protein carriers when small molecules, specifically haptens, were chemically attached to them. Antibodies raised against the conjugate including the attached small molecule, recognised the hapten in both its conjugated and native form<sup>459</sup>. Following this, Avery and Goebel showed conjugation of small saccharides to proteins resulted in enhanced immunogenicity, increasing polysaccharide specific antibody titres<sup>460</sup>. In addition, they were able to show protection against pneumococcal type III strain in rabbits following immunisation with capsular polysaccharide conjugates<sup>461</sup>, triggering glycoconjugate vaccine research.

The first glycoconjugate developed for human use was against Hib licensed between 1987 and 1990<sup>462–464</sup>. Since conception, glycoconjugate vaccines have been deemed the safest and most efficacious vaccines currently used in immunisation schedules across the globe, significantly decreasing globally meningitis and pneumonia disease in humans<sup>465–467</sup>.

### 1.10.2. *Mechanism of Action of Glycoconjugates*

Glycoconjugate vaccines composed of covalently attached protein and polysaccharide components provide the immune system with multiple triggers, increasing a vaccine's protective lifetime through B and T cells memory responses<sup>468</sup>. Avci *et al*<sup>469</sup> proposed a mechanism whereby conjugated polysaccharide is detected by innate immune cells and binds to BCRs on B cell surfaces. This triggers glycoconjugate internalisation and degradation into smaller fragments<sup>470</sup>, presented on MHC class II complexes interacting with  $\alpha\beta$  TCRs on T cell surfaces<sup>469</sup>. This may be peptide fragment presentation through a classical mechanism, and / or glycopeptide presentation through a novel mechanism leading to the generation of CD4<sup>+</sup> T cell help through cell-cell interactions and cytokine secretion. This in turn increases polysaccharide specific antibody levels, affinity maturation and proliferation of polysaccharide-specific B cells from memory pools, resulting in IgM to IgG isotype switching leading to higher antibody avidity<sup>469</sup>. Such B cells present in a memory pool mean that upon subsequent encounter of the specific antigen these cells can rapidly proliferate and differentiate into plasma cells secreting high anti-polysaccharide antibody titres with high antibody avidity<sup>457</sup> (Figure 1.3).



**Figure 1.3: Schematic representation of immunological processing and presentation of glycoconjugate vaccines by B cells to initiate T cell help effector functions.**

Glycoconjugate polysaccharide antigens bind to B cell receptors (BCR) of B cells **(A)**.

They are internalised into endolysosomes and processed by acid hydrolysis **(B)**.

Fragments of either polysaccharide conjugated to protein (glycopeptides – novel mechanism) or peptide fragment (classical mechanism) are presented on major histocompatibility complex (MHC) class II molecules on B cell surfaces **(C)**.

T cell receptors (TCR) on T cells recognise either the glycopeptide or peptide fragment and provide help to B cells by cell-cell interactions and cytokine secretion **(D)**.

B cells undergo rapid proliferation and trigger random mutation in Ig genes altering BCR and secreted antibody specificity **(E)**.




There is also isotype switching from IgM to IgG antibodies against polysaccharide and protein carrier components **(F)**, and induction of memory B cell pools.

Original figure created with BioRender.com. Figure adapted from Figure 6 Avci *et al.*, 2011<sup>467</sup>.

### 1.10.3. Glycoconjugate Vaccines Related to this Work

It is of great interest and importance to investigate vaccine approaches to control GAS disease. Glycoconjugate vaccines are a promising approach and will be applied to selected vaccine antigens summarised Figure 1.4 for this work.

Compared to bacterial polysaccharides, there has been less attention on protein carrier properties in glycoconjugate vaccines<sup>471</sup>. Recently, however, there has been a drive for new carriers to be investigated<sup>472</sup>, as data suggests that repeated immunisation with the same classical carrier for different glycoconjugate vaccines in some cases dampens immunological potency and efficacy<sup>473–476</sup>. This thesis therefore focuses on whether ‘double hit’ glycoconjugates, containing GAS protein antigens, can be produced and invoke an immune response. The potential benefit from additional protective antibodies directed against the protein component has currently not been fully exploited.

	Polysaccharide		MaIE	Proteins SpyCEP	SpyAD
	GAC Wildtype	GAS_Rha			
<b>Bacteriology</b> 	30 - 50 % cell mass Conserved Encoded by <i>gac</i> operon	Mutant from serial animal passage / genetic knockout Cells viable	Membrane exposed Conserved Binds maltodextrin	Cell wall anchored / Secreted Conserved Cleave IL-8 cytokine	Cell wall anchored Conserved Cell division and ECM binding
<b>Pathology</b> 	Essential for host survival High antibody titres = protection against strep throat	Reduce virulence in rabbit animal model Increased killing by human whole blood and neutrophils	Upregulated for survival in saliva Colonisation	Facilitate bacteremia and sepsis Immune evasion	Adhesion to epithelial and pharyngeal cells Colonisation
<b>Immunobiology</b> 	Anti GAC antibodies from natural infection and conjugates GlcNAc = immunodominant epitope	Polyrhamnose recognised by Anti GAS sera Hypothetical reduction in autoimmunogenic antibodies	Immunogenic - elevated in murine sera Antibodies reduce colonisation in animal challenge	Immunogenic - neutralising antibodies Protection from lethal challenge in animal model	Immunogenic Protection from lethal challenge in animal model

**Figure 1.4: Summary of glycoconjugate vaccine antigen selection.**

Advantageous properties include favourable immunogenicity and relevant pathogenic roles for GAS survival and infection of human hosts. Bacteriological, pathological and immunobiological roles of selected vaccine antigens used in glycoconjugate vaccine production are stated.

Original figure created with BioRender.com.



## 1.11. Research Aims, Objectives, and Hypotheses

The overall aim of my PhD project is to investigate whether novel glycoconjugate vaccines against GAS can be developed using both chemical and biological conjugation methodologies. In particular this work concerns two conjugation methods, traditional chemical EDC conjugation, and biological *N*-linked conjugation. The resulting glycoconjugates obtained from these methods were characterised in terms of their physicochemical and immunogenicity properties.

### *1.11.1. Chapter 3: Chemical Conjugation Objectives*

This experimental Chapter describes the approach to obtaining purified polysaccharide and GAS specific protein antigens for use in investigation of chemical conjugation reactions. The suitability and optimisation of the selected chemical reaction is also described using EDC Carbodiimide zero length linker.

### *1.11.2. Chapter 4: Biological Conjugation Objectives*

This experimental Chapter outlines the preparation of the *E. coli* produced recombinant polysaccharide components and GAS specific protein antigens for inclusion in biological conjugation reactions. It also depicts the two experimental methods for investigation of successful biological conjugation using PglB oligosaccharyltransferase enzyme.

### *1.11.3. Chapter 5: Immunogenicity Characterisation Objectives*

This experimental Chapter concerns glycoconjugate vaccine immunogenicity. Concerning the carrier protein, GAS proteins were investigated and compared to the common carrier, TT for chemical conjugates and NanA for biological conjugates. The role of the polysaccharide moiety has also been investigated through either inclusion of polysaccharide derived from GAS cells for chemical conjugation, or alternatively *E. coli* produced recombinant polysaccharide for biological glycoconjugate vaccine inclusion.

### *1.11.4. Hypotheses*

- Biological conjugation is a viable alternative to classic chemical conjugation.
- GAS carrier proteins will provide a 'double hit' approach to glycoconjugate vaccines. GAS carrier proteins can effectively produce antibodies directed against the polysaccharide moiety, in addition to protein specific antibodies.

## CHAPTER 2

### Materials and Methods

Generic materials and methods stated in this Chapter were used throughout this thesis. Experimental information relevant to addressing specific research questions are placed in the relevant thesis sections; Chapter 3, Chapter 4, and Chapter 5.

#### 2.1. Materials

##### 2.1.1. *Bacterial Strains*

*Streptococcus pyogenes* (GAS) NCTC-8198 strain is an M1 serotype from Public Health England (PHE) culture collection, isolated from an individual with scarlet fever, and will be used for this work. *Escherichia coli* (*E. coli*) strains used in this work are presented in relevant Chapters for either chemical (Chapter 3 method section 3.3.1), or biological conjugation (Chapter 4 method section 4.3.1). All strains were stored long term in 20% glycerol (Sigma-Aldrich™) at -80 °C.

##### 2.1.2. *Antibodies and Lectins*

Primary and secondary antibodies as well as lectins used throughout all Chapters are stated in Table 2.1. A short description states the purpose for use, usual working dilution and supplier.

**Table 2.1: List of antibodies and lectins used in this study.**

Name	Dilution	Primary antibodies	
		Description / Use	Supplier
Mouse anti-His	1:10,000	Detection of His tagged recombinant GAS antigens, NanA, and AcrA proteins (biological conjugation)	AbCam
Mouse anti-6x-His Tag (HIS.H8)	1:10,000		Thermo Fisher Scientific
Mouse 6x-His Tag (AD1.1.10) HRP	1:5,000 - 10,000	Detection of His tagged recombinant GAS antigens (chemical conjugation) conjugated to Horse Radish Peroxidase (HRP)	Thermo Fisher Scientific
Rabbit anti-His	1:5,000	Detection of His tagged recombinant GAS antigens, NanA, and AcrA proteins (biological conjugation)	AbCam
Goat anti-GAS pAb	1:1,000 - 5,000	Detection of Group A Carbohydrate from GAS cells (chemical conjugation) and recombinant rhamnose variants (biological conjugation)	Thermo Fisher Scientific
Rabbit anti-GAS pAb	1:5,000		AbCam
Guinea Pig 10/132 TT antitoxin	1:200 - 800	Detection of Tetanus Toxoid (chemical conjugation)	NIBSC
Horse 60/013 TT antitoxin	1:200	Detection of Tetanus Toxoid (chemical conjugation)	NIBSC

<b>Secondary antibodies</b>				
Name	Dilution	Description / Use		Supplier
Goat anti-mouse IgG IRDye® conjugated 660RD	1:10,000	Development of primary antibodies raised in Goat species	Fluorescent detection in red channel	LI-COR®
Rabbit anti-goat IgG IRDye® conjugated 680RD	1:10,000	Development of primary antibodies raised in Rabbit species	Fluorescent detection in red channel	LI-COR®
Donkey anti-Guinea Pig IgG IRDye® conjugated 680RD	1:10,000	Development of primary antibodies raised in Guinea Pig species	Fluorescent detection in red channel	LI-COR®
Goat anti-Rabbit IgG IRDye® conjugated 800CW	1:10,000	Development of primary antibodies raised in Rabbit species	Fluorescent detection in green channel	LI-COR®
Rabbit anti-Goat IgG IRDye® conjugated 800CW	1:10,000	Development of primary antibodies raised in Goat species	Fluorescent detection in green channel	LI-COR®
Rabbit anti-Goat IgG HRP	1:10,000	Development of primary antibodies raised in Goat species	HRP conjugated detection by Pierce ECL	Thermo Fisher Scientific
Streptavidin HRP	1:5,000	Biotin detection	HRP conjugated detection by Pierce ECL	BD Pharmingen™
IRDye® 680RD Streptavidin	1:10,000	Biotin detection	Fluorescent detection in red channel	LI-COR®
Rabbit anti-human IgG-HRP	1 : 20,000	Detection of human IgG in IVIG	HRP conjugated detection by TMB	Sigma-Aldrich™
Rabbit anti-Mouse IgG-HRP	1 : 20,000	Detection of mouse IgG	HRP conjugated detection by TMB	Sigma-Aldrich™
Goat anti-Guinea Pig	1 : 200	Detection of Guinea Pig IgG	HRP conjugated detection by TMB	Sigma-Aldrich™
<b>Lectins</b>				
Name	Dilution	Description / Use		Supplier
Biotin-conjugated sWGA	1 : 2,500	GlcNAc detection for GAC extracted from WT GAS and $\Delta$ <i>gacI</i> GAS cells (chemical conjugation)		Vector Laboratories
Biotin-conjugated SBA	1 : 2,500	GalNAc and Gal detection for <i>C. jejuni</i> heptasaccharide (biological conjugation)		Vector Laboratories

## 2.2. DNA Manipulation

### 2.2.1. Cloning Materials

Commercial and cloned plasmid constructs (Table 2.2) and primers (Table 2.3) used in this work are presented for either chemical or biological conjugation purposes. Oligonucleotides were ordered from Integrated DNA Technologies (IDT), Thermo Fisher Scientific or TwistBioscience. Primer properties were checked in terms of melting temperature and secondary structure using the NEB Tm calculator (<https://tmcalsculator.neb.com>). All restriction enzymes were sourced from New England Biolabs.

**Table 2.2: Plasmids used in this study.**

Plasmid	Description	Source
pET28a	High copy number protein expression vector with kanamycin resistance cassette and T7 promoter	Commercial
pET28a_ <i>malE</i>	pET28a carrying full length MalE	This study
pET28a_ <i>spyCEP</i>	pET28a carrying SpyCEP fragment	This study
pET28a_ <i>spyAD</i>	pET28a carrying full length SpyAD	This study
pET28a_ <i>plyC</i>	pET28a carrying PlyCA and PlyCB monomer	This study
pUC19	High copy number plasmid with ampicillin resistance cassette and lac promoter	Commercial
pUC19_ <i>kanR</i>	pUC19 harbouring kanamycin resistance cassette	This study
pUC19_ <i>kanR_gacI</i>	pUC19 harbouring kanamycin resistance cassette and <i>gac</i> locus homology arms	This study
pEXT21	Low copy number protein expression vector with spectinomycin resistance cassette and tac promoter	477
pEXT21_ <i>pglB</i>	pEXT21 carrying full length PglB	478
pEXT21_ <i>malE</i>	pEXT21 carrying full length MalE	This study
pEXT21_ <i>spyAD</i>	pEXT21 carrying full length SpyAD	This study
pEXT21_ <i>AcrA</i>	pEXT21 carrying full length <i>C. jejuni</i> AcrA	479
pACYC_ <i>pglΔpglB</i>	pACYC carrying <i>pgl</i> glycosylation locus from <i>C. jejuni</i> with inactivated PglB	480
pEC415	Medium copy number plasmid with ampicillin resistant cassette and PBAD promoter	481
pEC415_ <i>malE</i>	pEC415 carrying full length MalE	This study
pEC415_ <i>spyCEP</i>	pEC415 carrying SpyCEP fragment	This study
pEC415_ <i>spyAD</i>	pEC415 carrying full length SpyAD	This study

**Table 2.3: Oligonucleotides used in this study.**

F = forward and R = reverse 5' – 3'. RE = Restriction Enzyme site.

Primer Name (F/R)	Sequence 5'to3'	RE Site	T <sub>m</sub>
T7 F	TAATACGACTCACTATAGGG		53.2
T7 R	GCTAGTTATTGCTCAGCGGT		57.3
pUCmut F	<u>CCATGGG</u> GACGTCTAAGAAACCATTATTATCATG	NcoI	58.9
pUCmut R	<u>CTCGAGG</u> TAACTGTCAGACCAAGTTTACTC	XhoI	59.3
kan F	<u>CTCGAGG</u> GACGCTCAGTGGAACGAAAAC	XhoI	59.8
kan R	<u>CCATGGC</u> CGCTCATGAATTAATTCTTAGAAAAAC	NcoI	59.8
pUCseq	GGTGGCCTAACTACGG		54.3
pUCmcs F	GCGTAAGGAGAAAATACCG		54.5
pUCmcs R	GCCGATTCATTAATGCAGCT		55.3
spyAD1 F	GCAGCTCAAACGGCTAATG		56.7
spyAD2 F	CCTAATGCGCCTGTTTACC		56.7
spyAD1 R	GCGCTCACGTATCACTTG		56
spyAD2 R	CCATAATGCCCTGATACCC		56.7
gacHJK F	CAAGCCTATCACAGTGGTG		56.7
gacHJK R	GCTGTTGATAAACTTCCTAAC		54.3
gacI Internal F	GGTTGAGGTATTAGTTCAATC		54
gacC F	GCAAACAGTTAACGACTGG		59.8
gacD R	CCAATAAAAGCGCTACTGG		59.4
PglB F	GAGGGCGCAAGAGATATGAT		62.5
PglB R	CATCGCTTAAAACCACTCCGT		62.5
pEXT21 MSC F	GTGAAGGTGAGCCAGTGAG		63
pEXT21 MSC R	CCGACTGGAAGCGGG		62.6
pEXT21seq F	GGTATGGCTGTGCAGGTCGT		67.6
pEXT21seq R	GCTTAATTTGATGCCTGGCA		62.3
SpyCEP F	GCCGCTGCTACTGGAG		62.5
SpyCEP R	CTCCAGTAGCAGCGGC		62.5
pEC_seqF	GACGAAATCACTCGTCTGTG		57.3
pEC_seqR	GCCACCTGACGTCTAAGAAA		57.3

### 2.2.2. DNA Isolation

Bacterial DNA was isolated from 1 colony forming unit (CFU) from an agar plate in 50 µl nuclease free water boiled for 10 minutes at 100 °C and spun for 5 minutes at 13,000 x g to isolate DNA in the supernatant and stored at -20 °C for colony Polymerase Chain Reaction (cPCR) purposes. Alternatively, 1.5 ml GAS overnight cultures were pelleted at 14,000 x g for 2 minutes incubated with 300 µl 5% w/v Chelex® 100 Resin (Biorad). Thoroughly vortexed suspensions were boiled for 10 minutes at 100 °C and spun at 13,000 x g for 10 minutes. Approximately 80 µl of the supernatant was removed and the DNA preparation stored at -20 °C.

QIAquick PCR purification Kit (Qiagen) was used to purify DNA for downstream processing. For isolation of plasmid DNA from *E. coli* a QIAprep Spin Miniprep Kit (Qiagen) or Monarch Plasmid Miniprep Kit (NEB) were used. All kits were used to manufacturer's instructions. All DNA preparations concentration and purity were measured using a NanoDrop1000 spectrophotometer (NanoDrop Technologies, Inc), and stored at -20 °C.

### 2.2.3. Polymerase Chain Reaction (PCR)

Oligonucleotide primers (IDT or Thermo Fisher Scientific) used in this work were added to PCR reactions containing; 1 x Phusion HF buffer (20 mM Tris-HCl, 100 mM KCl, 1 mM dithiothreitol (DTT), 0.1 mM ethylenediaminetetraacetic acid (EDTA), 200 µg/ml bovine serum albumin (BSA), 50% Glycerol, 1X stabilisers) (NEB), 0.2 mM deoxyribonucleotide triphosphate (dNTPs), 0.2 mM primers, 1 unit Phusion DNA polymerase (NEB), 1 µl DNA template and molecular grade distilled water to a final reaction volume of 50 or 20 µl. Standard thermo-cycling conditions were as follows; 98 °C for 30 sec; 98 °C for 10 sec, annealing and extension temperature tailored to specific reactions for 30 sec and 1 min/kb respectively for 35 cycles; and 72 °C for 4 mins for a final extension. These were run in either a DNA Engine Tetrad 2 Thermal cycler (Bio-Rad Laboratories Inc.) or a Veriti 9902 Thermal Cycler (Applied Biosystems ABI).

### 2.2.4. Gel Electrophoresis

PCR products were separated on 1 – 1.5 % w/v agarose gels stained with 1:10,000 SafeView Nucleic Acid Stain (NBS Biologicals) or GelRed Nucleic Acid Gel Stain (Biotium) with a 1 Kb DNA Hyperladder (Bioline). Samples were mixed with x4 DNA Loading Gel Dye (NEB) and run routinely at 100 - 120 volts and 500 mAmps in 1x Tris-borate-EDTA (TBE) buffer. After 30 – 60 minutes PCR products were visualised using a Gene Genius Bio Imaging System Gel Doc System (Syngene PXi).

Amplified DNA in some cases was purified from extracted agarose bands using a QIAquick Gel Extraction Kit (Qiagen) according to manufacturer's instructions for downstream ligations.

#### 2.2.5. *Restriction Digestion*

Plasmids used each Chapter are stated in Table 2.2. Digestion reactions used for cloning contained 1x CutSmart reaction buffer (10 mM Bis-Tris-Propane-HCl, 10 mM MgCl<sub>2</sub>, 100 µg/ml BSA) (NEB), 50 – 500 ng/µl of insert or plasmid DNA, 20 unit of restriction enzyme and molecular grade distilled to a final volume of 20 – 50 µl. Reactions were incubated at 37 °C for either 1 hour for insert DNA or 3 hours for plasmid DNA. 5' phosphate was removed by adding 5 units of Antarctic phosphatase (NEB) to prevent self-ligation of plasmids in 1 x Antarctic phosphate buffer (50 mM Bis-Tris-Propane-HCl, 1 mM MgCl<sub>2</sub>, 0.1 mM Zn<sub>6</sub>Cl<sub>2</sub>) (NEB). This was added directly to digestion reactions and left for 15 minutes at 37 °C before increasing the temperature specific to inactivate the enzyme (65 – 80 °C). Digested plasmids and inserts were resolved on 1 – 1.5 % agarose gels before being added to ligation reactions.

#### 2.2.6 *DNA Ligation*

Ligation reactions were set up to contain insert to plasmid DNA ratios 3 : 1 or 5 : 1. Either 1 unit of Quick-Stick Ligase (Bioline) or T4 DNA ligase (NEB) buffer (10 mM Tris-HCl, pH 7.4, 50 mM KCl, 1 mM DTT, 0.1 mM EDTA, and 50% glycerol) with 1x Quick-Stick Ligase (Bioline) or T3 DNA ligase (NEB) respectively were added. Reactions were left for 15 minutes or 3 hours at room temperature for Quick-Stick Ligase or T3 DNA ligase respectively. DNA was used for *E. coli* transformations or frozen at -20 °C.

#### 2.2.7. *Transformation and Clone Screening*

Heat shock was generally used to transform *E. coli* cloning strains. In short, thawed chemically competent *E. coli* cells were chilled on ice for 30 minutes and mixed gently with 1 - 5 µl ligation reactions. Cells were incubated at 42 °C for 30 seconds and then chilled on ice for 2 – 5 minutes. 950 µl of Super Optimal broth with Catabolite repression (SOC) broth was added to cells and incubated at 37 °C, 180 rpm shaking for 1 hour. Following recovery 150 µl of culture was spread onto LB agar plate containing appropriate antibiotic. Specific electrorotation methods to transform glycoengineering *E. coli* strains in Chapter 4, or GAS cells to generate a mutant in Chapter 3 are stated in the specific Chapters.

Single colonies were screened for desired constructs through cPCR with oligonucleotide primers designed to anneal at regions flanking the plasmids MCS (Table 2.3). Alternatively, restriction digests were used as a tool to confirm positive clones. A single colony was grown overnight in LB supplemented with appropriate antibiotic for plasmid retention and digested with specific restriction endonucleases flanking the MCS insert site and visualised on a 1 – 1.5 % w/v agarose gel for specific fragmentation patterns. Any potential insert positive clones were confirmed with DNA sequencing. Purified DNA and plasmid samples were sequenced using SourceBioscience or Eurofins Genomics according to company requirements.

### 2.3. Protein Techniques

#### 2.3.1. *SDS-PAGE*

For sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) 5 - 10 µl of sample was mixed with 1x NuPAGE lithium dodecyl sulfate (LDS) sample buffer (141 mM Tris base, 106 mM Tris HCl, 2% LDS, 0.51 mM EDTA, 0.22 mM SERVA Blue G-250, 0.175 mM phenol red) pH 8.5 (Thermo Fisher Scientific), and heated at 70 °C for 10 minutes. Pre-cast 10% or 4 – 12 % NuPAGE Bis-Tris gels (Thermo Fisher Scientific) were used in 1x 3-(N-morpholino)propanesulfonic acid (MOPS) (Thermo Fisher Scientific) or 1x 2-(N-morpholino) ethanesulfonic acid (MES) (Thermo Fisher Scientific) buffer chosen according to size of proteins for resolution. Gels were run for 30 – 60 minutes at 180 V. PageRuler Plus Prestained Protein Ladder (Thermo Fisher Scientific) was used as a protein marker. Gels were visualised by Coomassie or western blotting.

#### 2.3.2. *Coomassie Staining*

Gels were incubated with Coomassie stain (0.01 % Coomassie Brilliant Blue stain R250 (Thermo Fisher Scientific), 10% Acetic Acid (Sigma-Aldrich™), 40 % Ethanol (Thermo Fisher Scientific)) at room temperature for 30 minutes shaking. Gels were then de-stained in 10% Acetic Acid, 40 % Ethanol solution, and visualised with the Sygene Pxi gel doc system.

#### 2.3.3. *Protein Purification – Affinity Chromatography*

Proteins and glycoproteins were purified from lysates through their Histidine tags. For protein antigens used in chemical conjugation (Chapter 3), and large scale bioconjugation *E. coli* strains (Chapter 4) samples were applied to a NTA-Ni 1 ml HisTrap™ IMAC column (Cytiva), attached to ÄKTA start protein purification system (Cytiva). In short *E. coli* cells were lysed by resuspending the pellet in lysis buffer containing 500 mM Tris-HCl, 50 mM NaCl, 4 µg/ml DNase (Sigma-Aldrich™), 100



µg/ml lysozyme (Sigma-Aldrich™), and 1x Bugbuster® (Merck), pH 7.5. Lysates were spun at 13,000 x *g* for 10 minutes and the supernatant containing soluble protein / glycoprotein was 0.2 µm filter sterilised and applied to the column. At a flow rate of 1 ml/min wash buffer (50 mM Tris-HCl, 500 mM NaCl, 20 mM Imidazole, pH 7.5) was applied and equilibrated the chromatography system. Elution buffer (50 mM Tris, 50 mM NaCl, 250 mM Imidazole, pH 7.5) applied at 1 ml/min was used to elute His-tagged proteins / glycoproteins. Elution fractions corresponding to a peak on the UV<sub>280nm</sub> chromatograph trace of the nickel affinity purification were analysed by SDS-PAGE and western blot.

For small scale bioconjugation reactions (Chapter 4) NiNTA beads (Qiagen) were used. NiNTA beads were added directly to lysate samples, left rotating for 1 hour at 4 °C in a microcentrifuge tube. Bound His tagged protein samples were harvested through washing of the resulting pellet from a 5 minute spin at 13,000 x *g* with 1 ml of wash buffer (50 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, pH 8) three times. Each wash step was left for 10 minutes at 4 °C, before another 5 minute 13,000 x *g* spin, before the supernatant was discarded and 75 µl of elution buffer (50 mM Tris-HCl, 300 mM NaCl, 250 mM imidazole, pH 8) was added to the pelleted resin. The tube was inverted multiple times at room temperature before a final 13,000 x *g* centrifugation for 1 minute, and 75 µl of the supernatant containing the eluted protein / glycoprotein was removed and stored at - 20 °C before analysis.

#### *2.3.4. Protein Purification – Size Exclusion Chromatography*

Unless started otherwise, pooled fractions were applied to a HiPrep 16/60 Sephacryl S-300 HR column (Cytiva) attached to an ÄKTA start protein purification system (Cytiva). The equilibrated column was run with size exclusion buffer (500 mM Tris, 150 mM NaCl pH 7.5, or Phosphate Buffered Saline (PBS) pH 7.4), and samples collected in 1 ml fractions over 1.5 column volumes. Eluted fractions with a peak on the UV<sub>280</sub> trace were analysed for purity by Coomassie stained SDS-PAGE and by western blot.

#### *2.3.5. Endotoxin Removal*

To remove residual endotoxin from preparations, Pierce™ High Capacity Endotoxin Removal Resin (Thermo Fisher Scientific) was used. Prior to sample application, resin was regenerated overnight with 0.2 N NaOH, washed and equilibrated with endotoxin-free size exclusion buffer (PBS pH 7.5) three times, before samples were applied and incubated for 4 – 16 hours, followed by elution by a spin at 500 x *g* for 1 minute. Manufactures instructions were followed.

## 2.4. Immunoassays

### 2.4.1. *Western Blotting and Dot Blots*

iBlot 2 Dry blotting system (Thermo Fisher Scientific) was used for transfer of samples from SDS-PAGE Bis-Tris gels to nitrocellulose membranes. Transfer stacks were assembled and loaded as per manufacturer's instructions onto the iBlot and transferred for 7 minutes (1 minute at 20 V, 4 minutes at 23 V, 2 minutes at 25 V). After transfer nitrocellulose membranes were blocked with blocking buffer 1x PBS, 0.1 % Tween20 (Sigma-Aldrich™) (PBS-T), and 5% skimmed milk (Marvel)) for 1 hour. When blotting with soybean agglutinin (SBA) lectin (Vector Laboratories) (Table 2.1), milk was not used in PBS-T blocking buffer due to non-specific binding of the lectin. However, in general for antibodies, blocking buffer was replaced with buffer containing PBS-T, 1 % milk and the primary antibody incubated for 1 hour (Table 2.1) or PBS-T alone for lectin staining. Membranes were washed three times with 1x PBS-T for 5 minutes before incubation with the secondary antibody in PBS-T alone or 1 % milk supplementation (Table 2.1). All incubations and wash steps were conducted at room temperature on a plate rocker, with incubation with the secondary antibody being performed in the dark for fluorescent antibodies (Table 2.1). The membrane was washed three more times with PBS-T for 5 minutes before the fluorescent signal in two channels, 700 and 800 nm, was detected using an Odyssey CLx LICOR detection system (LI-COR Biosciences, USA). This uses a solid-state diode laser at wavelengths 685 and 785 nm respectively to scan the image dependant on the IRdye conjugated to the secondary antibody. Alternatively, blots stained with horseradish peroxidase (HRP) conjugated antibodies were developed with Pierce ECL detection (Thermo Fisher Scientific) and imaged with Sygene Pxi gel doc system. Such methods were used to assess recombinant protein and polysaccharide production, native polysaccharide extraction as well as verify glycoconjugate production.

Alternatively, 5 µl of samples were routinely dotted onto pre-cut 2 µm pore size nitrocellulose membranes (Thermo Fisher Scientific) and allowed to dry before blocking and development with appropriate antibody combinations and dilutions as stated above.

### 2.4.2. *Enzyme Linked Immunoassays (ELISA)*

All ELISA followed a similar format coating wells of MaxiSORP 96-well plates (Thermo Fisher Scientific) with 50 µl/well specific antibodies or purified recombinant protein antigens or polysaccharide diluted in carbonate buffer pH 9.6 (1.59 g/L sodium carbonate, 2.93 g/L sodium hydrogen carbonate) left overnight at 4 °C.

Routinely plates were washed 7 times with PBS-T and blocked with 100 µl/well Assay Diluent (AD) (PBS supplemented with 0.01 M EDTA, 0.3 % Tween 20 and 1% BSA (Sigma-Aldrich™)) for 1 hour at 37 °C, before washing again seven times with PBS-T. Samples were added and titrated either across or down the plate, specific to each assay as stated in Chapters 3 – 5 in a final volume of 50 – 100 µl/well, left for 1 – 2 hours at either room temperature or 37 °C. 100 µl/well AD alone was routinely used to assess assay background. 100 µl/well of specific primary antibodies for detection were routinely diluted in AD buffer and added to washed wells left for 1 hour at room temperature. If necessary 100 µl/well secondary antibodies were applied to washed wells diluted in AD buffer for 1.5 hours at 37 °C. Plates were then washed 7 times, and plates submerged in PBS-T for 10 – 15 minutes where non-specific reactivity was observed to minimise background. Plates were developed by addition of 100 µl/well room temperature 3,3',5,5'-tetramethylbenzidine substrate (TMB) (Thermo Fisher Scientific) incubated at room temperature for 15 minutes before the reaction stopped by adding 100 µl/well of 1 M sulphuric acid. Plates were read at 450 nm wavelength on a VMax® Kinetic ELISA Microplate Reader (Molecular Devices) and antibody titres calculated by the serum dilution which gives OD<sub>450nm</sub> of 0.5 using the interpolation of a sigmoidal curve in GraphPad Prism.

## 2.5. Quantification Techniques

### 2.5.1. *Protein Quantification*

The bicinchoninic acid (BCA) assay (Thermo Fisher Scientific) was used for quantification of protein / glycoprotein samples in a 96-well plate format. According to manufacturer's instructions known concentrations of BSA were used to generate a standard curve to enable calculation of protein concentration by rearranging  $y=mx+c$  ( $y$  = absorbance of protein of interest,  $m$  = gradient,  $c$  =  $y$  intercept, and  $x$  = concentration of protein of interest). Standards and sample were diluted in the same buffer, often PBS pH 7.4, and measured at 570 nm wavelength using VMax® Kinetic ELISA Microplate Reader (Molecular Devices). Data was plotted in Microsoft Excel to generate a standard curve to determine protein concentration.

### 2.5.2. *Polysaccharide Quantification*

GAC variants (Chapter 3) and *E. coli* produced recombinant rhamnose variants (Chapter 4) were quantified using Anthrone colourimetric assay, use to detect hexoses through the formation of a yellow-green coloured complex. A 1 mM stock of polysaccharide repeating unit comprising of either 2 x L-rhamnose (Sigma-Aldrich™) 1 x GlcNAc (Sigma-Aldrich™), or just 2 x L-rhamnose (Sigma-Aldrich™),

was prepared in water and used to generate a standard curve (7.5 – 100 nM). 0.2% anthrone (Sigma-Aldrich™) was prepared in sulphuric acid (Sigma-Aldrich™). In triplicate, either standards or samples were diluted in the same buffer matrix, either water or PBS pH 7.4, in a final volume of 300 µl, with 600 µl of anthrone reagent added to the tubes before they were mixed by vortexing. Samples and standards were heated for 15 minutes at 95 °C in an oil bath before cooling to room temperature. 300 µl of each reaction was added to a flat bottomed 96 well plate before absorbances were read at 625 nm measuring hexose chromogenic cyclic-furfural intermediate generation by POLARstar Omega plate reader spectrophotometer (BMG LABTECH).

Buffer blank subtracted absorbances were used to determine the concentration of unknown polysaccharide containing samples based on the repeating unit standard concentrations and absorbances. A conversion factor based on these repeating units was also used to calculate resulting concentrations. Data was plotted in Microsoft Excel to generate a standard curve. Within the range of the assay the absorbance is directly proportional to the amount of hexose present within a sample. Note that small scale Anthrone assays were also carried out due to sample volume constraints. Assays were carried out with reduced total reaction volumes of either 600 µl (1 : 2 sample : anthrone), or 400 µl (1 : 4 sample : anthrone) with overall reduced sensitivity of 15 nMol repeating unit standard.

## 2.6. Data Analysis and Presentation

### 2.6.1. *Statistical Analysis*

For all statistical tests,  $p < 0.05$  was considered significant. One-way ANOVA, Kruskal-Wallis, Mann-Whitney U tests, and Unpaired T test were performed in GraphPad Prism 9.3.1 for Windows, GraphPad Software, San Diego, California USA, [www.graphpad.com](http://www.graphpad.com). All graphs were also plotted using GraphPad Prism Version 9.3.1.

### 2.6.2. *Scientific Diagrams*

All diagrams in this thesis are original and created in BioRender with a publication license as the proof of publication rights. These are identified with “Created with BioRender.com.” in such figure legends.

## CHAPTER 3

# Use of Chemical Conjugation to Manufacture GAS Glycoconjugate Vaccines

### 3.1. Introduction

Following Avery and Goedel's discovery, conjugation, the covalent attachment of two or more molecules has been routinely used to enhance polysaccharide based vaccine immunogenicity<sup>460</sup>. Licenced glycoconjugate vaccines, comprised of polysaccharide or capsule and protein components attached using chemical approaches, are successful due to such components providing the immune system with multiple triggers, increasing the vaccine's protective lifetime through memory responses<sup>468</sup>. The work in this Chapter investigates the suitability of wildtype and modified Group A Carbohydrate (GAC) as potential glycoconjugate polysaccharide vaccine components, for coupling to a panel of GAS specific protein carriers, using chemical conjugation methods.

#### 3.1.1. Group A Carbohydrate (GAC) Properties Related to Glycoconjugate Vaccine Production

GAC is a species defining conserved feature of all circulating GAS serotypes<sup>137,147,482</sup>. Antibody binding experiments have shown that GAC polymers are mainly localised to the outer surface of the GAS cell wall<sup>137</sup>, via covalent phosphate ester linkages to N-Acetyl-Muramic acid (MurNAc) within the repeating N-Acetyl-Glucosamine (GlcNAc) -MurNAc chain of peptidoglycan<sup>483,484</sup>. GAC can also be interconnected via phosphodiester bonds within the "mesh-like" structure of the peptidoglycan<sup>152,483</sup> being highly abundant making up 40 to 60% of the total cell wall mass<sup>137</sup>. This surface localisation makes GAC of interest as a glycoconjugate polysaccharide component from an immunogenicity perspective, as well as practically for extraction from GAS cells for chemical coupling reactions.

Immunological studies have revealed that both GAC<sup>409,445</sup> and synthetic GAC related oligosaccharide fragments of varying chain lengths<sup>139,485</sup> are immunogenic and accessible to antibody binding despite encapsulation<sup>139</sup>. Additionally, anti-GAC antibodies have been shown to be protective, highlighted in a study of Mexican children which showed high titres of GAC specific IgG antibodies correlated with protection against throat carriage<sup>409</sup>.

Therefore, GAC has been included as an antigen in GAS specific vaccines, but alone is unlikely to be strongly immunogenic, as polysaccharides fail to induce T cell memory responses. In order to boost GAC immunogenicity, a small number of groups have conjugated it to a small panel of protein carriers. These have included classical carrier proteins such as Tetanus Toxoid (TT)<sup>409</sup> and CRM<sub>197</sub>, a non-toxic mutant diphtheria toxin<sup>139</sup>. TT-GAC conjugates protected mice from intranasal colonisation and intraperitoneal lethal challenge of two different M GAS serotypes<sup>409</sup>, whereas GAC conjugated to CRM<sub>197</sub> investigated size dependent immunogenicity, using varying lengths of synthetic polymers, also generated protective antibodies<sup>139</sup>.

Despite promise with wildtype GAC vaccines, the GlcNAc sidechains on the polyrhamnose backbone has been associated with ARF development following serological specificity studies<sup>486</sup>. Experimentally both within this work, and the work of others, the discovery of the *gac* operon has allowed the generation of GAS mutants possessing a rhamnose polymer devoid of GlcNAc sidechains through knockout of the *gacI* gene<sup>141</sup>. Recent data has shown antibodies also recognise rhamnose epitopes within polymer backbones<sup>141,412</sup>. Therefore, such modified GAC variants have been evaluated as a universal GAS vaccine antigen, testing their immunogenicity, and potential to improve vaccine safety. When conjugated, modified GAC was able to generate anti-GAC antibody responses comparable to wildtype GAC<sup>141</sup>, showing disease reduction with partial protection from bacteraemia and skin infections<sup>396</sup>, without any cross-reactivity with human heart or brain tissue lysates<sup>412</sup>.

Such studies therefore show promise for both wildtype and mutant GAC polysaccharides as a glycoconjugate vaccine component, but there is limited information on the direct comparison of the performance of each polysaccharide, and the effects of their conjugation to novel protein carriers. This will be discussed in the context of using chemicals for linkage.

### 3.1.2. Glycoconjugate Protein Carriers

Protein carriers selected for inclusion in glycoconjugate vaccines are ideally themselves immunogenic, which when conjugated to polysaccharides antigens can enhance immunogenicity via long-lasting T-cell dependent and memory antibody responses. Suitable protein carriers must therefore enable induction of effective anti-polysaccharide immune responses and be compatible with conjugation techniques, being safe and produced at high yields and low cost<sup>471,472,487</sup>.

### 3.1.2.1. *Traditional Chemical Glycoconjugate Protein Carriers*

Traditionally carrier proteins included in currently licenced glycoconjugate vaccines are heterologous to the vaccine candidate organism<sup>66,472</sup>. There are currently five carrier proteins in licenced glycoconjugate vaccines; TT, DT, CRM<sub>197</sub>, recombinant *E. coli* produced *Haemophilus* protein D (PD), and outer membrane protein complex of serogroup B meningococcus (OMPC)<sup>66,472</sup> (Table 3.1). TT, DT and CRM<sub>197</sub> are detoxified bacterial toxins by chemical or genetic means, routinely used in glycoconjugate vaccines due to established safety records<sup>487</sup>, extensive vaccination history<sup>472</sup> and availability at industrial scale. OMPC and PD are less established, with OMPC used for a Hib<sup>488</sup> and a second generation pneumococcal conjugate vaccine<sup>489</sup>, and PD for a multivalent pneumococcal conjugate vaccine<sup>490,491</sup>. These carriers have been successful due to their compatibility with well-characterised conjugation chemistries, and their ability to induce effective long lasting anti-polysaccharide immune responses<sup>471</sup>. TT will be used in this work as a classical carrier protein, to validate the selected conjugation chemistry, as well as facilitate comparisons with GAS specific protein carriers.

### 3.1.2.2. *Novel and Pathogen Specific Protein Carriers*

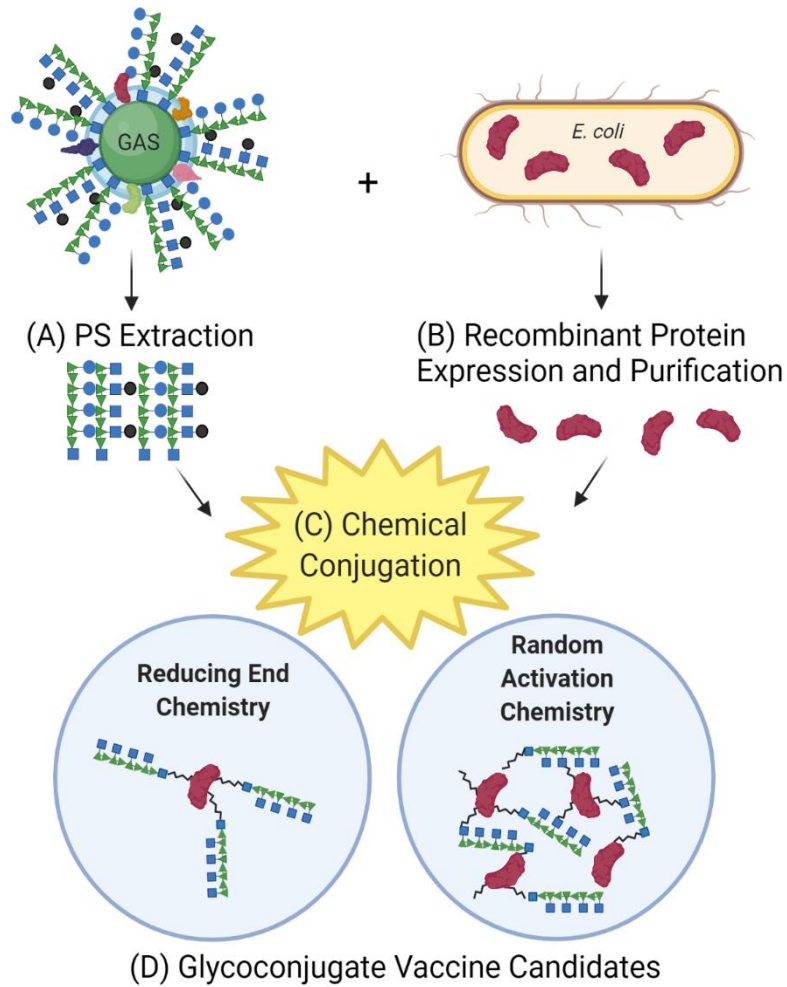
Compared to bacterial polysaccharides, there has been less attention on protein carrier properties in chemical glycoconjugate vaccines<sup>471</sup>. However, following evidence of sub-optimal performance with traditional protein carriers<sup>473-476</sup>, it is of interest to investigate alternative protein antigens for inclusion. This has been exploited for chemical coupling reactions with ExoA protein, a recombinant non-toxic form of *Pseudomonas aeruginosa* exotoxin A used in a chemical glycoconjugate in clinical trials<sup>492-494</sup>. In terms of GAS specific protein antigens selected as carriers within chemical glycoconjugate vaccines, a number have been investigated over the last few years. Specifically these include C5a peptidase protein<sup>413,447</sup>, SpyCEP and SLO<sup>411</sup>, as well as SpyAD<sup>412</sup>. In these studies, addition of GAS specific proteins have shown favourable immunogenicity, generating opsonophagocytic IgG antibodies, which can protect mice from disease, as well as reduce mortality following GAS challenge<sup>412,413,447</sup>.

### 3.1.3. Chemical Conjugation Approaches

Chemical conjugation approaches generally requires extraction and purification of polysaccharide directly from the organism's cell wall or capsule against which you wish to protect against before covalent attachment to purified recombinantly expressed protein carriers (Figure 3.1a-c). It may also entail recombinant production of both protein and polysaccharide components before chemical coupling.

Polysaccharides can be attached to proteins by applying a variety of different chemical approaches. Conjugation chemistries can utilise naturally present reactive groups or alternatively add in cross-linking reagents to artificially introduce compatible reactive groups for attachment. Such cross-linking reagents facilitate either site directed attachment, where the linker is released from the final product or alternatively can become part of the conjugate structure. Incorporation of linkers to either protein or polysaccharide components can render them more prone to conjugation improving attachment levels, however, it can also damage conformational epitopes in the process. Alternatively, or in addition to linker attachment, polysaccharide or proteins can be chemically modified through activation approaches to yield compatible reactive groups for attachment. This can be through reducing end selective terminal activation, or random multiple activation along polysaccharide chains<sup>495</sup> (Figure 3.1d). The choice of chemical conjugation approach, and whether to include linkers and / or activation / derivation steps, is often governed by structure, size, and composition of vaccine components<sup>496</sup>. For example larger polysaccharides are usually randomly activated, and smaller polysaccharides activated at the reducing end to preserve protective epitopes<sup>495</sup>.





**Figure 3.1: Simplified schematic representation of chemical conjugation.**

Polysaccharide (PS) is extracted from the native organism, e.g. GAC from GAS cells **(A)**, and recombinant protein carriers expressed and purified from *E. coli* cells **(B)**. Extracted polysaccharide and recombinant protein carriers containing compatible reactive groups are conjugated together using a compatible cross-linking chemical reagent **(C)** yielding heterogenous glycoconjugate vaccine candidates depending on the selected approach **(D)**. Schematic shows different approaches yielding different glycoconjugate species, with reducing end chemistries leading to terminal single ended glycoconjugate products, termed “sun-like” structures, and random activation chemistries yielding cross-linked “mesh-like” structures of higher molecular weights with several attached protein - polysaccharide molecules.

Original figure created with BioRender.com.

No single chemical conjugation method is preferred in manufacturing. Table 3.1 highlights the variety of chemical approaches used in licenced commercially available glycoconjugate vaccines. There are however common reactive groups targeted by these chemistries, for example amine groups on carrier proteins, usually in the form of surface exposed lysine residues. Carboxyl groups are also commonly targeted by these chemistries on either protein's glutamic or aspartic acid residues, or along polysaccharide chains. Such groups can be conjugated to amino or hydrazido derivatives by carbodiimide mediated condensation, while amino groups can be coupled with introduced groups such as succinmido ester, cyano-ester and squarate or direct reductive amination on polysaccharides<sup>458,497</sup>.

Generally, linkage of polysaccharide to carrier proteins results in heterogenous populations with random displays of glycoconjugate molecules stimulating the immune system. Some selectivity can be achieved by modulating component stoichiometry<sup>498</sup>, or targeting or introducing specific non-natural amino acids on the protein facilitating site directed attachment<sup>499</sup>. Such selective or milder approaches can be important for carrier proteins with a dual purpose, such as GAS antigens used in this study, to maintain protective B and T cell epitopes. For this reason, this work investigates N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC) chemistry as a zero-length linker, reducing end targeting chemical approach, based on compatible groups predicted to be present on enzymatically extracted GAC variants (-COOH) and recombinant protein carriers containing surface exposed lysine residues (-NH<sub>3</sub>).

**Table 3.1: Choice of chemical conjugation strategies in licensed commercially available glycoconjugate vaccines.**

The table denotes the vaccine manufacturer, carrier protein, polysaccharide and conjugation chemistry used to produce glycoconjugate vaccines against particular pathogens.

Pathogen	Manufacturer	Carrier Protein	Polysaccharide	Conjugation Chemistry
<b><i>Haemophilus influenzae</i> type B (Hib)</b>	ActHIB/Sanofi Pasteur	TT	Native polysaccharide	-
	Hiberix/GSK vaccines	TT	Size-reduced polysaccharide	-
	Quinvaxem/GSK vaccines	CRM <sub>197</sub>	Depolymerised polysaccharide	Active ester chemistry
	PedvaxHIB/Merck	OMPC	Native polysaccharide	-
<b><i>Neisseria meningitidis</i> serogroup C</b>	NeisVac-C/Pfizer	TT	Native polysaccharide	Reductive amination
	Meningitec/Nuron Biotech	CRM <sub>197</sub>	Native polysaccharide	Reductive amination
	Menjugate/GSK vaccines	CRM <sub>197</sub>	Depolymerised polysaccharide	Active ester chemistry
	Menitorix/GSK vaccines (with Hib)	TT	Size-reduced polysaccharides	-
<b><i>Neisseria meningitidis</i> serogroup CY</b>	MenHibrix/GSK vaccines (with Hib)	TT	Native polysaccharides	-
<b><i>Neisseria meningitidis</i> serogroup ACWY</b>	Menactra/Sanofi Pasteur	DT	Depolymerised polysaccharide	-
	Menveo/GSK	CRM <sub>197</sub>	Depolymerised polysaccharide	Active ester chemistry
	Nimenrix/Pfizer	TT	Size-reduced polysaccharide	Active ester chemistry
<b><i>Streptococcus pneumoniae</i> serogroup 4, 6B, 9V, 14, 18C, 19F, 23</b>	Prevnar/Pfizer	CRM <sub>197</sub>	Native polysaccharide	Reductive amination
<b><i>Streptococcus pneumoniae</i> serogroup 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F, 23F</b>	Synflorix/GSK	NTHi PD DT TT	Size-reduced polysaccharide	Reductive amination
<b><i>Streptococcus pneumoniae</i> serogroup 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 23F</b>	Prevnar13/Pfizer	CRM <sub>197</sub>	Native polysaccharide	-

- Information not available.

## 3.2. Aim, Hypothesis and Objectives

This Chapter describes work to develop GAS specific glycoconjugates using a traditional chemical conjugation method. The glycoconjugates developed using this approach will be compared to bioconjugates manufactured in *E. coli* host cells using *C. jejuni* PglB oligosaccharyltransferase enzyme described in Chapter 4, with their immunogenicity investigated in Chapter 5.

This work utilises EDC conjugation chemistry to attach Group A Carbohydrate (GAC), extracted from wildtype and mutant GAS cells to novel recombinant GAS protein antigens as carriers. These will be compared to the classical protein carrier Tetanus Toxoid (TT) conjugated to GAC.

**Aim:** To investigate the suitability of EDC conjugation for attachment of GAC variants to novel GAS antigens (MalE, SpyCEP, and SpyAD) and classical TT as carrier proteins.

### **Hypothesis:**

- EDC chemical conjugation will be a suitable method for attachment of GAC variants to both novel GAS antigens (MalE, SpyCEP, and SpyAD) and TT as carrier proteins to build glycoconjugate vaccines.

### **Objectives:**

- 1) Select appropriate GAS proteins to act as carriers based on pathogenesis function, cellular localisation, and immunological properties.
- 2) Clone, express and purify a select panel of GAS proteins and prepare monomeric TT classical carrier for comparison purposes.
- 3) Produce a GAS *gacI* mutant for GlcNAc deficient GAC and enzymatically extract and purify wildtype and mutant GAC (GAS\_Rha) from GAS cells.
- 4) Assess purified proteins and GAC variants reactive compatibility and perform EDC based conjugation chemistry reactions to conjugate protein carriers (GAS antigens and TT) to wildtype and GAS\_Rha variants.
- 5) Determine successful conjugation and quantify the conjugated / unconjugated components to calculate reaction efficiency and conjugate molar ratios.
- 6) Determine size and physiochemical properties of conjugate components and resulting glycoconjugate preparations.

### 3.3. Materials and Methods

#### 3.3.1. Bacterial Strains

GAS and *E. coli* strains description and source used are listed in Table 3.2. All strains were routinely stored in 20% glycerol (Sigma-Aldrich™) at -80 °C.

**Table 3.2: List of *E. coli* and GAS strains used in chemical conjugation work.**

Strains	Description	Source
<b>Commercial <i>E. coli</i></b>		
<b>NEB® 5α Stable Competent</b>	Commercially available DH5 derivative strain, chemically competent for cloning and plasmid propagation.	NEB
<b>BL21 (DE3)</b>	Commercially available DE3 derivative strain, competent for expression of His-tagged GAS protein antigens. F-ompT hsdSB(rB - mB-) gal dcm(DE3) phenotype.	NEB
<b><i>Streptococcus pyogenes</i></b>		
<b>NCTC-8198</b>	<i>S. pyogenes</i> (GAS) strain 8198 from NCTC used to extract wildtype Group A Carbohydrate (GAC).	Bacterial culture collections, Public Health England (PHE).
<b>Δ<i>gacI</i> NCTC-8198 Mutant</b>	<i>S. pyogenes</i> (GAS) strain 8198 from NCTC with <i>gacI</i> gene deletion used to extract polyrhamnose (GAS_Rha).	This study

#### 3.3.2. Bacterial Growth Conditions

*E. coli* strains were routinely cultured in Luria-Bertani (LB) (Invitrogen™) broth at 37 °C in an aerobic atmosphere (Innova-32). For recombinant GAS protein antigen expression to be used in chemical conjugations, BL21 (DE3) *E. coli* were grown in OverNight Express instant TB media (Merck) supplemented with 10 ml/L glycerol. Agar plates and liquid cultures of *E. coli* were supplemented with 50 µg/ml kanamycin for propagation of pET28a plasmids and modified pUC19 derivatives for Δ*gacI* GAS mutant generation.

*S. pyogenes* M1 strain NCTC-8198 was routinely cultured in Todd Hewitt (TH) broth (Sigma-Aldrich™) supplemented with 0.2% yeast extract (Sigma-Aldrich™), or on TH agar plates (TH broth supplemented with 1 % agar) (Oxoid Thermo Scientific). GAS was incubated at 37 °C in a static 5% CO<sub>2</sub> incubator (Binder).

#### 3.3.3. GAS Carrier Protein Selection, Cloning, Expression and Purification

##### 3.3.3.1. *Bioinformatic Analysis*

Artemis Release 17.0.1 (<https://sanger.ac.uk/science/tools/artemis>) was used for the visualisation of whole genome sequences, and Geneious Prime Version

2020.0.3 was routinely used for cloning design and analysis of DNA sequence reads. DNA and protein sequences were extracted from genome files and used for further analysis, including BLAST against DNA Sequences selected from the NCBI Database.

For protein analysis, SignalP-5.0 topology prediction server (<http://www.cbs.dtu.dk/services/SignalP/>) was used for signal peptide detection and the overall protein structure was modelled using Phyre<sup>2</sup> (Protein Homology/analogy Recognition Engine V 2.0) <http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id>)<sup>500</sup>.

Whole protein sequences were tested by VaxiJen ([www.ddg-pharmfac.net/vaxijen](http://www.ddg-pharmfac.net/vaxijen/)), and Immune Epitope Database and Analysis Resource (IEDB) (<https://www.iedb.org/>) indicating protein immunological properties based on sequence auto cross covariance, and catalogues of experimental data on B and T cell epitopes respectively.

#### 3.3.3.2. *GAS Protein Carrier Cloning Strategy*

GAS protein carrier DNA sequences were synthesised as gBlocks® Gene Fragments (IDT) with cloned sequences stated in Appendix section 7.1.1 . N-terminal secretion signals and C- terminal sortase or hydrophobic domains identified using SignalP-5.0 topology prediction server were removed along with stop codons. SpyCEP was cloned as a predicted immunogenic fragment (amino acid residues 35 – 587) of the full-sized protein rendering it functionally inactive<sup>208</sup>.

The gene sequences were engineered to contain NcoI (CCATGG) and XhoI (CTCGAG) restriction sites at the 5' and 3' termini respectively for ligation into pET28a plasmid (Novagen) for the addition of a C-terminal x6 Histidine tag. Additional GC bases were added after the compatible start codon ATG within NcoI to transcribe the gene in frame. Restriction digest cloning was conducted as stated in general methods section 2.2.5.

#### 3.3.3.3. *GAS Antigen Expression and Purification*

Sequence confirmed plasmids were transformed into BL21 (DE3) *E. coli* via heat shock methods (general methods section 2.2.7). Small scale expression studies comprised of 2 ml overnight cultures of transformed BL21 (DE3) *E. coli* in either autoinduction Overnight Express Instant TB Medium (Merck) or in LB as a negative control without induction supplemented with kanamycin antibiotic. Cell pellets were obtained by centrifugation at 4,000 x g for 10 minutes and standardised to OD<sub>600nm</sub> = 20/ml in cell lysis buffer (50 mM Tris pH 7.5, 50 mM NaCl, 4 µg/ml DNase (Sigma-

Aldrich™), 100 µg/ml lysozyme (Sigma-Aldrich™), and x1 Bugbuster® (Merck)). Samples were incubated at room temperature for 15 minutes with rotation, then 10 µl of total whole cell lysate was put to one side and the remainder spun at 13,000 x g for 10 minutes. The soluble fraction was present in the supernatant and the insoluble fraction in the pellet which was resuspended in 100 µl of PBS supplement with 1 % SDS. Recombinant protein expression localisation was determined by Coomassie staining SDS-PAGE (general methods section 2.3.1) with induction compared to the uninduced controls.

Large scale cultures were grown in 300 ml autoinduction Overnight Express Instant TB Medium (Merck) in baffled flasks at 37 °C until stationary phase (>16 hours). Cells were harvested at 4,000 x g for 10 minutes and pellets stored at -20 °C. Thawed cells were treated with cell lysis buffer as above for 30 minutes at room temperature with rotation followed by 30 minutes at 37 °C. Lysates were spun at 13,000 x g for 10 minutes and supernatant containing soluble protein was removed and 0.2 µm filter sterilised for immediate downstream purification. His-tag purification followed by size exclusion chromatography was conducted as stated in general methods sections 2.3.3 and 2.3.4.

Amicon® Ultra-15 Centrifugal Filters (Merck) with a Molecular Weight Cut Off (MWCO) of 3,000 Da were used to concentrate pooled recombinant protein fractions through spinning 4,000 x g for 15 minutes to reduce sample volume (< 1 ml). 100 µl aliquots of recombinant purified proteins were stored at -80 °C and concentration assessed by Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific) (general methods section 2.5.1). Limulus Amebocyte Lysate (LAL) test was performed by the Pyrogens Team (NIBSC) assessing endotoxin levels following treatment of recombinant antigens with Pierce™ High Capacity Endotoxin Removal Resin (Thermo Fisher Scientific) (general methods section 2.3.5).

#### 3.3.4. Preparation of Tetanus Toxoid (TT) Monomer

The WHO 2nd IS for TT (code 04/150) was obtained from the Centre for Biological Reference Materials (see <https://www.nibsc.org/products>). Monomeric TT was obtained following resuspension of TT in 0.1 M PBS pH 7.1. One ml (0.026 g, 690 Lf units) of TT was applied to a HiPrep 16/60 Sephacryl S-300 HR column (Cytiva) equilibrated with 0.2 µm filtered 0.1 M PBS pH 7.1 attached to ÄKTA start protein purification system (Cytiva) (general methods section 2.3.4). Eluted proteins were analysed for purity by Coomassie stained SDS-PAGE and fractions containing purified protein pooled.

### 3.3.5. *ΔgacI* Mutant Generation for Polyrhamnose Polysaccharide (GAS Rha) Production

#### 3.3.5.1. *Production of Chemically Competent GAS*

GAS cells were grown in THY broth supplemented with 20 mM glycine overnight, sub-cultured and grown to early log-phase ( $OD_{600nm}$  0.2 – 0.4) before harvest and incubation on ice for 30 minutes. Cells were pelleted following 5,000 x *g* spin and washed three times with ice cold 10% sterile glycerol solution. Cells were resuspended in ~ 600  $\mu$ l 10% glycerol and used in electroporation reactions.

#### 3.3.5.2. *Electroporation of GAS*

Electroporation cuvettes were chilled before 200  $\mu$ l GAS electrocompetent cells were added to 30  $\mu$ g plasmid DNA incubated on ice for 2 minutes. Gene Pulser system was set to 1.75 kV voltage, 400 ohms resistance, 25  $\mu$ F capacitance, giving a time constant between 8 – 10. Following electroporation, GAS cells were outgrown in THY broth without antibiotic selection, grown statically for 2 hours at 37 °C 5% CO<sub>2</sub>, before plating on THY plates supplemented with kanamycin. Colonies which grew on kanamycin were assumed to be plasmid single crossover integrated clones. Colonies were re-streaked on non-selective plates four times, then used in genomic PCR screens, obtained by 5% Chelex treatment (general methods section 2.2.2) to determine plasmid loss and gene deletion (double crossover events).

### 3.3.6. Polysaccharide Extraction and Purification

#### 3.3.6.1. *PlyC Purification for Enzymatic Digestion of GAS Cell Wall*

An *E. coli* strain harbouring pET28a plasmid encoding the genetic cluster required for PlyC enzymatic activity (*plyCA* active site and *plyCB* cell wall binding domain, functioning as an octamer)<sup>501</sup> was cloned by Dr. Shaw. Recombinant PlyC *E. coli* were lysed in lysis/wash buffer (5 mM sodium phosphate, 1 mM EDTA, 1 mM DTT, pH 6.1) supplemented with 4  $\mu$ g/ml DNase (Sigma-Aldrich™), 100  $\mu$ g/ml lysozyme (Sigma-Aldrich™) at 37 °C for 1 hour. Suspensions were spun at 17,000 x *g* 10 minutes and supernatant was 0.2  $\mu$ m filter sterilised and applied to an HiTrap Q HP column (Cytiva) attached to ÄKTA start protein purification system (Cytiva). Lysis/wash buffer equilibrated the chromatography system, and elution buffer (5mM sodium phosphate, 1 mM EDTA, 1 mM DTT, 1 M NaCl, pH 6.1) eluted purified PlyC protein. Elution fractions were analysed by Coomassie stained SDS-PAGE and pooled before being applied to a HiPrep 16/60 Sephacryl S-300 HR column (Cytiva) equilibrated with size exclusion buffer (general methods section 2.3.4). Pooled fractions purity was analysed by Coomassie stained SDS-PAGE.



### 3.3.6.2. *GAC / GAS\_Rha Enzymatic Extraction*

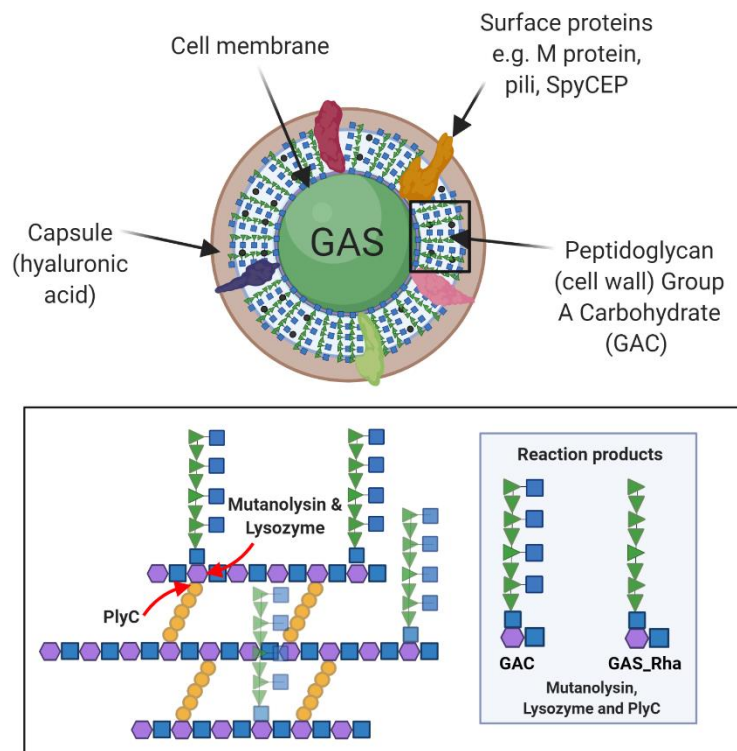
GAS cells were harvested from overnight cultures (>16 hours) through centrifugation at 5,000 x g for 10 minutes. Pelleted cells were resuspended and boiled in 50 mM Tris-HCl, 150 mM NaCl, 1% SDS, pH 7.4 for 10 minutes, and centrifuged at 5,000 x g for 10 minutes to re-pellet cells. Residual SDS was removed by 50 mM Tris-HCl, 150 mM NaCl, pH 7.4 washes, and remaining pellets resuspended in a reduced volume of 50 mM Tris-HCl, pH 7.4 buffer. Cells were mechanically disrupted through 0.2 mm glass beads in a ribolyser FastPrep®-24 for 3 – 5 rounds of 60 seconds 6.4 m/s and supernatants removed once beads settled. Beads were washed by resuspension in buffer and supernatants removed once beads were settled. This was repeated until added buffer was transparent following bead resuspension. Pooled washes were added to 20 mM MgSO<sub>4</sub>, 50 µg/ml DNase, 50 µg/ml RNase, 30 µg/ml hyaluronidase, incubated at 37 °C for 2 hours, before 20 µg/ml proteinase K was added and left at 37 °C overnight to remove cell components such as DNA, RNA, and proteins.

Samples were spun at 5,000 x g for 20 minutes and the peptidoglycan pellet resuspended in PlyC buffer (5 mM sodium phosphate, 1 mM EDTA, 1 M NaCl, pH 6) with 2 mg/ml lysozyme (Sigma-Aldrich™), 200 U/ml Mutanolysin (Sigma-Aldrich™), 1 cComplete™ EDTA-free Protease Inhibitor Cocktail (Roche), and 20 U/mg recombinant PlyC cell wall hydrolase incubated at 37 °C for 24 hours. Digested cell walls were centrifuged at 5,000 x g for 20 minutes to pellet undigested contaminants and the digested cell walls in the supernatant removed for GAC purification (Figure 3.2).

### 3.3.6.3. *GAC / GAS\_Rha Purification*

Released cell wall components were 0.2 µm filtered and buffer exchanged in Amicon® Ultra-15 Centrifugal Filters (Merck) devices with a 3,000 Da MWCO equilibrated with loading buffer (5 mM sodium phosphate, 20% EtOH pH 8.0). Buffer exchanged polysaccharide was applied to a 5 ml HiTrap™ Canto™ adhere column (Cytiva), attached to ÄKTA start protein purification system (Cytiva), using the sample pump following system equilibration with 5x column volumes of loading buffer. Pump B was equilibrated with elution buffer (5mM sodium phosphate, 20% EtOH, 200 mM NaCl pH 8.0) before 3x column volumes of wash buffer was applied at a flow rate of 1 ml/minute to remove unbound material such as protein impurities. Bound polysaccharide was eluted from the column with 100 % elution buffer containing salt eluent. Based on UV<sub>280nm</sub> chromatographs the resulting elution fractions containing purified polysaccharide were collected and pooled before

dialysed using a Slide-A-Lyzer® Dialysis Cassette (Thermo Fisher Scientific) with 20,000 Da MWCO three times in dH<sub>2</sub>O overnight.



**Figure 3.2: Enzymatic extraction method of GAC and GAS\_Rha from wildtype M1 NCTC-8198 strain and associated mutant  $\Delta gacI$  GAS cells.**

Schematic of a GAS cell indicating location of cell wall anchored GAC. Boxed structure shows GAC attached to the peptidoglycan layer (cell wall) as a schematic of enzymatic targeting of the GAS peptidoglycan layer. Mutanolysin cleaves  $\beta$ 1,4 glycosidic bonds between GlcNAc (blue square) and MurNAc (purple hexagon), and PlyC cleaves between the lactyl moiety from MurNAc and the first L-Ala of the interconnecting peptide chain (yellow circles) as indicated by red arrows. Inset box shows resulting reaction products used for conjugation reactions.

Original figure created with BioRender.com.

### 3.3.7. Polysaccharide Characterisation

Dialysed GAC / GAS\_Rha samples were routinely visualised by SDS-PAGE and immunoblot, and quantified by anthrone assay as stated in general methods section 2.5.2.

#### 3.3.7.1. *Latex Agglutination*

Streptococcal Lancefield group antigens were identified by latex agglutination tests (Remel PathoDx) performed according to the manufacturer's instructions on GAS colonies grown on THY agar. Bacterial colonies were incubated with specific beads with Remel™ Streptex™ Acid Extraction reagents.

#### 3.3.7.2. *Lectin Staining*

Overnight cultures were centrifuged at 4,000 x g for 20 minutes, resuspended in PBS, pH 7.4 supplemented with 30 µg/ml hyaluronidase and incubated at 37 °C for 2 hours to remove the HA capsule. Treated GAS cultures were pelleted at 4,000 x g for 20 minutes and washed twice more in the original culture volume of PBS before being normalised to an OD<sub>600nm</sub> of 0.4 or 0.8, for early or late logarithmic growth phase cultures respectively. Suspensions were dotted onto nitrocellulose membranes and blocked in 1X Carbo-Free™ Blocking Solution (Vector Laboratories) for 30 minutes at room temperature. Biotinylated succinylated wheat germ agglutinin (sWGA) (Vector Laboratories) was added for 30 minutes at room temperature. The nitrocellulose membrane was washed three times in PBS, before addition of IRDye® 680RD Streptavidin (LiCor) diluted in PBS (general methods section 2.1.2). The nitrocellulose membrane was washed a subsequent three times in PBS and imaged using Odyssey® Imaging System (Li-Cor).

#### 3.3.7.3. *Microscopy*

Wildtype and mutant  $\Delta gacI$  GAS cells were visualised using light microscopy by cell permeable nucleic acid-selective fluorescent dye Acridine Orange (Invitrogen™) performed in collaboration with Dr. Rob Francis at NIBSC. GAS cells were prepared from overnight cultures centrifuged at 4,000 x g for 20 minutes and washed with PBS, pH 7.4 before serial dilution. In tubes 0.5 µl Acridine Orange was added to 45 µl of serially diluted GAS cultures and mixed by pipetting. 10 µl of stained GAS cells were pipetted onto each slide with a coverslip on placed on top for imaging. Samples were protected from the light and imaged using 63x/1.2 NA H<sub>2</sub>O immersion lens on the Leica SP8X confocal laser scanning microscope (Leica Microsystems). Images presented are maximum intensity projections of confocal Z stacks.

#### 3.3.7.4. *NMR Analysis*

Extracted and purified GAC and GAS\_Rha structures were determined by Nuclear Magnetic Resonance (NMR) spectroscopy performed by Dr. Tim Rudd at NIBSC, using a Bruker 700 MHz Avance NEO spectrometer (Bruker, UK) fitted with a QCI-F cryoprobe. Samples were washed three times by lyophilisation and resuspended in 590  $\mu$ l D<sub>2</sub>O (Sigma-Aldrich™). The pulse sequence noesypr1d was used, and 512 scans were measured, with a D1 of 2 s, at 298 K.

#### 3.3.7.5. *Molecular Sizing using SEC-MALS*

High Pressure Liquid Chromatography coupled with Size Exclusion Chromatography (HPLC-SEC) and Multi Angle Light Scattering (MALS) system was used to determine molecular sizing in collaboration with Dr. Kay Lockyer (NIBSC). The system used consisted of a Dionex ICS-5000<sup>+</sup> system pump (Thermo Fisher Scientific (Dionex)), Dionex ICS-5000<sup>+</sup> Thermal Compartment (Thermo Fisher Scientific), and Dionex autosampler (Thermo Fisher Scientific) fitted with a 200  $\mu$ l injection loop. A Tosoh Bioscience TSK PWXL guard column was fitted to protect the separation of the column to trap highly adsorptive sample components and residual particulates in either the sample or mobile phase. TSKgel® G3000<sub>PWXL</sub> and G2000<sub>PWXL</sub> analytical columns were connected in series to obtain good separation of polysaccharides as detected by chromatographic signals by a multi-wavelength UV detector (Thermo Fisher Scientific) measuring 214<sub>nm</sub> and 280<sub>nm</sub> signals, in addition to an Optilab T-rEX refractometer (Wyatt Technology Corp.), a DAWN-HELEOS II 18-angle light scattering detector (Wyatt Technology Corp.) and a Visco-Star-II Viscometer (Wyatt Technology Corp.). 95% PBS pH 7.4 and 5% HPLC grade Acetonitrile (ACN) (VWR chemicals) were ultra-filtered through MF-Millipore® 0.10  $\mu$ m pore size membrane filters and used as eluent at a flow rate of 0.3 ml/minute. Generally, 50  $\mu$ g of polysaccharide at a concentration of 500  $\mu$ g/ml was injected in a volume of 100  $\mu$ l per sample in duplicate. Samples were run at a flow rate of 0.5 ml/minute.

After establishing the baseline, ensuring the column is performing well, recorded signals were aligned using BSA monomer standard (Thermo Fisher Scientific), as a baseline test mix. Data from 11 detectors between the angles of 57.0° and 141.0° were used and weight-average molecular mass (*M<sub>w</sub>*) and percentage recoveries were determined using Astra for Windows 7.3.2™ software from Wyatt Technology Corp. Signals were aligned from the main peak and a calibration constant of 4.3380 x 10<sup>-5</sup> was used. A refractive index value of 0.150 ml/g was used for

polysaccharides. The weight-average molecular mass ( $M_w$ ), polydispersity ( $M_w/M_n$ ) and hydrodynamic radius ( $R_h(v)h$ ) were determined by the Zimm method.

### 3.3.8. EDC Carbodiimide Conjugation

#### 3.3.8.1. *GAS Protein Antigen $\epsilon$ -amine Reactivity*

Free amine group availability was tested on GAS antigens through the generation of an N-acetyl lysine standard curve (0.5 – 5  $\mu\text{g/ml}$ ). Protein carriers were diluted accordingly in size exclusion buffer (50 mM Tris, 150 mM NaCl, pH 7.5) and added to 200 mM saturated sodium borate and 10 mM 2,4,6-trinitrobenzene sulfonic acid (TNBSA) solution respectively (ratio 1 : 4 : 1). Samples were incubated 20 minutes at 37 °C shaking before transfer to a 96-well plate for measurement of 420 nm absorbance read by POLARstar Omega plate reader spectrophotometer (BMG LABTECH) detecting chromogenic trinitrophenyl derivatives.

#### 3.3.8.2. *GAC Carboxylic Acid Reactivity*

Polysaccharides GAC and GAS\_Rha were mixed with 100-fold excess EZ-Link™ Amine-PEG<sub>3</sub>-Biotin (Thermo Fisher Scientific), and 10-fold excess 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) (Thermo Fisher Scientific) linking agent at room temperature for 2 hours in 100 mM MES pH 4.7 buffer. Samples were centrifuged 4,000 x  $g$  and desalted using 0.5 ml Zeba™ Spin Desalting Columns with 7,000 Da MWCO (Thermo Fisher Scientific) before analysis by Coomassie stained SDS-PAGE and immunoblot using Streptavidin-HRP conjugated antibody (BD Pharmingen) detection (general methods section 2.1.2) imaged with Sygene Pxi gel doc system.

#### 3.3.8.3. *EDC Carbodiimide Chemical Conjugation Reactions*

EDC conjugation was performed in a sequential two step reaction. First, carboxyl containing GAC / GAS\_Rha were activated with EDC and the desired O-acylisourea reactive intermediate stabilised with N-hydroxysulfosuccinimide (sulfo-NHS) (Thermo Fisher Scientific) in 100 mM MES (Sigma-Aldrich™) buffer at pH 5 – 6. Secondly, conjugation with stabilised intermediates was conducted with amine containing proteins in phosphate buffer at pH 7.2 - 7.5 to favour amide bond formation.

GAC / GAS\_Rha was lyophilised using a freeze drier (Thermo Fisher Scientific) overnight. Lyophilised samples were reconstituted in a volume of 100 mM MES pH 5 and left shaking for 30 minutes to ensure full resuspension. EDC and Sulfo-NHS were equilibrated to room temperature before 1 mg/ml and 1.5 mg/ml stocks were made respectively in 100 mM MES pH 5 and used immediately to prevent

hydrolysis and unfavourable side reactions. EDC and sulfo-NHS were added slowly to the solution to obtain a final concentration of 10-fold molar excess EDC to GAC / GAS\_Rha, and an EDC to Sulfo-NHS ratio of 2 : 5 *m/m* was used to prevent precipitation. The polysaccharide reaction with linking agents and stabilisers was left to react for 2 hours at room temperature with agitation. 2-Mercaptoethanol (2-ME) (Sigma-Aldrich™) was diluted to a 20 mM stock as a thiol containing compound to inactive unreacted EDC, prevent protein – protein conjugation and maximise active ester functionality. The reaction was mixed and incubated for 10 minutes at room temperature with agitation before the pH was increased to ~ 7.2 using 100 mM sodium bicarbonate (Sigma-Aldrich™) as measured by a small electrode (Mettler Toledo Seven Multi). Purified recombinant protein carriers dissolved in PBS pH 7.2 were then added to the activated polysaccharide solution. The reaction was left for 3 hours at room temperature followed by overnight incubation at 4 °C with agitation. After ~ 16 hours the reaction was equilibrated to room temperature before reactants (e.g. EDC and sulfo-NHS) and water soluble reaction by-product (e.g. isourea) were removed by buffer exchange using a Vivaspin® 2 Hydrosart® membrane MWCO 5,000 Da (Sartorius) centrifugal filter into PBS pH 7.2. Following buffer exchange, samples were aliquoted and those required in short term were stored at 4 °C, and long term - 20 °C.

#### *3.3.8.4. Purification of Glycoconjugates by Ultracentrifugation*

Glycoconjugates containing larger proteins such as SpyAD and TT were purified from small molecular weight free polysaccharide using Amicon® Ultracel-50 regenerated cellulose 2 ml centrifugal spin filters with a 50,000 Da MWCO (Merck Millipore). As per the manufacturer's instructions, membranes were washed with 2 ml of ultra-pure water before equilibrating with 2 ml sterile PBS pH 7.2 through spinning 4,000 x *g* for 5 minutes at each step. Buffer exchanged conjugate samples were applied to the membrane and concentrated to a volume of ~ 500 µl followed by 5 washes in sterile PBS pH 7.2. The conjugate was brought to a final volume of 1 ml and collected. Flow through, washes and retentate were all collected for downstream analysis.

#### 3.3.9. Glycoconjugate Analytical Methods

Anthrone and BCA were used to determine total polysaccharide and protein content in glycoconjugate samples (general method sections 2.5.1 – 2.5.2.).

### 3.3.9.1. Polysaccharide Quantification using High-Performance Anion-Exchange Chromatography with Pulsed Amperometric Detection (H-PAD)

Polysaccharide containing samples were diluted in ultra-pure water to fit within the prepared monosaccharide rhamnose (Sigma-Aldrich™) standard curve (0.5 – 10 µg/ml) according to their concentration calculated by anthrone assay. Diluted samples were acid hydrolysed by the addition of 4 M Trifluoroacetic acid (TFA) (Sigma-Aldrich™) and heated to 100 °C in a heat block for 3 hours to generate monosaccharide. TFA was removed by spinning samples in an Eppendorf™ Concentrator Plus speedvac fitted with a 48 x 1.5/2 mL microcentrifuge tube rotor overnight, setting V-AQ (vacuum, aqueous) at ambient temperature. The resulting pellets were resuspended in their original volume in ultra-pure water and spiked with 4 µg/ml fucose internal standard. 50 µl of each sample at an approximate concentration of 2.5 µg/ml were injected at a flow rate of 1 ml/min onto a CarboPac PA-10 (4 x 250 mm, Dionex) chromatographic column with an AminoTrap™ pre-column in collaboration with Dr. Arif Felek (NIBSC). Chromatography was performed using an Dionex ICS-5000+ system (Thermo Fisher Scientific). Samples were run on an isocratic gradient for 20 minutes followed by a cleaning and equilibration cycle with a total run time of 40 minutes. The elution buffers were prepared by filtering 400 mM NaOH (Thermo Fisher Scientific) and ultra-pure water with MF-Millipore® 0.20 µm pore size membrane filter. Elution conditions stated in Table 3.3.

**Table 3.3: H-PAD elution conditions.**

Time (Minutes)	Percentage NaOH in Buffer B
0 – 19	3.5 %
20 – 25	100 %
26 – 40	3.5 %

Similarly treated rhamnose (Sigma-Aldrich™) and GlcNAc (Sigma-Aldrich™) standards were run separately to establish peak elution times. The detection is performed by applying a potential wave form standard quad for polysaccharides with Ag/AgCl as the reference electrode. Chromatographs display signals detected by the gold working electrode, conductivity (nC) on the Y-axis and time on the X-axis. Peak area and retention times are noted after peak integration using Chromeleon™ software (Version 7, Thermo Fisher Scientific). Polysaccharide

concentration was calculated by using the monosaccharide conversion factor for rhamnose - 0.597. This enabled extracted GAC composition analysis through comparisons with monosaccharide mixtures replicating the polysaccharide repeating unit comprised of 2 x L-rhamnose and 1 x GlcNAc for wildtype GAC.

#### 3.3.9.2. *Immunoprecipitation of Unconjugated Protein Components*

500 µl of total TT-conjugate samples were mixed with 2.5 ml of anti-GAS antibodies (Thermo Fisher Scientific) diluted to 2 µg/ml in Fetal Bovine Serum (FBS) (Gibco™) and incubated at 100 rpm shaking at 37 °C for 2 hours. After incubation, 1.5 ml of 50% polyethylene glycol (PEG) 6000 (Sigma-Aldrich™) was added to the glycoconjugate sample and incubated for a further 2 hours shaking at 37 °C. The samples were then centrifuged at room temperature for 30 minutes at 13,000 rpm, and the supernatant was pipetted into a second tube. Supernatant was diluted 1 : 1 in assay diluent (AD) before being tested by ELISA (general methods section 2.4.2). MaxiSORP plates (Thermo Fisher Scientific) coated with Horse anti-TT (NIBSC 60/013) were diluted to 1.6IU /ml, samples and controls were titrated across the plate 1 : 2 with negative control wells used to monitor background, and TT in samples detected by Guinea pig anti-TT antibodies (NIBSC10/132) diluted 1 : 200 in AD buffer, followed by anti-Guinea Pig conjugated HRP (Sigma-Aldrich™) secondary antibodies (general methods section 2.1.2).

#### 3.3.9.3. *Capture ELISA*

Standard ELISA conditions were followed (general methods section 2.4.2). For TT-conjugates ELISA plates were coated overnight at 4 °C with 100 µl/well Guinea Pig anti-TT antibodies (NIBSC 10/132) diluted 1 : 200 in carbonate buffer.

Glycoconjugate samples and controls were standardised to glycoconjugate protein concentration of 100 µg/ml were then applied to plates and diluted 1 in 2 across the plate in AD buffer. Samples were incubated for 1 hour at 37° C and detected using Goat anti-GAS antibody (Thermo Fisher Scientific) diluted 1 : 2,000 in AD buffer incubated for 1 hour at 37° C followed by Rabbit anti-Goat IgG conjugated to HRP (Thermo Fisher Scientific) incubated for 1.5 hours at 37 ° C (general methods section 2.1.2).

For SpyAD glycoconjugates plates were coated with Goat anti-GAS (Thermo Fisher Scientific) diluted 1 : 1,000 in carbonate buffer incubated overnight at 4° C. Samples were standardised to 50 µg/ml polysaccharide calculated by anthrone reagent, and applied to the plate diluted 1 in 2 across the plate in AD buffer. Samples were left for



1 hour at 37° C, before washing and detection with Anti-His antibodies conjugated to HRP diluted 1 : 5,000 in AD buffer, incubated for 1.5 hours at 37 ° C.

#### 3.3.9.4. *Molecular Sizing of Glycoconjugates*

HPLC-SEC was used to determine size differences between glycoconjugate molecules and protein only molecules in collaboration with Dr. Arif Felek. Separation was accomplished using sodium phosphate buffer (66.6 mM sodium phosphate dibasic heptahydrate, 33.3 mM sodium phosphate monobasic monohydrate) pH 7.4 as eluent using a bioZen 1.8 µm SEC-2, LC 300 x 4.6 mm column (Phenomenex®) with 150 pore size (Å) attached to a Dionex UHPLC system (Thermo Fisher Scientific). Samples prepared at a concentration of 1 mg/ml were injected at a flow rate of 0.5 ml/min with isocratic elution over 25 minutes. Injected samples of 50 µl volume were detected by UV at wavelengths 214<sub>nm</sub> and 280<sub>nm</sub>.

#### 3.3.9.5. *Hydrophobicity Testing of Glycoconjugates*

Reverse Phase-High Performance Liquid Chromatography (RP-HPLC) was used to assess glycoconjugate and component hydrophobicity in collaboration with Dr. Arif Felek. A Gemini C18 3 µm 3 x 100 mm column (Phenomenex®) with a 110 pore size (Å) was used for separation and chromatography using a Dionex UHPLC system (Thermo Fisher Scientific). The mobile phase consisted of 60% HPLC grade ACN (VWR Chemicals) with 0.01% TFA (VWR Chemicals), applied to the system at a flow rate of 0.5 ml/min (Buffer B). Samples were diluted in mobile phase prior to injection (50µl/injection) and chromatography was performed with an isocratic elution of buffer B as stated in Table 3.4. Samples eluting from the column were detected by UV at wavelengths 214<sub>nm</sub> and 280<sub>nm</sub>.

**Table 3.4: RP-HPLC elution conditions.**

Time (Minutes)	Percentage ACN in Buffer B
0 – 15	60 %
16 – 18	100 % (cleaning)
19 – 35	60 % (equilibrate)

### 3.3.10. Immunogenicity Analysis

#### 3.3.10.1. *Animals and Immunisation Schedule*

Animal studies were conducted according to the UK Home Office regulations and were approved by the local ethics committee. For investigation of the immunogenicity of GAS protein antigens, 5 mice per group of female 6-8-week-old BALB/c mice (Charles River Laboratories, Saffron Walden, UK) were immunised

subcutaneously. Each mouse received 20 µg of purified recombinant protein antigen adsorbed onto Alum adjuvant (Alhydrogel® 2 %, InvivoGen) at 100 µg/dose. Antigens were adsorbed to Alum overnight at 4 °C with rotation. Each mouse received 2 doses at 14-day intervals, with serum collected prior to immunisation (pre-immune - day 0), 13 days after the initial immunisation (test bleed), followed by terminal bleeds collected two weeks after the second immunisation (day 24). Note that for the second immunisation mice in the group receiving SpyAD protein antigen received protein alone without Alum adjuvant. Serum for use in immunoassays was isolated from red blood cells pelleted following a spin at 4,000 x g for 10 minutes at room temperature. Pre-immune sera collected from each group were pooled together.

### 3.3.10.2. *ELISA to Measure IgG Antibody Titres Against Vaccine*

#### *Components*

ELISA (general methods section 2.4.2) was used to quantify IgG antibody titres in mouse sera or pooled IVIG (Ig Vena, Kedrion Biopharma) (Batch Numbers: 157002, 157705, and 157713). MaxiSORP 96-well plates (Thermo Fisher Scientific) were coated with 50 µl/well of 0.5 µg/mL individual recombinant proteins or purified GAC in carbonate buffer overnight at 4 °C. Serum and IVIG dilutions were prepared in AD buffer; pooled pre-immune and test bleeds diluted 1 : 50, terminal bleeds diluted 1 : 100 (SpyCEP and SpyAD) or 1 : 30 (MalE) and IVIG diluted 1:100. Samples were diluted 3- or 5- fold across the plate and incubated at 37 °C for 2 hours before incubation with 100 µl/well anti-mouse IgG-HRP whole molecule (Sigma-Aldrich™), or anti-human IgG-HRP whole molecule (Sigma-Aldrich™) diluted 1 : 20,000 in AD at room temperature for 1 hour at 37 °C.

### 3.3.10.3. *GAS Cell Compartment Separation*

GAS overnight cultures were sub-cultured to OD<sub>600nm</sub> 0.05 in THY media and grown to mid-log phase (OD<sub>600nm</sub> ~ 0.7). GAS cells were pelleted at 4000 x g and resuspended in cell wall extraction buffer (10 mM Tris-HCl, pH 8.0, 30% (w/v) raffinose (Sigma-Aldrich™), 100-U/mL mutanolysin, 1-mg/mL lysozyme and 20-U/mL recombinant PlyC with 1 cComplete™ EDTA-free Protease Inhibitor Cocktail (Roche)) to OD<sub>600nm</sub> 20. Suspensions were incubated for 3 hours at 37 °C with agitation before pelleting cells at 6,000 x g for 20 minutes at room temperature. The supernatant contained the cleaved cell wall, whilst the pellet was resuspended in 10 mM Tris-HCl, pH 7.4 supplemented with 10 µg/ml DNase, and underwent three freeze thaw cycles between 37 °C and - 80 °C for 5 minutes per step to obtain cytoplasm fractions and insoluble fractions after a 4,000 x g spin. Culture

supernatants were filtered through a 0.2 µm filter and concentrated by precipitating proteins with 10% TCA overnight on ice. Proteins were pelleted by a spin at 17,000  $\times g$  for 20 minutes and washed three times with 90% acetone (Sigma-Aldrich™) buffer before drying in the speedyVac (Eppendorf™) set at 60 °C under a vacuum. The resulting pellets were resuspended in 10 mM Tris-HCl, pH 8.0 to OD<sub>600nm</sub> 20.

#### 3.3.10.4. *IL-8 Cleavage Assay*

To assess anti-SpyCEP antibody functionality an IL-8 cleavage assay was performed using an IL-8 DuoSET ELISA kit (RnD, DY208-05). GAS cultures were grown to late mid-log (OD<sub>600nm</sub> 0.7) in THY broth, and cells were pelleted by a 4,000  $\times g$  spin for 20 minutes at room temperature. GAS supernatants were 0.2 µm filter sterilised, diluted 1 : 40 in PBS buffer, pH 7.4 and pre-incubated with mouse sera dilutions (1 : 20) at 37 °C for 1 hour shaking in a total volume of 240 µl before addition of 125 pg/ml IL-8 (RnD, 208-IL-010/CF). Samples were incubated at 37 °C for 16 hours with agitation, and maxiSORP 96-well plates (Thermo Fisher Scientific) coated with 100 µl/well of 4 µg/ml mouse anti-human IL-8 capture antibody incubated overnight at 4 °C. 100 µl/well of samples were then added to blocked plates, incubated at room temperature for 2 hours. Human IL-8 detection antibody was added 100 µl/well incubated at room temperature for a further 2 hours, before detected by 100 µl/well of streptavidin conjugated HRP incubated for 20 minutes at room temperature (general methods section 2.4.2).

## 3.4. Results

### 3.4.1. Design and Analysis of GAS Specific Carrier Proteins

Two important steps within GAS infection and disease onset are initial colonisation followed by invasive disease. Therefore, to explore both aspects of the host-pathogen interaction, three GAS antigens were selected from a list of other GAS proteins (Appendix section 7.1), for involvement in colonisation (MalE and SpyAD) and invasive disease onset (SpyCEP). SpyCEP and SpyAD have been previously investigated as vaccine candidates in pre-clinical studies, while MalE is a novel protein antigen, yet to be included in a tested vaccination preparation. Protein antigen properties are outlined in Table 3.5.

The amino acid sequences of the three GAS antigens were analysed using VaxiJen predictive software based on pre-processing of amino acids<sup>502</sup>. All three proteins have an antigenic score above the probable antigen threshold (> 0.5) whereas interestingly TT showed the lowest antigenic score below the threshold, predicted to not be an antigen based on *in silico* data (Table 3.6). The Immune Epitope Database and Analysis Resource (IEDB) was also used to assess the candidates for predicted protective cell epitopes. Table 3.6 shows the *in silico* tool revealed predicted antigenic B and T cell epitopes for SpyAD, SpyCEP, and TT but did not detect any B/T cell epitopes for MalE contrary to the literature.

**Table 3.5: GAS protein antigens selected for use as carriers for chemical conjugation.**

Protein common names and Uniprot identifiers are shown, with reference strain SF370 Gene ID in brackets. Specific pathogenic function in the host, and current vaccine inclusion with associated protective mechanism (P = Protective, O = Opsonic, N = Neutralising) are also stated.

Protein Name	Uniprot Identity	Pathogenesis Function	Vaccine Inclusion	Protective mechanism	References
<b>SpyCEP</b> (ScpC) (Spy0416)	Q3HV58	<b>Immune evasion</b> – Calcium-dependent cleavage of human IL-8 chemokine reducing neutrophil recruitment <b>Invasion</b> – High expression associated with GAS dissemination	GSK Combo5 GSK Conjugate	P, N	208,395,411,503–505
<b>MaIE</b> (Spy1058) (Spy1294)	J7MAC3	<b>Survival</b> – Binding of maltodextrin as a carbon source in saliva <b>Colonisation</b> – binding to the oropharynx	-	-	124,365,506,507
<b>SpyAD</b> (PrgA) (Spy0269) (Spy0299)	Q9A1H3	<b>Colonisation</b> – Adhesion to pharyngeal and epithelial ECM components in initial weak binding events <b>Survival</b> – Protein localised in divisome important in cell division	GSK Combo Two Distinct Glycoconjugates	P	365,395,411,412,508,509

**Table 3.6: Predicted immunological properties of selected GAS antigens.**

Protein carrier antigenic scores and number of predictive protective cell epitopes are shown based VaxiJen and IEDB software. Specific protective regions and predicted cross protective capability are also stated based on a literature search.

Protein Name (PDB Identifier) NCTC-8198 GAS	Antigenic score <sup>A</sup>	Predicted protective cell epitopes ( <i>in silico</i> ) <sup>B</sup>	Predicted protective epitopes (literature)	Cross protective capability	References
<b>SpyCEP*</b> (ScpC) (Spy0416)	0.592	29: 29 B cell 0 T cell	NTD Neutralising protection from lethal GAS challenge	Conserved	150,203,208,395,407,510,511
<b>MaIE</b> (Spy1294) (Spy1058)	0.527	0	Unknown epitopes Antibodies present in humans and mouse convalescent sera	Conserved	156,506
<b>SpyAD</b> (PrgA) (Spy0269)	0.592	8: 7 B cell 1 T cell	Mainly NTD, some CTD regions Protection from intranasal GAS challenge	Conserved	150,395,407,508,512
<b>TT</b> (1A8D) <i>C. tetani</i> (strain E88)	0.434	16: 0 B cell 16 T cell	C fragment of heavy chain - TT04 and TT10 linear epitope Protection from lethal challenge with TT in mice (meningococcal and pneumococcal TT containing glycoconjugates)	Not applicable from GAS vaccine perspective	513–516

\*Denotes a cloned fragment of full-sized protein.

<sup>A</sup> Threshold >0.5 = probable antigen.

<sup>B</sup> Experimentally tested antigen or epitope with 70% BLAST match to specific GAS antigen.



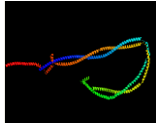

The candidate antigens were analysed with structural *in silico* tools using cloned amino acid sequences. Signal peptides and C-terminal cell anchoring domains were removed ensuring *E. coli* cytoplasmic expression. MalE and SpyAD were cloned as full-length proteins, whereas SpyCEP was cloned as a protective fragment<sup>203</sup>. This was decided due to vaccine safety concerns associated with functional enzymatic activity from the full-sized SpyCEP protein<sup>21</sup> (Figure 3.3a). SpyCEP was truncated to recombinantly express the N-terminal part of the protein containing 2 of the 3 active sites<sup>208</sup>. This was selected due to observations that the fragment can induce antibodies capable of neutralising native SpyCEP's IL-8 cleaving activity<sup>208</sup>. Cloned nucleic acid and amino acid sequences for all proteins/fragments are stated in Appendix section 7.1.1.

Proteins were experimentally shown in the literature to be between 40 - 100 kDa in size, and surface localised, either through identification of a C-terminal peptidoglycan cell wall anchoring (LPXTG) domain for SpyCEP<sup>517</sup>, a putative C-terminal hydrophobic transmembrane domain followed by a charged tail for SpyAD<sup>508</sup>, or an N-terminal lipid-modified membrane anchoring domain (MKSWQKVIVGGASLTLASTLLVGC) for MalE as a lipoprotein<sup>156</sup>. *In silico* tool Phyre2 was used to predict protein structure based on the amino acid sequence used for recombinant protein production in *E. coli*. All three antigens are multidomain proteins with coiled coil regions, most notably for SpyAD protein (Table 3.7). In comparison, TT multidomain protein shows little alpha helical structure, with a high level of beta strands and some disordered regions.

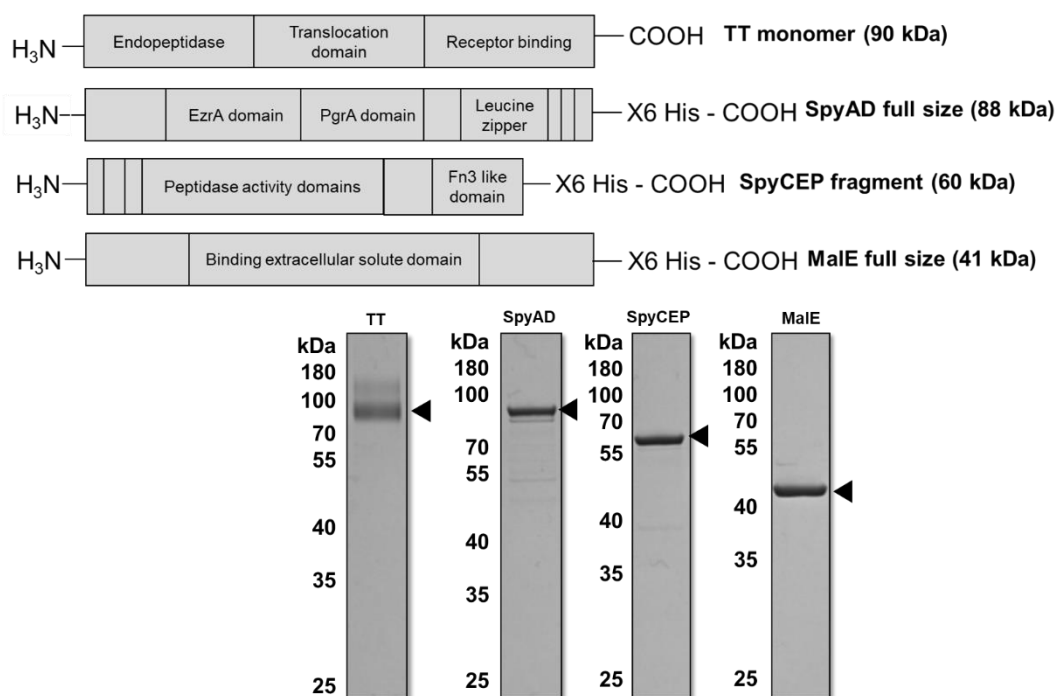
**Table 3.7: Predicted physicochemical properties of selected GAS antigens.**

Molecular weight, protein structure, modelled by Phyre<sup>2</sup> and cell surface localisation (LPXTG motif or C terminal domain) are highlighted. Properties correct for full size proteins for MaIE, SpyAD, and TT, and a cloned fragment for SpyCEP. Phyre2 web portal used for protein modelling, prediction, and analysis<sup>498</sup>.

N/A = Not Applicable.

Protein Name	Molecular weight (kDa)	Protein localisation	Protein structure	Phyre <sup>2</sup> structural prediction software
<b>SpyCEP</b>	60.1 (cloned fragment) 170 (full size)	<b>Cell wall</b> – LPXTG anchoring domain <sup>517</sup> <b>Extracellular Secretion</b> <sup>518</sup>	Multidomain modular – nine structural domains and transmembrane domain region <sup>519,520</sup>	 <p>Alpha helix (16%) Beta strand (21%) TM helix (3%) Disordered (37%) Confidence: 100% Coverage:71%</p>
<b>MaIE</b>	41.1 (full size)	<b>Membrane</b> <sup>156,521</sup>	Likely multidomain modular <sup>507</sup>	 <p>Alpha helix (43%) Beta strand (16%) Disordered (8%) Confidence: 100% Coverage:96%</p>
<b>SpyAD</b>	88.3 (full size)	<b>Membrane and Cytoplasmic</b> – Moonlighting protein <sup>508</sup>	Likely multidomain modular containing coiled coil CTD region <sup>508</sup>	 <p>Alpha helix (85%) Beta strand (0%) Disordered (24%) Confidence: 96.9% Coverage: 86%</p>
<b>TT</b>	180 (dimer) 90 (monomer)	N/A	Multidomain modular <sup>522</sup>	 <p>Alpha helix (2%) Beta strand (49%) Disordered (23%) Confidence:100% Coverage:100%</p>





**Figure 3.3: Classical protein carrier Tetanus Toxoid (TT) and GAS protein carriers used for chemical conjugation reactions.**

(A) Schematic representation of cloned domains of GAS proteins (SpyAD and MalE are full size and SpyCEP is a fragment) and monomeric TT.

(B) Coomassie-stained SDS-PAGE of 1 µg of recombinantly expressed and purified GAS antigens and monomeric TT.

#### 3.4.1.1. Recombinant Expression and Purification of GAS Antigens in *E. coli*

GAS antigens were cloned for expression and purification from BL21 (DE3) *E. coli*. SpyCEP, MalE, and SpyAD were successfully cloned into the pET28a expression vector to contain C-terminal His-tags and were shown to express well in BL21 (DE3) *E. coli* soluble fraction. Proteins were purified by nickel column and size exclusion chromatography to a high purity. MalE purified to the highest yield at 30 – 40 mg, followed by SpyCEP at 10 – 14 mg and SpyAD around 0.5 – 1.5 mg per 50 ml culture cell pellets grown in autoinduction media, as quantified by BCA assay (data not shown).

Monomeric TT was purified using size exclusion chromatography from the internal WHO Standard 2nd International Standard for Tetanus Toxoid for use in Flocculation Test (NIBSC code: 04/150). The purity of purified recombinant antigens and TT was assessed by SDS-PAGE stained by Coomassie (Figure 3.3b). All three recombinant proteins and monomeric TT showed minimal contaminants following

purification, confirmed further through tandem mass spectrometry analysis stating purity as well as amino acid sequence. The amino acid identity was found to be 100 % for MalE, SpyCEP, and SpyAD based on the amino acid sequences used for gBlocks® Gene Fragments synthesis as a reference (Appendix section 7.1.1.). Purified TT was shown to have 98 % amino acid identity to Uniprot identifier TT, and sequences obtained from UniProt.P04958. The purity of samples for use in chemical conjugations were found to be between 97.97 and 99.82 % for TT, SpyCEP and MalE, and 83.54 % for SpyAD. To reduce toxicity for future immunogenicity studies, protein preparations were further purified to remove *E. coli* endotoxin and analysed by LAL to quantify remaining endotoxin. Final endotoxin concentrations were deemed a suitable level for subsequent mouse immunisations; SpyCEP has 302.4 endotoxin units EU/mg, MalE 288.6 EU/mg and SpyAD 27.3 EU/mg, correlating to less than 6 EU per 100 µl immunisation dose, lower than the limit of 10 EU for a 20 g mouse<sup>523</sup>.

### 3.4.2. Proposed Vaccine Components Are Immunogenic

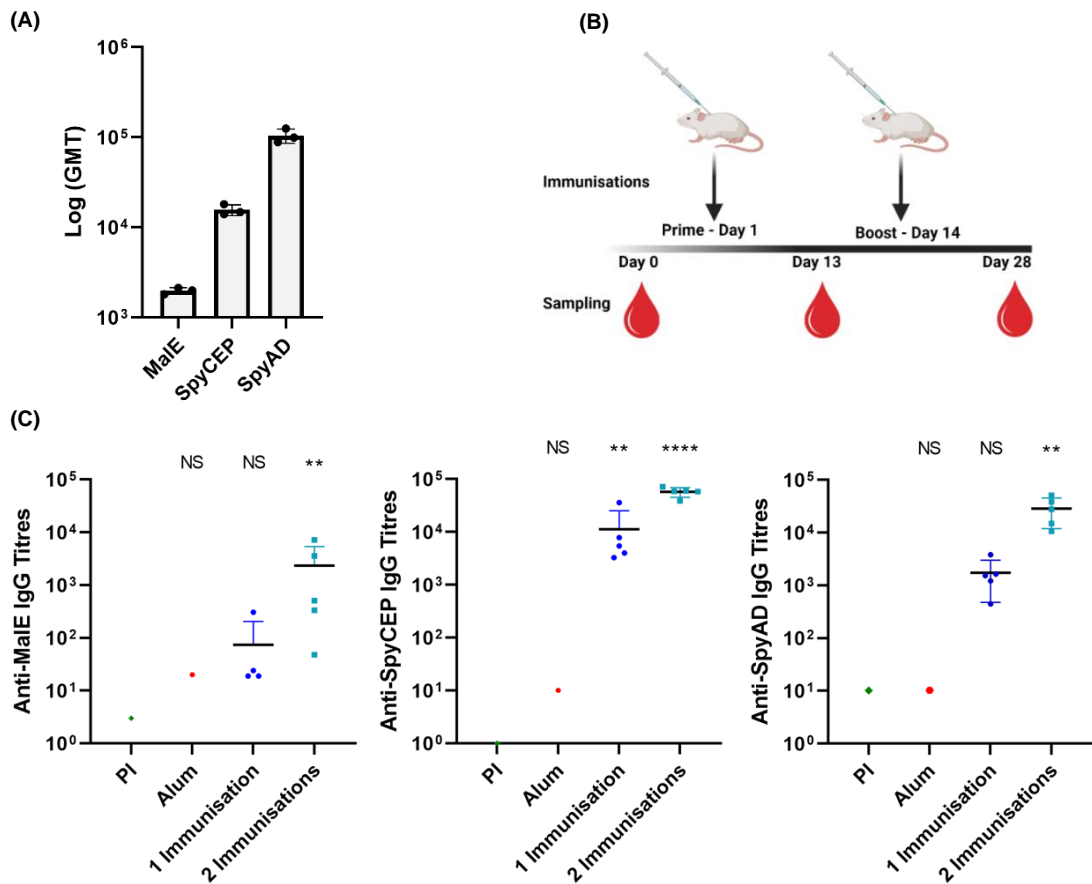
#### 3.4.2.1. *Naturally Occurring Antibodies Recognise GAS Proteins*

Following *in silico* analysis, the antigenicity of three GAS antigens was tested with three separate batches of pooled human IgG (IVIG) by ELISA to confirm their suitability as conjugate carrier proteins. Antibodies against all three protein antigens were found in pooled IVIG batches, confirming an immunogenic response to these proteins following natural human GAS infections (Table 3.8). High geometric mean titres (GMT) were observed for antibodies recognising recombinant SpyCEP and SpyAD proteins in IVIG, 15,611 and 104,354, respectively, with lower titres observed for MalE with GMT 1,973 (Figure 3.4a). Nevertheless, although considerably lower than the other two antigens, the data shows the presence of MalE specific antibodies.

#### 3.4.2.2. *Immunogenicity of GAS Proteins*

The immunogenicity of the GAS proteins was assessed following subcutaneous immunisation of mice in the presence of Alum adjuvant. The prime-boost immunisation schedule consisted of two immunisations 2 weeks apart (Figure 3.4b). The IgG response to immunisation with each GAS antigen was assessed and compared to a group of mice which received PBS and Alum adjuvant only, and pre-bleeds before any immunisation, termed pre-immune (PI) for each protein antigen. IgG titres were quantified by ELISA with the data presented as the GMT for each group (Table 3.8). Following the first immunisation, relatively small anti-protein responses were observed for MalE (GMT: 74, range: 1 - 307) and SpyAD (GMT:

1,732, range: 446 - 3,817), compared to SpyCEP (GMT: 11,286, range: 4,017 - 35,919). For all three GAS antigens there was a significant boosting effect with the second immunisation improving antibody titres (SpyCEP  $p = < 0.0001$ , MaIE  $p = 0.0024$ , and SpyAD  $p = 0.002$ ). Although MaIE showed a modest IgG titre with the greatest mouse to mouse variation observed, the boosting effect appeared to be most noteworthy with 31.5 times more antibodies present after the second immunisation compared to the first which was not significantly above the pre-immune level. Of all three antigens, SpyCEP showed the highest immunogenicity with a GMT of 57,325 after the second immunisation, almost double that of SpyAD with a GMT of 28,517, and 25 times greater than MaIE with a GMT of 2,333 after two immunisations (Figure 3.4c).



**Figure 3.4: Antibody response to GAS protein antigens in pooled human IgG (IVIG) and following subcutaneous immunisation of BALB/c mice.**

**(A)** OD 0.5 Geometric Mean Antibody Titre from titration of IVIG against recombinant proteins MalE, SpyCEP and SpyAD as coating antigens. Detection of human IgG with 1 : 20,000 Rabbit anti-human HRP (whole molecule). Three batches of IVIG were tested.

**(B)** Schematic representation of immunisation schedule with mice receiving two doses (days 1 and 14) of 20  $\mu$ g of individual recombinantly expressed GAS antigens MalE, SpyCEP or SpyAD or PBS control adsorbed onto Alum adjuvant (n = 5 per group). Day 0 = pre-immune sera, Day 13 = test sera (1 immunisation), Day 28 = terminal sera (2 immunisations). Original figure created with BioRender.com.

**(C)** OD 0.5 Geometric Mean Antibody Titre for each individual mouse within each group receiving individual recombinant proteins MalE, SpyCEP or SpyAD after the first and second immunisations. Titres are compared to pooled pre-immune (PI) sera (green circle), and pooled sera from PBS and adjuvant group (Alum) as a negative control (red circle).

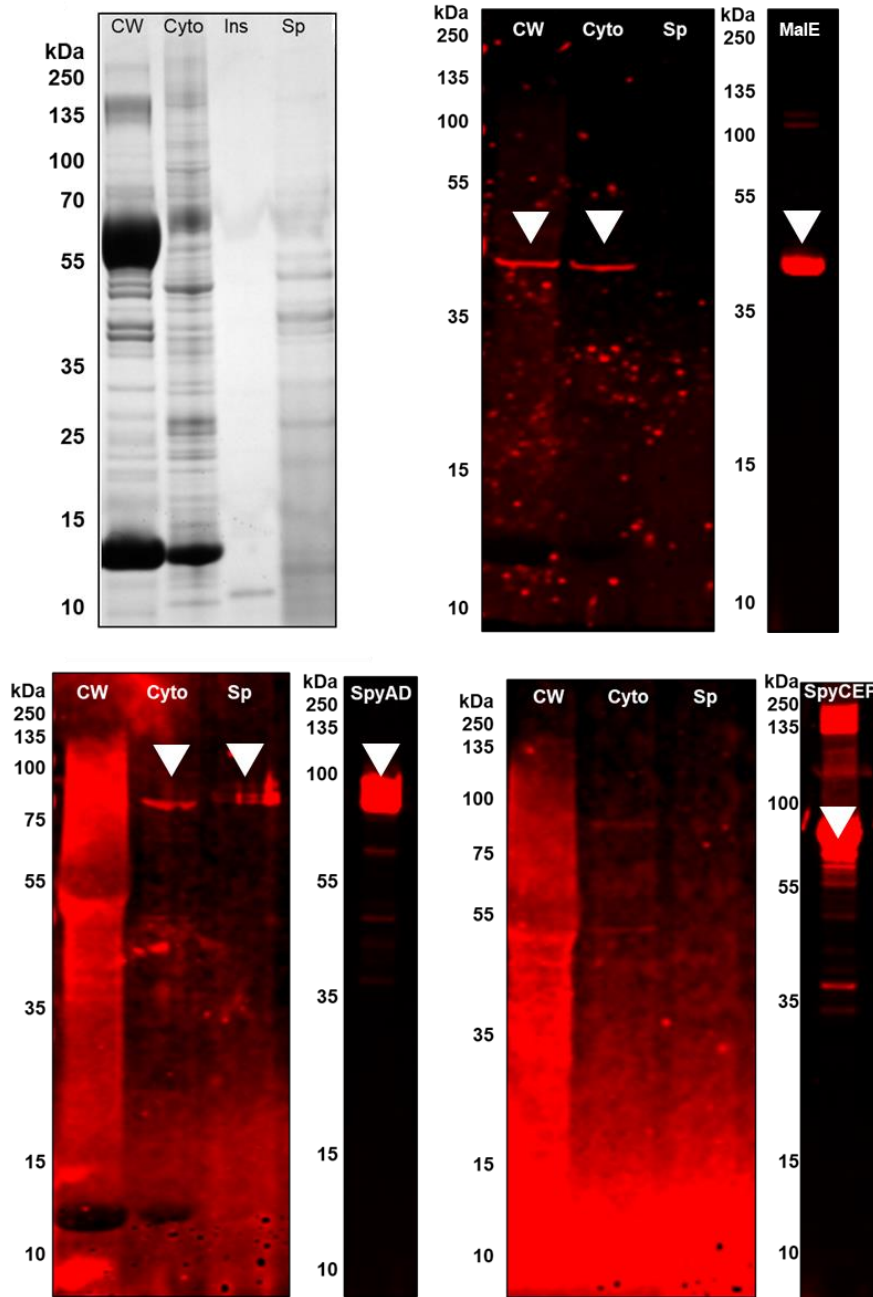
Significance compared to pre-immune control group determined by Kruskal-Wallis test, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , and \*\*\*\*  $p < 0.0001$ .

**Table 3.8: Antibody responses to GAS protein antigens in pooled human IgG (IVIG) and IgG titres following subcutaneous immunisation of BALB/c mice.**

Two immunisations of 20 µg of recombinantly expressed GAS antigens MalE, SpyCEP and SpyAD adsorbed onto Alum adjuvant (n = 5 per group). Titres are expressed as the serum dilution giving an OD<sub>450 nm</sub> 0.5 using the interpolation of a sigmoidal curve in GraphPad Prism.

Group	Geometric Mean Titre (GMT) of IgG Response against Individual GAS Antigens		
	MalE	SpyCEP	SpyAD
IVIG	1,973 (1,809 - 2,118)	15,611 (14,015 - 18,073)	104,354 (88,582 - 125,066)
Pre-immune (Day 0)	3	1	10
After First Immunisation (Day 13)	74 (1 - 307)	11,286 (4,017 - 35,919)	1,732 (446 - 3,817)
After Second Immunisation (Day 28)	2,333 (48 - 7,192)	57,325 (38,686 - 71,545)	28,517 (14,823 - 50,738)

To confirm surface exposure of the native proteins, mouse immune sera generated against recombinant proteins were used to demonstrate binding to GAS cell compartments by western blot. Cell wall extracts (CW), cytoplasm (Cyto), and culture supernatants (Sp) were obtained from late exponential-phase M1 GAS cultures and run on SDS-PAGE before staining with Coomassie (Figure 3.5a) or transferred to nitrocellulose membrane and probed by Western blotting with antigen-specific immune sera. Two of the three antigens could be detected with the immune sera at their predicted molecular weights. MalE was detected in the cell wall and cytoplasmic fractions (Figure 3.5b) and SpyAD was present in the cytoplasm and supernatant (Figure 3.5c). The recombinant SpyCEP fragment used as the immunogen could be detected, however limited banding, indicative of the full sized protein (170 kDa), was observed when blotting immune sera against GAS cell extracts (Figure 3.5d). MalE and SpyAD recombinant protein antigens used as immunogens were also recognised by mouse immune sera.

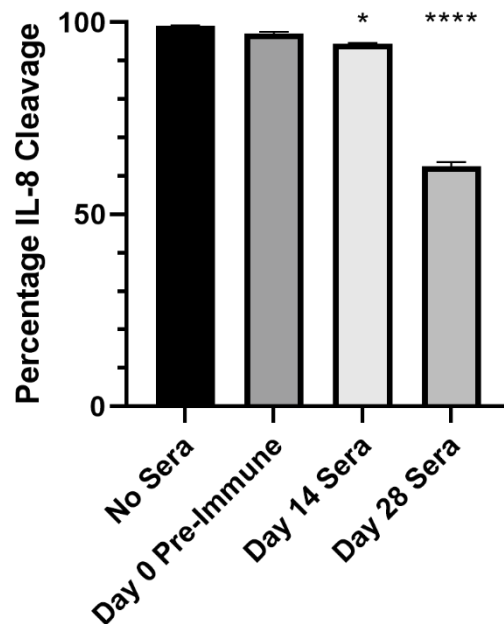


**Figure 3.5: Surface exposure and recognition of GAS antigens by protein specific mouse antisera.**

**(A)** Concentrated cell wall (CW) extracts, cytoplasm (Cyto), insoluble fraction (INS) and culture supernatants (Sp) from M1 NCTC-8198 GAS cells were separated by SDS-PAGE and stained with Coomassie blue.

Separated cell compartments and purified recombinant protein positive control preparations were transferred onto nitrocellulose membranes and incubated with antigen-specific mouse sera from day 28 (terminal). Mouse sera dilutions; 1 : 500 Anti MalE **(B)**, 1 : 750 Anti SpyAD **(C)**, and 1 : 1,000 Anti SpyCEP **(D)**. Detected with Goat anti-Mouse IgG IRDye® 680RD. White arrows indicate recognised native antigen or recombinant protein.

SpyCEP can cleave the IL-8 cytokine as an immune evasion strategy during natural GAS infection. Therefore, the neutralising activity of anti-SpyCEP antibodies was evaluated in an IL-8 cleavage assay using late-log GAS supernatants with and without immune sera. GAS supernatants were incubated with IL-8 alone or IL-8 in the presence of pooled immune sera from SpyCEP immunisation at day 14 or day 28, or pre-immune serum as a negative control. Pre-incubation of GAS supernatant with immune sera following a single immunisation showed a modest yet statistically significant reduction in IL-8 cleavage compared with nonimmune sera (6 % reduction,  $p = 0.0299$ ) equating to 94 % cleavage (Figure 3.6). The SpyCEP specific neutralising capability of sera was drastically improved following a second immunisation, with sera at day 28 resulting in 37 % reduction ( $p = < 0.0001$ ) compared with nonimmune sera, equating to 63 % cleavage (Figure 3.6). This demonstrates that a second immunisation is required to generate a significant difference in functional immunity, specifically the generation of neutralising antibodies specific to SpyCEP.



**Figure 3.6: Neutralising capability of anti-SpyCEP immune sera for IL-8 cleavage activity during incubation of GAS culture supernatants with recombinant IL-8 cytokine.**

Assessment of IL-8 cytokine degradation by SpyCEP in GAS culture supernatants with and without sera (no sera, day 0 pre-immune sera, or immune sera day 14 and day 28). Mouse sera and GAS supernatants were incubated 1:1 prior to addition of IL-8. Following a further incubation overnight IL-8 was detected by ELISA and percentage cleavage calculated as compared with day 0 pre-immune sera.

NS, not significant, \*  $p = 0.0299$ , and \*\*\*\*  $p < 0.0001$  (one-way ANOVA).

### 3.4.3. Extraction and Characterisation of Polysaccharide from GAS cells for Use in Chemical Conjugation Reactions

#### *3.4.3.1. Naturally Occurring Antibodies Recognise GAC*

In addition to protein antigens, purified wildtype (WT) GAC antigenicity was also tested with the same three batches of pooled human IgG (IVIG) using ELISA and purified GAC coated plates. Geometric mean antibody titre (GMT), expressed as the IVIG dilution giving an OD<sub>450nm</sub> 0.5 using the interpolation of a sigmoidal curve in GraphPad Prism, was determined to be 77,289 (range: 50,006 - 127,907). This confirms that the GAC extraction process preserves epitopes that are recognised by antibodies generated from natural human infections.

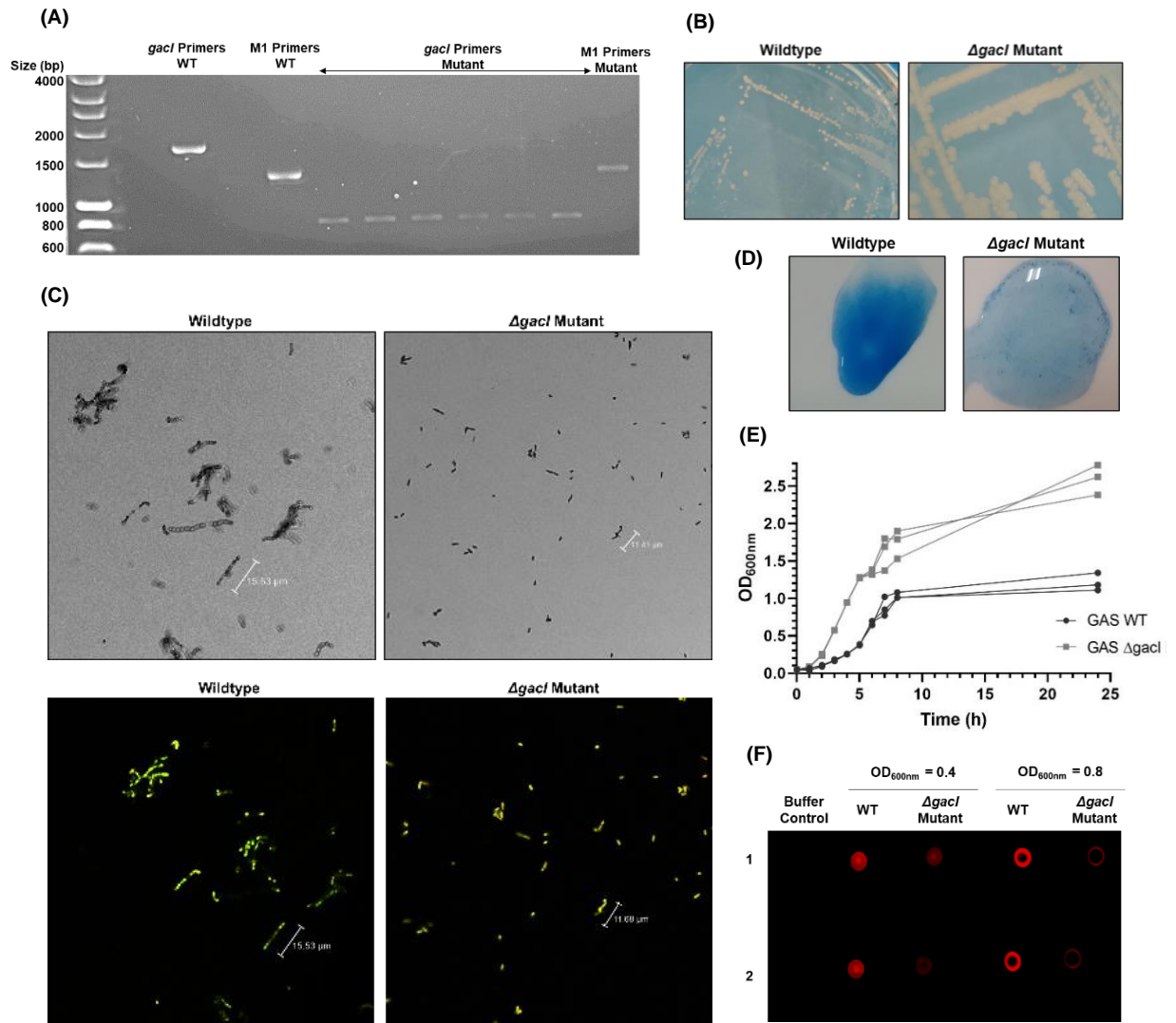
#### *3.4.3.2. Generation and Characterisation of $\Delta gacI$ Mutant*

To circumvent any potential autoimmune reactivity of a WT GAC glycoconjugate vaccine, a mutant strain expressing only the polyrhamnose backbone was generated. Towards this aim, a non-replicating kanamycin resistant derivative of pUC19 plasmid compatible for GAS transformation via electroporation was generated by replacing pUC19 ampicillin cassette with pET28a kanamycin cassette by restriction digest cloning (general methods section 2.2.5). The construct was cloned to contain 750 bp homology arms (totalling 1.5 kb) flanking the *gacI* coding region targeting this gene for deletion. After initial kanamycin selection for plasmid integration, and subsequent passage on non-selective plates encouraging plasmid loss, a successful markerless chromosomal deletion of *gacI* within the GAC 12-gene operon was achieved as confirmed by PCR amplification using primers flanking *gacI* gene (Figure 3.7a).

After targeted gene deletion, the  $\Delta gacI$  GAS mutant strain showed more mucoidal colony morphology on THY agar compared to the parent wildtype M1T1 strain suggesting that the gene deletion could be affecting the cell surface (Figure 3.7b). Acridine orange staining of the cells was carried out to investigate the mutant cell morphology and chain formation using microscopy. Images show that the mutant cells appear to be less circular in shape, and form shorter chains compared to the wildtype (Figure 3.7c). This was an interesting and unexpected observation, therefore, to confirm the mutant's identity as GAS, the strain was subsequently tested by the diagnostic latex agglutination test (Strep Test) and grown on blood agar to confirm retention of the characteristic  $\beta$ -haemolytic phenotype. The latex agglutination test, which uses the surface carbohydrate as part of the identification process, showed a partial loss in reactivity for  $\Delta gacI$  GAS mutant cells compared to



wildtype parental strain to GAC specific beads which showed strong agglutination (Figure 3.7d). Despite such a distinctive change in phenotype the *gacI* gene deletion appeared to not negatively impact  $\Delta gacI$  GAS mutant grown in THY broth compared to the parent strain, and rather appears to positively impact growth over the tested 24-hour period, with final measured optical density being almost double for the mutant compared to the wildtype (Figure 3.7e). To confirm that deletion of the *gacI* gene does indeed prevent GlcNAc addition to the rhamnose polymer backbone, both wildtype and  $\Delta gacI$  GAS mutant whole cells were dotted onto nitrocellulose membranes and stained with biotinylated Succinylated Wheat Germ Agglutinin (sWGA) lectin which specifically recognises GlcNAc. Compared to the parent wildtype strain, the  $\Delta gacI$  GAS cells showed reduced GlcNAc staining when detected using streptavidin-IRDye 800CW conjugate at both the early and late log phases of growth (Figure 3.7f).



**Figure 3.7: Generation and phenotypic characterisation of *gacI* gene deletion in GAS M1 NCTC-8198.**

**(A)** PCR amplification with primers flanking *gacI* show a 1.8 kb product with WT GAS and a 780 bp product upon deletion of the *gacI* gene in screened mutants (C8 – C13). Primers flanking M1 protein were also included to demonstrate GAS origin.

**(B)** Colony morphology of  $\Delta gacI$  mutant and wildtype GAS M1 NCTC-8198 grown on THY agar.

**(C)** Confocal microscopy images of acetone fixed wildtype and  $\Delta gacI$  mutant GAS bacteria stained with acridine orange.

**(D)** Latex agglutination reaction with GAC-specific beads as indicated on GAS WT and  $\Delta gacI$  mutant.

**(E)** Growth of  $\Delta gacI$  mutant and wildtype M1 GAS in THY broth over a 24-hour time period.

**(F)** Dot blot of two separate colonies of  $\Delta gacI$  and WT GAS M1 NCTC-8198 whole cells stained with GlcNAc-specific sWGA lectin conjugated to biotin at OD = 0.4 and 0.8.

#### 3.4.3.3. *Extraction of GAC Variants from GAS Cells*

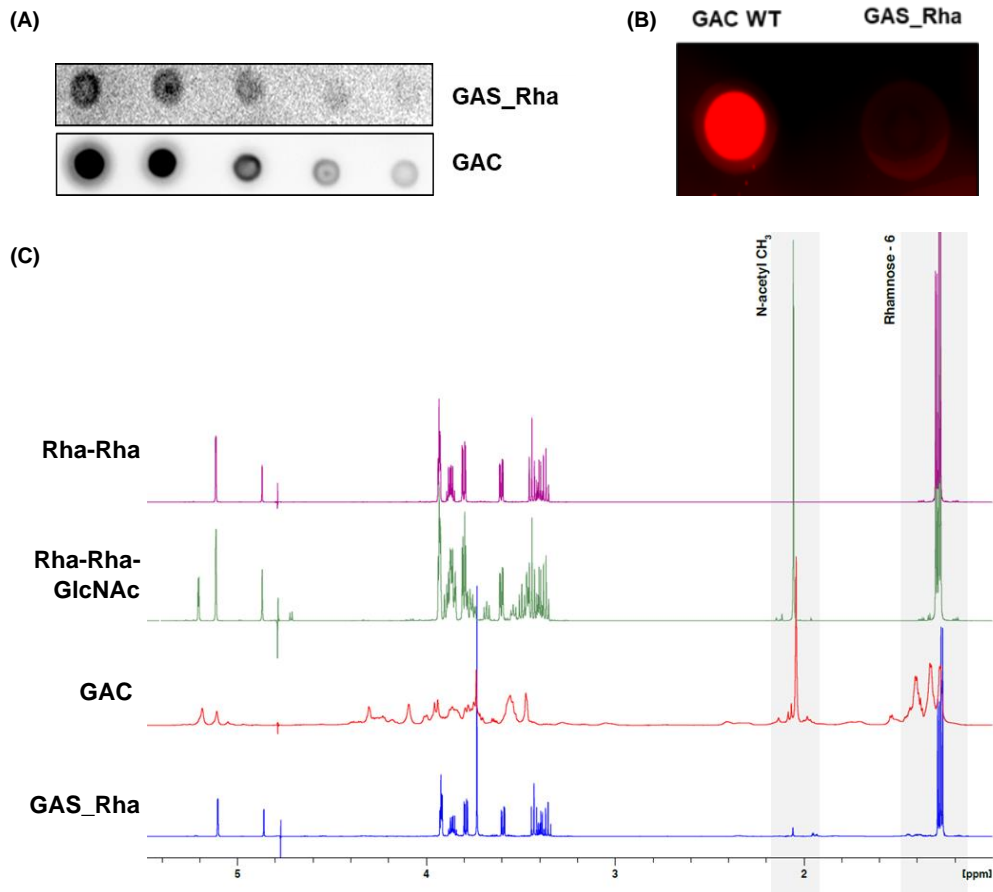
Wildtype GAC and mutant polysaccharide (GAS\_Rha) were enzymatically extracted from GAS bacterial cultures as described in the materials and methods (section 3.3.6). Extracted GAC or GAS\_Rha material was buffer exchanged and purified using HiTrap Capto Adhere ion exchange column, binding and eluting GAC / GAS\_Rha whilst removing PlyC and other contaminating protein species. Two commercially available polyclonal antibodies raised against GAS (Thermo Fisher Scientific or AbCam) (general methods section 2.1.2) were used to indicate the presence of GAC or GAS\_Rha within starting materials, check for material lost in buffer exchange and flow through samples, as well as purified chromatography elution fractions. Both GAC and GAS\_Rha can be detected by the polyclonal antiserum (Figure 3.8a). SDS-PAGE showed the successful removal of contaminating PlyC among other protein species from the purified GAC / GAS\_Rha elution fractions (Appendix Figure 7.2).

Extracted and purified GAC and GAS\_Rha preparations were quantified using Anthrone colourimetric assay. Routinely from large scale 2 litre cultures, yields were ~ 10-fold lower for GAS\_Rha compared to WT GAC, despite similar bacterial growth (data not shown). This may have been attributable to PlyC sensitivity, with  $\Delta gacI$  mutant cells appearing to be less sensitive to PlyC lysis compared to wildtype cells based on optical density after co-incubation (Appendix Table 7.1). Purified WT GAC and GAS\_Rha were standardised to polysaccharide concentration of 100  $\mu\text{g/ml}$  based on anthrone analysis and analysed by sWGA-specific lectin staining. WT GAC showed a strong sWGA signal, whereas GAS\_Rha as expected did not show any signal (Figure 3.8b).

#### 3.4.3.4. *GAC / GAS\_Rha Physicochemical Characterisation*

The characteristics of the wildtype GAC and mutant GAS\_Rha polysaccharide structures were confirmed by NMR and Size Exclusion Chromatography with Multi Angled Light Scattering (SEC-MALS). Whole cell analysis by dot blot indicated some GlcNAc signal was still present within  $\Delta gacI$  GAS mutant cells, therefore, to confirm loss of GlcNAc from the GAC polymer chains, extracted GAS\_Rha was analysed by NMR and compared to wildtype GAC. The proton NMR analysis confirmed spectra characteristic of polysaccharides, and the extraction and purification methods were deemed successful for removal of contaminating products in GAC preparations. Wildtype GAC and GAS Rha were compared to their expected repeating units Rha-Rha-GlcNAc and Rha-Rha respectively using commercially purchased monosaccharide prepared as mixtures. GAC spectra showed chemical

shifts representing rhamnose (1.3 ppm) and N-acetyl CH<sub>3</sub> species (2.1 ppm) such as GlcNAc which are also observed in the repeating unit (Figure 3.8c red and green). GAS\_Rha and rhamnose repeating unit only showed species in the 1.3 ppm region only (Figure 3.8c blue and purple) showing successful removal of GlcNAc sidechain modification by *gacI* gene deletion.



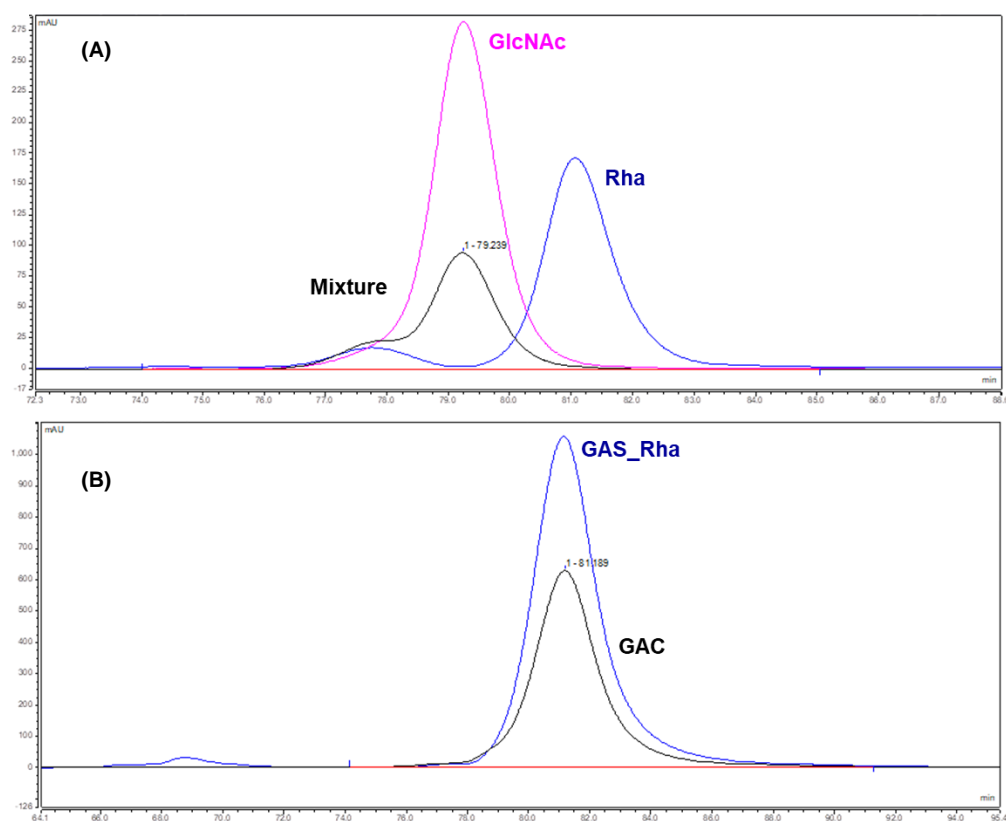
**Figure 3.8: Analysis of extracted and purified wildtype GAC and mutant GAS\_Rha polysaccharides.**

**(A)** Eluted polysaccharide following purification were dotted onto nitrocellulose membrane. Both GAC and GAS\_Rha are recognised by polyclonal anti-GAS (Thermo Fisher Scientific) antibodies detected by rabbit anti-Goat IgG HRP.

**(B)** sWGA lectin staining of extracted and purified 100 µg/ml GAC WT and GAS\_Rha.

**(C)** <sup>1</sup>H NMR bi-dimensional spectra of wildtype GAC (red) and GAS\_Rha (blue) polysaccharide. These were compared to monosaccharides making up repeating units either rhamnose and GlcNAc (green) or rhamnose alone (purple). Rhamnose region and N-Acetyl CH<sub>3</sub> region indicated in greyed out sections at 1.3 and 2.1 ppm respectively.

Wildtype GAC and GAS\_Rha polymers were sized using SEC-MALS detection, as the method of choice for polysaccharide molar mass characterisation. A marker consisting of DNA, thyroglobulin, BSA, carbonic anhydrase and tyrosine were used as calibrants (data not shown). Initially, the UV<sub>214nm</sub> signal was tested for GAC components, Rhamnose, GlcNAc and a mixture as present in the wildtype repeating unit (Rha – Rha – (GlcNAc)). UV signals showed that polysaccharide components eluted close to the TSK gel G3000<sub>PWXL</sub> and G2000<sub>PWXL</sub> analytical columns (used in tandem) total retention volume (Vt) at 82.9 minutes, with main peaks observed at 81.1 minutes for rhamnose, 79.2 minutes for GlcNAc, as well as a mixture, observed as a main peak at 79.2 minutes and a fronting shoulder at 77.5 minutes (Figure 3.9a). Wildtype GAC and GAS\_Rha were then run in the same manner as the polysaccharide constituents initially analysed by UV<sub>214nm</sub> signals. WT GAC eluted at 81.2 minutes compared to GAS\_Rha which eluted at 81.1 minutes, again close to the Vt at 83.0 minutes for this particular run (Figure 3.9b). Both samples have a similar elution time to each other, suggesting that both polysaccharides are small molecules due to their closeness to the Vt with similar molecular weights. However, WT GAC does show a greater degree of peak symmetry, compared to GAS\_Rha which shows some peak trailing, suggesting the presence of lower molecular weight material.



**Figure 3.9: Purified polysaccharide wildtype GAC and GAS\_Rha signals at UV<sub>214nm</sub> wavelength.**

**(A)** Polysaccharide components analysed using Size Exclusion Chromatography (SEC) at wavelength 214<sub>nm</sub> showing rhamnose (blue), GlcNAc (pink) and 2 : 1 Rhamnose to GlcNAc mixture (black).

**(B)** SEC chromatograms of wildtype GAC (black) and GAS\_Rha (blue) at wavelength 214<sub>nm</sub>.

Due to poor signal using UV<sub>214nm</sub> detection, polysaccharide components and samples were then analysed using MALS detectors. Accurate molecular mass determination of BSA was used to determine system suitability, however, rhamnose and GlcNAc components alone gave poor refractive index (RI) signals. Consequently, determinations of molecular mass of polysaccharide components, rhamnose and GlcNAc were inaccurate, although the known molecular masses for each could be used in calculations. Analysis of purified polysaccharide samples also proved difficult due to the associated small molecular masses and low signals obtained. Although not quantitative, in comparison GAS\_Rha appears to be smaller than WT GAC (Table 3.9). The average molecular weight (*mw*) (g/mol) was found to be 9.837 kDa for WT GAC which is within the size range stated in the literature<sup>139</sup>, whereas the *mw* of GAS\_Rha was found to be 4.892 kDa, which is slightly smaller than the reported average molar mass for native GAC without GlcNAc sidechains<sup>412</sup>.

In addition to average molecular weight, polydispersity ( $M_w/M_n$ ) which shows heterogeneity in molar mass distributions, defined as the weight average ( $M_w$ ) divided by the number average molecular weight ( $M_n$ ), was also determined. GAS\_Rha is more uniform in its average size ( $mw/mn = 1.847$ ) compared to GAC sample ( $mw/mn = 3.317$ ) which may have polysaccharides of varying molecular weights. Structural conformation was also determined through analysis of the sample hydrodynamic radius. Similar values were obtained for both samples in terms of hydrodynamic radius (4.469 and 4.586), suggesting both polysaccharides have similar conformations and physical properties in solution (Table 3.9).

**Table 3.9: SEC-MALS analysis of purified polysaccharide wildtype GAC and GAS\_Rha.**

Sample	Molecular mass <sup>1</sup> (g/mol)	Molecular mass (Mw) <sup>2</sup> (g/mol)	Polydispersity (Mw/Mn)	Hydrodynamic radius (Rh(v)n) (nm)
<b>BSA</b>	66,463	69,480 (1.4)	1.308 (2.0)	7.101 (34.4)
<b>WT GAC</b>	7,000 – 9,900 <sup>139,411</sup>	9,837 (54.2)	3.317 (62.3)	4.469 (28.2)
<b>GAS_Rha</b>	7,200 <sup>412</sup>	4,892 (34.7)	1.847 (57.7)	4.586 (32.3)

<sup>1</sup> Accepted molecular mass.

<sup>2</sup> Molecular mass determined by MALS, using RI of 0.185 for protein and 0.150 for polysaccharide samples. (% error included as calculated by Astra Software).

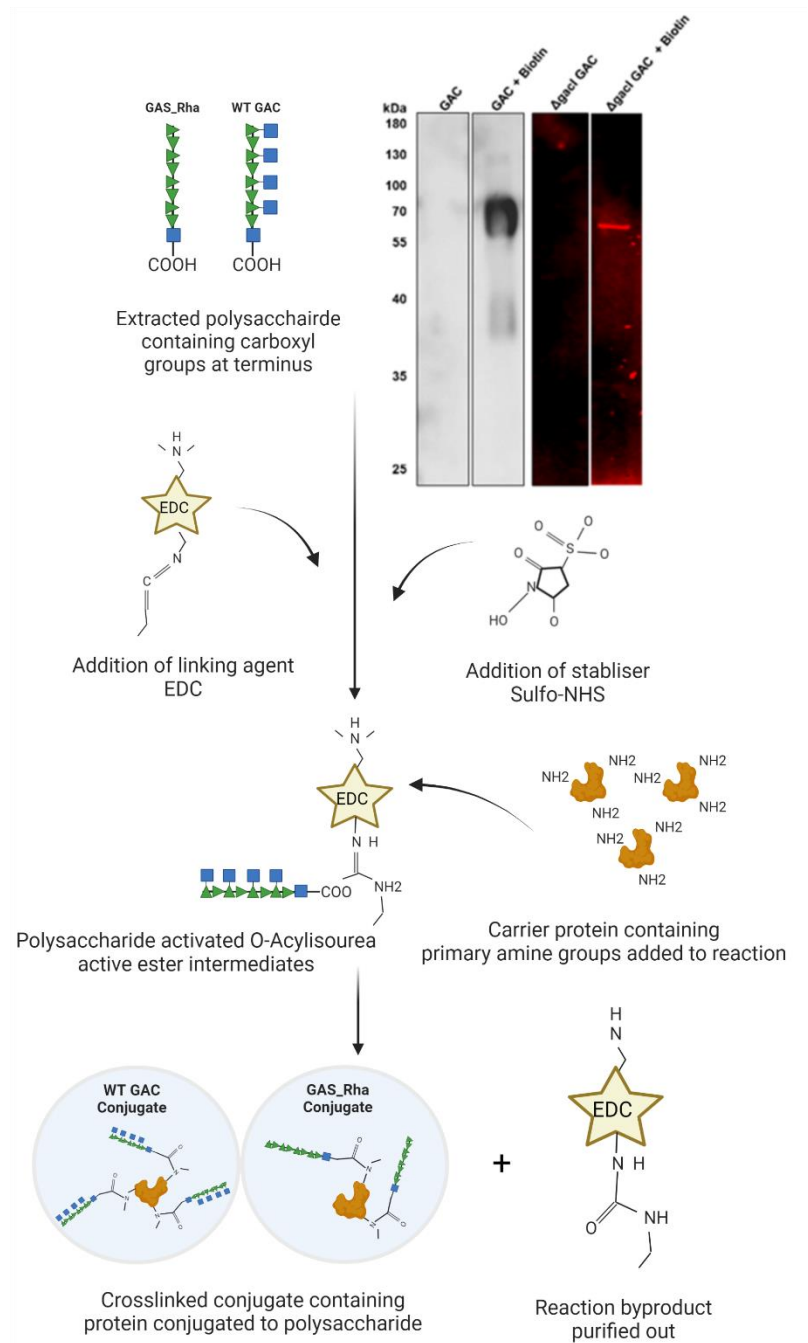
#### 3.4.4. Selection of EDC Conjugation Chemistry Approach

EDC carbodiimide chemistry was selected as the conjugation approach as a well-established, safe, and simple procedure<sup>524</sup>, utilising amine and carboxylic acid groups for linkage (Figure 3.10). EDC suitability for linking GAC variants with GAS protein antigens and the classical protein carrier TT was experimentally tested using *in silico* and *in vitro* assays. These were conducted to confirm the availability of primary amines on carrier proteins and carboxylic acids on the MurNAc residue of peptidoglycan dimers on extracted GAC variants before proceeding with the conjugation reactions.

##### 3.4.4.1. *Demonstration of Available Carboxylic Acid Group on Extracted Polysaccharide*

It is predicted that enzymatically extracted polysaccharide should contain a free carboxylic acid group present on the MurNAc following cleavage from the peptide chain of the cell wall by PlyC. To confirm this experimentally, EDC conjugation of amine-PEG<sub>3</sub>-biotin molecules was conducted with wildtype GAC and GAS\_Rha. Reactions were analysed by SDS-PAGE and western blotting probed with streptavidin conjugated to either HRP or IRDye® 680RD dye to detect conjugated biotin to an exposed carboxylic group present on the GAC variants (Figure 3.10). Both wildtype and GAS\_Rha showed a band ~ 60 kDa present in the sample incubated with EDC and biotin, but not in the polysaccharide only control sample. Despite the same amount of polysaccharide added to each reaction, the more sensitive IRDye® 680RD Streptavidin was required for the detection of biotin attached to GAS\_Rha. Nevertheless, this not only validates that both GAC variants contain the necessary reactive group for targeting, but also EDC can successfully conjugate a small molecule to it.





**Figure 3.10: Chemical conjugation strategy and validation of polysaccharide carboxylic group availability.**

Mechanism of direct coupling of carboxylic acid containing molecules to amine containing molecules using EDC carbodiimide chemistry coupled with Sulfo-NHS stabiliser. Possible products and intermediates are shown resulting in crosslinked conjugate containing polysaccharide and protein. Immunoblot of enzymatically extracted wildtype and GAS\_Rha conjugated with biotin by EDC to confirm carboxylic groups present, detected with streptavidin antibodies conjugated to either HRP for biotin-GAC detection, or IRDye® 680RD conjugate for biotin-GAS\_Rha detection.

Original figure created with BioRender.com adapted from Bioconjugate Techniques, 3rd Edition (2013) by Greg T. Hermanson<sup>520</sup>.

### 3.4.4.2. Available Amine Groups on Carrier Proteins

When selecting prospective carrier proteins, *in silico* methods were used to assess the total lysine (K) availability of candidates. The number of K residues on the molecular surface is an important factor as it can significantly affect the polysaccharide antigen loading efficiency. Total K, as a measure for amine group availability for chemical conjugation was predicted *in silico* based on translated codon sequences, as well as quantified experimentally using TNBSA assay (Table 3.10).

GAS carrier protein amino acid sequences were analysed to calculate the percentage K of total amino acids. TT as a classical protein carrier, known to be a good acceptor for polysaccharide attachment using EDC based chemistry<sup>525–528</sup> was used as control and comparator. Results for TT showed 7.7 % total K content. This was similar to SpyCEP and SpyAD proteins total K with 7.9 % and 7.5 % respectively. Despite being the smallest protein tested, MalE had the most K at 10.8 % of its total structure based on its amino acid sequence (Table 3.10).

**Table 3.10: L-Lysine (K) availability as a measure of amine groups on GAS protein carriers and monomeric TT for targeted EDC chemical conjugation.**

K availability was determined using both *in silico* tool Geneious Prime Version 2020.0.3, and experimentally using TNBS solution. Surface exposed amine groups were determined and expressed as nmol free amine / mg of protein. Molar concentrations of total protein were calculated, and free amine concentration calculated as a percentage of total protein and total lysine availability.

Protein name	Total K Content ( <i>in silico</i> )	Experimental K availability (nmol/mg protein)	Molar conc. of protein (nmol/mg protein)	Percentage available K in total protein	Available K (experimental) in total K ( <i>in silico</i> )	Theoretical number of K available
TT	7.7 % 102/1315	0.42	16.90	2.5 %	32 %	33
SpyCEP	7.9 % 44/588	0.43	16.64	2.6 %	33 %	15
MalE	10.8 % 41/384	0.36	24.33	1.5 %	13 %	5
SpyAD	7.5 % 61/813	0.53	11.33	4.7 %	62 %	38

*In silico* analysis, however, did not indicate how many of these lysine residues could be targeted by the EDC linking agent. This was further investigated experimentally using the TNBSA assay to quantify the amount of available lysine residues containing the required  $\epsilon$ -amine groups. TNBSA reacts with amine groups in a trinitrophenylation reaction to form highly chromogenic derivatives which can be quantified<sup>529–531</sup>. Experimental K nmol / mg of protein was determined and expressed as a percentage of available K in total protein, as well as percentage of total K residues which were surface exposed within the protein. Experimental findings suggest that despite a similar number of K residues in each antigen based on *in silico* data, SpyAD contains the most available K residues, followed by SpyCEP and then MalE (Table 3.10). In comparison, monomeric TT showed a similar number of surface exposed K residues to SpyAD, indicating that there are a large range of residues that can be targeted for EDC conjugation.

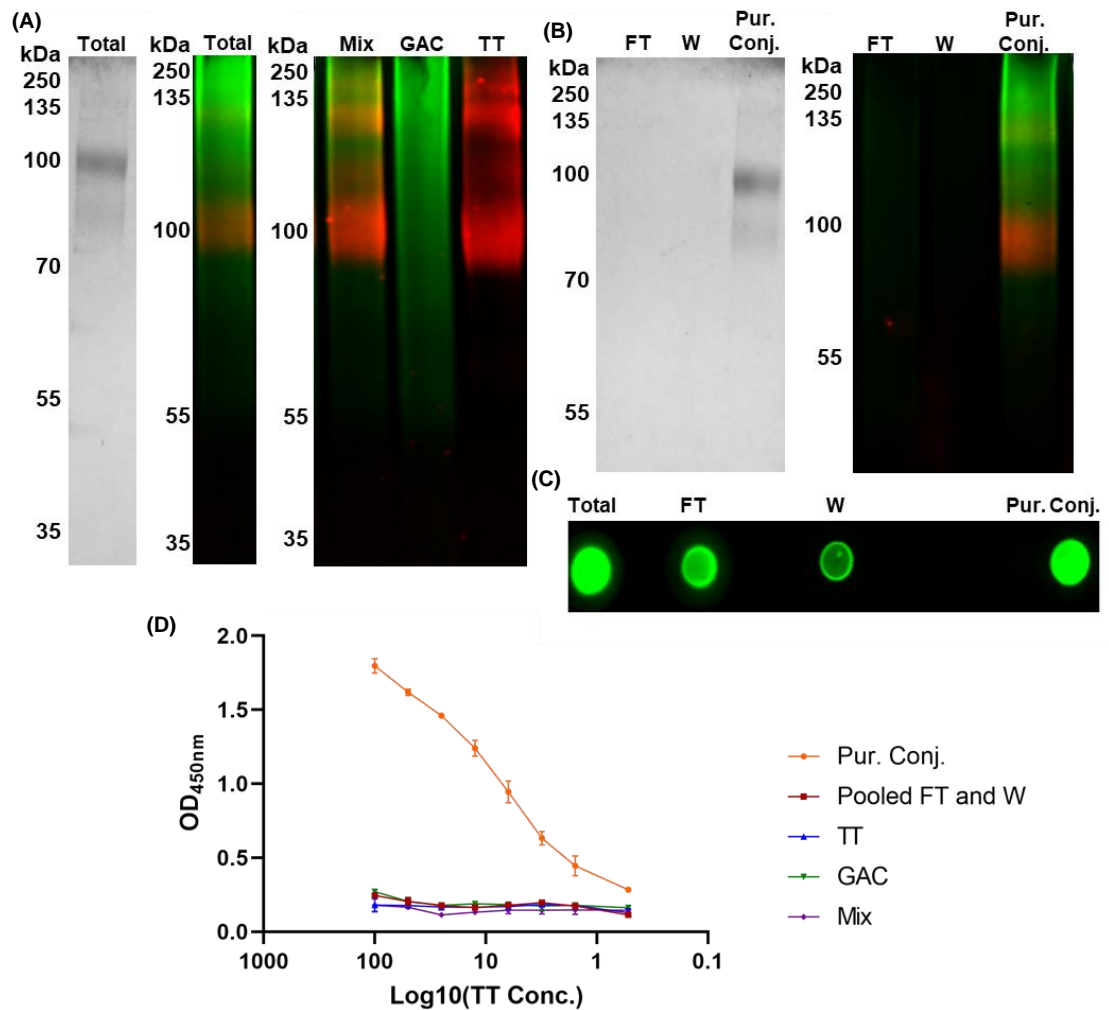
#### 3.4.5. Chemical Conjugation of TT Classical Protein Carrier and Wildtype GAC

As TT has been demonstrated as a good carrier for GAC in previous studies with different manufacturing chemistries<sup>409,485</sup> (Chapter 1 Table 1.2) it was the first carrier protein tested with EDC chemistry in this study as a control to optimise conditions. TT and GAC were incubated with EDC and sulfo-NHS and recovered protein and polysaccharide components from this sample, termed total reaction, were quantified using BCA and Anthrone respectively. Total protein recovery was found to be consistently between ~ 60 – 80 % of the calculated protein added to each reaction. This is likely due to the observation of some protein precipitation when setting up the reaction. Recovery for GAC was consistently higher than protein recovery at >95 % of the amount of calculated GAC added into each reaction.

The total reaction sample was then passed through a 50 kDa centrifugal spin filter to separate unconjugated GAC from the conjugate sample and unconjugated TT. This sample was termed purified conjugate (pur. conj.). Both an aliquot of the total reaction and purified conjugate sample along with reaction components were analysed by SDS-PAGE Coomassie and dual stained for TT and GAC by western blot using Guinea Pig anti-TT and Rabbit anti-GAS antibodies respectively (Figure 3.11a/b). GAC was also detected by dot blot showing some GAC present in the flowthrough (FT) and wash (W) samples (Figure 3.11c). This was likely unconjugated material as expected, with Figure 3.11b showing a loss of lower molecular weight GAC species in the purified conjugate sample compared to the total reaction sample (Figure 3.11a). Purified GAC alone was also passed through

the ultracentrifugation filter to ensure the absence of non-specific interactions, confirming specific separation (Appendix Figure 7.4a).

The resultant purified conjugate sample therefore contained the TT-GAC conjugate, in addition to unconjugated TT molecules. However, Figure 3.11a/b demonstrates that TT mixed with GAC in a 1 to 1 w/w ratio does not look dissimilar from the proposed conjugate sample. There is no clear molecular weight increase for TT which may be expected if TT was decorated with multiple GAC molecules. Therefore, an ELISA format was adopted to show physical attachment of TT and GAC molecules. Plates were coated with Guinea Pig anti-TT antibodies overnight before the addition and incubation of samples. These included the purified conjugate sample believed to contain TT-GAC, as well as unconjugated GAC pooled from FT and W samples from the purification step (Figure 3.11c), as well as TT and GAC separately, and as a mixture as controls. Detection of conjugate binding was evaluated using Goat anti-GAS antibodies followed by Rabbit anti-Goat IgG conjugated to HRP. The purified conjugate sample gave a 450<sub>nm</sub> absorbance signal ~ 1.8, titratable down the plate from 100 µg/ml protein concentration, whereas the control samples, most notably the mixture of GAC and TT gave minimal absorbance signals (Figure 3.11d). This suggests that the sample retained after purification contains molecules of TT conjugated to GAC molecules.



**Figure 3.11: Analysis of TT-GAC conjugation reaction.**

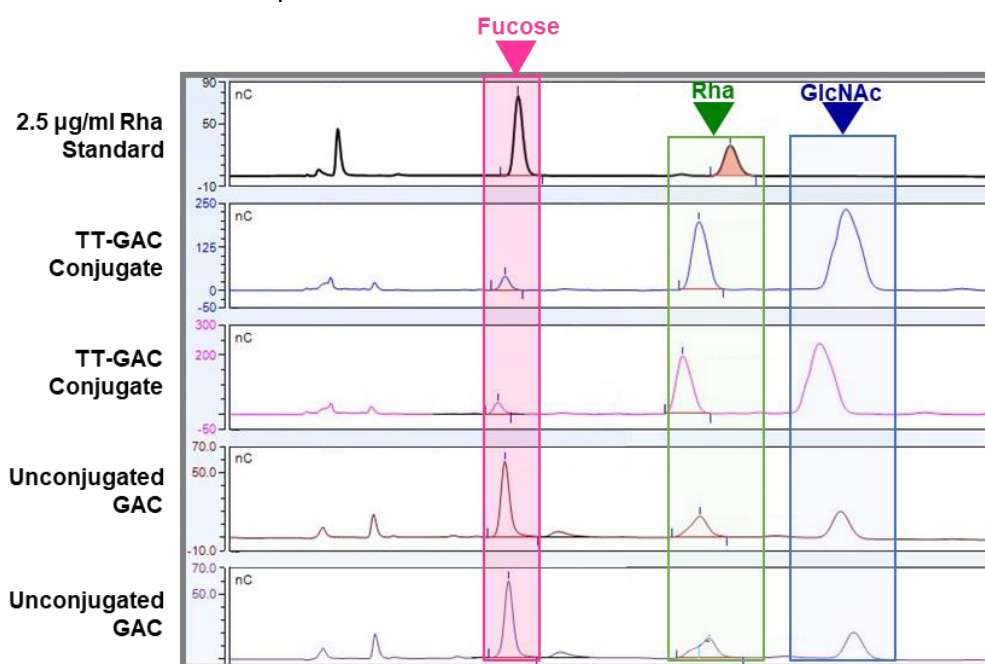
**(A)** Coomassie or dual stained by western blot of buffer exchange of conjugation reaction to remove reactants produce total reaction (total). Western blot analysis of total reaction (total) compared to a mixture (mix) or individual components (GAC or TT).

**(B / C)** Total reaction was then purified with a 50 kDa MWCO spin step to remove low MW components through flow through (FT) and wash (W) steps to produce purified conjugate (pur.conj.). **B:** Coomassie or dual staining by western blot and **C:** Dot blot with anti-GAC detection only.

Dual stained western blots incubated with Rabbit anti-GAS developed with Goat anti-Rabbit IgG IRDye® -800CW (GAC = green), and Guinea Pig anti-TT detected with Donkey anti-Guinea Pig IgG IRDye® - 680RD (TT = red).

**(D)** Capture ELISA to confirm covalent attachment. Plates captured conjugate samples using Guinea Pig anti-TT antibodies detected with Goat anti-GAS antibodies followed by Rabbit anti-Goat IgG - HRP. Samples were titrated based on TT concentration or equivalent GAC concentration for GAC samples. Samples include purified conjugate (orange), unconjugated GAC pooled from FT and W samples (red), and TT (blue) and GAC (green) alone, and a component mixture 1 : 1 (w/w) (purple).

High-Performance Anion-Exchange Chromatography with Pulsed Amperometric Detection (H-PAD) was also used as a more accurate method of polysaccharide quantification using known monosaccharide concentrations (Figure 3.12, top panel). Purified GAC could be resolved and detected by the gold working electrode, showing two distinct peaks for rhamnose and GlcNAc monosaccharides making up the repeating unit. Based on rhamnose monosaccharide standards, a rhamnose peak at 11 minutes, and a major peak representing GlcNAc at 13 minutes retention time could be identified. H-PAD showed that the purified conjugate sample contained 127  $\mu\text{g}$  conjugated GAC, equating to 39 % of 328  $\mu\text{g}$  recovered GAC from the total reaction based on anthrone analysis. The pooled FT and W samples contained 115  $\mu\text{g}$  unconjugated GAC, equating to 35 % of the recovered GAC being unconjugated (Table 3.11). Based on this data 26 % of GAC was unaccounted for perhaps due to losses during the purification process. This data was consistent with polysaccharide quantification based on anthrone reagent for purified conjugate and pooled FT and W samples.



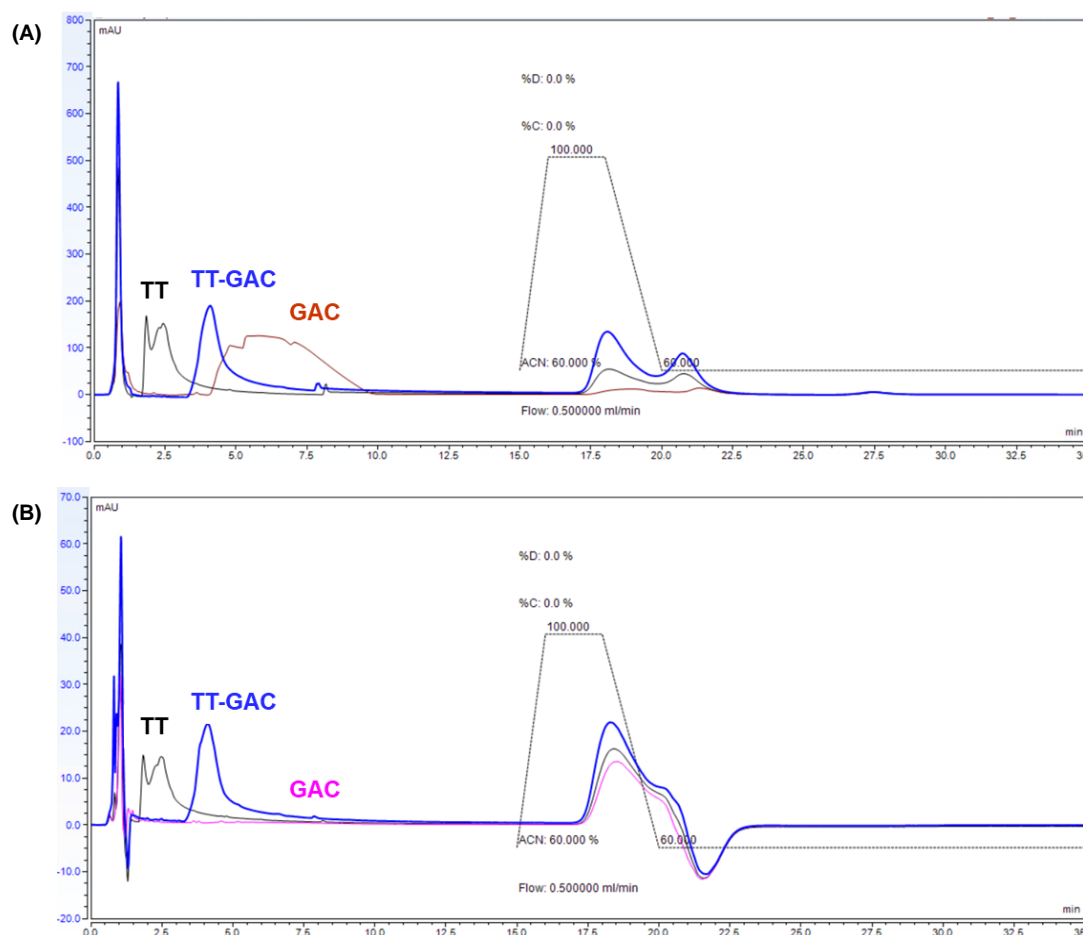
**Figure 3.12: H-PAD polysaccharide quantification of purified TT-GAC conjugation reaction.**

A rhamnose standard curve (0.5 – 10  $\mu\text{g}/\text{ml}$ ), TT-GAC conjugate, unconjugated GAC pooled from flow through and washes were run on a CarboPac PA-10 column with an AminoTrap™ pre-column with a fucose internal reference (pink shaded box). Two distinct peaks can be resolved for rhamnose (Rha) (green shaded box) and GlcNAc (blue shaded) for all samples. 2.5  $\mu\text{g}/\text{ml}$  rhamnose standard is shown as an example to quantify GAC concentrations within samples based on peak area, relative to the fucose internal reference.

The protein content within the purified conjugate sample was determined to be 170 µg, equating to 68 % of the 250 µg of TT recovered in the total reaction (Table 3.11). The molecular weight spin-based separation method chosen to purify the conjugate from unconjugated GAC would not have removed unconjugated TT. Therefore, the purified conjugate sample likely contains TT-GAC conjugate and unconjugated TT. To quantify conjugated versus unconjugated protein, immunoprecipitation analysis was performed. Using anti-GAS antibodies and PEG to precipitate GAC molecules, unconjugated TT could be separated from the conjugated TT for quantification. Supernatant containing free TT and resuspended conjugate precipitant were tested for TT concentration by ELISA based on the OD<sub>450nm</sub> signal from a generated TT standard curve ranging from 0.01 to 25 µg/ml. Results showed that 58% of TT in the purified conjugate sample was conjugated (Table 3.11). Based on this, the protein to polysaccharide ratio for the TT-GAC conjugate could be calculated. The reaction achieved between 1 : 1.06 and 1 : 1.30 w/w TT to GAC (TT = 98 µg and GAC = 104 / 127 µg). However, when considering the molecular weight of TT and GAC at 90,000 Da and 9,837 Da (Table 3.9) respectively, this can be converted into a molar ratio of between 1 : 9.72 and 1 : 11.87 m/m TT (1.08 µM) to GAC (10.57 – 12.91 µM) (Table 3.12).

To further understand these ratios, additional TT-GAC characteristics were analysed. Initially traditional HPLC-SEC could not demonstrate a discernible size shift between TT-GAC and TT protein alone (Appendix Figure 7.6a). Therefore, an alternative method, Reverse Phase-High Performance Liquid Chromatography (RP-HPLC) was used to assess whether conjugation alters the hydrophobicity of TT when conjugated to GAC. An experiment was designed to observe any shift in the hydrophobic elution profiles by assessing differences in retention time when using an isocratic elution acetonitrile gradient. The aim was to compare conjugate to components TT and GAC alone. TT and TT-GAC conjugate were diluted to 210 µg/ml based on protein concentration as calculated by BCA. This equated to 273 µg/ml GAC within the glycoconjugate sample based on the conjugate ratio 1 : 1.30 w/w (Table 3.12). GAC alone was injected at 50 µg/ml as the highest concentration achievable based on material constraints. The chromatograph in Figure 3.13a showing UV<sub>214nm</sub> signal indicates a clear shift in retention time suggesting different hydrophobic characteristics of TT-GAC glycoconjugate (blue line), compared to components TT (black line) and GAC (brown line) alone. The consistency of the separation of RP-HPLC standard injected both before and after test samples confirmed the integrity of the Gemini C-18 110Å 3 µm (3 x 100 mm) column

(Phenomenex) (Appendix 7.6b). Results also indicated the relative absence of unconjugated GAC in the glycoconjugate sample, providing further confirmation of conjugation. As can be seen in UV<sub>214nm</sub> chromatogram there is a prominent peak that can be observed for the purified GAC sample (Figure 3.13a brown line). The identity of this peak as a GAC peak can be verified by its absence in the UV<sub>280nm</sub> chromatograph which is specific to proteins (Figure 3.13b pink line). As almost 6 times more GAC was present in the TT-GAC conjugate sample, if unconjugated GAC was present it would be expected to generate a matching if not larger peak to the unconjugated GAC sample. Therefore, although difficult to prove directly, this suggests that conjugation of TT with GAC changes the hydrophobicity of the molecule. TT likely becomes more hydrophobic, causing it to interact more with the C-18 column, explaining the differences in retention time, and proving successful conjugation.



**Figure 3.13: Reverse Phase HPLC chromatograms of chemical glycoconjugate TT-GAC.**

Samples were injected into a Gemini C-18 110Å 3 μm 3x100 mm column (Phenomenex®) and analysed by UV<sub>214 nm</sub> (A), and UV<sub>280nm</sub> (B). TT-GAC glycoconjugate sample (blue line) were compared to TT protein (black line), and GAC (brown / pink line) only samples, to observe differences in retention time indicating hydrophobicity changes between components and the glycoconjugate sample.



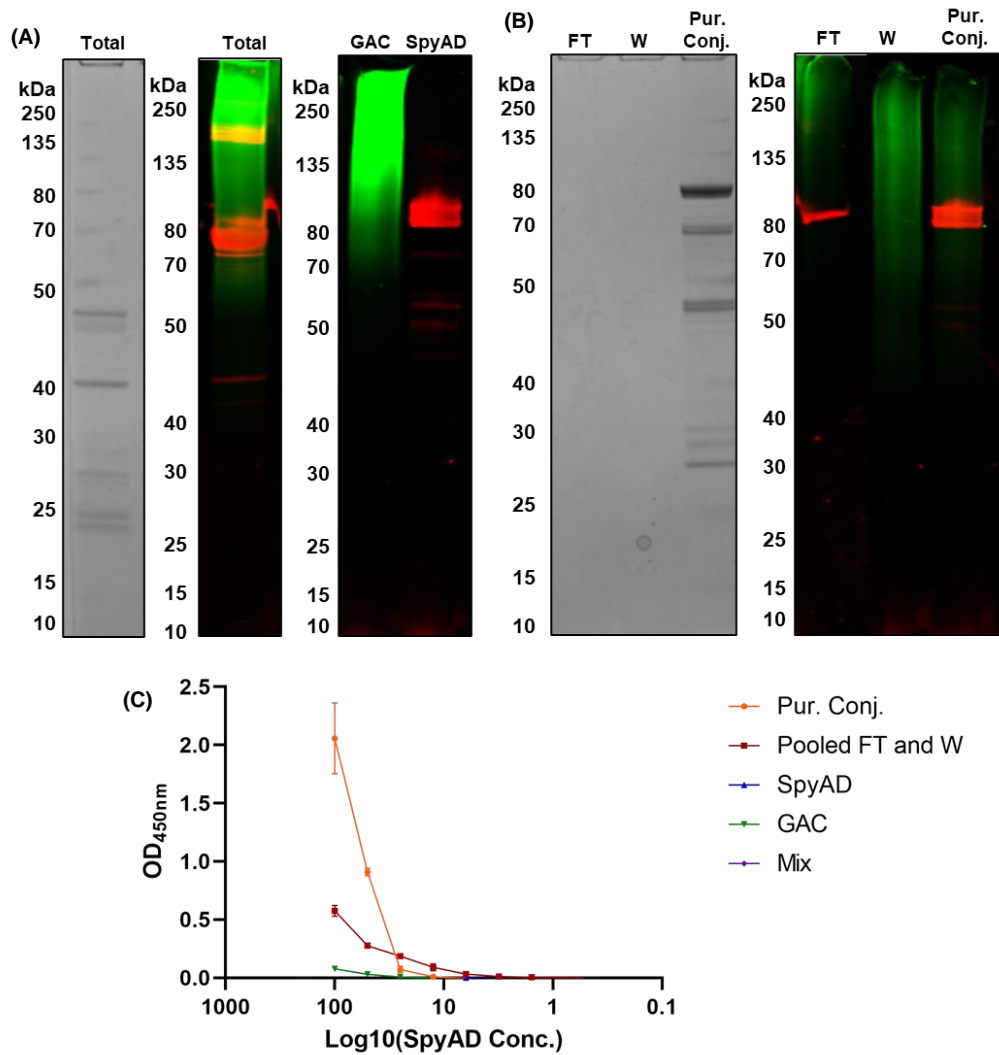
### 3.4.6. Chemical Conjugation of GAS Protein Antigens and Wildtype GAC

#### *3.4.6.1. Successful Conjugation of SpyAD to Wildtype GAC*

The optimised TT methodology of 16 hours incubation followed by buffer exchange into PBS pH 7.2 using a 5 kDa MWCO ultracentrifugation filter was undertaken for SpyAD (Figure 3.14a). The total reaction was then passed through a 50 kDa MWCO centrifugal spin filter, and the retentate sample predicted to contain SpyAD-GAC glycoconjugate and unconjugated SpyAD was termed purified conjugate. In this sample SpyAD was seen to be successfully decorated with wildtype GAC following analysis by SDS-PAGE Coomassie and dual stained western blot for the anti-His signal on SpyAD and anti-GAS for GAC (Figure 3.14b). Total protein recovery following the reaction was found to be 71 % of the calculated starting protein material and GAC recovery found to be 82 % of the total calculated GAC added into the initial reaction.

Unconjugated GAC and some SpyAD appeared to be lost in the flow through (FT) and wash (W) steps during conjugate purification (Figure 3.14b). As with TT-GAC, despite normalised protein loading concentrations there was no clear molecular weight increase for SpyAD in the purified conjugate sample, however there is a reduction in the lower molecular weight GAC species following unconjugated GAC removal.

To confirm that the purified conjugate sample contains GAC attached to SpyAD, ELISA plates were coated with anti-GAS antibodies and conjugate binding was detected for SpyAD with mouse anti-His antibody. SpyAD-GAC conjugate was compared to unconjugated GAC in the FT and W samples Figure 3.14b, and the glycoconjugate components separately, and as a physical mixture (Figure 3.14c). The conjugation sample gave an OD<sub>450nm</sub> absorbance signal around 2, titratable down the plate until a 1 in 4 dilution, correlating to 25 µg/ml of SpyAD before the signal plateaued to baseline. A slight signal was observed in the unconjugated GAC sample likely related to the detection of SpyAD in the FT sample as shown by western blot during the purification step (Figure 3.14b). This suggests that some conjugate may have passed through the spin filters. However, notably, the control GAC/SpyAD mixture samples gave minimal absorbance signal, and the purified conjugate sample gave the highest signal level.



**Figure 3.14: Analysis of SpyAD-GAC conjugation reaction.**

**(A)** Coomassie and dual stained western blot of buffer exchange of conjugation reaction to remove reactants produce total reaction (total). Western blot analysis of total reaction (total) compared to individual components (GAC or SpyAD).

**(B)** Total reaction was then purified with a 50 MWCO spin step to remove low MW components through flow through (FT) and wash (W) steps to produce purified conjugate (pur.conj.) Coomassie or dual stained by western blot.

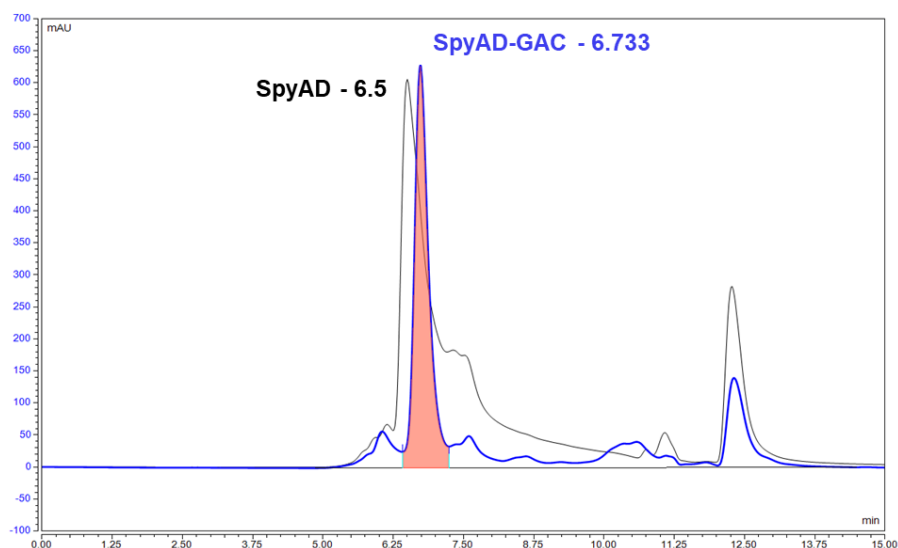
Dual stained western blots incubated with Rabbit anti-GAS developed with Goat anti-Rabbit IgG IRDye® -800CW (GAC = green), and Mouse anti-His detected with Rabbit anti-Mouse IgG IRDye® - 680RD (SpyAD = red).

**(C)** Capture ELISA to confirm covalent attachment. Plates captured conjugate samples using Goat anti-GAS antibodies and detected with anti-His HRP antibodies. Samples were titrated based on SpyAD concentration or equivalent GAC concentration for GAC samples. Samples include purified conjugate (orange), unconjugated GAC pooled from FT and W samples (red), and SpyAD (blue) and GAC (green) alone, and a component mixture 1 : 1 (w/w) (purple).

Based on H-PAD analysis the purified conjugate sample, made up of SpyAD-GAC and unconjugated SpyAD contained 195 µg conjugated GAC, equating to 48 % of the GAC present in the total reaction sample before purification. The pooled FT and W samples were also tested and found to contain 122 µg GAC, equating to 30 % of the total GAC added being unconjugated (Table 3.11), leaving 22 % of GAC unaccounted for compared to the total reaction polysaccharide content, suggesting potential lost in the MWCO purification.

Protein concentration determined by BCA showed that the purified conjugate sample contained 210 µg SpyAD, equating to 57 % of the protein present with the total reaction sample (370 µg SpyAD recovered) (Table 3.11). Immunoprecipitation to determine unconjugated from conjugated SpyAD was not possible in this instance due to material constraints. Therefore, when considering total protein in the purified conjugate sample, SpyAD-GAC conjugation reactions were calculated to contain 1 : 0.99 *w/w* SpyAD to GAC, converted to a molar ratio of between 1 : 8.84 and 1 : 8.33 *m/m* SpyAD (2.38 µM) to GAC (19.82 – 21.04 µM) based on their molecular weights of 88,000 Da and 9,837 Da (Tables 3.9 and 3.12).

In addition to understanding component ratios, SpyAD-GAC glycoconjugate was successfully analysed using HPLC-SEC techniques. A protein size shift is influenced not only by the protein to polysaccharide ratio, but also the hydrodynamic radius of the conjugated molecule. Using a BioZen 1.8 µm SEC-2 LC column successful conjugation could be demonstrated with SpyAD protein alone eluting at 6.5 minutes and SpyAD conjugated to GAC eluting at 6.733 minutes (Figure 3.15). SpyAD and SpyAD-GAC consistently eluted at different retention times with different runs validating the small difference in retention properties (Appendix section 7.5b).



**Figure 3.15: HPLC-SEC chromatogram of chemical glycoconjugate SpyAD-GAC.**

Glycoconjugate samples were injected into a bioZen 1.8  $\mu\text{m}$  SEC-2 LC column (Phenomenex®) with 150 pore size ( $\text{\AA}$ ) and analysed by SEC at  $\text{UV}_{214\text{nm}}$ . SpyAD-GAC injected glycoconjugate samples (blue line) were compared to protein only samples (black line), to observe a size change indicative of successful conjugation. Red shading shows the peak area for glycoconjugate samples.

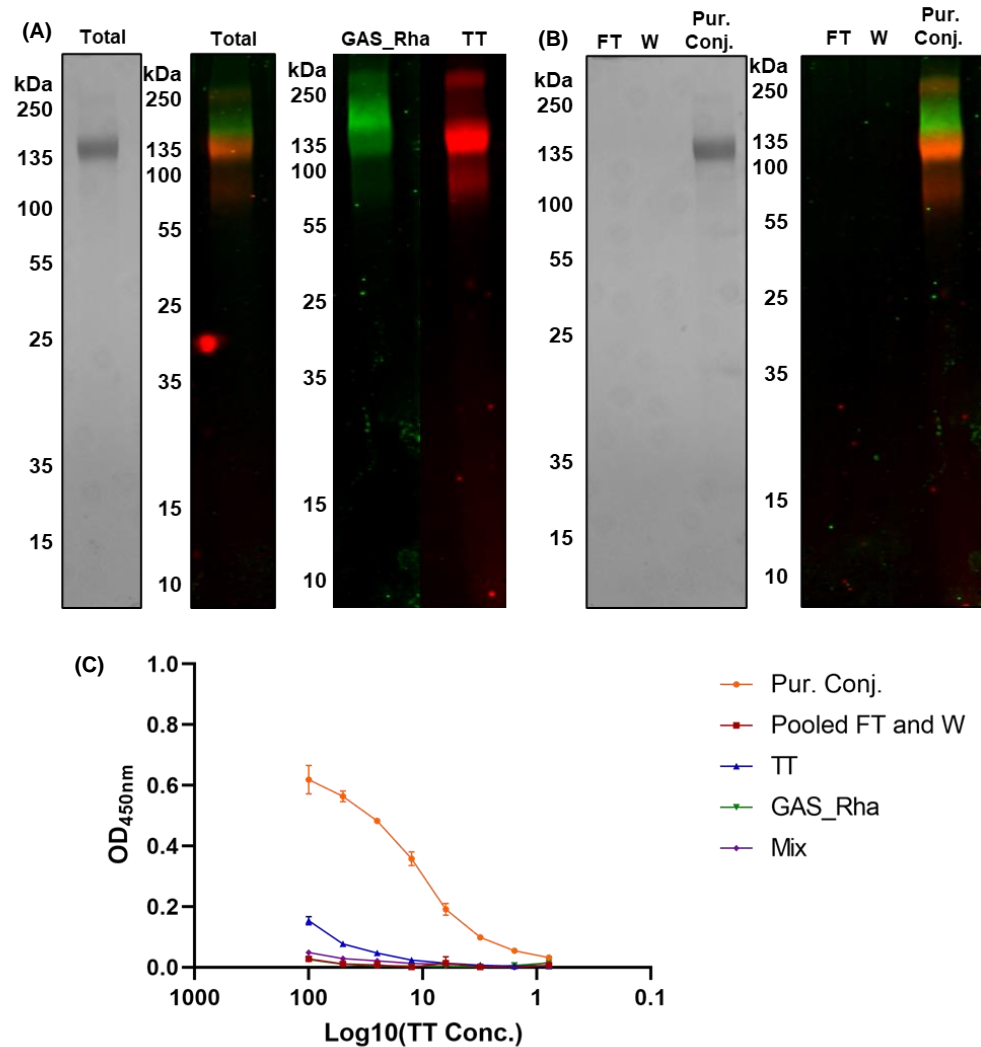
#### 3.4.6.2. *MalE and SpyCEP Failed to Conjugate to Wildtype GAC*

Attempts to conjugate MalE and SpyCEP GAS antigens to wildtype GAC using the same chemical conjugation reactions as SpyAD and TT did not lead to the generation of successful glycoconjugates. A variety of conditions were tested using both EDC alone, as well as EDC and Sulfo NHS stabiliser, with a variety of reactant ratios, and reaction conditions such as pH and buffer systems. Additionally, a variety of purification methods were tested to separate unconjugated molecules from proposed glycoconjugates. These techniques mostly focused on using affinity chromatography, as opposed to the separation of components based on size which was feasible for larger proteins SpyAD and TT only. Table 7.3 in the Appendix summarises the reactions, conditions, and purification methods tested.

#### 3.4.7. Chemical Conjugation of TT Classical Protein Carrier and GAS\_Rha

The same conjugation experiment as above was performed with rhamnose polymers enzymatically extracted and purified from mutant  $\Delta gacI$  NCTC-8198 strain (Chapter 3 section 3.3.6) and monomeric TT. The conjugation reaction was initially buffer exchanged showing good retention of conjugation components with 98 % TT and 88% GAS\_Rha recovered based on the initial calculated reaction input.

The total reaction was then purified using a 50 kDa MWCO spin filter to remove unconjugated GAS\_Rha. The retentate sample, termed purified conjugate, was analysed by SDS-PAGE Coomassie and dual stained western blot for TT and GAS\_Rha along with an aliquot of the total reaction sample (Figure 3.16a/b). The buffer exchanged total reaction did not look dissimilar to the purified conjugate sample. Therefore, a similar ELISA format was used to confirm TT attachment to GAS\_Rha as stated above (Chapter 3 section 3.4.5). Plates were coated with anti-TT antibodies and glycoconjugate binding was detected with Goat anti-GAS antibodies. TT-GAS\_Rha gave an  $OD_{450nm}$  absorbance signal around 0.6, titratable down the plate from 100  $\mu g/ml$  TT protein concentration (Figure 3.16c). All control samples, except TT alone which showed low non-specific cross reactivity with the detection antibodies, gave no observable signal. However, in comparison to TT-GAC conjugate, the absorbance signal observed from TT-GAS\_Rha conjugates was lower, despite similar GAC recovery and standardised sample application (Chapter 3 section 3.4.5).



**Figure 3.16: Analysis of TT-GAS\_Rha conjugation reaction.**

**(A)** Coomassie and dual stained western blot of buffer exchange of conjugation reaction to remove reactants produce total reaction (total). Western blot analysis of total reaction (total) compared to individual components (GAS\_Rha or TT).

**(B)** Total reaction was then purified with a 50 MWCO spin step to remove low MW components through flow through (FT) and wash (W) steps to produce purified conjugate (pur.conj.) Coomassie or dual stained by western blot.

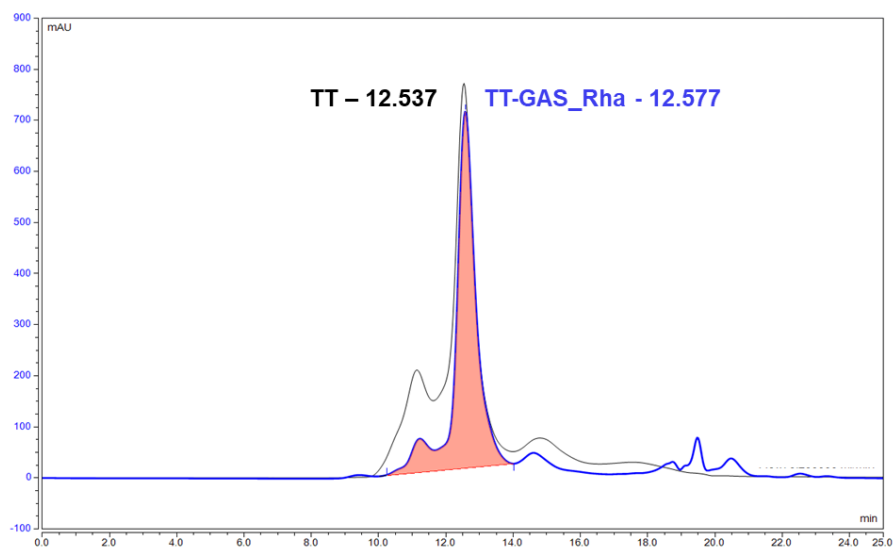
Dual stained western blots incubated with Rabbit anti-GAS developed with Goat anti-Rabbit IgG IRDye® -800CW (GAS\_Rha = green), and Mouse anti-His detected with Rabbit anti-Mouse IgG IRDye® - 680RD (TT = red).

**(C)** Capture ELISA to confirm covalent attachment. Plates captured conjugate samples using Guinea Pig anti-TT antibodies detected with Goat anti-GAS antibodies followed by Rabbit anti-Goat IgG - HRP. Samples were titrated based on TT concentration or equivalent GAS\_Rha concentration for the GAS\_Rha sample. Samples include purified conjugate (orange), unconjugated GAS\_Rha pooled from FT and W samples (red), and TT (blue) and GAS\_Rha (green) alone, and a component mixture 1 : 1 (w/w) (purple).

The calculated protein to polysaccharide ratio for TT\_GAS\_Rha conjugate was found to be lower than TT-GAC. Based on H-PAD analysis there were some discrepancies, with GAS\_Rha appearing to be poorly detected compared to WT GAC. H-PAD found that 11 % of GAS\_Rha was present in the purified conjugate sample (33 µg), and 21 % of GAS\_Rha was unconjugated in the pooled FT and W samples (61 µg). This means compared to polysaccharide in the total reaction 68 % of GAS\_Rha was unaccounted for using H-PAD analysis. On the other hand, the Anthrone assay on the same purified conjugate sample showed 68 µg conjugated GAS\_Rha, equating to 23 % of the polysaccharide in the total reaction before purification (295 µg GAS\_Rha). The tested pooled FT and W samples from the purification step were found to contain 154 µg GAS\_Rha, equating to 52 % of the total GAS\_Rha being unconjugated, reducing the amount of unaccounted polysaccharide (Table 3.11).

Protein concentration was determined with 340 µg of unconjugated and conjugated TT in the TT-GAS\_Rha sample, equating to 74 % of the TT protein present in the total reaction (460 µg) (Table 3.11). Similar to TT-GAC analysis, immunoprecipitation analysis separated unconjugated TT from conjugated TT. Based on OD<sub>450nm</sub> absorbance signal, 184 µg of TT was conjugated and 127 µg of TT unconjugated, equating to 54 % and 37 % of the purified conjugate sample TT content (Table 3.11). When considering conjugated TT and GAS\_Rha concentrations, the protein to polysaccharide ratio was calculated to be between 1 : 0.18 and 1 : 0.37 *w/w* TT to GAS\_Rha. This equates to a molar ratio of between 1 : 3.30 and 1 : 6.81 *m/m* TT (2.04 µM) to GAS\_Rha (6.75 – 13.90 µM) (Table 3.12) when considering the respective molecular weights of 90,000 Da and 4,892 Da (Table 3.9).

Further analysis of TT-GAS\_Rha was carried out in a similar manner to SpyAD-GAC (Chapter 3 section 3.4.6) using HPLC-SEC techniques. The aim was to show differences in retention time of TT-GAS\_Rha glycoconjugate and the TT component alone. TT-GAS\_Rha showed a minor size shift at 12.577 minutes compared to TT alone at 12.537 minutes (Figure 3.17). This was a consistent finding, with TT eluting at the same retention time and TT-GAS\_Rha conjugate eluting later following multiple runs (Appendix Figure 7.5a).



**Figure 3.17: HPLC-SEC chromatogram of chemical glycoconjugate TT-GAS\_Rha.** Glycoconjugate samples were injected into a bioZen 1.8  $\mu\text{m}$  SEC-2 LC column (Phenomenex®) with 150 pore size ( $\text{\AA}$ ) and analysed by SEC at  $\text{UV}_{214\text{nm}}$ . Injected glycoconjugate samples, TT-GAS\_Rha (blue line) were compared to protein only samples, TT (black line), to observe a size change indicative of successful conjugation. Red shading shows the peak area for glycoconjugate samples.



**Table 3.11: Quantification of chemical glycoconjugate vaccine components.**

Unconjugated (UC) component analysis. Polysaccharide concentration ( $\mu\text{g/ml}$ ) based on Anthrone<sup>1</sup> and H-PAD<sup>2</sup>, and protein concentration ( $\mu\text{g/ml}$ ) based BCA<sup>3</sup> analysis and immunoprecipitation – ELISA<sup>4</sup> analysis. NT = Not Tested.

Total PS and protein is following buffer exchange (total reaction) and the conjugate / unconjugated is following MWCO purification to generate the purified conjugate sample.

Reaction	Polysaccharide					Protein		
	Total <sup>1</sup>	Conjugated <sup>1</sup>	Conjugated <sup>2</sup>	UC <sup>1</sup>	UC <sup>2</sup>	Total <sup>3</sup>	UC and conjugated <sup>3</sup>	Conjugated <sup>4</sup>
<b>TT-GAC</b>	328	104 (32 %)	127 (39 %)	157 (49%)	115 (35%)	250	170 (68 %)	98 (58 %)
<b>SpyAD-GAC</b>	409	207 (51 %)	195 (48 %)	129 (32 %)	122 (30 %)	370	210 (57 %)	NT
<b>TT-GAS_Rha</b>	295	68 (23 %)	33 (11 %)	154 (52 %)	61 (21 %)	460	340 (74 %)	184 (54 %)

**Table 3.12: Total and conjugate ratios of chemical glycoconjugate vaccines.**

Weight ( $w/w$ ) and molar ( $m/m$ ) ratios of protein (TT or SpyAD) to polysaccharide (PS) (GAC or GAS\_Rha). Molar ratios were converted assuming protein molecular weight ( $mw$ ) to be 90,000 or 88,000 g/mol for TT or SpyAD respectively, and WT GAC and GAS\_Rha average polymer  $mw$  to be 9,837 and 4,892 g/mol respectively. NT = not tested. These are based on purified conjugate samples only.

Reaction	Conjugated PS <sup>1/2</sup> / Unconjugated & Conjugated Protein <sup>3</sup>		Conjugated PS <sup>1/2</sup> / Conjugated Protein <sup>3</sup>	
	Protein : PS $w/w$	Protein : PS $m/m$	Conjugate Protein : PS $w/w$	Conjugate Protein : PS $m/m$
<b>TT-GAC</b>	1 : 0.61 $w/w^{1,3}$ 1 : 0.75 $w/w^{2,3}$	1 : 5.59 $m/m^{1,3}$ 1 : 6.83 $m/m^{2,3}$	1 : 1.06 $w/w^{1,4}$ 1 : 1.30 $w/w^{2,4}$	1 : 9.72 $m/m^{1,4}$ 1 : 11.87 $m/m^{2,4}$
<b>SpyAD-GAC</b>	1 : 0.99 $w/w^{1,3}$ 1 : 0.99 $w/w^{2,3}$	1 : 8.84 $m/m^{1,3}$ 1 : 8.33 $m/m^{2,3}$	NT	NT
<b>TT-GAS_Rha</b>	1 : 0.2 $w/w^{1,3}$ 1 : 0.10 $w/w^{2,3}$	1 : 3.68 $m/m^{1,3}$ 1 : 1.79 $m/m^{2,3}$	1 : 0.37 $w/w^{1,4}$ 1 : 0.18 $w/w^{2,4}$	1 : 6.81 $m/m^{1,4}$ 1 : 3.30 $m/m^{2,4}$

## 3.5. Discussion

This Chapter demonstrates the successful generation of three chemically derived glycoconjugate vaccines, two containing classical protein carrier TT, and one 'double hit' glycoconjugate vaccine containing GAS antigen SpyAD. All three glycoconjugates were built using the same approach, a two-step reaction using EDC (Carbodiimide) and sulfo NHS stabiliser.

### 3.5.1. Glycoconjugate Antigen Selection

#### 3.5.1.1. *Polysaccharide Glycoconjugate Vaccine Components (Wildtype GAC and GAS\_Rha)*

GAC as an attractive vaccine component, was successfully extracted, purified, and conjugated to two different proteins (TT and SpyAD). However, WT GAC contains GlcNAc epitopes which may result in autoimmune antibody generation, therefore this work also investigated polysaccharide polymers containing rhamnose only (GAS\_Rha) as an alternative antigen. GAS\_Rha could be successfully conjugated to classical protein carrier TT. Testing of IVIG by ELISA showed antibodies present recognised WT GAC whereas testing for GAS\_Rha was inconclusive (data not shown). This confirms the presence of anti-GAC antibodies in the human population and that GAC is a viable vaccine candidate, in accordance with the literature<sup>140</sup>.

Prior to conjugation, to obtain GAS\_Rha, a deletion was introduced into the GAS genome to knockout the *gacI* gene within the *gac* operon. A previous publication generated a similar mutant GAS strain through plasmid integrational mutagenesis knocking out the *gacI* gene by inserting *cat*, Chloramphenicol Acetyltransferase gene, in GAS strain M1T1 5448<sup>141</sup>. This mutant did not show differences in colony morphology, nor changes in growth compared to the parent strain over a 7 hour time period in TH broth without yeast supplementation<sup>141</sup>. Rather they determined that the *gacI* deletion affects cell separation of stationary phase cultures<sup>141</sup>. Such differences were less pronounced in exponential phase cultures, perhaps explaining the apparent shorter GAS chains observed in exponential phase cultures by microscopy in this work. Such phenotypic differences may also be explained by differences in strain background leading to altered gene expression, particularly considering the spatial differences within the *gac* operon between an in-frame gene deletion and an insertional knockout.

Mutant GAS colonies showed reduced agglutination with Group A reagents compared to WT GAS. This test relies on latex particles interacting with group-specific antibodies, agglutinating in the presence of homologous antigen, therefore,

this was an expected result due to the loss of immunodominant GlcNAc residues. Additionally, *gacI* mutant whole cells treated with hyaluronidase, as well as purified GAS\_Rha polysaccharide samples showed reduced sWGA lectin blotting compared to WT GAS. Hyaluronidase treatment was required to remove the capsule which also contains GlcNAc residues to ensure GAC GlcNAc analysis wasn't confounded by this factor. These are similar observations, along with agglutination reactions, to the characterisation performed in a previous study<sup>141</sup>. Additionally, both WT and  $\Delta gacI$  strains were found to share similar whole cell lysate total protein profiles as stained by Coomassie blue (Appendix Figure 7.3a) and purified polysaccharides samples were found to react with polyclonal goat and rabbit anti-GAS sera, suggesting some antibodies found in these products are directed at the rhamnose backbone. Murine anti-protein serum also reacted with  $\Delta gacI$  GAS cell walls (Appendix Figure 7.3b).

However, normalised samples showed that purified GAS\_Rha consistently gave weaker anti-GAS antibody signals compared with WT GAC, likely explained by the loss of the immunodominant GlcNAc epitope, and the branching structure deemed to be important in epitope recognition<sup>445,532</sup>. In addition to immunological techniques, NMR was used to determine the identity of both polysaccharide variants. Rhamnose and GlcNAc specific regions could be distinguished for WT GAC, with a clear loss of the region resolving N-acetyl CH<sub>3</sub> (GlcNAc) in GAS\_Rha samples demonstrating that the *gacI* deletion results in the production of polysaccharide containing rhamnose only. Such polysaccharide identity differences may have been complimented by additional characterisation to determine the mole percentage amount of individual sugars. This type of HPLC tracing and linkage analysis has been previously performed on GAS rhamnose polysaccharides<sup>141</sup>, however, was not possible in the present study because of material constraints associated with low polysaccharide extraction yields following cell wall enzymatic digestion.

Along with identity verification, when using polysaccharides as vaccine antigens their size, specifically relating to polymer chain lengths, is important to analyse as this may be relevant for the vaccine's physicochemical and immunogenic properties. Average polymer chains were found to be 18 repeating units for WT GAC and 15 repeating units for GAS\_Rha. WT GAC is within the molar mass range observed in the literature. One group producing synthetic GAC determined it to be 8.9 +/- 1 kDa consisting of 18 repeating units<sup>139</sup>, whereas another group which chemically extracted GAC using acid treatment, reported an average polymer length of 14 repeating units, based on the average molecular size of 7.0 kDa, sized based on

dextran standards<sup>411</sup> rather than MALS analysis used in this work. Such minor differences may be possibly due to strain differences, material loss through the selection extraction method, or from the chosen measurement methodology.

The extracted GAS\_Rha polymers in this work showed shorter average chain lengths compared to the limited data available in the literature. Published studies characterising extracted rhamnose polysaccharides found the average chain size to be 7.2 kDa from a similar mutant GAS strain<sup>412</sup>. Differences may be due to changes in expression levels across the operon through the different mutation methods. For example, the gene deletion may have affected the expression of downstream genes within the *gac* operon due to spatial differences in the promoter region. Alternatively, as the  $\Delta gacI$  GAS mutant strain genome was not sequenced, it is possible there could be undetected mutations within the chain length regulator region. Generally, however, small size discrepancies between this work and the work of others may be explained by different polysaccharide extraction techniques. For example, acid treatment can hydrolyse bonds within the polysaccharide chain, observed in a study showing loss of the glycerol phosphate moiety in acid extracted WT GAC compared with enzymatically extracted GAC<sup>149</sup>.

In this study applying a traditional sizing workflow based on molecular weight and UV detection was not possible due to poor UV signals observed at the recorded wavelengths (214<sub>nm</sub> and 280<sub>nm</sub>) for rhamnose monosaccharides. In fact, when testing a range of UV wavelengths, it appeared rhamnose showed the best optical activity at lower wavelengths, such as 187<sub>nm</sub> (Appendix Table 7.2). Therefore, MALS was selected as an alternative detection method to determine molecular weight directly, independent of elution time. It also proved useful due to the lack of commercially available GAC related reference material. WT GAC and GAS\_Rha polymers were found to be relatively small and close to each other in size making the separation of these molecules difficult. MALS detection is only effective if good separation and resolution is achieved on the SEC column, with the accuracy for lower molecular weight species found to be influenced by sample load and chosen eluent. A variety of columns and mobile phases were tested in this work to ensure analytes were within the useful range of SEC, however even with an optimised system, high percentage errors were attributed to polysaccharide samples molar mass, polydispersity, and hydrodynamic radius readings. This shows that further optimisation of columns, compatibility with eluents<sup>533</sup>, as well as detection methods, may have yielded more accurate sizing data, as the UV detector on the MALS system was not capable of going below 190<sub>nm</sub> wavelength.

### 3.5.1.2. *Protein Glycoconjugate Vaccine Components (MaIE, SpyCEP and SpyAD)*

In addition to polysaccharide antigens, the selection of an appropriate carrier protein should not be underestimated. The protein component actively participates in immunological processing which is critical for induction of high avidity antibodies, and immune memory particularly in infants and the elderly. A number of GAS specific carrier protein antigens have recently been investigated for attachment to GAC<sup>411,412,447</sup>, due to increased interest in improving vaccine immunogenicity<sup>472</sup>.

Prior to conjugation, this work selected protein antigens from GAS based on the literature demonstrating essential roles in pathogenesis and ability to generate immune responses. Selected proteins are present across the range of GAS infection stages, such as initial host adhesion through MaIE and SpyAD, to invasive disease for SpyCEP, with the aim that blocking such functions may stop colonisation and pathogenicity. The proteins vary in their characterisation with SpyCEP, followed by SpyAD being most characterised, currently included in experimental vaccines at various stages<sup>387,411,412</sup>, and MaIE as the most novel candidate not currently included in any vaccine despite identification as a potential target<sup>156,506</sup>.

Experimentally, the immunogenicity of purified recombinant protein antigens on their own was investigated using *in silico* tools, IVIG, and in a murine model, assessing the generation of IgG antibodies and antibody function. Although all protein carriers appear to induce some level of immunogenicity, SpyAD and SpyCEP are more immunogenic compared to MaIE. Such improved immune responses for SpyAD and SpyCEP raised in mice corroborates the *in silico* and IVIG ELISA data showing that both proteins have more predicted protective cell epitopes and higher IgG GMT compared to MaIE. As stated in the literature, immunogenic regions of SpyAD and SpyCEP are more defined, appearing to be localised to the NTD regions. These antigens produce more robust humoral responses, as well as induce similar levels of protection against GAS mouse challenge from a number of GAS isolates<sup>395</sup>. A previous study investigating SpyAD and SpyCEP as carrier proteins in GAC glycoconjugates and as antigens alone showed comparable anti-protein total IgG levels similar to this work<sup>411</sup>. However, as the vaccines were standardised to the polysaccharide component, the SpyCEP dose was in fact 2.7 times lower than SpyAD, suggesting SpyCEP is more immunogenic than SpyAD, similar to mouse data in this work. In comparison, MaIE has been found to be less immunogenic in this work as well as in other immunogenicity studies<sup>433</sup>. Some studies however have

shown that specific antibodies can be generated in mouse models<sup>156</sup>, indicating that although not fully apparent following the *in silico* analysis of the antigen structure, the protein in fact encodes accessible immunogenic epitopes. This, along with TT, determined to not be an antigen based on this analysis, despite inclusion in licenced vaccines and shown to be immunogenic experimentally<sup>534</sup>, demonstrates *in silico* tools cannot be solely relied upon for antigen selection.

Western blot analysis confirmed the presence of all antigens in their predicted cellular compartments, but high concentrations of antibodies were needed for the detection of MalE and SpyCEP, based on low anti-protein antibody titres and poor detection respectively. Cells were not treated with hyaluronidase, therefore capsule may have interfered with native protein detection in this instance. Additionally, in the case of SpyCEP antigen, protein specific antibodies are functional at reducing IL-8 cleavage, which may be important during infection to prevent immune evasion strategies. In the literature, functional antibodies that prevent IL-8 cleavage were demonstrated to confer protection against infection<sup>208,511</sup>. Conjugation chemistry will prove to be essential for retaining this functionality as “mesh-like” conjugates have previously been shown to prevent this functional immune response against SpyCEP despite good anti-protein IgG titres<sup>411</sup>.

### 3.5.2. EDC Carbodiimide Conjugation Chemistry Suitability

Preservation of polysaccharide, as well as carrier protein epitopes, are important in generating immunogenicity from glycoconjugate vaccines, particularly for ‘double hit’ vaccines. Further, GAC as a small and simple polysaccharide, means that any minor disruption of epitopes has the potential to have a major impact on immune system presentation, compared with larger and more complex polysaccharides<sup>535</sup>. For this reason, the conjugation process itself should be uncomplicated and mild, to limit masking or damage to protective epitopes on both components.

While there are several conjugation chemistries, only a few have been used in licensed vaccines. Reductive amination is the most broadly used conjugation method, requiring aldehyde groups on polysaccharide antigens and lysine residues on carrier proteins<sup>139,411</sup>. Despite having been used to conjugate GAC polysaccharides in previous studies, it was decided to not proceed with reductive amination in this work. This was due to observed batch to batch inconsistency with GAC containing reactions<sup>411</sup>, and the requirement of polysaccharide oxidation through sodium periodate treatment to render it more reactive, which can impact

epitope availability through ring-opening<sup>536</sup>. This work investigated suitability of EDC as a zero-length linker method to conjugate GAC variants and protein carriers.

EDC itself does not form part of the final conjugate product. This is beneficial when using small polysaccharides as it preserves the polymer structure and conformational epitopes. The EDC conjugation method requires carboxylic acids for linkage, with both GAC variants shown to contain such reactive groups through polymer biotinylation using EDC. This semi-selective method generates “sun-like” glycoconjugate structures, with polysaccharide polymers attached from their reducing ends, distinct from other preclinical studies using acid hydrolysis methods<sup>139,409</sup>. Enzymatically extracted GAC variants were predicted to have preserved polymer chains due to the extraction method using PlyC, enabling reducing end targeting.

Other groups investigating GAS specific glycoconjugate vaccines such as Kabanova *et al.*,<sup>139</sup> have suggested that native GAC extraction methods may lead to contaminants such as M protein or hyaluronic acid. They therefore investigated synthetic GAS core oligosaccharides, proposed as superior alternative antigen in terms of purity. To mitigate the risk of contaminants in preparations, in this study GAC extraction samples underwent thorough purification to remove contaminants, validated by NMR analysis to confirm the absence of these potential contaminants within preparations.

#### 3.5.2.1. EDC Carbodiimide Conjugation Procedure Optimisation

Initial attempts to achieve Male or SpyCEP conjugates had failed using the readily available one step EDC reaction protocols. Hence, some optimisation steps were taken to increase the chances of conjugation for all target proteins. EDC chemistry can result in two unfavourable reactions, inactivation of ester intermediates and hydrolysis of EDC itself prior to linkage<sup>524</sup>. The first step was to increase the amount of polysaccharide as much as possible to increase the reaction chances, mitigating unfavourable side reactions which can be difficult to limit, control, and experimentally observe. Hydrolysis and inactivation of O-acylisourea ester reaction intermediates occurs rapidly in aqueous solutions particularly when active carboxylates do not find the required amine group for conjugation. This is prevalent when target molecules for conjugation are in low concentrations compared to water within reaction buffers, resulting in isourea formation and regeneration of the carboxylic group, which cannot partake in the coupling reaction<sup>537</sup>. Higher polysaccharide concentrations can therefore reduce this occurrence so this was

tested with GAC variants, however this was limited by the polysaccharide yields which could be obtained from bacterial cultures (1 – 3 L), limiting the possibility to scale up reactions to obtain milligrams of polysaccharide.

The second step was to prevent the hydrolysis of EDC linking agent, which is prevalent in certain reaction conditions. The literature is consistent in the use of MES buffer for the initial carboxylic acid activation step using EDC, to stabilise the pH within the acidic region. At acidic pH, EDC reacts quickly, reducing active ester intermediate hydrolysis, therefore increasing active species availability<sup>537</sup>. However, at acidic pH hydrolysis of EDC can occur more rapidly, which often occurs when in excess compared to the carboxylic acid containing molecule<sup>537</sup>. Despite attempts at pH optimisation to balance EDC activation and stability using 100 mM MES buffer pH 5.5, limited conjugation between TT, MalE and SpyCEP with WT GAC was observed in the single step reaction (Appendix Table 7.3). This may be because such pH conditions are non-optimal for amide bond formation, observed to be between pH 6 and 7.5<sup>524</sup>.

Additionally for successful conjugation to occur both protein and polysaccharide components must be soluble and stable in reaction buffers at the required concentrations. Throughout this work protein precipitation was observed when added directly to 100 mM MES buffer pH 5.5, therefore reducing the amount of carrier protein available for covalent attachment. To overcome this, and improve amide bond formation, a two-step method was tested increasing the pH to 7.2 before carrier protein addition. However, some protein precipitation was still observed even at more neutral reaction pH, particularly for SpyCEP protein discussed below.

Finally, a defined two-step method was selected but with the introduction of sulfo-NHS stabiliser. This process first activated and stabilised the polysaccharide carboxylic acid, followed by optimising the pH required for amide bond formation on protein carriers. Sulfo-NHS generates sulfo-NHS esters, a stable intermediate, from a reaction between the GAC variants carboxylic acids and NHS in the presence of EDC. Generated intermediates contained hydrophilic active groups that couple rapidly with amine nucleophiles, attacking the electron deficient carbonyl of the active ester. This results in the sulfo-NHS group leaving to create a stable amide linkage with the protein carrier, without being part of the final glycoconjugate. Additionally, this method provided a means to remove excess EDC by quenching



with thiol containing 2-ME compound after the GAC activation step, to prevent protein – protein conjugation.

Care was taken when carrying out the two-step sulfo-NHS conjugation method as NHS ester intermediate stability is improved in acidic conditions with low reaction water content. NHS activated molecules must also be used promptly for the addition of amine containing molecules. Therefore, 100 mM MES pH 5.5 was continued to be used for activation of GAC variants incubated with EDC and sulfo-NHS, before immediate buffer exchange to sodium bicarbonate and PBS for amide bond formation. The latter therefore provided a more alkaline pH and salts to maintain protein stability in solution, considerably reducing protein precipitation for all reactions, successfully generating TT and SpyAD conjugates. Storage of chemical glycoconjugate reactions was also assessed in these buffer systems over a one month period (Appendix Figure 7.1).

However, for SpyCEP, although having a structure most similar to TT based on *in silico* analysis, and available surface exposed lysine residues for attachment successful conjugation remained elusive. It repeatedly precipitated in the conjugation reaction buffers tested, and it was prone to degradation (Appendix Figure 7.4b). Previous work developing successful SpyCEP containing glycoconjugates used the full sized protein<sup>411</sup> rather than a cloned subunit like the one used in this work and others<sup>208</sup>. Full sized SpyCEP will have contained more lysine residues available for conjugation. It would also have different physicochemical properties, which would likely have made it more stable in solution. MalE could also not be conjugated perhaps due to showing the lowest percentage of surface exposed lysine residues across all the proteins tested and smaller protein size. Additionally, differences in protein structure and therefore lysine availability may have been observed due to the more acidic conditions of EDC conjugation reactions.

Nevertheless, three glycoconjugate vaccines were successfully generated, which suggests that the suitability and optimisation of EDC reactions may be specific to each protein carrier. It is likely influenced by a combination of protein characteristics such as molecular weight, conformation linked to amine group exposure, as well as protein pI in specific pH reaction buffer conditions linked to stability. Specifically, lysine reactivity has been linked to both solvent and reagent accessibility and amino acid sequence<sup>538</sup>, sometimes leading to preferential conjugation to certain lysine residues which are positioned in more favourable protein or peptide regions<sup>498</sup>. It

has also been shown that at slightly basic pH, the lysine with the lowest pKa is the kinetically favoured residue for reactions<sup>539</sup>, therefore optimised buffer systems may have improved reactive group availability and thus conjugation success for unsuccessful protein candidates.

#### 3.5.2.2. *Confirmation of EDC Carbodiimide Conjugations*

Initially, conjugation reactions were analysed using western blotting analysis, probing for both GAC and protein carriers simultaneously as a quick and simple analysis. In all three purified glycoconjugate samples, there was little evidence of a size increase. However, western blots demonstrated that extracted and purified GAC variants do not appear to migrate at a defined size. This was perhaps due to the extraction process, and polysaccharide properties such as charge and structure, or potential polymer chain length heterogeneity within the samples. This, therefore, made the analysis of glycoconjugate size ambiguous; however, blotting was able to successfully demonstrate the separation of lower molecular weight species of unconjugated GAC from the glycoconjugate sample particularly for TT-GAC.

Therefore, to initially confirm successful conjugation before extensive physicochemical analysis, alternative methods not based on size separation were investigated. A capture ELISA was developed for both TT and SpyAD testing, successfully showing a signal with glycoconjugate samples, not observed in controls. However, this method purely confirms linkage of protein to the polysaccharide, and due to GAC signal amplification, it does not give any information regarding glycoconjugate physicochemical properties.

Further analysis was carried out to confirm conjugation and determine conjugates physicochemical properties such as size shift following conjugation. HPLC-SEC is the gold standard for confirming conjugation in the literature. Conjugate size and polydispersity determination is also important for understanding the observed immunogenicity of vaccines downstream. This work showed differences in run times for protein components alone (TT or SpyAD) compared with the proteins when conjugated to either WT GAC or GAS\_Rha using UV<sub>214nm</sub> detection. This suggests that conjugation of polysaccharide alters the molecules hydrodynamic radius and therefore interaction with the column and elution properties. Nevertheless, baseline separation was difficult to achieve due to poor resolution, despite selecting a porous column with a separation range of 1 - 450 kDa, suitable for low molecular weight species. Similarities in hydrodynamic radius between unconjugated and conjugated proteins, and the low molecular weight of polysaccharides (4.9 - 9.8 kDa) compared

to the large molecular weight of proteins (88 – 90 kDa) was likely the cause for difficulty in achieving distinct peaks. This may have been exacerbated by the presence of unconjugated protein in the glycoconjugate samples, resulting in broader peaks due to mixed protein populations, similar to the issues observed in SDS-PAGE separation and western blotting. Additionally, separation may have been influenced by the linear nature of polysaccharides, arranged around the protein in a “sun-like” structure due to reducing end conjugation. Larger “mesh-like” structures are likely to have more distinct elution times than these small linear polymer structures<sup>411</sup>. This concept may explain the small differences observed between unconjugated and conjugated protein, particularly for TT-GAS\_Rha containing shorter GAS\_Rha polymers compared to WT GAC within SpyAD-GAC glycoconjugate.

Using HPLC-SEC methods TT-GAC did not show any distinct size differences compared to TT protein carrier alone (Appendix Figure 7.6a). Rather for this conjugate, changes in physicochemical properties such as protein hydrophobicity and polarity were investigated using reverse-phase chromatography (RP-HPLC). This approach was selected as TT has hydrophobic surfaces that facilitate RP-HPLC<sup>540,541</sup>, a gold standard technique used for protein / peptides separation<sup>542</sup>. TT-GAC showed different hydrophobicity than TT alone suggesting that conjugation alters the overall TT structure, influencing how it interacts and elutes from the C-18 column using an ACN gradient. However, the TT-GAC sample failed to show a peak representing unconjugated TT predicted to be within the sample preparation. This may be due to lack of column resolution to distinguish the two protein populations, especially given more TT is conjugated than unconjugated. Though interesting, showing changes to protein physicochemical properties after conjugation, quantitative glycoconjugate sizing was not possible due to low resolution, resulting in small differences in retention times. Future analysis may therefore benefit from the use of more sensitive markers, or rather condition optimisation to achieve better peak resolution through different mobile and stationary phases to suit protein physicochemical properties.

### 3.5.3. Determination of Polysaccharide : Protein Ratio in Glycoconjugates

An important consideration in the optimisation of chemical conjugation reaction is to limit the percentage of unconjugated components. It is essential to quantify and remove these components where possible as evidence suggests that these can negatively influence immune response stimulation<sup>543,544</sup>. There are several approaches which can be used to experimentally quantify unconjugated

components in order to determine the ratio of conjugated protein to conjugated polysaccharide within a purified glycoconjugate.

The simplest and quickest methods to quantify components in a purified sample are colourimetric assays (e.g. anthrone and BCA). This work also tested H-PAD as an alternative chromogenic method for WT GAC quantification. For the most part, H-PAD and Anthrone polysaccharide quantification for WT GAC containing glycoconjugate samples were shown to be fairly consistent with each other, both used in calculations for weight / weight (*w/w*) and molar (*m/m*) calculations. However, GAS\_Rha detection and quantification by H-PAD proved more problematic, with minimal signal observed for purified GAS\_Rha polysaccharide only samples, as well as conjugate sample TT-GAS\_Rha (data not shown). GAS\_Rha was reactive by anthrone, similar to a recent study quantifying polysaccharide content of rhamnose specific glycoconjugate vaccines<sup>412</sup>. Final optical density readings however were found to be lower when using rhamnose only standards compared to a rhamnose and GlcNAc mixture for WT GAC quantification, suggesting rhamnose is less sensitive to reacting with anthrone reagent than GlcNAc. An adjusted standard concentration curve was required, therefore lowering the sensitivity of the assay for GAS\_Rha quantification in glycoconjugate samples.

For determining the ratio of conjugated protein to conjugated polysaccharide immunoprecipitation of TT-GAC and TT-GAS\_Rha glycoconjugates was necessary. Samples showed a similar percentage of conjugated TT but in the context of ratios to polysaccharide, the TT-GAC reaction appeared to be more efficient with higher GAC decoration compared to TT-GAS\_Rha. Such molar calculations consider the molecular weight of both protein, and polysaccharide from SEC-MALS analysis of pre-conjugation material. The apparent lower level of decoration of TT with GAS\_Rha compared to WT GAC may be explained by the unique characteristics of GAS\_Rha compared to WT GAC, such as GAS\_Rha showing lower yields following extraction and purification, due to suspected reduced PlyC activity on cell wall peptidoglycan digestion (Appendix Table 7.1). This would have hindered carboxylic acid availability for attachment, a theory backed up by reduced biotin signal for GAS\_Rha compared to WT GAC. The calculations also assumed that all free polysaccharide was removed by spin purification, however a more accurate ratio would be achieved if the unconjugated polysaccharide was tested for conjugates, but this was not possible with the material and time constraints of this project.

### 3.6. Summary

Based on the objectives, three GAS protein antigens, specifically MalE, SpyAD, and SpyCEP were selected and deemed appropriate to investigate as carriers. These were subsequently cloned, expressed, and purified, along with monomeric TT for comparison purposes. Additionally, a GAS *gacI* mutant strain was developed producing polysaccharides which could be enzymatically extracted and purified along with WT GAC from GAS cells.

Both extracted and purified glycoconjugate vaccine components were subsequently shown to contain reactive compatible groups for EDC chemical conjugation reactions. Reactions were optimised to achieve ideal conditions for conjugation by modifying component ratios, reaction pH, and the inclusion of a stabiliser. Successful conjugation was determined for three reaction combinations, specifically TT-GAC, SpyAD-GAC and TT-GAS\_Rha. Quantification of conjugation components were calculated for each glycoconjugate, and the ratio of conjugated protein to conjugated polysaccharide was calculated in some instances. Glycoconjugate vaccines were also separated based on size, showing differences in retention times, as well as differences in physiochemical properties compared to protein components alone. In short, this Chapter describes the manufacture of novel GAS specific glycoconjugate vaccines using an optimised chemical conjugation approach. Such vaccines will be compared to glycoconjugates generated using biological conjugation for antibody generation and function.

## CHAPTER 4

# Use of Bioconjugation to Manufacture GAS Glycoconjugate Vaccines

### 4.1. Introduction

Despite extensive research there is currently no effective licensed GAS vaccine<sup>545</sup>, therefore alternative vaccine design approaches are imperative. Traditionally, glycoconjugates composed of extracted polysaccharide and recombinantly produced protein are conjugated chemically in a complex and costly multistep process described in Chapter 3. Chemical conjugation has limitations such as technical challenges, low product yields, and batch to batch variation<sup>546</sup>. Therefore, in recent years the development of Protein Glycan Coupling Technology (PGCT) has provided an alternative and in some cases superior method to chemical conjugation<sup>547</sup>.

PGCT has the potential to simplify glycoconjugate production, and has since been adopted as a glycoconjugate manufacturing method by a number of pharmaceutical companies, investigating the use of this technology in vaccine clinical trials<sup>548–550</sup>. PGCT exploits and manipulates bacterial polysaccharide biosynthesis pathways, including the O-antigen, capsular polysaccharide, protein linked glycosylation pathways, as well as conjugation to lipids. The term bioconjugation covers a number of related PGCT glycosylation systems, with two characterised by the amino acid within specific protein acceptor sequons for polysaccharide attachment, namely asparagine (*N*-linked) or serine/threonine (*O*-linked) glycosylation<sup>551,552</sup>. *N*-linked glycosylation, specifically using oligosaccharyltransferase (OST) PglB (protein glycosylation B) from *Campylobacter jejuni*, and is arguably the most well characterised bioconjugation approach<sup>552</sup>. This will form the basis of the work in this Chapter describing glycoengineering biotechnology in the workhorse bacterium *Escherichia coli* aiming to develop GAS specific glycoconjugate vaccines.

#### 4.1.1. Discovery of Protein Glycan Coupling Technology (PGCT)

PglB specific *N*-linked glycosylation machinery is encoded by the *pgl* operon, first discovered within the *C. jejuni* (strain NCTC 11168) genome in early 2000<sup>553</sup>, independent of the lipopolysaccharide (LPS) and flagella *O*-linked glycosylation loci<sup>554</sup>. The gene cluster is responsible for glycosylation of proteins with the *C. jejuni* heptasaccharide<sup>480</sup> (Figure 4.1). Genes for heptasaccharide synthesis are encoded

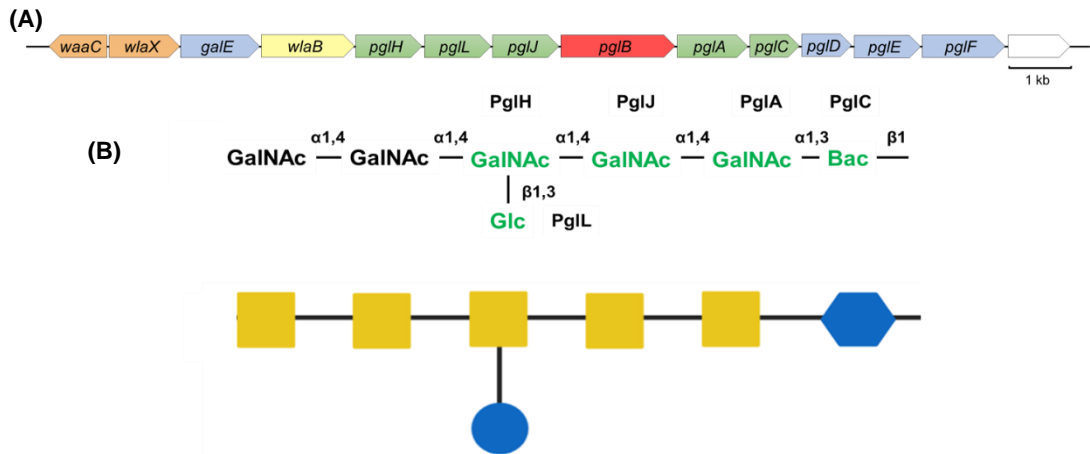
along with the *pglB* gene encoding the OST enzyme pivotal to bioconjugation function, transferring fully assembled polysaccharides from lipid linkers to protein substrates<sup>555</sup>.

Successful cloning and expression of the genetic loci required for functional reconstitution and biosynthesis of *C. jejuni* heptasaccharide within *E. coli* cells has opened up the possibility of conjugating recombinant polysaccharides to proteins for both research and industrial applications<sup>550,553,556,557</sup>. Such *E. coli* cell expression systems have since been used as “vaccine factories”, exploited at both the basic research level, as well as in product development<sup>548,549</sup>. PGCT is therefore becoming a feasible alternative to chemical conjugation<sup>558</sup>, with benefits such as *E. coli* systems producing inexhaustible fully synthesised recombinant polysaccharide sources, and readily purified glycoproteins at reduced costs and improved yields<sup>559</sup>. Currently the leading bioconjugate vaccine candidates manufactured with this technology are multivalent *Shigella* bioconjugate vaccines which have proceeded successfully to phase 1 (e.g. *S. dysenteriae* or *S. flexneri 2a*, ClinicalTrials.gov: NCT01069471 or NCT02388009 respectively)<sup>548,550</sup>, as well as phase 2b clinical trials for *S. flexneri* (ClinicalTrials.gov: NCT02646371)<sup>560</sup>, in addition to a tetravalent Extraintestinal Pathogenic *E. coli* (ExPEC) bioconjugate vaccine in phase 2 (ClinicalTrials.gov: NCT02546960)<sup>561</sup>.

#### 4.1.2. *C. jejuni* pgl Operon Biosynthesis Mechanism and Expression

Expression of the 11 genes assembles and transports the *C. jejuni* heptasaccharide polysaccharide<sup>553,557</sup> (Figure 4.1) to the periplasm for attachment onto a number of native protein substrates using PglB OST<sup>555</sup> by *N*-linked glycosylation<sup>555,562,563</sup>. The heptasaccharide is composed of GalNAc $\alpha$ 1,4-GalNAc $\alpha$ 1,4-[Glc $\beta$ 1,3-]GalNAc $\alpha$ 1,4-GalNAc $\alpha$ 1,4-GalNAc $\alpha$ 1,3-Bac $\beta$ 1, where GalNAc is N-Acetyl-Galactosamine, Glc is glucose, and Bac is bacillosamine. The heptasaccharide added to such protein substrates, is first synthesised onto an undecaprenol-pyrophosphate (Und-PP) lipid linker attached to the cytoplasmic leaflet of the inner membrane. A flippase enzyme is then responsible for translocation of the fully synthesised polysaccharide across to the periplasm, where it can be subsequently recognised by PglB and attached onto a recognition sequon on a given acceptor protein<sup>558,564</sup>. The protein recognition sequon for *N*-linked glycosylation is D/E-X-N-X-S/T, where X represents any amino acid except proline, and positive D/E and S/T amino acids are at the  $-/+ 2$  positions, pivotal in locating asparagine (N) as the acceptor amino acid<sup>565</sup>. Bacterial *N*-linked glycosylation occurs primarily after protein folding but PglB can also act co-translationally<sup>566,567</sup>. Protein glycotags must be accessible to PglB, and if not natively

present should be engineered into flexible open regions<sup>568</sup>, such as the N- and C-termini<sup>567</sup>, extended out from the proteins native conformation and structure.



**Figure 4.1: The *pgl* locus in *Campylobacter jejuni* (NCTC 11168) produces a heptasaccharide for bioconjugation onto carrier proteins by PglB OST enzyme.**

(A) Schematic representation of the *pgl* genetic locus in *C. jejuni*. Horizontal arrows represent each gene designation, with colour denoting predicted gene function. Green, glycosyltransferase; blue, sugar biosynthesis; red, oligosaccharyltransferase; yellow, ABC transporter, and orange LPS transferases.

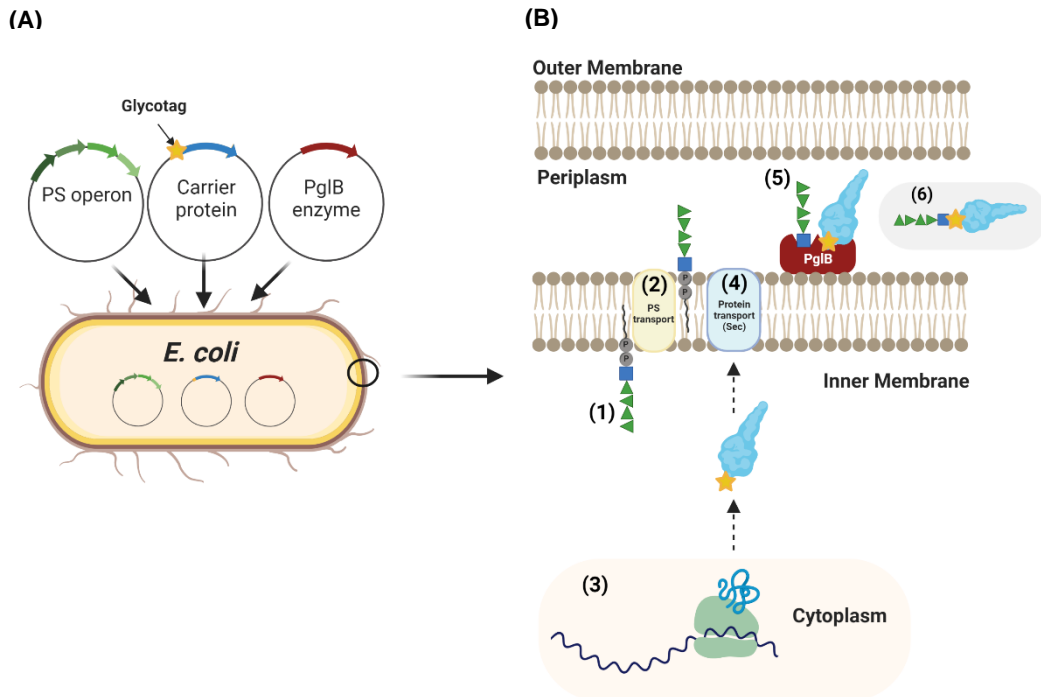
(B) Polysaccharide product from the resulting *pgl* encoded proteins with monomer linkages, and a visual representation based on nomenclature for glycans (SNFG)<sup>703,704</sup>. Square = GalNAc, circle = glucose, hexagon = bacillosamine. Green text refers to the products of genes in Figure 4.1a highlighted in green.

Original figure adapted from Linton *et al.*<sup>480</sup>. Sequence data assembled from Y11648 and AJ131360 acquisition numbers in NCTC 11168 strain.



#### 4.1.3. Glycoconjugate Production using *C. jejuni* PglB Bioconjugation Inside *E. coli* Cells

Successful *N*-linked glycosylation of proteins inside *E. coli* glycoengineering host cells involves recombinant expression of three components, a polysaccharide biosynthesis locus, an acceptor protein with OST specific acceptor motif/s, and an OST enzyme source. These components are then targeted to the periplasmic face of the inner membrane where conjugation occurs (Figure 4.2).



**Figure 4.2: Glycoconjugate production using *C. jejuni* PglB bioconjugation inside *E. coli* host cells.**

**(A)** Plasmids containing genes encoding polysaccharide (PS) biosynthesis (green), the protein acceptor (blue) containing a glycotag (yellow star), and the oligosaccharyltransferase (OST) PglB (red) are co-transformed into *E. coli* host cells.

**(B)** Bioconjugates are produced as follows; **(1)** The polysaccharide biosynthesis locus is expressed and built onto undecaprenol-pyrophosphate (Und-PP) lipid linkers within the inner membrane. **(2)** The polysaccharide is flipped from the cytoplasm to the periplasm by a specific flippase enzyme. **(3)** The carrier protein is expressed and synthesised in the cytoplasm. **(4)** Synthesised carrier proteins are exported to the periplasm through the Sec secretion system. **(5)** In the periplasm both the polysaccharide and the carrier protein containing a specific glycotag can be recognised by PglB OST enzyme. PglB transfers the polysaccharide from Und-PP onto the asparagine residue within the glycotag D/E-X-N-X-S/T motif on the fully folded carrier protein, resulting in protein glycosylation. **(6)** An inexhaustible supply of glycoproteins can be subsequently purified from *E. coli* cells.

Original figure created with BioRender.com based on Cuccui *et al.*<sup>705</sup>.

Wacker *et al.*<sup>569</sup>, first demonstrated expression and conjugation of the *C. jejuni* heptasaccharide to AcrA, a native *C. jejuni* periplasmic lipoprotein, ~ 40 kDa in size with two naturally present asparagine residues for PglB targeting using a glycoengineering *E. coli* host. This proof of concept study showed successful glycosylation of AcrA protein following transfer of the *pgl* genetic locus into *E. coli*, a bacterium that does not natively possess any *N*-linked glycosylation machinery<sup>569</sup>. Such findings opened the possibility of *E. coli* cells expressing PglB to be used to transfer a diverse range of polysaccharide substrates to a variety of acceptor proteins to generate glycoproteins.

As the bioconjugation field has developed, several recombinant polysaccharide and protein carrier combinations have been tested for different pathogens. A 'one pot' *in vitro* assay has been devised as a simplified alternative bioconjugation testing approach. This method, originally developed by Jaroentomeechai *et al.*,<sup>570</sup> is a novel cell-free glycoprotein synthesis approach. In short it works by obtaining cell extracts which are selectively enriched with glycosylation components including the oligosaccharyltransferase enzyme as well as lipid linked polysaccharide and recombinant protein. These are then co-incubated in single tube where the linking reaction occurs outside of the cell, offering benefits such as more control over relative concentrations and overall reaction times. Both the traditional cell-based (Figure 4.2) and 'one pot' *in vitro* methods will be assessed for their relative merits towards the generation of GAS specific bioconjugates.

#### 4.1.4. Recombinant Polyrhamnose Expression in *E. coli* Based on Group A Carbohydrate (GAC) and Homologs in *Streptococcus mutans*

Cell surface Group A Carbohydrate (GAC) consists of an  $\alpha$ -L-(1 $\rightarrow$ 2) and  $\alpha$ -L-(1 $\rightarrow$ 3) linked rhamnose backbone structure, which is shared by the rhamnose-glucose polysaccharide (RGP) in *Streptococcus mutans*<sup>571</sup>. The major difference between the polysaccharide expressed in *S. pyogenes* (GAC) to *S. mutans* (RGP) are their side chain modifications, with GAC decorated with GlcNAc and RGP decorated with glucose. *S. mutans* *rgp* genes encoding RGP share > 50 % identity with *S. pyogenes* homologs in the *gac* operon<sup>571</sup>. Previously, the *S. mutans* rhamnose backbone has been successfully synthesised through *rgpA* – *rgpF* (except for *rgpE*) gene expression<sup>572</sup>, and transported to the *E. coli* cell surface, altering LPS profiles<sup>571</sup>.

Shibata *et al.*<sup>571</sup>, successfully showed similarity between native rhamnose biosynthesis pathways in *S. mutans* and the reconstituted pathways expressed and

synthesised in *E. coli* host cells. Both pathways required genes to first produce a lipid linked reducing end, being the monosaccharide directly linked to the lipid carrier, specifically a GlcNAc monosaccharide, followed by activated rhamnose monosaccharides. Some of these genes, for example those synthesising GlcNAc and rhamnose sugars, are readily available in *E. coli* in the form of *wecA* and *rmID* genes respectively. In the absence of the *S. mutans wecA* homolog (*rgpG*), the *E. coli wecA* gene is essential to build lipid linked GlcNAc to form the reducing end at the start of the polysaccharide polymer<sup>571,573</sup>. Rhamnose polymers are then built onto the GlcNAc reducing end by *rgpA* the first rhamnosyltransferase, followed by *rgpB* and *rgpF* as the second and third rhamnosyltransferases. Rhamnosyl polymers are then transported to the periplasmic space by essential genes *rgpC* and *rgpD* forming an ATP dependent ABC transporter, known as a flippase<sup>571,572</sup>.

These studies provide evidence that *Streptococcal* rhamnose polymers can be expressed in *E. coli*, indicating great promise for successful reconstitution of the related *S. pyogenes* GAC biosynthesis pathway for application in PGCT. This will be discussed in the context of investigating the generation of GAS specific bioconjugates in this Chapter.

## 4.2. Aim, Hypothesis and Objectives

The aim of this work was to adopt the principles of PGCT to develop a GAS specific bioconjugate, providing a novel mechanism for glycoconjugate vaccine biosynthesis. These will be compared to traditional methods of chemical conjugation described in Chapter 3.

This work explores the substrate specificity and coupling potential of PglB for *E. coli* derived rhamnose polymers with various acceptor proteins. Additionally, this work investigates assembly of a single *E. coli* host strain, containing chromosomal encoded PglB OST, and plasmids encoding genes required for recombinant rhamnose biosynthesis and transport, as well as model and GAS specific carrier proteins.

**Aim:** Investigate *N*-linked glycosylation, specifically PglB OST capability, to glycosylate characterised (AcrA and NanA) and novel GAS specific (MalE, SpyCEP, and SpyAD) carrier proteins with *E. coli* produced recombinant rhamnose variants and GAS produced GAC.

### **Hypothesis:**

- Bioconjugation will provide a viable alternative approach to chemical conjugation for manufacturing of GAS specific glycoconjugate vaccine candidates.

### **Objectives:**

- 1) Express recombinant rhamnose variants (rpWTRha, rpMRha, and rpORha) in conjunction with collaborator Dr. Dorfmueller, University of Dundee.
- 2) Design, clone, and express GAS specific carrier proteins with PglB compatible terminal glycotags suitable for both glycosylation approaches.
- 3) Investigate the 'one pot' *in vitro* approach to conjugate model carrier protein (AcrA) and novel GAS specific carrier proteins with *E. coli* produced recombinant rhamnose.
- 4) Investigate cell-based glycosylation in CLM24cedA::pglB *E. coli* cells to conjugate streptococcal specific carrier protein (NanA) and novel GAS specific carrier proteins with *E. coli* produced recombinant rhamnose (rpORha) using chromosomal encoded PglB OST enzyme.

## 4.3. Methods

### 4.3.1. Bacterial Strains

*S. pyogenes* (GAS) and *E. coli* strains used in this work are listed in Table 4.1. A DH5 $\alpha$  derivative NEB5 $\alpha$  *E. coli* (NEB) and DH10 $\beta$  *E. coli* (NEB) strains were used for plasmid cloning and propagation (Chapter 4 section 4.3.3). BL21 (DE3) *E. coli* was used for protein expression testing (Chapter 4 section 4.3.7). Glycoengineering CLM24 *E. coli* strain was primarily used for 'one pot' *in vitro* glycosylation studies (Chapter 4 sections 4.3.8 – 4.3.9). This strain is a K12 W3110 variant with inactivated  $\Delta waaL$ <sup>558</sup> and  $\Delta wbbL$ <sup>574</sup> genes. In some instances, for example assessing recombinant rhamnose expression, the CS2775 *E. coli* strain was used with and without the *rfaS* gene (Chapter 4 section 4.3.6). CS2775 *E. coli* strain lacks endogenous rhamnose modifications on the surface LPS<sup>571,575–577</sup>. In cell-based glycosylation studies, CLM24*cedA::pglB* was used, denoting CLM24 as a W3110-derivative with *pglB* gene integrated into *cedA* on the chromosome<sup>578,579</sup> (Chapter 4 section 4.3.10).

GAS M1 strain NCTC - 8198 and  $\Delta gacI$ , a derivative of NCTC - 8198 strain with an in-frame deletion of *gacI* producing polyrhamnose (Chapter 3 section 3.3.5) were used for lipid linked polysaccharide extraction for use in 'one pot' *in vitro* glycosylation studies to assess PglB OST substrate specificity (Chapter 4 section 4.3.9).

**Table 4.1: List of *E. coli* and GAS strains used in bioconjugation work.**

Strains	Description	Source
<b>Commercial <i>E. coli</i></b>		
<b>NEB5<math>\alpha</math> Stable Competent</b>	Commercially available DH5 derivative strain, chemically competent for plasmid propagation.	NEB
<b>DH10<math>\beta</math></b>	Commercially available DH10 derivative strain, chemically competent for high quality plasmid preparations.	NEB
<b>BL21 (DE3)</b>	Commercially available DE3 derivative strain, competent for protein expression. F-ompT hsdSB(rB - mB-) gal dcm(DE3) phenotype.	NEB
<b>Glycoengineering <i>E. coli</i></b>		
<b>W3110</b>	K12 strain - three mutations $\lambda$ -, IN(rrnD-rrnE)1 and rph-1.	580
<b>CLM24</b>	K12 W3110 derivative with inactivated <i>waaL</i> ligase and reduced LPS toxicity due to inactivated <i>lpxM</i> .	558
<b>CLM24<i>cedA::pglB</i></b>	K12 CLM24-derivative, $\Delta$ <i>waaL</i> with genomic integration of <i>C. jejuni</i> PglB in <i>cedA</i> on the chromosome.	578,579
<b>CS2775 (ligase negative)</b>	CS2767 derivative with <i>rfaS</i> knocked out with Kan resistance cassette to reduce surface rhamnose. Inactivated <i>waaL</i> ligase.	581
<b>CS2775 (ligase positive)</b>	CS2767 derivative with <i>rfaS</i> knocked out with Kan resistance cassette to reduce surface rhamnose.	151
<b><i>Streptococcus pyogenes</i></b>		
<b>NCTC - 8198</b>	<i>Streptococcus pyogenes</i> (GAS) strain 8198 from NCTC.	Bacterial culture collections, Public Health England (PHE)
<b><math>\Delta</math><i>gacI</i> NCTC - 8198 Mutant</b>	<i>Streptococcus pyogenes</i> (GAS) strain 8198 from NCTC with <i>gacI</i> gene deletion.	This study – Chapter 3

#### 4.3.2. Bacterial Growth Conditions

Commercial *E. coli* strains were grown in LB Lennox (5 g/L yeast extract, 10 g/L tryptone, 5 g/L NaCl), and glycoengineering *E. coli* strains were grown in a variety of media including; SSOB (5 g/L yeast extract, 20 g/L tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, 2YT (10 g/L yeast extract, 16 g/L tryptone, 5 g/L NaCl, pH 7.2), 2YTPG (10 g/L yeast extract, 16 g/L tryptone, 5 g/L NaCl, 7 g/L K<sub>2</sub>HPO<sub>4</sub>, 3g/L KH<sub>2</sub>PO<sub>4</sub>, 18 g/L glucose, pH 7.2), and Brain Heart Infusion (BHI) (5 g/L beef heart infusion, 12.5 g/L calf brains, 2.5 g/L disodium hydrogen phosphate). Media was supplemented with appropriate antibiotics for transformed constructs at concentrations stated in Table 4.2.

**Table 4.2: List of antibiotics used in transformed *E. coli* strains.**

Antibiotic	Final Concentration <i>E. coli</i> (µg/ml)	Plasmid Selection
Ampicillin	25	pHD0677 / pEXT20
Chloramphenicol	25	pACYC
Erythromycin	300	pHD0249/ pHD0689 / pHD0256
Kanamycin	50	<i>cedA::pglB</i>
Spectinomycin	250	pEXT21

#### 4.3.3. Cloning of GAS antigens in Commercial *E. coli* Strains

GAS antigens MalE, SpyCEP and SpyAD DNA gene sequences were synthesised as gBlocks® Gene Fragments (IDT). Gene sequences were engineered to contain NheI (GCTAGC) and NdeI (CATATG) restriction sites at the 5' and 3' termini for restriction ligation into pEC415 construct or XbaI (TCTAGA) and NcoI (CCATGG) for ligation into pEXT21 construct for 'one pot' and cell glycosylation methods respectively. Ligations were performed using T3 DNA ligase in T4 DNA ligase buffer (NEB) (general methods section 2.2.6) and the mixture transformed into chemically competent cloning NEB5α *E. coli* (NEB) or DH10β *E. coli* (NEB) via heat shock methods (general methods section 2.2.7) (Table 4.1). Potential insert positive clones were screened using colony PCR (cPCR) (general methods section 2.2.2) using primers flanking the pEC415 or pEXT21 MCS region, as well as protein sequence specific primers (general methods section Table 2.3).

#### 4.3.4. Production of Electrocompetent Glycoengineering *E. coli* Strains for Polysaccharide Expression Testing and 'One Pot' *In Vitro* Testing

Overnight *E. coli* host strain (CLM24 and CS2775) cultures were made electrocompetent by appropriate dilution into 10 ml LB broth to an OD<sub>600nm</sub> of 0.05, supplemented with appropriate antibiotic (Tables 4.1 – 4.2), shaking at 180 rpm at 37 °C. When cells reached mid-log phase at an OD<sub>600nm</sub> of 0.4 – 0.8, cultures were chilled on ice for 15 minutes before cells pelleted through centrifugation at 4,000 x *g* for 10 minutes at 4 °C. Supernatants were discarded and cells resuspended in 10 ml of pre-chilled 10% sterile glycerol and centrifuged again at 4,000 x *g* for 10 minutes at 4 °C, repeated three times. Supernatants were discarded each time and finally cells were re-suspended in 100 - 200 µl of 10 % sterile glycerol, often left on ice and used immediately or stored at - 80 °C in 50 µl aliquots.

#### 4.3.5. Production of Multi-plasmid PGCT Systems via Electroporation Transformation of CLM24cedA::pglB *E. coli* Cells

Electrocompetent CLM24cedA::pglB *E. coli* cells were used immediately, kept on ice, and mixed with an appropriate amount of plasmid DNA (50 – 500 ng/µl). Purified plasmids were extracted from commercial cloning NEB5α *E. coli* (NEB) or DH10β *E. coli* (NEB) strains, obtained via Monarch plasmid Miniprep Kits (NEB) or QIAprep Spin Miniprep Kits (Qiagen), selected based on kit availability (general methods section 2.2.2).

To produce strains containing all three necessary components for glycoconjugate production, CLM24cedA::pglB *E. coli* cells were first transformed with recombinant rhamnase constructs (pHD0249 alone, or pHD0677 followed by pHD0689 or pHD0256 for rpWTRha, rpMRha, and rpORha respectively) (Figure 4.5). Chilled electrocompetent CLM24cedA::pglB *E. coli* cells were pipetted into 0.2 mm electroporation cuvettes (Bio-Rad Laboratories Inc.). Plasmid DNA was added to the cells and electroporated at 2.5 kV in a Gene Pulser Xcell (Bio-Rad Laboratories Inc.) electroporator before recovery in 950 µl Super Optimal Broth (SOC) (0.5 % yeast extract, 2 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgSO<sub>4</sub>, 20 mM glucose) for 1 hour at 37 °C, 200 rpm shaking. Following recovery, a suitable dilution and volume of culture was spread onto LB agar plates containing appropriate antibiotic for recombinant rhamnase construct selection (Table 4.2), or alternatively if poor transformation efficiency was expected, cells were pelleted at 13,000 x *g* for one minute, and the total cell pellet was resuspended in a reduced volume of SOC before plating without dilution. Potential insert positive clones were



screened using cPCR (general methods section 2.2.2) with primers internal to the *gac* locus for rpWTRha and rpORha selection (general methods Table 2.3).

Confirmed CLM24*cedA::pglB* *E. coli* strains containing recombinant rhamnose constructs (rpWTRha, rpMRha or rpORha) were made electrocompetent for transformation with mid-copy number plasmid pEXT21 encoding *nanA* or GAS antigens (*malE* or *spyAD*) in the same manner. During SOC media recovery, ampicillin and erythromycin were added as appropriate to maintain rhamnose encoding constructs during outgrowth (Table 4.2). Spectinomycin in addition to ampicillin and erythromycin antibiotics were then used to isolate protein carrier positive transformants on LB agar plates. Potential insert positive clones were screened using cPCR using primers flanking the pEXT21 MCS region, as well as protein sequence specific primers (general methods Table 2.3). Expression tests were undertaken to determine inclusion and expression of the carrier proteins in addition to polyrhamnose structures (Chapter 4 section 4.3.10).

#### 4.3.6. Optimisation of Recombinant Rhamnose Expression

Glycoengineering CLM24 in addition to CS2775 *E. coli* cells were used to test for expression of recombinant polyrhamnose structures. One colony was used to inoculate 5 ml of various media (LB, SSOB, 2YT, 2YTPG) grown at 37 °C 180 rpm in the presence of ampicillin and / or erythromycin overnight (Table 4.2). Overnight cultures were used to inoculate 10 ml of fresh media to an OD<sub>600nm</sub> of 0.05 and grown at 37 °C 180 rpm until late-log phase (OD<sub>600nm</sub> = 0.6 – 0.8) and routinely induced with 0.2 % L-arabinose (BioBasic Canada Inc.) for strains containing pHD0677 ('acceptor stem' construct). Following induction, cells were either left at 37 °C, or dropped to 30 °C at 200 rpm. This occurred in the case of suspected cell toxicity, complimented with spiking of cultures with 1% glucose (Sigma-Aldrich™) at the point of subculture and again at the point of induction to improve growth. Cultures were incubated for 3 hours and overnight (~ 18 hours) post induction before the cells were pelleted by a spin at 13,000 x *g* for 10 minutes. Cells containing polyrhamnose encoding constructs pHD0249 or pHD0689 required no inducer. Cell pellets were then resuspended in a volume of PBS equating to an OD<sub>600nm</sub> of 10, and loaded into Lysing Matrix B – 2 ml Bead Beating Tubes (MP Biomedicals) before ribolysing in a FastPrep-24 Classic Instrument (MP Biomedicals) on programme 1 (2 m/s for 30 seconds) for 5 cycles. Resulting ribolyser lysates, termed whole cell lysates (WCL) were centrifuged at 13,000 x *g* for 10 minutes and the supernatants removed before analysis or storage at - 20 °C.

Alternatively, when investigating periplasmic location of recombinant rhamnose variants, cell pellets were resuspended to an OD<sub>600nm</sub> of 10 in periplasmic lysis buffer (30 mM Tris-HCl, 20% sucrose, 1 mM EDTA, and 1 mg/ml lysozyme, pH 7.4) and left for 1 hour rotating at 4 °C, before a spin at 13,000 x g for 10 minutes to separate cellular compartments. The supernatant contains the periplasm, whilst the cytoplasm, membrane and inclusion bodies are present in the pellet<sup>558</sup>. The periplasm was removed and aliquoted into a fresh tube and the pelleted fraction was resuspended in the same volume of PBS buffer supplemented with 1 % SDS (Sigma-Aldrich™). The resuspended pellets were then treated to 3 rounds of snap freeze-thaw on dry ice with added ethanol, and centrifuged at 13,000 x g for 10 minutes obtaining the cytoplasmic fraction in the supernatant from the pelleted fraction containing membrane portions and inclusion bodies.

For CS2775 *E. coli* cells with an intact WaaL phenotype, containing recombinant rhamnose exported to the LPS surface, an acid lipid A extraction was performed. Pelleted whole *E. coli* cells were heated in 2 % acetic acid (Sigma-Aldrich™) at 100 °C in a heat block for 1.5 hours. Treated cells were spun for 10 minutes at 13,000 x g to separate the recombinant rhamnose attached to the O-antigen chain in the supernatant from the lipid A attached to the bacterial membranes pelleted by centrifugation. The pH of the supernatant was increased from 3.5 to 8 by 4 M NaOH solution before application to a 5 ml HiTrap™ Capto™ adhere column (Cytiva), attached to an ÄKTA Start protein purification system (Cytiva). The chromatography system was equilibrated with 5x column volumes of loading buffer (5 mM sodium phosphate, 20% EtOH, pH 8.0) and pump B equilibrated with elution buffer (5mM sodium phosphate, 20% EtOH, 200 mM NaCl, pH 8.0). Recombinant rhamnose samples were applied using the sample pump before the column was washed with 5x column volumes of wash buffer to remove unbound material. Bound recombinant rhamnose was eluted from the column with 100 % elution buffer containing salt eluent. Based on UV<sub>280nm</sub> chromatographs the resulting elution fractions containing purified recombinant rhamnose were collected.

Optical density normalised WCL, periplasmic and cytoplasmic fractions, in addition to LPS extracted polysaccharide samples were routinely dotted onto nitrocellulose membranes (general method section 2.4.1) or separated by SDS-PAGE and analysed by western blot (general method section 2.4.1), to enable semi-quantitative comparisons between samples.

#### 4.3.7. Optimisation of GAS Antigen Protein Expression

For recombinant GAS antigen expression, transformed commercial BL21 (DE3) *E. coli* cells, and glycoengineering *E. coli* strains (CLM24 and CLM24*cedA::pglB*) were grown in either LB or BHI supplemented with appropriate antibiotics (Table 4.2) induced with 0.2% arabinose (BioBasic Canada Inc.) for pEC415 encoded proteins, or 1 mM Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) (Roche) for pEXT21 encoded proteins both at mid-log phase ( $OD_{600nm} = 0.4 - 0.6$ ). One ml of culture was removed 3 hours post induction, and the remainder left overnight to grow to stationary phase (~ 18 hours). Cells were harvested at 13,000 x g for 10 minutes and pellets analysed immediately, or stored at - 20 °C. Similar to cells expressing recombinant rhamnose variants, cells expressing protein antigens were treated with periplasmic lysis buffer normalised to an  $OD_{600nm}$  of 10 for 1 hour at 4 °C (Chapter 4 section 4.3.6). Protein expression and localisation in the periplasm was detected using SDS-PAGE and western blots (general methods sections 2.4.1), normalised by optical density enabling semi-quantitative comparisons.

#### 4.3.8. Preparation of Cell Free Lysate Extracts for Inclusion in 'One Pot' *In Vitro* Glycosylation Reactions

The 'one pot' *in vitro* reaction requires cell free lysates from each of the three components needed for glycosylation; the lipid linked polysaccharide, the carrier protein, and the PglB OST. Separate bacterial cultures of transformed CLM24 *E. coli* carrying the required genetic information for protein carriers AcrA encoded on pEXT20 and GAS antigens MalE, SpyAD and SpyCEP encoded on pEC415 plasmid were grown in 2YTPG nutrient rich media. Transformed CLM24 *E. coli* was also used for expression of PglB expressed on pEXT21 plasmid and the native PglB heptasaccharide expressed on pACYC to be used as a positive control polysaccharide. All plasmids are stated in general method section Table 2.2. Recombinant rhamnose variants were expressed in CLM24 in addition to CS2775 *E. coli* (Table 4.1). Cultures were grown overnight at 30 °C after induction with either 1 mM IPTG (Geationwneron) for the cultures containing plasmids encoding PglB, model carrier protein AcrA, or *C. jejuni* heptasaccharide, or alternatively 0.2% arabinose (BioBasic Canada Inc.) for cultures containing recombinant polysaccharide construct pHD0677 or GAS carrier proteins (MalE, SpyAD, and SpyCEP) at mid log phase ( $OD_{600} = 0.4 - 0.6$ ). Alternatively, wildtype and mutant lipid linked oligosaccharides (LLO-GAC / LLO-GAS\_Rha) were extracted from 1 L GAS cultures, either wildtype or  $\Delta gacI$  NCTC - 8198 strain (Chapter 3 section 3.3.6) grown as stated previously (Chapter 3 section 3.3.2).

Both *E. coli* and GAS cells were harvested with a 4,000 x *g* spin for 20 minutes and washed in S30 buffer (10 mM tris acetate, 14 mM magnesium acetate, 60 mM potassium acetate, pH 8.2) three times in a total 50 ml volume, centrifuged for 10 minutes at 4,000 x *g* between each wash. Supernatants were discarded, and after the final wash step, pellets were resuspended in 10 - 50 ml of S30 buffer and lysed as follows.

*E. coli* cell pellets obtained from large scale bacterial cultures containing recombinant rhamnase constructs, carrier proteins and PglB were lysed separately in a cell homogeniser applying high pressure to break open cells. The system was washed with 20% ethanol followed by 70% ethanol then distilled H<sub>2</sub>O, before the chamber of the homogeniser was equilibrated with 50 ml of S30 buffer. The resuspended pellets were added and run through the system 3 – 5 times until the lysates became less opaque and viscous. The lysates were kept on ice between each run and finally centrifuged at 30,000 x *g* for 1 hour at 4 °C to pellet debris and separate cell free lysates present in the supernatant.

Mechanical cell disruption was applied by ribolyser to small scale *E. coli* cell pellets (e.g., < 300 ml initial culture volume) and GAS cell preparations. Washed *E. coli* cells were transferred to 15 ml lysing tubes containing lysing matrix B beads (MP Biomedicals) and ribolysed on a FastPrep-24 Classic Instrument (MP Biomedicals), set to programme 1 (2 m/s for 30 seconds) for 5 cycles. Resuspended S30 buffer washed GAS cells were also transferred into 15 ml lysing tubes before mechanical disruption via 5 cycles on programme 3 (6.4 m/s for 60 seconds) for 5 cycles to extract lipid linked polysaccharide. Cell free lysate extracts were centrifuged at 4,000 x *g* for 10 minutes, further clarified with another spin at 4,000 x *g* for 10 minutes, until the lipid linked polysaccharide in the supernatant was separated from the lysing beads settled in the bottom of the lysing tubes.

Resulting cell free lysates were either used directly in reactions, or alternatively concentrated, a step particularly necessary for samples containing LLO-GAC / LLO-GAS\_Rha extracted from GAS cells. Approximately 10 ml of polysaccharide cell free lysates were concentrated through high-speed centrifugation at 100,000 x *g* for 1 hour at 4 °C, with the resulting pellet resuspended in 1 ml of S30 buffer, 1 / 10<sup>th</sup> of the starting volume. Clarified and concentrated cell free lysates where necessary were aliquoted into 100 – 500 µl volumes to limit freeze thawing, either used in reactions immediately, or stored short term at - 20 °C or long term at - 80 °C for PglB containing lysates.

#### 4.3.9. 'One Pot' *In Vitro* Glycosylation Assay

Reactions were set up to contain titrated volumes of acceptor protein lysate, PglB lysate, and LLO-GAC/ LLO-GAS\_Rha lysates or recombinant rhamnose lysates. Volumes were specific to the batch of each component, ranging from 10  $\mu$ l to 800  $\mu$ l. The mixture of lysates were incubated with PglB cofactors, 0.1% n-dodecyl- $\beta$ -d-maltopyranoside (DDM) (Sigma-Aldrich™) and 10 mM manganese(II) chloride tetrahydrate ( $MnCl_2$ ) (ACROS Organics™) in a total volume of 1 ml topped up with S30 buffer. Reactions were incubated overnight for ~ 18 hours at 30 °C shaking at 100 – 140 rpm. Reaction supernatants following a 13,000 x g spin for 10 minutes were purified with 50  $\mu$ l of Ni-NTA agarose beads (Qiagen) (general methods section 2.3.3), and His-purified samples were separated by SDS-PAGE and analysed by western blot.

#### 4.3.10. Cell-Based Glycoprotein Production Inside *E. coli* Cells

The co-transformed strains containing chromosomal PglB (*cedA::pglB*) and optimised recombinant rhamnose structures (rpORha - pHD0677 and pHD0256) and protein antigens expressed by pEXT21 were grown in various media (Chapter 4 sections 4.3.6 – 4.3.7), supplemented with appropriate antibiotics for plasmid maintenance, namely spectinomycin (protein), ampicillin and erythromycin (recombinant rhamnose) and kanamycin (chromosomal PglB integration) (Table 4.2).

Overnight cultures of the co-transformed strain for glycosylation testing, in addition to the protein only chromosomal PglB control strain were grown in 2YT, 2YTPG or BHI media. A volume of the overnight culture was added to the selected broth, inoculated to achieve  $OD_{600nm}$  of 0.05. Cultures were grown at 37 °C at 180 rpm until reaching late-log phase,  $OD_{600nm}$  of 0.6 – 0.8, before induction where appropriate with 0.2 % L-arabinose (BioBasic Canada Inc.) for expression of genes in recombinant polysaccharide construct pHD0677, and 1 mM IPTG for protein antigen expression. Following induction, the growth temperature was dropped to 30 °C at 200 rpm, and nutrients were added to the culture medium. Specifically, 10 mM glycerol (Sigma-Aldrich™), 10 mM  $MgSO_4$  (Sigma-Aldrich™), 10 mM GlcNAc (Sigma-Aldrich™) were added to the cultures to improve polysaccharide expression, and 4 mM  $MnCl_2$  was added as a PglB cofactor to improve glycosylation<sup>582</sup>. Cultures were then incubated overnight (~ 18 hours) before obtaining cell pellets from a 20 minute spin at 4,000 x g, standardised to the same OD either in PBS buffer for lysis with a ribolyser or in periplasmic lysis buffer (Chapter 4 section 4.3.7). Samples were then His-purified using Ni-NTA beads (Qiagen) for small scale glycosylation

testing, and nickel affinity chromatography using an NTA-Ni 1 ml HisTrap™ IMAC column (Cytiva) attached to an ÄKTA Start protein purification system (Cytiva) for large scale cultures (general methods section 2.3.3).

For large scale cultures, following His tag purification, potential glycoconjugate samples pooled from collected fractions, were further purified by Size Exclusion Chromatography (SEC) to remove any contaminating protein or polysaccharide species from the glycoconjugate sample (general methods section 2.3.4). Purified samples were incubated with Pierce™ High Capacity Endotoxin Removal Resin (Thermo Fisher Scientific) overnight at room temperature with end to end rotation (general method section 2.3.5), and sample endotoxin levels were analysed by Limulus Amebocyte Lysate (LAL) test performed by the Pyrogens Team (NIBSC).

#### 4.3.11. Glycoprotein Analysis

##### 4.3.11.1. *Soybean Agglutinin Lectin Staining*

Glycoprotein samples were either electroblotted or spotted directly onto nitrocellulose membranes blocked in 1X Carbo-Free™ Blocking Solution (Vector Laboratories) for 30 minutes at room temperature. Biotin-conjugated Soybean Agglutinin (SBA) lectin (Vector Laboratories) was added (general method section 2.1.2) to assess conjugation of *C. jejuni* heptasaccharide for 30 minutes at room temperature. Blots were washed three times in PBS with 0.1% Tween-20 (PBS-T), before addition of IRDye® 680RD Streptavidin (LiCor) diluted in PBS. Nitrocellulose membranes were washed a further three times in PBS-T, and imaged using Odyssey® Imaging System (Li-Cor).

##### 4.3.11.2. *Proteinase K Treatment*

To determine successful decoration of proteins with recombinant rhamnose polymers 50 µg/mL of proteinase K (QIAGEN) was added to periplasmic extracts, or to pooled His purified samples (Chapter 4 section 4.3.10). Samples were incubated with and without proteinase K in a water bath at 55 °C for 1 hour, after which they were analysed by SDS-PAGE and western blot.

##### 4.3.11.3. *Biconjugate Capture ELISA*

Wells of MaxiSORP™ 96-well plates (Thermo Fisher Scientific) were coated with 1 : 1,000 diluted Goat anti-GAS (Thermo Fisher Scientific) antibody (general method section 2.1.2) in carbonate buffer (general method section 2.4.2). Plates were left to coat overnight at 4 °C for capture of recombinant rhamnose polysaccharide. The following morning, plates were washed 3 times with PBS-T and blocked with 100 µl/well AD for 1 hour at 37 °C, before washing again three times with PBS-T.

Purified glycoprotein and protein only control samples were diluted to 100 µg/ml of protein, calculated by BCA assay (general method section 2.5.1) and serially diluted across the plate in AD buffer before being left at room temperature with plate agitation for 1 hour. AD buffer was used as a negative no sample control. Plates were washed a further three times with PBS-T before addition of 100 µl/well of mouse anti-His HRP conjugate diluted 1 : 5,000 in AD buffer incubated at room temperature for 1.5 hours. Plates were washed 3 times again before development with TMB and sulphuric acid (general method section 2.4.2). Plates were read at an absorbance of 450 nm on a SoftMax Pro plate reader (Molecular Devices, UK) and graphing was performed using GraphPad Prism (general method section 2.6.2).

## 4.4. Results

### 4.4.1. Protein Carriers Used for Bioconjugation Approaches

Towards the aim of PglB specific bioconjugation, model proteins were selected, and GAS protein antigens were designed, cloned, and expressed, with the suitability for 'one pot' *in vitro* and cellular glycosylation approaches assessed.

#### 4.4.1.1. *GAS Specific Protein Carrier Cloning, Expression and Cellular Localisation in the Periplasmic Space*

To explore their compatibility with PGCT, GAS protein antigens, MalE, SpyCEP and SpyAD were modified to include an N-terminal Sec secretion signal, DsbA (MKKIWLALAGLVLAFSASAAQ) for periplasmic localisation<sup>583</sup>, a C-terminal PglB glycotag (DQNAT) and a decahistidine tag (Figure 4.3a). Full gene sequences for each antigen, with the exception of SpyCEP cloned as an inactive enzymatic fragment<sup>208</sup>, were synthesised for cloning into pEC415 expression vector (general method Table 2.2).

GAS protein antigen expression levels and cellular localisation were initially tested in BL21 (DE3) *E. coli*. Cells were induced with 0.2 % arabinose and grown for 3 hours or overnight post induction at mid-log phase. Pelleted cells were fractionated into insoluble, cytoplasmic, and periplasmic compartments for detection by mouse anti-His antibodies via western blotting. Expression of MalE, SpyCEP and SpyAD was observed at 3 hours post induction and were observed in all tested cellular compartments. Figure 4.3b shows the periplasmic cell compartment, with full length SpyAD and MalE observed after overnight growth, but both full length and a lower molecular weight species for the cloned SpyCEP fragment. This demonstrates that all three GAS antigens are successfully exported to the periplasm and can therefore be theoretically targeted by PglB for glycosylation.

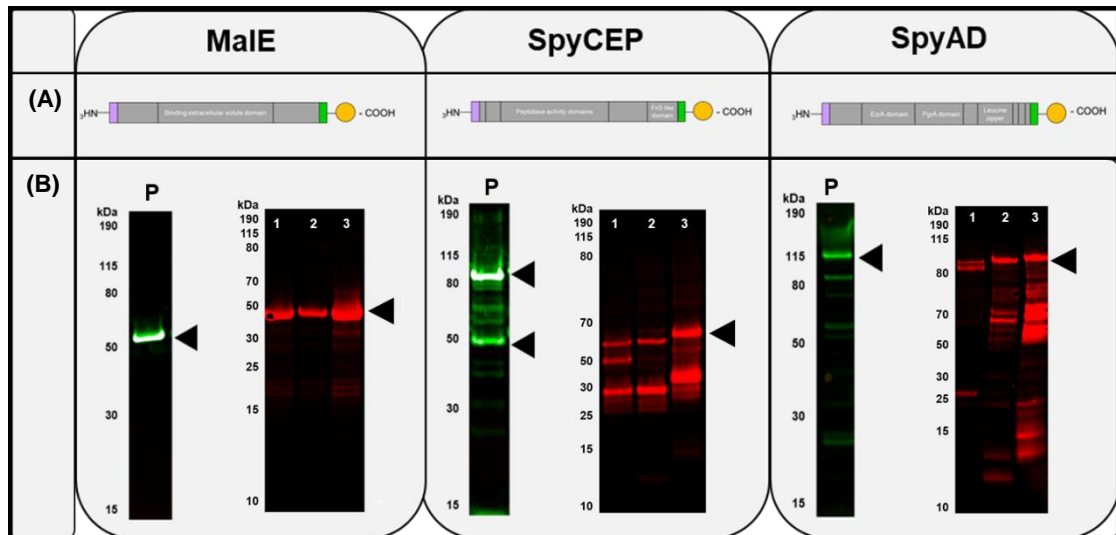
#### 4.4.1.2. *GAS Specific Protein Carrier Glycotag Availability Using C. jejuni Heptasaccharide as a Model Polysaccharide*

Following confirmation of expression and periplasmic localisation of GAS target antigens, glycotag availability was assessed. This was investigated using *C. jejuni* heptasaccharide as a model polysaccharide, being a well characterised native PglB substrate<sup>555</sup>. The *pgl* operon containing all the necessary machinery for *C. jejuni* heptasaccharide biosynthesis was expressed on pACYC plasmid in CLM24 *E. coli* cells. Initially to save time, and avoid the fine-tuning of expression of the three



components required for glycosylation in a single *E. coli* cell, the 'one pot' *in vitro* approach was used to test glycotag targeting.

Purified GAS protein antigens were separated by SDS-PAGE and analysed by western blot for changes to protein mass following glycosylation with the 1.2 kDa *C. jejuni* heptasaccharide. Figure 4.3b shows that all three GAS carriers undergo an increase in size in the presence of PglB OST and lipid linked *C. jejuni* heptasaccharide, absent in the no PglB controls with unglycosylated protein sizes as expected (MalE 41 kDa, SpyCEP fragment 60 kDa, and SpyAD 88 kDa). Glycotag targeting of GAS antigens with PglB appears to be most efficient with SpyCEP and SpyAD proteins, whereas with MalE most of the protein species appear to be unglycosylated. All cell lysates show evidence of degradation products or rather non-specific cell components which have co-eluted following His-tag purification.



**Figure 4.3: Expression, localisation, and *C. jejuni* heptasaccharide glycosylation of recombinant GAS carrier proteins using the 'one pot' *in vitro* bioconjugation system.**

Cloned GAS antigens MalE, SpyCEP and SpyAD were expressed in BL21 (DE3) *E. coli*. At mid-log, cultures were induced with 0.2 % arabinose and grown for a further 3 or 18 hours.

**(A)** Schematic representation of selected GAS protein domains engineered to contain an N-terminal periplasmic DsbA secretion tag (purple), C-terminal glycotag (green) and x10 Histidine tag (yellow circle).

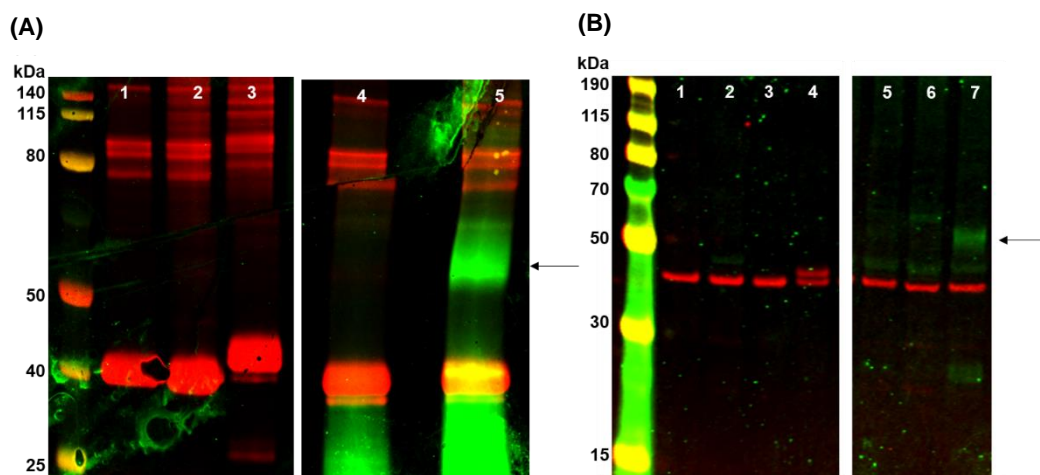
**(B)** Overnight GAS protein expression and localisation in the periplasm (P) of BL21 (DE3) *E. coli* cells detected by mouse anti-His antibodies developed by Goat anti-mouse IgG IRDye®-800CW (green channel). 'One pot' *in vitro* test glycosylation of GAS antigens in the presence of *C. jejuni* heptasaccharide and PglB cell-free lysates. Arrows indicate protein size shift in the presence of all three components, detected by mouse anti-His antibodies developed by Goat anti-mouse IgG IRDye®-680CW (red channel). Control reactions 1) protein only, 2) protein and polysaccharide only, and test reaction 3) protein, polysaccharide and PglB co-incubated.

#### 4.4.2. Polysaccharide Extracted From GAS Cells as Compatible Substrates Recognised and Transferred by PglB to Model Carrier Protein AcrA

PglB has relaxed substrate specificity, but demonstrates reduced affinity for polysaccharides that lack N-acetyl group at their reducing end and have  $\beta$ -1,4 linkage between their first and second monosaccharide<sup>584–586</sup>. Wildtype GAC contains a GlcNAc at its reducing end but a  $\beta$ -1,4 linkage which suggests it may be a poor substrate for PglB.

To test whether GAC, a non-native polysaccharide, can be recognised and transferred to a carrier protein by PglB, the same 'one pot' *in vitro* assay was used. The *C. jejuni* heptasaccharide was substituted for lipid linked GAC (LLO-GAC) extracted from GAS cells as the assay is compatible with both recombinantly produced polysaccharides and native lipid linked polysaccharides. AcrA was selected as a model carrier protein as it has been used in a wide range of bacterial glycosylation studies<sup>478,568,582,587</sup>, and contains a single functional glycotag for targeted glycosylation. The three components required for glycoconjugate production (polysaccharide, carrier protein and OST) were independently prepared, and cell free lysates were incubated in a single tube. These reactions comprised of LLO-GAC, and *E. coli* lysates containing recombinant carrier protein AcrA and *C. jejuni* PglB enzyme. A positive control reaction with AcrA, PglB, and *C. jejuni* heptasaccharide was also included to confirm reaction conditions, PglB activity, and reliability of the 'one pot' *in vitro* assay.

Following overnight incubation and His-purifications glycoproteins were separated by SDS-PAGE and immunoblotted with anti-His antibodies to detect AcrA, and anti-GAS antibodies to detect conjugated LLO-GAC. Glycosylation of AcrA with the *C. jejuni* heptasaccharide shows a size shift in AcrA protein (Figure 4.4a, lane 3), yet following LLO-GAC incubation there was no shift in the anti-protein signal. This was despite an anti-GAC signal observed at 60 kDa, absent when PglB was excluded from the reaction (Figure 4.4a, lanes 4 and 5), suggesting successful conjugation but with the majority of protein being unglycosylated.



**Figure 4.4: ‘One pot’ *in vitro* glycosylation of wildtype (LLO-GAC) and mutant (LLO-GAS\_Rha) polysaccharides as PglB OST substrates for conjugation to AcrA carrier protein.**

AcrA protein, lipid linked GAC polysaccharides (LLO-GAC/ -GAS\_Rha), and PglB OST cell free lysates were incubated overnight, his-purified and reactions separated by SDS-PAGE before protein and polysaccharide detection by Western blot. AcrA was detected with Mouse anti-His developed with Goat anti-mouse IgG IRDye®-680RD (protein = red) and GAC variants by Rabbit anti-GAS antibody developed with Goat anti-rabbit IRDye®-800CW (polysaccharide = green). Arrow indicates a potential glycosylation product at ~ 60 kDa.

Lanes standardised to contain the same volumes of AcrA and PglB components for each blot.

**(A) LLO-GAC Wildtype:** Lanes 1 – 3, control reactions: 1 – AcrA alone; 2 – AcrA + PglB; 3 – AcrA + PglB + *C. jejuni* heptasaccharide, and 4 – AcrA + LLO-GAC; Lane 5 is a test reaction: - AcrA + LLO-GAC + PglB. Lanes 4 and 5 have matched LLO-GAC lysate volumes.

**(B) LLO-GAS\_Rha Mutant:** Lanes 1 – 4, control reactions: 1 – AcrA alone; 2 – AcrA + LLO-GAS\_Rha; 3 – AcrA + PglB; 4 – AcrA + PglB + *C. jejuni* heptasaccharide. Lanes 5 – 7, test reactions containing increasing amounts of LLO-GAS\_Rha. Lane 3 LLO-GAS\_Rha lysate volume matched to lane 7 test sample.

LLO-GAC, however, contains GlcNAc side chains with  $\alpha 1 - 2$  and  $\alpha 1 - 3$  linkages at position 3 in addition to the reducing end GlcNAc<sup>139</sup>. Therefore, to confirm that putative PglB activity observed in Figure 4.4a was targeted to the reducing end GlcNAc, and not the sidechain  $\alpha$ -linked GlcNAc residues, the GAS mutant generated for chemical conjugation purposes, producing polyrhamnose backbones without GlcNAc side chain attachment was tested (Chapter 3). Lipid linked polysaccharides were extracted from the mutant GAS strain in the same manner as the wildtype GAS isolate. This lipid linked GAC variant, referred to as LLO-GAS\_Rha, was added into ‘one pot’ reactions. The volume of LLO-GAS\_Rha lysates were titrated and co-incubated with AcrA carrier protein and PglB OST cell free lysates, prior to purification, and analysis as stated above. Similar to LLO-GAC,

results show no obvious size change in the glycosylated AcrA carrier protein. However, a similar 60 kDa species detected by GAS antisera is also evident for the LLO-GAS\_Rha containing samples (Figure 4.4b, lanes 5 – 7), absent from the no PglB controls when incubated at the highest volume of LLO-GAS\_Rha to standardised volumes of AcrA and PglB lysates, matched for test and control samples (Figure 4.4b, lane 2). These proof of principal studies validated the use of GAC in PGCT, therefore, plasmids enabling recombinant polysaccharide expression in *E. coli* to resemble part of the GAC structure were sourced from Dr. Dorfmueller as part of a collaboration with the University of Dundee.

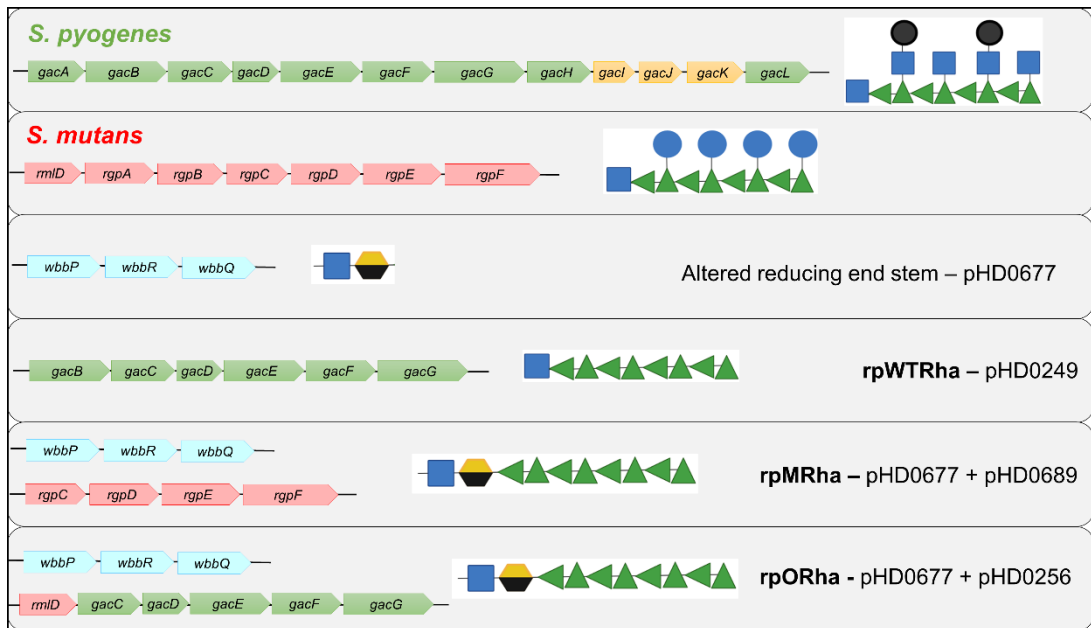
#### 4.4.3. *E. coli* Expressed Recombinant Rhamnose for Use in Bioconjugation

##### 4.4.3.1. *Construction of E. coli Produced Recombinant Rhamnose (rpWTRha, rpMRha, and rpORha) Plasmids*

One of the first and most important steps to consider in the construction of recombinant bioconjugate vaccines is the ability to express biosynthesis genes which encode a polysaccharide of interest inside a non-native host cell. This is traditionally and most simply achieved through cloning of an exogenous polysaccharide synthesis locus onto a compatible plasmid for expression in a suitable glycoengineering *E. coli* strain<sup>588</sup>.

In collaboration with Dr. Dorfmueller from the University of Dundee, expression of recombinant polymers was assessed in *E. coli* host cells. Successful cloning and expression of such large genetic regions is not without challenges and such initial trials to express the entire *gac* locus resulted in mutations leading to expression of polyrhamnose structures without GlcNAc side chains (Patent Reference: WO2020249737A1). Three recombinant polyrhamnose systems were tested with the predicted polysaccharide structures shown in Figure 4.5 and biosynthesis genes in Tables 4.3 and 4.4.

- 1) Polyrhamnose from the *gac* locus (pHD0249) (Table 4.3) referred to as rpWTRha (Recombinant Poly-Wildtype Rhamnose)
- 2) Polyrhamnose from closely related *Streptococcus mutans* genes (pHD0689) (Table 4.4) referred to as rpMRha (Recombinant Poly-Modified Rhamnose)
- 3) Polyrhamnose using an optimised system composed of a combination of *S. pyogenes* genes within the *gac* locus, and a *S. mutans* gene (pHD0256) (Table 4.4) referred to as rpORha (Recombinant Poly-Optimised Rhamnose)



**Figure 4.5: Schematic representation of genetic loci and recombinant rhamnose structures encoded on plasmids pHD0249, pHD677 + pHD0689, and pHD0677 + pHD0256.**

Horizontal arrows represent each gene designation, with colour denoting predicted gene function. Green and Red, rhamnose polymerisation (GAS or *S. mutans* respectively); Orange, GlcNAc side chain modification in GAS cells; Blue, altered reducing end biosynthesis based on *Shigella* genes. Visual representation of native and recombinant polysaccharides based on nomenclature for glycans (SNFG)<sup>704</sup>. Green triangle; rhamnose, blue square; GlcNAc, black circle; phosphate modification, blue circle; glucose, and yellow and black hexagon; galactose.

The *gac* genetic operon encodes all genes necessary to synthesise GAC, and the *rgp* genetic operon encodes all genes necessary to synthesise rhamnose-glucose polysaccharide (RGP) in *S. mutans*.

pHD0249 denotes the plasmid used to express rhamnose with native  $\beta$ -1,4 GlcNAc to rhamnose reducing end linkage producing rpWTRha.

pHD0677 denotes the plasmid used to synthesise an altered reducing end sugar stem containing GlcNAc – Galactose  $\alpha$ -1,3 linkage. Either pHD0689 or pHD0256 plasmids are used to polymerise rhamnose based on *S. mutans* or GAS genes respectively. When co-expressed either rpMRha or rpORha is produced.

**Table 4.3: Genes encoding rpWTRha (pHD0249).**

Genes present on the plasmid produces *S. pyogenes* derived rhamnose polymers with a native  $\beta$ 1,4 GlcNAc – Rha reducing end for PglB recognition.

pHD0249		
Gene Name	Gene Function	Description
<i>gacB</i>	$\alpha$ -D-GlcNAc $\alpha$ -1,2-L-rhamnosyltransferase	Attaches the first TDP- $\beta$ -L-activated rhamnose onto acceptor und-PP-GlcNAc 3-OH group forming a $\beta$ -1,4 glycosidic bond.
<i>gacC</i>	$\alpha$ -L-Rha $\alpha$ -1,3-L-rhamnosyltransferase	Elongation of rhamnose backbone working with GacF and GacG.
<i>gacD</i>	ABC-transporter (permease protein)	Translocation of rhamnose backbone across the inner membrane using ATP dependent ABC transporter, working with GacE.
<i>gacE</i>	ABC-transporter (ATP-binding protein)	Translocation of rhamnose backbone across the inner membrane using ATP dependent ABC transporter, working with GacD.
<i>gacF</i>	$\alpha$ -L-Rha $\alpha$ -1,3-L-rhamnosyltransferase	Elongation of rhamnose backbone working with GacC and GacG.
<i>gacG</i>	$\alpha$ -L-Rha $\alpha$ -1,3-L-rhamnosyltransferase	Elongation of rhamnose backbone working with GacC and GacF.

To improve *C. jejuni* PglB OST recognition, pHD0677 plasmid was generated encoding *Shigella dysenteriae* O-antigen genes. Expression enabled D-galactose attachment to GlcNAc at polysaccharide reducing ends as a permissible PglB substrate<sup>558</sup>. Specifically, *wbbP* (Q53982\_SHIDY) generates activated UDP-D-galactose<sup>589</sup>, which is attached to lipid linked GlcNAc on the inner membrane generated by *wecA* gene from *E. coli* host cells. This is then further elongated by *S. dysenteriae wbbR* (Q32EG0\_SHIDS) and *wbbQ* (A0A090NIC3\_SHIDY) genes which encode TDP-  $\alpha$ 1,3-L-rhamnose<sup>590,591</sup>, forming an  $\alpha$ -1,3 GlcNAc-Galactose at the reducing end, and rhamnose as the third monosaccharide in the chain. The rhamnose polymer can then be further elongated by either pHD0256 or pHD0689 to build the entire backbone before translocation across the membrane for PglB targeting.

The pHD0689 construct contains *S. mutans* genes similar to genes in the *gac* operon (*rgpC* – *rgpF*), whereas pHD0256 construct contains *S. pyogenes* genes, specifically *gacC* to *gacG*, along with *rmID* from *S. mutans* to enable this polymerisation step from the modified acceptor stem (Figure 4.5). Therefore, when either of these plasmids are co-expressed with pHD0677 in *E. coli* cells, the resulting polysaccharide structure is a modified stem of GlcNAc and galactose, followed by the rhamnose backbone. pHD0689 *S. mutans* gene derived polymers are referred to as rpMRha (Recombinant Poly-Modified Rhamnose), and pHD0256

GAS gene derived polymers are referred to as rpORha (Recombinant Poly-Optimised Rhamnose) (Table 4.4).

**Table 4.4: Genes encoding the modified acceptor stem (pHD0677), rpMRha (pHD0689) and rpORha (pHD0256).**

Genes present on the two plasmids which when co-expressed produces rhamnose (either *S. mutans* or *S. pyogenes* derived) with a modified reducing end containing an 'acceptor stem' for PglB recognition, specifically an  $\alpha$ 1,3- GlcNAc – Galactose linkage.

pHD0677		
Gene Name	Gene Function	Gene Description
<i>wbbP</i>	Glycosyltransferase	Generation of D-galactose for attachment to lipid linked und-PP-GlcNAc attached to the inner membrane.
<i>wbbR</i>	Glycosyltransferase	Attachment of TDP- $\alpha$ 1,3-L-Rhamnose to D-galactose attached to lipid linked GlcNAc by $\alpha$ -1-3 linkage.
<i>wbbQ</i>	Glycosyltransferase	
pHD0689		
Gene Name	Gene Function	Description
<i>rgpC</i> ( <i>gacD</i> homolog)	ABC-transporter (permease protein)	Translocation of rhamnose backbone across the inner membrane using ATP dependent ABC transporter, working with RgpD.
<i>rgpD</i> ( <i>gacE</i> homolog)	ABC-transporter (ATP-binding protein)	Translocation of rhamnose backbone across the inner membrane using ATP dependent ABC transporter, working with RgpC.
<i>rgpE</i> ( <i>gacF</i> homolog)	Glycosyltransferase	Elongation of rhamnose backbone.
<i>rgpF</i> ( <i>gacG</i> homolog)	Rhamnosyltransferase	Elongation of rhamnose backbone.
pHD0256		
Gene Name	Gene Function	Description
<i>rmID</i>	dTDP-4-dehydrorhamnose reductase	Reduces dTDP-6-deoxy-L-lyxo-4-hexulose to dTDP-L-rhamnose. Gene from <i>S. mutans</i> .
<i>gacC</i>	$\alpha$ -L-Rha $\alpha$ -1,3-L-rhamnosyltransferase	
<i>gacD</i>	ABC-transporter (permease protein)	See Table 4.3 for gene function and Chapter 1 figure 1.1 for schematic of native GAC polymer biosynthesis.
<i>gacE</i>	ABC-transporter (ATP-binding protein)	
<i>gacF</i>	$\alpha$ -L-Rha $\alpha$ -1,3-L-rhamnosyltransferase	
<i>gacG</i>	$\alpha$ -L-Rha $\alpha$ -1,3-L-rhamnosyltransferase	

All three *E. coli* plasmid systems are predicted to utilise the *E. coli wecA* gene to provide lipid linked GlcNAc at the reducing end attached to the inner membrane<sup>558,592</sup>, due to the absence of GAS or *S. mutans wecA* homologs (*gacO* and *rgpG* respectively) in the constructs. Figure 4.5 demonstrates the intended polysaccharide structures produced from these plasmid combinations, with pHD0249 thought to build rhamnose straight onto und-PP-GlcNAc, and pHD0689 and pHD0256 thought to have an additional galactose monosaccharide before rhamnose polymerisation. All three polysaccharide variants should then be translocated fully assembled by a flippase enzyme in the form of an ABC transporter, provided by *gacD* and *gacE* GAS genes encoded on pHD0249 and pHD0256<sup>141,145,152</sup>, or *rgpC* and *rgpD* *S. mutans* genes encoded on pHD0689<sup>571,572</sup>.

#### 4.4.3.2. Selection of *E. coli* Strains for Recombinant Rhamnose Expression

To improve bioconjugation efficiency *E. coli* cells are often modified to optimally synthesise non-native polysaccharide structures. Recombinant rhamnose constructs, rpWTRha, rpMRha, or rpORha, were transformed via electroporation into either electrocompetent CLM24<sup>558</sup> or CS2775<sup>571</sup> *E. coli* host cells.

CLM24 *E. coli* was initially selected as the host cell as the W3110 variant glycoengineering strain does not build its native O-antigen (O16) due to an IS5 element insertional inactivation of the *wbbL* gene, necessary for its synthesis<sup>574</sup>. This is thought to improve monosaccharide substrate availability for recombinant polysaccharide biosynthesis. Additionally, CLM24 does not possess a functional O-antigen ligase,  $\Delta waaL$  genotype, such that recombinant polysaccharide structures are not ligated onto lipid A for incorporation into the outer membrane. This in turn increases the availability of synthesised polysaccharide substrate in the periplasm where the PglB mediated conjugation reaction occurs. CLM24 has also been engineered to contain a mutation in the *lpxM* gene to remove a myristoyl group from the lipid A component<sup>578</sup> to reduce downstream toxicity as a desirable glycoconjugate vaccine property<sup>593,594</sup>.

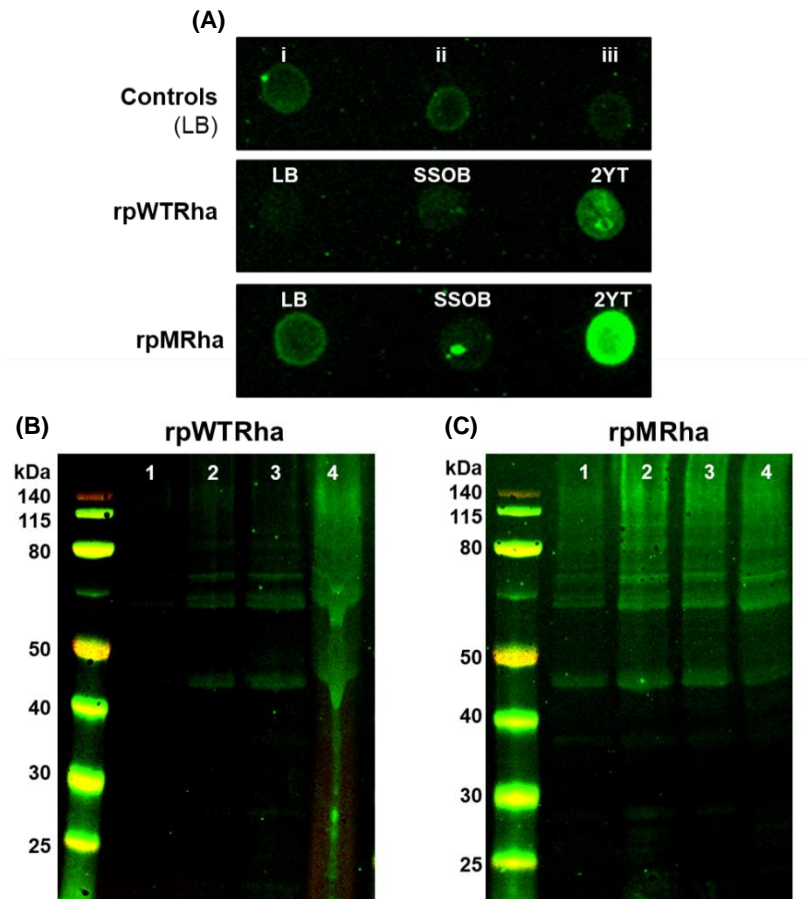
To investigate and optimise expression of these gene clusters to produce recombinant rhamnose structures, initially CLM24 *E. coli* containing either rpWTRha or rpMRha plasmids were grown in LB, SSOB, or 2YT. Whole cell lysates (WCL) were dotted onto nitrocellulose membranes, probed with Rabbit anti-GAS and detected with Goat anti-Rabbit IgG IRDye® conjugated 800CW (Figure 4.6a). Expression of both recombinant rhamnose constructs in WCL samples was poor in LB media particularly for rpWTRha which was not above the background level of



negative controls, CLM24 *E. coli* and empty plasmid constructs also grown in LB media treated in the same manner (Figure 4.6a, dots i - iii). Improved polyrhamnose production was observed with a stronger anti-GAS signal than background in CLM24 *E. coli* cells grown in 2YT nutrient rich broth compared to other media for both rpWTRha and rpMRha production (Figure 4.6a).

Next, the duration of expression was investigated for efficiency of recombinant polysaccharide production at two timepoints, 3 hours and overnight (~ 18 hours) in CLM24 *E. coli* grown in 2YT media. Periplasmic fractions were prepared in lysis buffer and compared to WCLs from lysis buffer and ribolyser methods to assess localisation of PglB OST activity and influence of lysis methods on recombinant rhamnose yields. Figures 4.6b and 4.6c show recombinant rhamnose species ran at a high molecular weight, likely based on charge and structure rather than actual polysaccharide size, with profiles looking similar for both polysaccharide variants. Detection of both rpWTRha and rpMRha polymers, with the same commercial anti-GAS antibody confirms the similarity and cross reactivity of rhamnose backbones produced from *S. mutans* *rgp* homologs and GAS *gac* genes (Figure 4.6b/c). In fact, *E. coli* expressing rpMRha gave a stronger anti-GAS signal than rpWTRha, with both constructs showing better expression when cultures were left overnight. Furthermore, WCL samples obtained from ribolyser lysis to mechanically break open cells showed stronger anti-GAS signal for the strain expressing rpWTRha (Figure 4.6b, lane 4).

Based on dot blot results refinement of growth media improved production of recombinant polysaccharide. However, expression of rpWTRha and rpMRha in separated CLM24 *E. coli* cell lysates was deemed to be suboptimal, confounded by background rhamnose signals detected by the polyclonal anti-GAS antibody. This could be due to presence of endogenous rhamnose in the host strain, observed in periplasmic control samples obtained at mid-log prior to induction of expression, particularly for lysates containing rpMRha (Figure 4.6c, lane 1). However, without empty plasmid controls in Figure 4b/c leaky expression of rhamnose genes could also not be ruled out. Nevertheless, such signals confounded western blot detection of recombinant expression, and also may be negatively impacting rhamnose monosaccharide availability for recombinant polymer production.



**Figure 4.6: CLM24 *E. coli* cells express recombinant rhamnose variants rpWTRha and rpMRha in different media and induction time points.**

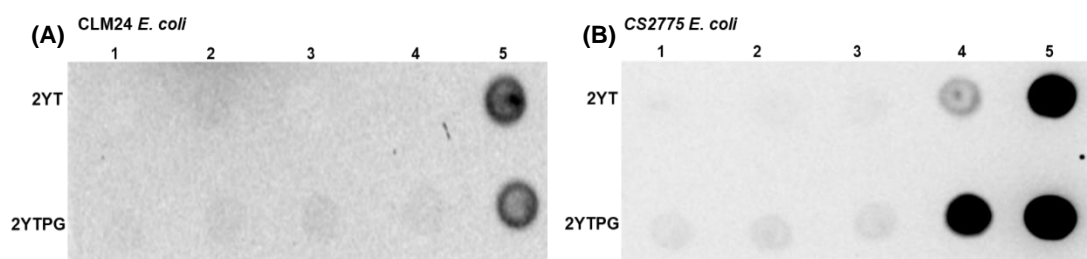
CLM24 *E. coli* expressing rpWTRha or rpMRha were grown in selected media and induced with 0.2 % arabinose (for pHD0677 expression) for indicated time scales. Samples were lysed and either spotted onto nitrocellulose membranes **(A)** or separated by SDS-PAGE **(B and C)** for western blot analysis. Rabbit anti-GAS and Goat anti-Rabbit IgG IRDye@-800CW antibodies were used to detect rhamnose expression.

**(A)** Dot blot of CLM24 *E. coli* whole cell lysates (WCL). Negative controls; i – CLM24 ; ii – CLM24 + empty plasmid for rpWTRha (pHD0139) ; iii – CLM24 + empty plasmids for rpMRha (pHD0131 + pHD0139), and test samples; rpWTRha and rpMRha in LB, SSOB, and 2YT media.

**(B and C)** CLM24 *E. coli* WCLs and periplasm cell compartments separated by SDS-PAGE and immunoblotted. Lane 1, periplasm obtained prior to induction; Lane 2, three hours post induction periplasm; Lane 3, ~ 18 hours post induction periplasm; compared to Lane 4, WCL FastPrep ribolyser cell lysate ~ 18 hours post induction. **(B)** rpWTRha and **(C)** rpMRha. Samples grown in 2YT media.

To investigate this, and avoid non-specific rhamnose detection, strain CS2775 *E. coli* were used as alternative host cells, possessing disrupted native endogenous rhamnose polysaccharide production<sup>151</sup> (Table 4.1). In this strain background, previously demonstrated nutrient rich media 2YT and 2YTPG (a similar media to 2YT with additional nutrients) were investigated for recombinant polymer production for all three plasmid systems.

There was little observed rpWTRha expression in either strain or media conditions above background levels of the strain only and empty plasmid control samples (Figure 4.7a/b, dots 1 – 3). For this reason, low expression levels, as well as a predicted suboptimal reducing end for PglB transfer, the rpWTRha rhamnose variant were not included in any subsequent glycosylation testing. However, compared with CLM24 *E. coli*, CS2775 *E. coli* gave superior rhamnose production from rpMRha in 2YT when left overnight post arabinose induction which was further enhanced by growth in 2YTPG (Figure 4.7a/b, dot 4). rpORha, expressed from *gac* genes with the modified acceptor stem, showed similar production in 2YT and 2YTPG media, with expression in both medias improved in CS2775 *E. coli* (Figure 4.7b, dot 5). For this reason rpORha polymers were used in subsequent characterisation studies and bioconjugation reactions.



**Figure 4.7: Superior recombinant rhamnose polymer expression in CS2775 *E. coli* cells compared to CLM24 *E. coli* cells in different media.**

CS2775 *E. coli* expressing rpWTRha, rpMRha, and rpORha grown in nutrient rich media (2YT and 2YTPG) induced with 0.2 % arabinose at mid log phase and incubated overnight. OD normalised samples were lysed using a ribolyser and whole cell lysates (WCL) spotted onto nitrocellulose membranes. Goat anti-GAS and Rabbit anti-Goat IgG-HRP antibodies were used to detect rhamnose expression.

Dots 1 – 2 negative controls; 1 = empty plasmid for rpWTRha (pHD0139), 2 = empty plasmids for rpM/ORha (pHD0131 + pHD0139), 3 rpWTRha, 4 rpMRha, and 5 rpORha in either 2YT or 2YTPG media. **(A)** CLM24 and **(B)** CS2775 *E. coli*.

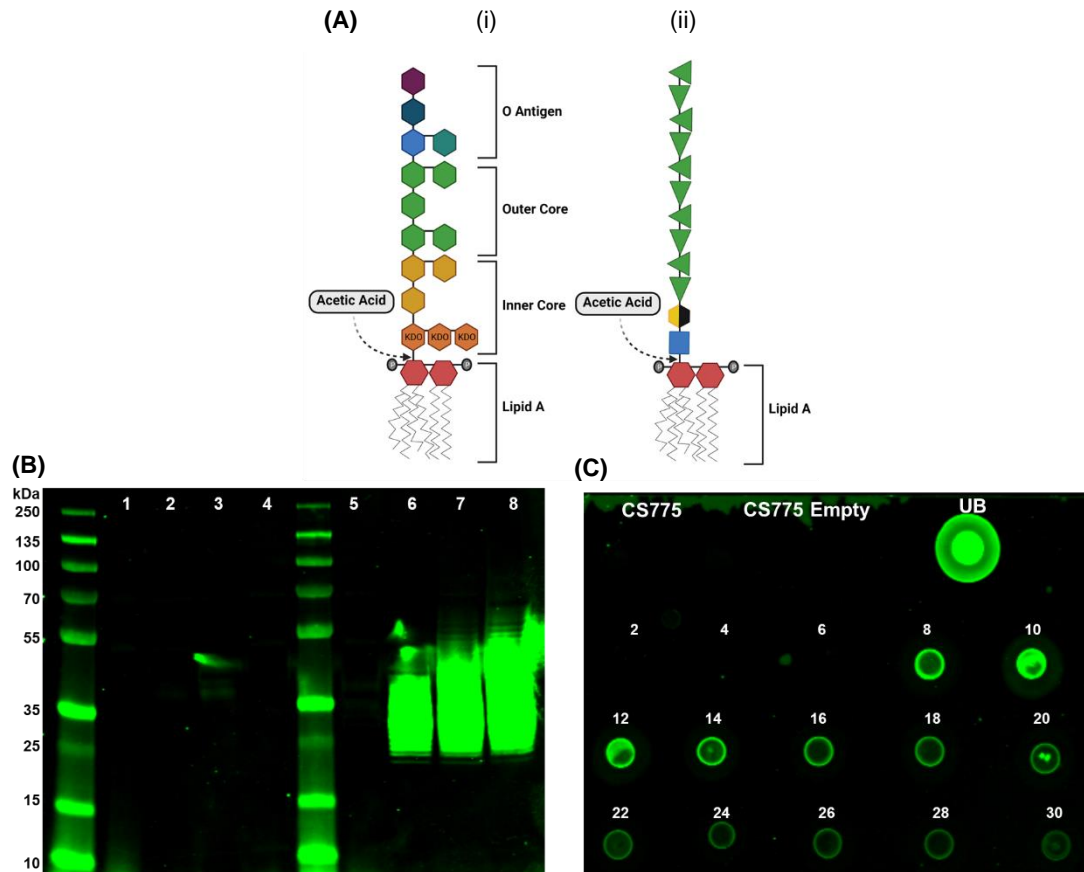
#### 4.4.3.3. Characterisation of *rpORha* Assembly and Identity

In addition to validating efficient expression, the correct assembly, structural integrity, and therefore identity of recombinant *E. coli* produced polysaccharides is pivotal when using glycoproteins as vaccine candidates. Towards this aim, *rpORha* constructs and empty plasmid controls were transformed into an *E. coli* strain related to CS2775 without *waaL* gene inactivation (Table 4.1). In this strain heterogenous polysaccharide precursors are transferred by the functional O-antigen ligase, WaaL, to the outer core of LPS on the cell surface (Figure 4.8a). The rationale being that to improve downstream analysis surface presentation of lipid A linked polysaccharide is easier to isolate than recombinant polysaccharide inside the *E. coli* cells periplasm, with less co-purified contaminating sugars, previously detected by western blot.

Strains were induced and incubated for 3 hours or overnight and lysed through incubation at 100 °C to release the outer membrane. Surface expression of lipid A linked recombinant rhamnose polymers were initially tested by western blot probing with anti-GAS antibodies (Figure 4.8b). With time there is increased recombinant rhamnose polymer production in the *rpORha* containing strain not evident in the uninduced sample (Figure 4.8a, lanes 5 and 8), nor in the strain containing empty plasmids only (Figure 4.8b, lanes 1 – 4). This demonstrates that the recombinant gene expression of *rpORha* has modified LPS profiles of CS2775 *E. coli*, with polymers in schematic Figure 4.8a (ii) expected.

For downstream analysis of *rpORha* polymer identity, such whole cell lysate samples required further purification, and detachment from the LPS molecule. Recombinant rhamnose containing samples were incubated with acetic acid, which has been shown to cleave a labile linkage between the core oligosaccharide terminal sugar 3-deoxy-D-mannooctulosonic acid / ketodeoxyoctonic acid (KDO) and lipid A, to release the O-antigen chain attached to the core sugars<sup>595</sup> (Figure 4.8a (i)). Heated acetic acid treated samples were spun to obtain a fraction containing the released O-antigen, and attached recombinant rhamnose polymers, which were applied to an ÄKTA Start system to purify the sample using a multimodal strong anion exchanger Captopadhere column used for GAC / GAS\_Rha purification in Chapter 3 section 3.3.6.3. Analysis of unbound material and eluted fractions with anti-GAS antibodies showed the strain background and empty plasmid controls did not show any signal, however there was signal observed in the unbound fractions, as well as purified eluted fractions (Figure 4.8c). The eluted

fractions showing most anti-GAS signal, 8 – 18 were pooled and buffer exchanged before lyophilisation. The freeze dried polysaccharide sample was reconstituted and hydrolysed for testing by High-Performance Anion-Exchange Chromatography with Pulsed Amperometric Detection (H-PAD) but sample impurities obscured monosaccharide detection (data not shown).



**Figure 4.8: Purification of rpORha polymers isolated from CS2775 *E. coli* cell surfaces.**

Heated acetic acid treated CS2775 *E. coli* cells were used to isolate and purify surface rpORha polymers using a captodhere column.

(A) (i) LPS structure schematic showing acetic acid liable bond between KDO of the inner core and the lipid A component on K12 derivative *E. coli*. (ii) Hypothesised rpORha polymer extracted from lipid A component. Original figure created with BioRender.com.

Western blots probed with Rabbit anti-GAS developed with Goat anti-Rabbit IRDye®-800CW antibodies.

(B) CS2775 *E. coli* (with intact Waal) WCLs. Lanes 1 – 4 empty plasmids (pHD0131 + pHD0139); lanes 5 – 8 rpORha plasmids (pHD0677 + pHD0256); Lanes 1 / 5 – Uninduced, lanes 2 / 6 – 3 hrs post induction, lanes 3 / 7 – overnight post induction, and lanes 4 / 8 – concentrated overnight post induction 1 / 100.

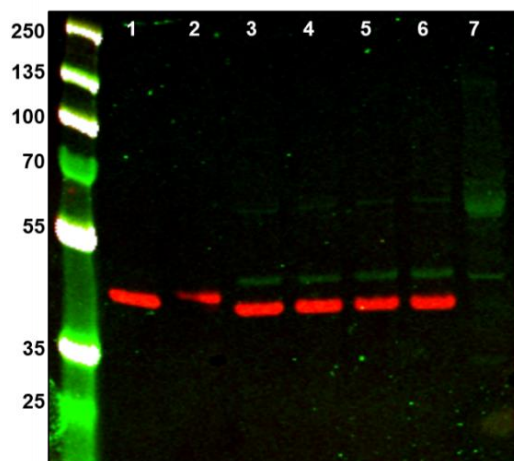
(C) Purified and released recombinant rpORha polymers eluted from multimodal captodhere column using salt eluent. Numbers relate to eluted fractions on UV<sub>280 nm</sub> chromatograph.

#### 4.4.4. 'One Pot' In Vitro Reactions to Assemble Bioconjugates

After demonstration of superior expression of rpORha recombinant rhamnose in CS2775 *E. coli* strain (Figure 4.7b), and preference of using constructs encoding rhamnose from *gac* cluster genes, the ability of PglB to recognise and transfer the modified reducing end rhamnose polymers to the well characterised carrier protein and native PglB substrate *C. jejuni* AcrA protein was assessed. AcrA used previously was selected to allow comparisons of transfer efficiency between native lipid linked polysaccharides extracted from GAS cells (Figure 4.4) and the recombinant *E. coli* produced counterpart. Due to the simplicity and previous success using the 'one pot' *in vitro* method this was initially tested.

Lipid linked rpORha polymers were extracted from CS2775 *E. coli* cells with inactive WaaL after growth in 2YTPG media by mechanical cell lysis, whereas carrier protein AcrA and PglB OST were expressed in CLM24 *E. coli*. A positive control reaction containing *C. jejuni* heptasaccharide was also included by incubating cell free lysates from CLM24 *E. coli* expressing the *pgl* locus on pACYC plasmid (Figure 4.1).

Overnight reactions were set up, purified, and analysed as previously stated for detection of glycosylated AcrA by anti-His antibodies, and rpORha by anti-GAS antibodies. Purified samples showed that AcrA could be consistently decorated with model *C. jejuni* heptasaccharide, as indicated by a clear increase in molecular weight compared to the protein alone control (Figure 4.9, lane 2). No AcrA increase in molecular weight was observed when co-incubated with PglB and rpORha (Figure 4.9, lanes 4 - 6). Some non-specific polysaccharide signal was observed in test samples, however, this may not truly be conjugated to AcrA due to the presence of a similar signal in the reactions that lacked PglB (Figure 4.9, lane 3). Therefore, although glycosylation of AcrA with *C. jejuni* heptasaccharide and GAS polysaccharide was observed using the 'one pot' method, no convincing evidence of glycosylation with rpORha was observed in the conditions tested.

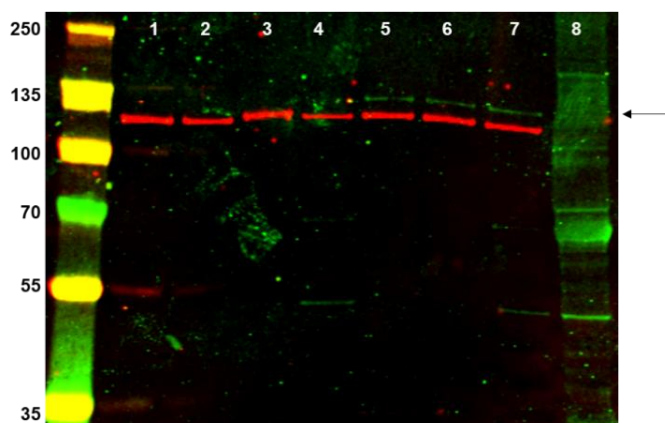


**Figure 4.9: ‘One pot’ *in vitro* glycosylation testing of rpORha for conjugation to AcrA carrier protein.**

AcrA protein, lipid linked rpORha, or control polysaccharide, and PglB OST cell free lysates were incubated overnight, and reactions separated by SDS-PAGE before detecting for protein and polysaccharide by Western blot. AcrA was detected with Mouse anti-His developed with Goat anti-Mouse IgG IRDye®-680RD (protein = red) and rpORha by Rabbit anti-GAS developed with Goat anti-Rabbit IRDye®-800CW (LLO-Polysaccharide = green) antibodies.

Lanes 1 - 3 are control reactions: 1 – AcrA alone; 2 – AcrA + PglB + *C. jejuni* heptasaccharide; 3 – AcrA + rpORha. Lanes 4 - 6 are test reactions containing increasing amounts of LLO-rpORha with PglB. Lane 7 is unconjugated LLO-rpORha from lane 4 test reaction containing the lowest LLO-rpORha lysate volume.

Following successful decoration of GAS antigens with *C. jejuni* heptasaccharide Figure 4.3, ‘one pot’ *in vitro* bioconjugation reactions were set up for simple investigation of glycotag targeting with the more relevant polysaccharide structure rpORha. When incubated with SpyCEP and MalE GAS carrier proteins neither a polysaccharide signal nor a protein size shift was detected by western blot (data not shown). However, when incubated with SpyAD carrier protein, a faint polysaccharide band was observed above the protein at ~ 135 kDa (Figure 4.10, lanes 4 - 7). Although the ‘one pot’ approach demonstrated successful transfer of model *C. jejuni* heptasaccharide to GAS protein antigens, as well as rpORha polymers to SpyAD GAS protein, the more traditional single cell-based conjugation method was subsequently investigated with both a relevant model carrier protein as well as GAS carrier proteins.



**Figure 4.10: ‘One pot’ *in vitro* glycosylation testing of rpORha for conjugation to GAS protein SpyAD.**

SpyAD protein, lipid linked rpORha, or control polysaccharide, and PglB OST cell free lysates were incubated overnight, and reactions separated by SDS-PAGE before detecting for protein and polysaccharide by Western blot. SpyAD was detected with Mouse anti-His developed with Goat anti-mouse IgG IRDye®-680RD (SpyAD = red), and rpORha by Rabbit anti-GAS developed with Goat anti-Rabbit IRDye®-800CW (LLO-rpORha = green) antibodies.

Lanes 1 - 3 are control reactions: 1 – SpyAD alone; 2 – SpyAD + LLO-rpORha; 3 – SpyAD + PglB + *C. jejuni* heptasaccharide; Lanes 4 - 7 are test reactions containing increasing amounts of LLO-rpORha with PglB. Lanes 2 control reaction is matched to lane 7 with the highest LLO-rpORha lysate volume tested. Lane 8 is unconjugated LLO-rpORha obtained from lane 4 containing the lowest LLO-Rha lysate volume. Arrow indicates suspected weak glycosylation product at ~ 135 kDa.

#### 4.4.5. Cell-Based Glycosylation to Assemble Bioconjugates

Following the apparent lack of glycosylation using the ‘one pot’ *in vitro* approach, the aim was to test recombinant rhamnose glycosylation of proteins within *E. coli* cells. rpORha plasmids were transformed into CLM24*cedA::pglB* *E. coli* strain containing the *pglB* gene inserted onto the chromosome (*cedA::pglB*)<sup>578</sup>. This strain was selected due to plasmid incompatibility reasons, as well as evidence to suggest that keeping PglB expression low reduces cellular toxicity and metabolic burden, thus improving polysaccharide production and glycosylation efficiency<sup>596</sup>.

##### 4.4.5.1. *Preparation of GAS Protein Carriers For Cellular Glycosylation*

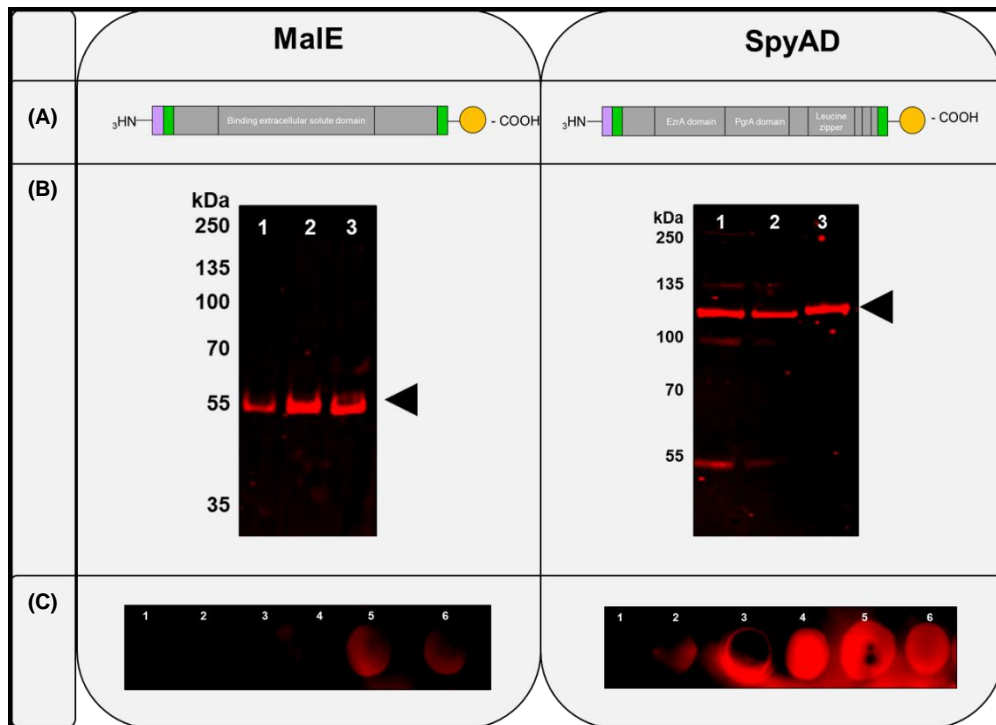
GAS specific protein antigens were cloned into low copy number plasmid pEXT21 to avoid incompatibility clashes with the plasmids encoding the polyrhamnose biosynthetic pathway genes within a single host strain. The same GAS antigens nucleic acid sequences, excluding their native secretion signal and cell anchoring domains, used previously in the ‘one pot’ *in vitro* system, were cloned into pEXT21



using XbaI and EcoRI restriction sites. Proteins were engineered to contain the same N-terminal Sec system periplasmic secretion signal sequence, DsbA (MKKIWLALAGLVLAFSASAAQ). However, an additional glycotag (DQNAT) was engineered onto the protein sequence at the N-terminus in addition to the C-terminus, aiming to improve targeted N-linked glycosylation. As previously stated, the proteins were also engineered to contain a 10xHis tag for purification purposes (Figure 4.11a). MalE and SpyAD were successfully cloned into this system.

#### 4.4.5.2. *GAS Protein Carrier Glycotag Availability Using C. jejuni Heptasaccharide as a Model Polysaccharide*

Initially, *C. jejuni* heptasaccharide, encoded by the *pgl* operon on pACYC plasmid, was transformed into cells containing GAS antigens expressed on pEXT21 plasmid. The aim of this was to validate the cellular glycosylation approach, as well as inspect site occupancy of the two glycotags present on the cloned GAS antigens. MalE and SpyAD GAS proteins showed an increase in molecular weight in the presence of PglB and the *C. jejuni* heptasaccharide compared to the protein only control strains as detected by anti-His antibodies (Figure 4.11b, lane 3). Successful conjugation was further confirmed by blotting with soybean agglutinin (SBA) lectin to detect the heptasaccharide (Figure 4.11c). This shows that glycosylation occurs in both the 'one pot' *in vitro* reactions (Figure 4.3b) and cell-based assay using chromosomal PglB (Figure 4.11b).



**Figure 4.11: Expression, localisation, and *C. jejuni* heptasaccharide glycosylation of recombinant GAS carrier proteins using the cell-based bioconjugation system.**

MaIE and SpyAD were cloned and expressed in CLM24*cedA::pglB* *E. coli* strain containing the *pglB* gene inserted onto the chromosome (*CedA::pglB*). At mid-log, cultures were induced with 1 mM IPTG and grown for a further 3 or 18 hours.

**(A)** Schematic representation of selected GAS protein domains engineered to contain an N-terminal periplasmic DsbA secretion tag (purple), glycotags at both the N- and C-terminus (green) and a 10x Histidine tag (yellow circle).

**(B)** Lanes 1 – 2: GAS protein expression and localisation in the periplasm of CLM24*cedA::pglB* *E. coli* cells at 3 hrs and 18 hrs post induction respectively. Lane 3 - Test glycosylation of GAS antigens in the presence of *C. jejuni* heptasaccharide. Arrows indicate protein size shift in the presence of all three components after 18 hrs post induction. Proteins detected by mouse anti-His and Goat anti-mouse IgG IRDye®-680RD antibodies.

**(C)** Dots 1 – 3: Protein only control strains, Dots 4 – 6: Test strains containing *C. jejuni* locus; uninduced (1 / 4), 3 hrs post induction (2 / 5), and 18 hrs post induction (3 / 6). Dotted periplasmic fractions were incubated with SBA lectin conjugated to biotin detected with streptavidin- IRDye® 680RD.

#### 4.4.5.3. Cell-Based Glycosylation of Streptococcal Model Carrier Protein NanA with rpORha

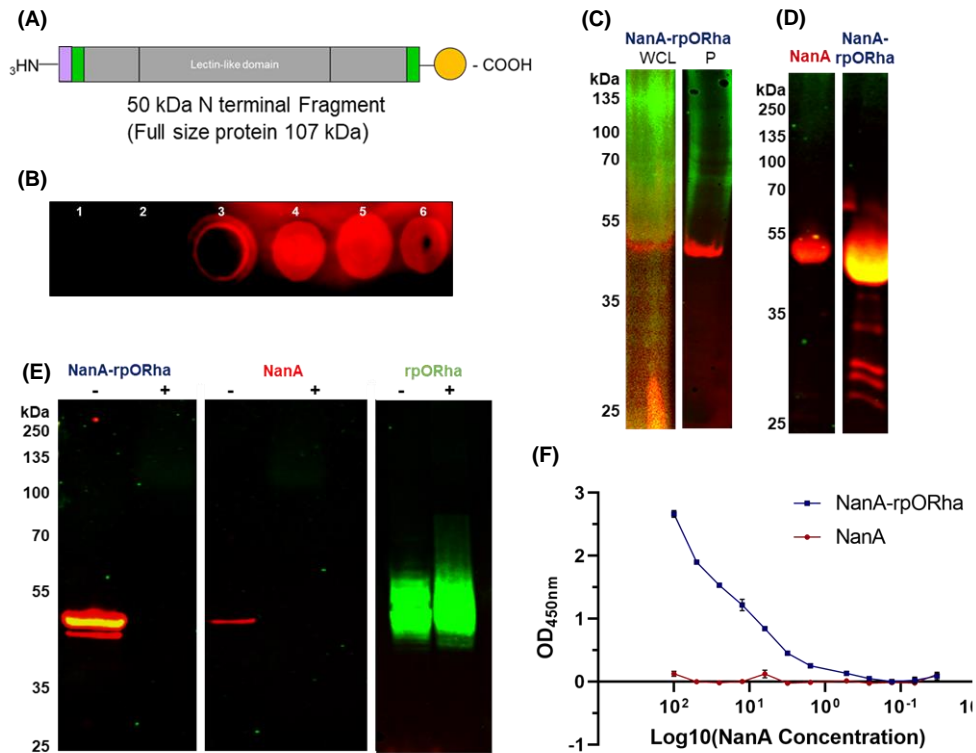
*E. coli* produced rpORha was tested using the same glycosylation system within the CLM24*cedA::pglB E. coli* strain. Neuraminidase A (NanA, Spd1504), a *S. pneumoniae* protein, was used as a model carrier protein for rpORha conjugation due to its previous success in PglB systems with the *S. pneumoniae* serotype 4 capsule<sup>597</sup>, indicating its suitability as a carrier for a streptococcal polysaccharide. NanA is secured to the cell wall<sup>598</sup>, expressed in all strains<sup>599</sup>, as a multifaceted virulence factor, shown to be immunogenic, and protective against nasopharyngeal colonisation<sup>600</sup>. In this work an N-terminal fragment, essential to protein function as a lectin like domain<sup>601</sup> was expressed (Figure 4.12a) due to the full-length size significantly reducing recombinant protein yields in previous work<sup>597</sup>. NanA was engineered to contain 2 glycotags at either terminus which could be targeted with *C. jejuni* model heptasaccharide using cell-based chromosomal PglB (Figure 4.12b).

CLM24*cedA::pglB* cells expressing rpORha were transformed with *nanA* in the same low copy number plasmid used for GAS antigen expression (pEXT21). Initial small-scale cultures (10 – 50 ml) showed no evidence of glycoprotein production in either 2YT or 2YTPG media (data not shown). A weak glycosylation signal was evident in *E. coli* cells grown in BHI media, therefore, to test for successful glycosylation, cultures were scaled up to 1 L volumes. Expression of both rpORha polymers and NanA protein was observed in WCLs and periplasmic fractions (P) isolated from large scale cultures grown in BHI media incubated overnight at 30 °C post induction (Figure 4.12c). Purified His tagged NanA from WCL samples was tested from both a NanA protein only control strain and the conjugation strain (NanA-rpORha). Western blot analysis showed purified NanA at a similar predicted fragment molecular weight (~ 50 kDa) from both strains, but with a glycosylation signal detected by anti-GAS antibodies present in the conjugation strain only (Figure 4.12c/d).

As there was no increase in NanA molecular weight, a proteinase K assay was used to confirm successful glycoprotein production in the conjugation strain. Proteinase K was added to the His purified conjugate sample, as well as to the components alone to digest the protein carrier such that subsequent loss of a polysaccharide signal should demonstrate attachment of polysaccharide to protein. Western blot analysis with anti-GAS antibodies shows the disappearance of the glycosylation signal for NanA-rpORha after Proteinase K treatment in the conjugate sample, confirming that rpORha is indeed conjugated to NanA (Figure 4.12e). Glycoconjugate components

alone showed that NanA is sensitive to proteinase K digest as expected, whereas rpORha polymers are unaffected unless conjugated to a susceptible protein carrier.

To further to demonstrate physical attachment of NanA to rpORha polymers, a sandwich ELISA was developed whereby anti-GAS antibodies were used to coat plates to capture and recognise rpORha, and anti-His conjugated to HRP antibodies used to detect the attached His-tagged NanA protein. Prior to application of NanA-rpORha conjugate and protein only control samples to the plate, the samples were further purified using size exclusion chromatography and processed through an endotoxin removal resin to remove interfering cell components. During the sandwich ELISA thorough washing of the plate was performed to also ensure no residual non-specific reactivity. NanA-rpORha conjugate samples showed a titratable signal diluted 1 in 2 across the plate according to a starting NanA protein concentration of 100 µg/ml (Figure 4.12f), as calculated by BCA assay (data not shown). This signal was not observed in the matched protein only sample containing NanA only at the same protein concentration. Therefore, proteinase K digestion and ELISA reactivity provides additional evidence of successful conjugation of rpORha to NanA carrier protein occurring inside *CLM24cedA::pgIB* cells due to PglB enzymatic activity.



**Figure 4.12: Conjugation of *S. pneumoniae* NanA carrier protein with rpORha in chromosomal PglB *E. coli* strain, CLM24cedA::pglB.**

1 litre cultures of CLM24cedA::pglB expressing either both NanA and rpORha together (conjugate), or NanA alone (protein only), in BHI media were induced with 0.2 % arabinose and 1 mM IPTG at mid-log phase and grown overnight.

**(A)** Schematic representation of cloned NanA fragment engineered to contain an N-terminal periplasmic DsbA secretion tag (purple), glycotags at both the N and C-terminus (green) and a x10 Histidine tag (yellow circle).

**(B)** Dots 1 – 3: protein only control strains, Dots 4 – 6: test strains containing *C. jejuni* locus; uninduced (1 / 4), 3 hrs (2 / 5), and 18 hrs post induction (3 / 6). Dotted periplasmic fractions were incubated with SBA lectin conjugated to biotin detected with Streptavidin IRDye® 680RD.

C - E Western blot analysis of conjugations. Red signal shows NanA carrier detected by Mouse anti-His developed with Goat anti-mouse IgG IRDye®-680RD. Green shows rpORha detected by Rabbit anti-GAS developed with Goat anti-rabbit IRDye®-800CW .

**(C)** Whole cell lysate (WCL) and periplasmic (P) fractions from conjugate strain.

**(D)** His purified WCLs to detect NanA and NanA-rpORha glycoprotein produced.

**(E)** Treatment of NanA-rpORha conjugate and components alone (NanA and rpORha) with (+) or without (-) proteinase K.

**(F)** Sandwich ELISA coated with anti-GAS antibodies, detected by anti-His antibodies conjugated to HRP for NanA protein only (red) and purified NanA-rpORha conjugate (blue).

Total NanA protein and rpORha concentrations were subsequently determined using BCA and Anthrone assays respectively within the purified NanA-rpORha conjugate sample. When considering component concentrations, a ratio of 2.5 : 1 (*w/w*) NanA to rpORha was calculated (Table 4.5). To convert to molar ratios rpORha was assumed to be 4.9 kDa, based on the size of GAS\_Rha extracted from mutant GAS cells in Chapter 3 Table 3.9, and calculated to NanA's molecular weight of 50 kDa. The molar ratio was calculated to be 1 : 4.05 (*m/m*) NanA to rpORha (Table 4.5). Such molarity calculations however are approximate as these include the concentration as a sum of both unconjugated and conjugated NanA protein, as well as assumes rpORha size to be 4.9 kDa without any experimental sizing data on isolated rpORha polymers.

**Table 4.5: Characteristics of NanA-rpORha bioconjugate.**

Table denotes rpORha concentration ( $\mu\text{g/ml}$ ) based on Anthrone, and NanA protein concentration ( $\mu\text{g/ml}$ ) based on BCA analysis. Weight / weight (*w/w*) and molar (*m/m*) ratios are stated as the preparation used for immunogenicity studies (Chapter 5).

rpORha Conc ( $\mu\text{g/ml}$ )	Total NanA Conc ( $\mu\text{g/ml}$ )	NanA-rpORha Ratio ( <i>w/w</i> )	NanA-rpORha Ratio ( <i>m/m</i> )
71.5	180	2.5 : 1	1 : 4.05 (NanA – 3.6 $\mu\text{M}$ rpORha – 14.6 $\mu\text{M}$ )

## 4.5. Discussion

Research undertaken in this Chapter aimed to utilise PglB dependent bioconjugation with *E. coli* modified glycoengineering hosts to produce GAS specific bioconjugate vaccines. *N*-linked PglB OST glycosylation was selected as an approach as it has been successful in the development and testing of vaccines against Gram-negative *Francisella tularensis*<sup>602,603</sup> and importantly in the context of this work, Gram-positive bacteria such as *S. pneumoniae* Serotype 4<sup>597,604</sup> and *Staphylococcus aureus*<sup>556</sup>.

Although holding great promise, the PglB approach is not suitable for all polysaccharides, for example those with glucose as the reducing end monosaccharide<sup>588,605</sup>, and a multitude of polysaccharides and saccharide linkages remain to be tested using this system. It was therefore of interest to assess PglB transfer capabilities of polysaccharide extracted from GAS cells in an 'one pot' *in vitro* approach as an initial proof of principle experiment. This was in addition to engineering *E. coli* to produce recombinant rhamnose variants, to test with model and GAS specific protein antigens in both 'one pot' and cellular glycosylation approaches. This work describes and evaluates the application of PGCT as an alternative GAS specific glycoconjugate manufacturing method.

### 4.5.1. Successful Conjugation Testing of Proteins with Polysaccharide Variants using PglB OST

An area initially investigated in the *N*-linked PGCT system was the specificity of the *C. jejuni* PglB OST to study the polysaccharide substrate. PglB OST enzyme's ability to recognise and transfer polysaccharide substrate is thought to be determined by the identity of the reducing end monosaccharide, as well as the linkage between the reducing end and the following monosaccharide. Only polysaccharides with a reducing end containing an acetamido group at the C2 position are currently thought to be permissive<sup>565,569</sup>. Initially to negate the need for labour intensive cloning, required for recombinant polysaccharide biosynthesis, glycosylation tests were first trialled using the 'one pot' method containing lipid linked polysaccharide extracted from GAS bacterial cells. GAC has a GlcNAc at the reducing end, suggesting it would be suitably recognised by PglB, however the following sugar in the chain, rhamnose, is attached to the GlcNAc by a  $\beta$ -1,4 linkage by the *gacB* gene in the cluster<sup>151</sup>. This particular linkage may not be recognised or transferred by PglB efficiently, as demonstrated by studies assessing transfer of a GlcNAc-GlcNAc disaccharide attached to either a non-native eukaryotic isoprene

lipid linker<sup>584</sup>, or a native und-PP lipid linker<sup>585,586</sup>. PglB transfer assessed in this work using the 'one pot' system, showed that GAC extracted from wildtype and mutant GAS cells, containing the  $\beta$ -1,4 rhamnose-GlcNAc linkage, may be compatible with PglB, however transfer seemed to occur at low efficiencies with model carrier protein AcrA, and was not confirmed beyond western blot analysis. Evidence also suggests that PglB can be forced to accept unfavourable linkages but only when providing substrate excess or forcing the components together in a concentrated one pot (Wren lab, personal communication).

Based on this data, it was of interest to investigate recombinant rhamnose polymer biosynthesis. Exploiting PGCT has a major advantage over native polysaccharide methods as it negates the need to work with pathogenic bacteria to generate glycoconjugate vaccine components. The gene clusters and polymer biosynthesis expression systems, devised by Dr. Dorfmüller at Dundee University, generated three distinct polymers for testing. This was based on three different construct combinations, including exogenous GAS specific *gac* and *S. mutans rpg* encoded rhamnose polymers alongside genes to produce either native or modified 'acceptor stem' structures. From these three, only one polymer, rpORha, could be successfully conjugated, encoded by *gac* genes containing the modified 'acceptor stem' for improved PglB recognition. Using the 'one pot' approach a GAS specific protein antigen SpyAD could be successfully conjugated to rpORha polymers, albeit at low reaction efficiencies. Therefore, in an attempt to improve reaction efficiency and product yield, the same recombinant polymer, rpORha was tested in a cellular glycosylation approach using a host strain containing PglB integrated onto the chromosome (*cedA::pglB*). Successful conjugation could be demonstrated with streptococcal specific carrier protein NanA, however despite occurring inside *E. coli* cells, bioconjugation product yields did not improve. Additionally, SpyAD carrier protein, despite showing promise in the 'one pot' approach could not be targeted using the cellular glycosylation approach (data not shown).

This Chapter has therefore shown that although PglB-mediated bioconjugation is possible with certain recombinant rhamnose and protein carrier combinations, the conditions tested in this work may have been sub-optimal for efficient glycoconjugate production. Factors influencing expression of bioconjugate components, for example host strains, expression systems, and culture conditions, as well as PglB OST efficiency, and protein carrier effects, are now discussed to address limitations within the thesis, aiming to improve future GAS specific bioconjugate production.



#### 4.5.2. Optimisation and Characterisation of Recombinant Rhamnose

##### Polymers

Prior to testing bioconjugation reactions, the developed recombinant rhamnose expression systems were tested in isolation to determine the suitability and availability of polysaccharide for PglB mediated attachment. This was of interest, given that the reliable recombinant polysaccharide production in the correct cellular location is one of the major rate limiting steps in efficient bioconjugation reactions<sup>588</sup>.

Recombinant polysaccharide expression was initially tested in CLM24 *E. coli* host strain, with favourable phenotypes for bioconjugation, such as disruption of pathways involved in O antigen biosynthesis to increase substrate availability for protein conjugation<sup>574</sup>. Whole cell lysates and periplasmic fractions from cells expressing recombinant polysaccharides showed low anti-GAS signal suggesting poor expression yields. Some signal was also observed in non-induced cells, likely through leaky expression of constructs, or endogenous rhamnose present in native polysaccharides in CLM24 *E. coli* strains cross reacting with anti-GAS antibodies.

Given these limitations, in an attempt to improve recombinant rhamnose yields and reduce non-specific rhamnose signal, alternative *E. coli* host strain, CS2775 was tested due to an inactive *rfaS* gene, preventing endogenous rhamnose modifications on surface LPS<sup>571,575–577</sup>. CS2775 *E. coli* cells gave superior yields of recombinant rhamnose variants, specifically rpMRha and rpORha, as well as reduced background in non-induced negative controls, enabling more clear analysis of recombinant expression levels. In addition to host strain background, the literature also shows that factors such as growth media, induction conditions, and nutrient and co-factor supplementation can improve and optimise bioconjugate component expression yields<sup>606</sup>. Broths rich in nutrients, namely 2YT and 2YTPG, were found to improve the production of rhamnose polymers in both strains compared to LB media, with an increase in cell density of *E. coli* grown in 2YT thought to be due to the breakdown of the media components leading to increased metabolite availability<sup>607</sup>.

In this study however, only rhamnose expression in isolation was optimised, but future work should focus on improving cell-based glycosylation as a whole. This is particularly relevant when considering the availability of certain components needed for polysaccharide biosynthesis such as glycosyltransferases and flippase enzyme activity, as well as lipid linkers (und-PP). These are required for transport of assembled polysaccharides to the correct cellular location, with their availability

shown to be a major rate limiting factor in recombinant polysaccharide expression<sup>578</sup>. Und-PP lipid linkers particularly must be recycled in the cell as the supply is limited, and multiple endogenous polysaccharides such as peptidoglycan rely on this linker for its biosynthesis<sup>608,609</sup>. After assembly recombinant polysaccharide polymers must be removed from und-PP, and added to either the cell surface in cases where WaaL is active, or to a carrier protein in the presence of PglB OST. Bioconjugation strains used in this work had inactivated *waaL*, and where there is no PglB source, limited rhamnose production may therefore be related to reduced und-PP availability. This theory was confirmed when rpORha was expressed in CS2775 *E. coli* with an active *waaL* gene to transfer polymers from und-PP to the cell surface showing superior recombinant rhamnose expression. Therefore, rpORha polymers are being synthesised efficiently in CS2775 *E. coli* cells, suggesting that it may be more beneficial to test expression of all the components together prior to condition optimisation.

Polysaccharide transfer from the membrane seems to be a rate limiting step in recombinant production linked to *C. jejuni* PglB activity levels. Therefore, bioconjugation may also have improved by introducing *pglB* into the chromosome of the CS2775 strain to enable its use in cellular glycosylation systems to improve the observed low bioconjugate yields. Exogenous polysaccharide biosynthesis may have also benefited from an inducible expression system, using promoter activity to fine tune and boost expression levels. This is likely to become more achievable in future expression and bioconjugation testing, owed to advances in molecular cloning techniques<sup>610-612</sup> and their application to PGCT.

#### 4.5.3. Suitability of Proteins as Bioconjugate Carriers Using Both Glycosylation Methods

GAS carrier proteins, MalE, SpyCEP and SpyAD were cloned into two plasmid systems compatible with the two bioconjugation approaches. To be compatible with polysaccharide plasmids for cell-based glycosylation the low copy number plasmid pEXT21 was used and proved difficult for cloning steps. For example, SpyCEP, which showed the clearest glycosylation signal using the 'one pot' system, could not be tested in the cell-based system due to cloning challenges. Therefore, alternative plasmids could be tested in future to streamline the process. All cloned carrier proteins in both constructs were successfully secreted into the periplasm where the *C. jejuni* PglB reaction occurs. MalE antigen looked to be stable, whereas SpyCEP and SpyAD in pEC415 used for 'one pot' analysis showed multiple bands, potentially from degradation, perhaps due to instability of the cloned peptide

fragment in the case of SpyCEP. Alternative SpyCEP fragments should be generated to obtain a more stable peptide as instability will have downstream repercussions for vaccine use.

The three GAS antigens showed differences in their suitability as protein carriers. SpyCEP and SpyAD showed some level of glycosylation with *C jejuni* model polysaccharide, through molecular weight increase. However, SpyCEP was limited by the presence of degradation products, and SpyAD showed a less clear glycosylation signal due to its large molecular weight, failing to show distinct separation between glycosylated and unglycosylated protein forms. However, this could be rectified by further separation or using a polysaccharide with a higher molecular weight. MalE showed low glycosylation compared to the other antigens, suggesting its protein folding and structure may be less optimal for PglB transfer. To improve this, additional glycotags could be added to extend away from the protein, or protein surface loops could be targeted.

'One pot' testing with GAS specific recombinant rhamnose demonstrated no clear glycosylation, even for model carrier protein AcrA. This was unlikely to be due to polysaccharide levels due to optimised expression conditions and increasing quantities of all lipid-linked recombinant polysaccharides having minimal effects on bioconjugation. Interestingly, an exception was the reaction containing SpyAD with rpORha, demonstrating successful but low conjugation efficiency. SpyAD, the largest protein tested in this study, may have therefore provided a folded structure with greater accessibility for binding to PglB's active site, enabling the attachment of polysaccharide substrate containing a reducing end optimised for PglB recognition.

Although useful for initial proof of concept studies, in this instance, small scale 'one pot' reaction volumes containing finite amounts of the required three components may have contributed to poor bioconjugate yields. Therefore, to determine whether glycosylation yields could be improved, particularly for SpyAD protein testing, the cell-based glycosylation approach was adopted. This was aiming to address 'one pot' limitations, where components are artificially introduced in close proximity in a single tube, compared to cell-based methods where protein carriers, polysaccharide and PglB OST are being continually synthesised and exported to the same cellular location, enabling spatial association.

Prior to testing of GAS antigens as carriers, CLM24*cedA::pglB* host cells containing chromosomal PglB were transformed with rpORha plasmids and NanA as a streptococcus specific carrier protein. Scaling up of cultures was necessary to

observe evidence of the successful decoration of NanA with rpORha polymers by ELISA and proteinase K digestion, suggesting low product yield and poor reaction efficiency. Western blot analysis failed to show conclusive bioconjugation of NanA by a protein molecular weight increase. An absence of a protein size shift has been observed with transfer of other polysaccharides such as *Salmonella* polysaccharide (Elizabeth Atkins, personal communication), and may be challenging to observe when small polysaccharide polymers, such as rpORha, are attached to a single protein molecule in a mixture containing a large proportion of unconjugated protein. Separation by SDS-PAGE may also not show polysaccharides at their true size due to physicochemical properties. This is observed for native GAC extracted from GAS cells shown to be ~ 9.8 kDa, corresponding to ~ 18 repeating units, yet blots show it as a high molecular weight smear. With this in mind SDS-PAGE separation may not have been accurate for recombinant rhamnose polymer size, and rpORha polymers extracted from KDO could not be experimentally characterised by NMR nor other methods such as SEC-MALS due to purity issues. Further work is therefore required to obtain intact recombinant rhamnose samples with a high enough purity and yield for physicochemical analysis and accurate sizing.

Despite this, low polysaccharide polymerisation was suspected, leading to attachment of shorter rpORha polymers, which would not provide a detectable separation in molecular weight from unconjugated protein. Since *C. jejuni* heptasaccharide showed a minor yet detectable protein carrier size shift, it suggests that the attached rpORha polymer consists of a single repeating unit or a disaccharide. Interestingly, bioconjugation data generated by Dr. Dorfmueller showed a distinct NanA protein specific molecular weight increase via anti-His antibody detection following rpORha attachment. This was likely due to longer chains, increasing the final molecular weight to a detectable level through a size shift. It is unclear why this could not be reproduced in this work; however, conjugation efficiency seems to deteriorate with storage of transformed glycoengineering strains, therefore fresh transformations may be necessary with each test. This is a limitation to cell-based bioconjugation and would be helped by improved expression systems in optimal strain and growth conditions.

For NanA-rpORha analysis, ELISA provided a method to show attachment whereby the polysaccharide component was captured, and the protein component detected. This method showed NanA decorated with rpORha but required removal of endotoxin contaminants to obtain an acceptable assay background level to detect such glycosylation signal. This was despite the generation of samples in an  $\Delta/pXM$

strain background, preventing KDO myristate attachment reducing LPS toxicity<sup>594,613,614</sup>. Therefore, care must be taken to prevent LPS endotoxin and other native polysaccharides such as O antigens, from providing a source of contamination when detecting and analysing glycoconjugate identity and for immunogenicity testing. Contaminating host cell components were also suspected to have interfered with NanA-rpORha bioconjugate component quantification such as non-specific cross reactive hexose polysaccharides detected in anthrone assays.

NanA-rpORha's calculated glycoconjugate ratios were approximate based on total protein in both glycosylated and unglycosylated forms, and appeared to be largely inaccurate when considering glycotag availability. This in part was due to the selected His-tag purification method, which retains unconjugated protein material in the formulation. Optimal separation conditions based on anionic exchange or size exclusion chromatography may have purified samples further for more accurate component quantification, however conditions were not achieved in this work due to NanA-rpORha bioconjugate properties such as cell contaminants interfering with analysis. Future studies would also benefit from improving glycosylation to maximise bioconjugate yield so that separation of conjugated protein can be achieved based on size exclusion methods.

When testing MalE and SpyAD GAS antigens with *C. jejuni* heptasaccharide and chromosomal PglB, there was some level of glycosylation through molecular weight increases detected via anti-His antibodies, and lectin staining specific to the heptasaccharide. This was still very limited for MalE consistent with 'one pot' reactions indicating lower glycosylation efficiency compared to SpyAD, which showed more distinct attachment with the largest molecular weight increase. However, it was unclear using these detection methods whether glycosylation of both engineered glycotags at either terminus was achieved in the cell-based system due to the absence of a double shift in the protein's molecular weight. Therefore, it remains unclear whether there is merit in the addition of a second glycotag to positively effect glycoprotein production in this work and would require further investigation.

MalE and SpyAD were tested for attachment to rpORha polymers using chromosomal PglB OST. Despite a number of conditions tested, such as variable inducer concentrations, expression duration, and temperature reduction from 37 to 30 °C to slow bacterial growth and negate any potential cellular toxicity effects, neither GAS antigen could be decorated with rpORha polymers. Given the

successful attachment of *C. jejuni* heptasaccharide to SpyAD using the 'one pot', the glycotag availability and PglB function do not appear to be problematic in this instance, and rather unsuccessful bioconjugation may relate to different protein specific effects. For example, MalE showed poor glycosylation with *C. jejuni* heptasaccharide, suggesting there may be protein steric limitations effecting enzyme or polysaccharide attachment. There may also be differences in the physiochemical properties of expressed *C. jejuni* heptasaccharide and rpORha polymers such as size, structure, and charge which are not compatible with the particular protein's physiochemical properties.

#### 4.5.4. Characterising and Troubleshooting Bioconjugate Vaccine Production

Overall bioconjugate component levels may have been a contributing factor. A major constraint when using the cell-based glycosylation approach is the restriction of using plasmids with compatible origins of replication and selective antibiotic resistance markers. In the current system, there was three origins of replication (pBR322, p15A and IncW) and resistance markers (ampicillin, erythromycin and spectinomycin), limiting the use of alternative plasmid systems, which may have boosted expression levels. A method to overcome this is to rationally design the host strain to integrate components onto the chromosome to streamline bioengineering pathways and optimise the balance between expression and cell metabolic burden. This has been achieved in some instances by chromosomal integration of glycosyltransferase genes, streamlining recombinant polysaccharide production, and lowering expression levels<sup>615</sup>. This underpinned the use of *cedA::pglB* to facilitate conjugation whilst reducing cell burden caused by high expression of the multi-transmembrane PglB enzyme<sup>478</sup>. This is because higher PglB expression does not necessarily correlate with improved glycosylation efficiency<sup>596</sup>. In fact, chromosomal PglB integration has been shown to enhance *S. pneumoniae* serotype 4 capsule attachment to AcrA<sup>604,615,616</sup>.

Despite this, rpORha expression levels may have been negatively affected by the use of *cedA::pglB E. coli* as this strain contains endogenous rhamnose polymers. Future studies may therefore benefit from using the CS2775 *E. coli* strain, which showed superior rpORha expression, for OST integration, as well as high throughput metabolite screening to achieve optimal growth conditions to improve yields. Generally, host strain compatibility and metabolite screening are areas which require optimisation, with increased emphasis placed on rationally altering and improving bacterial host cells for bioconjugate production as the PGCT field develops. A good example of this is the development of 11 *E. coli* strains distinct in

terms of components either rationally added or removed by CRISPR technology to the parent *E. coli* W3110 strain<sup>578</sup>. This particular study is a landmark for PGCT, offering a strain catalogue for the expression of several different recombinant polysaccharides, as well as carrier proteins, and linking OST enzyme combinations. Therefore, given that there is currently no consensus on *E. coli* strain or associated growth conditions, and that the optimal phenotypic background and expression was unlikely achieved in this work, the development and use of such strains will be important tools in the future generation of GAS specific bioconjugates.

In addition to strain background, the compatibility of the vaccine components and OST enzyme is essential. One approach to improve OST enzyme recognition is to modify native *C. jejuni* PglB, which has demonstrated great promise in a recent patent from GSK (Patent WO 2021/028303 A1). The group showed a 4,343-fold increase in *S. pneumoniae* serotype 8 capsular polysaccharide conjugated to ExoA model carrier protein by point mutating *C. jejuni* PglB in seven distinct mutation rounds. The reducing end linkage of glucose was  $\beta$ -1,4, similar to native GAC and rpWTRha GlcNAc linkage in this work, and may prove useful if the mutations within PglB improves recognition of monosaccharide linkages. Such increased glycosylation efficiency through *C. jejuni* PglB engineering has also been observed in other studies broadening substrate specificity<sup>617,618</sup>.

Additionally, an alternative approach to improve OST and substrate compatibility is to investigate different OST enzymes for *N*-linked glycosylation, such as PglB paralogues and orthologues, which can display increased substrate specificity<sup>551,564,619,620</sup>. Beyond the scope and timescale of this work, *O*-linked OST enzymes such as PglL and PglS are also being investigated in certain instances, usually for increased reducing end promiscuity towards galactose<sup>621,622</sup> and glucose<sup>605,623</sup>, in addition to cytoplasmic *N*-linked bacterial glycosylation pathways<sup>624–627</sup>.

## 4.6. Summary

The PglB OST mediated *N*-linked glycosylation system holds great promise for the future production of glycoconjugate vaccines. The technology is still in its infancy and the full potential of bioconjugation vaccine production has yet to be realised. Findings in this Chapter provide an example of recombinant rhamnose production, in addition to novel GAS antigens engineered to be compatible with two approaches testing bioconjugation technology in modified *E. coli* glycoengineering strains. Glycosylation was observed for SpyAD with recombinant rhamnose using the 'one

pot' system and for NanA using the cell-based system, but further work will validate the use of rpORha with species specific carrier proteins for 'double hit' glycoconjugates. Other tested GAS proteins were also generally able to be conjugated to model *C. jejuni* polysaccharide but highlighted the need to study availability of glycotags as MalE showed only partial size shifts. However, the combination of selected proteins and recombinant rhamnose variants for PglB recognition and transfer had not previously been investigated, adding knowledge to the field with data useful to guide subsequent compatibility testing.

Overall, it appears that the pivotal step towards *N*-linked bioconjugation success will be to fully understand PglB enzyme requirements, specifically recombinant polysaccharide structures and carrier protein compatibility inside modified and streamlined *E. coli* host cells. It is hoped as technological understanding and PglB OST requirements develop, more appropriate systems and compatible enzymes can be used for increased promiscuity and transfer of problematic polysaccharides such as GAC. It will also be beneficial to fully understand the intricate and complex relationships between exogenous and native pathways inside glycoengineering *E. coli* strains to optimise expression levels, likely to be tailored to the selected recombinant polysaccharide, carrier protein and OST enzyme for bioconjugation.



## CHAPTER 5

# Immunological Analysis of Chemical Glycoconjugate and Bioconjugate Vaccines

### 5.1. Introduction

Glycoconjugate manufacturing design and production influences a vaccine's immunogenicity and presumed efficacy to prevent disease. Glycoconjugate vaccines are diverse in their polysaccharide and protein components, as well as the variety of conjugation methods used to attach them together. This results in a variety of molecular structures, influencing exposed immunogenic epitopes, and how antigens are displayed and processed by the immune system. Despite this, the design and manufacture of glycoprotein antigens has often been empirical, associated with variation in manufacturing processes, yielding incompletely characterised antigens and variable immunogenicity profiles<sup>628</sup>.

Glycoconjugate vaccines are processed by the immune system through B and T cell interactions to induce adaptive immune responses (Chapter 1 Figure 1.3). It is thought that a peptide fragment is presented by B cells to encourage helper T cells engagement<sup>458,629</sup>, however, a new model has recently been proposed, suggesting engagement of polysaccharide specific T cells, termed Tcarb cells, through B cell presentation of glycopeptides<sup>535,630,631</sup>. It is hypothesised that polysaccharide antigens in their native form do not bind MHC class II on B cell surfaces. However, attachment to a carrier protein allows the presentation of the polysaccharide antigen through binding of the peptide portion to MHC class II following processing within antigen presenting cells (APC) endolysosomes. This emphasises the importance of protein carriers in establishing long-term polysaccharide directed humoral responses. However, the exact mechanism and intricate relationship between immune responses and antigen processing, is currently inadequately understood. It is believed that polysaccharide and carrier protein properties, related to the selected conjugation method, are important factors to consider during development, with unfavourable properties attributable to clinical underperformance such as lower than expected antibody titres<sup>632</sup>.

#### 5.1.1. Polysaccharide Antigens Relating to Glycoconjugate Immunogenicity

Polysaccharide identity, size, structure, and conformation alter glycoconjugate vaccine immune responses. For example, in some instances increased

polysaccharide chain lengths are thought to improve vaccine polysaccharide to protein ratios<sup>458</sup>, in addition to preserving repetitive structural immunogenic B cell epitopes<sup>633</sup>. However, the relationship between polysaccharide chain length and resulting immunogenicity is not clear, with some protective epitopes thought to be associated with branched structures for certain polysaccharides<sup>634–640</sup>. Additionally, glycoconjugates made with both small oligosaccharides as well as large polysaccharides have been shown to be successful in clinical trials, as well as efficacy testing of licenced vaccines<sup>628</sup>. Therefore, although polysaccharide chain length and structure influences adaptive immune responses, specifically antigen presentation by B cells<sup>535</sup>, the optimum number of repeat units and protective conformations required to achieve long lasting antibody responses is currently unknown. This is likely something that must be determined in the context of each specific glycoconjugate vaccine.

GAC, in comparison to other streptococcal polysaccharides, has a simple structure composed of only two monosaccharide constituents, rhamnose and GlcNAc, in a fairly linear unbranched structure (Chapter 1 Figure 1.1). GAC is also reasonably small in size and polysaccharide chain length. Chapter 3 Table 3.9 showed that extracted wildtype GAC, included in chemical conjugation reactions, is 9.8 kDa in size, correlating to an average of ~ 18 repeating units within the chain to stimulate the immune system. In comparison, the bioconjugate vaccine manufactured in Chapter 4 investigated recombinant rhamnose as an antigen encoded by the *gac* gene cluster. This includes *gacH*, a gene hypothesised to be involved in polysaccharide chain length regulation (Dr. Dorfmueller, personal communication), and may therefore produce polysaccharide chains of a similar length to native GAC. However, as discussed, this could not be experimentally determined, and recombinant rhamnose conformation and size may be different, as was seen for rhamnose polymers extracted from the *gacI* GAS mutant produced in Chapter 3.

#### 5.1.2. Protein Antigens Relating to Glycoconjugate Immunogenicity

Historically, when using classical carrier proteins such as TT or CRM<sub>197</sub>, the carriers of choice for most currently licensed glycoconjugate vaccines, the polysaccharide antigen holds greater immunogenic importance for vaccine induced protection<sup>464,638,641</sup>. Recently, due to a limited panel of classical protein carriers, immunogenic suppression of anti-polysaccharide responses has been observed after repeated immunisations with vaccines containing the same classical carrier proteins<sup>642,643</sup>. This is mediated by inhibition of polysaccharide antibody responses by carrier-specific B cells and suppressor T cells due to pre-existing carrier protein

immunity<sup>475</sup>. This is a concern for both multivalent vaccines containing the same carrier protein, as well as for co-administered vaccines in the infant immunisation schedule that also use the same carrier protein in different preparations<sup>487</sup>.

Recently, more interest has been directed towards investigation of whether it is beneficial to include protein antigens from the same pathogen to generate 'double hit' glycoconjugate vaccines. It is hoped that such protein carriers can contribute to protection through its own independent immunological properties, in addition to supporting the immune response generated against the attached polysaccharide antigen. For this to occur it is essential to preserve and understand the selected protein antigens conformation to ensure optimal T cell presentation and maintenance of B cell epitopes. Therefore approaches such as structural and reverse vaccinology are becoming more common during the selection of protein carriers in 'double hit' preparations<sup>395,644</sup>.

'Double hit' glycoconjugate vaccines have been experimentally tested for both chemical and biological conjugation methods<sup>411,412,447,597</sup>. Specifically, GAS protein antigens SpyAD, SpyCEP, SLO, and C5a peptidase have been investigated in several recent studies assessing chemical conjugation methods<sup>411,412</sup>. Generally these studies have found GAS antigens to be good carriers for GAC, inducing GAC specific antibodies<sup>411,447</sup>, and in some instances anti-protein antibodies<sup>447</sup>. Some 'double hit' vaccines have also shown protection in mice against GAS challenge when a combination of glycoconjugate and protein antigens are immunised together<sup>412</sup>. In addition to chemically derived glycoconjugates, 'double hit' bioconjugates have also been investigated for some bacterial species. Such studies have yet to test GAS specific antigens, however, studies have demonstrated strong opsonic antibody responses following immunisation with a *S. pneumoniae* bioconjugate vaccine containing protein carrier NanA conjugated to serotype 4 capsular polysaccharide<sup>597</sup>. NanA was also used in this work as a control carrier protein for rhamnose polysaccharide attachment (Chapter 4 section 4.4.5).

### 5.1.3. Manufacturing Methods Relating to Glycoconjugate Immunogenicity

Polysaccharide and protein carrier properties place constraint on the selected conjugation approach to covalently attach them together. A plethora of conjugation approaches are available; however, such variety adds a layer of complexity when assessing and comparing glycoconjugate immunogenicity. For example, differences in preparations can be observed in the site of attachment on carrier proteins, as well as the presence or absence of linker within the product. Different conjugation

methods result in varying degrees of protein decoration with polysaccharide antigens, and have in some instances yielded unexpected immunogenicity results<sup>632,645</sup>. This highlights that the selection of an optimal conjugation approach is not obvious for a given combination of vaccine components.

Semi-selective chemical conjugation approaches utilising terminal activation of the polysaccharide chain is possible<sup>646</sup>, and is similar to biological conjugation enabling single point enzymatic attachment of polysaccharides using the terminal reducing end monosaccharide<sup>569,647</sup>. However, specific point attachment of a protein using chemical linking agents is often more complicated, traditionally targeting many amino acid residues on protein surfaces<sup>458</sup>, whereas bioconjugate production allows more precise control over carrier protein attachment sites, and therefore distribution of glycoconjugate size. This is due to introduction of specific glycotags to known protein regions to control the level and position of polysaccharide attachment<sup>564,648</sup>. Despite this, there is currently no consensus nor threshold of protein carrier decoration with polysaccharide reported to induce strong anti-polysaccharide immune responses.

This Chapter will investigate whether both methods, chemical and biological conjugation, are able to generate immunogenic glycoconjugate vaccines. This is through testing of IgG antibody levels and functionality with GAS cell binding and GAS cell killing by opsonophagocytosis. Additionally, the immunogenicity of chemical glycoconjugate SpyAD-GAC containing two GAS antigens will be compared to TT-GAC glycoconjugate manufactured using the same methodology but with a classical protein carrier (Chapter 3). The level and functionality of anti-polysaccharide and anti-carrier protein IgG responses induced by the two glycoconjugates will be assessed, determining whether there is any additional protectivity capability when including a GAS antigen as a carrier protein.

## 5.2. Aim, Hypothesis and Objectives

Three chemical glycoconjugate vaccines were successfully generated; GAS protein antigen SpyAD conjugated to GAC as a 'double hit' GAS vaccine, and classical carrier protein Tetanus Toxoid (TT) conjugated to either GAC or GAS\_Rha (Chapter 3). One biological glycoconjugate vaccine was generated with NanA *Streptococcal* carrier protein conjugated to recombinant polyrhamnose (rpORha) (Chapter 4).

**Aim:** Investigate the immunogenicity of the resulting glycoconjugate vaccine candidates to determine a preferred glycoconjugate manufacturing method.

### **Hypothesis:**

- Chemical and biological derived glycoconjugate candidates will have different immunogenicity profiles, specifically antibody levels and functionality.

### **Objectives:**

- 1) Immunise mice with glycoconjugate preparations manufactured using either chemical or biological conjugation methodologies and different carrier proteins. Test immune sera by ELISA to determine anti-polysaccharide and anti-protein IgG antibody titres where appropriate.
- 2) Determine if antibodies are functional for binding to GAS cell surfaces or killing GAS cells by opsonophagocytosis assay.

## 5.3. Methods

### 5.3.1. Animals and Immunisations

As stated previously, animal studies were conducted according to the UK Home Office regulations and were approved by the local ethics committee. Groups of 5 female 8-10 week-old BALB/c mice (Charles River Laboratories, Saffron Walden, UK) were immunised subcutaneously with various conjugates containing 2 µg of polysaccharide (GAC / GAS\_Rha / recombinant rhamnose - rpORha) per dose. Vaccines were adsorbed onto aluminium hydroxide (Alum) adjuvant (Alhydrogel® 2 %, InvivoGen) at 100 µg/dose concentration overnight at 4 °C with agitation. Each mouse received 3 doses at 14 day intervals, with a group receiving PBS with Alum only as a negative control. Serum was collected 13 days after the initial immunisation (test bleed 1), and before the third immunisation on day 27 (test bleed 2), followed by terminal bleeds collected two weeks after the final immunisation (day 42).

### 5.3.2. Immunoassays

#### 5.3.2.1. *ELISA to Measure IgG Antibody Titres Against Polysaccharide and Protein Portions of Glycoconjugate Vaccines*

MaxiSORP 96-well plates (Thermo Fisher Scientific) were coated with 50 µl/well 0.5 µg/ml of purified polysaccharide (GAC or GAS\_Rha) or recombinant protein (TT or SpyAD) diluted in carbonate buffer overnight at 4 °C. Serum dilutions were prepared in assay diluent before being added to the plate and serially diluted. Test sera (day 13 / 27) were routinely diluted 1 : 50 and terminal sera diluted 1 : 100 before 2-fold or 3-fold titrations incubated at 37 °C for 2 hours. Plates were developed by addition of 100 µl/well of anti-mouse IgG-HRP (Sigma-Aldrich™) diluted 1 : 20,000 in assay diluent incubated at 37 °C for 1 hour before detection and data analysis as stated in general method sections 2.4.2 and 2.6.1.

#### 5.3.2.2. *Immunofluorescence Staining and Flow Cytometry (FAC) Analysis of Glycoconjugate Sera*

GAS cultures were grown to log-phase OD<sub>600nm</sub> 0.4 and cells harvested for 10 minutes at 4,000 x g. GAS cells were resuspended in 1 ml PBS with 10 % goat serum and incubated for 20 minutes at room temperature. GAS cells were harvested through a 4000 x g spin for 10 minutes before resuspension in 100 µl of staining buffer (PBS supplemented with 0.1% BSA and 10% goat serum). Washed cells were harvested again before incubation with immune mouse serum diluted 1 : 4 in staining buffer for 1 hour at 4 °C. GAS cells were washed with PBS

supplemented with 0.1 % BSA before incubation with 100 µl Goat anti-Mouse-FITC (Thermo Fisher Scientific) diluted 1 : 3,000 in PBS supplemented with 0.1% BSA for 45 minutes at 4 °C. GAS cells were washed with PBS supplemented with 0.1% BSA before suspension in 1 ml of fixative buffer (2% formaldehyde and 50% PBS). Sample data was collected on a BD FACS Canto II flow cytometer, with FlowJo Software used for data analysis.

### 5.3.2.3. *Opsonophagocytosis Assay (OPA) to Assess GAS Killing by Immune Sera*

To test opsonophagocytic activity of immune sera against GAS cells, HL60 cells (ATCC CCL-240) were grown in RPMI 1640 media (Sigma-Aldrich™), 10% FBS (Gibco®), 1% glutamine (Sigma-Aldrich™) and 1% penicillin–streptomycin (Sigma-Aldrich™). For the assay, HL60 cells were differentiated in 100 ml culture medium supplemented with 0.8% dimethylformamide (DMF) (Sigma-Aldrich™) at  $4 \times 10^5$  cells/ml for 6 days without penicillin-streptomycin. Frozen GAS cultures were tested with differentiated HL60 cells in the presence of 2 % baby rabbit serum as a source of complement (Pel-Freez Biologicals) at a ratio of 1 : 400, GAS : HL60 cells. A control sample received heat inactivated baby rabbit serum. GAS frozen cultures were thawed, washed, and diluted to calculate input dilution in sterile opsonisation buffer B (OBB) (1 x Hanks' Balanced Salt Solution (HBSS) with salts (Gibco®), 0.1% gelatin and 5% FBS), to yield ~ 60 – 100 CFU / 5 µl spot. Ten µl of diluted GAS was added to 20 µl of immune sera dilutions or OBB in 96-well round-bottom plates incubated for 30 minutes at room temperature shaking at 700 rpm. During the incubation, differentiated HL60 cells were harvested at 350 x g for 5 minutes and the media supernatant removed. The pelleted cells were washed in 1 x HBSS without salts (Gibco®) harvested by a 5-minute 350 x g spin followed by a subsequent wash with 1 x HBSS with salts (Gibco®) before being harvested again by centrifugation. Pelleted HL60 cells were resuspended in OBB to  $1 \times 10^7$  cells/ml and mixed with 2% complement or heat-inactivated complement at a ratio of 4 : 1, cells to complement. 50 µl of the cell/complement mixture was then added to opsonised GAS before a 90-minute incubation at 37 °C, 5% CO<sub>2</sub> with shaking at 200 rpm. Incubated plates were then placed on ice for 20 minutes to stop phagocytosis before the cultures were resuspended and 5 µl was drip-plated on THY agar plates for visual enumeration of GAS CFUs.

## 5.4. Results

### 5.4.1. Antibody Titres Generated from Glycoconjugate Vaccine

#### Immunisation

Three chemically derived glycoconjugates manufactured in Chapter 3; TT-GAC, SpyAD-GAC and TT-GAS\_Rha, as well as one biological glycoconjugate, NanA-rpORha manufactured in Chapter 4 were tested for immunogenicity (Figure 5.1). Two test bleeds (days 13 and 27) as well as a terminal bleed (day 42) were collected to determine IgG antibody production in test groups. Antiserum responses were compared to those of mice which received GAC-CRM<sub>197</sub> (as a positive control), PBS with Alum, or polysaccharide antigens alone (as negative controls), as described in Table 5.1.

**Table 5.1: GAS specific glycoconjugate vaccine groups and antigen components.**

The table denotes mice groupings (n = 5/group), polysaccharide (PS) and protein identity and description.

Group	PS Identity	Carrier Protein	Description
1	N/A	N/A	Adjuvant Only
2	GAC WT*	CRM <sub>197</sub>	Positive Control
3	GAC WT	N/A	PS only
4	GAS_Rha	N/A	PS only
5	GAC WT	TT	Chemical Conjugate
6	GAC WT	SpyAD	Chemical Conjugate
7	GAS_Rha	TT	Chemical Conjugate
8	rpORha	NanA	Biological Conjugate

N/A = not applicable,

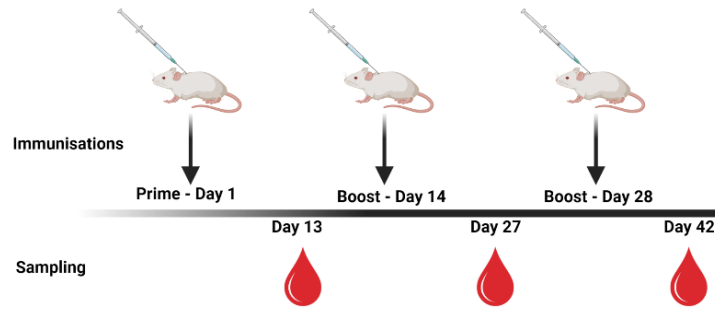
WT = wildtype,

CRM<sub>197</sub> = genetically detoxified form of diphtheria toxin,

TT = Tetanus Toxoid,

\* = Wildtype GAC extracted used a different method to those described in this thesis, produced externally.





**Figure 5.1: Subcutaneous immunisation schedule of BALB/c mice for GAS specific glycoconjugate immunogenicity study.**

Schematic representation of immunisation schedule of mice receiving three doses of 2 µg of polysaccharide conjugated to protein carriers adsorbed onto Alum adjuvant – days 1, 14 and 28 (n = 5 mice/group). Day 13 = test sera 1 (one immunisation), day 27 = test sera 2 (two immunisations), and day 42 = terminal sera (three immunisations).

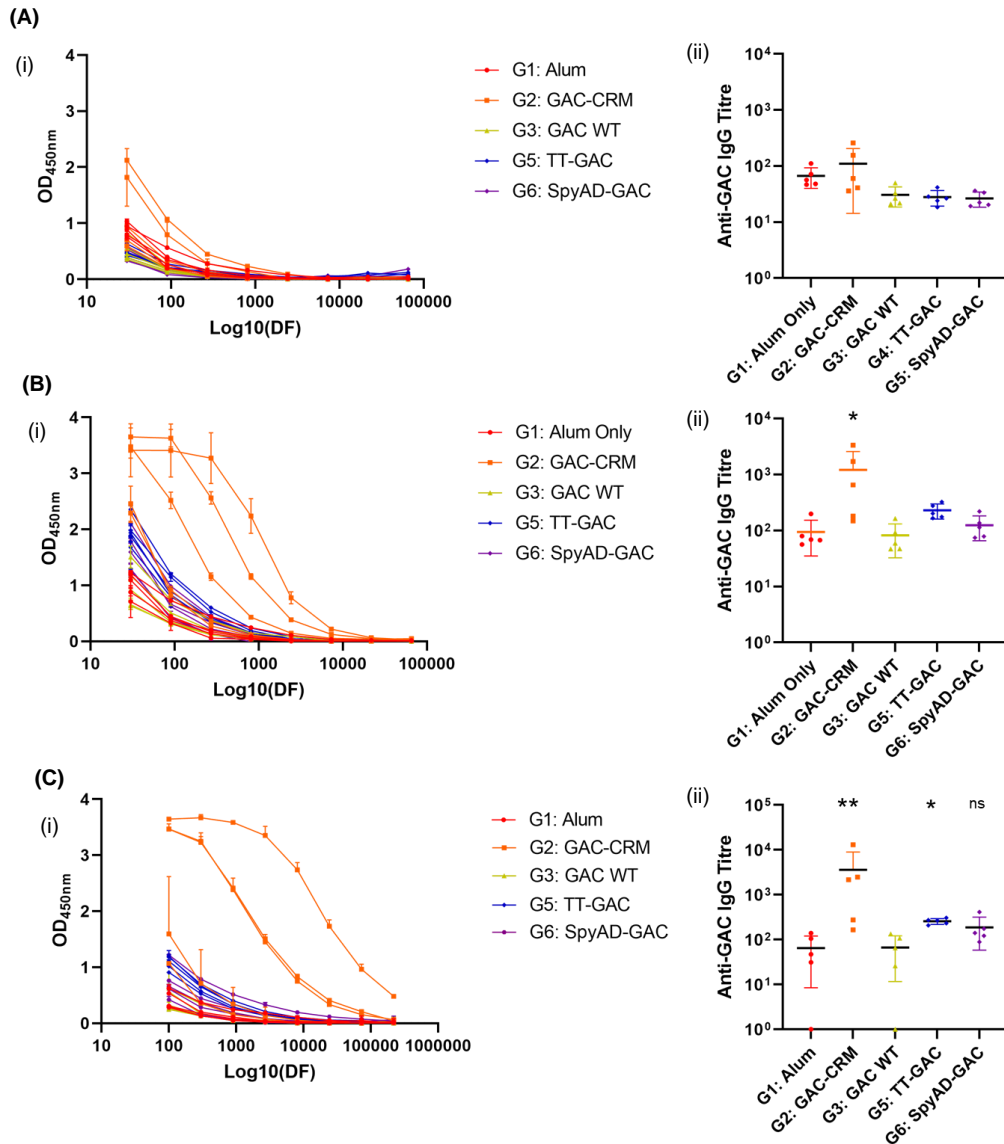
Original figure created with BioRender.com.

#### 5.4.1.1. Anti-Polysaccharide IgG Levels

In order to test glycoconjugate vaccine immunogenicity WT GAC was used to coat ELISA plates as the most clinically relevant of the polysaccharide preparations, observed during natural GAS infection. Background signal was established from sera obtained from mice immunised with Alum only, with geometric mean titre (GMT) = 38 (range: 1 - 139). As expected, polysaccharide alone did not induce IgG antibodies above the background signal in either test bleeds (GMT = 31 (range: 21 - 50), and GMT = 82 (range: 46 - 164)) or terminal bleed sera (GMT = 66 (range: 1 - 107)) (Figure 5.2a - c – yellow dot). The group of mice immunised with the positive control (GAC-CRM<sub>197</sub>) showed high variability in the IgG antibody level, with three mice responding well with significantly elevated anti-GAC IgG responses after the second immunisation compared to the adjuvant baseline (GMT = 1,213 (range: 148 – 3,360) (p = 0.0351) (Figure 5.2b – orange dot). By the end of the experiment at day 42 anti-GAC IgG responses were significantly increased in mice immunised with GAC-CRM<sub>197</sub> glycoconjugate vaccine compared to the adjuvant baseline (GMT = 50,722 (range: 222 – 13,148)) (p = 0.0047) (Figure 5.2c – orange dot).

Mice primed with one dose of either TT-GAC or SpyAD-GAC glycoconjugates showed no response above the background signal for Alum adjuvant and polysaccharide only groups (Figure 5.2a – blue and purple dots) with GMT of anti-GAC IgG antibodies for TT-GAC; 28 (range: 18 - 41) and SpyAD-GAC; 26 (range: 19 - 36). After boosting 2 weeks later with a second dose of glycoconjugate vaccine,

although there was a slight increase in the anti-GAC IgG antibody responses for both vaccines, this was still not significantly above baseline for either glycoconjugate (Figure 5.2b – blue and purple dots). However, terminal sera collected after a third immunisation (day 42) showed that mice immunised with TT-GAC had significantly elevated anti-GAC IgG titres compared to Alum background (GMT = 257 (range: 208 – 304),  $p = 0.0373$ ) (Figure 5.2c – blue dots). SpyAD-GAC did also show a slight increase in anti-GAC IgG titres, similar to titres from TT-GAC immunisation, however, this was found to not be statistically significant compared to Alum only background (GMT = 187 (range: 88 - 411),  $p = 0.4670$ ) (Figure 5.2c – purple dots). This is despite one mouse generating higher anti-GAC antibody titres than mice immunised with TT-GAC. In fact, following Mann-Whitney analysis no statistical difference was found between anti-GAC antibody titres generated by three doses of TT-GAC and SpyAD-GAC glycoconjugates.



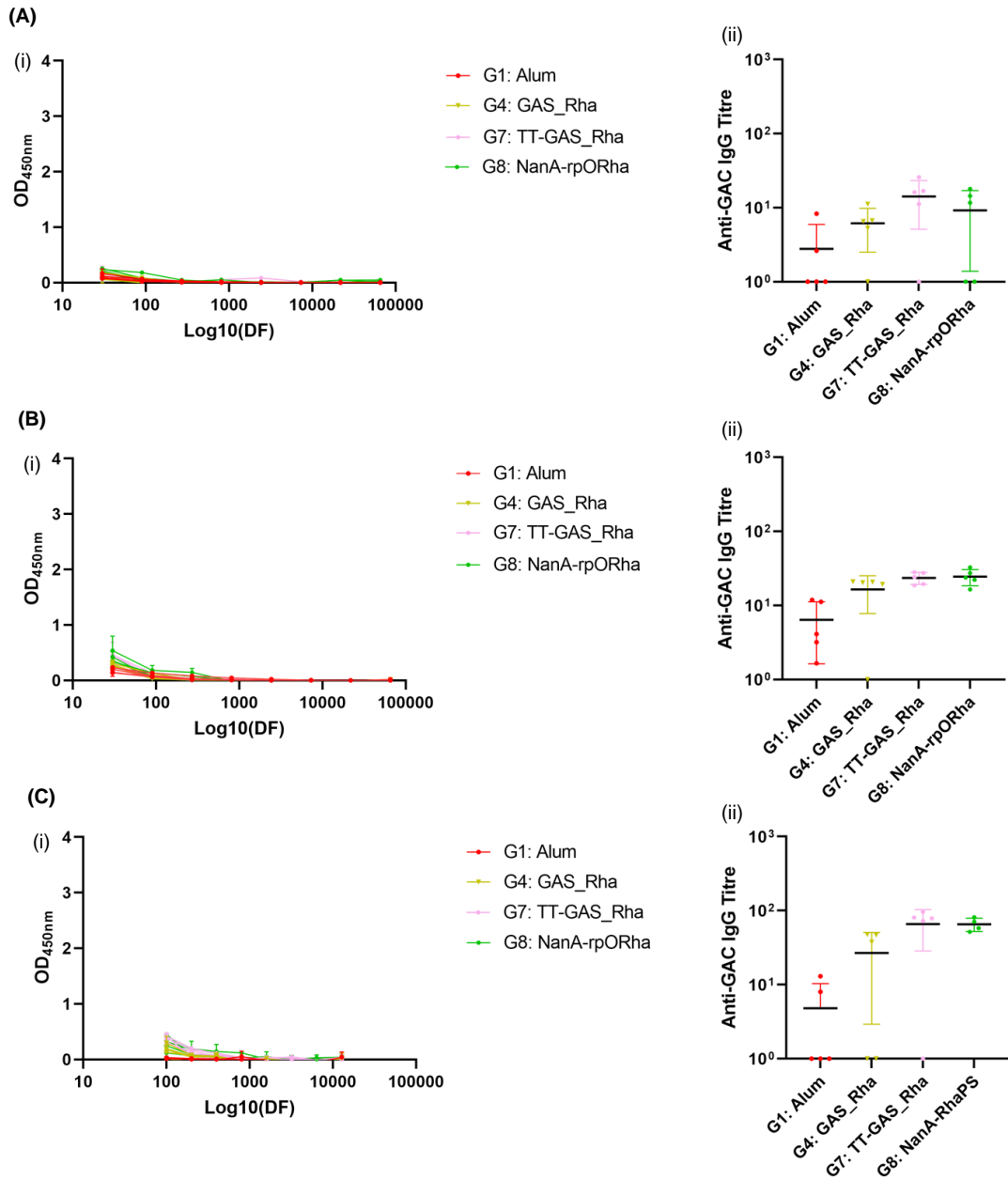
**Figure 5.2: ELISA analysis of anti-GAC IgG antibody titres in day 13, 27 and 42 antisera of mice immunised with wildtype GAC containing glycoconjugates.**

BALB/c mice were immunised with 3 doses of polysaccharide or glycoconjugate vaccines, positive control GAC-CRM (group 2 - orange), WT GAC (group 3 - yellow), TT-GAC (group 4 - blue) and SpyAD-GAC (group 6 - purple). Sera were analysed for anti-GAC response by ELISA. Antiserum titration curves (i) and individual IgG titre and Geometric Mean Titres of the group (GMT; horizontal bars with vertical standard deviation error bars) (ii) for individual mice were tested and calculated for (A) test bleed 1, day 13, (B) test bleed 2, day 27 and (C) terminal bleed, day 48 on plates coated with purified WT GAC polysaccharide. DF = serum dilution factor.

Significance compared to group 1 (Alum) determined by Kruskal-Wallis test, \*  $p < 0.05$ , and \*\*  $p < 0.01$ , ns = not significant.

Similar to GAC WT, GAS\_Rha polysaccharide alone did not elicit anti-GAC IgG responses above the Alum only group signal in either test or terminal bleeds after three immunisations (Figure 5.3a - c – yellow dot). Although slightly higher, glycoconjugate TT-GAS\_Rha antisera also failed to show any significant anti-GAC IgG antibody responses above the Alum antisera baseline for either test or terminal bleeds (Figure 5.3a - c – pink dot). Antisera from the third immunisation showed limited antibody boosting in 4 mice only, and more variability between the background signal (Figure 5.3 c – pink dot). Additionally, compared to mice receiving GAS\_Rha polysaccharide only, there was no significant difference in anti-GAC IgG after any immunisation.

Bioconjugate NanA-rpORha also failed to induce a significant anti-GAC IgG antibody response similar to TT-GAS\_Rha. Again, there was limited IgG boosting with each immunisation, and the anti-GAC IgG titres were found to not be significantly above the Alum only baseline after the third immunisation, nor above the group receiving GAS\_Rha polysaccharide only (Figure 5.3a - c – green dot). This suggests limited humoral responses against native GAC with this modified polysaccharide antigen.

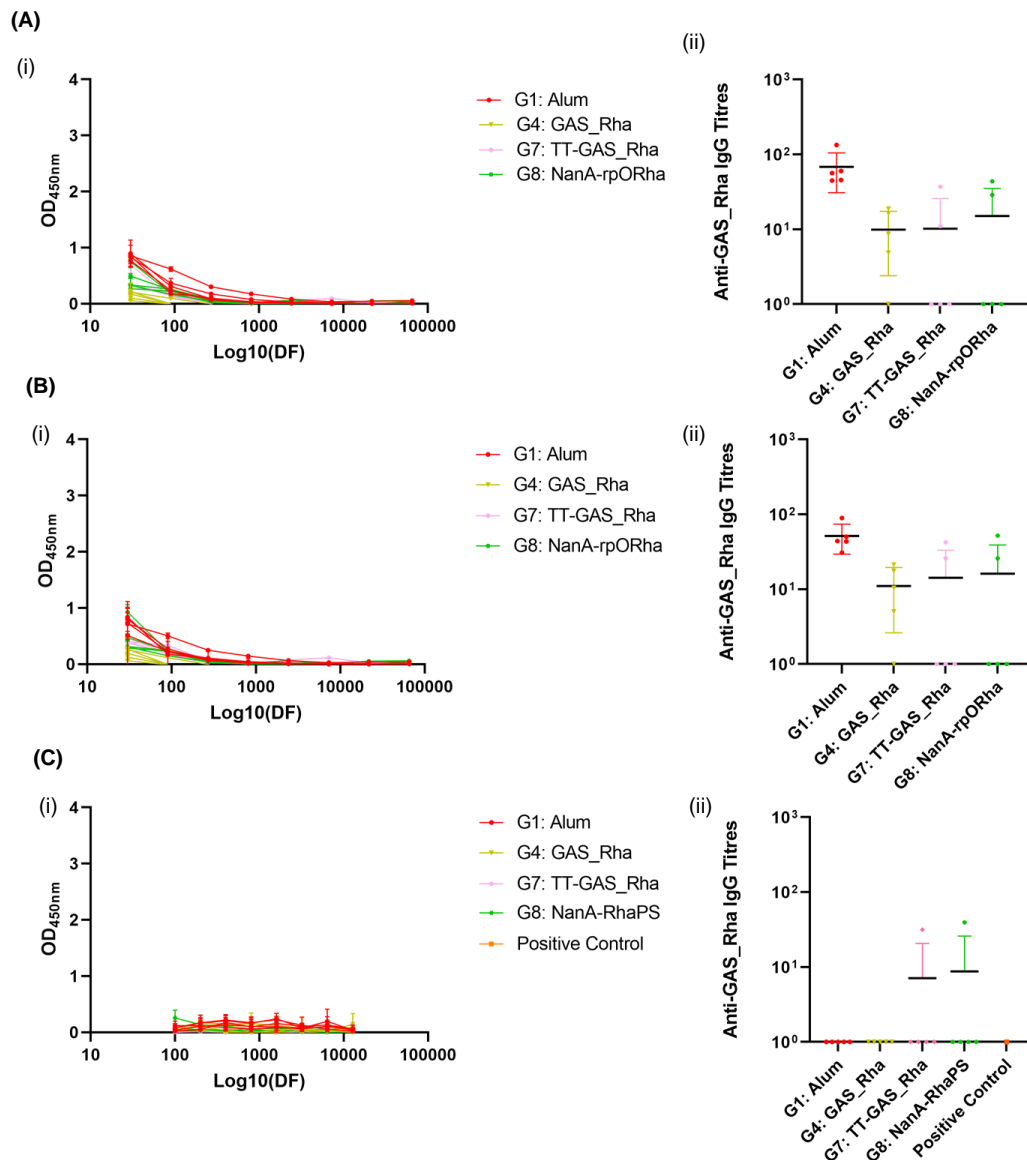


**Figure 5.3: ELISA analysis of anti-GAC IgG antibody titres in day 13, 27 and 42 antisera of mice immunised with modified GAC containing glycoconjugates.**

BALB/c mice were immunised with 3 doses of polysaccharide or glycoconjugates, GAS\_Rha (group 4 - yellow), TT-GAS\_Rha (group 7 - pink) and NanA-rpORha (group 8 - green). Sera were analysed for anti-GAC response by ELISA. Antiserum dilution titrations (i) and OD 0.5 Geometric Mean Titres (GMT; horizontal bars with vertical standard deviation error bars) (ii) for individual mice were tested and calculated for (A) test bleed 1, day 13, (B) test bleed 2, day 27 and (C) terminal bleed, day 48, compared to adjuvant only group (group 1 - red) on plates coated with purified WT GAC polysaccharide. DF = serum dilution factor.

ELISA optical density signal readings appeared to be generally low for murine antiserum from the modified polysaccharide conjugate groups when tested on WT GAC coated plates. Therefore, to test whether the observed low signal was due to polysaccharide modification, or absence of critical polysaccharide epitopes, an attempt was made to subsequently coat plates with GAS\_Rha polysaccharide, composed of rhamnose epitopes only to check possible recognition of immune sera antibodies by native antigens. Neither rhamnose containing glycoconjugate vaccine, TT-GAS\_Rha or NanA-rpORha gave an elevated signal above the Alum only group for this coating antigen in either test sera (Figure 5.4a - b pink and green dots), and in fact a small reduction in titres was observed after the third immunisation (Figure 5.4c - pink and green dots).

To verify if anti-GAC antibodies can recognise GAS\_Rha polysaccharide as coating antigens, a 'positive control' serum, used to assess plate to plate variation, was prepared from pooled sera containing antibodies against WT GAC from a previous immunisation. Results showed that anti-GAC positive serum did not recognise the GAS\_Rha polysaccharide coating antigen, unlike when WT GAC was used as a coating antigen (Figure 5.4c – orange dot). This may suggest that anti-GAC antibodies do not strongly recognise epitopes present on rhamnose polysaccharides.



**Figure 5.4: ELISA analysis of anti-GAS\_Rha IgG antibody titres in day 13, 27 and 42 antisera of mice immunised with modified GAC containing glycoconjugates.**

BALB/c mice were immunised with 3 doses of polysaccharide or glycoconjugates, GAS\_Rha (group 4 - yellow), TT-GAS\_Rha (group 7 - pink) and NanA-rpORha (group 8 - green). Sera were analysed for anti-GAS\_Rha response by ELISA. Antiserum dilution titrations (i) and OD 0.5 Geometric Mean Titres (GMT; horizontal bars with vertical standard deviation error bars) (ii) for individual mice were tested and calculated for (A) test bleed 1, day 13, (B) test bleed 2, day 27 and (C) terminal bleed, day 48, compared to adjuvant only group (group 1 - red) on plates coated with purified GAS\_Rha polysaccharide. DF = serum dilution factor.

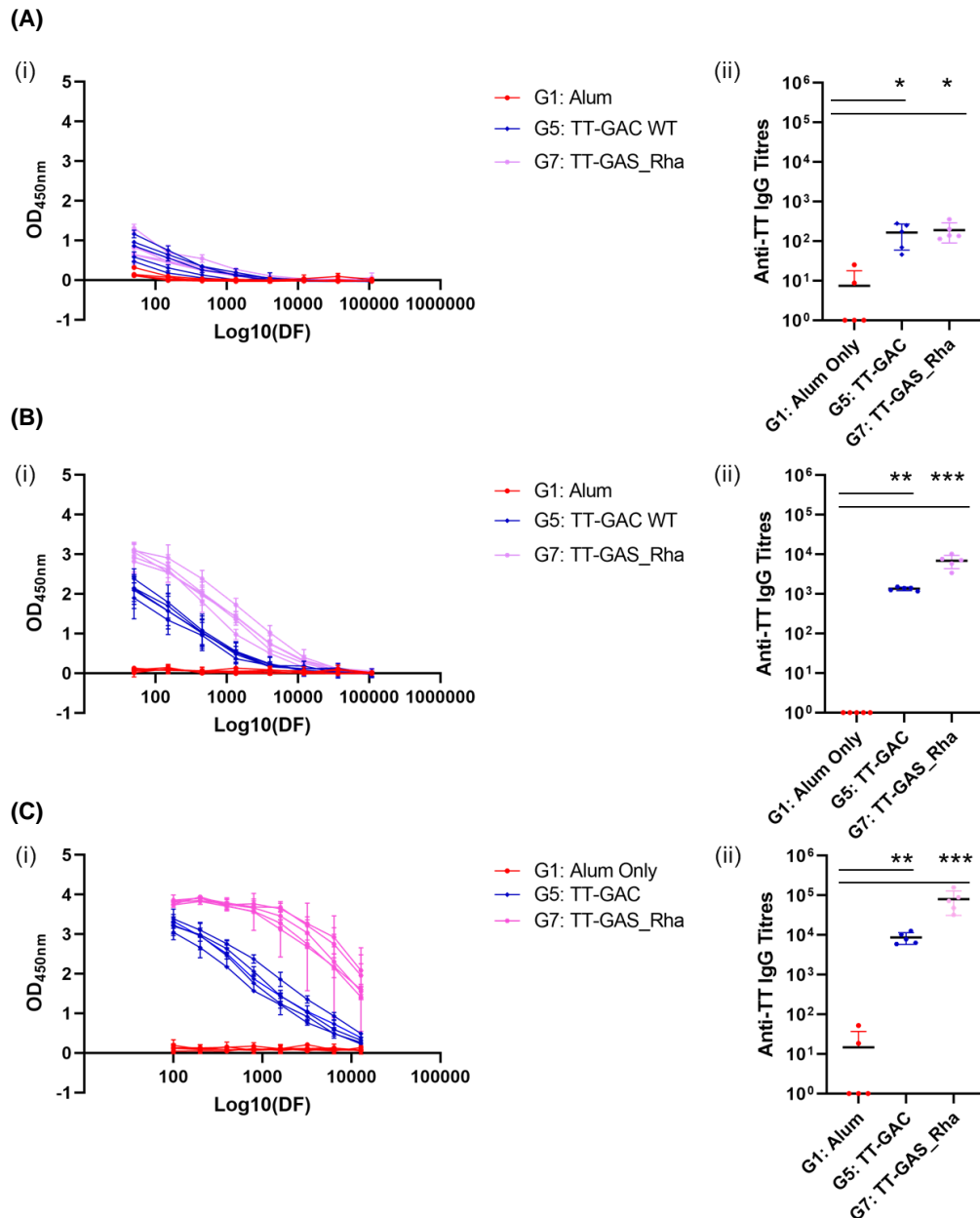
#### 5.4.1.2. *Anti-Protein IgG Levels*

In addition to testing anti-polysaccharide IgG responses, the immunogenic responses directed against the protein portion of the glycoconjugate vaccines were also evaluated in mice immunised with TT-GAC, TT-GAS\_Rha and SpyAD-GAC chemical glycoconjugates.

After priming mice with one dose of glycoconjugate samples (TT-GAC or TT-GAS\_Rha) anti-TT IgG titres were similar for both groups, significantly elevated compared to mice receiving Alum only (TT-GAC; GMT = 164 (range: 46 - 249),  $p = 0.0213$ , and TT-GAS\_Rha; GMT = 190 (range: 115 - 359),  $p = 0.0113$ ) (Figure 5.5a). Anti-TT IgG antibody titres were elevated further after the second immunisation, ~ 10-fold for TT-GAC (GMT = 1,351 (range: 1,163 - 1,527),  $p = 0.079$ ), and 36-fold for TT-GAS\_Rha (GMT = 6,858 (range: 3,424 - 10,235),  $p = 0.030$ ) (Figure 5.5b). A further increase in anti-TT IgG antibodies was observed after the third dose of both TT-GAC (GMT = 8,615 (range: 5,906 - 12,717),  $p = 0.0079$ ), and TT-GAS\_Rha (GMT = 79,558 (range: 32,044 - 157,224),  $p = 0.0008$ ) (Figure 5.5c). The final anti-TT titres were almost 10-fold higher for mice receiving TT-GAS\_Rha compared to TT-GAC glycoconjugate ( $p = 0.0079$ ).

The GAS specific protein response was also tested in sera following immunisation with 'double hit' glycoconjugate vaccine SpyAD-GAC. Data showed that the glycoconjugate induced a good level of anti-SpyAD IgG with a boosting effect from each immunisation compared to mice receiving Alum. The first immunisation showed a GMT = 266 (range: 167 – 426 ( $p = 0.0079$ )), the second immunisation a GMT = 935 (range: 536 – 1,955 ( $p = 0.0079$ )), and the third immunisation a GMT = 5,037 (range: 3,280 - 7,058 ( $p = 0.001$ )). This equated to approximately a 19-fold increase in anti-SpyAD IgG antibody titres between the first and third immunisation. Anti-SpyAD responses were significantly above background level from mice immunised with Alum only ( $p = <0.0001$ ) (Figure 5.6 a - c – purple dot).

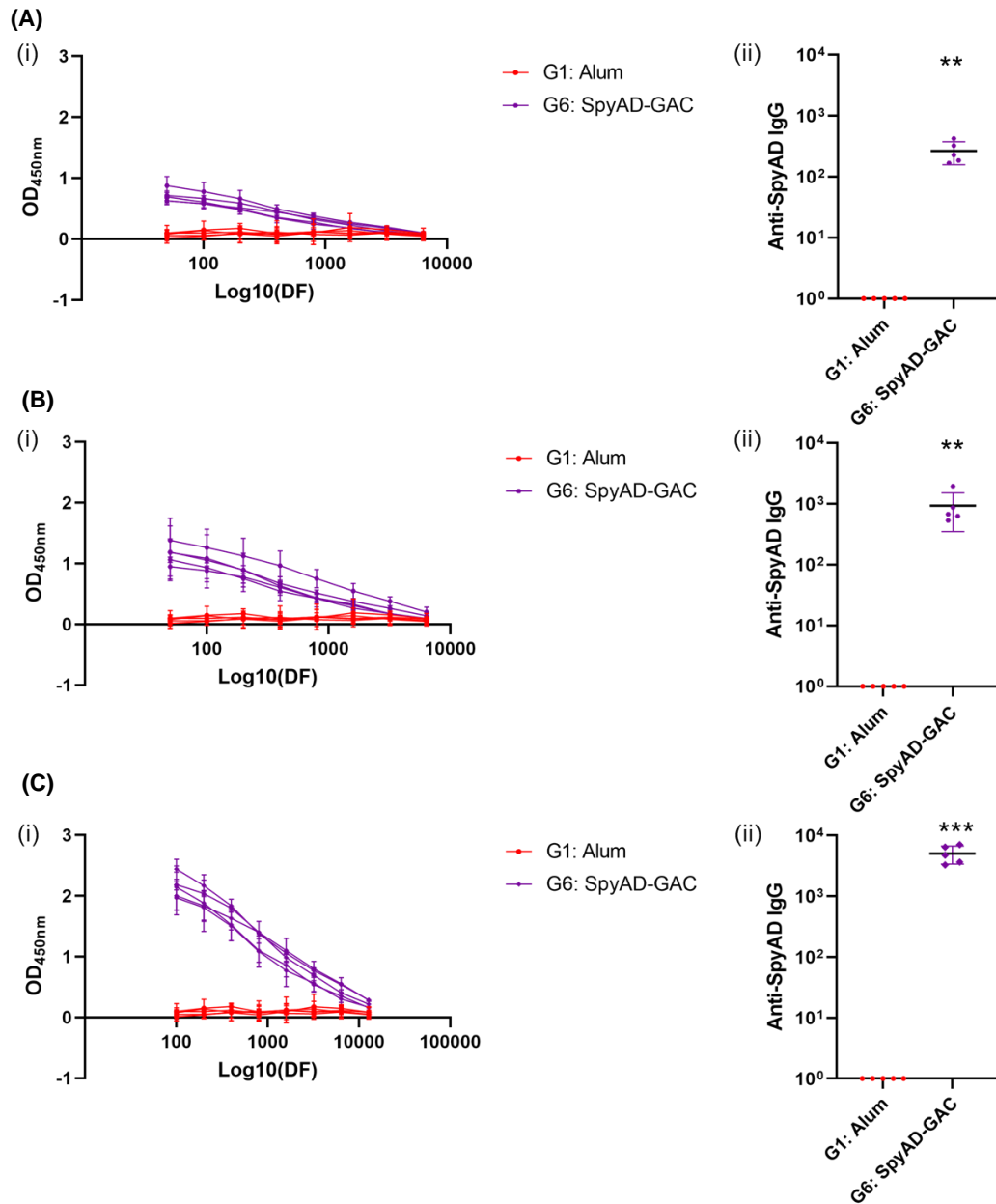




**Figure 5.5: ELISA analysis of anti-TT antibody titres in day 13, 27 and 42 antisera of mice immunised with TT containing glycoconjugates.**

BALB/c mice were immunised with 3 doses of glycoconjugates, TT-GAC (group 5 - blue), or TT-GAS\_Rha (group 7 - pink). Sera were analysed for anti-TT response by ELISA. Antiserum dilution titrations (i) and OD 0.5 Geometric Mean Titres (GMT; horizontal bars with vertical standard deviation error bars) (ii) for individual mice were tested and calculated for (A) test bleed 1, day 13, (B) test bleed 2, day 27 and (C) terminal bleed, day 48, compared to adjuvant only group (group 1 - red) on plates coated with purified TT monomer. DF = serum dilution factor.

Significance compared to group 1 (Alum) determined by Kruskal-Wallis test, \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ .



**Figure 5.6: ELISA analysis of anti-SpyAD antibody titres in day 13, 27 and 42 antisera of mice immunised with SpyAD containing glycoconjugate.**

BALB/c mice were immunised with three doses of glycoconjugate, SpyAD-GAC (group 6 - purple). Sera were analysed for anti-SpyAD response by ELISA. Antiserum dilution titrations (i) and OD 0.5 Geometric Mean Titres (GMT; horizontal bars with vertical standard deviation error bars) (ii) for individual mice were tested and calculated for (A) test bleed 1, day 13, (B) test bleed 2, day 27 and (C) terminal bleed, day 48, compared to adjuvant only group (group 1 - red) on plates coated with purified SpyAD protein. DF = serum dilution factor.

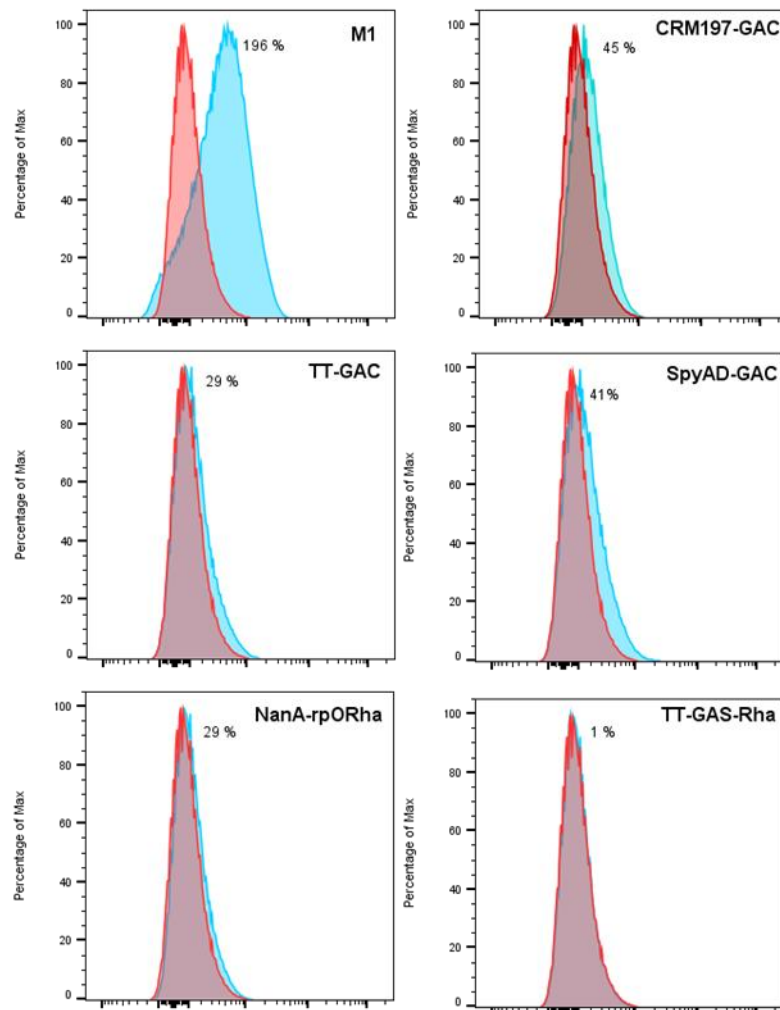
Significance compared to group 1 (Alum) determined by Mann-Whitney U test, \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

## 5.4.2. Function of Antibodies Generated from Glycoconjugate Vaccine Immunisation

### 5.4.2.1. *Glycoconjugate Directed Antibodies Binding GAS Cell Surfaces*

It is important to assess immune response functionality as well as antibody titres. First, binding to whole GAS cells was assessed to demonstrate whether antibodies were opsonising due to cell surface GAC inclusion, as well as SpyAD as a 'moon lighting' protein<sup>508</sup> for SpyAD-GAC antisera. Pooled terminal sera (day 42) from mice immunised with glycoconjugates TT-GAC, SpyAD-GAC, TT-GAS\_Rha, and NanA-rpORha were incubated with live M1 GAS strain NCTC 8189. Antibody binding was detected with FITC-labelled secondary antibodies and analysed by flow cytometry. Immune sera were compared with nonimmune sera from mice receiving Alum adjuvant only (Figure 5.7, red histogram).

Although limited GAS cell surface binding was observed for all glycoconjugate sera compared to M1 immune sera, used as a positive control, slight GAS cell surface binding was observed for some samples (Figure 5.7, blue histograms). Immune sera against glycoconjugate vaccines containing WT GAC, specifically TT-GAC and SpyAD-GAC appeared to show better GAS cell surface binding compared to glycoconjugate vaccines containing modified GAC extracted from GAS cells (TT-GAS\_Rha). The percentage of increased mean fluorescence intensity was found to be 29 % and 41 % for TT-GAC and SpyAD-GAC IgG specific antibodies respectively, when compared to the mean FITC signal of sera from mice receiving Alum only. The mean percentage increase in fluorescence from binding of SpyAD-GAC specific IgG antibodies was also found to be similar to positive control glycoconjugate GAC-CRM<sub>197</sub> which showed a 45 % increase in mean fluorescence. TT-GAS\_Rha glycoconjugate antisera appeared to show the same FITC signal compared to GAS cells incubated with nonimmune sera, with just 1 % increase in mean fluorescence. Interestingly, a smaller shift was observed for antibody binding to GAS cells from mice immunised with NanA-rpORha bioconjugate containing modified rhamnose polymers. This antiserum showed an increase of 29 % in mean fluorescence intensity compared to Alum antisera, however whether this was attributable to anti-polysaccharide or anti-protein, or indeed a combination of antibody responses is unknown.



**Figure 5.7: Flow cytometry analysis of GAS M1 NCTC 8189 cells stained with mouse antisera against glycoconjugate vaccines.**

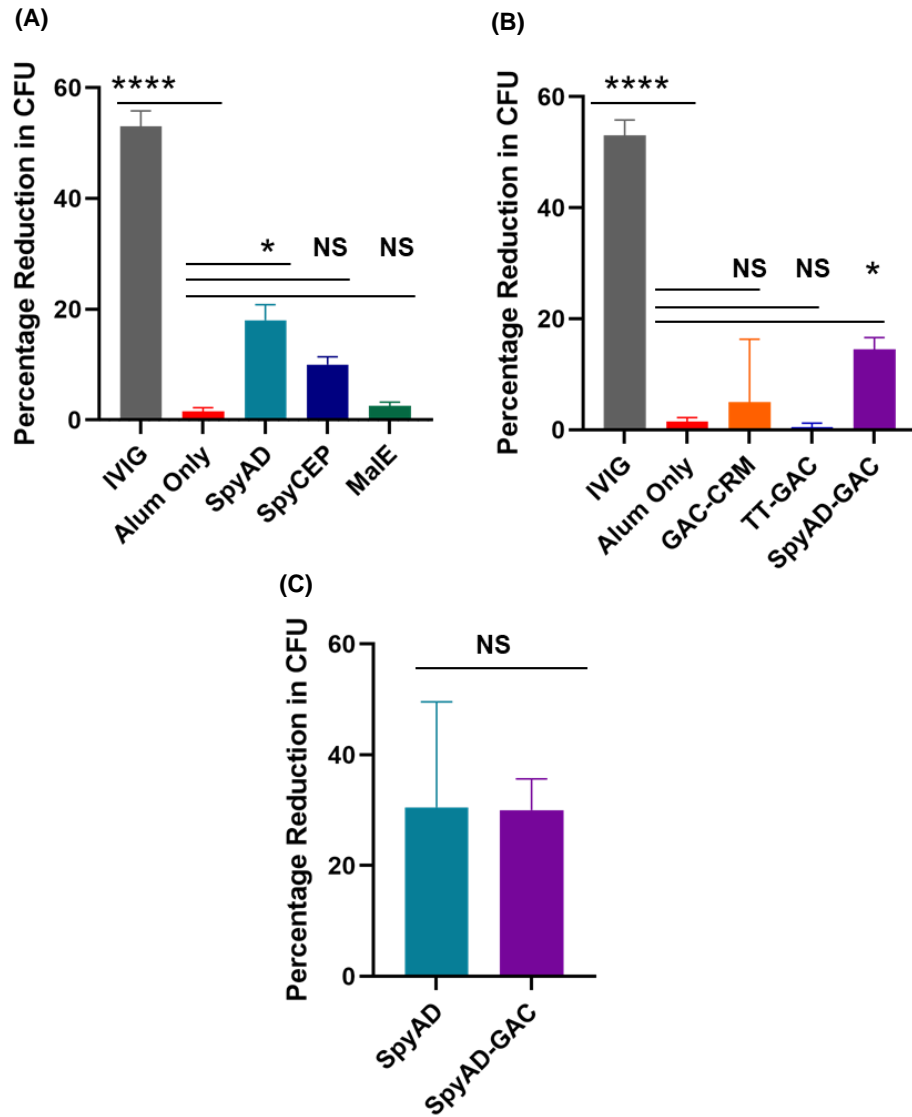
Histograms show specific IgG binding by a shift in the FITC fluorescent signals. Mouse antisera was diluted 1 : 4 and compared to diluted nonimmune sera (group 1 - Alum). Red = nonimmune sera, and blue = immune sera. Numbers indicate the calculated mean percentage of increased signal of immunised mouse serum over the nonimmune serum signal.

#### 5.4.2.2. *Glycoconjugate Directed Antibodies Inducing GAS Opsonophagocytic Killing*

In addition to cell surface binding, functionality through opsonophagocytic GAS killing was investigated. Initially immune sera against protein antigens (Chapter 3 section 3.4.2.2), were tested for presence of opsonic antibodies compared to sera obtained from mice receiving Alum only, and pooled human IgG (IVIG) was used as a positive control. Out of the three tested GAS antigens, only sera against SpyAD showed some level of killing with 18 % reduction in GAS CFUs, significantly greater than baseline ( $p = 0.0207$ ) (Figure 5.8a). Anti-SpyCEP antibodies showed a 10 %

reduction in GAS CFUs but this was not statistically different from baseline, and MalE directed antibodies failed to initiate any functional opsonophagocytic killing above nonimmune sera (Figure 5.8a).

Antisera from mice immunised with glycoconjugate vaccines which generated some level of anti-GAC antibodies, specifically TT-GAC and SpyAD-GAC, were also tested for opsonophagocytic GAS killing compared to nonimmune sera from mice immunised with Alum only. Antisera from positive control glycoconjugate vaccine, GAC-CRM<sub>197</sub> (Figure 5.8b – orange bar), and test glycoconjugate TT-GAC (Figure 5.8b – dark blue bar) failed to show any opsonophagocytic killing above baseline levels. Of all the tested glycoconjugate vaccines, SpyAD-GAC glycoconjugate performed the best with 17.5 % reduction in GAS CFUs (Figure 5.8b – purple bar), statistically significant above baseline levels ( $p = 0.0194$ ). To compare killing activity, SpyAD-GAC conjugate and SpyAD protein alone antisera were normalised to one another and showed similar levels of GAS killing (Figure 5.8c). This suggests killing is mediated by anti-SpyAD antibodies, as the effect was not observed in mice immunised with classical carrier protein (TT or CRM<sub>197</sub>) glycoconjugate vaccines.



**Figure 5.8: Glycoconjugate and protein antigen immune sera can induce neutrophil-mediated GAS killing assessed by a HL60-based opsonophagocytosis assay (OPA).**

GAS cell killing by immune mouse sera was quantified relative to the controls containing nonimmune sera (group 1 – alum). Pooled human IgG (IVIg) positive control, Alum only group as baseline, and sera against protein antigens **(A)**, glycoconjugate vaccines **(B)** and normalised SpyAD containing protein antisera **(C)** were tested.

Unpaired T test comparison with Alum sera is indicated: NS, not significant, \*  $p < 0.05$ , and \*\*\*\*  $p < 0.0001$ .

## 5.5. Discussion

The immunogenicity of glycoconjugate vaccines is influenced by a multitude of factors, including but not limited to polysaccharide and protein carrier characteristics, and the selected conjugation approach<sup>458</sup>. Understanding in molecular chemistry, glycobiology, and glycoimmunology have led to substantially improved glycoconjugate vaccine efficacy, however the relationship between polysaccharide and protein epitopes in preparations remains fairly uncharacterised<sup>628</sup>. This Chapter aimed to assess immune stimulation by glycoconjugates derived from two conjugation methods, through chemical attachment of purified antigens, or by bacterial *E. coli* glycoengineering. Additionally, previous studies have only investigated GAC or rhamnose containing polysaccharide antigens in isolation, distinct from this study with a direct comparison of polymers enzymatically extracted from GAS cells. Overall, from the four candidates, only TT-GAC glycoconjugate was immunogenic with respect to the GAC response after three immunisations in mice. Comparisons and potential reasons for the limited immunogenicity will now be discussed.

### 5.5.1. Immune Response to Chemical Glycoconjugate Vaccines – TT-GAC and SpyAD-GAC

Protein moieties in both semi-selective chemical and site-specific biological glycoconjugate preparations provide T cell help, traditionally through classical protein carriers known to be efficient in supporting anti-polysaccharide IgG antibody generation. In this study, relatively similar anti-polysaccharide IgG levels were induced by both TT-GAC and SpyAD-GAC glycoconjugates, however, TT-GAC showed slightly higher responses compared to SpyAD-GAC terminal sera, the latter of which was deemed not to be significantly elevated above baseline.

Protein to polysaccharide ratio is a key aspect to understanding conjugation reaction efficiency and resulting immunogenicity. In this work total molar ratios were calculated due to challenges encountered to separate free protein from conjugates, therefore both conjugated and unconjugated protein forms were present in preparations. Total molar ratios were calculated to be similar for TT-GAC (1 : 5.59 *m/m*) and SpyAD-GAC (1 : 8.84 *m/m*) protein to polysaccharide molecules. Quantification of unconjugated versus conjugated TT was possible, showing its conjugate molar ratio to be around 1 : 9.72 *m/m* TT to GAC but SpyAD-GAC was constrained by the amount of remaining sample for testing so the conjugate ratio could not be calculated. However, based on the assumption that not all SpyAD

protein in the purified reaction will be conjugated to GAC molecules, SpyAD may have in fact been more glycosylated than TT, despite not inducing significant anti-GAC antibody titres. This may be linked to data generated in this thesis suggesting that SpyAD has more targetable surface lysine residues compared to TT. Despite this however, based on slightly improved anti-GAC IgG titres, classical carrier TT may be a superior carrier to SpyAD, but further work would be required to validate this claim. TT and SpyAD are similar in size, 90 kDa and 88 kDa respectively, therefore such differences may be due to protein structural properties, specifically in relation to the way the GAC epitopes are presented. When assessing protein structures using *in silico* tools, SpyAD showed increased alpha helical conformation compared to TT which may impact GAC polymer extension. However, such theories cannot be proven based on this work due to the limited number of reactions performed, and minor difference in anti-polysaccharide titres. Rather such observations may be due to batch to batch variations between reactions<sup>546</sup> and preparation composition, rather than the physiochemical properties of the conjugation components themselves. These confounding factors should be further explored to determine consistency between conjugation reactions in future studies through more accurate ratio quantification.

In addition to anti-GAC antibody titres, assessing whether GAS antigen SpyAD was an efficient carrier for GAC, it was also of interest to determine whether conjugation improves the overall glycoconjugate immunogenicity. This was through determining GAS specific anti-protein antibody responses. High anti-SpyAD IgG titres were observed in both test and terminal sera with a boosting effect with each SpyAD-GAC immunisation. A previous study chemically conjugating GAC to SpyAD using random derivatisation and activation along the entire GAC chain generated glycoconjugates with an alternative “mesh-like” structure<sup>411</sup>. This vaccine generated lower anti-SpyAD protein antibodies following conjugation than a comparable dose of SpyAD protein alone<sup>411</sup>, likely due to heavy polysaccharide decoration masking protein specific B cell epitopes. This observation gives merit to less disruptive conjugation approaches forming “sun-like” glycoconjugate structures, specifically for ‘double hit’ glycoconjugates, preventing shielding of immune epitopes, and disruption to protein structure and conformation. The SpyAD-GAC glycoconjugate generated in this work was able to present vaccine components in such conformations to allow the generation of specific antibodies capable of recognising the native protein. This strongly suggests that it is important to select conjugation chemistry suitable to preserve its role as a B cell epitope carrier, as well as induce



its own protein directed immune response. However, such “sun-like” conjugation methods may reduce overall polysaccharide attachment due to limiting the region of the chain which is activated and available for attachment<sup>467</sup>. It may therefore not be suitable for polysaccharides which are difficult to activate and target without chemical modification, or those that require specific conformations to stimulate strong immune responses. Calculated conjugate ratios suggest a good level of polysaccharide decoration for TT-GAC and SpyAD-GAC using terminal linkage, however based on the modest anti-polysaccharide titres the position and structure of the polysaccharide chains in relation to protein may not have been optimal.

Antibody titres are only meaningful if they convey functional immunity. Towards this within the time constraints of this project, antibody function was assessed by *in vitro* assays comparing terminal immune sera from SpyAD-GAC and TT-GAC immunisations. SpyAD-GAC terminal sera showed superior antibody function compared to TT-GAC immune sera. This was predicted to be due to the additional immunological valency generated from presenting the immune system with two antigens in SpyAD-GAC absent in TT-GAC glycoconjugate. However, sera following immunisation with SpyAD protein alone also demonstrated killing by opsonophagocytosis to a similar level as SpyAD-GAC antiserum when normalised according to anti-SpyAD antibody titres. This level of GAS killing was found to be comparable to a study immunising rabbits with other SpyAD chemical glycoconjugates<sup>412</sup>, however, suggests that anti-SpyAD antibodies are contributing more than anti-GAC antibodies to opsonophagocytic killing. This is complimented by the lack of significant GAS killing observed with incubation of TT-GAC or CRM<sub>197</sub>-GAC positive control immune sera, despite the latter having the highest anti-GAC IgG titres of all tested glycoconjugates and presence of antibodies capable of binding to GAS cell surfaces. A number of studies have encountered difficulties in observing killing attributable to anti-GAC mouse sera. This therefore justifies the use of ‘double hit’ vaccines as the anti-protein response may be more significant towards killing activity, as well as making use of other animals such as rabbits which are capable of generating higher titres of opsonic anti-GAC antibodies<sup>141,412</sup>.

A major limitation to this analysis is that SpyAD-GAC preparations also contain unconjugated SpyAD protein capable of generating its own anti-protein response in addition to the glycoconjugate molecules. The presence of unconjugated SpyAD may be the cause of the relatively low anti-GAC antibody titres due to expansion of protein specific B cell populations, discussed further below in section 5.5.2 in the context of TT containing glycoconjugates. The total SpyAD protein in both

conjugated and unconjugated forms when immunising with 2 µg of GAC in the glycoconjugate sample was calculated to be 2.15 µg per 200 µl dose, which is approximately 20-fold lower than the 20 µg dose used to immunise mice with protein antigens only. Therefore, antibodies against the protein component of the glycoconjugate, as well as antibodies against unconjugated protein may be contributing to the antibody function. However, future studies would benefit from immunising with more highly purified glycoconjugate samples, to remove unconjugated protein, and assess the generated immune responses more clearly.

#### 5.5.2. Immune Response to Chemical Glycoconjugate Vaccines – TT-GAC and TT-GAS Rha

As previously described, the GlcNAc on GAC has been implicated in autoimmune responses to GAS infection. Such antibodies do not recognise the rhamnose component of GAC<sup>299,649,650</sup>, therefore, the evaluation of rhamnose polymers as alternative antigens to aid safe vaccine manufacture have been long recognised<sup>141,412,416</sup>.

The presented work investigated the immunogenicity of modified GAC which does not contain GlcNAc epitopes and went further to compare with wildtype GAC conjugates. This was initially tested in the form of GAS\_Rha extracted from mutant GAS cells, conjugated in the same manner as WT GAC to classical carrier protein TT. The TT-GAS\_Rha glycoconjugate vaccine appeared not to be immunogenic in this study, with all mouse test and terminal sera not found to contain anti-GAC antibodies significantly above negative control levels (Alum adjuvant). TT-GAS\_Rha sera also showed limited binding to M1 GAS NCTC 8189 cell surfaces.

Conjugate ratios could be calculated for TT-GAC and TT-GAS\_Rha based on quantification following immunoprecipitation of unconjugated TT. TT-GAS\_Rha showed less polysaccharide decoration at 1 : 6.81 *m/m*, compared to TT-GAC at 1 : 9.72 *m/m*, TT to polysaccharide molecules. Lower polysaccharide decoration of TT with GAS\_Rha compared to WT GAC may have been responsible for the lack of immunogenicity. Additionally, polysaccharide compositional differences may also have been responsible. For example, prior to conjugation, purified GAS\_Rha polymers were found to have a lower mass of ~ 4.9 kDa relating to loss of side chains as well as fewer repeating unit reducing the overall polymer size, limiting the number of available epitopes for immune stimulation. GAS\_Rha polymers also do not contain the immunodominant GlcNAc residues, resulting in the loss of the binding conformation groove, thought to be important in antibody generation and

recognition<sup>144,442,443</sup>. This theory is also complimented by a recent study finding the branching point within WT GAC repeating unit to be important<sup>416</sup>. However, interestingly the identity of the monosaccharide located at this branching point does not necessarily need to be a GlcNAc residue, highlighting recognition of structure rather than monosaccharide identity<sup>416</sup>.

The apparent lack of immunogenicity generated by rhamnose containing polymers in TT-GAS\_Rha glycoconjugate is dissimilar to studies in the literature. One study used GAS derived rhamnose polymers conjugated to a recombinant pneumococcal protein<sup>141</sup>. They demonstrated robust immune responses capable of WT GAC recognition, binding to a number of disease-associated GAS isolates, as well as inducing GAS bacterial killing following 5 immunisations (10 – 20 µg/dose) in a rabbit animal model. Another, more relevant study immunising with glycoconjugates containing rhamnose polymers attached to GAS antigen SpyAD, had a similar regime to this work with 3 immunisations at a slightly higher dose of 5 µg of polysaccharide, also in a rabbit model<sup>412</sup>. This group did not numerically quantify generated anti-GAC titres, and rather showed functionality by WT GAS cell binding of glycoconjugate antisera to a variety of isolates<sup>412</sup>. This may have therefore been mediated by anti-SpyAD rather than anti-GAC antibodies. Additionally, both studies likely used rabbits to generate more opsonophagocytic antibodies compared to mouse sera as shown for a *Pseudomonas aeruginosa* glycoconjugate<sup>651</sup>.

A major limitation to this work generally was the inability to remove unconjugated protein from glycoconjugate preparations used in immunisations. Limited immunogenicity was therefore likely related to high levels of unconjugated protein directing immune responses towards the carrier protein rather than the polysaccharide. This has been shown experimentally with high protein concentrations at required polysaccharide doses possibly resulting in limited accessibility to polysaccharide specific B cell epitopes for efficient antibody generation<sup>473</sup>. Specifically, this phenomenon has been observed in a study assessing the immunogenicity of pertussis-diphtheria-tetanus (DTaP) components administered alone and in combination with Hib glycoconjugate vaccine. The study found reduced anti-polysaccharide responses when co-immunising Hib glycoconjugate with the TT component of DTaP<sup>652</sup>. In this instance this effect was thought to be due to TT overloading, interfering with polysaccharide binding to B cells. This effect can also be mediated by competition for T cell help following antigen processing by B cells specific for carrier protein epitopes and B cells specific for polysaccharide epitopes. This leads to suppression of anti-

polysaccharide responses, due to the clonal expansion of carrier protein specific B cells<sup>473</sup>. Such clonal dominance and expansion through intra-molecular antigenic competition decreases the probability of polysaccharide specific B cells interacting with antigen, and has been observed particularly with TT immunisation<sup>473</sup>.

These theories may explain the lower-than-expected anti-polysaccharide responses of TT-GAC and TT-GAS\_Rha compared to highly purified positive control glycoconjugate, GAC-CRM<sub>197</sub> which does not contain unconjugated protein. TT-GAC and TT-GAS\_Rha showed high anti-TT carrier protein antibodies in both test and terminal sera, with TT-GAS\_Rha having ~ 10-fold higher anti-TT antibodies than TT-GAC after the second immunisation. This observation links with the unconjugated TT concentration separated and determined by immunoprecipitation ELISA quantification. TT-GAS-Rha had a higher concentration of unconjugated TT protein in the preparation. In a 2 µg polysaccharide dose there was 0.6 µg/ml unconjugated TT (equivalent to 16 % total protein) in TT-GAC and 3.8 µg/ml unconjugated TT (28 % total protein) in TT-GAS\_Rha preparations. Therefore, based on unconjugated TT concentrations and GAS\_Rha's smaller molecular mass, it was likely that TT-GAS\_Rha contained more accessible TT (both conjugated and unconjugated) compared to TT-GAC. Taking this into account, it is plausible that the presence of dominant anti-TT immune responses, leading to the clonal expansion of B cells recognising TT, may have masked any potential small B cell response to GAS\_Rha in the current experimental setup.

### 5.5.3. Immune Response to Bioconjugate Vaccine – NanA-rpORha

Similar to TT-GAS\_Rha, NanA-rpORha bioconjugate containing rhamnose polymers, also failed to induce an anti-GAC IgG response above Alum background, with no boosting effects observed after each immunisation. For both NanA-rpORha and TT-GAS\_Rha polysaccharide chain lengths were predicted to be short based on western blot and SEC-MALS analysis respectively. Shorter chain lengths may therefore reduce the density of accessible B cell epitopes, hindered by the lack of structural branching points within the repeating units as discussed previously.

Studies have found the minimal immunogenic epitope displaying the required structural motifs to induce anti-GAC antibodies to be a hexasaccharide polymer, equating to two repeating units (Rha-Rha-GlcNAc)<sup>139</sup>. Analysis of glycoconjugates in this work failed to assign molar masses, and therefore did not provide polysaccharide chain length information to aid immunogenicity understanding. In theory however all investigated polysaccharides, either GAS or *E. coli* derived

should have provided at least a minimal epitope in terms of chain length. Chain length beyond the minimal epitope is thought to influence the magnitude of immune response to a certain extent, with one study showing a synthetic dodecasaccharide, equating to four trisaccharide GAC repeating units was required to generate antibody titres equivalent to purified native GAC<sup>139</sup>. Despite this, increased polysaccharide chain length does not always correlate with increased immunogenicity. For example, one study found the inclusion of WT trisaccharide GAC polysaccharide generated elevated anti-GAC antibodies than corresponding hexa- and nona-saccharide glycoconjugates attached to a variety of protein carriers<sup>447</sup>. This however was not the case when assessing the immunogenicity of rhamnose only polysaccharides attached to a lipid-core peptide, showing that increasing the number of rhamnose residues in the immunogenic epitope improved immunogenicity<sup>416</sup>. An optimal immunogenic epitope relating to chain length and polysaccharide identity is therefore hard to determine, with the latter study investigating the immunogenicity of a single repeating unit<sup>447</sup> compared to a polymer consisting of multiple repeating units<sup>139</sup>. This is an area of future study important to decipher the relationship between polysaccharide structural properties and resulting polysaccharide specific antibody production.

NanA-rpORha had a total molar ratio of 1 : 4.05 (*m/m*) NanA to rpORha molecules, based on the protein value accounting for both conjugated and unconjugated NanA. The total molar ratio for NanA-rpORha conjugation components is unlikely to be accurate given that a single NanA molecule can only be targeted in two places where the glycotags are positioned. This may be due to detection of contaminating endotoxin by anthrone polysaccharide determination. Based on western blotting it was unclear whether both sites were occupied with rpORha polymers, and SEC-MALS could not be conducted on conjugate samples due to material constraints. Therefore, the actual size of polymers on the final glycoconjugate is only hypothesised based on expression data in the absence of proteins and may be different for the conjugate. Regardless, the optimal number of glycotags and precise protein attachment sites related to bioconjugation yields and resulting immunogenicity remains unclear, however efficiency has been shown to vary depending on glycotag location<sup>653</sup>. This work only investigated glycotags added to protein termini similar to other studies<sup>567</sup>, however it may be beneficial to place glycotags within protein loops to encourage attachment within the protein itself<sup>564</sup>. This may produce heavily glycosylated products<sup>654</sup>, which may be beneficial to create more potent immunogens, and vaccines which physically look similar to

chemical glycoconjugates. Such approaches may be beneficial if the poor immunogenicity generated from NanA-rpORha was related to sub-optimal glycotag positioning within regions of the NanA protein which are distant to useful epitopes for antigen processing by B cells.

Compared to TT-GAS\_Rha antisera, slightly improved GAS cell binding was observed with NanA-rpORha antisera, despite both rhamnose containing glycoconjugates not generating anti-GAC titres above baseline. Such differences in GAS cell binding may be associated with the inclusion of streptococcal protein carrier in NanA-rpORha bioconjugate, absent in TT-GAS\_Rha chemical glycoconjugate. The colonisation factor NanA protein is natively involved in adherence to epithelial and nasopharyngeal cells<sup>655</sup>, as well as biofilm formation<sup>656</sup> for *S. pneumoniae*. NanA has been shown to possess a catalytic domain located at the C terminal, and a lectin like domain at the N terminal involved in binding endothelial cell surface receptors<sup>601</sup>. In this work NanA-rpORha contains the lectin like domain of NanA only, which may share antigenic epitopes with proteins on the surface of GAS cells. Although BLAST analysis failed to show any significant similarities between NanA and the GAS genome, functional redundancy and similarities between protein extracellular matrix binding domains have been observed. For example Pava fibronectin binding pneumococcal protein displays high similarity to GAS adhesin FBP54<sup>657</sup>, so this may be possible with NanA epitopes. This would therefore explain the apparent slight improved antibody affinity for GAS cell bacterial binding through the generation of assumed anti-protein antibody responses, although not numerically quantified in this work.

#### 5.5.4. Additional Functional Immunological Analysis of Glycoconjugate Vaccines

This Chapter also investigated opsonophagocytic killing, as opsonic antibodies against M proteins that mediate phagocytic uptake are thought to be a potential human correlate of protection in serotype-specific immunity<sup>345</sup>, and demonstrate protection against symptomatic pharyngitis<sup>658</sup>. Additionally, the GAS vaccine development roadmap states the requirement of high-throughput assays to measure functional opsonic antibodies to assess effectiveness<sup>545</sup>. As discussed, the best GAS cell killing was observed for SpyAD-GAC glycoconjugate antisera despite having lower anti-GAC titres compared to TT-GAC glycoconjugate antisera. This is thought to be mediated primarily by protein specific antibodies.

The overall low immune response coupled with time and material constraints meant that further characterisation of the immunogenicity generated by these glycoconjugate vaccine candidates was not investigated. However, characterisation of the quality of the immune response elicited would be of use to better understand glycoconjugate mechanism of action. For example, T cell helper activity, important in the overall immune response, is believed to be dependent on the nature of the polysaccharide antigen as well as the carrier protein<sup>535</sup>. T helper responses can themselves be evaluated indirectly by assessing IgG subclasses, such as IgG1 and IgG2a which can indicate Th1/2 type responses. For example, vaccines containing synthetic or WT GAC conjugated to CRM<sub>197</sub><sup>139</sup> or Generalised Modules for Membrane Antigens (GMMA)<sup>659</sup> formulated with Alum have been shown to induce predominantly IgG1 antibodies suggesting a Th2 response. Th2 responses appear to be consistent with Alum adjuvant<sup>660,661</sup>, and therefore may have been observed as the predominant subclass following immunisation with glycoconjugates formulated with Alum in this study. Alternative adjuvants may improve the response to GAC, as seen for numerous GAS protein antigens<sup>397</sup>, and should be further explored.

Another approach to investigate B and T cell responses is the use of nude mice devoid of T cells in glycoconjugate immunisations. This approach has been applied to GAS vaccines, used for immunisation with GMMA containing GAC, discussed as alternative antigens in the overall discussion<sup>662</sup>. This particular study found that both B and T cell activity is necessary for strong humoral responses towards GAC<sup>662</sup>. Additionally, nude mice have also been used to show that different mechanisms of presentation, specifically polysaccharide structure and chain length dictating the dominance of T cell responses. For example, based on a *S. Typhi* capsule containing glycoconjugate, shorter polysaccharide antigens appear to require more T cell help for robust immune stimulation compared to glycoconjugates containing longer polysaccharide antigens<sup>663</sup>. Such T cell activity is important in immune response magnitude and duration, as without T cell effects apoptosis and depletion of polysaccharide-specific B cells in the spleen and bone marrow can occur, leading to hyporesponsiveness<sup>663</sup>. Such impact of polysaccharide size on B and T cell interactions is known to be antigen-specific<sup>458,467</sup>, requiring future studies to focus on deciphering the best protein carrier to induce T cell activity and enable GAC directed memory responses. This may entail measuring memory T cell response in spleen proliferation assays by pre-incubating cells with the carrier protein and measure proliferation.

In addition to nude mice, enzyme-linked immunospot (ELISPOT) assays are useful for assessing cytokine levels as an indirect measurement of T cell helper function. This approach proved useful in the discovery glycopeptide presentation by B cells for recognition with polysaccharide-specific CD4<sup>+</sup> T cell clones through IL-2 / -4 generation<sup>470</sup>. This approach therefore showed CD4<sup>+</sup> T cell expansion recognising the polysaccharide portion of the glycoconjugate, absent in protein carriers alone. It may therefore have been useful to quantify IL-2/4 cytokine levels following immunisation to assess the relationship between anti-GAC IgG titres and induction of T cell helper responses for these glycoconjugate candidates.

Once immune response mechanisms have been fully characterised, antibody function linked to clinical efficacy is important for testing. Efficacy targets based on the GAS vaccine roadmap are the reduction of pharyngitis and skin infections, due to associated difficulties in demonstrating vaccine efficacy against RHD<sup>56</sup>. Towards this aim, despite no representative animal model, studies testing pre-clinical GAS vaccine candidates have made use of non-human primates to recapitulate human pharyngeal infection<sup>434,664,665</sup>, as well as other animal models<sup>422,666,667</sup> to aid selection of efficacious vaccine candidates. Additionally, to accelerate vaccine licensure, controlled human infection models (CHIMS) are being developed to overcome the limited applicability of animal models, and bridge the gap between animal and human GAS immune responses<sup>668</sup>. Although a lengthy process, such models and correlate of protection testing are imperative towards human safety and efficacy testing for preclinical vaccines to make it to market.

#### 5.5.5. Vaccine Safety Testing

A requirement, additional to characterising immune responses, is to test for autoreactive vaccine induced antibodies, as stated in the GAS vaccine roadmap<sup>56</sup>. Studies to date have confirmed no observed autoimmune antibodies generated after vaccination with modified GAC polymers devoid of GlcNAc residues<sup>141,412</sup>. However, interestingly, neither have such antibodies been observed for WT GAC containing glycoconjugate vaccines<sup>140</sup>, which may be immunologically more beneficial to encourage native recognition of circulating GAS isolates. Such testing has investigated antibody cross reactivity with human tissues such as heart, brain, kidney and liver, or more specifically cytoskeletal proteins present in mammalian tissue using ELISA and immunofluorescence<sup>140</sup>. This can also be achieved more simply through incubating antisera with human heart and brain lysates for western blot analysis<sup>412</sup>. Since this work was proof of concept and associated with generally low anti-GAC antibody titres, the generation of autoimmune antibodies was not



tested. However, it is likely such testing will be required for the eventual development and licensure of any GAS specific vaccine, especially those containing polysaccharide antigens.

## 5.6. Summary

This Chapter investigated the immune response generated in mice following three subcutaneous immunisations with glycoconjugate vaccines manufactured using chemical and biological conjugation methodologies. Immune sera were tested by determining anti-polysaccharide serum IgG antibody titres for all groups, and anti-protein serum IgG antibody titres for three groups immunised with different glycoconjugate vaccines, specifically TT-GAC, TT-GAS\_Rha, and SpyAD-GAC. TT-GAC was the only test glycoconjugate able to generate both significant anti-polysaccharide and anti-protein directed responses, whereas SpyAD-GAC showed strong anti-protein response without significant anti-polysaccharide antibody generation. Rhamnose polysaccharide antigens (GAS\_Rha and rpORha) failed to show any polysaccharide directed immune responses regardless of protein carrier. Serum antibody function was also assessed through GAS cell surface binding and opsonophagocytosis killing. SpyAD-GAC was the only test glycoconjugate able to show a functional immune response through GAS killing, mediated by SpyAD protein directed antibodies, supporting the investigation of 'double hit' glycoconjugate vaccines. This study among others has somewhat contributed to the understanding of the mechanisms underlying the immune responses to GAS specific glycoconjugates, providing a novel example of the comparison of WT GAC and rhamnose polysaccharide antigens. However, further immune response characterisation, and in-depth studies to understand the effect of glycoconjugate physiochemical properties and attachment methods related to immunogenicity is crucial for the future production of knowledge-based glycoconjugate vaccines.

# CHAPTER 6

## Overall Discussion

### 6.1. Broad Themes and Thesis Scope

Glycoconjugate vaccines are safe and efficacious, minimising the disease burden from bacterial infections<sup>462,669–671</sup>. Their success dates back over 40 years, but the full potential of glycoconjugate vaccinology remains to be fully exploited. Practically the development of new glycoconjugate vaccines is challenging<sup>672,673</sup>, often impeded by conflicting evidence on optimal vaccine components and production methods. This study, generating novel glycoconjugate vaccine candidates, using two distinct conjugation methodologies, is important towards the generation of a desperately needed Group A Streptococcus (GAS) vaccine. The following Chapter presents the main conclusions of the thesis, as well as discussing remaining questions and areas of further study.

#### 6.1.1. Identifying Novel Immunogenic Antigens for Glycoconjugate Vaccine Inclusion

The focus on 'double hit' glycoconjugate vaccines is an exciting development in the field of vaccinology. Until recently the main focus of GAS glycoconjugates included classical protein carriers only, generating opsonic anti-GAC antibodies<sup>139</sup>, aiming to reduce nasopharynx colonisation in humans<sup>409</sup>. GAC related polysaccharides without GlcNAc sidechains have also been investigated as alternative antigens<sup>141,412</sup>. Determining whether rhamnose polymers are feasible alternative antigens to wildtype GAC, may have wider implications for not only vaccine safety but also antibody cross reactivity with other streptococcal species. This is due to similarities in polysaccharide biosynthetic pathways in other bacteria<sup>152</sup>, and cross reactivity of anti-GAC antibodies with rare Streptococcal group C variant strains such as *Streptococcus dysgalactiae* subsp. *Equisimilis*<sup>674,675</sup>, in addition to *Streptococcus castoreus*, a Group A bacterium isolated from beavers<sup>676</sup>. Both wildtype and modified GAC (GAS\_Rha) variants tested in this work showed differences in their physiochemical properties and immune characteristics. Specifically, western blots showed GAS\_Rha is antigenic, however, in the current experimental set up GAS\_Rha did not appear to be immunogenic, with wildtype GAC generating higher serum IgG responses, and improved GAS cell surface binding, compared to modified rhamnose (GAS\_Rha / rpORha) when included in glycoconjugate preparations.

The approach in this thesis compared a classic carrier protein (TT) to species specific carrier proteins forming a 'double hit' glycoconjugate vaccine. The aim was to improve immune responses against GAC through inclusion of species-specific protein antigens. This led to the selection of GAS proteins MalE, SpyCEP and SpyAD, based on their own favourable immunogenicity and physiochemical properties for polysaccharide attachment. As expected, some differences were observed such as immunogenicity profiles, with SpyAD and SpyCEP generating strong anti-protein IgG titres compared to MalE. SpyAD protein directed immune responses were shown to be functional, with no loss of GAS killing activity following conjugation of SpyAD to GAC. SpyAD was able to generate a modest anti-GAC IgG response following chemical conjugation, similar to but less significant than TT-GAC compared with Alum controls. Although various reaction conditions were tested glycoconjugates could not be obtained for all selected protein carriers. This demonstrates that although protein candidates showed promising immunogenicity, not all proteins are suitable for conjugation despite predicted favourable properties. Protein stability and reactivity, especially for bioconjugation, likely rendered MalE and the SpyCEP fragment tested here as unsuitable candidates for glycoconjugate building.

#### 6.1.2. Overall Glycoconjugate Structure and Immunogenicity from Different Manufacturing Methods

This interesting study is unique comparing two distinct conjugation methods which has not previously been attempted to our knowledge for any pathogen or vaccine. Many previous chemical linkage methods have yielded "mesh-like" glycoconjugates, such as TT attached to large polysaccharides<sup>527</sup>. With respect to a GAS vaccine, GAC polymers are smaller in size compared to other polysaccharides such as Hib and *Neisseria meningitidis* group C capsules in licenced glycoconjugates<sup>677</sup>. Although the linkage method was different for chemical and enzymatic conjugation both methods here required minimal modification of GAC and proteins. Both methods encouraged polysaccharide attachment to protein carriers through the terminal monosaccharide, predicted to present the polymer in "sun-like" structures. Based on a previous study, conjugating GAC to CRM<sub>197</sub>, "mesh-like" structures were superior to "sun-like" structures for an anti-GAC response but was inferior and masked protein directed immune responses for 'double hit' targets<sup>411</sup>. The aim was therefore to preserve immunogenic B cell epitopes along the entire polysaccharide chain length, as well as B and T cell epitopes on protein carriers, important for the immunogenicity of 'double hit' glycoconjugates.

Although both methods here produce “sun-like” conjugates, the decoration location and level of polysaccharide attachment was predicted to be different between methods. EDC chemical conjugation targeted abundant surface exposed lysine residues, but the quantity exposed varies substantially between proteins due to size and folding. Therefore, although highly reactive, the attachment point of a particular lysine is difficult to control leading to random attachment and heterogenous glycoconjugate populations<sup>458,678</sup>. Bioconjugate vaccines meanwhile are more defined due to the specific introduction of glycotags, enabling terminal targeting by recombinant polysaccharides. Such site-specific attachment is thought to overcome batch to batch variation, and the complicated and often costly manufacturing and purification steps attributable to chemical conjugation reactions<sup>679,680</sup>.

From a consistency perspective more defined and characterised vaccines such as bioconjugates are also attractive candidates. Therefore approaches to more closely control chemical conjugation reactions such as modification and unnatural amino acid incorporation<sup>681,682</sup>, as well as introduction of targetable linkers to modify exposing chemical groups (e.g. hydrazides, maleimides, azides, alkynes)<sup>538</sup> have also been investigated. This is in addition to synthetic polysaccharide antigens with defined chain lengths, and reducing end compositions to yield more defined antigen populations<sup>538,683</sup>. Conversely, the heterogeneity of bioconjugates has also been investigated, thought to be superior in some instances due to providing the immune system with multiple epitope presentations. This has been achieved more simply through cloning carrier proteins to contain multiple glycotags, at the peptide terminal, as well as within the structure to produce heavily glycosylated proteins<sup>654</sup>. In these instances, more heterogenous bioconjugate populations may be observed, similar to chemical glycoconjugates, due to conformational preference for particular glycotags over other glycotags depending on the position within the protein structure.

A fuller understanding of antigen specific immunogenicity, relating to overall glycoconjugate structure is required before determining whether heterogenous or homogenous vaccine preparations are preferred. Effective immune stimulation requires preservation of B cell epitopes for recognition of native antigen by B cells for polysaccharide and protein carrier directed responses, which may be different with defined and more random heterogenous preparations<sup>677</sup>. Licenced chemical glycoconjugate vaccines have shown that random heterogenous populations produce effective and protective polysaccharide immune responses<sup>684–686</sup>. Yet homogenous site specific vaccines containing recombinant polysaccharides have

also been shown to be effective in clinical trials<sup>439,548,687</sup>. Greater characterisation of structure–immunogenicity relationships resulting from different conjugation manufacturing methods is necessary to fully understand molecular and cellular immune mechanisms critical for vaccine efficacy. Future attempts should therefore aim to produce like-for-like combinations of protein and polysaccharide antigens, not possible in the current thesis timeframe, to enable better comparisons between manufacturing methods.

### 6.1.3. Limitations of Glycoconjugate Production

All tested glycoconjugate vaccines would have benefited from further purification to remove unconjugated protein and any potential remaining polysaccharide. Separation would aid physicochemical analysis to accurately size glycoconjugates, determine true conjugate polysaccharide to protein ratios, as well as assess whether a boost in anti-polysaccharide immune responses could be observed when limiting interfering anti-protein responses. Improved yields could have enabled investigation of preparative and analytical size exclusion chromatography (SEC) to separate distinct populations containing glycoconjugate molecules from unconjugated protein. In this study using analytical SEC it was possible to show differences in retention times for chemical glycoconjugates SpyAD-GAC and TT-GAS\_Rha, suggesting differences in hydrodynamic radii. However, such differences could not be seen for TT-GAC nor NanA-rpORha. Differences in physicochemical properties, specifically hydrophobicity could be demonstrated for TT-GAC, however limited data could be obtained for NanA-rpORha characterisation due to an inability to obtain stable baselines. Such accurate size determination of the bioconjugate was thought to have been impacted by potential carry over of contaminants from the *E. coli* host cells such as lipopolysaccharide. This was also a general observation for other physicochemical analysis attempted such as H-PAD which was possible for chemical glycoconjugates and their components but not for bioconjugate characterisation. It therefore appears that the current experimental setup, including the selected *E. coli* host strain and purification methods for bioconjugate assembly could not yield concentrated samples pure enough for accurate downstream analysis.

## 6.2. A Preferred Glycoconjugate Manufacturing Method

This work shows for the first time that both glycoconjugate production methods can be applied successfully to GAS antigens, but some general practical limitations were shared by both methods. These included low reactant and product yields

placing constraint on reaction conditions, downstream purification testing, and immunogenicity analysis. For example, for chemical conjugation reactions the concentration of extracted and purified GAC, GAS\_Rha and recombinant GAS protein antigens reduced the scalability of reactions for testing. Similarly, although the need for independent polysaccharide and protein extraction and purification was negated with bioconjugation, scalability issues were also observed with poor NanA-rpORha product yields despite large bacterial culture volumes. This limited the downstream analyses such as physicochemical characterisation, as well as purification testing to remove unconjugated components, a predicted key limitation for lower-than-expected immunogenicity profiles. For both methods there were issues with protein antigen compatibility, either due to instability for chemical conjugation or glycotags being potentially obscured for biological conjugation.

From the perspective of difficulties in downstream analysis mentioned previously biological conjugation appeared to be logistically more difficult to achieve, producing a smaller panel of bioconjugates from the selected component combinations (i.e. NanA-rpORha only). This may be expected given that the technology is in its infancy, with all testing and application to GAS vaccine development completely novel. This is compared to chemical conjugation which has been carried out routinely in currently licenced glycoconjugate vaccines, as well as preclinical GAS specific glycoconjugate vaccine studies<sup>139,141,409,411</sup>. Chemical conjugation methods in this work provided a larger panel of vaccines for testing, demonstrating better compatibility of conjugation components. It also enabled direct immunological differences to be observed, with TT-GAC giving the best anti-polysaccharide response compared to TT-GAS\_Rha and the SpyAD-GAC 'double hit' glycoconjugate after three immunisations.

Despite being a relatively new vaccinology tool, further optimisation of polysaccharide antigens related to enzyme requirements and condition testing holds promise for bioconjugation technology. Such developments could reduce vaccine manufacturing costs as an important consideration in the development of a GAS vaccine where disease burdens are greatest in low- and middle-income countries (LMICs). Additionally, bioconjugation offers the opportunity of tailor-made and rationally designed vaccine candidates that can be quickly assembled and tested. The other advantage is that once tested the cloned vaccines production is consistent, negating the need to extract and purify each component as is the case for chemical conjugation. Therefore, although based on this work and the selected antigens for testing, chemical linkage appears to be the preferred manufacturing

method, bioconjugation has the potential to be not only a valid alternative, but also offer a superior manufacturing method.

## 6.3. Future Perspectives and Remaining Questions

### 6.3.1. Current Limitations to Glycoconjugate Vaccine Analysis

Work in this thesis has contributed not only to the understanding of glycoconjugate manufacturing, but also the GAS vaccine research field more generally. Tested vaccine components have not previously been conjugated chemically using EDC. Neither have GAC derived polymers and GAS antigens been tested in bioconjugation technology. MalE and SpyCEP protein antigens, which although did not show conjugation with either method, provided some insight into PglB reaction requirements. The panel of three antigens provided a good starting point, however, a general study restraint is related to the limited capability to test component compatibility. The vaccinology field more widely would benefit from high throughput reliable testing of vaccine components to identify the most optimal and compatible antigen candidates. Reverse vaccinology approaches have been adopted successfully for antigen selection<sup>395</sup>, as well as recently Design of Experiment (DoE) for finding optimal conjugation conditions<sup>411</sup>. A combinational approach may be most appropriate for obtaining optimal glycoconjugate yields containing the most compatible and immune stimulating antigens.

Successful antigen investigation is also dependent on improved general glycoconjugate characterisation methods. This is to assess compatibility and physicochemical properties, important in determining consistent manufacturing and vaccine stability<sup>646</sup>. The field has seen some improvements in glycoconjugate physicochemical analysis in recent years, such as high-field NMR, particle-size distribution analyses, and cryo-electron microscopy to reveal glycoconjugate structural details<sup>628</sup>. However, due to their nature characterisation can be challenging and a wider breadth of analytical tools are necessary. Challenges are often encountered due to heterogenous populations composed of two or more distinct antigens, with different properties attributable to protein and polysaccharides. This study used a variety of analytical methods stated in Appendix Table 7.4 for glycoconjugate property analysis. For example, for the polysaccharide component, HPLC sizing methods with different detectors, as well as H-PAD for quantification, and NMR for characterisation were used in Chapter 3. However, information on polysaccharide epitope mapping, glycoconjugate molecular weight, and vaccine conformations could not be obtained. Such challenges are faced more

generally when studying the glycome, as well as glycoconjugate vaccine manufacture<sup>628</sup>. Specifically in this work data generation was hindered by low product yields and the inability to remove unconjugated protein species, and bioconjugate contaminants interfering with analysis.

To overcome this obstacle new technologies and scientific techniques are required to further analyse vaccine components and glycoconjugate molecules, to determine properties such as unconjugated and glycosylation levels. Further analysis could include mass spectrometry analysis to correlate polysaccharide composition with glycosylation site to further understand structure – function relationships<sup>688–690</sup>. Sensitivity and selectivity can also be improved with matrix-assisted laser desorption ionisation (MALDI) with time-of-flight (TOF) analyser to determine glycoconjugate molecular weights<sup>691</sup>, component ratios<sup>692</sup>, as well as glycosylation site attachment<sup>693</sup>. Such approaches are in addition to high pressure liquid chromatography techniques, which are also useful in glycoconjugate vaccine quality and stability analysis, particularly through size and conformation determinations<sup>694–696</sup>. Likely a combination of techniques will be necessary for a true understanding of structure and how it relates to immunogenicity, to inform future glycoconjugate design and production more generally.

### 6.3.2. Glycoconjugate Antigen Inclusion Towards Clinical Efficacy Targets.

The overall goal of a GAS vaccine is to reduce GAS pharyngitis and skin infections, as set out by the World Health Organisation's (WHO) roadmap<sup>56</sup>. Vaccine efficacy testing should work towards these clinical endpoints, determining not only optimal vaccine components and characterisation to achieve this, but also assessment of vaccine immunogenicity. Chapter 5 showed immunological differences in the tested glycoconjugate vaccines, as well as variability in polysaccharide specific IgG titres between individual mice, even with immunisation with the positive control glycoconjugate (GAC-CRM<sub>197</sub>). Future studies may therefore benefit from investigating optimal dosing regimens, with new administration routes and effective adjuvants, aiming to improve anti-polysaccharide responses.

This work showed worth in the investigation of 'double hit' glycoconjugate vaccines through SpyAD generated anti-protein responses demonstrating antibody function through GAS cell killing and GAS cell surface binding by flow cytometry. The antibody function from including SpyAD had benefits over TT inclusion through the anti-protein response. Further testing of a widened panel of GAS protein antigens will benefit future investigations as discussed. From an immunogenicity perspective,



given the target clinical endpoints, GAS proteins involved in initial host colonisation and adhesion to cells may be better suited than those implicated in systemic GAS infection which occurs later in host - pathogen interactions. Additionally, the work of others has shown SpyAD conjugated to rhamnose polysaccharide provided minimal protection until combined with other protein antigens (e.g. SLO and C5a peptidase) in a multicomponent formulation<sup>412</sup>. Therefore, the focus should be on carriers to provide most optimal anti-polysaccharide responses whilst retaining protein antibodies, as other proteins can be included in the formulation for general broader GAS strain protection. This also needs to be considered in terms of the practicalities of using such proteins in the manufacturing process, as this work demonstrated that specific conjugation conditions are not suitable for all proteins tested. Specifically, MalE and SpyCEP proteins appeared to be less suited to both chemical and bioconjugation methods, with MalE also showing less favourable immunogenicity profiles.

A greater understanding of what constitutes a good carrier protein is therefore required from both a physicochemical and robust immunogenicity perspective. This will require a deeper understanding of B and T cell activity following immunisation to determine how protein carrier properties relates to favourable vaccine efficacy. Specifically, it is still unclear whether polysaccharide-specific and carrier-specific T cell clones expand following antigen presentation and processing by B cells. It is also unclear whether small peptides or rather glycopeptide fragments are presented to T cells for glycoconjugates containing GAC. Glycoconjugate vaccines containing *S. Typhi* Vi, and Hib polysaccharides conjugated to CRM<sub>197</sub> demonstrated the importance of both the peptide and polysaccharide portions for presentation to induce B - T cell interactions<sup>535</sup>. However, group C *N. meningitidis* polysaccharide conjugated to CRM<sub>197</sub> required peptide presentation only for T cell helper responses<sup>535</sup>. This shows that different presentation mechanisms to induce T cell responses can occur based on polysaccharide structures and possibly conjugation method. Therefore, the mechanism of GAC containing glycoconjugate presentation by APCs needs to be further characterised. This will enable not only better understanding of T cell dependent polysaccharide specific humoral responses, but also more generally allow the production of knowledge-based GAS glycoconjugate vaccines with optimal structure – immunogenicity profiles.

### 6.3.3. Alternative Approaches to Glycoconjugate Vaccines

In recent years novel alternatives to classical protein carriers have been investigated to generate anti-polysaccharide responses. These have included

nanoparticles, virus-like particles, liposomes, Outer Membrane Vesicles (OMVs), and their derivatives Generalized Modules for Membrane Antigens (GMMA)<sup>697</sup>. The general aim of nanotechnology is to improve overall vaccine efficacy through optimising antigen delivery, specifically enhancing antigen stability and better APC targeting<sup>698,699</sup>. Such technologies have provided advantages over classical protein carriers such as being optimally sized for immune system stimulation, possessing their own intrinsic innate immunostimulatory properties triggering toll-like receptors<sup>700,701</sup> as well as favourable B cell activation through presenting multiple copies of polysaccharide antigens<sup>702</sup>. Similar to GAS antigens used as protein carriers in this work, GMMA and OMVs can also immunologically function as both the carrier for polysaccharide epitopes as well as an antigen in its own right when produced from the target organism. They can also be manipulated to express recombinant proteins on their surface in addition to polysaccharide antigens, providing the immune system with alternative antigen conformations compared to traditional glycoconjugates.

Even without a 'double hit' nature, they are proving to be effective immunostimulants and have been investigated as polysaccharide carrier vesicles. GAC antigen presentation in GAS specific GMMA based vaccine platforms through GAC terminally linked via chemical agents to oxidised GMMA lipooligosaccharides have been investigated<sup>415</sup>. The GAC-GMMA vaccine showed equal and in some instances boosted anti-polysaccharide immunogenicity when compared to glycoconjugates containing CRM<sub>197</sub> carrier protein<sup>415,659</sup>. GAS specific nanotechnology has also shown favourable immunogenicity profiles. For example, a lipopeptide based vaccine containing a synthetic single repeat unit of GAC (rhamnose-GlcNAc trisaccharide), forming self-adjuvating lipid-core complexes, demonstrated opsonic and protective anti-GAC antibodies comparable to subunit peptide vaccine J8-lipopeptide<sup>416</sup>. Additionally synthetic GAC derived antigens (tetra and hexa-rhamnose chains) have been conjugated to gold nanoparticles, as an alternative delivery system, showing favourable antibody binding affinity<sup>414</sup>. Such alternative vaccine manufacturing approaches will be important to consider in the future landscape of GAS vaccine development, especially should they prove superior to proteins to induce T cell help and therefore humoral responses required for GAS protection.

## 6.4. Concluding Remarks

Given the limited number of successful, reliable, and affordable glycoconjugate production technology platforms, it is of great interest to investigate the suitability of alternative glycoconjugate vaccine manufacturing approaches. This landmark study has expanded our understanding of the suitability of GAC related polysaccharides, as well as GAS protein antigens, relating to the requirements of two distinct conjugation methodologies. Such work can also be more generally applied to the ongoing efforts to address the huge global presence and large disease burden attributable to GAS, for which a vaccine is most desperately needed. Beyond GAS, vaccines are not available for most bacterial infections. Significantly, the glycoconjugate vaccine technology described in this thesis could also be applied to multiple antibiotic resistant bacteria - the slow pandemic.

## References

1. Lancefield, R. A serological differentiation of human and other groups of hemolytic *Streptococci*. *J Exo Med* **57**, 571–595 (1933).
2. WHO. *The Current Evidence for the Burden of Group A Streptococcal Diseases*. (2005).
3. Beaton, A. *et al.* The American Heart Association’s Call to Action for Reducing the Global Burden of Rheumatic Heart Disease: A Policy Statement From the American Heart Association. *Circulation* **142**, 358–368 (2020).
4. Cunningham, M. W. Pathogenesis of Group A Streptococcal infections. *Clin. Microbiol. Rev.* **13**, 470–511 (2000).
5. Henningham, A., Barnett, T. C., Maamary, P. G. & Walker, M. J. Pathogenesis of Group A Streptococcal infections. *Discov. Med.* **13**, 329–342 (2012).
6. Steer, A. C., Danchin, M. H. & Carapetis, J. R. Group A Streptococcal infections in children. *J. Paediatr. Child Health* **43**, 203–213 (2007).
7. Carapetis, J. R., Steer, A. C., Mulholland, K. E. & Weber, M. Global Burden of Group A Streptococcal disease. *Lancet Infect Dis* **5**, 685–694 (2005).
8. Valery, P. C. *et al.* Skin infections among Indigenous Australians in an urban setting in far North Queensland. *Epidemiol. Infect.* **136**, 1103–1108 (2008).
9. Cole, C. & Gazewood, J. Diagnosis and treatment of impetigo. *Am. Fam. Physician* **75**, 859–864 (2007).
10. O’Sullivan, C. E., Baker, M. G. & Zhang, J. Increasing hospitalizations for serious skin infections in New Zealand children 1990–2007. *Epidemiol. Infect.* **139**, 1794–1804 (2011).
11. Pfoh, E., Wessels, M. R., Goldmann, D. & Lee, G. M. Burden and Economic Cost of Group A Streptococcal Pharyngitis. *Pediatrics* **121**, 229–234 (2008).
12. Cunningham, M. W. Pathogenesis of group A Streptococcal infections and their sequelae. *Adv. Exp. Med. Biol.* **609**, 29–42 (2008).
13. Parks, T., Smeesters, P. & Andrew, S. Streptococcal skin infection and rheumatic heart disease. *Curr. Opin. Infect. Dis.* **25**, 145–153 (2012).
14. Lee, J. L., Naguwa, S. M., Cheema, G. S. & Gershwin, M. E. Acute rheumatic fever and its consequences: a persistent threat to developing nations in the 21st century. *Autoimmun. Rev.* **9**, 117–123 (2009).
15. Auala, T., Lauro, B., Zavale, G., Mbakwem, A. Ç. & Mocumbi, A. O. Acute Rheumatic Fever and Rheumatic Heart Disease: Highlighting the Role of Group A Streptococcus in the Global Burden of Cardiovascular Disease. *Pathog. Dis.* **11**, 1–12 (2022).
16. Diao, M. *et al.* Pregnancy in women with heart disease in sub-Saharan Africa. *Arch. Cardiovasc. Dis.* **104**, 370–374 (2011).
17. Zühlke, L. *et al.* Characteristics, complications, and gaps in evidence-based interventions in rheumatic heart disease: the Global Rheumatic Heart Disease Registry (the REMEDY study). *Eur. Heart J.* **36**, 1115–1122 (2015).
18. Coffey, P. M., Ralph, A. P. & Krause, V. L. The role of social determinants of

- health in the risk and prevention of Group A Streptococcal infection, acute rheumatic fever and rheumatic heart disease: A systematic review. *PLoS Negl. Trop. Dis.* **12**, e0006577 (2018).
19. Watkins, D. A. *et al.* Global, Regional, and National Burden of Rheumatic Heart Disease 1990–2015. *N. Engl. J. Med.* **377**, 713–722 (2017).
  20. Parnaby, M. G. & Carapetis, J. R. Rheumatic fever in indigenous Australian children. *J. Paediatr. Child Health* **46**, 527–533 (2010).
  21. Walker, M. J. *et al.* Disease manifestations and pathogenic mechanisms of Group A *Streptococcus*. *Clin. Microbiol. Rev.* **27**, 264–301 (2014).
  22. Bisno, A. L. Nonsuppurative poststreptococcal sequelae: rheumatic fever and glomerulonephritis. in *Principles and Practice of Infectious Diseases* 2611–2622 (2010).
  23. Marshall, C. S. *et al.* Acute post-streptococcal glomerulonephritis in the Northern Territory of Australia: a review of 16 years data and comparison with the literature. *Am. J. Trop. Med. Hyg.* **85**, 703–710 (2011).
  24. Masuyama, T. *et al.* Outbreak of acute glomerulonephritis in children: observed association with the T1 subtype of Group A Streptococcal infection in northern Kyushu, Japan. *Acta Paediatr. Jpn. Overseas Ed.* **38**, 128–131 (1996).
  25. Zheng, M. *et al.* Genetic Analysis of Group A *Streptococcus* Isolates Recovered during Acute Glomerulonephritis Outbreaks in Guizhou Province of China. *J. Clin. Microbiol.* **47**, 715 LP – 720 (2009).
  26. Garvey, M. A., Giedd, J. & Swedo, S. E. PANDAS: the search for environmental triggers of pediatric neuropsychiatric disorders. Lessons from rheumatic fever. *J. Child Neurol.* **13**, 413–423 (1998).
  27. Shulman, S. T. & Tanz, R. R. Group A streptococcal pharyngitis and immune-mediated complications: from diagnosis to management. *Expert Rev. Anti. Infect. Ther.* **8**, 137–150 (2010).
  28. Katz, A. R. & Morens, D. M. Severe streptococcal infections in historical perspective. *Infect. Dis. Soc. Am.* **14**, 298–307 (1992).
  29. Nelson, G. E. *et al.* Epidemiology of Invasive Group A Streptococcal Infections in the United States 2005-2012. *Clin. Infect. Dis.* **63**, 478–486 (2016).
  30. Ralph, A. P. & Carapetis, J. R. Group A Streptococcal diseases and their global burden. *Curr. Top. Microbiol. Immunol.* **368**, 1–27 (2013).
  31. Quinn, R. W. Comprehensive review of morbidity and mortality trends for rheumatic fever, streptococcal disease, and scarlet fever: the decline of rheumatic fever. *Rev. Infect. Dis.* **11**, 928–953 (1989).
  32. Cliff, A. D., Haggett, P., Smallman-Raynor, M., Stroup, D. F. & Williamson, G. D. The importance of long-term records in public health surveillance: the US weekly sanitary reports 1888-1912 revisited. *J. Public Health Med.* **19**, 76–84 (1997).
  33. Chen, M. *et al.* Outbreak of scarlet fever associated with emm12 type Group A *Streptococcus* in 2011 in Shanghai, China. *Pediatr. Infect. Dis. J.* **31**, 158–162 (2012).

34. Tse, H. *et al.* Molecular Characterization of the 2011 Hong Kong Scarlet Fever Outbreak. *J. Infect. Dis.* **206**, 341–351 (2012).
35. Chalker, V. *et al.* Genome analysis following a national increase in Scarlet Fever in England 2014. *BMC Genomics* **18**, 4–13 (2017).
36. Lamagni, T. L. *et al.* Predictors of death after severe *Streptococcus pyogenes* infection. *Emerg. Infect. Dis.* **15**, 1304–1307 (2009).
37. Lamagni, T. L. *et al.* Epidemiology of severe *Streptococcus pyogenes* disease in Europe. *J. Clin. Microbiol.* **46**, 2359–2367 (2008).
38. O'Grady, K. A. F. *et al.* The epidemiology of invasive Group A Streptococcal disease in Victoria, Australia. *Med. J. Aust.* **186**, 565–569 (2007).
39. O'Loughlin, R. E. *et al.* The epidemiology of invasive Group A Streptococcal infection and potential vaccine implications: United States, 2000-2004. *Clin. Infect. Dis.* **45**, 853–862 (2007).
40. Lappin, E. & Ferguson, A. J. Gram-positive toxic shock syndromes. *Lancet Infect. Dis.* **9**, 281–290 (2009).
41. Public Health England. Group A Streptococcal infections: seasonal activity, 2017/18: third report. *Heal. Prot. Rep.* **12**, (2018).
42. British Columbia Centre for Disease Control. *Invasive Group A Streptococcal disease (iGAS) in British Columbia 2017: annual summary background.* (2017).
43. Centers for Disease Control and Prevention. *Active Bacterial Core surveillance (ABCs) report Emerging Infections Program Network: Group A Streptococcus.* (2016).
44. Stevens, D. L. *et al.* Severe Group A Streptococcal infections associated with a toxic shock-like syndrome and scarlet fever toxin A. *N. Engl. J. Med.* **321**, 1–7 (1989).
45. Cleary, P. P. *et al.* Clonal basis for resurgence of serious *Streptococcus pyogenes* disease in the 1980s. *Lancet (London, England)* **339**, 518–521 (1992).
46. Tyrrell, G. J. *et al.* Epidemic of Group A Streptococcus M/emm59 causing invasive disease in Canada. *Clin. Infect. Dis.* **51**, 1290–1297 (2010).
47. Turner, C. E. *et al.* Molecular analysis of an outbreak of lethal postpartum sepsis caused by *Streptococcus pyogenes*. *J. Clin. Microbiol.* **51**, 2089–2095 (2013).
48. Nouri, B. Z., Venturini, C., Beatson, S. A. & Walker, M. J. Analysis of a *Streptococcus pyogenes* puerperal sepsis cluster by use of whole-genome sequencing. *J. Clin. Microbiol.* **50**, 2224–2228 (2012).
49. Raymond, J., Schlegel, L., Garnier, F. & Bouvet, A. Molecular characterization of *Streptococcus pyogenes* isolates to investigate an outbreak of puerperal sepsis. *Infect. Control Hosp. Epidemiol.* **26**, 455–461 (2005).
50. Berkley, J. A. *et al.* Bacteremia among children admitted to a rural hospital in Kenya. *N. Engl. J. Med.* **352**, 39–47 (2005).
51. Fernandes, G. R. *et al.* Genomic comparison among lethal invasive strains of

- Streptococcus pyogenes* Serotype M1. *Front. Microbiol.* **8**, 1–10 (2017).
52. Steer, A. C. *et al.* Prospective surveillance of invasive Group A Streptococcal disease, Fiji, 2005–2007. *Emerg. Infect. Dis.* **15**, 216–222 (2009).
  53. Steer, A. C. *et al.* High burden of invasive beta-haemolytic streptococcal infections in Fiji. *Epidemiol. Infect.* **136**, 621–627 (2008).
  54. Carapetis, J. R., Walker, A. M., Hibble, M., Sriprakash, K. S. & Currie, B. J. Clinical and epidemiological features of Group A Streptococcal bacteraemia in a region with hyperendemic superficial streptococcal infection. *Epidemiol. Infect.* **122**, 59–65 (1999).
  55. Whitehead, B. D., Smith, H. V & Nourse, C. Invasive Group A Streptococcal disease in children in Queensland. *Epidemiol. Infect.* **139**, 623–628 (2011).
  56. Vekemans, J. *et al.* The path to Group A *Streptococcus* vaccines: WHO research and development technology roadmap and preferred product characteristics. *Clin. Infect. Dis.* **16**, 877–883 (2019).
  57. Hartley, G., Enders, J. F., Mueller, J. H. & Schoenbach, E. B. Absence Of Clinical Disease In Spite Of A High Incidence Of Carriers Of Group A Hemolytic *Streptococci* Of A Single Type; Failure Of Tyrothricin To Influence The Carrier Rate. *J. Clin. Invest.* **24**, 92–96 (1945).
  58. DeMuri, G. P. & Wald, E. R. The Group A Streptococcal Carrier State Reviewed: Still an Enigma. *J. Pediatric Infect. Dis. Soc.* **3**, 336–342 (2014).
  59. Roberts, A. L. *et al.* Detection of Group A Streptococcus in tonsils from pediatric patients reveals high rate of asymptomatic streptococcal carriage. *BMC Pediatr.* **12**, 1–9 (2012).
  60. Shaikh, N., Leonard, E. & Martin, J. M. Prevalence of Streptococcal Pharyngitis and Streptococcal Carriage in Children: A Meta-analysis. *Pediatrics* **126**, 557–564 (2010).
  61. Centers for Disease Control and Prevention. Nosocomial Group A Streptococcal infections associated with asymptomatic health-care workers--Maryland and California, 1997. *MMWR. Morb. Mortal. Wkly. Rep.* **48**, 163–166 (1997).
  62. Brook, I. & Gober, A. E. Recovery of interfering and beta-lactamase-producing bacteria from Group A beta-haemolytic *Streptococci* carriers and non-carriers. *J. Med. Microbiol.* **55**, 1741–1744 (2006).
  63. Brook, I. The role of beta-lactamase-producing bacteria in the persistence of streptococcal tonsillar infection. *Rev. Infect. Dis.* **6**, 601–607 (1984).
  64. Brook, I. Overcoming penicillin failures in the treatment of Group A streptococcal pharyngo-tonsillitis. *Int. J. Pediatr. Otorhinolaryngol.* **71**, 1501–1508 (2007).
  65. Brook, I. Failure of penicillin to eradicate group A beta-hemolytic streptococci tonsillitis: causes and management. *J. Otolaryngol.* **30**, 324–329 (2001).
  66. Pichichero, M. E. Characteristics, development, and clinical trials. *Hum. Vaccin. Immunother.* **9**, 2505–2523 (2013).
  67. Crowe, C., Sanders, W. E. J. & Longley, S. Bacterial interference. Role of the normal throat flora in prevention of colonization by Group A *Streptococcus*. *J. Infect. Dis.* **128**, 527–532 (1973).

68. Osterlund, A., Popa, R., Nikkilä, T., Scheynius, A. & Engstrand, L. Intracellular reservoir of *Streptococcus pyogenes in vivo*: a possible explanation for recurrent pharyngotonsillitis. *Laryngoscope* **107**, 640–647 (1997).
69. Osterlund, A. & Engstrand, L. An intracellular sanctuary for *Streptococcus pyogenes* in human tonsillar epithelium studies of asymptomatic carriers and *in vitro* cultured biopsies. *Acta Otolaryngol.* **117**, 883–888 (1997).
70. Kaplan, E. L., Chhatwal, G. S. & Rohde, M. Reduced ability of penicillin to eradicate ingested Group A *Streptococci* from epithelial cells: clinical and pathogenetic implications. *Clin. Infect. Dis.* **43**, 1398–1406 (2006).
71. Marouni, M. J., Barzilai, A., Keller, N., Rubinstein, E. & Sela, S. Intracellular survival of persistent Group A *Streptococci* in cultured epithelial cells. *Int. J. Med. Microbiol.* **294**, 27–33 (2004).
72. Conley, J. *et al.* Biofilm formation by Group A *Streptococci*: is there a relationship with treatment failure? *J. Clin. Microbiol.* **41**, 4043–4048 (2003).
73. Ogawa, T. *et al.* Biofilm formation or internalization into epithelial cells enable *Streptococcus pyogenes* to evade antibiotic eradication in patients with pharyngitis. *Microb. Pathog.* **51**, 58–68 (2011).
74. Veasy, L. G. *et al.* Resurgence of acute rheumatic fever in the intermountain area of the United States. *N. Engl. J. Med.* **316**, 421–427 (1987).
75. Wald, E. R., Dashefsky, B., Feidt, C., Chiponis, D. & Byers, C. Acute rheumatic fever in western Pennsylvania and the tristate area. *Pediatrics* **80**, 371–374 (1987).
76. Hamburger Jr., M., Puck, T. T., Hamburger, V. G. & Johnson, M. A. Studies on the Transmission of Hemolytic *Streptococcus* Infections: Hemolytic *Streptococci* in the Air, Floor Dust, and Bedclothing of Hospital Wards and their Relation to Cross Infection. *J. Infect. Dis.* **75**, 71–78 (1944).
77. Hamburger Jr., M. & Robertson, O. H. Expulsion of Group A hemolytic *Streptococci* in droplets and droplet nuclei by sneezing, coughing and talking. *Am. J. Med.* **4**, 690–701 (1948).
78. Shelburne, S. A. 3rd *et al.* Growth characteristics of and virulence factor production by Group A *Streptococcus* during cultivation in human saliva. *Infect. Immun.* **73**, 4723–4731 (2005).
79. Colebrook, L., Maxted, W. R. & Johns, A. M. The presence of haemolytic and other *Streptococci* on human skin. *J. Pathol. Bacteriol.* **41**, 521–527 (1935).
80. Wasserzug, O. *et al.* A cluster of ecthyma outbreaks caused by a single clone of invasive and highly infective *Streptococcus pyogenes*. *Clin. Infect. Dis.* **48**, 1213–1219 (2009).
81. Centers for Disease Control and Prevention. Outbreak of Group A *Streptococcal* pneumonia among Marine Corps recruits-California, November 1-December 20, 2002. *MMWR. Morb. Mortal. Wkly. Rep.* **52**, 106–109 (2003).
82. Sumbly, P. *et al.* Evolutionary origin and emergence of a highly successful clone of serotype M1 Group A *Streptococcus* involved multiple horizontal gene transfer events. *J. Infect. Dis.* **192**, 771–782 (2005).



83. Aziz, R. K. & Kotb, M. Rise and persistence of global M1T1 clone of *Streptococcus pyogenes*. *Emerg. Infect. Dis.* **14**, 1511–1517 (2008).
84. Yang, P. *et al.* Characteristics of Group A *Streptococcus* strains circulating during scarlet fever epidemic, Beijing, China, 2011. *Emerg. Infect. Dis.* **19**, 909–915 (2013).
85. Lau, E. H. Y., Nishiura, H., Cowling, B. J., Ip, D. K. M. & Wu, J. T. Scarlet fever outbreak, Hong Kong, 2011. *Emerging infectious diseases* **18**, 1700–1702 (2012).
86. Hsieh, Y. C. & Huang, Y. C. Scarlet fever outbreak in Hong Kong, 2011. *Journal of Microbiology, Immunology and Infection* **44**, 409–411 (2011).
87. Smeesters, P. R. *et al.* Differences between Belgian and Brazilian Group A *Streptococcus* epidemiologic landscape. *PLoS One* **1**, 1–5 (2006).
88. Lancefield, R. C. The Antigenic Complex Of *Streptococcus Haemolyticus* : I. Demonstration Of A Type-Specific Substance In Extracts Of *Streptococcus Haemolyticus*. *J. Exp. Med.* **47**, 91–103 (1928).
89. Fischetti, V. A. Streptococcal M protein: molecular design and biological behavior. *Clin. Microbiol. Rev.* **2**, 285–314 (1989).
90. Bisno, A. L., Brito, M. O. & Collins, C. M. Molecular basis of Group A Streptococcal virulence. *Lancet Infect. Dis.* **3**, 191–200 (2003).
91. Colman, G., Tanna, A., Efstratiou, A. & Gaworzewska, E. T. The serotypes of *Streptococcus pyogenes* present in Britain during 1980-1990 and their association with disease. *J. Med. Microbiol.* **39**, 165–178 (1993).
92. Johnson, D. R., Stevens, D. L. & Kaplan, E. L. Epidemiologic analysis of Group A Streptococcal serotypes associated with severe systemic infections, rheumatic fever, or uncomplicated pharyngitis. *J. Infect. Dis.* **166**, 374–382 (1992).
93. Schwartz, B., Facklam, R. R. & Breiman, R. F. Changing epidemiology of Group A Streptococcal infection in the USA. *Lancet Infect Dis* **336**, 1167–1171 (1990).
94. Beall, B., Facklam, R. & Thompson, T. Sequencing *emm*-specific PCR products for routine and accurate typing of Group A *Streptococci*. *J. Clin. Microbiol.* **34**, 953–958 (1996).
95. Whatmore, A. M., Kapur, V., Sullivan, D. J., Musser, J. M. & Kehoe, M. A. Non-congruent relationships between variation in *emm* gene sequences and the population genetic structure of Group A *Streptococci*. *Mol. Microbiol.* **14**, 619–631 (1994).
96. Kaufhold, A. *et al.* Rapid typing of Group A *Streptococci* by the use of DNA amplification and non-radioactive allele-specific oligonucleotide probes. *FEMS Microbiol. Lett.* **119**, 19–25 (1994).
97. Relf, W. A., Martin, D. R. & Sriprakash, K. S. Identification of sequence types among the M-nontypeable Group A *Streptococci*. *J. Clin. Microbiol.* **30**, 3190–3194 (1992).
98. McMillan, D. J. *et al.* Updated model of Group A *Streptococcus* M proteins based on a comprehensive worldwide study. *Clin. Microbiol. Infect.* **19**, 222–229 (2013).

99. Sanderson-Smith, M. *et al.* A systematic and functional classification of *Streptococcus pyogenes* that serves as a new tool for molecular typing and vaccine development. *J. Infect. Dis.* **210**, 1325–1338 (2014).
100. Koller, T. *et al.* Typing of the pilus-protein-encoding FCT region and biofilm formation as novel parameters in epidemiological investigations of *Streptococcus pyogenes* isolates from various infection sites. *J Med Microbiol* **59**, 442–452 (2010).
101. Falugi, F. *et al.* Sequence Variation in Group A *Streptococcus* Pili and Association of Pilus Backbone Types with Lancefield T Serotypes. *J. Infect. Dis.* **198**, 1834–1841 (2008).
102. Kratovac, Z., Manoharan, A., Luo, F., Lizano, S. & Bessen, D. E. Population genetics and linkage analysis of loci within the FCT region of *Streptococcus pyogenes*. *J. Bacteriol.* **189**, 1299–1310 (2007).
103. Beall, B. *et al.* *emm* and *sof* gene sequence variation in relation to serological typing of opacity-factor-positive Group A Streptococci. *Microbiology* **146**, 1195–1209 (2000).
104. Cole, J. N., Henningham, A., Gillen, C. M., Ramachandran, V. & Walker, M. J. Human pathogenic streptococcal proteomics and vaccine development. *Proteomics Clin. Appl.* **2**, 387–410 (2008).
105. Smeesters, P. R., McMillan, D. J., Sriprakash, K. S. & Georgousakis, M. M. Differences among Group A *Streptococcus* epidemiological landscapes: Consequences for M protein-based vaccines? *Expert Rev. Vaccines* **8**, 1705–1720 (2009).
106. Bessen *et al.* Molecular epidemiology and genomics of Group A Streptococcus. *Infect. Genet. Evol.* **33**, 393–418 (2015).
107. Kaplan, E. L., Wotton, J. T. & Johnson, D. R. Dynamic epidemiology of Group A Streptococcal serotypes associated with pharyngitis. *Lancet* **358**, 1334–1337 (2001).
108. Steer, A. C., Law, I., Matatolu, L., Beall, B. W. & Carapetis, J. R. Global *emm* type distribution of Group A *Streptococci*: systematic review and implications for vaccine development. *Lancet Infect. Dis.* **9**, 611–616 (2009).
109. Ekelund, K. *et al.* Variations in *emm* Type among Group A Streptococcal Isolates Causing Invasive or Noninvasive Infections in a Nationwide Study. *J. Clin. Microbiol.* **43**, 3101–3109 (2005).
110. Espinosa, L. E. *et al.* M Protein Gene Type Distribution among Group A Streptococcal Clinical Isolates Recovered in Mexico City, Mexico, from 1991 to 2000, and Durango, Mexico, from 1998 to 1999: Overlap with Type Distribution within the United States. *J. Clin. Microbiol.* **41**, 373–378 (2003).
111. Liang, Y. *et al.* Epidemiological and molecular characteristics of clinical isolates of *Streptococcus pyogenes* collected between 2005 and 2008 from Chinese children. *J. Med. Microbiol.* **61**, 975–983 (2012).
112. Shea, P. R. *et al.* Group A Streptococcus *emm* gene types in pharyngeal isolates, Ontario, Canada, 2002–2010. *Emerg. Infect. Dis.* **17**, 2010–2017 (2011).
113. Wozniak, A. *et al.* M-protein gene-type distribution and hyaluronic acid capsule in Group A *Streptococcus* clinical isolates in Chile: association of

- emm* gene markers with *csrR* alleles. *Epidemiol. Infect.* **140**, 1286–1295 (2012).
114. Dhanda, V., Vohra, H. & Kumar, R. Group A *Streptococcus* virulence factors genes in north India & their association with *emm* type in pharyngitis. *Indian J. Med. Res.* **133**, 110–115 (2011).
  115. Kumar, R. *et al.* *Streptococcus pyogenes* pharyngitis & impetigo in a rural area of Panchkula district in Haryana, India. *The Indian journal of medical research* **135**, 133–136 (2012).
  116. Kumar, R. *et al.* Epidemiology of Group A Streptococcal pharyngitis & impetigo: a cross-sectional & follow up study in a rural community of northern India. *Indian J. Med. Res.* **130**, 765–771 (2009).
  117. Tartof, S. Y. *et al.* Factors associated with Group A *Streptococcus emm* type diversification in a large urban setting in Brazil: a cross-sectional study. *BMC Infect. Dis.* **10**, 327 (2010).
  118. Smoot, J. C. *et al.* Genome sequence and comparative microarray analysis of serotype M18 Group A *Streptococcus* strains associated with acute rheumatic fever outbreaks. *Proc. Natl. Acad. Sci.* **99**, 4668 LP – 4673 (2002).
  119. Bessen, D. E., Sotir, C. M., Readdy, T. L. & Hollingshead, S. K. Genetic correlates of throat and skin isolates of Group A Streptococci. *J. Infect. Dis.* **173**, 896–900 (1996).
  120. Hollingshead, S. K., Arnold, J., Readdy, T. L. & Bessen, D. E. Molecular evolution of a multigene family in group A streptococci. *Mol. Biol. Evol.* **11**, 208–219 (1994).
  121. Hollingshead, S. K., Fischetti, V. A. & Scott, J. R. Complete nucleotide sequence of type 6 M protein of the Group A *Streptococcus*. *J. Biol. Chem.* **261**, 1677–1686 (1986).
  122. Banks, D. J. *et al.* Progress toward characterization of the Group A *Streptococcus* metagenome: complete genome sequence of a macrolide-resistant serotype M6 strain. *J. Infect. Dis.* **190**, 727–738 (2004).
  123. Beres, S. B. *et al.* Genome sequence of a serotype M3 strain of Group A *Streptococcus*: Phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proc. Natl. Acad. Sci.* **99**, 10078–10083 (2002).
  124. Ferretti, J. J. *et al.* Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 4658–4663 (2001).
  125. Green, N. M. *et al.* Genome sequence of a serotype M28 strain of Group A *Streptococcus*: potential new insights into puerperal sepsis and bacterial disease specificity. *J. Infect. Dis.* **192**, 760–770 (2005).
  126. McShan, W. M. *et al.* Genome Sequence of a Nephritogenic and Highly Transformable M49 Strain of *Streptococcus pyogenes*. *J. Bacteriol.* **190**, 7773–7785 (2008).
  127. Nakagawa, I. *et al.* Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. *Genome Res.* **13**, 1042–1055 (2003).
  128. Sumbly, P. *et al.* Evolutionary Origin and Emergence of a Highly Successful

- Clone of Serotype M1 Group A Streptococcus Involved Multiple Horizontal Gene Transfer Events. *J. Infect. Dis.* **192**, 771–782 (2005).
129. Tsatsaronis, J. A. *et al.* Streptococcal collagen-like protein A and general stress protein 24 are immunomodulating virulence factors of Group A *Streptococcus*. *FASEB J.* **27**, 2633–2643 (2013).
  130. Langshaw, E. L., Pandey, M. & Good, M. F. Cellular interactions of CovR/S mutant Group A Streptococci. *Microbes Infect.* **20**, 531–535 (2018).
  131. Courtney, H. S., Hasty, D. L. & Dale, J. B. Molecular mechanisms of adhesion, colonization, and invasion of Group A Streptococci. *Ann. Med.* **34**, 77–87 (2002).
  132. Albertí, S., Ashbaugh, C. D. & Wessels, M. R. Structure of the *has* operon promoter and regulation of hyaluronic acid capsule expression in Group A *Streptococcus*. *Mol. Microbiol.* **28**, 343–353 (1998).
  133. Cole, J. N. *et al.* M protein and hyaluronic acid capsule are essential for *in vivo* selection of CovRS mutations characteristic of invasive serotype M1T1 Group A *Streptococcus*. *MBio* **1**, (2010).
  134. Ashbaugh, C. D., Warren, H. B., Carey, V. J. & Wessels, M. R. 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**, 550–560 (1998).
  135. Maclennan, A. P. The production of capsules, hyaluronic acid and hyaluronidase by Group A and Group C Streptococci. *J. Gen. Microbiol.* **14**, 134–142 (1956).
  136. Flores, A. R. *et al.* Asymptomatic carriage of Group A streptococcus is associated with elimination of capsule production. *Infect. Immun.* **82**, 3958–3967 (2014).
  137. Mccarty, M. The lysis of Group A hemolytic *Streptococci* by extracellular enzymes of *Streptomyces albus*. Nature of the cellular substrate attacked by the lytic enzymes. *J. Exp. Med.* **96**, 569–580 (1952).
  138. McCarty, M. & Lancefield, R. Variation in the group-specific carbohydrate of Group A *Streptococci*. Immunochemical studies on the carbohydrates of variant strains. *J. Exp. Med.* **102**, 11–28 (1955).
  139. Kabanova, A. *et al.* Evaluation of a Group A *Streptococcus* synthetic oligosaccharide as vaccine candidate. *Vaccine* **29**, 104–114 (2010).
  140. Sabharwal, H. *et al.* Group A *Streptococcus* (GAS) Carbohydrate as an Immunogen for Protection against GAS Infection. *J. Infect. Dis.* **193**, 129–135 (2006).
  141. van Sorge, N. M. *et al.* The Classical Lancefield Antigen of Group A *Streptococcus* is a Virulence Determinant with Implications for Vaccine Design. *Cell Host Microbe* **15**, 729–740 (2014).
  142. van der Beek, S. L. *et al.* GacA is essential for Group A *Streptococcus* and defines a new class of monomeric dTDP-4-dehydrorhamnose reductases (RmlD). *Mol. Microbiol.* **98**, 946–962 (2015).
  143. Henningham, A. *et al.* Virulence role of the GlcNAc side chain of the Lancefield cell wall carbohydrate antigen in non-M1-serotype Group A

- Streptococcus*. *MBio* **9**, 1–12 (2018).
144. Stuike-Prill, R. & Pinto, B. M. Conformational analysis of oligosaccharides corresponding to the cell-wall polysaccharide of the *Streptococcus* Group A by Metropolis Monte Carlo simulations. *Carbohydr. Res.* **279**, 59–73 (1995).
  145. Rush, J. S. *et al.* The molecular mechanism of N-acetylglucosamine side-chain attachment to the Lancefield Group A carbohydrate in *Streptococcus pyogenes*. *J. Biol. Chem.* **292**, 19441–19457 (2017).
  146. Kreis, U. C., Varma, V. & Pinto, B. M. 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. *Int. J. Biol. Macromol.* **17**, 117–130 (1995).
  147. McCarty, M. & Lancefield, R. Variation in the group-specific carbohydrate of Group A Streptococci I. Immunochemical studies on the carbohydrates of variant strains. *J. Exp. Med.* **102**, 11–28 (1955).
  148. Hamada, S. *et al.* Isolation and immunochemical characterization of carbohydrate antigens prepared from Group A *Streptococcus pyogenes*. *Zentralbl Bakteriol Mikrobiol Hyg A* **256**, 37–48 (1983).
  149. Edgar, R. J. *et al.* Discovery of glycerol phosphate modification on Streptococcal rhamnose polysaccharides. *Nat. Chem. Biol.* **15**, 463–471 (2019).
  150. Davies, M. R. *et al.* Atlas of Group A Streptococcal vaccine candidates compiled using large-scale comparative genomics. *Nat. Genet.* **51**, 1035–1043 (2019).
  151. Zorzoli, A. *et al.* Group A, B, C, and G *Streptococcus* Lancefield antigen biosynthesis is initiated by a conserved  $\alpha$ -D-GlcNAc- $\beta$ -1,4-L-rhamnosyltransferase. *J. Biol. Chem.* **294**, 15237–15256 (2019).
  152. Mistou, M. Y., Sutcliffe, I. C. & Van Sorge, N. M. Bacterial glycobiology: Rhamnose-containing cell wall polysaccharides in Gram-positive bacteria. *FEMS Microbiol. Rev.* **40**, 464–479 (2016).
  153. Barnett, T. C. & Scott, J. R. Differential recognition of surface proteins in *Streptococcus pyogenes* by two sortase gene homologs. *J. Bacteriol.* **184**, 2181–2191 (2002).
  154. Scott, J. R. & Barnett, T. C. Surface proteins of Gram-positive bacteria and how they get there. *Annu. Rev. Microbiol.* **60**, 397–423 (2006).
  155. Barnett, T. C., Patel, A. R. & Scott, J. R. A novel sortase, SrtC2, from *Streptococcus pyogenes* anchors a surface protein containing a QVPTGV motif to the cell wall. *J. Bacteriol.* **186**, 5865–5875 (2004).
  156. Lei, B., Liu, M., Chesney, G. L. & Musser, J. M. Identification of New Candidate Vaccine Antigens Made by *Streptococcus pyogenes*: Purification and Characterization of 16 Putative Extracellular Lipoproteins. *J. Infect. Dis.* **189**, 79–89 (2007).
  157. Nobbs, A. H., Lamont, R. J. & Jenkinson, H. F. *Streptococcus* adherence and colonization. *Microbiol. Mol. Biol. Rev.* **73**, 407–50 (2009).
  158. Hastly, D. L., Ofek, I., Courtney, H. S. & Doyle, R. J. Multiple adhesins of Streptococci. *Infect. Immun.* **60**, 2147–2152 (1992).

159. Mora, M. *et al.* Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. *Proc. Natl. Acad. Sci.* **102**, 15641–15646 (2005).
160. Falugi, F. *et al.* Sequence Variation in Group A *Streptococcus* Pili and Association of Pilus Backbone Types with Lancefield T Serotypes. *J. Infect. Dis.* **198**, 1834–1841 (2008).
161. Kang, H. J., Coulibaly, F., Proft, T. & Baker, E. N. Crystal Structure of Spy0129, a *Streptococcus pyogenes* Class B Sortase Involved in Pilus Assembly. *PLoS One* **6**, 159–169 (2011).
162. Quigley, B. R., Zähler, D., Hatkoff, M., Thanassi, D. G. & Scott, J. R. Linkage of T3 and Cpa pilins in the *Streptococcus pyogenes* M3 pilus. *Mol. Microbiol.* **72**, 1379–1394 (2009).
163. Abbot, E. L. *et al.* Pili mediate specific adhesion of *Streptococcus pyogenes* to human tonsil and skin. *Cell. Microbiol.* **9**, 1822–1833 (2007).
164. Crotty Alexander, L. E. *et al.* M1T1 Group A *Streptococcal* pili promote epithelial colonization but diminish systemic virulence through neutrophil extracellular entrapment. *J. Mol. Med. (Berl)*. **88**, 371–381 (2010).
165. Smith, W. D. *et al.* Roles of minor pilin subunits Spy0125 and Spy0130 in the serotype M1 *Streptococcus pyogenes* strain SF370. *J. Bacteriol.* **192**, 4651–4659 (2010).
166. Kreikemeyer, B. *et al.* *Streptococcus pyogenes* Collagen Type I-Binding Cpa Surface Protein: Expression Profile, Binding Characteristics, Biological Functions, And Potential Clinical Impact. *J. Biol. Chem.* **280**, 33228–33239 (2005).
167. Lizano, S., Luo, F. & Bessen, D. E. Role of *Streptococcal* T antigens in superficial skin infection. *J. Bacteriol.* **189**, 1426–1434 (2007).
168. Nakata, M. *et al.* Mode of expression and functional characterization of FCT-3 pilus region-encoded proteins in *Streptococcus pyogenes* serotype M49. *Infect. Immun.* **77**, 32–44 (2009).
169. Smeesters, P. R., McMillan, D. J. & Sriprakash, K. S. The *Streptococcal* M protein: A highly versatile molecule. *Trends Microbiol.* **18**, 275–282 (2010).
170. Ellen, P. R. & Gibbons, R. J. Parameters Affecting the Adherence and Tissue Tropisms of *Streptococcus pyogenes*. *Infect. Immun.* **9**, 85–91 (1974).
171. Beachey, E. H. & Simpson, W. A. The adherence of Group A *Streptococci* to oropharyngeal cells: the lipoteichoic acid adhesin and fibronectin receptor. *Infection* **10**, 107–111 (1982).
172. Rohde, M. & Cleary, P. P. Adhesion and invasion of *Streptococcus pyogenes* into host cells and clinical relevance of intracellular *Streptococci*. in *Streptococcus pyogenes: Basic Biology to Clinical Manifestations* (eds. Ferretti, J. J., Stevens, D. L. & Fischetti, V. A.) 1–30 (2016).
173. Hanski, E. & Caparon, M. Protein F, a fibronectin-binding protein, is an adhesin of the Group A *Streptococcus*. *Proc. Natl. Acad. Sci.* **89**, 6172–6176 (1992).
174. Talay, S. R., Valentin-Weigand, P., Jerlström, P. G., Timmis, K. N. & Chhatwal, G. S. Fibronectin-binding protein of *Streptococcus pyogenes*:

- sequence of the binding domain involved in adherence of Streptococci to epithelial cells. *Infect. Immun.* **60**, 3837–3844 (1992).
175. Hanski, E., Horwitz, P. A. & Caparon, M. G. Expression of protein F, the fibronectin-binding protein of *Streptococcus pyogenes* JRS4, in heterologous Streptococcal and enterococcal strains promotes their adherence to respiratory epithelial cells. *Infect. Immun.* **60**, 5119–5125 (1992).
  176. Okada, N., Pentland, A. P., Falk, P. & Caparon, M. G. M protein and protein F act as important determinants of cell-specific tropism of *Streptococcus pyogenes* in skin tissue. *J. Clin. Invest.* **94**, 965–977 (1994).
  177. Margarit, I. *et al.* Capturing host-pathogen interactions by protein microarrays: identification of novel Streptococcal proteins binding to human fibronectin, fibrinogen, and C4BP. *FASEB J.* **23**, 3100–3112 (2009).
  178. Walden, M. *et al.* An internal thioester in a pathogen surface protein mediates covalent host binding. *Elife* **4**, e06638 (2015).
  179. Rocha, C. L. & Fischetti, V. A. Identification and Characterization of a Novel Fibronectin-Binding Protein on the Surface of Group A Streptococci. *Infect. Immun.* **67**, 2720–2728 (1999).
  180. Courtney, H. S., Dale, J. B. & Hasty, D. I. Differential effects of the Streptococcal fibronectin-binding protein, FBP54, on adhesion of Group A *Streptococci* to human buccal cells and HEp-2 tissue culture cells. *Infect. Immun.* **64**, 2415–2419 (1996).
  181. Jeng, A. *et al.* Molecular genetic analysis of a Group A *Streptococcus* operon encoding serum opacity factor and a novel fibronectin-binding protein, SfbX. *J. Bacteriol.* **185**, 1208–1217 (2003).
  182. Schwarz-Linek, U. *et al.* Pathogenic bacteria attach to human fibronectin through a tandem beta-zipper. *Nature* **423**, 177–181 (2003).
  183. Norris, N. C. *et al.* Structural and functional analysis of the tandem  $\beta$ -zipper interaction of a Streptococcal protein with human fibronectin. *J. Biol. Chem.* **286**, 38311–38320 (2011).
  184. Lukomski, S., Bachert, B. A., Squeglia, F. & Berisio, R. Collagen-like proteins of pathogenic Streptococci. *Mol. Microbiol.* **103**, 919–930 (2017).
  185. Zutter, M. M. & Santoro, S. A. Widespread histologic distribution of the alpha 2 beta 1 integrin cell-surface collagen receptor. *Am. J. Pathol.* **137**, 113–120 (1990).
  186. Popova, S. N., Lundgren-Akerlund, E., Wiig, H. & Gullberg, D. Physiology and pathology of collagen receptors. *Acta Physiol.* **190**, 179–187 (2007).
  187. Humtsoe, J. O. *et al.* A Streptococcal Collagen-like Protein Interacts with the Alpha 2 beta 1 Integrin and Induces Intracellular Signaling. *J. Biol. Chem.* **280**, 13848–13857 (2005).
  188. Caswell, C. C. *et al.* Identification of the first prokaryotic collagen sequence motif that mediates binding to human collagen receptors, integrins alpha 2 beta 1 and alpha 11 beta 1. *J. Biol. Chem.* **283**, 36168–36175 (2008).
  189. Oliver-Kozup, H. *et al.* The Group A Streptococcal collagen-like protein-1, Scl1, mediates biofilm formation by targeting the extra domain A-containing variant of cellular fibronectin expressed in wounded tissue. *Mol. Microbiol.* **87**,

672–689 (2013).

190. Caswell, C. C., Oliver-Kozup, H., Han, R., Lukomska, E. & Lukomski, S. Sc11, the multifunctional adhesin of Group A *Streptococcus*, selectively binds cellular fibronectin and laminin, and mediates pathogen internalization by human cells. *FEMS Microbiol. Lett.* **303**, 61–68 (2010).
191. Horstmann, R. D., Sievertsen, H. J., Knobloch, J. & Fischetti, V. A. Antiphagocytic activity of Streptococcal M protein: selective binding of complement control protein factor H. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 1657–1661 (1988).
192. Bisno, A. L. Alternate complement pathway activation by Group A *Streptococci*: role of M-protein. *Infect. Immun.* **26**, 1172–1176 (1979).
193. Akesson, P., Moritz, L., Truedsson, M., Christensson, B. & von Pawel-Rammingen, U. IdeS, a highly specific immunoglobulin G (IgG)-cleaving enzyme from *Streptococcus pyogenes*, is inhibited by specific IgG antibodies generated during infection. *Infect. Immun.* **74**, 497–503 (2006).
194. Haapasalo, K. *et al.* Acquisition of complement factor H is important for pathogenesis of *Streptococcus pyogenes* infections: evidence from bacterial *in vitro* survival and human genetic association. *J. Immunol.* **188**, 426–435 (2012).
195. Buffalo, C. Z. *et al.* Conserved patterns hidden within Group A *Streptococcus* M protein hypervariability recognize human C4b-binding protein. *Nat. Microbiol.* **1**, 16155 (2016).
196. Suvilehto, J. *et al.* Binding of complement regulators factor H and C4b binding protein to Group A Streptococcal strains isolated from tonsillar tissue and blood. *Microbes Infect.* **10**, 757–763 (2008).
197. Podbielski, A., Schnitzler, N., Beyhs, P. & Boyle, M. D. P. M-related protein (Mrp) contributes to Group A Streptococcal resistance to phagocytosis by human granulocytes. *Mol. Microbiol.* **19**, 429–441 (1996).
198. Stenberg, L., O'Toole, P. & Lindahl, G. Many Group A Streptococcal strains express two different immunoglobulin-binding proteins, encoded by closely linked genes: characterization of the proteins expressed by four strains of different M-type. *Mol. Microbiol.* **6**, 1185–1194 (1992).
199. Pawel-Rammingen, U. von. Streptococcal IdeS and its impact on immune response and inflammation. *J. Innate Immun.* **4**, 132–140 (2012).
200. Fernie-King, B. A. *et al.* Streptococcal inhibitor of complement (SIC) inhibits the membrane attack complex by preventing uptake of C5b67 onto cell membranes. *Immunology* **103**, 390–398 (2001).
201. Cleary, P. P., Prahbu, U., Dale, J. B., Wexler, D. E. & Handley, J. Streptococcal C5a peptidase is a highly specific endopeptidase. *Infect. Immun.* **60**, 5219–5223 (1992).
202. Wexler, D. E., Chenoweth, D. E. & Cleary, P. P. Mechanism of action of the Group A Streptococcal C5a inactivator. *Proc. Natl. Acad. Sci.* **82**, 8144–8148 (1985).
203. Turner, C. E., Kurupati, P., Jones, M. D., Edwards, R. J. & Sriskandan, S. Emerging Role of the Interleukin-8 Cleaving Enzyme SpyCEP in Clinical *Streptococcus pyogenes* Infection. *J. Infect. Dis.* **200**, 555–563 (2009).



204. Kapur, V., Majesky, M. W., Li, L. L., Black, R. A. & Musser, J. M. Cleavage of interleukin 1 beta (IL-1 beta) precursor to produce active IL-1 beta by a conserved extracellular cysteine protease from *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci.* **90**, 7676–7680 (1993).
205. Herwald, H., Collin, M., Müller-Esterl, W. & Björck, L. Streptococcal cysteine proteinase releases kinins: a virulence mechanism. *J. Exp. Med.* **184**, 665–673 (1996).
206. Yukino, W. *et al.* Cysteine Protease Activity and Histamine Release from the Human Mast Cell Line HMC-1 Stimulated by Recombinant Streptococcal Pyrogenic Exotoxin B/Streptococcal Cysteine Protease. *Infect. Immun.* **70**, 3944–3947 (2002).
207. Collin, M. & Olsén, A. Effect of SpeB and EndoS from *Streptococcus pyogenes* on Human Immunoglobulins. *Infect. Immun.* **69**, 7187–7189 (2001).
208. Turner, C. E., Kurupati, P., Wiles, S., Edwards, R. J. & Sriskandan, S. Impact of immunization against SpyCEP during invasive disease with two streptococcal species: *Streptococcus pyogenes* and *Streptococcus equi*. *Vaccine* **27**, 4923–4929 (2009).
209. Sierig, G., Cywes-Bentley, C., Wessels, M. & Ashbaugh, C. Cytotoxic Effects of Streptolysin O and Streptolysin S Enhance the Virulence of Poorly Encapsulated Group A Streptococci. *Infect. Immun.* **71**, 446–455 (2003).
210. Manetti, A. G. O. *et al.* *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. *Mol. Microbiol.* **64**, 968–983 (2007).
211. Becherelli, M. *et al.* The ancillary protein 1 of *Streptococcus pyogenes* FCT-1 pili mediates cell adhesion and biofilm formation through heterophilic as well as homophilic interactions. *Mol. Microbiol.* **83**, 1035–1047 (2012).
212. Kimura, K. R. *et al.* Involvement of T6 pili in biofilm formation by serotype M6 *Streptococcus pyogenes*. *J. Bacteriol.* **194**, 804–812 (2012).
213. Courtney, H. S. *et al.* Relationship between expression of the family of M proteins and lipoteichoic acid to hydrophobicity and biofilm formation in *Streptococcus pyogenes*. *PLoS One* **4**, 4166–4166 (2009).
214. Akiyama, H., Morizane, S., Yamasaki, O., Oono, T. & Iwatsuki, K. Assessment of *Streptococcus pyogenes* microcolony formation in infected skin by confocal laser scanning microscopy. *J Dermatol Sci* **32**, 193–199. (2003).
215. Caparon, M. G., Stephens, D. S., Olsén, A. & Scott, J. R. Role of M protein in adherence of Group A Streptococci. *Infect. Immun.* **59**, 1811–1817 (1991).
216. Norrby-Teglund, A., Lustig, R. & Kotb, M. Differential induction of Th1 versus Th2 cytokines by Group A Streptococcal toxic shock syndrome isolates. *Infect. Immun.* **65**, 5209–5215 (1997).
217. Norrby-Teglund, A., Norgren, M., Holm, S. E., Andersson, U. & Andersson, J. Similar cytokine induction profiles of a novel streptococcal exotoxin, MF, and pyrogenic exotoxins A and B. *Infect. Immun.* **62**, 3731–3738 (1994).
218. Baker, M., Gutman, D. M., Papageorgiou, A. C., Collins, C. M. & Acharya, K. R. Structural features of a zinc binding site in the superantigen streptococcal pyrogenic exotoxin A (SpeA1): Implications for MHC class II recognition. *Protein Sci.* **10**, 1268–1273 (2001).

219. Fast, D., Schlievert, P. & Nelson, R. Toxic shock syndrome-associated Staphylococcal and Streptococcal pyrogenic toxins are potent inducers of tumor necrosis factor production. *Infect. Immun.* **57**, 291–294 (1989).
220. Smoot, L. *et al.* Characterization of Two Novel Pyrogenic Toxin Superantigens Made by an Acute Rheumatic Fever Clone of *Streptococcus pyogenes* Associated with Multiple Disease Outbreaks. *Infect. Immun.* **70**, 7095–7104 (2002).
221. Unnikrishnan, M. *et al.* The Bacterial Superantigen Streptococcal Mitogenic Exotoxin Z Is the Major Immunoactive Agent of *Streptococcus pyogenes*. *J. Immunol.* **169**, 2561–2569 (2002).
222. You, Y. *et al.* Scarlet Fever Epidemic in China Caused by *Streptococcus pyogenes* Serotype M12: Epidemiologic and Molecular Analysis. *EBioMedicine* **28**, 128–135 (2018).
223. Lynskey, N. N. *et al.* Emergence of dominant toxigenic M1T1 *Streptococcus pyogenes* clone during increased scarlet fever activity in England: a population-based molecular epidemiological study. *Lancet Infect. Dis.* **19**, 1209–1218 (2019).
224. Rümke, L. W. *et al.* Dominance of M1UK clade among Dutch M1 *Streptococcus pyogenes*. *Lancet Infect. Dis.* **20**, 539–540 (2020).
225. Marrack, P. & Kappler, J. The staphylococcal enterotoxins and their relatives. *Science (80-. )*. **248**, 705–711 (1990).
226. Sundberg, E. J., Deng, L. & Mariuzza, R. A. TCR recognition of peptide/MHC class II complexes and superantigens. *Semin. Immunol.* **19**, 262–271 (2007).
227. Herrmann, T., Baschieri, S., Lees, R. K. & MacDonald, H. R. *In vivo* responses of CD4+ and CD8+ cells to bacterial superantigens. *Eur. J. Immunol.* **22**, 1935–1938 (1992).
228. Shiseki, M. *et al.* Comparison of pathogenic factors expressed by Group A *Streptococci* isolated from patients with streptococcal toxic shock syndrome and scarlet fever. *Microb. Pathog.* **27**, 243–252 (1999).
229. Hiesch, J. G., Bernheimer, A. W. & Weissmann, G. Motion Picture Study Of The Toxic Action Of Streptolysins On Leucocytes. *J. Exp. Med.* **118**, 223–228 (1963).
230. Launay, J. M. & Alouf, J. E. Biochemical and ultrastructural study of the disruption of blood platelets by streptolysin O. *Biochim. Biophys. Acta* **556**, 278–291 (1979).
231. Alouf, J. E. Streptococcal toxins (streptolysin O, streptolysin S, erythrogenic toxin). *Pharmacol. Ther.* **11**, 661–717 (1980).
232. Ofek, I., Zafiriri, D., Goldhar, J. & Eisenstein, B. I. Inability of toxin inhibitors to neutralize enhanced toxicity caused by bacteria adherent to tissue culture cells. *Infect. Immun.* **58**, 3737–3742 (1990).
233. Betschel, S. D., Borgia, S. M., Barg, N. L., Low, D. E. & De Azavedo, J. C. Reduced virulence of Group A Streptococcal Tn916 mutants that do not produce streptolysin S. *Infect. Immun.* **66**, 1671–1679 (1998).
234. Siggins, M. K. *et al.* Extracellular bacterial lymphatic metastasis drives *Streptococcus pyogenes* systemic infection. *Nat. Commun.* **11**, 1–12 (2020).

235. Lynskey, N. N. *et al.* Rapid Lymphatic Dissemination of Encapsulated Group A *Streptococci* via Lymphatic Vessel Endothelial Receptor-1 Interaction. *PLoS Pathog.* **11**, e1005137 (2015).
236. LaPenta, D., Rubens, C., Chi, E. & Cleary, P. P. Group A *Streptococci* efficiently invade human respiratory epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12115–12119 (1994).
237. Cue, D., Dombek, P. E., Lam, H. & Cleary, P. P. *Streptococcus pyogenes* serotype M1 encodes multiple pathways for entry into human epithelial cells. *Infect. Immun.* **66**, 4593–4601 (1998).
238. Hagman, M. M., Dale, J. B. & Stevens, D. L. Comparison of adherence to and penetration of a human laryngeal epithelial cell line by Group A *Streptococci* of various M protein types 1. *FEMS Immunol. Med. Microbiol.* **23**, 195–204 (1999).
239. Brouwer, S., Barnett, T. C., Rivera-Hernandez, T., Rohde, M. & Walker, M. J. *Streptococcus pyogenes* adhesion and colonization. *FEBS Lett.* **590**, 3739–3757 (2016).
240. Ozeri, V., Rosenshine, I., Mosher, D. F., Fässler, R. & Hanski, E. Roles of integrins and fibronectin in the entry of *Streptococcus pyogenes* into cells via protein F1. *Mol. Microbiol.* **30**, 625–637 (1998).
241. Cywes, C. & Wessels, M. R. Group A *Streptococcus* tissue invasion by CD44-mediated cell signalling. *Nature* **414**, 648–652 (2001).
242. Schragar, H. M., Albertí, S., Cywes, C., Dougherty, G. J. & Wessels, M. R. Hyaluronic acid capsule modulates M protein-mediated adherence and acts as a ligand for attachment of Group A *Streptococcus* to CD44 on human keratinocytes. *J. Clin. Invest.* **101**, 1708–1716 (1998).
243. Cywes, C., Stamenkovic, I. & Wessels, M. R. CD44 as a receptor for colonization of the pharynx by Group A *Streptococcus*. *J. Clin. Invest.* **106**, 995–1002 (2000).
244. Fluckiger, U. & Fischetti, V. A. Immunoglobulins inhibit adherence and internalization of *Streptococcus pyogenes* to human pharyngeal cells. *Adv Exp Med Biol* **418**, 909–911 (1997).
245. Molinari, G., Talay, S. R., Valentin-Weigand, P., Rohde, M. & Chhatwal, G. S. The fibronectin-binding protein of *Streptococcus pyogenes*, SfbI, is involved in the internalization of Group A *Streptococci* by epithelial cells. *Infect. Immun.* **65**, 1357–1363 (1997).
246. Terao, Y. *et al.* Fba, a novel fibronectin-binding protein from *Streptococcus pyogenes*, promotes bacterial entry into epithelial cells, and the *fba* gene is positively transcribed under the Mga regulator. *Mol. Microbiol.* **42**, 75–86 (2001).
247. Sumitomo, T. *et al.* Streptolysin S contributes to Group A Streptococcal translocation across an epithelial barrier. *J. Biol. Chem.* **286**, 2750–2761 (2011).
248. Bastiat-Sempe, B., Love, J. F., Lomayesva, N. & Wessels, M. R. Streptolysin O and NAD-Glycohydrolase Prevent Phagolysosome Acidification and Promote Group A *Streptococcus* Survival in Macrophages. *MBio* **5**, 1614–1690 (2014).

249. Molinari, G. & Chhatwal, G. S. Invasion and Survival of *Streptococcus pyogenes* in Eukaryotic Cells Correlates with the Source of the Clinical Isolates. *J. Infect. Dis.* **177**, 1600–1607 (1998).
250. O'Neill, A. M., Thurston, T. L. M., Holden, D. W., Nizet, V. & Gilmore, M. S. Cytosolic Replication of Group A *Streptococcus* in Human Macrophages. *MBio* **7**, 16–20 (2021).
251. Liu, Z., Treviño, J., Ramirez-Peña, E. & Sumbly, P. The small regulatory RNA FasX controls pilus expression and adherence in the human bacterial pathogen Group A *Streptococcus*. *Mol. Microbiol.* **86**, 140–154 (2012).
252. Patenge, N., Fiedler, T. & Kreikemeyer, B. Common regulators of virulence in Streptococci. *Curr. Top. Microbiol. Immunol.* **368**, 111–153 (2013).
253. Federle, M., McIver, K. & Scott, J. A Response Regulator That Represses Transcription of Several Virulence Operons in the Group A *Streptococcus*. *J. Bacteriol.* **181**, 3649–3657 (1999).
254. Graham, M. R. *et al.* Virulence control in Group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and *in vivo* infection modeling. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 13855–13860 (2002).
255. Federle, M. J. & Scott, J. R. Identification of binding sites for the Group A Streptococcal global regulator CovR. *Mol. Microbiol.* **43**, 1161–1172 (2002).
256. Sumbly, P., Whitney, A. R., Graviss, E. A., DeLeo, F. R. & Musser, J. M. Genome-wide analysis of group A streptococci reveals a mutation that modulates global phenotype and disease specificity. *PLoS Pathog.* **2**, 0041–0049 (2006).
257. Kansal, R. G. *et al.* Dissection of the molecular basis for hypervirulence of an *in vivo*-selected phenotype of the widely disseminated M1T1 strain of Group A *Streptococcus* bacteria. *J. Infect. Dis.* **201**, 855–865 (2010).
258. Ribardo, D. A. & McIver, K. S. Defining the Mga regulon: Comparative transcriptome analysis reveals both direct and indirect regulation by Mga in the Group A *Streptococcus*. *Mol. Microbiol.* **62**, 491–508 (2006).
259. Hondorp, E. R. & McIver, K. S. The Mga virulence regulon: infection where the grass is greener. *Mol. Microbiol.* **66**, 1056–1065 (2007).
260. McIver, K. S. & Myles, R. L. Two DNA-binding domains of Mga are required for virulence gene activation in the Group A *Streptococcus*. *Mol. Microbiol.* **43**, 1591–1601 (2002).
261. Okada, N., Geist, R. T. & Caparon, M. G. Positive transcriptional control of *mry* regulates virulence in the Group A *Streptococcus*. *Mol. Microbiol.* **7**, 893–903 (1993).
262. Perez-Casal, J., Caparon, M. G. & Scott, J. R. Mry, a trans-acting positive regulator of the M protein gene of *Streptococcus pyogenes* with similarity to the receptor proteins of two-component regulatory systems. *J. Bacteriol.* **173**, 2617–2624 (1991).
263. Fogg, G. C., Gibson, C. M. & Caparon, M. G. The identification of *rofA*, a positive-acting regulatory component of *prtF* expression: use of an  $\text{my}\delta$ -based shuttle mutagenesis strategy in *Streptococcus pyogenes*. *Mol. Microbiol.* **11**, 671–684 (1994).

264. VanHeyningen, T., Fogg, G., Yates, D., Hanski, E. & Caparon, M. Adherence and fibronectin binding are environmentally regulated in the Group A Streptococci. *Mol. Microbiol.* **9**, 1213–1222 (1993).
265. Chaussee, M. S. *et al.* Rgg Influences the Expression of Multiple Regulatory Loci To Coregulate Virulence Factor Expression in *Streptococcus pyogenes*. *Infect. Immun.* **70**, 762–770 (2002).
266. Kreikemeyer, B., Boyle, M. D. P., Buttaro, B. A. (Leonard), Heinemann, M. & Podbielski, A. 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. *Mol. Microbiol.* **39**, 392–406 (2001).
267. Kaplan, M. H. & Frengley, J. D. Autoimmunity to the heart in cardiac disease. Current concepts of the relation of autoimmunity to rheumatic fever, postcardiotomy and postinfarction syndromes and cardiomyopathies. *Am. J. Cardiol.* **24**, 459–473 (1969).
268. Zabriskie, J. B., Hsu, K. C. & Seegal, B. C. Heart-reactive antibody associated with rheumatic fever: characterization and diagnostic significance. *Clin. Exp. Immunol.* **7**, 147–159 (1970).
269. Kaplan, M. Immunologic Relation Of Streptococcal And Tissue Antigens. I. Properties Of An Antigen In Certain Strains Of Group A *Streptococci* Exhibiting An Immunologic Cross-Reaction With Human Heart Tissue. *J. Immunol.* **90**, 595–606 (1963).
270. Kaplan, M. H., Bolande, R., Rakita, L. & Blair, J. Presence Of Bound Immunoglobulins And Complement In The Myocardium In Acute Rheumatic Fever. Association With Cardiac Failure. *N. Engl. J. Med.* **271**, 637–645 (1964).
271. Kaplan, M. H. & Suchy, M. L. Immunologic Relation Of Streptococcal And Tissue Antigens. Cross-Reaction Of Antisera To Mammalian Heart Tissue With A Cell Wall Constituent Of Certain Strains Of Group A Streptococci. *J. Exp. Med.* **119**, 643–650 (1964).
272. Zabriskie, J. B. Mimetic relationships between Group A *Streptococci* and mammalian tissues. *Adv. Immunol.* **7**, 147–188 (1967).
273. Cunningham, M. W., Krisher, K. & Graves, D. C. Murine monoclonal antibodies reactive with human heart and Group A Streptococcal membrane antigens. *Infect. Immun.* **46**, 34–41 (1984).
274. Cunningham, M. W. & Russell, S. M. Study of heart-reactive antibody in antisera and hybridoma culture fluids against Group A Streptococci. *Infect. Immun.* **42**, 531–538 (1983).
275. Cunningham, M. W. & Swerlick, R. A. Polyspecificity of antistreptococcal murine monoclonal antibodies and their implications in autoimmunity. *J. Exp. Med.* **164**, 998–1012 (1986).
276. Cunningham, M. W., Hall, N. K., Krisher, K. K. & Spanier, A. M. A study of anti-Group A Streptococcal monoclonal antibodies cross-reactive with myosin. *J. Immunol.* **136**, 293–298 (1986).
277. Krisher, K. & Cunningham, M. W. Myosin: a link between *Streptococci* and heart. *Science* **227**, 413–415 (1985).

278. Cunningham, M. W. *et al.* Human monoclonal antibodies reactive with antigens of the Group A *Streptococcus* and human heart. *J. Immunol.* **141**, 2760–2766 (1988).
279. Dale, J. B. & Beachey, E. H. Epitopes of Streptococcal M proteins shared with cardiac myosin. *J. Exp. Med.* **162**, 583–591 (1985).
280. Dale, J. B. & Beachey, E. H. Human cytotoxic T lymphocytes evoked by Group A Streptococcal M proteins. *J. Exp. Med.* **166**, 1825–1835 (1987).
281. Dale, J. B. & Beachey, E. H. Multiple, heart-cross-reactive epitopes of Streptococcal M proteins. *J. Exp. Med.* **161**, 113–122 (1985).
282. Barnett, L. A. & Cunningham, M. W. A new heart-cross-reactive antigen in *Streptococcus pyogenes* is not M protein. *J. Infect. Dis.* **162**, 875–882 (1990).
283. Galvin, J. E., Hemric, M. E., Ward, K. & Cunningham, M. W. Cytotoxic mAb from rheumatic carditis recognizes heart valves and laminin. *J. Clin. Invest.* **106**, 217–224 (2000).
284. Delunardo, F. *et al.* Streptococcal–vimentin cross-reactive antibodies induce microvascular cardiac endothelial proinflammatory phenotype in rheumatic heart disease. *Clin. Exp. Immunol.* **173**, 419–429 (2013).
285. Cox, C. J. *et al.* Brain Human Monoclonal Autoantibody from Sydenham Chorea Targets Dopaminergic Neurons in Transgenic Mice and Signals Dopamine D2 Receptor: Implications in Human Disease. *J. Immunol.* **191**, 5524–5541 (2013).
286. Kirvan, C. A., Cox, C. J., Swedo, S. E. & Cunningham, M. W. Tubulin Is a Neuronal Target of Autoantibodies in Sydenham’s Chorea. *J. Immunol.* **178**, 7412–7421 (2007).
287. Cunningham, M. W., Antone, S. M., Gulizia, J. M., McManus, B. A. & Gauntt, C. J.  $\alpha$ -Helical Coiled-Coil Molecules: A Role in Autoimmunity against the Heart. *Clin. Immunol. Immunopathol.* **68**, 118–123 (1993).
288. Manjula, B. N. & Fischetti, V. A. Sequence homology of Group A Streptococcal Pep M5 protein with other coiled-coil proteins. *Biochem. Biophys. Res. Commun.* **140**, 684–690 (1986).
289. Markowitz, A. S. & Lange, C. F. J. Streptococcal Related Glomerulonephritis. Isolation, Immunochemistry And Comparative Chemistry Of Soluble Fractions From Type 12 Nephritogenic *Streptococci* And Human Glomeruli. *J. Immunol.* **92**, 565–575 (1964).
290. Cunningham, M. W., Antone, S. M., Smart, M., Liu, R. & Kosanke, S. Molecular analysis of human cardiac myosin-cross-reactive B- and T-cell epitopes of the Group A Streptococcal M5 protein. *Infect. Immun.* **65**, 3913–3923 (1997).
291. Delvig, A. A. & Robinson, J. H. Two T Cell Epitopes from the M5 Protein of Viable *Streptococcus pyogenes* Engage Different Pathways of Antigen Processing in Mouse Macrophages. *J. Immunol.* **160**, 5267–5272 (1998).
292. Roberts, S. *et al.* Pathogenic Mechanisms in Rheumatic Carditis: Focus on Valvular Endothelium. *J. Infect. Dis.* **183**, 507–511 (2001).
293. Bisno, A. L. *Streptococcus pyogenes*. in *Principles and Practice of Infectious Diseases, Vol. 2* 1786–1799 (1995).

294. Robinson, J. H. *et al.* Mapping T-cell epitopes in Group A Streptococcal type 5 M protein. *Infect. Immun.* **59**, 4324–4331 (1991).
295. Pruksakorn, S. *et al.* Identification of T cell autoepitopes that cross-react with the C-terminal segment of the M protein of Group A Streptococci. *Int. Immunol.* **6**, 1235–1244 (1994).
296. Gorton, D., Govan, B., Olive, C. & Ketheesan, N. B- and T-Cell Responses in Group A *Streptococcus* M-Protein- or Peptide-Induced Experimental Carditis. *Infect. Immun.* **77**, 2177–2183 (2009).
297. Adderson, E. E., Shikhman, A. R., Ward, K. E. & Cunningham, M. W. Molecular analysis of polyreactive monoclonal antibodies from rheumatic carditis: human anti-N-acetylglucosamine / anti-myosin antibody V region genes. *J. Immunol.* **161**, 2020–2031 (1998).
298. Shikhman, A. R., Cunningham, M. W., Hildebrand, W., Fischetti, V. A. & Cunningham, M. W. Immunological mimicry between N-acetyl-beta-D-glucosamine and cytokeratin peptides. Evidence for a microbially driven anti-keratin antibody response. *J. Immunol.* **152**, 4375–4387 (1994).
299. Dudding, B. A. & Ayoub, E. M. Persistence of Streptococcal Group A antibody in patients with rheumatic valvular disease. *J. Exp. Med.* **128**, 1081–1098 (1968).
300. Sikder, S., Rush, C., Govan, B., Alim, M. A. & Ketheesan, N. Anti-Streptococcal antibody and T-cell interactions with vascular endothelial cells initiate the development of rheumatic carditis. *J. Leukoc. Biol.* **107**, (2019).
301. Faé, K. C. *et al.* CXCL9/Mig Mediates T cells Recruitment to Valvular Tissue Lesions of Chronic Rheumatic Heart Disease Patients. *Inflammation* **36**, 800–811 (2013).
302. Toor, D. & Sharma, N. T cell subsets: an integral component in pathogenesis of rheumatic heart disease. *Immunol. Res.* **66**, 18–30 (2018).
303. Toor, D. & Vohra, H. Immune responsiveness during disease progression from acute rheumatic fever to chronic rheumatic heart disease. *Microbes Infect.* **14**, 1111–1117 (2012).
304. Bas, H. D. *et al.* A Shift in the Balance of Regulatory T and T Helper 17 Cells in Rheumatic Heart Disease. *J. Investig. Med.* **62**, 78 LP – 83 (2014).
305. Barnett, L. A. & Cunningham, M. W. Evidence for actin-like proteins in an M protein-negative strain of *Streptococcus pyogenes*. *Infect. Immun.* **60**, 3932–3936 (1992).
306. Kil, K. S., Cunningham, M. W. & Barnett, L. A. Cloning and sequence analysis of a gene encoding a 67-kilodalton myosin-cross-reactive antigen of *Streptococcus pyogenes* reveals its similarity with class II major histocompatibility antigens. *Infect. Immun.* **62**, 2440–2449 (1994).
307. Cunningham, M. W. *et al.* Cytotoxic and viral neutralizing antibodies crossreact with Streptococcal M protein, enteroviruses, and human cardiac myosin. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 1320–1324 (1992).
308. Quinn, A., Ward, K., Fischetti, V. A., Hemric, M. & Cunningham, M. W. Immunological relationship between the class I epitope of Streptococcal M protein and myosin. *Infect. Immun.* **66**, 4418–4424 (1998).

309. Quinn, A., Adderson, E. E., Shackelford, P. G., Carroll, W. L. & Cunningham, M. W. Autoantibody germ-line gene segment encodes VH and VL regions of a human anti-Streptococcal monoclonal antibody recognizing Streptococcal M protein and human cardiac myosin epitopes. *J. Immunol.* **154**, 4203–4212 (1995).
310. Guilherme, L. & Kalil, J. Rheumatic Heart Disease: Molecules Involved in Valve Tissue Inflammation Leading to the Autoimmune Process and Anti-S. *pyogenes* Vaccine. *Front. Immunol.* **4**, 1–6 (2013).
311. Diamantino Soares, A. C. *et al.* Circulating cytokines predict severity of rheumatic heart disease. *Int. J. Cardiol.* **289**, 107–109 (2019).
312. Rehman, S. *et al.* A study on the association of TNF- $\alpha$ , IL-6, IL-10 and IL-1 VNTR gene polymorphisms with rheumatic heart disease in Pakistani patients. *Cytokine* **61**, 527–531 (2013).
313. Lyang, K. M. *et al.* Dysregulated IL-1 $\beta$ -GM-CSF Axis in Acute Rheumatic Fever That Is Limited by Hydroxychloroquine. *Circulation* **138**, 2648–2661 (2018).
314. Guilherme, L. *et al.* Heart-directed Autoimmunity: the Case of Rheumatic Fever. *J. Autoimmun.* **16**, 363–367 (2001).
315. Guilherme, L. *et al.* Rheumatic Heart Disease: Proinflammatory Cytokines Play a Role in the Progression and Maintenance of Valvular Lesions. *Am. J. Pathol.* **165**, 1583–1591 (2004).
316. Faé, K. C. *et al.* How an autoimmune reaction triggered by molecular mimicry between Streptococcal M protein and cardiac tissue proteins leads to heart lesions in rheumatic heart disease. *J. Autoimmun.* **24**, 101–109 (2005).
317. Bryant, P. A., Robins-Browne, R., Carapetis, J. R. & Nigel, C. Some of the People, Some of the Time: susceptibility to acute rheumatic fever. *Circulation* **119**, 742–753 (2009).
318. French, K. A. & Poppas, A. Rheumatic Heart Disease in Pregnancy. *Circulation* **137**, 817–819 (2018).
319. Carapetis, J. R. *et al.* Acute rheumatic fever and rheumatic heart disease. *Nat Rev Dis Prim.* **2**, 1134–1137 (2018).
320. Bryant, P. A. *et al.* Susceptibility to Acute Rheumatic Fever Based on Differential Expression of Genes Involved in Cytotoxicity, Chemotaxis, and Apoptosis. *Infect. Immun.* **82**, 753–761 (2014).
321. Beltrame, M. H., Catarino, S. J., Goeldner, I., Boldt, A. B. W. & de Messias-Reason, I. J. The Lectin Pathway of Complement and Rheumatic Heart Disease. *Frontiers in Pediatrics* **2**, 148 (2015).
322. Guilherme, L. & Kalil, J. Rheumatic fever and rheumatic heart disease: cellular mechanisms leading autoimmune reactivity and disease. *J. Clin. Immunol.* **30**, 17–23 (2010).
323. Guilherme, L., Köhler, K. F. & Kalil, J. Rheumatic heart disease. Mediation by complex immune events. *Adv. Clin. Chem.* **53**, 31–50 (2011).
324. Muhamed, B., Parks, T. & Sliwa, K. Genetics of rheumatic fever and rheumatic heart disease. *Nat. Rev. Cardiol.* **17**, 145–154 (2020).
325. Berdeli, A., Celik, H. A., Özyürek, R., Dogrusoz, B. & Aydin, H. H. TLR-2



- gene Arg753Gln polymorphism is strongly associated with acute rheumatic fever in children. *J. Mol. Med.* **83**, 535–541 (2005).
326. Settin, A., Abdel-Hady, H., El-Baz, R. & Saber, I. Gene Polymorphisms of TNF- $\alpha$ , IL-10, IL-6, and IL-1 VNTR Related to Susceptibility and Severity of Rheumatic Heart Disease. *Pediatr. Cardiol.* **28**, 363–371 (2007).
  327. Abdallah, A. M. *et al.* IL10 Promoter Polymorphisms are Associated with Rheumatic Heart Disease in Saudi Arabian Patients. *Pediatr. Cardiol.* **37**, 99–105 (2016).
  328. Azevedo, P. M. *et al.* Association study involving polymorphisms in IL-6, IL-1RA, and CTLA4 genes and rheumatic heart disease in New Zealand population of Māori and Pacific ancestry. *Cytokine* **85**, 201–206 (2016).
  329. Gray, L.-A. *et al.* Genome-Wide Analysis of Genetic Risk Factors for Rheumatic Heart Disease in Aboriginal Australians Provides Support for Pathogenic Molecular Mimicry. *J. Infect. Dis.* **216**, 1460–1470 (2017).
  330. Guilherme, L., Weidebach, W., Kiss, M. H., Snitcowsky, R. & Kalil, J. Association of human leukocyte class II antigens with rheumatic fever or rheumatic heart disease in a Brazilian population. *Circulation* **83**, 1995–1998 (1991).
  331. Guilherme, L., Köhler, K., Postol, E. & Kalil, J. Genes, autoimmunity and pathogenesis of rheumatic heart disease. *Ann. Pediatr. Cardiol.* **4**, 13 (2011).
  332. Messias-Reason, I. J., Schafranski, M. D., Kremsner, P. G. & Kun, J. F. J. Ficolin 2 (FCN2) functional polymorphisms and the risk of rheumatic fever and rheumatic heart disease. *Clin. Exp. Immunol.* **157**, 395–399 (2009).
  333. Li, Y. *et al.* A normal polymorphism site of TLR2 3' untranslated region is related to rheumatic heart disease by up-regulating TLR2 expression. *Ann. Clin. Biochem.* **52**, 470–475 (2014).
  334. Brandt, E. R. *et al.* Opsonic human antibodies from an endemic population specific for a conserved epitope on the M protein of Group A *Streptococci*. *Immunology* **89**, 331–337 (1996).
  335. Lancefield, R. Persistence of type-specific antibodies in man following infection with Group A *Streptococci*. *J. Exo Med* **110**, 271–292 (1959).
  336. World Health Organisation. *Correlates of Vaccine-Induced Protection: Methods and Implications. WHO Technical Report Series* (2013).
  337. Lamagni, T. L. *et al.* Severe *Streptococcus pyogenes* infections, United Kingdom, 2003-2004. *Emerg. Infect. Dis.* **14**, 202–209 (2008).
  338. Kaul, R., McGeer, A., Low, D. E., Green, K. & Schwartz, B. Population-based surveillance for Group A *Streptococcal* necrotizing fasciitis: Clinical features, prognostic indicators, and microbiologic analysis of seventy-seven cases. Ontario Group A *Streptococcal* Study. *Am. J. Med.* **103**, 18–24 (1997).
  339. Pruksakorn, S. *et al.* Towards a vaccine for rheumatic fever: identification of a conserved target epitope on M protein of Group A *Streptococci*. *Lancet* **344**, 639–642 (1994).
  340. Lewnard, J. A., Whittles, L. K., Rick, A.-M. & Martin, J. M. Naturally Acquired Protection Against Upper Respiratory Symptoms Involving Group A *Streptococcus* in a Longitudinal Cohort Study. *Clin. Infect. Dis.* **71**, 244–254

(2020).

341. Campbell, P. T. *et al.* Longitudinal Analysis of Group A *Streptococcus emm* Types and *emm* Clusters in a High-Prevalence Setting: Relationship between Past and Future Infections. *J. Infect. Dis.* **221**, 1429–1437 (2020).
342. Smeesters, P. R., Mardulyn, P., Vergison, A., Leplae, R. & Van Melderen, L. Genetic diversity of Group A *Streptococcus* M protein: Implications for typing and vaccine development. *Vaccine* **26**, 5835–5842 (2008).
343. Lancefield, R. C. Current knowledge of type-specific M antigens of Group A Streptococci. *J. Immunol.* **89**, 307–313 (1962).
344. Pandey, M. *et al.* Streptococcal Immunity Is Constrained by Lack of Immunological Memory following a Single Episode of Pyoderma. *PLoS Pathog.* **12**, 1006–1022 (2016).
345. Tsoi, S. K., Smeesters, P. R., Frost, H. R. C., Licciardi, P. & Steer, A. C. Correlates of Protection for M Protein-Based Vaccines against Group A *Streptococcus*. *J. Immunol. Res.* **2015**, 1–11 (2015).
346. Ozberk, V., Pandey, M. & Good, M. F. Contribution of cryptic epitopes in designing a Group A Streptococcal vaccine. *Hum. Vaccin. Immunother.* **14**, 2034–2052 (2018).
347. Hall, M. A. *et al.* Intranasal immunization with multivalent Group A Streptococcal vaccines protects mice against intranasal challenge infections. *Infect. Immun.* **72**, 2507–2512 (2004).
348. Steer, A. C., Dale, J. B. & Carapetis, J. R. Progress toward a global Group A Streptococcal vaccine. *Pediatr. Infect. Dis. J.* **32**, 180–182 (2013).
349. Dooling, K., Shapiro, D., Beneden, C., Hersh, A. & Hicks, L. Overprescribing and inappropriate antibiotic selection for children with pharyngitis in the United States. *JAMA Pediatr* **168**, 1073 (2014).
350. Ganesan, K. & M., M. B. Is Primary Prevention of Rheumatic Fever the Missing Link in the Control of Rheumatic Heart Disease in Africa? *Circulation* **120**, 709–713 (2009).
351. Del Mar, C. B., Glasziou, P. P. & Spinks, A. B. Antibiotics for sore throat. *Cochrane database Syst. Rev.* **23**, 1–43 (2006).
352. Shulman, S. T. *et al.* Clinical Practice Guideline for the Diagnosis and Management of Group A Streptococcal Pharyngitis: 2012 Update by the Infectious Diseases Society of America. *Clin. Infect. Dis.* **55**, 86–102 (2012).
353. Hasenbein, M. E. *et al.* Detection of Multiple Macrolide- and Lincosamide-Resistant Strains of *Streptococcus pyogenes* from Patients in the Boston Area. *J. Clin. Microbiol.* **42**, 1559–1563 (2004).
354. Seppälä, H. *et al.* Resistance to erythromycin in Group A Streptococci. *N. Engl. J. Med.* **326**, 292–297 (1992).
355. Zimbelman, J., Palmer, A. & Todd, J. Improved outcome of clindamycin compared with beta-lactam antibiotic treatment for invasive *Streptococcus pyogenes* infection. *Pediatr. Infect. Dis. J.* **18**, 1096–1100 (1999).
356. Varaldo, P. E. *et al.* Nationwide survey in Italy of treatment of *Streptococcus pyogenes* pharyngitis in children: influence of macrolide resistance on clinical and microbiological outcomes. *Clin. Infect. Dis.* **29**, 869–873 (1999).

357. Vannice, K. S. *et al.* *Streptococcus pyogenes* *pbp2x* Mutation Confers Reduced Susceptibility to  $\beta$ -Lactam Antibiotics. *Infect. Dis. Soc. Am.* **71**, 201–204 (2020).
358. Chochua, S. *et al.* Invasive Group A Streptococcal Penicillin Binding Protein 2 $\times$  Variants Associated with Reduced Susceptibility to  $\beta$ -Lactam Antibiotics in the United States, 2015-2021. *Antimicrob. Agents Chemother.* **66**, e0080222 (2022).
359. Grebe, T. & Hakenbeck, R. Penicillin-binding proteins 2b and 2x of *Streptococcus pneumoniae* are primary resistance determinants for different classes of beta-lactam antibiotics. *Antimicrob. Agents Chemother.* **40**, 829–834 (1996).
360. Heggie, A. *et al.* Prevalence and Characteristics of Pharyngeal Group A  $\beta$ -Hemolytic *Streptococci* in US Navy Recruits Receiving Benzathine Penicillin Prophylaxis. *J. Infect. Dis.* **166**, 1006–1013 (1992).
361. Morita, J. Y. *et al.* Impact of azithromycin on oropharyngeal carriage of Group A *Streptococcus* and nasopharyngeal carriage of macrolide-resistant *Streptococcus pneumoniae*. *Pediatr. Infect. Dis. J.* **19**, 41–46 (2000).
362. Tanz, R. R., Shulman, S. T., Barthel, M. J., Willert, C. & Yogev, R. Penicillin plus rifampin eradicates pharyngeal carriage of group A streptococci. *J. Pediatr.* **106**, 876–880 (1985).
363. Ebell, M. H., Smith, M. A., Barry, H. C., Ives, K. & Carey, M. The rational clinical examination. Does this patient have strep throat? *JAMA* **284**, 2912–2918 (2000).
364. Barnett, M. L. & Linder, J. A. Antibiotic Prescribing to Adults with Sore Throat in the United HHS Public Access. *JAMA Intern Med* **174**, 138–140 (2014).
365. Reglinski, M., Lynskey, N. N., Choi, Y. J., Edwards, R. J. & Sriskandan, S. Development of a multicomponent vaccine for *Streptococcus pyogenes* based on the antigenic targets of IVIG. *J. Infect.* **72**, 450–459 (2016).
366. Bonanni, P. Demographic impact of vaccination. *Vaccine* **17**, 120–125 (1999).
367. Sheel, M., Moreland, N. J., Fraser, J. D. & Carapetis, J. Development of Group A Streptococcal vaccines: an unmet global health need. *Expert Rev. Vaccines* **15**, 227–238 (2016).
368. Gordis, L. The virtual disappearance of rheumatic fever in the United States: lessons in the rise and fall of disease. *Circulation* **72**, 1155–1162 (1985).
369. World Health Organisation. Immunization, vaccines, and biologicals. Available at: <http://www.who.int/immunization/research/development/en/>.
370. Young, D. Failure of type specific *Streptococcus pyogenes* vaccine to prevent respiratory infections. *U. S. Nav. Med. Bull.* **46**, 709–718 (1946).
371. Wolfe, C. K. J., Hayashi, J. A., Walsh, G. & Barkulis, S. S. Type-specific antibody response in man to injections of cell walls and M protein from Group A, type 14 *Streptococci*. *J. Lab. Clin. Med.* **61**, 459–468 (1963).
372. Potter, E. V., Stollerman, G. H. & Siegel, A. C. Recall of type specific antibodies in man by injections of Streptococcal cell walls. *J. Clin. Invest.* **41**, 301–310 (1962).

373. Massell, B. F., Michael, J. G., Amezcua, J. & Siner, M. Secondary and apparent primary antibody responses after Group A Streptococcal vaccination of 21 children. *Appl. Microbiol.* **16**, 509–518 (1968).
374. Massell, B. F., Honikman, L. H. & Amezcua, J. Rheumatic Fever Following Streptococcal Vaccination. *JAMA* **207**, 1151–1119 (1969).
375. Food and Drug Administration. *Status of specific products; Group A Streptococcus*. (2005).
376. Bisno, A. L., Rubin, F. A., Cleary, P. P. & Dale, J. B. Prospects for a Group A Streptococcal Vaccine: Rationale, Feasibility, and Obstacles - Report of a National Institute of Allergy and Infectious Diseases Workshop. *Clin. Infect. Dis.* **41**, 1150–1156 (2005).
377. Pastural, É. *et al.* Safety and immunogenicity of a 30-valent M protein-based Group A Streptococcal vaccine in healthy adult volunteers: A randomized, controlled phase I study. *Vaccine* **38**, 1384–1392 (2020).
378. Sekuloski, S. *et al.* Evaluation of safety and immunogenicity of a Group A *Streptococcus* vaccine candidate (MJ8VAX) in a randomized clinical trial. *PLoS One* **13**, 1–14 (2018).
379. Fox, E. N., Waldman, R. H., Wittner, M. K., Mauceri, A. A. & Dorfman, A. Protective Study with a Group A Streptococcal M Protein Vaccine. Infectivity Challenge Of Human Volunteers. *J. Clin. Invest.* **52**, 1885–1892 (1973).
380. Polly, S. M. *et al.* Protective Studies with a Group A Streptococcal M Protein Vaccine. Challenge of Volunteers after Local Immunization in the Upper Respiratory Tract. *J. Infect. Dis.* **131**, 217–224 (1975).
381. Beachey, E. H., Stollerman, G. H., Johnson, R. H., Ofek, I. & Bisno, A. L. Human immune response to immunization with a structurally defined polypeptide fragment of Streptococcal M protein. *J. Exp. Med.* **150**, 862–877 (1979).
382. Dale, J. B., Chiang, E. Y. & Lederer, J. W. Recombinant tetravalent Group A Streptococcal M protein vaccine. *J. Immunol.* **151**, 2188–2194 (1993).
383. Dale, J. B., Simmons, M., Chiang, E. C. & Chiang, E. Y. Recombinant, octavalent Group A Sstreptococcal M protein vaccine. *Vaccine* **14**, 944–948 (1996).
384. Kotloff, K. L. *et al.* Safety and Immunogenicity of a Recombinant Multivalent Group A Streptococcal Vaccine in Healthy Adults Phase 1 Trial. *Am. Meidical Assoc.* **292**, 709–715 (2004).
385. McNeil, S. A. *et al.* Safety and Immunogenicity of 26-Valent Group A *Streptococcus* Vaccine in Healthy Adult Volunteers. *Clin. Infect. Dis.* **41**, 1114–1122 (2005).
386. Hu, M. C. *et al.* Immunogenicity of a 26-Valent Group A Streptococcal Vaccine Immunogenicity of a 26-Valent Group A Streptococcal Vaccine. *Infect. Immun.* **70**, 2171–2177 (2002).
387. Dale, J. B., Penfound, T. A., Chiang, E. Y. & Walton, W. J. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of Group A *Streptococci*. *Vaccine* **29**, 8175–8178 (2011).
388. Wozniak, A. *et al.* Protective immunity induced by an intranasal multivalent

- vaccine comprising 10 *Lactococcus lactis* strains expressing highly prevalent M-protein antigens derived from Group A *Streptococcus*. *Microbiol. Immunol.* **62**, 395–404 (2018).
389. García, P. C. *et al.* Clinical and microbiological response of mice to intranasal inoculation with *Lactococcus lactis* expressing Group A *Streptococcus* antigens, to be used as an anti-Streptococcal vaccine. *Microbiol. Immunol.* **62**, 711–719 (2018).
  390. Guilherme, L. *et al.* Towards a vaccine against rheumatic fever. *Clin. Dev. Immunol.* **13**, 125–132 (2006).
  391. Postol, E. *et al.* Group A *Streptococcus* Adsorbed Vaccine: Repeated Intramuscular Dose Toxicity Test in Minipigs. *Sci. Rep.* **9**, 9733 (2019).
  392. Batzloff, M. R. *et al.* Protection against Group A *Streptococcus* by Immunization with J8–Diphtheria Toxoid: Contribution of J8- and Diphtheria Toxoid–Specific Antibodies to Protection. *J. Infect. Dis.* **187**, 1598–1608 (2003).
  393. Hayman, W. A. *et al.* Mapping the minimal murine T cell and B cell epitopes within a peptide vaccine candidate from the conserved region of the M protein of Group A *Streptococcus*. *Int. Immunol.* **9**, 1723–1733 (1997).
  394. Bi, S. *et al.* A multicomponent vaccine provides immunity against local and systemic infections by Group A *Streptococcus* across serotypes. *MBio* **10**, 1–12 (2019).
  395. Bensi, G. *et al.* Multi High-Throughput Approach for Highly Selective Identification of Vaccine Candidates: the Group A *Streptococcus* Case. *Mol. Cell. Proteomics* **11**, M111.015693 (2012).
  396. Rivera-Hernandez, T. *et al.* Differing Efficacies of Lead Group A Streptococcal Vaccine Candidates and Full-Length M Protein in Cutaneous and Invasive Disease Models. *MBio* **7**, 1–9 (2016).
  397. Rivera-Hernandez, T. *et al.* Vaccine-induced Th1-type response protects against invasive Group A *Streptococcus* infection in the absence of opsonizing antibodies. *MBio* **11**, 1–11 (2020).
  398. Reglinski, M. & Sriskandan, S. The contribution of Group A Streptococcal virulence determinants to the pathogenesis of sepsis. *Virulence* **5**, 127–136 (2014).
  399. Reglinski, M., Gierula, M., Lynskey, N. N., Edwards, R. J. & Sriskandan, S. Identification of the *Streptococcus pyogenes* surface antigens recognised by pooled human immunoglobulin. *Sci. Rep.* **5**, 1–9 (2015).
  400. Cleary, P. P., Matsuka, Y. V., Huynh, T., Lam, H. & Olmsted, S. B. Immunization with C5a peptidase from either Group A or B *Streptococci* enhances clearance of group a *Streptococci* from intranasally infected mice. *Vaccine* **22**, 4332–4341 (2004).
  401. Kawabata, S. *et al.* Systemic and mucosal immunizations with fibronectin-binding protein FBP54 induce protective immune responses against *Streptococcus pyogenes* challenge in mice. *Infect. Immun.* **69**, 924–930 (2001).
  402. Courtney, H. S., Hasty, D. L. & Dale, J. B. Serum opacity factor (SOF) of *Streptococcus pyogenes* evokes antibodies that opsonize homologous and

- heterologous SOF-positive serotypes of Group A *Streptococci*. *Infect. Immun.* **71**, 5097–5103 (2003).
403. Roggiani, M. *et al.* Toxoids of Streptococcal pyrogenic exotoxin A are protective in rabbit models of streptococcal toxic shock syndrome. *Infect. Immun.* **68**, 5011–5017 (2000).
  404. Kapur, V. *et al.* Vaccination with Streptococcal extracellular cysteine protease (interleukin-1B convertase) protects mice against challenge with heterologous group challenge with Group A *Streptococci*. *Microb. Pathog.* **16**, 443–450 (1994).
  405. McCormick, J. K. *et al.* Development of Streptococcal Pyrogenic Exotoxin C Vaccine Toxoids That Are Protective in the Rabbit Model of Toxic Shock Syndrome. *J. Immunol.* **165**, 2306–2312 (2000).
  406. Young, P. G., Proft, T., Harris, P. W. R., Brimble, M. A. & Baker, E. N. Structure and activity of *Streptococcus pyogenes* SipA: A signal peptidase-like protein essential for pilus polymerisation. *PLoS One* **9**, e99135 (2014).
  407. Fritzer, A. *et al.* Novel conserved Group A Streptococcal proteins identified by the antigenome technology as vaccine candidates for a non-M protein-based vaccine. *Infect. Immun.* **78**, 4051–4067 (2010).
  408. McMillan, D. J. *et al.* Identification and assessment of new vaccine candidates for Group A Streptococcal infections. *Vaccine* **22**, 2783–2790 (2004).
  409. Sabharwal, H. *et al.* Group A *Streptococcus* (GAS) carbohydrate as an immunogen for protection against GAS infection. *J. Infect.* **193**, 129–135 (2006).
  410. Malkiel, S., Liao, L., Cunningham, M. W. & Diamond, B. T-cell-dependent antibody response to the dominant epitope of Streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive with cardiac myosin. *Infect. Immun.* **68**, 5803–5808 (2000).
  411. Benedetto, R. Di *et al.* Rational design of a glycoconjugate vaccine against Group A *Streptococcus*. *Int. J. Mol. Sci.* **21**, 1–21 (2020).
  412. Gao, N. J. *et al.* Site-Specific Conjugation of Cell Wall Polyrihamnose to Protein SpyAD Envisioning a Safe Universal Group A Streptococcal Vaccine. *Infect. Microbes Dis.* **3**, 87–100 (2021).
  413. Wang, S. *et al.* Group A *Streptococcus* Cell Wall Oligosaccharide-Streptococcal C5a Peptidase Conjugates as Effective Antibacterial Vaccines. *ACS Infect. Dis.* **6**, 281–290 (2020).
  414. Pitirollo, O. *et al.* Gold nanoparticles morphology does not affect the multivalent presentation and antibody recognition of Group A *Streptococcus* synthetic oligorhamnans. *Bioorg. Chem.* **99**, 1–8 (2020).
  415. Micoli, F. *et al.* GMMA Is a Versatile Platform to Design Effective Multivalent Combination Vaccines. *Vaccines* **8**, 540 (2020).
  416. Khatun, F. *et al.* Immunogenicity Assessment of Cell Wall Carbohydrates of Group A *Streptococcus* via Self-Adjuvanted Glyco-lipopeptides. *ACS Infect. Dis.* **7**, 390–405 (2021).
  417. Dale, J. B. Multivalent Group A Streptococcal vaccine designed to optimize

- the immunogenicity of six tandem M protein fragments. *Vaccine* **17**, 193–200 (1999).
418. Croucher, N. J. *et al.* Population genomics of post-vaccine changes in Pneumococcal epidemiology. *Nat. Genet.* **45**, 656–663 (2013).
  419. Bart, M. J. *et al.* Global Population Structure and Evolution of *Bordetella pertussis* and Their Relationship with Vaccination. *MBio* **5**, 1–13 (2014).
  420. Giffard, P. M., Tong, S. Y. C., Holt, D. C., Ralph, A. P. & Currie, B. J. Concerns for efficacy of a 30-valent M-protein-based *Streptococcus pyogenes* vaccine in regions with high rates of rheumatic heart disease. *PLoS Negl. Trop. Dis.* **13**, 1–20 (2019).
  421. Good, M. F., Pandey, M., Batzloff, M. R. & Tyrrell, G. J. Strategic development of the conserved region of the M protein and other candidates as vaccines to prevent infection with Group A *Streptococci*. *Expert Rev. Vaccines* **14**, 1459–1470 (2015).
  422. Batzloff, M. R., Hartas, J., Zeng, W., Jackson, D. C. & Good, M. F. Intranasal vaccination with a lipopeptide containing a conformationally constrained conserved minimal peptide, a universal T cell epitope, and a self-adjuvanting lipid protects mice from Group A *Streptococcus* challenge and reduces throat colonization. *J. Infect. Dis.* **194**, 325–330 (2006).
  423. Jones, K. F. & Fischetti, V. A. The importance of the location of antibody binding on the M6 protein for opsonization and phagocytosis of Group A M6 *Streptococci*. *J. Exp. Med.* **167**, 1114–23 (1988).
  424. Sandin, C., Carlsson, F. & Lindahl, G. Binding of human plasma proteins to *Streptococcus pyogenes* M protein determines the location of opsonic and non-opsonic epitopes. *Mol. Microbiol.* **59**, 20–30 (2006).
  425. Pandey, M. *et al.* Combinatorial Synthetic Peptide Vaccine Strategy Protects against Hypervirulent CovR/S Mutant *Streptococci*. *J. Immunol.* **196**, 3364–3374 (2016).
  426. Pandey, M. *et al.* A Synthetic M Protein Peptide Synergizes with a CXC Chemokine Protease To Induce Vaccine-Mediated Protection against Virulent Streptococcal Pyoderma and Bacteremia. *J. Immunol.* **194**, 5915–5925 (2015).
  427. Batzloff, M. R. *et al.* Preclinical immunogenicity and safety of a Group A Streptococcal M protein-based vaccine candidate. *Hum. Vaccines Immunother.* **12**, 3089–3096 (2016).
  428. Good, M. F., Batzloff, M. R. & Pandey, M. Strategies in the development of vaccines to prevent infections with Group A *Streptococcus*. *Hum. Vaccines Immunother.* **9**, 2393–2397 (2013).
  429. Nordström, T. *et al.* Enhancing Vaccine Efficacy by Engineering a Complex Synthetic Peptide To Become a Super Immunogen. *J. Immunol.* **199**, 2794–2802 (2017).
  430. Schulze, K. *et al.* Bivalent mucosal peptide vaccines administered using the LCP carrier system stimulate protective immune responses against *Streptococcus pyogenes* infection. *Nanomedicine* **13**, 2463–2474 (2017).
  431. Pandey, M., Batzloff, M. R. & Good, M. F. Vaccination against rheumatic heart disease: A review of current research strategies and challenges. *Curr.*

*Infect. Dis. Rep.* **14**, 381–390 (2012).

432. Dale, J. B. *et al.* Group A streptococcal vaccines: Paving a path for accelerated development. *Vaccine* **31**, 216–222 (2013).
433. Shaw, H. A., Ozanne, J., Burns, K. & Mawas, F. Multicomponent Vaccines against Group A *Streptococcus* Can Effectively Target Broad Disease Presentations. *Vaccines* **9**, 1–16 (2021).
434. Rivera-Hernandez, T. *et al.* An Experimental Group A *Streptococcus* Vaccine That Reduces Pharyngitis and Tonsillitis in a Nonhuman Primate Model. *MBio* **10**, 1–10 (2019).
435. Shet, A., Kaplan, E., Johnson, D. & Cleary, P. P. Human immunogenicity studies on Group A Streptococcal C5a peptidase (SCPA) as a potential vaccine against Group A Streptococcal infections. *Indian J. Med. Res.* **119**, 95–98 (2004).
436. Henningham, A. *et al.* Conserved anchorless surface proteins as Group A Streptococcal vaccine candidates. *J. Mol. Med.* **90**, 1197–1207 (2012).
437. Henningham, A. *et al.* Structure-informed design of an enzymatically inactive vaccine component for Group A *Streptococcus*. *MBio* **4**, 509–513 (2013).
438. Dale, J. B. & Walker, M. J. Update on Group A Streptococcal vaccine development. *Curr. Opin. Infect. Dis.* **33**, 244–250 (2020).
439. Micoli, F., Costantino, P. & Adamo, R. Potential targets for next generation antimicrobial glycoconjugate vaccines. *FEMS Microbiol. Rev.* **42**, 388–423 (2018).
440. Pitner, J. B. *et al.* Bivalency and epitope specificity of a high-affinity IgG3 monoclonal antibody to the *Streptococcus* Group A carbohydrate antigen. Molecular modeling of a Fv fragment. *Carbohydr. Res.* **324**, 17–29 (2000).
441. Johnson, M. A. & Pinto, B. M. Saturation transfer difference 1D-TOCSY experiments to map the topography of oligosaccharides recognized by a monoclonal antibody directed against the cell-wall polysaccharide of Group A *Streptococcus*. *J. Am. Chem. Soc.* **124**, 15368–15374 (2002).
442. Reimer, K. B., Gidney, M. A. J., Bundle, D. R. & Mario Pinto, B. Immunochemical characterization of polyclonal and monoclonal *Streptococcus* Group A antibodies by chemically defined glycoconjugates and synthetic oligosaccharides. *Carbohydr. Res.* **232**, 131–142 (1992).
443. Michon, F. *et al.* Doubly branched hexasaccharide epitope on the cell wall polysaccharide of Group A *Streptococci* recognized by human and rabbit antisera. *Infect. Immun.* **73**, 6383–6389 (2005).
444. Weimar, T., Harris, S. L., Pitner, J. B., Bock, K. & Pinto, B. M. Transferred nuclear overhauser enhancement experiments show that the monoclonal antibody strep 9 selects a local minimum conformation of a *Streptococcus* Group A trisaccharide-hapten. *Biochemistry* **34**, 13672–13681 (1995).
445. Salvadori, L. G., Blake, M. S. & Mccarty, M. Group A *Streptococcus*-Liposome ELISA Antibody Titers to Group A Polysaccharide and Opsonophagocytic Capabilities of the Antibodies. *J. Immunol.* **171**, 593–600 (1995).
446. Auzanneau, F. I., Borrelli, S. & Pinto, B. M. Synthesis and immunological



- activity of an oligosaccharide-conjugate as a vaccine candidate against Group A *Streptococcus*. *Bioorganic Med. Chem. Lett.* **23**, 6038–6042 (2013).
447. Zhao, Y. *et al.* 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).
  448. Macleod, C. M., Hodges, R. G., Heidelberger, M. & Bernhard, W. G. Prevention Of Pneumococcal Pneumonia By Immunization With Specific Capsular Polysaccharides. *J. Exp. Med.* **82**, 445–465 (1945).
  449. Finland, M. & Sutliff, W. D. Specific Antibody Response Of Human Subjects To Intracutaneous Injection Of *Pneumococcus* Products. *J. Exp. Med.* **55**, 853–865 (1932).
  450. Ada, G. & Isaacs, D. Carbohydrate-protein conjugate vaccines. *Clin. Microbiol. Infect.* **9**, 79–85 (2003).
  451. Peltola, H., Käyhty, H., Sivonen, A. & Mäkelä, H. *Haemophilus influenzae* type b capsular polysaccharide vaccine in children: a double-blind field study of 100,000 vaccinees 3 months to 5 years of age in Finland. *Pediatrics* **60**, 730–737 (1977).
  452. Koskela, M., Leinonen, M., Häivä, V. M., Timonen, M. & Mäkelä, P. H. First and second dose antibody responses to Pneumococcal polysaccharide vaccine in infants. *Pediatr. Infect. Dis.* **5**, 45–50 (1986).
  453. Davies, J. The response of infants to inoculation with type I *Pneumococcus* carbohydrate. *J Immunol* **33**, 1–7 (1937).
  454. Barrett, D. J. Human immune responses to polysaccharide antigens: an analysis of bacterial polysaccharide vaccines in infants. *Adv. Pediatr.* **32**, 139–158 (1985).
  455. Weintraub, A. Immunology of bacterial polysaccharide antigens. *Carbohydr. Res.* **338**, 2539–47 (2003).
  456. Kelly, D. F., Pollard, A. J. & Moxon, E. R. Immunological Memory The Role of B Cells in Long-term Protection Against Invasive Bacterial Pathogens. *JAMA* **294**, 3019–3023 (2005).
  457. Pollard, A. J., Perrett, K. P. & Beverley, P. C. Maintaining protection against invasive bacteria with protein-polysaccharide conjugate vaccines. *Nat. Rev. Immunol.* **9**, 213–220 (2009).
  458. Costantino, P., Rappuoli, R. & Berti, F. The design of semi-synthetic and synthetic glycoconjugate vaccines. *Expert Opin. Drug Discov.* **6**, 1045–1066 (2011).
  459. Landsteiner, K. Experiments On Anaphylaxis To Azoproteins. *J. Exp. Med.* **39**, 631–637 (1924).
  460. Avery, O. T. & Goebel, W. F. Chemo-Immunological Studies on Conjugated Carbohydrate-Proteins: Immunological Specificity of Synthetic Sugar-Protein Antigens. *J. Exp. Med.* **50**, 533–550 (1929).
  461. Goebel, W. F. & Avery, O. T. Chemo-Immunological Studies On Conjugated Carbohydrate-Proteins: The Synthesis Of Thep-Aminobenzyl Ether Of The Soluble Specific Substance Of Type iii *Pneumococcus* And Its Coupling With Protein. *J. Exp. Med.* **54**, 431–436 (1931).

462. Schneerson, R., Barrera, O., Sutton, A. & Robbins, J. B. Preparation, characterization, and immunogenicity of *Haemophilus influenzae* type b polysaccharide-protein conjugates. *J. Exp. Med.* **152**, 361–76 (1980).
463. Eskola, J. *et al.* Efficacy of *Haemophilus influenzae* type b polysaccharide-diphtheria toxoid conjugate vaccine in infancy. *N. Engl. J. Med.* **317**, 717–722 (1987).
464. Anderson, P. W. *et al.* Vaccines consisting of periodate-cleaved oligosaccharides from the capsule of *Haemophilus influenzae* type b coupled to a protein carrier: structural and temporal requirements for priming in the human infant. *J. Immunol.* **137**, 1181–1186 (1986).
465. Costantino, P. *et al.* Development and phase 1 clinical testing of a conjugate vaccine against *Meningococcus A* and *C*. *Vaccine* **10**, 691–698 (1992).
466. De Gregorio, E. & Rappuoli, R. From empiricism to rational design: A personal perspective of the evolution of vaccine development. *Nat. Rev. Immunol.* **14**, 505–514 (2014).
467. Jones, C. Vaccines based on the cell surface carbohydrates of pathogenic bacteria. *An. Acad. Bras. Cienc.* **77**, 293–324 (2005).
468. Pace, D. Glycoconjugate vaccines. *Expert Opin.* **13**, 12–33 (2013).
469. Avci, F. Y., Li, X., Tsuji, M. & Kasper, D. L. A novel mechanism for glycoconjugate vaccine activation of the adaptive immune system. *Nat Med* **17**, 1602–1609 (2012).
470. Avci, F. Y., Li, X., Tsuji, M. & Kasper, D. L. A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nat. Med.* **17**, 1602–1609 (2011).
471. Bröker, M., Berti, F., Schneider, J. & Vojtek, I. Polysaccharide conjugate vaccine protein carriers as a “neglected valency” – Potential and limitations. *Vaccine* **35**, 3286–3294 (2017).
472. Micoli, F., Adamo, R. & Costantino, P. Protein carriers for glycoconjugate vaccines: History, selection criteria, characterization and new trends. *Molecules* **23**, 1–18 (2018).
473. Schutze, M. P., Leclerc, C., Jolivet, M., Audibert, F. & Chedid, L. Carrier-induced epitopic suppression, a major issue for future synthetic vaccines. *J. Immunol.* **135**, 2319–2322 (1985).
474. Schutze, M.-P., Leclerc, C., Vogel, F. R. & Chedid, L. Epitopic suppression in synthetic vaccine models: Analysis of the effector mechanisms. *Cell. Immunol.* **104**, 79–90 (1987).
475. Etlinger, H. M. *et al.* Use of prior vaccinations for the development of new vaccines. *Science* **249**, 423–425 (1990).
476. Falugi, F. *et al.* Rationally designed strings of promiscuous CD4+ T cell epitopes provide help to *Haemophilus influenzae* type b oligosaccharide: a model for new conjugate vaccines. *Eur. J. Immunol* **31**, 3816–3824 (2001).
477. Dykxhoorn, D. M., Pierre, R. S. & Linn, T. Promoter Expression Vectors. *Gene* **177**, 133–136 (1996).
478. Ihssen, J. *et al.* Production of glycoprotein vaccines in *Escherichia coli*. *Microb. Cell Fact.* **9**, 1–13 (2010).

479. Hall, C. Investigating novel vaccine candidates for *Clostridium difficile*. (2020).
480. Linton, D. *et al.* Functional analysis of the *Campylobacter jejuni* N-linked protein glycosylation pathway. *Mol. Microbiol.* **55**, 1695–1703 (2005).
481. Schulz, H., Hennecke, H. & Thöny-Meyer, L. Prototype of a heme chaperone essential for cytochrome c maturation. *Science* **281**, 1197–1200 (1998).
482. Mccarty, M. The lysis of Group A Hemolytic *Streptococci* by extra-cellular enzymes of *Streptomyces albus*. *J. Exp. Med.* **96**, 569–581 (1952).
483. Heymann, H., Manniello, J, M. & Barkulis, S, S. Structure of Streptococcal Cell Walls. Phosphate esters in the walls of Group A *Streptococcus pyogenes*. *Biochem. Biophys. Res. Commun.* **26**, 486–491 (1967).
484. Silhavy, T. J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb. Perspect. Biol.* **2**, 1–16 (2010).
485. Auzanneau, F.-I., Borrelli, S. & Pinto, B. M. Synthesis and immunological activity of an oligosaccharide-conjugate as a vaccine candidate against Group A *Streptococcus*. *Bioorg. Med. Chem. Lett.* **23**, 6038–6042 (2013).
486. Shikhman, A. R., Greenspan, N. S. & Cunningham, M. W. Cytokeratin peptide SFGSGFGGGY mimics N-acetyl-beta-D-glucosamine in reaction with antibodies and lectins, and induces *in vivo* anti-carbohydrate antibody response. *J. Immunol.* **153**, 5593–5606 (1994).
487. Dagan, R., Poolman, J. & Siegrist, C. A. Glycoconjugate vaccines and immune interference: A review. *Vaccine* **28**, 5513–5523 (2010).
488. Donnelly, J. J., Deck, R. R. & Liu, M. A. Immunogenicity of a *Haemophilus influenzae* polysaccharide-*Neisseria meningitidis* outer membrane protein complex conjugate vaccine. *J. Immunol.* **145**, 3071–3079 (1990).
489. Kilpi, T. *et al.* Protective efficacy of a second Pneumococcal conjugate vaccine against Pneumococcal acute otitis media in infants and children: randomized, controlled trial of a 7-valent Pneumococcal polysaccharide-Meningococcal outer membrane protein complex conjugate. *Clin. Infect. Dis.* **37**, 1155–1164 (2003).
490. Forsgren, A., Riesbeck, K. & Janson, H. Protein D of *Haemophilus influenzae*: a protective nontypeable *H. influenzae* antigen and a carrier for Pneumococcal conjugate vaccines. *Clin. Infect. Dis.* **46**, 726–731 (2008).
491. Prymula, R. *et al.* Pneumococcal capsular polysaccharides conjugated to protein D for prevention of acute otitis media caused by both *Streptococcus pneumoniae* and non-typable *Haemophilus influenzae*: a randomised double-blind efficacy study. *Lancet* **367**, 740–748 (2006).
492. Cohen, D. *et al.* Double-blind vaccine-controlled randomised efficacy trial of an investigational *Shigella sonnei* conjugate vaccine in young adults. *Lancet* **349**, 155–159 (1997).
493. Fattom, A. *et al.* Laboratory and clinical evaluation of conjugate vaccines composed of *Staphylococcus aureus* type 5 and type 8 capsular polysaccharides bound to *Pseudomonas aeruginosa* recombinant exoprotein A. *Infect. Immun.* **61**, 1023–1032 (1993).
494. Kossaczka, Z. *et al.* Safety and immunogenicity of Vi conjugate vaccines for

- typhoid fever in adults, teenagers, and 2- to 4-year-old children in Vietnam. *Infect. Immun.* **67**, 5806–5810 (1999).
495. Kuberan, B. & Linhardt, R. J. Carbohydrate Based Vaccines. *Curr. Org. Chem.* **4**, 653–677 (2000).
  496. Zarei, A. E., Almehdar, H. A. & Redwan, E. M. Hib Vaccines: Past, present, and future perspectives. *J. Immunol. Res.* **2016**, 1–18 (2016).
  497. Xu, P. *et al.* Conjugate Vaccines from Bacterial Antigens by Squaric Acid Chemistry: A Closer Look. *ChemBioChem* **18**, 799–815 (2017).
  498. Crotti, S. *et al.* Defined Conjugation of Glycans to the Lysines of CRM197 Guided by their Reactivity Mapping. *ChemBioChem* **15**, 836–843 (2014).
  499. Kapoor, N. *et al.* Malaria Derived Glycosylphosphatidylinositol Anchor Enhances Anti-Pfs25 Functional Antibodies That Block Malaria Transmission. *Biochemistry* **57**, 516–519 (2018).
  500. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Pyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
  501. Nelson, D., Schuch, R., Chahales, P., Zhu, S. & Fischetti, V. A. PlyC : A multimeric bacteriophage lysin. *PNAS* **103**, 1–6 (2006).
  502. Doytchinova, I. A. & Flower, D. R. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* **8**, 1–7 (2007).
  503. Zinkernagel, A. S. *et al.* The IL-8 Protease SpyCEP/ScpC of Group A *Streptococcus* Promotes Resistance to Neutrophil Killing. *Cell Host Microbe* **4**, 170–178 (2009).
  504. Kurupati, P. *et al.* Chemokine-cleaving *Streptococcus pyogenes* protease SpyCEP is necessary and sufficient for bacterial dissemination within soft tissues and the respiratory tract. *Mol. Microbiol.* **76**, 1387–1397 (2010).
  505. Zingaretti, C. *et al.* *Streptococcus pyogenes* SpyCEP: a chemokine-inactivating protease with unique structural and biochemical features. *FASEB J.* **24**, 2839–2848 (2010).
  506. Shelburne, S. A. *et al.* Maltodextrin utilization plays a key role in the ability of Group A *Streptococcus* to colonize the oropharynx. *Infect. Immun.* **74**, 4605–4614 (2006).
  507. Shelburne, S. A. *et al.* MalE of Group A *Streptococcus* participates in the rapid transport of maltotriose and longer maltodextrins. *J. Bacteriol.* **189**, 2610–2617 (2007).
  508. Gallotta, M. *et al.* SpyAD, a Moonlighting Protein of Group A *Streptococcus* Contributing to Bacterial Division and Host Cell Adhesion. *Infect. Immun.* **82**, 2890–2901 (2014).
  509. Mortensen, R. *et al.* Identifying protective *Streptococcus pyogenes* vaccine antigens recognized by both B and T cells in human adults and children. *Sci. Rep.* **6**, 1–11 (2016).
  510. Sumby, P. *et al.* A chemokine-degrading extracellular protease made by Group A *Streptococcus* alters pathogenesis by enhancing evasion of the innate immune response. *Infect. Immun.* **76**, 978–985 (2008).

511. Rodríguez-Ortega, M. J. *et al.* Characterization and identification of vaccine candidate proteins through analysis of the Group A *Streptococcus* surface proteome. *Nat. Biotechnol.* **24**, 191–197 (2006).
512. Grandi, G. Bacterial surface proteins and vaccines. *F1000 Biol. Rep.* **3**, 1–3 (2010).
513. Schultz, D. Pneumococcus polysaccharide conjugates for use as vaccine against tetanus and diphtheria. 11–27 (2003).
514. Bröker, M. Potential protective immunogenicity of tetanus toxoid, diphtheria toxoid and Cross Reacting Material 197 (CRM197) when used as carrier proteins in glycoconjugates. *Hum. Vaccin. Immunother.* **12**, 664–667 (2016).
515. Xing, D. K., McLellan, K., Corbel, M. J. & Sesardic, D. Estimation of antigenic tetanus toxoid extracted from biodegradable microspheres. *Biol. J. Int. Assoc. Biol. Stand.* **24**, 57–65 (1996).
516. Sheppard, A. J., Cussell, D. & Hughes, M. Production and characterization of monoclonal antibodies to tetanus toxin. *Infect. Immun.* **43**, 710–714 (1984).
517. Edwards, R. J. *et al.* Specific C-Terminal Cleavage and Inactivation of Interleukin-8 by Invasive Disease Isolates of *Streptococcus pyogenes*. *J. Infect. Dis.* **192**, 783–790 (2005).
518. Lei, B., Mackie, S., Lukomski, S. & Musser, J. M. Identification and immunogenicity of Group A *Streptococcus* culture supernatant proteins. *Infect. Immun.* **68**, 6807–6818 (2000).
519. Jobichen, C. *et al.* Structure of ScpC, a virulence protease from *Streptococcus pyogenes*, reveals the functional domains and maturation mechanism. *Biochem. J.* **475**, 2847–2860 (2018).
520. McKenna, S. *et al.* Structure, dynamics and immunogenicity of a catalytically inactive CXC chemokine-degrading protease SpyCEP from *Streptococcus pyogenes*. *Comput. Struct. Biotechnol. J.* **18**, 650–660 (2020).
521. Sutcliffe, I. C. & Harrington, D. J. Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. *Microbiology* **148**, 2065–2077 (2002).
522. Masuyer, G., Conrad, J. & Stenmark, P. The structure of the tetanus toxin reveals pH-mediated domain dynamics. *EMBO Rep.* **18**, 1306–1317 (2017).
523. McIntyre, C. & Reinin, G. Reduction in Endotoxin Levels After Performing the Prepare for Aseptic Sort Procedure on the BD FACSAria II Flow Cytometer. *BD Biosci. Appl. Note* 12 (2009).
524. Hermanson, G. T. Zero-Length Crosslinkers. in *Bioconjugate Techniques* 259–273 (2013).
525. Bergquist, C., Lagergard, T. & Holmgren, J. Antibody responses in serum and lung to intranasal immunization with *Haemophilus influenzae* type b polysaccharide conjugated to cholera toxin B subunit and tetanus toxoid. *Apmis* **106**, 800–806 (1998).
526. Kashef, N. *et al.* Synthesis and characterization of *Pseudomonas aeruginosa* alginate-tetanus toxoid conjugate. *J. Med. Microbiol.* **55**, 1441–1446 (2006).
527. Lockyer, K. *et al.* Higher mass Meningococcal group C-tetanus toxoid vaccines conjugated with carbodiimide correlate with greater immunogenicity.

- Vaccine* **38**, 2859–2869 (2020).
528. Shen, X. *et al.* Group B *Streptococcus* capsular polysaccharide-cholera toxin B subunit conjugate vaccines prepared by different methods for intranasal immunization. *Infect. Immun.* **69**, 297–306 (2001).
  529. Cayot, P. & Tainturier, G. The quantification of protein amino groups by the trinitrobenzenesulfonic acid method: A reexamination. *Anal. Biochem.* **249**, 184–200 (1997).
  530. Goodwin, J. F. & Choi, S. Quantification of Protein Solutions with Trinitrobenzenesulfonic Acid. *Clin. Chem.* **16**, 24–31 (1970).
  531. Snyder, S. L. & Sobocinski, P. Z. An improved 2,4,6-trinitrobenzenesulfonic acid method for the determination of amines. *Anal. Biochem.* **64**, 284–288 (1975).
  532. Zabriskie, J. B., Poon-King, T., Blake, M. S., Michon, F. & Yoshinaga, M. Phagocytic, serological, and protective properties of Streptococcal Group A carbohydrate antibodies. *Adv. Exp. Med. Biol.* **418**, 917–919 (1997).
  533. Hitri, K. *et al.* O-acetylation of typhoid capsular polysaccharide confers polysaccharide rigidity and immunodominance by masking additional epitopes. *Vaccine* **37**, 3866–3875 (2019).
  534. Schauer, U. *et al.* Levels of antibodies specific to tetanus toxoid, *Haemophilus influenzae* type b, and Pneumococcal capsular polysaccharide in healthy children and adults. *Clin. Diagn. Lab. Immunol.* **10**, 202–207 (2003).
  535. Sun, X., Stefanetti, G., Berti, F. & Kasper, D. L. Polysaccharide structure dictates mechanism of adaptive immune response to glycoconjugate vaccines. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 193–198 (2019).
  536. Kristiansen, K. A., Potthast, A. & Christensen, B. E. Periodate oxidation of polysaccharides for modification of chemical and physical properties. *Carbohydr. Res.* **345**, 1264–1271 (2010).
  537. Nakajima, N. & Ikada, Y. Mechanism of Amide Formation by Carbodiimide for Bioconjugation in Aqueous Media. *Bioconjug. Chem.* **6**, 123–130 (1995).
  538. Berti, F. & Adamo, R. Antimicrobial glycoconjugate vaccines: an overview of classic and modern approaches for protein modification. *Chem. Soc. Rev.* **47**, 9015–9025 (2018).
  539. Matos, M. J. *et al.* Chemo- and Regio- selective Lysine Modification on Native Proteins. *J. Am. Chem. Soc.* **140**, 4004–4017 (2018).
  540. Boquet, P., Duflot, E. & Hautecoeur, B. Low pH induces a hydrophobic domain in the tetanus toxin molecule. *Eur. J. Biochem.* **144**, 339–344 (1984).
  541. Stojićević, I. *et al.* Tetanus toxoid purification: chromatographic procedures as an alternative to ammonium-sulphate precipitation. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **879**, 2213–2219 (2011).
  542. Mant, C. T. & Hodges, R. S. High performance liquid chromatography of peptides and proteins. *Sep. Anal. Conform.* 2–4 (1991).
  543. Peeters, C. C. A. M., Tenbergen-Meekes, A. M. J., Poolman, J. T., Zegers, B. J. M. & Rijkers, G. T. Immunogenicity of a *Streptococcus pneumoniae* type 4 polysaccharide-protein conjugate vaccine is decreased by admixture of high

- doses of free saccharide. *Vaccine* **10**, 833–840 (1992).
544. Rodriguez, M. E. *et al.* Immunogenicity of *Streptococcus pneumoniae* type 6B and 14 polysaccharide-tetanus toxoid conjugates and the effect of uncoupled polysaccharide on the antigen-specific immune response. *Vaccine* **16**, 1941–1949 (1998).
  545. Steer, A. C. *et al.* Status of research and development of vaccines for *Streptococcus pyogenes*. *Vaccine* **34**, 2953–2958 (2016).
  546. Frasch, C. E. Preparation of bacterial polysaccharide-protein conjugates: Analytical and manufacturing challenges. *Vaccine* **27**, 6468–6470 (2009).
  547. Langdon, R. H., Cuccui, J. & Wren, B. W. N-linked glycosylation in bacteria: an unexpected application. *Futur. Microbiol* **4**, 401–412 (2009).
  548. Riddle, M. S. *et al.* Safety and Immunogenicity of a Candidate Bioconjugate Vaccine against *Shigella flexneri* 2a Administered to Healthy Adults: a Single-Blind, Randomized Phase I Study. *Clin. Vaccine Immunol.* **23**, 908–917 (2016).
  549. Huttner, A. *et al.* Safety, immunogenicity, and preliminary clinical efficacy of a vaccine against extraintestinal pathogenic *Escherichia coli* in women with a history of recurrent urinary tract infection: a randomised, single-blind, placebo-controlled phase 1b trial. *Lancet Infect. Dis.* **17**, 528–537 (2017).
  550. Hatz, C. F. R. *et al.* Safety and immunogenicity of a candidate bioconjugate vaccine against *Shigella dysenteriae* type 1 administered to healthy adults: A single blind, partially randomized Phase I study. *Vaccine* **33**, 4594–4601 (2015).
  551. Nothaft, H. & Szymanski, C. M. Protein glycosylation in bacteria: Sweeter than ever. *Nat. Rev. Microbiol.* **8**, 765–778 (2010).
  552. Iwashkiw, J. A., Vozza, N. F., Kinsella, R. L. & Feldman, M. F. Pour some sugar on it: the expanding world of bacterial protein O-linked glycosylation. *Mol. Microbiol.* **89**, 14–28 (2013).
  553. Szymanski, C. M., Ruijin, Y., Ewing, C. P., Trust, T. J. & Guerry, P. Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* **32**, 1022–1030 (1999).
  554. Parkhill, J. *et al.* The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* **403**, 665–668 (2000).
  555. Young, N. M. *et al.* Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. *J. Biol. Chem.* **277**, 42530–42539 (2002).
  556. Wacker, M. *et al.* Prevention of *Staphylococcus aureus* infections by glycoprotein vaccines synthesized in *Escherichia coli*. *J. Infect. Dis.* **209**, 1551–1561 (2014).
  557. Wacker, M. *et al.* N-Linked Glycosylation in *Campylobacter jejuni* and Its Functional Transfer into *E. coli*. *Science (80- )*. **298**, 1790–1793 (2002).
  558. Feldman, M. F. *et al.* Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **102**, 3016–3021 (2005).

559. Dow, J. M., Mauri, M., Scott, T. A. & Wren, B. W. Improving Protein Glycan Coupling Technology (PGCT) for glycoconjugate vaccine production. *Expert Rev. Vaccines* **19**, 507–527 (2020).
560. Talaat, K. R. *et al.* Human challenge study with a *Shigella* bioconjugate vaccine: Analyses of clinical efficacy and correlate of protection. *EBioMedicine* **66**, 1–10 (2021).
561. Frenck, R. W. J. *et al.* Safety and immunogenicity of a vaccine for extra-intestinal pathogenic *Escherichia coli* (ESTELLA): a phase 2 randomised controlled trial. *Lancet. Infect. Dis.* **19**, 631–640 (2019).
562. Ding, W., Nothaft, H., Szymanski, C. M. & Kelly, J. Identification and Quantification of Glycoproteins Using Ion-Pairing Normal-phase Liquid Chromatography and Mass Spectrometry. *Mol. Cell. Proteomics* **8**, 2170–2185 (2009).
563. Scott, N. E. *et al.* Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome. *Mol. Cell. Proteomics* **10**, 201–218 (2011).
564. Kowarik, M. *et al.* Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J.* **25**, 1957–1966 (2006).
565. Lizak, C., Gerber, S., Numao, S., Aebi, M. & Locher, K. P. X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474**, 350–356 (2011).
566. Shrimal, S., Cherepanova, N. A. & Gilmore, R. Cotranslational and posttranslational N-glycosylation of proteins in the endoplasmic reticulum. *Semin. Cell Dev. Biol.* **41**, 71–78 (2015).
567. Fisher, A. C. *et al.* Production of Secretory and Extracellular N-Linked Glycoproteins in *Escherichia coli*. *Appl. Environ. Microbiol.* **77**, 871–881 (2011).
568. Kowarik, M. *et al.* N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science (80-. )*. **314**, 1148–1150 (2006).
569. Wacker, M. *et al.* Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc. Natl. Acad. Sci.* **103**, 7088–7093 (2006).
570. Jaroentomeechai, T. *et al.* Single-pot glycoprotein biosynthesis using a cellfree transcription-translation system enriched with glycosylation machinery. *Nat. Commun.* **9**, 1–11 (2018).
571. Shibata, Y., Yamashita, Y., Ozaki, K., Nakano, Y. & Koga, T. Expression and characterization of Streptococcal *rgp* genes required for rhamnan synthesis in *Escherichia coli*. *Infect. Immun.* **70**, 2891–2898 (2002).
572. Yamashita, Y., Tsukioka, Y., Tomihisa, K., Nakano, Y. & Koga, T. Genes involved in cell wall localization and side chain formation of rhamnose-glucose polysaccharide in *Streptococcus mutans*. *J. Bacteriol.* **180**, 5803–5807 (1998).
573. Yamashita, Y. *et al.* A novel gene required for rhamnose-glucose polysaccharide synthesis in *Streptococcus mutans*. *J. Bacteriol.* **181**, 6556–6559 (1999).



574. Liu, D. & Reeves, P. R. *Escherichia coli* K12 regains its O antigen. *Microbiology* **140**, 49–57 (1994).
575. Ozaki, K. *et al.* A novel mechanism for glucose side-chain formation in rhamnose-glucose polysaccharide synthesis. *FEBS Lett.* **532**, 159–163 (2002).
576. Klena, J. D. & Schnaitman, C. A. Genes for TDP-rhamnose synthesis affect the pattern of lipopolysaccharide heterogeneity in *Escherichia coli* K-12. *J. Bacteriol.* **176**, 4003–4010 (1994).
577. Pradel, E., Parker, C. T. & Schnaitman, C. A. Structures of the *rfaB*, *rfaI*, *rfaJ*, and *rfaS* genes of *Escherichia coli* K-12 and their roles in assembly of the lipopolysaccharide core. *J. Bacteriol.* **174**, 4736–4745 (1992).
578. Kay, E. J. *et al.* Engineering a suite of *E. coli* strains for enhanced expression of bacterial polysaccharides and glycoconjugate vaccines. *Microb. Cell Fact.* **21**, 1–15 (2022).
579. Mauri, M. *et al.* Multivalent poultry vaccine development using Protein Glycan Coupling Technology. *Microb. Cell Fact.* **20**, 1–26 (2021).
580. Bachmann, B. J. Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriol. Rev.* **36**, 525–557 (1972).
581. Bauer, M. E. & Welch, R. A. Pleiotropic effects of a mutation in *rfaC* on *Escherichia coli* hemolysin. *Infect. Immun.* **65**, 2218–2224 (1997).
582. Garcia-Quintanilla, F., Iwashkiw, J. A., Price, N. L., Stratilo, C. & Feldman, M. F. Production of a recombinant vaccine candidate against *Burkholderia pseudomallei* exploiting the bacterial *N*-glycosylation machinery. *Front. Microbiol.* **5**, 1–10 (2014).
583. Schierle, C. F. *et al.* The DsbA Signal Sequence Directs Efficient, Cotranslational Export of Passenger Proteins to the *Escherichia coli* Periplasm via the Signal Recognition Particle Pathway. *J. Bacteriol.* **185**, 5706–5713 (2003).
584. Chen, M. M., Glover, K. J. & Imperiali, B. From peptide to protein: Comparative analysis of the substrate specificity of *N*-linked glycosylation in *C. jejuni*. *Biochemistry* **46**, 5579–5585 (2007).
585. Valderrama-Rincon, J. D. *et al.* An engineered eukaryotic protein glycosylation pathway in *Escherichia coli*. *Nat. Chem. Biol.* **8**, 434–436 (2012).
586. Glasscock, C. J. *et al.* A flow cytometric approach to engineering *Escherichia coli* for improved eukaryotic protein glycosylation. *Metab. Eng.* **47**, 488–495 (2018).
587. Iwashkiw, J. A. *et al.* Exploiting the *Campylobacter jejuni* protein glycosylation system for glycoengineering vaccines and diagnostic tools directed against brucellosis. *Microb. Cell Fact.* **11**, 1–11 (2012).
588. Kay, E., Cuccui, J. & Wren, B. W. Recent advances in the production of recombinant glycoconjugate vaccines. *npj Vaccines* **4**, 1–8 (2019).
589. Göhmann, S., Manning, P. A., Alpert, C.-A., Walker, M. J. & Timmis, K. N. Lipopolysaccharide O-antigen biosynthesis in *Shigella dysenteriae* serotype 1: analysis of the plasmid-carried *rfp* determinant. *Microb. Pathog.* **16**, 53–64

- (1994).
590. Klena, J. D. & Schnaitman, C. A. Function of the *rfb* gene cluster and the *rfe* gene in the synthesis of O antigen by *Shigella dysenteriae* 1. *Mol. Microbiol.* **9**, 393–402 (1993).
  591. Sturm, S., Jann, B., Jann, K., Fortnagel, P. & Timmis, K. N. Genetic and biochemical analysis of *Shigella dysenteriae* 1 O antigen polysaccharide biosynthesis in *Escherichia coli* K-12: structure and functions of the *rfb* gene cluster. *Microb. Pathog.* **1**, 307–324 (1986).
  592. Lukose, V., Walvoort, M. T. C. & Imperiali, B. Bacterial phosphoglycosyl transferases: Initiators of glycan biosynthesis at the membrane interface. *Glycobiology* **27**, 820–823 (2017).
  593. Mamat, U. *et al.* Detoxifying *Escherichia coli* for endotoxin-free production of recombinant proteins. *Microb. Cell Fact.* **14**, 1–15 (2015).
  594. Cognet, I. *et al.* Expression of recombinant proteins in a lipid A mutant of *Escherichia coli* BL21 with a strongly reduced capacity to induce dendritic cell activation and maturation. *J. Immunol. Methods* **272**, 199–210 (2003).
  595. Micoli, F. *et al.* A scalable method for O-antigen purification applied to various *Salmonella* serovars. *Anal. Biochem.* **434**, 136–145 (2013).
  596. Strutton, B. Engineering *Escherichia coli* to improve its N-linked glycosylation capabilities and the development of a new method of quantifying production. (2016).
  597. Reglinski, M. *et al.* A recombinant conjugated Pneumococcal vaccine that protects against murine infections with a similar efficacy to Prevnar-13. *npj Vaccines* **3**, 1–11 (2018).
  598. Cámara, M., Boulnois, G. J., Andrew, P. W. & Mitchell, T. J. A neuraminidase from *Streptococcus pneumoniae* has the features of a surface protein. *Infect. Immun.* **62**, 3688–3695 (1994).
  599. King, S. J., Whatmore, A. M. & Dowson, C. G. NanA, a neuraminidase from *Streptococcus pneumoniae*, shows high levels of sequence diversity, at least in part through recombination with *Streptococcus oralis*. *J. Bacteriol.* **187**, 5376–5386 (2005).
  600. Tong, H. H., Li, D., Chen, S., Long, J. P. & DeMaria, T. F. Immunization with recombinant *Streptococcus pneumoniae* neuraminidase NanA protects chinchillas against nasopharyngeal colonization. *Infect. Immun.* **73**, 7775–7778 (2005).
  601. Uchiyama, S. *et al.* The surface-anchored NanA protein promotes pneumococcal brain endothelial cell invasion. *J. Exp. Med.* **206**, 1845–1852 (2009).
  602. Cuccui, J. *et al.* Exploitation of bacterial N-linked glycosylation to develop a novel recombinant glycoconjugate vaccine against *Francisella tularensis*. *Open Biol.* **3**, 130002 (2013).
  603. Chen, L. *et al.* Outer membrane vesicles displaying engineered glycotopes elicit protective antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 3609–3618 (2016).
  604. Herbert, J. A. *et al.* Production and efficacy of a low-cost recombinant

- Pneumococcal protein polysaccharide conjugate vaccine. *Vaccine* **36**, 3809–3819 (2018).
605. Harding, C. M. *et al.* A platform for glycoengineering a polyvalent Pneumococcal bioconjugate vaccine using *E. coli* as a host. *Nat. Commun.* **10**, 1–11 (2019).
606. Kämpf, M. M. *et al.* In vivo production of a novel glycoconjugate vaccine against *Shigella flexneri* 2a in recombinant *Escherichia coli*: Identification of stimulating factors for in vivo glycosylation. *Microb. Cell Fact.* **14**, 1–12 (2015).
607. Kram, K. E. & Finkel, S. E. Rich Medium Composition Affects *Escherichia coli* Survival, Glycation, and Mutation Frequency during Long-Term Batch Culture. *Appl. Environ. Microbiol.* **81**, 4442–4450 (2015).
608. Jorgenson, M. A., Kannan, S., Laubacher, M. E. & Young, K. D. Dead-end intermediates in the enterobacterial common antigen pathway induce morphological defects in *Escherichia coli* by competing for undecaprenyl phosphate. *Mol. Microbiol.* **100**, 1–14 (2016).
609. Jorgenson, M. A. & Young, K. D. Interrupting Biosynthesis of O Antigen or the Lipopolysaccharide Core Produces Morphological Defects in *Escherichia coli* by Sequestering Undecaprenyl Phosphate. *J. Bacteriol.* **198**, 3070–3079 (2016).
610. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
611. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647 (2008).
612. de Kok, S. *et al.* Rapid and reliable DNA assembly via ligase cycling reaction. *ACS Synth. Biol.* **3**, 97–106 (2014).
613. Carty, S. M., Sreekumar, K. R. & Raetz, C. R. H. Effect of cold shock on lipid A biosynthesis in *Escherichia coli*. Induction At 12 degrees C of an acyltransferase specific for palmitoleoyl-acyl carrier protein. *J. Biol. Chem.* **274**, 9677–9685 (1999).
614. Clementz, T., Zhou, Z. & Raetz, C. R. Function of the *Escherichia coli* *msbB* gene, a multicopy suppressor of *htrB* knockouts, in the acylation of lipid A. Acylation by MsbB follows laurate incorporation by HtrB. *J. Biol. Chem.* **272**, 10353–10360 (1997).
615. Yates, L. E. *et al.* Glyco-recoded *Escherichia coli*: Recombineering-based genome editing of native polysaccharide biosynthesis gene clusters. *Metab. Eng.* **53**, 59–68 (2019).
616. Strutton, B., Jaffé, S. R. P., Pandhal, J. & Wright, P. C. Producing a glycosylating *Escherichia coli* cell factory: The placement of the bacterial oligosaccharyl transferase *pglB* onto the genome. *Biochem. Biophys. Res. Commun.* **495**, 686–692 (2018).
617. Ihssen, J. *et al.* Increased efficiency of *Campylobacter jejuni* N - oligosaccharyltransferase PglB by structure-guided engineering. *Open Biol.* **5**, (2015).
618. Ollis, A. A. *et al.* Substitute sweeteners: Diverse bacterial oligosaccharyltransferases with unique N-glycosylation site preferences. *Sci.*

- Rep.* **5**, 1–13 (2015).
619. Jervis, A. J. *et al.* Characterization of *N*-linked protein glycosylation in *Helicobacter pullorum*. *J. Bacteriol.* **192**, 5228–5236 (2010).
  620. Mills, D. C. *et al.* Functional analysis of *N*-linking oligosaccharyl transferase enzymes encoded by deep-sea vent proteobacteria. *Glycobiology* **26**, 398–409 (2016).
  621. Pan, C. *et al.* Biosynthesis of Conjugate Vaccines Using an *O*-Linked Glycosylation System. *MBio* **7**, 1–11 (2016).
  622. Sun, P. *et al.* Design and production of conjugate vaccines against *S. Paratyphi A* using an *O*-linked glycosylation system *in vivo*. *npj Vaccines* **3**, 1–9 (2018).
  623. Feldman, M. F. *et al.* A promising bioconjugate vaccine against hypervirulent *Klebsiella pneumoniae*. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 18655–18663 (2019).
  624. Grass, S., Lichti, C. F., Townsend, R. R., Gross, J. & St. Geme III, J. W. The *Haemophilus influenzae* HMW1C Protein Is a Glycosyltransferase That Transfers Hexose Residues to Asparagine Sites in the HMW1 Adhesin. *PLOS Pathog.* **6**, 1–9 (2010).
  625. Cuccui, J. *et al.* The *N*-linking glycosylation system from *Actinobacillus pleuropneumoniae* is required for adhesion and has potential use in glycoengineering. *Open Biol.* **7**, 1602–1612 (2017).
  626. Naegeli, A. *et al.* Molecular analysis of an alternative *N*-glycosylation machinery by functional transfer from *Actinobacillus pleuropneumoniae* to *Escherichia coli*. *J. Biol. Chem.* **289**, 2170–2179 (2014).
  627. Choi, K.-J., Grass, S., Paek, S., St Geme, J. W. 3rd & Yeo, H.-J. The *Actinobacillus pleuropneumoniae* HMW1C-like glycosyltransferase mediates *N*-linked glycosylation of the *Haemophilus influenzae* HMW1 adhesin. *PLoS One* **5**, e15888 (2010).
  628. Avci, F. *et al.* Glycoconjugates: What It Would Take To Master These Well-Known yet Little-Understood Immunogens for Vaccine Development. *mSphere* **4**, 1–8 (2019).
  629. Avci, F. Y. & Kasper, D. L. How Bacterial Carbohydrates Influence the Adaptive Immune System. *Annu. Rev. Immunol.* **28**, 107–130 (2009).
  630. Avci, F. Y., Li, X., Tsuji, M. & Kasper, D. L. Isolation of carbohydrate-specific CD4+ T cell clones from mice after stimulation by two model glycoconjugate vaccines. *Nat. Protoc.* **7**, 2180–2192 (2012).
  631. Middleton, D. R., Sun, L., Paschall, A. V & Avci, F. Y. T Cell Mediated Humoral Immune Responses to Type 3 Capsular Polysaccharide of *Streptococcus pneumoniae*. *J. Immunol.* **199**, 598–603 (2017).
  632. Egan, W., Frasch, C. E. & Anthony, B. F. Lot-Release Criteria, Post licensure Quality Control and the *Haemophilus influenzae* Type b Conjugate Vaccines. *JAMA J. Am. Med. Assoc.* **273**, 888–889 (1995).
  633. Paoletti, L. C. *et al.* Effects of chain length on the immunogenicity in rabbits of Group B *Streptococcus* type III oligosaccharide-tetanus toxoid conjugates. *J. Clin. Invest.* **89**, 203–209 (1992).

634. Laferrière, C. A., Sood, R. K., de Muys, J. M., Michon, F. & Jennings, H. J. The synthesis of *Streptococcus pneumoniae* polysaccharide-tetanus toxoid conjugates and the effect of chain length on immunogenicity. *Vaccine* **15**, 179–186 (1997).
635. Michon, F. *et al.* Group B Streptococcal Type II and III Conjugate Vaccines: Physicochemical Properties That Influence Immunogenicity. *Clin. Vaccine Immunol.* **13**, 936–943 (2006).
636. Rana, R. *et al.* Development and characterization of *Haemophilus influenzae* type b conjugate vaccine prepared using different polysaccharide chain lengths. *Vaccine* **33**, 2646–2654 (2015).
637. Carmenate, T. *et al.* Effect of conjugation methodology on the immunogenicity and protective efficacy of Meningococcal group C polysaccharide-P64k protein conjugates. *FEMS Immunol. Med. Microbiol.* **40**, 193–199 (2004).
638. Micoli, F. *et al.* Vi-CRM 197 as a new conjugate vaccine against *Salmonella Typhi*. *Vaccine* **29**, 712–720 (2011).
639. Tontini, M. Characterization of carbohydrate based vaccines. (Universite de Cergy Pontoise, 2012).
640. Arcuri, M. *et al.* The influence of conjugation variables on the design and immunogenicity of a glycoconjugate vaccine against *Salmonella typhi*. *PLoS One* **12**, e0189100 (2017).
641. Anderson, P. W. *et al.* Effect of oligosaccharide chain length, exposed terminal group, and hapten loading on the antibody response of human adults and infants to vaccines consisting of *Haemophilus influenzae* type b capsular antigen unterminally coupled to the diphtheria . *J. Imm* **142**, 2464–2468 (1989).
642. Dagan, R., Eskola, J., Leclerc, C. & Leroy, O. Reduced response to multiple vaccines sharing common protein epitopes that are administered simultaneously to infants. *Infect. Immun.* **66**, 2093–2098 (1998).
643. Renjifo, X. *et al.* Carrier-induced, hapten-specific suppression: A problem of antigen presentation? *J. Immunol.* **161**, 702–706 (1998).
644. Bottomley, J., Rappuoli, R. & Lambert, P.-H. Vaccine design in the 21st century. in *The vaccine book* 45–65 (2016).
645. Lee, L. H. & Blake, M. S. Effect of increased CRM197 carrier protein dose on meningococcal C bactericidal antibody response. *Clin. Vaccine Immunol.* **19**, 551–556 (2012).
646. Biemans, R., Micoli, F. & Romano, M. R. Glycoconjugate vaccines, production and characterization. in *Recent Trends in Carbohydrate Chemistry* 285–313 (2020).
647. Nita-Lazar, M., Wacker, M., Schegg, B., Amber, S. & Aebi, M. The N-X-S/T consensus sequence is required but not sufficient for bacterial N-linked protein glycosylation. *Glycobiology* **15**, 361–367 (2005).
648. Gerber, S. *et al.* Mechanism of bacterial oligosaccharyltransferase: *in vitro* quantification of sequon binding and catalysis. *J. Biol. Chem.* **288**, 8849–8861 (2013).

649. Shulman, S. *et al.* Differences in Antibody Response to Streptococcal Antigens in Children with Rheumatic and Non-rheumatic Mitral Valve Disease. *Circulation* **50**, 1244–1251 (1974).
650. Ayoub, E. M., Taranta, A. & Bartley, T. D. Effect of valvular surgery on antibody to the Group A Streptococcal carbohydrate. *Circulation* **50**, 144–150 (1974).
651. Theilacker, C. *et al.* Construction and Characterization of a *Pseudomonas aeruginosa* Mucoic Exopolysaccharide-Alginate Conjugate Vaccine. *Infect. Immun.* **71**, 3875–3884 (2003).
652. Mawas, F. *et al.* Immune interaction between components of acellular Pertussis-Diphtheria- Tetanus (DTaP) vaccine and *Haemophilus influenzae* b (Hib) conjugate vaccine in a rat model. *Vaccine* **24**, 3505–3512 (2006).
653. Silverman, J. M. & Imperiali, B. Bacterial N-glycosylation efficiency is dependent on the structural context of target sequons. *J. Biol. Chem.* **291**, 22001–22010 (2016).
654. Marshall, L. E. *et al.* An O-Antigen Glycoconjugate Vaccine Produced Using Protein Glycan Coupling Technology Is Protective in an Inhalational Rat Model of Tularemia. *J. Immunol. Res.* **2018**, 1–12 (2018).
655. Brittan, J. L., Buckeridge, T. J., Finn, A., Kadioglu, A. & Jenkinson, H. F. Pneumococcal Neuraminidase A: an essential upper airway colonization factor for *Streptococcus pneumoniae*. *Mol. Oral Microbiol.* **27**, 270–283 (2012).
656. Soong, G. *et al.* Bacterial neuraminidase facilitates mucosal infection by participating in biofilm production. *J. Clin. Invest.* **116**, 2297–2305 (2006).
657. Holmes, A. R. *et al.* The *pavA* gene of *Streptococcus pneumoniae* encodes a fibronectin-binding protein that is essential for virulence. *Mol. Microbiol.* **41**, 1395–1408 (2001).
658. Wannamaker, L. W., Denny, F. W., Perry, W. D., Siegel, A. C. & Rammelkamp, C. H. J. Studies on immunity to Streptococcal infections in man. *Am. J. Dis. Child.* **86**, 347–348 (1953).
659. Palmieri, E. *et al.* GMMA as an Alternative Carrier for a Glycoconjugate Vaccine against Group A *Streptococcus*. *Vaccines* **10**, 1–17 (2022).
660. Stewart-Tull, D. E. S. Adjuvant formulations for experimental vaccines. *Methods Mol. Med.* **87**, 175–194 (2003).
661. Weniger, B. G. The Vaccine Book. *Emerg. Infect. Dis.* **10**, 1347–1348 (2004).
662. Micoli, F. *et al.* Generalized Modules for Membrane Antigens as Carrier for Polysaccharides: Impact of Sugar Length, Density, and Attachment Site on the Immune Response Elicited in Animal Models. *Front. Immunol.* **12**, 1–12 (2021).
663. Micoli, F. *et al.* Short Vi-polysaccharide abrogates T-independent immune response and hyporesponsiveness elicited by long Vi-CRM197 conjugate vaccine. *Proc. Natl. Acad. Sci.* **117**, 24443–24449 (2020).
664. Sumby, P., Tart, A. H. & Musser, J. M. A non-human primate model of acute Group A *Streptococcus* pharyngitis. *Methods Mol. Biol.* **431**, 255–267 (2008).
665. Skinner, J. M. *et al.* Comparison of rhesus and cynomolgus macaques in a

- Streptococcus pyogenes* infection model for vaccine evaluation. *Microb. Pathog.* **50**, 39–47 (2011).
666. Gorton, D., Blyth, S., Gorton, J. G., Govan, B. & Ketheesan, N. An alternative technique for the induction of autoimmune valvulitis in a rat model of rheumatic heart disease. *J. Immunol. Methods* **355**, 80–85 (2010).
  667. Quinn, A., Kosanke, S., Fischetti, V. A., Factor, S. M. & Cunningham, M. W. Induction of autoimmune valvular heart disease by recombinant Streptococcal M protein. *Infect. Immun.* **69**, 4072–4078 (2001).
  668. Osowicki, J. *et al.* A controlled human infection model of *Streptococcus pyogenes* pharyngitis (CHIVAS-M75): an observational, dose-finding study. *The Lancet Microbe* **5247**, 291–299 (2021).
  669. Pace, D. & Pollard, A. J. Meningococcal A, C, Y and W-135 polysaccharide-protein conjugate vaccines. *Arch. Dis. Child.* **92**, 909–915 (2007).
  670. Acharya, I. L. *et al.* Prevention of typhoid fever in Nepal with the Vi capsular polysaccharide of *Salmonella typhi*. A preliminary report. *N. Engl. J. Med.* **317**, 1101–1104 (1987).
  671. Geno, K. A. *et al.* Pneumococcal Capsules and Their Types: Past, Present, and Future. *Clin. Microbiol. Rev.* **28**, 871–899 (2015).
  672. Plotkin, S., Robinson, J. M., Cunningham, G., Iqbal, R. & Larsen, S. The complexity and cost of vaccine manufacturing - an overview. *Vaccine* **35**, 4064–4071 (2017).
  673. Rappuoli, R. & Hanon, E. Sustainable vaccine development: a vaccine manufacturer's perspective. *Curr. Opin. Immunol.* **53**, 111–118 (2018).
  674. McCarty, M. Variation in the group-specific carbohydrate of group A *Streptococci*. Studies on the chemical basis for serological specificity of the carbohydrates. *J. Exp. Med.* **104**, 629–643 (1956).
  675. Shulman, S. T. & Ayoub, E. M. Serologic cross-reactions among Streptococcal Group A, A-variant, and C polysaccharides. *Clin. Immunol. Immunopathol.* **28**, 229–242 (1983).
  676. Lawson, P. A., Foster, G., Falsen, E., Markopoulos, S. J. & Collins, M. D. *Streptococcus castoreus* sp. nov., isolated from a beaver (*Castor fiber*). *Int. J. Syst. Evol. Microbiol.* **55**, 843–846 (2005).
  677. Anish, C., Beurret, M. & Poolman, J. Combined effects of glycan chain length and linkage type on the immunogenicity of glycoconjugate vaccines. *npj Vaccines* **6**, 150 (2021).
  678. Khatun, F., Stephenson, R. J. & Toth, I. An Overview of Structural Features of Antibacterial Glycoconjugate Vaccines That Influence Their Immunogenicity. *Chemistry* **23**, 4233–4254 (2017).
  679. Lee, C. *et al.* Quality Improvement of Capsular Polysaccharide in *Streptococcus pneumoniae* by Purification Process Optimization. *Front. Bioeng. Biotechnol.* **8**, 1–11 (2020).
  680. Jefferies, J. M. C., Macdonald, E., Faust, S. N. & Clarke, S. C. 13-valent Pneumococcal conjugate vaccine (PCV13). *Hum. Vaccin.* **7**, 1012–1018 (2011).
  681. Wals, K. & Ovaa, H. Unnatural amino acid incorporation in *E. coli*: current and

- future applications in the design of therapeutic proteins. *Front. Chem.* **2**, 1–12 (2014).
682. Nilo, A. *et al.* Tyrosine-Directed Conjugation of Large Glycans to Proteins via Copper-Free Click Chemistry. *Bioconjug. Chem.* **25**, 2105–2111 (2014).
683. Adamo, R. *et al.* Synthetically defined glycoprotein vaccines: current status and future directions. *Chem. Sci.* **4**, 2995–3008 (2013).
684. Peeters, C. C. *et al.* A comparative study of the immunogenicity of Pneumococcal type 4 polysaccharide and oligosaccharide Tetanus Toxoid conjugates in adult mice. *J. Immunol.* **146**, 4308–4314 (1991).
685. Steinhoff, M. C. *et al.* A randomized comparison of three bivalent *Streptococcus pneumoniae* glycoprotein conjugate vaccines in young children: effect of polysaccharide size and linkage characteristics. *Pediatr. Infect. Dis. J.* **13**, 368–372 (1994).
686. Stefanetti, G., Okan, N., Fink, A., Gardner, E. & Kasper, D. L. Glycoconjugate vaccine using a genetically modified O antigen induces protective antibodies to *Francisella tularensis*. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 7062–7070 (2019).
687. Ravenscroft, N. *et al.* Purification and characterization of a *Shigella* conjugate vaccine, produced by glycoengineering *Escherichia coli*. *Glycobiology* **26**, 51–62 (2016).
688. Mariño, K., Bones, J., Kattla, J. J. & Rudd, P. M. A systematic approach to protein glycosylation analysis: a path through the maze. *Nat. Chem. Biol.* **6**, 713–723 (2010).
689. Morelle, W. & Michalski, J.-C. Analysis of protein glycosylation by mass spectrometry. *Nat. Protoc.* **2**, 1585–1602 (2007).
690. Zhu, Z. & Desaire, H. Carbohydrates on Proteins: Site-Specific Glycosylation Analysis by Mass Spectrometry. *Annu. Rev. Anal. Chem.* **8**, 463–483 (2015).
691. Giménez, E., Benavente, F., Barbosa, J. & Sanz-Nebot, V. Towards a reliable molecular mass determination of intact glycoproteins by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun. mass Spectrom.* **21**, 2555–2563 (2007).
692. Jahouh, F., Hou, S., Kováč, P. & Banoub, J. H. Determination of the glycation sites of *Bacillus anthracis* neoglycoconjugate vaccine by MALDI-TOF/TOF-CID-MS/MS and LC-ESI-QqTOF-tandem mass spectrometry. *J. Mass Spectrom.* **46**, 993–1003 (2011).
693. Jahouh, F., Xu, P., Vann, W. F., Kováč, P. & Banoub, J. H. Mapping the glycation sites in the neoglycoconjugate from hexasaccharide antigen of *Vibrio cholerae*, serotype Ogawa and the recombinant Tetanus Toxin C-fragment carrier. *J. Mass Spectrom.* **48**, 1083–1090 (2013).
694. Giannelli, C. *et al.* Determination of free polysaccharide in Vi glycoconjugate vaccine against typhoid fever. *J. Pharm. Biomed. Anal.* **139**, 143–147 (2017).
695. Rajendar, B., Reddy, M. V. N. J., Suresh, C. N. V, Bheemaraju, S. & Matur, R. V. Determination of free carrier protein in glycoconjugate vaccines by size exclusion chromatography–high performance liquid chromatography with fluorescence detector. *Sep. Sci. PLUS* **4**, 280–285 (2021).



696. Beresford, N. J. *et al.* Quality, immunogenicity and stability of meningococcal serogroup ACWY-CRM(197), DT and TT glycoconjugate vaccines. *Vaccine* **35**, 3598–3606 (2017).
697. Stefanetti, G., Borriello, F., Richichi, B., Zanoni, I. & Lay, L. Immunobiology of Carbohydrates: Implications for Novel Vaccine and Adjuvant Design Against Infectious Diseases. *Front. Cell. Infect. Microbiol.* **11**, 1–23 (2022).
698. Koshy, S. T. & Mooney, D. J. Biomaterials for enhancing anti-cancer immunity. *Curr. Opin. Biotechnol.* **40**, 1–8 (2016).
699. Zhang, Y., Lin, S., Wang, X.-Y. & Zhu, G. Nanovaccines for cancer immunotherapy. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **11**, e1559 (2019).
700. van der Pol, L., Stork, M. & van der Ley, P. Outer membrane vesicles as platform vaccine technology. *Biotechnol. J.* **10**, 1689–1706 (2015).
701. Mancini, F., Rossi, O., Necchi, F. & Micoli, F. OMV Vaccines and the Role of TLR Agonists in Immune Response. *Int. J. Mol. Sci.* **12**, 1–20 (2020).
702. Berti, F. & Micoli, F. Improving efficacy of glycoconjugate vaccines: from chemical conjugates to next generation constructs. *Curr. Opin. Immunol.* **65**, 42–49 (2020).
703. Varki, A. *et al.* Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **25**, 1323–1324 (2015).
704. Neelamegham, S. *et al.* Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology* **29**, 620–624 (2019).
705. Cuccui, J. & Wren, B. Hijacking bacterial glycosylation for the production of glycoconjugates, from vaccines to humanised glycoproteins. *J. Pharm. Pharmacol.* **67**, 338–350 (2015).