Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.
Structural and Immunological analysis of the T antigen: a vaccine candidate for *Streptococcus pyogenes*

Jeremy Raynes

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

Doctor of Philosophy

Department of Molecular Medicine and Pathology

The University of Auckland

2019
Abstract

*Streptococcus pyogenes* is a globally important pathogen causing a broad range of human disease and significant morbidity and mortality. No vaccines for *S. pyogenes* are currently available, but vaccine candidates based on the T antigen, which forms the backbone of the *S. pyogenes* pilus, are in pre-clinical development. Pilus proteins have previously been shown to have protective properties and exhibit relatively low antigenic variation. This project aimed to investigate antibody-T antigen interactions and provide 3-dimensional structural data of protective epitopes to inform structure-led, vaccine design.

Large Fab-based antibody phage-display libraries were generated from mice vaccinated with the T18.1 pilus and were biopanned to identify individual Fab that bound to T18.1. The strain-specificity and cross-reactivity of the isolated Fab were determined using a panel of T antigens that covers all of the major circulating strain types. The binding affinity of the Fab to T18.1 as a monomer and in the polymerised pilus was assessed. Biopanning identified 20 unique Fab that bound to T18.1 with high affinity. Four promising Fab were selected for further characterisation. Three of these were T18-specific Fab while one also cross-reacted with T3.2.

The epitopes of the four selected Fab were mapped using Fab-T18.1 co-crystallography and peptide tiling. The atomic structures of T18.1 in isolation and in complex with one of the Fab were determined by x-ray crystallography. The co-crystal and peptide tiling data mapped the epitopes of all four antibodies to a single 35-residue region on the N-domain of T18.1. This region was also targeted by antibodies in human and animal sera. Competition ELISAs and flow cytometry identified multiple epitopes within this region and indicated that they have varying accessibility in the polymerised pilus.
The protective capacity of the four T18.1-specific monoclonal antibodies was measured in vitro using traditional whole blood bactericidal assays and a new HL-60-based opsonophagocytic killing assay. All four of the selected antibodies were found to be protective.

This thesis has identified four protective antibodies that bind to a single region on the T antigen. This fundamental knowledge on how antibodies interact with T antigens to elicit protection will inform T antigen vaccine development.
Acknowledgments

First and foremost I would like to thank Dr Nikki Moreland for her knowledge, guidance and support provided over the course of this project. I am sincerely grateful for her mentorship over the last several years.

I would also like to thank Dr Paul Young for all of the time he spent helping me with all things structural biology. Also Dr Jace Loh, Dr Reuben MacGregor and Fiona Clow for their various technical expertise.

I am particularly grateful for the opportunity that arose that allowed me to do a section of this work in Professor David Goldblatt’s laboratory at University College London. I would like to thank David and Dr Marta Zancolli for taking me under their wings. Thank you for all of the advice, the always interesting chats, and the introduction to prosecco.

Thank you to the people in the lab and in the office that make every day enjoyable. I would especially like to thank Mei Lin and Sam for their always entertaining company in the lab, office and on conferences. Vinaka.

I would also like to thank my pilin PhD buddy Lia for being there to celebrate all of the highs and commiserate all of the lows. I must also thank her for introducing me to the many buffet restaurants of Auckland.

I am deeply grateful to my parents and siblings for their unconditional love, support, and encouragement in all of my pursuits. Thank you.

Last, and certainly not least, I would like to thank Brett. I am eternally grateful for your unwavering love and support in all things.
Table of Contents

Abstract ........................................................................................................................................ ii

Acknowledgments ........................................................................................................................ iv

List of Tables .................................................................................................................................. x

List of Figures .................................................................................................................................. xii

Abbreviations ................................................................................................................................. xvi

1 Introduction .................................................................................................................................. 20

1.1 *Streptococcus pyogenes* in disease .................................................................................. 20

1.2 Epidemiology ......................................................................................................................... 22

1.2.1 The burden of *Streptococcus pyogenes* disease in New Zealand ...................... 23

1.3 The host defence against *Streptococcus pyogenes* ...................................................... 24

1.4 *Streptococcus pyogenes* virulence factors .................................................................. 25

1.4.1 The hyaluronic acid capsule ....................................................................................... 27

1.4.2 Secretion and anchoring of surface expressed virulence factors to the cell-wall ... 29

1.4.3 The M protein ............................................................................................................... 30

1.4.4 The *Streptococcus pyogenes* Pilus ........................................................................... 34

1.5 *Streptococcus pyogenes* vaccinology .............................................................................. 40

1.5.1 Obstacles to the development of a vaccine for *Streptococcus pyogenes* .......... 41

1.5.2 The M protein as a vaccine candidate ....................................................................... 45

1.5.3 The T antigen as a vaccine candidate ........................................................................ 48
2.2.8 Biosensor analysis of Fab binding to T18.1 .................................................. 102
2.2.9 Flow cytometry analysis of IgG1 binding to the polymerised T18.1 pilus ..... 103
2.2.10 In vitro protection assays ............................................................................. 103

3 Structural analysis of T18.1 ................................................................................. 107
  3.1 Introduction ........................................................................................................ 107
  3.2 Results ................................................................................................................ 108
    3.2.1 Cloning of tee18.1 ...................................................................................... 108
    3.2.2 Expression and purification of T18.1 ............................................................. 109
    3.2.3 Crystallisation of T18.1 ............................................................................... 110
    3.2.4 Structural analysis of T18.1 ......................................................................... 111
  3.3 Discussion ............................................................................................................ 120

4 Generation and characterisation of T antigen antibodies from mouse-human phagemid libraries .......................................................................................... 124
  4.1 Introduction ........................................................................................................ 124
  4.2 Results ................................................................................................................ 126
    4.2.1 Animal vaccinations ..................................................................................... 126
    4.2.2 Preparation of genetic material from mouse spleens ..................................... 128
    4.2.3 Generation of chimeric mouse/human Fab libraries ....................................... 129
    4.2.4 Preparation of biotinylated T18.1 for biopanning .......................................... 134
    4.2.5 Biopanning the phagemid library against T18.1 ............................................ 137
    4.2.6 Screening individual clones from the pil18 mouse phagemid library .......... 140

vii
4.2.7 Assessing the diversity of the Fab library .............................................................. 141
4.2.8 Expression of recombinant Fab ............................................................................. 144
4.2.9 Estimating the affinity of T18.1 binding Fab ......................................................... 146
4.2.10 Determining the reactivity of the T18.1 binding Fab to a T antigen panel ....... 148
4.2.11 Selection of Fab for epitope mapping and protection studies ......................... 153
4.3 Discussion ............................................................................................................... 154

5 Functional characterisation of T18.1 binding antibodies ....................................... 157
5.1 Introduction .............................................................................................................. 157
5.2 Results .................................................................................................................... 159

5.2.1 Characterisation of full-length IgG1 produced from the selected Fab .......... 159
5.2.2 Analysis of the affinity and kinetics of anti-T18.1 Fab using Surface Plasmon Resonance .................................................................................................................. 163
5.2.3 Analysis of binding to the polymerised pilus using flow cytometry .......... 166
5.2.4 Epitope mapping of the selected Fab and T18.1 by co-crystallography .......... 168
5.2.5 Binding of the αE3 Fab does not prevent pilus assembly ................................. 182
5.2.6 Epitope mapping of the selected Fab and T18.1 using an overlapping peptide library ........................................................................................................................................ 184
5.2.7 Animals vaccinated with T18.1-containing vaccines generate circulating antibodies that bind to the same immunogenic region of T18.1 ................................. 194
5.2.8 Human paediatric patients who generate an immune response to T18.1 produce circulating antibodies which bind to the same immunogenic region of T18.1 .......... 196
5.2.9 Analysing the protective capacity of T18.1 antibodies ........................................ 197

5.2.10 Discussion ............................................................................................................ 204

6 Discussion .................................................................................................................. 210

6.1 Project Summary ...................................................................................................... 210

6.2 Understanding Antibody-T antigen interactions .................................................... 211

6.2.1 Insight into T antigen cross-reactivity ................................................................. 211

6.2.2 The heavy chain dominant binding mode of αE3 ................................................ 215

6.2.3 Understanding T antigen-based protection ......................................................... 216

6.3 Implications for Vaccinology .................................................................................... 219

6.4 Implications for the *S. pyogenes* pilus model ....................................................... 221

6.5 Conclusion ................................................................................................................ 223

7 Appendix A: Primers for the construction of the chimeric mouse/human Fab library . 224

8 Appendix B: Papers published during this PhD ......................................................... 230

9 References .................................................................................................................... 232
List of Tables

Table 1.1 Proteins in the SPy7 vaccine.................................................................50
Table 2.1 Commercial antibodies used in this thesis........................................69
Table 2.2 Buffers and solutions used for SDS-polyacrylamide gel electrophoresis.............71
Table 2.3 Primers used for the amplification of tee18.1 from pET32a 3C-tee18.1..............73
Table 2.4 PCR parameters used for the amplification of tee18.1 from pET32a 3C-tee18.1...73
Table 2.5 Double-digest of the purified tee18.1 PCR products and pProEXHTa-Avitag with Kas1 and Xba1..............................................................74
Table 2.6 Primer pairs for the amplification of mouse V_H and V_L genes....................85
Table 2.7 PCR parameters used for the amplification of mouse V_H and V_L genes..........86
Table 2.8 PCR parameters for the amplification of the human C_H1γ and C_κ regions ..........87
Table 2.9 Parameters for overlap extension PCR to generate chimeric mouse/human heavy and light chains.................................................................88
Table 2.10 Parameters for overlap extension PCR to generate chimeric mouse/human Fab constructs.................................................................89
Table 2.11 Restriction digestion of Fab and pComb3XSS with Sfi1 ...............................90
Table 2.12 Biopanning conditions...........................................................................95
Table 2.13 The concentrations of the Fab and peptides used in the competition ELISA.....100
Table 3.1 Data collection and refinement statistics for T18.1 ......................................114
Table 4.1 Summary of solution-phase biopanning against T18.1 ...............................139
Table 4.2 Sequence analysis of the 23 Fab clones identified by phage display.............143
Table 4.3 Estimated affinity of the eighteen recombinantly expressed Fab for T18.1........148
Table 4.4 Properties of the recombinantly expressed T antigens used in this study. ..........150
Table 4.5 Summary of the Fab selected for epitope mapping and protection studies. ........153
Table 5.1 Collection and analysis parameters for αE3-T18.1........................................176
Table 5.2 Hydrogen bonds between the αE3 Fab and T18.1........................................180
Table 5.3 Properties of the T18.1 overlapping peptide library used in this study. ..........186
Table 5.4 Summary of S. pyogenes killing achieved with the anti-T18.1 rabbit serum in the
HL-60 OPKA. These data are representative of at least two individual experiments. .......200
Table 5.5 Testing the specificity of killing in the S. pyogenes OPKA. ..............................201
List of Figures

Figure 1.1 *S. pyogenes* produce an array of secreted and surface-bound virulence factors that enable colonisation and interfere with the host immune system. .................................................27

Figure 1.2 Model of the dimeric M protein from M1T1 *S. pyogenes* with the partial structure of the A and B regions shown in ribbon form. .................................................................31

Figure 1.3 The structure of the *S. pyogenes* pili and T antigens. .............................................35

Figure 1.4 Phylogenetic tree showing the full range of tee gene diversity organised into six sub clades. ............................................................................................................................39

Figure 1.5 Schematic illustrating the *S. pyogenes* OPKA .............................................................44

Figure 1.6 Schematic of an IgG1 antibody ..................................................................................53

Figure 1.7 Flow diagram outlining the construction and panning of the mouse Fab library...58

Figure 3.1 Immobilised metal affinity chromatography of T18.1 using a HisTrap HP column charged with 100 mM Nickel Chloride .........................................................................................109

Figure 3.2 Size Exclusion Chromatography of T18.1 using a Superdex 200 10/300GL column.................................................................................................................................110

Figure 3.3 Typical crystal morphologies of T18.1 crystals .........................................................111

Figure 3.4 Data collection of T18.1 .............................................................................................112

Figure 3.5 The structure of T18.1 presented as a ribbon diagram .............................................115

Figure 3.6 Sequence alignment of mature T18.1 and T1 proteins .............................................117

Figure 3.7 Structural overlay of the polypeptide backbone of T1 and T18.1 .........................118
Figure 3.8 Identification of sequence conservation between T18.1 and T1 mapped onto the T1 structure ................................................................. 119

Figure 3.9 Visualisation of sequence conservation between T18.1 and T1 mapped onto the T1 surface ........................................................................................................ 120

Figure 4.1 Display of a Fab fragment on the pIII coat protein of M13 filamentous phage... 126

Figure 4.2 Schematic showing the T18.1 pilus and the two T18.1-containing vaccines used in this thesis............................................................................................................. 127

Figure 4.3 Antibody endpoint titres of vaccinated animal sera to T18.1......................... 128

Figure 4.4 Amplification of immunoglobulin variable genes from pil18 mouse cDNA.....130

Figure 4.5 PCR amplification of human constant regions from pComb3XTT............... 130

Figure 4.6 Overlap extension of variable and constant gene regions from mouse B1 .......131

Figure 4.7 Overlap extension of light-chain and heavy-chain fragments from mouse B1 .....132

Figure 4.8 SfiI digestion of Fab and pComb3XSS ...................................................... 133

Figure 4.9 Purification of T18.1-biotin........................................................................ 136

Figure 4.10 Pull-down of biotinylated T18.1 on streptavidin-coupled beads............... 137

Figure 4.11 Polyclonal phage ELISA ........................................................................ 140

Figure 4.12 Monoclonal phage ELISA........................................................................ 141

Figure 4.13 Analysis of the CDR3 regions in the 20 unique Fab clones......................... 144

Figure 4.14 Immobilised metal affinity chromatography of αH3 Fab using a HisTrap HP column............................................................................................................. 145

Figure 4.15 Binding of the eighteen Fab to T18.1...................................................... 147
Figure 4.16 SDS-PAGE analysis of the T antigen panel .......................................................... 151

Figure 4.17 Specificity of the eighteen recombinantly expressed Fab and animal sera to the T antigen panel ........................................................................................................................................ 152

Figure 4.18 Protein alignment T3.2, T18.1 and T18.2 ............................................................ 156

Figure 5.1 Comparison of IgG1 and Fab binding to T18.1 ................................................ 161

Figure 5.2 Specificity of the IgG1 clones to the T antigen panel ....................................... 162

Figure 5.3 Analysis of βC1 binding to T18.1 using surface plasmon resonance ............... 164

Figure 5.4 Comparison of the kinetics of Fab binding to T18.1 ....................................... 166

Figure 5.5 Binding of anti-T18.1 IgG1 to the polymerised pilus ..................................... 168

Figure 5.6 Size exclusion chromatography of Fab-T18.1 complex formation and stability. 171

Figure 5.7 SDS-PAGE analysis of eluted αE3-T18.1 complex after SEC ....................... 172

Figure 5.8 Data collection of the Fab-T18.1 co-crystals .................................................... 174

Figure 5.9 The structure of the αE3-T18.1 complex presented as a ribbon diagram ........ 178

Figure 5.10 ProFunc calculation of the αE3 epitope on T18.1 .......................................... 179

Figure 5.11 The αE3 epitope mapped onto the surface of T18.1 ...................................... 181

Figure 5.12 Binding of the αE3 Fab induces the movement of three loops in T18.1 .......... 182

Figure 5.13 Pilus assembly in the presence of αE3 Fab ..................................................... 183

Figure 5.14 Immunoblot analysis of the selected Fab binding to T18.1 ......................... 185

Figure 5.15 Mapping the linear epitopes of the four selected Fab using the T18.1 overlapping peptide library ........................................................................................................... 188
Figure 5.16 Capture ELISA competing the four Fab and the four IgG1 to identify overlapping epitopes. .................................................................................................................................................191

Figure 5.17 Competition ELISA competing the four Fab and the peptides spanning their epitopes ................................................................................................................................................193

Figure 5.18 Competition ELISA competing the Fab with sera from the vaccinated animals ..................................................................................................................................................195

Figure 5.19 Mapping the linear epitopes of three human patients using the T18.1 overlapping peptide library. ..............................................................................................................................................197

Figure 5.20 OPKA killing curve for M217/T18.1 after 45 min with rabbit anti-T18.1 serum and matched pre-immune serum ........................................................................................................................................199

Figure 5.21 The vaccinated animals and monoclonal IgG1 have bactericidal activity against M217/T18.1 .................................................................................................................................................204

Figure 5.22 The immunogenic region mapped onto the surface of T18.1 .........................206

Figure 6.1 Protein alignment of the 35-residue immunogenic region (A111-D145) in all 2-domain T antigens in the panel ........................................................................................................................................212

Figure 6.2 Protein alignment of the CDRs of the four antibody clones identified in this project. .................................................................................................................................................214

Figure 6.3 The binding of the αE3 Fab to the polymerised T18.1 pilus indicates significant flexibility between T antigen monomers. ........................................................................................................223
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP1</td>
<td>Ancillary Protein 1</td>
</tr>
<tr>
<td>AP2</td>
<td>Ancillary Protein 2</td>
</tr>
<tr>
<td>ARF</td>
<td>Acute Rheumatic Fever</td>
</tr>
<tr>
<td>BRC</td>
<td>Baby Rabbit Complement</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>CDR</td>
<td>Complementarity Determining Region</td>
</tr>
<tr>
<td>CFU</td>
<td>Colony Forming Units</td>
</tr>
<tr>
<td>DMF</td>
<td>Dimethylformamide</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>ECM</td>
<td>Extra-cellular matrix</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>ESR</td>
<td>The Institute of Environmental Science and Research</td>
</tr>
<tr>
<td>Fab</td>
<td>Fragment antigen binding</td>
</tr>
<tr>
<td>FDA</td>
<td>The Food and Drug Administration</td>
</tr>
<tr>
<td>fmoc</td>
<td>Fluorenylmethoxycarbonyl chloride</td>
</tr>
<tr>
<td>GBS</td>
<td>Group B Streptococcus</td>
</tr>
<tr>
<td>HA</td>
<td>Hyaluronic acid</td>
</tr>
<tr>
<td>HCDR3</td>
<td>CDR3 of the heavy chain</td>
</tr>
</tbody>
</table>
LCDR3: CDR3 of the light chain
His6: Polyhistidine tag
HVR: Hyper Variable Region
IN: Intranasal
IdeS: IgG degrading enzyme of *S. pyogenes*
IFA: Incomplete Freund’s adjuvant
Ig: Immunoglobulin
IMAC: Immobilized Metal Ion Affinity Chromatography
IPTG: Isopropyl β-D-1-thiogalactopyranoside
IVIG: Intravenous Immunoglobulin
*L. lactis*: *Lactococcus lactis*
MBP: Maltose Binding Protein
MFI: Mean Fluorescence Intensity
MOPS: 3-(N-Morpholino) propanesulfonic acid
MR: Molecular Replacement
NET: Neutrophil Extracellular Trap
OPKA: Opsonophagocytic Killing Assay
PCR: Polymerase Chain Reaction
PDB: Protein Data Bank
PEG: Polyethylene Glycol
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFU</td>
<td>Plaque Forming Units</td>
</tr>
<tr>
<td>PSGN</td>
<td>Poststreptococcal Glomerulonephritis</td>
</tr>
<tr>
<td>RHD</td>
<td>Rheumatic Heart Disease</td>
</tr>
<tr>
<td>Rmsd</td>
<td>Root Mean Square Deviation</td>
</tr>
<tr>
<td>RocA</td>
<td>Regulator of Cov A</td>
</tr>
<tr>
<td>rTEV</td>
<td>Recombinant <em>Tobacco Etch Virus</em> protease</td>
</tr>
<tr>
<td>RU</td>
<td>Response Unit</td>
</tr>
<tr>
<td><em>S. pyogenes</em></td>
<td><em>Streptococcus pyogenes</em></td>
</tr>
<tr>
<td>S.C</td>
<td>Subcutaneous</td>
</tr>
<tr>
<td>scFV</td>
<td>Single chain variable fragment</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>SEC</td>
<td>Size Exclusion Chromatography</td>
</tr>
<tr>
<td>SIC</td>
<td>Streptococcal Inhibitor of Complement</td>
</tr>
<tr>
<td>SLO</td>
<td>Streptolysin O</td>
</tr>
<tr>
<td>SLS</td>
<td>Streptolysin S</td>
</tr>
<tr>
<td>SpeB</td>
<td>Streptococcal pyrogenic exotoxin B</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface Plasmon Resonance</td>
</tr>
<tr>
<td>SpyCEP</td>
<td><em>S. pyogenes</em> Cell Envelope Proteinase</td>
</tr>
<tr>
<td>SRP</td>
<td>Signal Recognition Particle</td>
</tr>
<tr>
<td>V&lt;sub&gt;H&lt;/sub&gt;</td>
<td>Variable region of the heavy chain</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>$V_L$</td>
<td>Variable region of the heavy chain</td>
</tr>
<tr>
<td>$V_\kappa$</td>
<td>Variable region of the kappa light chain</td>
</tr>
<tr>
<td>$V_\lambda$</td>
<td>Variable region of the lambda light chain</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Streptococcus pyogenes in disease

*Streptococcus pyogenes*, also known as Group A Streptococcus, is a non-motile, gram-positive coccus that exclusively colonises and infects humans causing over 500,000 deaths globally per year (Carapetis, Steer, Mulholland, & Weber, 2005). *S. pyogenes* cause a vast array of clinical diseases including non-invasive, invasive and post infection autoimmune pathologies of varying severity. The majority of *S. pyogenes*-related deaths are attributed to rheumatic heart disease (RHD), a chronic condition caused by repeated episodes of acute rheumatic fever (ARF) (Carapetis et al., 2016).

*S. pyogenes* primarily colonise the skin and oropharynx and are carried asymmetrically in approximately 2-20% of children (Berry et al., 2015; Delpech et al., 2017; Shaikh, Leonard, & Martin, 2010) and up to 5% of adults (Pearson et al., 2017; Steer, Lamagni, Curtis, & Carapetis, 2012). It is the leading cause of bacterial pharyngitis estimated to cause 37% of childhood pharyngitis and over 600 million cases each year (Carapetis et al., 2005; Shaikh et al., 2010). *S. pyogenes* strains expressing streptococcal pyrogenic exotoxins can cause scarlet fever, a historically important disease showing signs of global re-emergence (Lamagni et al., 2018; Lee, Cowling, & Lau, 2017; Park et al., 2017).

Superficial *S. pyogenes* skin infections cause pyoderma, such as impetigo, when limited to the epidermis but if the skin barrier is perturbed *S. pyogenes* can colonise deeper tissues. This can lead to the development of clinically severe invasive diseases. Cellulitis (infection of the subcutis) and bacteraemia are the most common invasive *S. pyogenes* diseases making up 40–80% of cases (Walker et al., 2014). Necrotising fasciitis and streptococcal toxic shock
syndrome (STSS) are less common but life-threatening with mortality rates above 80% if untreated (Davies et al., 1996; Low, 2013; O'Grady et al., 2007; O'Loughlin et al., 2007).

Superficial *S. pyogenes* infections are associated with two major post-infection autoimmune sequelae: Post Streptococcal Glomerulonephritis (PSGN) and ARF (potentially leading to RHD). PSGN is characterised by inflammation of the glomerulus, caused by the deposition of immune complexes, following a superficial *S. pyogenes* infection (Cunningham, 2008; Zheng et al., 2009). ARF is an autoimmune disease causing inflammation of the joints, brain, skin and heart valves that typically develops 1-5 weeks after infection with *S. pyogenes* in those 5-19 years of age (Gurney, Stanley, Baker, Wilson, & Sarfati, 2016; Lawrence, Carapetis, Griffiths, Edwards, & Condon, 2013; Parnaby & Carapetis, 2010). While the molecular pathogenesis of ARF remains poorly understood, it has been postulated that repeated untreated infections with *S. pyogenes* may prime the immune system of susceptible individuals for the development of ARF (Carapetis et al., 2016; Raynes et al., 2016; Zabriskie, Hsu, & Seegal, 1970). Subsequent infection with another, potentially ‘rheumatogenic’, strain of *S. pyogenes* causes a loss of tolerance and autoimmunity.

The mechanism behind this loss of tolerance is unknown but has been suggested to be due to molecular mimicry. The molecular mimicry hypothesis suggests that antibodies that bind to particular regions on coiled-coil antigens, (such as B repeats on the *S. pyogenes* M protein), may cross-react with human coiled-coil proteins causing autoimmunity (Henningham, Barnett, Maamary, & Walker, 2012). An alternative hypothesis suggests that the interaction between certain *S. pyogenes* M proteins and human collagen IV can lead to the development of non-cross-reactive anti-collagen antibodies which cause cardiac damage (Dinkla et al., 2003; Dinkla et al., 2009; Tandon et al., 2013). Repeated episodes of ARF cause increasing damage to the myocardium and heart valves (usually the aortic and mitral valves), which leads to the
development of RHD. Complications of RHD include endocarditis, embolic stroke and cardiac failure (Carapetis et al., 2016), and RHD is the most common cause of paediatric heart failure worldwide (Lee, Naguwa, Cheema, & Gershwin, 2009).

The classical ARF paradigm considers autoimmunity to be a consequence of untreated *S. pyogenes* pharyngitis. However, a mounting body of evidence suggests that ARF may also develop following skin infections, particularly in locations with high rates of impetigo and comparatively low rates of pharyngitis such as Fiji and the Northern Territories of Australia (McDonald et al., 2006; Parks, Smeesters, & Steer, 2012; Steer, Jenney, et al., 2009). Very recent data from New Zealand also suggests a role for *S. pyogenes* skin infections in ARF aetiology in more temperate climates (O’Sullivan, Moreland, Webb, Upton, & Wilson, 2017; Williamson et al., 2015).

### 1.2 Epidemiology

Globally, *S. pyogenes* disease is among the top ten infectious causes of death, causing an estimated 500,000 deaths each year (Carapetis et al., 2005). RHD in the developing world is the major source of mortality, causing an estimated 297,000 to 337,000 deaths per year (Watkins et al., 2017). The global prevalence of RHD (the number of people living with RHD) is estimated to be over 33 million. This is likely an underestimate of the true prevalence due the dearth of data from sub-Saharan Africa and the fact that this estimate does not include cases of subclinical RHD (Watkins et al., 2017). For comparison, this is similar to the prevalence of HIV/AIDS which is considered a global pandemic (World Health Organisation, 2017). The importance of RHD as a global disease has recently been recognised, with the World Health Organisation adopting a global resolution which calls for a “co-ordinated global response” to ARF and RHD (World Health Organisation, 2018).
Severe invasive *S. pyogenes* disease is also globally important, with an incidence of approximately 1.5 million cases per year. This causes 150,000 deaths each year, and has been increasing globally since the 1980s (Ralph & Carapetis, 2012).

While superficial disease is usually benign, the high case number (>600 million cases of pharyngitis and >100 million pyoderma per year) results in extensive antibiotic use and economic burden (Pfoh, Wessels, Goldmann, & Lee, 2008).

All *S. pyogenes* diseases are associated with socioeconomic status with >95% of invasive disease, PSGN and RHD, estimated to occur in the developing world (World Health Organisation, 2005). Even within developed countries such as Australia and New Zealand there are marked disparities in disease burden between at-risk and low-risk communities (Norton et al., 2004).

### 1.2.1 The burden of *Streptococcus pyogenes* disease in New Zealand

The overall burden of serious *S. pyogenes* disease in New Zealand decreased over the course of the 20th century, but the incidence of ARF (4.3/100,000) remains higher than in many developing countries (Gurney, Sarfati, Stanley, Wilson, & Webb, 2015). This burden of disease is not spread equally, with cases in Maori and Pacific Island populations accounting for approximately 90% of ARF cases in New Zealand (Jaine, Baker, & Venugopal, 2008). This substantial health inequality is focused in 5-14 year old Maori and Pacific Island children living in the upper North Island, with 60% of ARF cases occurring in the Auckland region (Milne, Lennon, Stewart, Vander Hoorn, & Scuffham, 2012). Within this group, the incidence of ARF is 40.2 and 81.2 per 100,000 (Maori and Pacific Island populations respectively) contributing to over 120 new ARF diagnoses and 150 RHD deaths each year (Jaine et al., 2008; Milne et al., 2012).
The incidence of ARF in New Zealand is associated with socioeconomic factors including deprivation status, crowded housing and access to primary healthcare (Gurney et al., 2015; Gurney et al., 2016; Jaine, Baker, & Venugopal, 2011; Milne et al., 2012). In order to address this health disparity, the New Zealand government has instigated primary health care initiatives as part of a national ‘Rheumatic Fever Prevention Programme’. This included school-based interventions that focused on the timely treatment of pharyngitis in high-risk areas, improving health literacy via health education campaigns, and attempts to improve housing quality (Anderson et al., 2016; Gray, Lennon, Anderson, Stewart, & Farrell, 2013; Lennon et al., 2017).

1.3 The host defence against Streptococcus pyogenes

The immune response to S. pyogenes is dependent on both the innate and adaptive arms of the immune system. Phagocytic neutrophils are an important early defence mechanism against extracellular bacteria and are the major cellular component implicated in clearance of S. pyogenes. Depletion of neutrophils (in mouse models) can render otherwise avirulent strains of S. pyogenes virulent (Hidalgo-Grass et al., 2006), and results in increased susceptibility to infection (Navarini et al., 2009). In addition to being phagocytic, neutrophils produce a number of antimicrobial products that act to either kill S. pyogenes, such as antimicrobial peptides and reactive oxygen species, or limit dissemination, such as Neutrophil Extracellular Traps (NETs) (Buchanan et al., 2006; Fieber & Kovarik, 2014; Sumby et al., 2005).

S. pyogenes is also susceptible to killing mediated by the complement system. Complement proteins are able to bind to the surface of S. pyogenes and form a pore in the bacterial membrane (the membrane-attack-complex). The membrane-attack-complex is a 100 Å pore that disrupts the integrity of the cell membrane, leading to osmotic lysis of the pathogen (Esser, 1994). Complement proteins that are bound to the bacterial surface also signal through complement
receptors on phagocytes, such as neutrophils, and increase the efficiency of phagocytosis (Erdei, Fust, & Gergely, 1991).

The adaptive immune response to *S. pyogenes* is largely characterised by the production of protective antibodies. Antibodies confer protection to extracellular bacteria via two main methods: opsonisation and neutralisation (Raynes et al., 2018). Opsonisation occurs when antibodies bind to their binding site (epitope) on an antigen and act as an immunological tag that stimulates other parts of the immune system. The Fc region or ‘tail’ of the antibody is recognised by receptors on phagocytes, including neutrophils, and binding to these receptors stimulates phagocytosis of the antibody-pathogen complex. Proteins of the complement system can also be recruited to the Fc region of antibodies and enhance phagocytosis or form the membrane-attack-complex as described above.

Antibody-mediated neutralisation occurs when the binding of an antibody to its epitope disrupts the function of the antigen. For example, if an antibody binds to an adhesin, it may inhibit bacterial adhesion (Loh, Lorenz, Tsai, Khemlani, & Proft, 2017), or, if an antibody binds to a secreted toxin, it may protect the host from the toxin’s effects (Bensi et al., 2012).

The development of an antibody-based response to *S. pyogenes* is associated with long-term immunity to the infecting strain (Lancefield, 1959). Over time, a robust memory response is thought to develop against antigens that are conserved across multiple *S. pyogenes* strains, and this provides broad, long-term protection in adults (Mortensen et al., 2015; Steer et al., 2016).

### 1.4 Streptococcus pyogenes virulence factors

In order to colonise and survive in the host while under attack from the immune system, *S. pyogenes* produce a vast array of virulence factors that aid in colonisation and interfere with host immune responses. Initial attachment is thought to be mediated by weak non-specific
interactions with host cells followed by specific attachments between bacterial adhesins and their specific targets. *S. pyogenes* produce a number of adhesins which bind to components of the extra-cellular matrix (ECM) and epithelial cells in the skin or oropharynx. These include the M protein, pili, and a variety of fibronectin and ECM binding proteins (Raynes et al., 2018; Walker et al., 2014). A vast range of secreted and surface-bound immune evasion factors then promote the survival of *S. pyogenes* in the host by interfering with multiple functions of the immune system (Figure 1.1). The recruitment of phagocytes to the site of infection is hindered by the proteases *S. pyogenes* cell envelope proteinase (SpyCEP) and C5a peptidase which cleave the chemoattractants Interleukin-8 (Edwards et al., 2005) and C5a respectively (Cleary, Prahbu, Dale, Wexler, & Handley, 1992). Phagocytic leukocytes such as macrophages and neutrophils are specifically lysed by the cytolysins Streptolysin S (SLS) (Miyoshi-Akiyama et al., 2005) and Streptolysin O (SLO) (Timmer et al., 2008) thereby inhibiting bacterial clearance. The proteases SpeB and IdeS cleave immunoglobulins (Åkesson, Moritz, Truedsson, Christensson, & von Pawel-Rammingen, 2006; Collin & Olsén, 2001) while the functions of antimicrobial peptides are inhibited by the binding of Streptococcal inhibitor of Complement (SIC) (Fernie-King, Seilly, & Lachmann, 2004). This non-exhaustive summary is simply to illustrate the vast range of immune functions that are disrupted, often with redundancy, by *S. pyogenes* virulence factors. The major virulence factors associated with this project and with *S. pyogenes* vaccinology are discussed in detail below.
S. pyogenes produce an array of secreted and surface-bound virulence factors that enable colonisation and interfere with the host immune system. Adhesins such as the M protein and pili bind to host proteins, promoting adhesion to cells and tissues. A range of immune evasion factors target multiple components of the immune system to inhibit immune function and support S. pyogenes survival. Based on (Walker et al., 2014).

1.4.1 The hyaluronic acid capsule

Many strains of S. pyogenes produce a capsule consisting of hyaluronic acid (HA). Capsular HA is an anionic polymer of N-acetylglucosamine and glucuronic acid and is identical to the HA found in the joints, skin, muscles and viscera of humans (Fraser, Laurent, & Laurent, 1997). Because of this molecular mimicry, the S. pyogenes capsule is considered to be a poor immunogen and strains that produce a capsule are resistant to phagocytosis (Fillit, McCarty, & Blake, 1986; Lancefield, 1928).

Indeed, acapsular mutants generated through directed mutation have repeatedly been shown to be rendered avirulent when compared to the encapsulated parent strains in both in vivo mouse...
models and *in vitro* bactericidal assays (Ashbaugh, Warren, Carey, & Wessels, 1998; Moses et al., 1997; Wessels, Goldberg, Moses, & DiCesare, 1994; Wessels, Moses, Goldberg, & DiCesare, 1991). Conversely, the expression of hyaluronic acid capsule in acapsular *S. pyogenes* strains enhances their resistance to phagocytosis (Moses et al., 1997). Encapsulated phenotypes have also been observed to arise *in vivo* through spontaneous mutations and these spontaneously encapsulated strains are more virulent than their wild type counterparts (Engleberg, Heath, Miller, Rivera, & DiRita, 2001). The increase in virulence seems to be due to the capsule conferring resistance to opsonophagocytosis although the presence of the capsule does not inhibit complement deposition or activation (Dale, Washburn, Marques, & Wessels, 1996). Recently, the hyaluronic acid capsule has been found to inhibit internalisation by macrophages, downregulate neutrophil production of both reactive oxygen species and neutrophil extracellular traps (NETs), and enhance bacterial survival within NETs (Cole et al., 2010; Schommer, Muto, Nizet, & Gallo, 2014; Secundino et al., 2016).

The capsule also plays a role in adhesion to epithelial cells in the skin and oropharynx, the two most common sites of *S. pyogenes* infection. The HA capsule binds to CD44 (a human HA-binding protein) expressed on the surface of epithelial cells (Cywes, Stamenkovic, & Wessels, 2000; Schrager, Albertí, Cywes, Dougherty, & Wessels, 1998; Wessels & Bronze, 1994). Acapsular mutants are unable to colonise the pharynx in murine models (although this may be due to increased immune clearance), and adhere less well to skin keratinocytes *in vitro* (Cywes et al., 2000; Schrager et al., 1998; Wessels & Bronze, 1994).

Furthermore, the *S. pyogenes* capsule has been shown to promote invasion through the epithelium in *in vitro* models of skin infection. Binding of the HA-capsule to CD44, expressed on keratinocyte monolayers, induces a signalling cascade within the keratinocytes that results
in rearrangement of the actin cytoskeleton, the opening of intercellular junctions, and invasion of \textit{S. pyogenes} into underlying tissue (Cywes & Wessels, 2001).

However, very high levels of capsule expression appear to inhibit the function of other adhesins expressed by \textit{S. pyogenes}. Hyper-encapsulated strains of \textit{S. pyogenes} show a decrease in their overall ability to interact with ECM components and adhere to epithelial cells (Darmstadt, Mentele, Podbielski, & Rubens, 2000; Hollands et al., 2010; Schrager et al., 1998). This is presumably due to the increased bulk of the capsule masking other \textit{S. pyogenes} adhesins and preventing their interactions with host receptors.

The expression of capsule is regulated by the CovR/S system which regulates approximately 15\% of the \textit{S. pyogenes} genome (Graham et al., 2002). The membrane-bound sensor kinase CovS phosphorylates CovR in response to external stimuli (Gryllos et al., 2007). CovR then binds to, and represses, the promoters of target genes including the hyaluronic acid synthesis operon (\textit{has}) that controls capsule synthesis (Heath, DiRita, Barg, & Engleberg, 1999). The RocA protein positively regulates the transcription of \textit{covR} and inactivating mutations in the \textit{rocA} gene have been shown to increase capsule production, particularly in \textit{S. pyogenes} of the M18 strain type (\textit{S. pyogenes} strain typing is discussed in section 1.4.3.2 and section 1.4.4.4) (Biswas & Scott, 2003; Lynskey et al., 2013).

### 1.4.2 Secretion and anchoring of surface expressed virulence factors to the cell-wall

\textit{S. pyogenes} secrete a large number of virulence factors to the cell wall where they are anchored through the activity of sortase enzymes. A signal peptide at the N-terminus of the unfolded polypeptide chain directs surface-expressed proteins to secretion pathways (such as the Sec system or Signal Recognition Particle pathway (SRP)) which translocate them across the bacterial membrane (Rosch, Vega, Beyer, Lin, & Caparon, 2008; Scott & Barnett, 2006). The
signal peptide is then removed by type I signal peptidases (van Roosmalen et al., 2004). Cell wall anchored proteins also possess a C-terminal cell wall sorting signal which contains a highly conserved LPXTG (where X is any amino acid) motif (Schneewind, Mihaylova-Petkov, & Model, 1993). The cell-wall anchored Sortase A enzyme cleaves this motif between the threonine and the glycine, removing the C-terminus of the secreted protein (Mazmanian, Liu, Ton-That, & Schneewind, 1999). Sortase A then catalyses the covalent linkage of the C-terminal threonine to lipid II (a peptidoglycan precursor in cell wall synthesis) resulting in incorporation of the secreted protein into the cell wall (Perry, Ton-That, Mazmanian, & Schneewind, 2002). Sortase A is regarded as a ‘housekeeping sortase’ and catalyses the anchoring of multiple LPXTG tagged proteins into the cell wall including the M protein and pili described in the following sections.

### 1.4.3 The M protein

The M protein is the major virulence factor of *S. pyogenes*. M proteins produced by different strains are highly variable, with individual M proteins exhibiting extensive diversity in structure and function. Mature M proteins form surface-expressed α-helical coiled-coil dimers. The first 50 amino acids at the N-terminus comprise the hypervariable region (HVR) which has low sequence identity between strains, and is relatively unstructured. However, recent findings suggest that there may be conserved patterns of amino acid sequence ‘hidden’ within the HVR that are important for the binding of host proteins by some M proteins (Ghosh, 2017). Outside of the HVR, the M protein consists of α-helical repeat regions which show increasing sequence conservation towards the C-terminus. These regions (A, B, C and D regions) feature a heptad repeat motif with hydrophobic residues at positions 1 and 4 forming the core of the coiled-coil and driving oligomerisation (McNamara et al., 2008). The crystal structure of the M protein from an M1 strain type indicates that the coiled-coil is not idealised with
destabilising residues in the B repeats disrupting the structure (McNamara et al., 2008; C. M. Stewart et al., 2016). These non-ideal coiled-coil regions appear to be necessary for interactions with human fibrinogen (C. M. Stewart et al., 2016). The A region and the B repeats in some strains have been implicated in the development of post S. pyogenes infection autoimmunity. Vaccination of animals with peptides from these regions results in the formation of lesions on the heart valves similar to those observed in human ARF and RHD (Faé et al., 2006; Kirvan, Galvin, Hilt, Kosanke, & Cunningham, 2014; Lymbury et al., 2003). The number and size of the repeat regions differs significantly between the M proteins expressed by different strains (Figure 1.2). The highly conserved C-terminus contains an LPXTG sorting motif and is anchored to the peptidoglycan cell wall by Sortase A.

![Model of the dimeric M protein from M1T1 S. pyogenes](PDB ID: 2OTO)

Figure 1.2 Model of the dimeric M protein from M1T1 S. pyogenes with the partial structure of the A and B regions shown in ribbon form. The sequence diversity of each region is shown as a gradient from dark blue (highly variable between strains) to light blue (conserved between strains). The repeat regions are represented as black bars. The locations of the peptides that the current M protein based vaccine candidates target are shown.
1.4.3.1 The M protein as a virulence factor

As a group, M proteins bind to an array of human proteins that enable adhesion, invasion through the mucosa, and immune evasion. The M protein has been shown to be necessary for adherence to HaCat keratinocytes (Perez-Casal, Okada, Caparon, & Scott, 1995) which is mediated by the binding of M protein to epithelial CD46 (a regulator of the complement system) (Okada, Liszewski, Atkinson, & Caparon, 1995). M protein expression is also necessary for colonisation of the pharynx in rat and primate models of pharyngitis (Ashbaugh et al., 2000; Hollingshead, Simecka, & Michalek, 1993). The binding of collagen IV by some M proteins has been associated with increased adherence to the ECM and epithelial basement membrane, as well as protection from phagocytosis (Dinkla et al., 2003; Dinkla et al., 2009). The M1 protein (M protein expressed by M1 strain types) has also been shown to promote intracellular invasion of epithelial cells by binding to laminin and fibronectin (Cue, Dombek, Lam, & Cleary, 1998; Cue, Lam, & Cleary, 2001; Dombek et al., 1999). In the case of fibronectin, the B repeats of M1 bind to fibronectin which itself binds to α5β1 integrins on the host cell. The binding of the M1-fibronectin complex to α5β1 integrins initiates a signalling cascade that promotes the internalisation of the S. pyogenes-fibronectin-integrin complex (Wang, Li, Dedhar, & Cleary, 2007). This is advantageous as it provides protection from some aspects of the host immune system such as antibody- and complement-mediated opsonophagocytosis, as well as antibiotic therapy (Medina, Goldmann, Toppel, & Chhatwal, 2003).

M proteins also promote immune evasion by binding to regulatory elements of the complement system and inhibiting antibody signalling. Factor H (Fischetti, Horstmann, & Pancholi, 1995; Gustafsson et al., 2013), factor H-like protein 1 (Johnsson et al., 1998) and complement protein C4 binding protein (C4BP) (Buffalo et al., 2016; Johnsson et al., 1996; Morfeldt et al., 2001) are all human proteins that negatively regulate complement activity. They induce the
inactivation or decay of complement proteins, which protects human cells from being damaged by complement activity. When captured by M proteins, they retain their physiological functions and inactivate complement proteins that deposit on the *S. pyogenes* surface. This, in turn, reduces complement-mediated lysis and opsonophagocytosis of *S. pyogenes*. M proteins also bind fibrinogen (Horstmann, Sievertsen, Leippe, & Fischetti, 1992), fibronectin (Cue et al., 2001) and serum albumin (Retnoningrum & Cleary, 1994) which may all confer passive resistance to complement deposition and phagocytosis by masking the bacterial surface. M proteins from some strains are also capable of preventing antibody-phagocyte signalling by binding to the Fc regions of IgG and IgA (Carlsson, Berggård, Stålhammar-Carlemalm, & Lindahl, 2003; Stenberg, O'Toole, & Lindahl, 1992) which, in turn, inhibits phagocytosis.

1.4.3.2 The M protein as a molecular marker for strain typing

*S. pyogenes* were traditionally strain-typed based on a serological reaction between sample extracts and standardised type-specific antisera (‘M typing’) (Lancefield, 1928). M typing sera was generated by vaccinating animals with killed *S. pyogenes* from known strains. The resulting immune sera were then adsorbed with heterologous strains to remove cross-reactive antibodies to other antigens, including heterologous M proteins and the Group A carbohydrate. The resulting sera were generally type-specific and had bactericidal activity against the homologous strain (Lancefield, 1928).

With the advent of molecular biology, and the recognition that the HVR contains the type-specific moiety recognised by M typing sera, sequencing of the HVR region of the *emm* gene (encoding the M protein; ‘*emm* typing’) has superseded serological M typing (Beachey, Seyer, Dale, Simpson, & Kang, 1981; Lancefield, 1962). Generally, strains within an *emm* type are defined as having >95% nucleotide sequence identity over the first 160bp of the *emm* gene (Facklam et al., 1999). Due to the high antigenic diversity of the HVR, this has resulted in the
identification of more than 200 unique emm types (Steer et al., 2016). *S. pyogenes* can also be classified by emm-pattern type. This groups strains by the chromosomal arrangement of the emm and emm-like genes. An isolate of *S. pyogenes* has between one and three emm and emm-like genes clustered on the chromosome and three major emm-patterns have been identified that are statistically associated with tissue tropism. A-C pattern strains tend to be associated with pharyngeal infections, D pattern strains tend to be associated with skin infections, while E pattern strains are associated with both and considered to be ‘generalists’ (Bessen, Fiorentino, & Hollingshead, 1997; McMillan et al., 2013).

Recently, an emm-cluster typing system has been proposed that organises *S. pyogenes* into phylogenetic clusters based on the sequence of the full-length M protein. This results in the >200 emm types grouping into 48 emm-clusters that each have similar structural and functional characteristics (Sanderson-Smith et al., 2014).

### 1.4.4 The *Streptococcus pyogenes* Pilus

#### 1.4.4.1 Pilus architecture and assembly

*S. pyogenes* pili are surface-expressed multiprotein filaments that are produced by the covalent polymerisation of subunit proteins (pilins) (Mora et al., 2005; Telford, Barocchi, Margarit, Rappuoli, & Grandi, 2006). They form a single chain usually comprised of a basal ancillary protein (AP2), the pilus shaft (made up of repeated T antigen monomers) and an apical ancillary protein (AP1) (Figure 1.3). If present, AP2 contains an LPXTG motif and is anchored to the cell wall by Sortase A (Linke et al., 2010). This covalently attaches the entire pilus structure to the cell wall. Some pili lack AP2 and in these strains, the pili are anchored to the cell wall by the C-domain of the most basal T antigen in the pilus instead (Nakata et al., 2011; Young, Moreland, et al., 2014). The pilus shaft is elongated by the formation of intermolecular isopeptide bonds between pilin subunits. The pilus-associated Sortase C catalyses the
formation of these covalent bonds in a similar manner to Sortase A (section 1.4.2). Sortase C cleaves between the threonine and glycine in LPXTG-like motifs found in the C-domains of T antigens. The new C-terminal threonine is then covalently linked to the side-chain of a conserved lysine (the ‘pilin lysine’) located in the N domain of the next T antigen in the pilus (Kang, Coulibaly, Clow, Proft, & Baker, 2007). AP1 forms the tip of the pilus (Quigley et al., 2010) and is also covalently linked to the shaft with an intermolecular isopeptide bond between the C-domain of AP1 and the N-domain of the following T antigen (Proft & Baker, 2009).

Figure 1.3 The structure of the S. pyogenes pili and T antigens. (A) Model of the pilus from M1T1 S. pyogenes shown in ribbon form. This is comprised of AP1 (PDB ID: 2XIC), AP2 (PDB ID: 3KLQ) and repeating units of the T1 T antigen (PDB ID: 3B2M). (B) The structure of T6 (PDB ID: 4P0D) shown in ribbon form (not to scale with (A)). The intramolecular isopeptide bonds are shown in orange. N = N-domain, M = Middle-domain, C = C-domain.
The pilus is thought to be assembled from the tip downwards. In this model, AP1 is secreted first and anchored to the cell wall. T antigen subunits are then added to the base by Sortase C until the pilus fibre is anchored to the cell wall via the AP2 subunit. The resulting pili can be up to 3 μm long but are only 2-3 nm thick (Figure 1.3) (Proft & Baker, 2009).

1.4.4.2 Structure of the T antigen

The T antigen is named for its trypsin-resistant properties. It was first isolated from streptococcal extracts and characterised as a serologically-active, variable antigen (Lancefield & Dole, 1946), but was not identified as being the protein component of the pilus shaft until 2005 (Mora et al.). Repeating monomers of the T antigen are polymerised to form the shaft of the pilus which can contain tens or even hundreds of T antigen monomers. The first crystal structure of a T antigen (T1, expressed by M1/T1 strains) was solved by x-ray crystallography in 2007 (Figure 1.3A) (Kang et al., 2007). T1 has two all-β domains with a modified immunoglobulin-fold. The two domains associate closely, with only one residue separating the N-domain from the C-domain. The interface between domains mostly consists of hydrophobic residues and buries approximately 1200 Å² of surface area. Successive T antigen monomers are predicted to be joined head-to-tail (by Sortase C) with a 120° rotation along the long axis. The interface between the N-domain of one T1 monomer and the C-domain of the next buries approximately 850 Å² of solvent-accessible surface (Kang et al., 2007).

Within the hydrophobic core of each domain is an auto-catalytic isopeptide bond that cross-links the side-chains of lysine and asparagine residues found on the first and last β-strand of each domain (Kang et al., 2007). These isopeptide bonds are vital for the correct folding of T antigens. Mutation of the surrounding catalytic residues prevents the formation of the isopeptide bond, and significantly decreases the proteolytic and thermal stability of T1 (Kang & Baker, 2009; Kang et al., 2007). The intramolecular isopeptide bonds also confer resistance
to mechanical extension, and this resistance is lost in mutants that cannot form these bonds (Alegre-Cebollada, Badilla, & Fernández, 2010). In this way, the isopeptide bonds stabilise the structure of individual T antigens which, owing to the entirety of the pilus shaft being formed of T antigens, is thought to confer strength to the pilus as a whole. Possessing a pilus that is mechanically strong and resistant to proteolysis may help colonising S. pyogenes adhere to the mucosa and remain bound in a harsh host environment.

The structure of T6 (expressed by M6/T6 strains) has also been solved by x-ray crystallography (Figure 1.3B). Overall the structures of T1 and T6 are dissimilar but they do share some conserved structural features. T6 is much larger and consists of three tandem domains which, like T1, possess variants of the immunoglobulin-fold (Young, Moreland, et al., 2014). Autocatalytic intramolecular isopeptide bonds are also found in the M-domain and C-domain of T6 but are not found in the N-domain. Unlike T1, the cores of the central three domains are highly decorated with loops and extensions.

While T1 and T6 share conserved features such as their extended structure of immunoglobulin-like domains and the presence of intradomain isopeptide bonds, they also exhibit significant structural diversity. This likely results from them being evolutionarily distinct from both each other, and all other T antigens (Figure 1.4). This suggests that they are not necessarily structurally representative of all T antigens.

1.4.4.3 The pilus as a virulence factor

S. pyogenes pili have important roles in adhesion and colonisation. The pilus from M1/T1 strains has been shown to facilitate adherence to tonsil epithelia (Manetti et al., 2007) and keratinocytes (Abbot et al., 2007), which represent cells from the two tissues most commonly colonised by S. pyogenes. This adherence is mediated through AP1 (Smith et al., 2010) which contains two thioester bonds that are involved in the formation of covalent bonds with host
cells (Linke-Winnebeck et al., 2013; Pointon et al., 2010). The same pilus has also been shown to bind to pharyngeal cells in vitro, and have a role in biofilm production (Manetti et al., 2007). The pilus from M6/T6 strains has also been associated with biofilm production (Kimura et al., 2011). Biofilms surrounding S. pyogenes have been observed in patients with impetigo (Akiyama, Morizane, Yamasaki, Oono, & Iwatsuki, 2003) and in vitro (Lembke et al., 2006). They can protect S. pyogenes from antibiotics and immune surveillance, and have been suggested as a cause of therapeutic failure (Conley et al., 2003). All AP1 proteins have been shown to bind to collagen I which promotes adhesion to the ECM (Kreikemeyer et al., 2005). Interestingly, it appears that the shaft of the pilus may function as an adhesin, at least in the case of M2/T2 strains, mediating binding to epithelial cells and serum proteins (Tsai, Loh, Clow, Lorenz, & Proft, 2017). However, expression of the pilus reduces virulence in some models of invasive disease (Nakata et al., 2009), seemingly through increasing neutrophil chemotaxis and NET production (Alexander et al., 2010). It has been suggested that its major role is in adherence and colonisation and that long-term expression of pili in invasive disease may be detrimental to bacterial survival (Alexander et al., 2010; Nakata et al., 2009).

1.4.4.4 The T antigen as a molecular marker for strain typing

In S. pyogenes, the pilus protein genes are found clustered within the highly variable Fibronectin-binding, Collagen-binding, T antigen (FCT) pathogenicity island. All S. pyogenes strains carry one of nine different FCT regions, and the precise complement of genes differs between FCT types (Falugi et al., 2008). The adhesins encoded by individual FCT regions are thought to partially determine a strain’s tissue tropism (Kratovac, Manoharan, Luo, Lizano, & Bessen, 2007). Generally, FCT regions encode the T antigen, AP1 (the collagen-binding protein), AP2, multiple fibronectin-binding proteins, and any other pilus associated enzymes that are required for assembly in that strain (such as Sortase C and SipA) (Falugi et al., 2008; Young, Proft, Harris, Brimble, & Baker, 2014; Zähner & Scott, 2008). The FCT 3, FCT 4, FCT
7 and FCT 8 types are most similar to each other (and the most common), while FCT1, FCT 2, FCT 5, FCT 6 and FCT 9 are genetically distinct (Figure 1.4) (Falugi et al., 2008).

Historically the T antigen (then an unidentified protein) was used in a similar manner as the M protein to serologically strain type S. pyogenes. The T antigen was known to be immunogenic and exhibit strain-variability, so agglutination reactions between sample extracts and standardised typing antisera were used to ‘T type’ strains (Lancefield & Dole, 1946). Once the T antigen was identified as the backbone protein of the pilus, T typing sera were confirmed to be largely type-specific for recombinant T antigens. Some T typing sera were also found to cross-react to multiple recombinant T antigens, such as T3 and T13, and these cross-reactivity patterns correlated well with prior T typing observations (Falugi et al., 2008; Johnson, Kaplan, VanGheem, Facklam, & Beall, 2006). T typing groups S. pyogenes strains into 21 T serotypes.
In the same way that *emm* typing has superseded M typing, a new genotyping system has been proposed using the genetic variability of the *tee* gene (which encodes the T antigen) to *tee* type *S. pyogenes* (Falugi et al., 2008). Sequencing of the *tee* gene correlates well with T serotyping, and has identified eighteen major *tee* types and three subtypes (Falugi et al., 2008; Steemson et al., 2014). Strains within a *tee* type are defined as having >97% sequence identity. There is a strong association between *emm* type and *tee* type, with *S. pyogenes* strains belonging to a given *emm* type usually, though not exclusively, carrying the same *tee* gene. In this way, over 200 *emm* types cluster into only 18 *tee* types (Steemson et al., 2014). Though these initial *tee* typing studies were small scale, they appear to have captured the majority of the variation within the *tee* gene. A recent study sequenced the whole genomes of over 1400 *S. pyogenes* isolates and identified 21 different *tee* types (Chochua et al., 2017). However, the *tee* typing annotation used in this study differed from the previously published systems (Falugi et al., 2008; Steemson et al., 2014).

In this thesis, *S. pyogenes* strains will be referred to by their strain name (for example MGAS8232, M18HS66 or *emm*217_007) and/or by their expression of the M protein and T antigen (for example M1/T1, M217/T18.1, M49/T18.2). Where *tee* typing notation is referred to, it will follow the original designation (Falugi et al., 2008; Steemson et al., 2014).

### 1.5 *Streptococcus pyogenes* vaccinology

Although *S. pyogenes* remains exquisitely sensitive to penicillin, antibiotic therapy is a resource intensive strategy for treatment and prevention that has failed to control severe disease at the population level (Bisno, Rubin, Cleary, & Dale, 2005; N. J. Moreland et al., 2014). This is evident by the increasing incidence and severity of invasive disease around the globe (Ralph & Carapetis, 2012), as well as the failure to prevent ARF and RHD in at-risk groups in developed countries. As *S. pyogenes* disease is most prevalent in communities with limited
access to healthcare, timely treatment and adherence to antibiotic regimens are significant issues. There is a clear need for a more effective method to control \textit{S. pyogenes} disease and prophylactic vaccination could be a practical strategy. This has been endorsed by the World Health Organization which prioritised the development of \textit{S. pyogenes} vaccines in 2014 (Sheel, Moreland, Fraser, & Carapetis, 2016). There are now a large number of \textit{S. pyogenes} vaccine candidates in preclinical development and early stage clinical trials (Dale et al., 2016; Raynes et al., 2018; Sheel et al., 2016; Steer et al., 2016). This section will discuss some of the major obstacles encountered during efforts to develop a vaccine for \textit{S. pyogenes}, and then discuss the M protein (currently the leading candidate) and the T antigen (as the subject of this thesis) as vaccine candidates.

\subsection*{1.5.1 Obstacles to the development of a vaccine for \textit{Streptococcus pyogenes}}

\textit{S. pyogenes} vaccine development has been ongoing for almost a century, with the first human trial (vaccinating patients with heat-inactivated \textit{S. pyogenes}) occurring in 1923 (Bloomfield & Felty, 1923). Throughout this period, progress has been made but significant obstacles have also been encountered. Most importantly, there have been safety concerns regarding the theoretical potential for a vaccine containing \textit{S. pyogenes} antigens to cause autoimmunity. In 1969, one vaccine trial (using large quantities of crudely purified M3 protein) was linked to the development of ARF in three of the volunteers (Massell, Honikman, & Amezcua, 1969). These volunteers were all siblings of patients who had been diagnosed with ARF. At the time it was thought, due to the lack of understanding of how \textit{S. pyogenes} infection induces autoimmunity, vaccination with streptococcal products could have caused ARF. Following this, the FDA banned the use of all \textit{S. pyogenes} products, including vaccines, in humans for nearly 30 years (Steer, Batzloff, Mulholland, & Carapetis, 2009). Since then, doubt has been cast on the association between the vaccine and the development of ARF. Numerous \textit{S. pyogenes} vaccine trials performed prior to Massell \textit{et al.} (1969) had not been associated with the development of
autoimmunity, and it is thought that including the siblings of ARF patients (who are pre-disposed to develop autoimmunity in some uncharacterised way) may have confounded the vaccine trial (Steer, Batzloff, et al., 2009). Furthermore, three M protein-based vaccines have now been shown to be safe in clinical trials (Good, Pandey, Batzloff, & Tyrrell, 2015; Kotloff et al., 2004; McNeil et al., 2005) which has somewhat alleviated the safety concerns. Also, the M protein-based vaccines that are currently in development (described below) only use small peptides of M proteins from the HVR and C repeats. This avoids the use of sequence from the B repeats that have been implicated in molecular mimicry and the development of autoimmunity (Faé et al., 2006; Kirvan et al., 2014; Lymbury et al., 2003).

A more technical obstacle in vaccine development is the lack of understanding of *S. pyogenes* correlates of protection. These are immunological markers that correlate with protective immunity that can be measured *in vitro*, such as antibody titer (Tsoi, Smeesters, Frost, Licciardi, & Steer, 2015). Correlates of protection provide an assay-based means of assessing the efficacy of vaccine candidates, as well as monitoring them post-licensure (preferably in a high-throughput manner). For example, an individual vaccinated with a *Streptococcus pneumoniae* conjugate vaccine who develops an antibody titer of >0.2 μg/mL (measured by ELISA), or has an opsonic titer >8 (measured in an opsonophagocytic killing assay with live bacteria) is considered to be protected from invasive disease (Jódar et al., 2003; Romero-Steiner et al., 2006). This provides a clear benchmark for efficacy studies where if a vaccine does not induce antibody titers of >0.2 μg/mL then it is not protective.

The identification of correlates of protection requires the use of robust and standardised immunoassays which have not yet been established for *S. pyogenes*. Immunity to *S. pyogenes* is dependent on both complement and antibody-mediated opsonophagocytosis (where complement proteins and protective antibodies are able to bind to *S. pyogenes* and increase the
efficiency of phagocytosis). However, ELISAs that measure antibody titers simply measure the total amount of antibodies that bind to an antigen, and do not necessarily measure functional antibodies or protective activity (Lorenz, Loh, Moreland, & Proft, 2017; Tsoi et al., 2015). For this reason, the current gold-standard for assessing protective antibody responses are bactericidal assays developed by Rebecca Lancefield and colleagues over 50 years ago. The *S. pyogenes* bactericidal assay has two major variations. The direct bactericidal assay measures the survival of bacteria in whole human blood from an immune individual. In this case, the immune blood is the source of test serum, phagocytes, and complement proteins (Lancefield, 1957). The indirect bactericidal assay measures the survival of bacteria in non-immune whole blood which is supplemented with test serum from an immune individual or animal. In this case, the donor non-immune blood is the source of phagocytes and complement proteins (Johnson et al., 1996). However, these assays have some widely accepted limitations. They are low-throughput, labour-intensive and the variation in phagocyte and complement activity between individuals (both immune and non-immune donors) makes the results of these assays difficult to reproduce and compare between laboratories (Sheel et al., 2016; Tsoi et al., 2015). Additionally, natural exposure to *S. pyogenes* can result in blood donors having endogenous immunity to the test strains. This necessitates the pre-screening of donors for immunity to each strain being used in the bactericidal assay (Reglinski, Lyncskey, & Sriskandan, 2016; Tsoi et al., 2015). Recently, efforts have been made to improve *S. pyogenes* bactericidal assays in a number of ways. Individually these modifications have addressed specific issues, such as increasing throughput (Lorenz et al., 2017) or reducing inter-donor variability (Reglinski, Lyncskey, & Sriskandan, 2016), but they have also resulted in a greater number of methods being used which further complicates inter-laboratory comparisons.

A new opsonophagocytic killing assay (OPKA) has been developed with the goal of overcoming all of these issues (Jones et al., 2018). The OPKA measures the reduction in the
number of viable bacteria in the presence of phagocytes, antibodies, and complement (Figure 1.5). This is a high-throughput assay that has been adapted from a validated OPKA that is routinely used in *S. pneumoniae* vaccine trials. The OPKA uses commercial baby rabbit complement (BRC) as a source of complement and dimethylformamide (DMF) -differentiated human promyelocytic leukemia cells (HL-60) as the source of phagocytes. This significantly reduces inter-assay variability. The OPKA has been used to characterise the killing of several clinically relevant *S. pyogenes* strains using M protein antisera. Unlike bactericidal assays, the specificity, precision, relative accuracy and linearity of the killing can all be readily quantified. This OPKA has the potential to become the standardised assay for the measurement of functional antibodies to *S. pyogenes* (Jones et al., 2018).

Figure 1.5 Schematic illustrating the *S. pyogenes* OPKA. The testing sera (or antibody) is incubated with *S. pyogenes* to allow the formation of antibody-antigen complexes. Commercial baby rabbit complement and differentiated HL-60 cells are added and incubated for a length of time that is dependent on the strain of *S. pyogenes* in the assay. Samples from the assay are plated and counted to determine killing.
1.5.2 The M protein as a vaccine candidate

The most advanced *S. pyogenes* vaccine candidates are subunit vaccines containing elements taken from M proteins (Batzloff et al., 2003; Dale, Penfound, Chiang, & Walton, 2011; Guilherme et al., 2006). They fall into two categories: type-specific vaccines which aim to elicit the production of strain-specific protective antibodies against the HVR, and conserved vaccines which aim to elicit the production of cross-protective antibodies to the C repeat regions (Figure 1.2). Both of these strategies ensure that elements from the B repeats (that have been implicated in molecular mimicry) are not included in the vaccine. These vaccine candidates have been extensively reviewed in the last few years, which reflects the activity in the *S. pyogenes* vaccine development field (Dale et al., 2016; Raynes et al., 2018; Sheel et al., 2016; Steer et al., 2016).

1.5.2.1 Type-specific vaccines

The most advanced candidate in the development pipeline is the type-specific ‘30-valent’ vaccine. It consists of four fusion proteins made-up of the HVRs from thirty strains that commonly cause disease in North America and Europe (Dale et al., 2011). It is the latest iteration of a multivalent HVR-based vaccine, being a successor to the ‘6-valent’ (Dale, 1999) and ‘26-valent’ (McNeil et al., 2005) vaccines. These vaccines have entered early phase clinical trials and have proven to be both immunogenic and safe (Kotloff et al., 2004; McNeil et al., 2005). Antibodies to the HVR have been shown to be protective in animal models (Dale, 1999; Dale et al., 2011; Penfound, Chiang, Ahmed, & Dale, 2010), but these may only develop rarely in humans during natural infection (Lannergård et al., 2011a; Lannergård et al., 2011b). Antibodies to the HVR were traditionally thought to only confer type-specific protection. This raised concerns regarding the global coverage of an HVR-based vaccine that had been designed to protect against strains circulating in the developed world (Steer, Law, Matatolu, Beall, &
Carapetis, 2009). Studies have shown that the strain diversity in these locations is generally much lower than in the developing world, and in high-risk children in developed countries - such as Maori and Pacific Island children in New Zealand (Steer, Magor, et al., 2009; Williamson et al., 2014; Williamson et al., 2015). This difference in strain diversity between populations has a significant effect on the predicted coverage provided by the 30-valent vaccine. The predicted coverage for pharyngitis-causing strains in low-risk children in New Zealand is 93%. However, in high-risk children it is only 48%. Vaccine coverage of S. pyogenes recovered from the skin of high-risk children is even lower at 34%. This suggested that the 30-valent vaccine would have limited coverage outside its target market. However, the 30-valent vaccine was also found to elicit the production of cross-reactive antibodies with some bactericidal activity against non-vaccine strains (Dale et al., 2011). This cross-reactivity occurs mostly, but not exclusively, between vaccine HVRs and non-vaccine HVRs that are within the same emm-cluster (Sanderson-Smith et al., 2014). While this has increased its predicted coverage (Engel et al., 2014; Williamson et al., 2015), it is likely that the vaccine would still require re-formulation to achieve global coverage of circulating strains (Williamson et al., 2015).

The potential cross-reactivity within emm-clusters could be exploited further. A recent structural vaccinology study has shown that vaccination with five rationally selected HVRs from a single emm-cluster evoked bactericidal responses against fifteen out of the seventeen strains within that cluster (Dale et al., 2017). This may provide a promising new angle for the design of a multivalent vaccine with cluster-based coverage.

1.5.2.2 Conserved region M protein vaccines

The alternative approach to developing an M protein based vaccine has been to use peptides found within the highly conserved C repeat regions of the M protein. While the peptides are
not entirely conserved among all strains, these vaccines have the potential to confer much broader protection than the type-specific vaccines (Steer, Law, et al., 2009).

The ‘J8’ vaccine is a short peptide containing a single B-cell epitope (as well as flanking residues which provide alpha helical structure). Vaccination with J8 and SpyCEP has been shown to elicit long-term antibody-mediated protection in mouse models (Batzloff et al., 2003; Batzloff, Yan, Davies, Hartas, & Good, 2004; Pandey, Mortensen, et al., 2016; Pandey, Wykes, Hartas, Good, & Batzloff, 2013). It has recently undergone a phase I clinical trial and found to be safe and immunogenic in humans (Sekuloski et al., 2018).

The SV1 vaccine is also based on C repeat peptides and combines five variants of the J14 peptide (some of which include the J8 sequence) to provide a predicted coverage of 97% of all S. pyogenes strains (McNeilly et al., 2016).

The StreptInCor vaccine contains a 55-residue sequence that incorporates both J8 and J14 sequences along with characterised B cell and T cell epitopes (Guilherme et al., 2010; Guilherme et al., 2006). Vaccination with StreptInCor elicits systemic IgG and IgA production in mice (Guilherme et al., 2009) and protects against intraperitoneal challenge (Postol et al., 2013). StreptInCor induced antibodies have been shown to be bactericidal in vitro against a small number of strains (De Amicis et al., 2014).

Although these conserved region vaccines are in the early stages of development, preclinical and phase I clinical trial results have been promising. However, questions regarding the efficacy of C repeat vaccines in humans remain, with some studies indicating that antibodies may not be able to bind to the C repeat region of some strains in human blood (McArthur & Walker, 2005; Penfound et al., 2010; Sandin, Carlsson, & Lindahl, 2006). This is hypothesised to be due to human albumin competing for binding to the same region.
1.5.3 The T antigen as a vaccine candidate

Since their discovery in 2005, *S. pyogenes* pili have been recognised as potential vaccine candidates. Mora *et al.* vaccinated mice with a mixture of the three pilus-forming proteins from the T1 pilus (recombinant T1, AP1 and AP2), and showed that it conferred a similar level of protection against intranasal challenge with M1/T1 *S. pyogenes* as an M1-based vaccine (73% and 82% respectively) (Mora *et al*., 2005). Vaccination with whole pili from M18/T18 and M28/T28 strains has also been shown to induce the production of antibodies that protect mice from intranasal challenge, partially neutralise adherence to keratinocytes and have *in vitro* bactericidal activity (Loh *et al*., 2017). The authors noted that the highest antibody titers were generated against the T antigen component of the vaccines, and that cross-reactivity patterns within their sera suggested that antibodies directed against the T antigen and not the AP proteins were likely to be bactericidal (Loh *et al*., 2017).

The expression and immunogenicity of T antigens during natural infection in humans has been confirmed indirectly. Sera from patients diagnosed with *S. pyogenes* pharyngitis (Manetti *et al*., 2007) and ARF (Young, Moreland, *et al*., 2014) have demonstrated robust antibody responses to recombinant T antigens.

Due to the repeating nature of the *S. pyogenes* pilus, T antigens are present tens or even hundreds of times per pilus structure. This makes them attractive vaccine candidates as any T antigen epitope will be presented to the immune system numerous times (Raynes *et al*., 2018). However, T antigens do exhibit a degree of antigenic diversity, with strains carrying eighteen *tee* types and three sub-types (which cluster into 9 FCT types) currently circulating (Figure 1.4) (Falugi *et al*., 2008; Steemson *et al*., 2014). While significant, this is markedly less diversity than exhibited by M proteins (which have >200 *emm* types and 48 *emm*-clusters). It has been proposed that a multivalent vaccine containing elements from only 12 T antigens
would provide 90% coverage (Falugi et al., 2008) while a vaccine containing 18 T antigens would provide near complete coverage (Steemson et al., 2014).

Cross-protection between closely related T antigens, such as those encoded in FCT3, FCT 4, FCT 7 and FCT 8, could further reduce this number. The other major advantage of a T antigen-based vaccine is that they have never been associated with the development of autoimmunity, so are not perceived to have the same safety concerns as M protein-vaccines.

### 1.5.4 Multicomponent vaccine candidates

Vaccines targeting single virulence factors (such as M proteins and T antigens) have thus far been the major focus of *S. pyogenes* vaccine development. However, vaccines that contain multiple antigens that are conserved across a wide number of strains are also being explored. Vaccines based on conserved antigens have the potential to provide improved efficacy and coverage. So far, two different approaches have been taken to identify antigens that are immunogenic, conserved and protective against a wide range of strains. The first approach used human Intravenous Immunoglobulin (IVIG) to identify the surface-expressed proteins targeted by the healthy adult population (Reglinski, Gierula, Lynskey, Edwards, & Sriskandan, 2015; Reglinski, Lynskey, Choi, Edwards, & Sriskandan, 2016). This process identified ten immunogenic proteins that were conserved across all genome-sequenced *S. pyogenes* strains. Vaccination with a combination of seven of these antigens (Spy7, Table 1.1) was shown to induce B cell and T cell responses, reduce bacterial dissemination and reduce the overall severity of disease in mouse models (Reglinski, Lynskey, Choi, et al., 2016).
Table 1.1 Proteins in the SPy7 vaccine.

<table>
<thead>
<tr>
<th>UniProt ID</th>
<th>Protein Product</th>
<th>Protein function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>C5AP_STRP1</td>
<td>C5a peptidase</td>
<td>Cleave chemoattractant (IL-8) (Edwards et al., 2005)</td>
<td></td>
</tr>
<tr>
<td>Q490V0_STRP1</td>
<td>Oligopeptide-binding protein</td>
<td>Peptide acquisition</td>
<td>(Reglinski et al., 2015)</td>
</tr>
<tr>
<td>Q491G2_STRP1</td>
<td>Nucleoside-binding protein</td>
<td>Nucleoside-binding</td>
<td>(Reglinski et al., 2015)</td>
</tr>
<tr>
<td>Q99XX8_STRP1</td>
<td>Putative pullulanase</td>
<td>Carbohydrate metabolism (Putative)</td>
<td>-</td>
</tr>
<tr>
<td>Q99ZW9_STRP1</td>
<td>Hypothetical membrane associated protein</td>
<td>Unknown</td>
<td>-</td>
</tr>
<tr>
<td>Q9A0C0_STRP1</td>
<td>Cell surface protein</td>
<td>Unknown</td>
<td>-</td>
</tr>
<tr>
<td>Q9A1H3_STRP1</td>
<td>SpyAD</td>
<td>Adhesion and cell division</td>
<td>(Gallotta et al., 2014)</td>
</tr>
</tbody>
</table>

The second approach used reverse vaccinology methodology where *S. pyogenes* genomes were analysed to identify proteins that were conserved across all genomes and predicted to be secreted or surface-expressed (Bensi et al., 2012). These proteins were then recombinantly expressed and screened for immunogenicity, expression level, and expression at the cell surface. These combined methods identified six protective antigens that met the criteria. Three of these: SpyCEP (Cleaves the chemoattractant IL-8), SPy0269 (‘putative surface exclusion protein’) and SLO (a cytolysin) were combined into the ‘Combo’ vaccine. Vaccination with Combo conferred cross-protection against four *S. pyogenes* strains in two mouse models of infection. This protection was comparable to the protection provided by vaccination with homologous M proteins (Bensi et al., 2012).

Perhaps the greatest advantage of these multicomponent vaccines is that they target conserved proteins which confer virulence in a number of ways. The proteins incorporated in the Spy7 vaccine have functions ranging from enabling *S. pyogenes* cell division to cleaving human IL-8 and inhibiting neutrophil chemotaxis (Table 1.1). The Combo vaccine incorporates two well
characterised virulence factors (in addition to highly conserved protein with unknown function). Thus, vaccination with a combination of antigens may result in a broader protective response that combats \textit{S. pyogenes} survival and virulence through multiple mechanisms rather than targeting a single antigen (Reglinski, Lynskey, Choi, et al., 2016).

1.5.5 \textit{Lactococcus lactis} as a vaccine delivery system

The two primary routes of \textit{S. pyogenes} infection are through the skin and through the mucosa. As such, it has been suggested that a vaccine that can induce the development of a protective mucosal immune response (characterised by the induction of IgA), could have the benefit of inhibiting colonisation of the oropharynx and preventing transmission of \textit{S. pyogenes} between hosts (Batzloff et al., 2005; D'Alessandri et al., 1978). The bacterium \textit{Lactococcus lactis} has previously been shown to be able to express the full pilus structures from Group B Streptococcus (GBS) and \textit{S. pyogenes}. When \textit{L. lactis} expressing these pili were used as intranasal vaccines in small animals (which mimics colonisation of the oropharynx in natural infection), they conferred protection against challenge with GBS and \textit{S. pyogenes} strains expressing the homologous pili, as well as inducing the production of bactericidal antibodies. (Buccato et al., 2006; Loh et al., 2017).

1.6 Antibodies and Phage Display

Antibodies are tetrameric glycoproteins made up of two heavy chains and two light chains. In humans, rabbits and mice the light chains can be either kappa or lambda light chains. In all three species kappa light chains are more commonly used (Haughton, Lanier, & Babcock, 1978; Molé, Béné, Montagne, Seilles, & Faure, 1994; Popkov et al., 2003). The constant regions of the heavy chains determine the class of the antibody with five classes existing in humans: IgA, IgD, IgE, IgG and IgM. These all have different overall structures and markedly different roles in immunity. The most common class in serum is IgG (Brüggemann et al., 1987)
which is further divided into four subclasses: IgG1, IgG2, IgG3, and IgG4. These all have different effector functions. For example IgG3 is the strongest activator of the complement cascade while IgG4 cannot bind to complement proteins (Vidarsson, Dekkers, & Rispens, 2014). This thesis will focus on IgG1 antibodies. In humans, these are the most abundant subclass in serum, comprising up to 60% of total IgG. IgG1 are both opsonic and potent activators of the complement system (Vidarsson et al., 2014). Both IgG1 and IgG3 have been associated with long-term protective immunity against *S. pyogenes* (Mortensen et al., 2015).

IgG1 are typically depicted as being Y shaped molecules (Figure 1.6). The heavy chain consists of one variable domain (V<sub>H</sub>) and three constant domains (C<sub>H1</sub>, C<sub>H2</sub> and C<sub>H3</sub>) while the light chain consists of one variable domain (V<sub>L</sub>) and one constant domain (C<sub>L</sub>). The variable regions in the heavy and light chains each contain three hypervariable complementarity determining regions (CDRs) (Townsend et al., 2016). The CDR3 region in the heavy chain is particularly diverse and is considered to be the main determinant of specificity (Xu & Davis, 2000). Canonically, the CDRs of both the heavy and light chains combine to form the antigen-binding surface or ‘paratope’. The molecular intricacies of the paratope determine the specificity of the antibody.

IgG1 can be divided into two fragments: the Fc fragment (Fragment crystallisable) and the Fab fragment (Fragment antigen binding) (Figure 1.6). The Fc fragment is the signalling portion of the IgG1 and consists of two constant domains from each heavy chain. The Fc region contains binding sites for complement proteins (Duncan & Winter, 1988; Kaul & Loos, 1997) and Fcγ receptors on phagocytes (Indik, Park, Hunter, & Schreiber, 1995; Kiyoshi et al., 2015). The Fc region also contains a conserved glycosylation site (at asparagine 297) which contributes to the structural conformation that is bound by Fcγ receptors and complement proteins. Different Fcγ glycosylation patterns alter the affinity of these proteins for the Fc region and influence their
effector functions (Higel, Seidl, Sörgel, & Friess, 2016; Pincetic et al., 2014). This represents the effector ‘tail’ of the bivalent IgG1 molecule. The Fab consists of one variable and one constant region from each of the heavy and light chains and contains the paratope. This represents one ‘binding arm’ of the bivalent IgG1 molecule.

Antibodies bind to their target epitope, using the paratope within the Fab region, and signal to effector components of the immune system through the Fc region. This signalling enhances the recognition of the antigen by phagocytes and complement proteins, and can result in opsonophagocytosis and activation of the complement cascade as described in section 1.3. It is this complex interplay between S. pyogenes, antibodies, complement proteins and phagocytes that bactericidal assays and OPKAs measure.

Figure 1.6 Schematic of an IgG1 antibody. The IgG1 molecule forms an approximately ‘Y’ shaped structure with two Fab ‘binding arms’ and an Fc ‘tail’. The Fab consists of one variable domain (V_L) and one constant domain (C_L) from the light chain (blue) and one variable domain (V_H) and one constant domain (C_H1) from the heavy chain (grey). The variable domains contain three CDRs (lines) which form the antigen binding surface. The heavy and light chains are linked by a disulphide bond (yellow) between the constant domains. The Fc consists of two constant domains from each heavy chain which interact with complement proteins and phagocyte Fc receptors.
1.6.1 Isolation of monoclonal antibodies

Monoclonal antibodies are antibodies that are produced from a single antibody-producing B cell or a clonal population of antibody-producing B cells. Thus, the antibodies are identical and bind to the same epitope. However, infection and vaccination stimulate the development of polyclonal humoral responses, in which heterogeneous populations of antibody-producing B cells activate. In order to study individual monoclonal antibodies, they must be isolated from the polyclonal response. There are several ways to do this, including the generation of hybridomas, cloning from single B cells and antibody phage display.

Hybridoma technology produces monoclonal antibodies by fusing an antibody-producing B cell with a non-antibody producing myeloma cell (immortalised B cell cancer cell) (Köhler & Milstein, 1975). The resulting ‘hybridoma’ produces the same antibody that was encoded by the B cell, and has the replicative capabilities of the immortal myeloma cell. The hybridoma can then be cultured, continuously secreting homogenous monoclonal antibodies that can be readily purified (Köhler & Milstein, 1975). However, the B cells used to make hybridomas are usually derived from mice vaccinated with an antigen of interest and this somewhat limits their applications. The choice of vaccine antigen is limited as the antigen has to be put into animals. This means that it must be stable in vivo and cannot be toxic. The antigen also needs to be immunogenic and not subject to tolerance (which prevents an immune response developing to the vaccine antigen). Murine antibodies also have therapeutic limitations as a strong anti-mouse immune response develops in humans after the administration of therapeutic mouse antibodies. This reduces the efficacy of subsequent administrations as the antibodies are rapidly cleared (Courtenay-Luck et al., 1986; Tjandra, Ramadi, & McKenzie, 1990). To address this issue, hybridoma antibodies can be engineered as chimeric mouse-human or fully humanised antibodies but, while this reduces their immunogenicity, it can also reduce their affinity and efficacy (Jain, Kamal, & Batra, 2007).
Monoclonal antibodies can also be generated by amplifying the variable regions of antibodies from single B cell clones. This process, which can be readily performed with human samples, involves the isolation of peripheral blood mononuclear cells from individuals and subsequent cell sorting of single B cells. The variable regions of the antibodies that are encoded by these single B cells are then amplified and cloned into eukaryotic expression vectors for *in vitro* expression (Hofmann & Lai, 2017). Cell sorting single B cells into 96 well plates and cloning in parallel can make this process fairly high throughput, especially when compared to hybridoma technology (Tiller et al., 2008). However, like hybridoma technology, this process isolates a single antibody, produced by a single B cell in response to vaccination or infection.

In contrast to these single cell methods, phage display technology (described in detail in section 1.6.2) uses extremely large antibody libraries to screen high numbers of clones against the target antigen *in vitro* (McCafferty, Griffiths, Winter, & Chiswell, 1990; G. Smith, 1985). The antibody libraries are generated by amplifying the variable regions of antibodies from a large number of antibody producing B cells in the tissues of naive or immune animals. As such, the libraries contain antibody genes that represent the entire polyclonal antibody repertoire of the donor, as opposed to a single clone. As the heavy and light chains are amplified separately and then randomly paired, greater diversity in heavy-light chain pairing can be obtained than exists in the donor (Proetzel & Ebersbach, 2012). This can produce antibodies that would otherwise be deleted during the development of tolerance *in vivo*. Antibody libraries can also be generated ‘synthetically’ (Bazan, Calkosiński, & Gamian, 2012). In this case, the construction of the antibody library does not require antibody-producing B cells; the variable regions of germline antibody genes are amplified and randomised using PCR to generate near unlimited diversity. Whichever method is used to construct a library, they generally contain $10^6$–$10^{11}$ different clones (Bazan et al., 2012). Each of these individual clones can then be displayed on filamentous phage particles and screened against the antigen of interest simultaneously.
Screening and selection are performed in vitro (‘biopanning’, described in section 1.6.2) which allows precise control over the selection conditions (Lee, Iorno, Sierro, & Christ, 2007). These can be altered in many ways to promote the selection of antibodies of interest. For example, the isolation of antibodies that bind to a specific epitope on an antigen can be enhanced by manipulating the orientation of the antigen to favourably present that epitope (Proetzel & Ebersbach, 2012).

However, the true power of phage display technology lies in the vast size of the antibody libraries. This project seeks to characterise how antibodies, generated as part of an in vivo polyclonal immune response, interact with a vaccine antigen. The ability to sample the whole antibody repertoire and select monoclonal antibodies from it is important for the characterisation of the response to the vaccine. For this reason, this thesis uses antibody phage display technology to generate and isolate monoclonal antibodies from immune libraries.

1.6.2 Antibody phage display

In this project, antibody phage display is used to investigate the antibody responses to the S. pyogenes T antigen. Phage display technology is built upon the ability of the pIII coat protein of M13 filamentous bacteriophage to display antibody variable domains without compromising their capacity to bind to the target of interest (McCafferty et al., 1990). Since the phage contains the genetic information of the antibody within its particle, the phenotype and genotype of the displayed fragment are physically linked. Antibodies with desirable binding properties (such as the ability to bind to T antigens) can then be selected and the genes encoding these antibodies can be recovered.

Many different forms of antibody fragments are able to be displayed on phage particles including single chain variable fragments (scFv), Fab fragments, single VH domains and camelids (Bradbury & Marks, 2004; Lee et al., 2007). Each of these formats has its own pros
and cons. For example, purified Fab fragments are highly stable but are not well tolerated by bacteria, so can be difficult to express in high yields. In contrast, scFVs are more easily expressed in bacteria but are less stable than Fab (Lee et al., 2007). Thus, the selection of the antibody format is usually dictated by the intended downstream applications. This thesis uses phage display of a Fab library to isolate Fab that bind to a T antigen from S. pyogenes. As part of the characterisation process, the epitopes of the Fab on the T antigen are mapped. The ‘gold-standard’ tool for epitope mapping is x-ray crystallography as it provides atomic detail of the antibody-antigen interface. While x-ray crystallography can be unpredictable, Fab have a well-known stabilising effect (Griffin & Lawson, 2011; Hunte & Michel, 2002; Röthlisberger, Pos, & Plückthun, 2004) and the use of them in this thesis increased the chances of successful antibody-T antigen co-crystallography.

Large combinatorial Fab libraries have been created in which light-chain and heavy-chain antibody fragments are PCR amplified separately, and then combined by overlap PCR to generate antibody libraries with similar diversity to the donor’s repertoire (Barbas, Kang, Lerner, & Benkovic, 1991; Bradbury & Marks, 2004). These libraries can be either naïve, using natural or synthetic genes, or immune following vaccination or infection. Immune libraries that are made following vaccination are strongly biased towards antibodies that bind to the vaccine and can be readily used to investigate the antibody responses to the vaccine antigen (Bradbury & Marks, 2004). Tissue from the spleen and bone marrow of vaccinated animals is a rich source of antibody-producing B cells. RNA extracted from these cells can be reverse transcribed to produce cDNA that encodes the antibodies expressed by the B cells. This cDNA template can then be used to generate a Fab library that is representative of the vaccinated animal’s antibody repertoire (Figure 1.7).
Antibody phage display libraries are commonly constructed using a phagemid system, in which the Fab to be displayed is cloned into a plasmid and expressed as a fusion protein fused to the pIII coat protein with the aid of a ‘helper phage’ (Bradbury & Marks, 2004). These helper phage have defective origins of replication which promotes the assembly of phagemid particles containing the Fab-phagemid instead of the helper phage genome.

Figure 1.7 Flow diagram outlining the construction and panning of the mouse Fab library.

In order to select the Fab-phagemid particles that bind to the target antigen from the large library, a screening process termed biopanning is performed (Lee et al., 2007). This uses
affinity purification to isolate the Fab-phagemid particles that bind to the target from those that do not. Typically, 3-4 rounds are sufficient to isolate binders against a target protein. In biopanning, the recombinant library is incubated with the target antigen to allow the binding of Fab-phagemid particles that are specific to the target. The non-binding Fab-phagemid particles are then washed away to remove them from the selection process. The bound Fab-phagemid particles are eluted and used to infect susceptible *Escherichia coli*. This allows the amplification of the selected Fab-phage and their enrichment in the Fab-phage pool. This enriched pool is then used in the following round of biopanning. Over the course of biopanning, the stringency of selection is increased by decreasing the amount of target antigen and increasing the washing steps. This promotes the selection of high affinity Fab-phagemid particles that bind specifically to the target antigen (Lee et al., 2007).
1.7 Project Outline and Aims

Antibiotic therapy has failed to control *S. pyogenes* disease at the global level despite *S. pyogenes* remaining sensitive to penicillin. Prophylactic vaccination has been proposed as a viable control strategy. However, efforts to develop a vaccine to prevent *S. pyogenes* infection stretch back almost a century and have not yet produced a safe and effective vaccine. The most advanced candidate is a multivalent vaccine based on the M protein. This has been designed to protect against the most prevalent strains circulating in North America and Europe but is predicted to only provide partial coverage in at-risk populations in New Zealand and the developing world (Steer, Law, et al., 2009; Williamson et al., 2015; Williamson et al., 2016).

The T antigen that forms the shaft of the pilus is an alternative target for the development of a *S. pyogenes* vaccine.

The overarching aim of this study is to characterise how antibodies interact with T antigens using T18.1 (the T antigen expressed by M18/T18.1 strains of *S. pyogenes*) as a model to understand the protective capacity of a T antigen-based vaccine. T18.1 has been selected as it is a representative FCT3 T antigen that clusters evolutionarily with all of the FCT3/4/7/8 T antigens (Figure 1.4). These are the most common T antigens and are expressed by >70% of clinical isolates (Loh et al., 2017; Steemson et al., 2014). In this way, T18.1 represents the cluster of conserved *S. pyogenes* T antigens. M18/T18.1 strains are also clinically important having historically been associated with outbreaks of severe invasive disease and ARF (Johnson, Stevens, & Kaplan, 1992; Marcon et al., 1988; Smoot, Korgenski, Daly, Veasy, & Musser, 2002; Veasy et al., 2004). M18/T18.1 strains have also recently been associated with ARF in NZ (Williamson et al., 2015).

A detailed understanding of where antibody epitopes are located on T antigens and how protective they are will inform the rational development of T antigen-based vaccines.
The objectives of this thesis are:

1. **To clone, express and purify recombinant T18.1 and use x-ray crystallography to determine its atomic structure.** The two structures of *S. pyogenes* T antigens that have been solved to date are evolutionarily distinct from the majority of T antigens. T18.1 is a representative FCT3 T antigen and determining its atomic structure will provide insight into the structural conservation of T antigens, as well as the structural basis for cross-reactive and type-specific immune responses.

2. **To construct Fab libraries from mice vaccinated with the T18.1 pilus and use phage display to identify Fab that bind to T18.1.** Phage display methodology allows the study of large repertoires of antibodies from vaccinated animals in a high-throughput manner. The antibodies in the Fab-phage library are representative of the in vivo immune response to the vaccine, and those that bind to T18.1 can be selected from the library using biopanning.

3. **To characterise the binding of selected antibodies to T18.1 and map their epitopes using x-ray crystallography and an overlapping peptide library.**

   A small number of Fab isolated in objective 2 will be selected for further analysis. The binding of the Fab to monomeric and polymeric T18.1 will be characterised using biochemical and biophysical techniques. Their specificity and cross-reactivity will be assessed using a panel of recombinant T antigens that represent the majority of circulating *S. pyogenes* strains. Mapping the epitopes of the selected Fab will enable the identification of immunogenic regions of T18.1 and provide essential data towards the wider vaccine development effort.

4. **To measure the protective capacity of anti-T18.1 antibodies using in vitro bactericidal assays and opsonophagocytic killing assays.** This will provide the first data on the protective capacity of T antigen antibodies in isolation. The identification
of protective epitopes will provide 3-dimensional structural data to inform structure-led, vaccine design.
2 Materials and Methods

2.1 Materials

2.1.1 Ethical approvals

The use of human samples from the Rheumatic Fever Risk Factors (RFRF) study was approved by the Health and Disabilities Committees (HDEC) (14/NTA/53/AM02).

The use of whole blood from healthy human volunteers was approved by the University of Auckland Human Participants Ethics Committee (UAHPEC) (021200).

The use of animal samples was approved by the University of Auckland Animal Ethics Committee (UAAEC) (001664).

2.1.2 Molecular biology

PBS

137 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4, 1.8 mM KH2PO4, pH 7.4

PBST

PBS + 0.1% (v/v) Tween-20

MT

PBST + 5% (w/v) skim milk powder

TAE buffer

0.1% (v/v) glacial Acetic acid, 2 mM EDTA, 40 mM Tris pH 8.0

TSB buffer

10% (w/v) PEG-4000, 10 mM MgSO4, 10 mM MgCl2, 5% (v/v) DMSO made up in LB

6x DNA loading dye

30% (v/v) Glycerol, 0.25% (w/v) Bromophenol blue, 0.25% (w/v) Xylene Cyanol

Phage-PEG solution

20% (w/v) PEG-6000, 2.5 M NaCl
2.1.2.1 Plasmids

pProEXHTa-Avitag

A modified form of pProEXHTa where a C-terminal Avitag has been added downstream of the Thrombin cleavage site. Modification made by Dr Paul Young, The University of Auckland. Contains a multiple cloning site for the production of His$_6$-tagged recombinant proteins. Contains an ampicillin resistance gene.

pACYC184-BirA

pACYC184 that has a cloned BirA gene (encoding *E. coli* biotin ligase). Used for the in vivo biotinylation of recombinant proteins with an Avitag. Contains a chloramphenicol resistance gene. Obtained from Dr Paul Young, The University of Auckland.

pComb3XSS

A phagemid containing a 1600 bp stuffer fragment which can be removed by Sfil digestion. The phagemid contains Sfi1 restriction sites for the production of HA- and His$_6$-tagged recombinant Fab. When a Fab is cloned into pComb3X, an amber stop codon is present between the heavy chain constant region and geneIII of the phagemid. This allows for the production of a Fab-geneIII fusion protein in suppressor *E. coli* and soluble Fab in non-suppressor *E. coli*. Contains an ampicillin resistance gene.

pComb3XTT

A phagemid containing a cloned human Fab to tetanus toxin. Used as a template for the amplification of the kappa light chain constant region and the heavy chain constant region. Contains ampicillin resistance gene.
2.1.3 Bacterial culture

2.1.3.1 Media

GM17 medium 3.725% (w/v) M17 powder, 0.5% (w/v) Glucose

THY medium 3% (w/v) Todd Hewitt powder, 1% (w/v) Yeast extract

STGG medium 3% (w/v) Bacto-Tryptone, 0.5% (w/v) D-Glucose, 10% (v/v) Glycerol

LB broth 1% (w/v) Bacto-Tryptone, 0.5% (w/v) Yeast extract, 1% (w/v) NaCl

2xTY medium 1.6% Bacto-Tryptone, 1% Yeast extract, 0.5% (w/v) NaCl

SB medium 1% (w/v) MOPS, 3% (w/v) Bacto-Tryptone, 2% Yeast extract, pH 7.0

SOB medium 2% (w/v) Bacto-Tryptone, 0.5% Yeast extract, 0.05% (w/v) NaCl, 0.018% (w/v) KCl, 1% (v/v) MgCl₂, pH 7.0

SOC medium SOB + 20mM Glucose

2.1.3.2 Bacterial strains

XL1-Blue E. coli F’ proA+B+lacIqA(lacZ)M15 Tn10/ recA1 endA1 gyrA96 thi-1 hsdRI7 supE44 relA1 lac

A suppressor strain. Purchased from Agilent Technologies.

BL21 (DE3) E. coli F- ompT hsdSB(rB-mB-) gal dcm met (DE3)
Used for recombinant protein expression. Purchased from Thermo Fisher Scientific.

**TOP10 E. coli**  
F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15  
Δ lacX74 recA1 araD139Δ(araleu)7697 galU galK rpsL(StrR)  
endA1 nupG  
Used for cloning recombinant proteins. Purchased from Thermo Fisher Scientific.

**TOP10F’ E. coli**  
F’ (lacIq, Tn10(TetR)) mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL(StrR)  
endA1 nupG  
A non-suppressor strain.

Used for the expression of recombinant Fab. Purchased from Thermo Fisher Scientific.

**TG1 E. coli**  
F’ traD36 proAB lacIqZ ΔM15] supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5(rK - mK -)  
Used for titering phage. Purchased from Lucigen.

**L. lactis-pil18**  
*L. lactis* MG1363 containing the complete pilus operon from MGAS8232 (M18/T18.1) *S. pyogenes* in the expression vector pLZ12-Km2 P23R. This modification was made by Dr Jacelyn Loh, The University of Auckland.

**MGAS8232**  
M18/T18.1 *S. pyogenes*. Reference strain. Obtained from Associate Professor Thomas Proft, The University of Auckland.

**M217_11574**  
M217/T18.1 *S. pyogenes*. Clinical isolate. Obtained from Associate Professor Thomas Proft, The University of Auckland.
M49_12339  M49/T18.2 *S. pyogenes*. Clinical isolate. Obtained from Associate Professor Thomas Proft, The University of Auckland.

M1 (43)  M1/T1 *S. pyogenes*. Clinical isolate. Obtained from Professor Shiranee Sriskandan, Imperial College London.

M6 (2)  M6/T6 *S. pyogenes*. Clinical isolate. Obtained from Professor Shiranee Sriskandan, Imperial College London.

### 2.1.3.3 Selective antibiotics

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Final working concentration</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin (AMP)</td>
<td>100 μg/mL</td>
<td>Gold Biotechnology</td>
</tr>
<tr>
<td>Chloramphenicol (CAM)</td>
<td>25 μg/mL</td>
<td>Duchefa</td>
</tr>
<tr>
<td>Kanamycin (KAN)</td>
<td>50 μg/mL for <em>E. coli</em></td>
<td>Gibco</td>
</tr>
<tr>
<td></td>
<td>200 μg/mL for <em>L. lactis</em></td>
<td></td>
</tr>
<tr>
<td>Tetracycline (TET)</td>
<td>10 μg/mL</td>
<td>Gold Biotechnology</td>
</tr>
</tbody>
</table>

### 2.1.4 Protein purification and analysis

| T antigen lysis buffer    | 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 2% (v/v) Glycerol, 0.1 μg/mL Lysozyme |
| T antigen wash buffer    | 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM Imidazole, 2% (v/v) Glycerol |
T antigen elution buffer 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 500 mM Imidazole, 2% (v/v) Glycerol

Crystallisation buffer 10 mM Tris pH 8.0, 100 mM NaCl

TES 0.2 M Tris pH 8.0, 0.5 mM EDTA, 0.5 M Sucrose

Fab wash buffer 50 mM Tris pH 8.0, 250 mM NaCl, 20mM Imidazole

Fab elution buffer 50 mM Tris pH 8.0, 100 mM NaCl, 500 mM Imidazole

Protoplast buffer 40% Sucrose, 10 mM MgCl2, 0.1 M KPO4 pH 6.2, 2 mg/ml Lysozyme, 400 U mutanolysin

Towbin Transfer Buffer 192 mM glycine, 0.37% (w/v) SDS, 20% (v/v) methanol, 25 mM Tris HCl pH 8.3

2.1.5 Functional assays

2.1.5.1 Reagents

FACS blocking buffer PBS, 3% FBS, 5 mM EDTA

FACS buffer PBS, 1% FBS, 5 mM EDTA

Opsonisation buffer 10% (v/v) FBS, 0.1% (w/v) gelatin, in HBSS with Ca/Mg
### 2.1.5.2 Antibodies

Table 2.1 Commercial antibodies used in this thesis.

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Species</th>
<th>Supplier</th>
<th>Working dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-human IgG (H+L)-HRP</td>
<td>Donkey</td>
<td>Jackson Immunoresearch</td>
<td>1:2500</td>
</tr>
<tr>
<td>Anti-Human IgG (H+L)- Alexa Fluor® 488</td>
<td>Donkey</td>
<td>Jackson Immunoresearch</td>
<td>1:1000</td>
</tr>
<tr>
<td>Anti-mouse IgG-HRP</td>
<td>Sheep</td>
<td>GE Healthcare</td>
<td>1:5000</td>
</tr>
<tr>
<td>Anti-rabbit IgG-HRP</td>
<td>Goat</td>
<td>Abcam</td>
<td>1:5000</td>
</tr>
<tr>
<td>Anti-HA-HRP</td>
<td>Mouse</td>
<td>Roche</td>
<td>1:500</td>
</tr>
<tr>
<td>Anti-M13-HRP</td>
<td>Mouse</td>
<td>GE Healthcare</td>
<td>1:5000</td>
</tr>
</tbody>
</table>

### 2.2 Methods

#### 2.2.1 General Methods

#### 2.2.1.1 DNA agarose gel electrophoresis

The DNA sample was mixed with the appropriate amount of 6x DNA loading dye and run at 70-100 V through a 1-2% (w/v) TAE agarose gel containing Redsafe Nucleic Acid Staining Solution (iNtRON) diluted 1:20,000. The DNA bands were visualised using a ChemiDoc Touch system (Bio-rad).
2.2.1.2 Preparation of electrocompetent cells

A single colony of *E. coli* was used to inoculate 5 mL of 2xTY media supplemented with strain appropriate antibiotics. This culture was incubated for 16 hours at 37°C, 180 rpm. Following incubation, the 5 mL culture was diluted 100-fold in 2xTY media supplemented with strain appropriate antibiotics and incubated at 37°C until the O.D$_{600}$ reached 0.5. The culture was cooled on ice for 20 min then centrifuged for 15 min at 4,000 g, 4°C. The cells then underwent three cycles of washing with ice-cold 10% Glycerol and centrifugation for 15 min at 4,000 g, 4°C. Each cycle, the pellet was resuspended in a lower volume of 10% Glycerol (500 mL, 250 mL then 20 mL). After the third spin, the cells were resuspended in 1.5 mL of 10% Glycerol, aliquoted and snap-frozen in liquid nitrogen. The cells were stored at -80°C until use.

2.2.1.3 Preparation of chemocompetent cells

A single colony of *E. coli* was used to inoculate 3 mL of 2xTY media supplemented with strain appropriate antibiotics. This culture was incubated for 16 hours at 37°C, 180 rpm. Following incubation, 40 μL of this culture was used to inoculate 40 mL of pre-warmed (37°C) 2xTY media (supplemented with strain appropriate antibiotics) which was incubated at 37°C until the O.D$_{600}$ reached 0.6. The culture was cooled on ice for 30 min then centrifuged for 10 min at 1,000 g, 4°C. The pellet was resuspended in 4 mL of TSB buffer (10% (w/v) PEG-4000, 10 mM MgSO$_4$, 10 mM MgCl$_2$, 5% (v/v) DMSO made up in LB) then aliquoted, snap-frozen in liquid nitrogen and stored at -80°C until use.

2.2.1.4 Denaturing SDS-polyacrylamide gel electrophoresis

Denaturing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was carried out using a Mini-PROTEAN® Tetra Cell System (Bio-Rad). SDS-PAGE gels were prepared by pouring resolving gel solution containing 8 - 15 % (w/v) acrylamide mix (Bio-Rad). Once the resolving gels were set, the stacking gel solution containing 5 % (w/v)
acrylamide was poured over the top of the resolving gel and combs were inserted to form wells. The solutions for making the gels and running SDS-polyacrylamide gel electrophoresis are shown in Table 2.2.

Protein samples (1–15 μl) were prepared by suspending in an equal volumes of SDS-gel loading buffer and heated at 95°C for 5 min. They were then cooled and electrophoresed beside molecular weight standards; either pre-stained or unstained SDS-PAGE page molecular markers (Bio-Rad). The samples were loaded onto the gel and electrophoresed in Tris-glycine buffer at room temperature under constant voltage of 120 V. Following electrophoresis, the gels were stained with Coomassie Blue staining solution for 45 min at RT with agitation. The Coomassie solution was then removed and the gel was rinsed with sterile water followed by soaking in destaining solution for 45 min with agitation.

Table 2.2 Buffers and solutions used for SDS-polyacrylamide gel electrophoresis.

<table>
<thead>
<tr>
<th>Resolving gel</th>
<th>Stacking gel</th>
<th>SDS-gel loading buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>375 mM Tris-HCl pH 8.8</td>
<td>125 mM Tris-HCl pH 6.8</td>
<td>50 mM Tris-HCl pH 6.8</td>
</tr>
<tr>
<td>10 – 15% (w/v) acrylamide</td>
<td>5% (w/v) acrylamide</td>
<td>100 mM DTT</td>
</tr>
<tr>
<td>1% (w/v) SDS</td>
<td>1% (w/v) SDS</td>
<td>2% (w/v) SDS</td>
</tr>
<tr>
<td>1% (w/v) (NH4)2S2O8</td>
<td>1% (w/v) (NH4)2S2O8</td>
<td>1% (w/v) bromophenol blue</td>
</tr>
<tr>
<td>0.04% (v/v) TEMED</td>
<td>0.1% (v/v) TEMED</td>
<td>10% (v/v) Glycerol</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tris – glycine buffer</th>
<th>Coomassie blue</th>
<th>Destaining solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 mM Tris</td>
<td>0.125% (w/v) Serva blue R</td>
<td>30% (v/v) methanol</td>
</tr>
<tr>
<td>250 mM glycine</td>
<td>30% (v/v) methanol</td>
<td>10% (v/v) acetic acid</td>
</tr>
<tr>
<td>0.1% (w/v) SDS</td>
<td>10% (v/v) acetic acid</td>
<td></td>
</tr>
</tbody>
</table>
2.2.1.5 Pull-down of biotinylated proteins

A volume of M280 Dynabeads with the binding capacity for the amount of protein to be pulled-down was added to a 1.5 mL tube and washed in 1 mL PBS. The tube was placed on a magnet for 3 min to separate the beads from solution and the PBS was removed. The biotinylated protein of interest was added to 500 μL PBS and the beads were resuspended in this solution. This mixture was incubated for 1 hour on a rotating wheel. The tube was then placed on a magnet for 3 min and the supernatant (containing unbound protein) was removed. The beads were then washed 6 times with 1 mL PBS and the first and last wash were collected. These washes and 10 μL of beads were analysed by SDS-PAGE (section 2.2.1.4).

2.2.2 Preparation of tee18.1 and tee18.1-Avitag

2.2.2.1 Cloning tee18.1 and tee18.1-Avitag

The tee18.1 and tee18.1-Avitag genes were PCR amplified from a plasmid (pET32a 3C-tee18.1) containing tee18.1 (kindly provided by Dr Loh) from the MGAS8232 strain. All primers were produced by Integrated DNA Technologies to a purity of desalted, diluted in MilliQ water (MQ) and stored at -20°C.

The tee18.1 and tee18.1-Avitag genes were PCR amplified in 100 μL reactions containing: 1 ng pET32a 3C-tee18.1, 0.3 μM sense primer, 0.3 μM antisense primer, 20 μL of 5x PCR buffer, 0.05 mM of each dNTP and 1 μL PrimeSTAR HS DNA polymerase (Takara). The primer pairs are shown in Table 2.3 and the cycle parameters are show in Table 2.4.
Table 2.3 Primers used for the amplification of $\textit{tee18.}1$ from pET32a 3C-$\textit{tee18.}1$.

<table>
<thead>
<tr>
<th>Primer sequence 5' to 3'</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\textit{tee18.}1$ sense primer</td>
<td>GTATTTTCAGGGGCGCAGACAGCAGGAGTGATTGATGGTTCA</td>
</tr>
<tr>
<td>$\textit{tee18.}1$ antisense primer</td>
<td>GACTGCAGGGCTCTAGATTAGTCACGCTTATTTGTGACAACGATTTCGTC</td>
</tr>
<tr>
<td>$\textit{tee18.}1$-Avitag antisense primer</td>
<td>GACTGCAGGGCTCTAGAGTCACGCTTATTTGTGACAAACGATTTCGTC</td>
</tr>
</tbody>
</table>

The Kas1 (sense primer) and Xba1 (antisense primers) restriction sites are shown in bold. The stop codon that prevents expression of the Avitag is shown in italics.

Table 2.4 PCR parameters used for the amplification of $\textit{tee18.}1$ from pET32a 3C-$\textit{tee18.}1$.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>98°C</td>
<td>5 min</td>
</tr>
<tr>
<td>10 cycles of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>98°C</td>
<td>20 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>70°C (-1°C/cycle)</td>
<td>30 s</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>90 s</td>
</tr>
<tr>
<td>25 cycles of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>98°C</td>
<td>20 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>60°C</td>
<td>30 s</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>90 s</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>10 min</td>
</tr>
</tbody>
</table>
The amplified products were analysed on a 1% agarose gel and then spin-column purified (Macherey Nagel) according to the manufacturer’s instructions.

The purified PCR products and the pProEXHTa-Avitag vector were double-digested with Kas1 (NEB) and Xba1 (NEB) restriction enzymes for 90 min at 37°C. The reaction mixture is shown in Table 2.5.

Table 2.5 Double-digest of the purified tee18.1 PCR products and pProEXHTa-Avitag with Kas1 and Xba1.

<table>
<thead>
<tr>
<th></th>
<th>tee18.1 PCR products</th>
<th>pProEXHTa-Avitag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template (μg)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Kas1 (U)</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Xba1 (U)</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>10x buffer NEB2.1 (μL)</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

The restriction digested products were spin-column purified (Macherey Nagel) as per the manufacturer’s instructions. The tee18.1 fragments were ligated into pProEXHTa-Avitag in 5 μL reactions that contained: 60 ng digested pProEXHTa-Avitag, 40 ng digested tee18.1 or tee18.1-Avitag, 1 μL 5x T4 ligase buffer (Invitrogen) and 1 U T4 DNA ligase. The ligation reaction was incubated for 16 hours at room temperature (RT).

The ligated products were transformed into chemically competent TOP10 E. coli as described in 2.2.2.2.

To identify positive clones, LB+AMP cultures (10 mL) were inoculated with a single colony from the plates containing transformed TOP10E. coli and incubated for 16 hours at 37°C, 180
rpm. The plasmids were purified from 4 mL of culture using the ZR Miniprep-Classic kit (Zymo Research) according to the manufacturer’s protocol.

The clones were sequence confirmed by Sanger sequencing. Purified plasmids were submitted for sequencing (carried out by Kristine Boxen, Centre for Genomics and Proteomics of the University of Auckland) using tee18.1 sense and tee18.1 antisense primers. The sequences were analysed using Geneious software (Biomatters Ltd).

The plasmids that were confirmed to contain the tee18.1 genes were stored at -20°C.

2.2.2.2 Transformation into *Escherichia coli* using the heat shock method

Plasmids were transformed into chemically competent *E. coli* (prepared as described in section 2.2.1.3) using the heat shock method. 1 μL of plasmid was added to 50 μL of chemically competent *E. coli* and incubated on ice for 30 min. The cells were heat shocked at 42°C for 1 min and immediately incubated on ice for 2 min. 950 μL of LB broth was added to each transformation mixture and they were incubated for in a Thermomixer (Eppendorf) for 1 hour at 37°C, 800 rpm. Following incubation, aliquots were spread plated onto LB+AMP plates (LB broth supplemented with 1.5% agar and 100 μg/mL AMP) and incubated for 16 hours at 37°C.

2.2.3 Expression and purification of recombinant proteins

2.2.3.1 Expression of recombinant T antigens

Plasmids containing the tee gene for the T antigen to be expressed were transformed into *E. coli* BL21 (DE3) using the heat shock method as described in section 2.2.2.2. The transformed cells were used to inoculate 10 mL of 2xTY+AMP (2xTY broth supplemented with 100 μg/mL AMP) and then incubated for 16 hours at 37°C. This ‘seeder’ culture was used to inoculate 1 L of 2xTY+AMP which was then incubated at 37°C, 200 rpm until the bacterial growth reached exponential phase (OD600 = 0.6-0.8). The culture was induced with 1 mM
Isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated for a further 4 hours at 37°C, 200 rpm. The pellet was collected by centrifugation at 4,000 g for 30 min and stored at -20°C.

### 2.2.3.2 Expression of biotinylated T18.1

The pProEXHTa-tee18.1-Avitag and pACYC184-BirA plasmids were co-transformed (at a 1:1 ratio) into *E. coli* BL21 (DE3) as described in section 2.2.2.2 except that after co-transformation, all media were supplemented with 100 μg/mL AMP and CAM 25 μg/mL. The transformed cells were used to inoculate 10 mL of 2xTY+AMP+CAM and then incubated for 16 hours at 37°C. This ‘seeder’ culture was used to inoculate 1 L of 2xTY+AMP+CAM which was then incubated at 37°C, 200 rpm until the bacterial growth reached exponential phase (O.D₆₀₀ = 0.6-0.8). The culture was induced with 1 mM IPTG and 20 μM D-Biotin was added at this time. The culture was incubated for a further 4 hours at 37°C, 200 rpm. The pellet was collected by centrifugation at 4,000 g for 30 min and stored at -20°C.

### 2.2.3.3 Purification of recombinant T antigens using Immobilised Metal Affinity Chromatography

The bacterial pellet was resuspended in 10% (w/v) T antigen lysis buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 2% (v/v) Glycerol, 0.1 μg/mL Lysozyme) supplemented with a complete protease inhibitor cocktail EDTA-free minitablet (Roche) and disrupted using a Qsonica Cell Disruptor sonicator (Misonix). The cell debris was pelleted by centrifugation for 25 min at 10,000 g, 4°C. The supernatant was taken and passed through a 0.2 μM filter (Sartorius) onto a Nickel charged-Nitrilotriacetic acid column (Ni-NTA, GE Healthcare) that had been pre-washed with 3 column volumes of T antigen wash buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM Imidazole, 2% (v/v) Glycerol). Unbound or non-specifically bound protein were removed by washing with 10 column volumes of T antigen wash buffer. The T antigen was eluted from the column in a continuous gradient (20 mL) of Imidazole from 0 mM to 500
mM by mixing T antigen wash buffer and T antigen elution buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 500 mM Imidazole, 2% (v/v) Glycerol) which was done using an AKTA FPLC system (GE Healthcare). The collected fractions were analysed by SDS-PAGE and the fractions containing T antigen were pooled.

The His$_6$ tag was removed from the T antigen by cleavage with rTEV-His$_6$ protease (kindly provided by Dr Young) at a ratio of 1:50 to the pooled protein (+ 5 mM 2- Mercaptoethanol, BME) for 4 hours, 4°C. The T antigen was then purified from the cleaved His$_6$ tags and rTEV-His$_6$ by passing it back over a pre-washed Ni-NTA column. The column was washed with 3 column volumes of T antigen wash buffer to wash through all of the purified untagged T antigen. This was collected and dialysed into PBS for 16 hours at 4°C. The recombinant protein was concentrated using a 30-kDa molecular-mass protein concentrator (VivaScience) and snap-frozen. Protein was stored at -80°C until use.

**2.2.3.4 Purification of recombinant T antigens using Size Exclusion Chromatography**

Size Exclusion Chromatography (SEC) was used as the final protein purification step before crystallography. All solutions used in SEC were freshly filtered through 0.2 μm filters (Sartorious) and degassed. Using an AKTA FPLC system, a Superdex 200 10/300GL column (GE Healthcare) was washed with 2 column volumes of MQ followed by 1 column volume of crystallisation buffer (10 mM Tris pH 8.0, 100 mM NaCl), according to the manufacturer’s instructions. The concentrated protein (see section 2.2.3.3) was injected onto the column which was run at 1 mL/min collecting 0.5 mL fractions. The collected protein, separated on the basis of size, was analysed by SDS-PAGE.
2.2.3.5 Expression of recombinant Fab

The pComb3X phagemids containing the 20 selected Fab clones were transformed into chemically competent TOP10F’ (Invitrogen) E. coli using the heat shock method as described in section 2.2.2.2.

A single colony from the plate of transformed cells was used to inoculate 10 mL of 2xTYAG (2xTY media supplemented with 100μg/mL of ampicillin and 2% Glucose) and incubated for 16 hours at 37°C, 200rpm. This ‘seeder’ culture was used to inoculate 1 L of 2xTY media supplemented with 100μg/mL of AMP and 0.1% Glucose. This culture was incubated at 30°C, 200 rpm until the bacterial growth reached exponential phase (O.D$_{600}$ = 0.6-0.8). Fab expression was induce with 1 mM IPTG and the culture was incubated for 16 hours at 18°C.

2.2.3.6 Purification of recombinant Fab

The pellet was harvested by centrifugation for 30 min at 4,000 g, 4°C. The Fab was purified from the pellet using a cold osmotic shock method. Briefly, the pellet was resuspended in 20 mL of chilled TES (0.2 M Tris pH 8.0, 0.5 mM EDTA, 0.5 M Sucrose) and incubated on ice for 30 minutes with occasional mixing. The periplasmic fraction (containing most of the recombinant Fab) was separated from the cells by centrifugation for 15 min at 13,000 g, 4°C. The periplasmic fraction was then removed and stored on ice. The pellet was then resuspended and in 20 mL of 1/5 TES (supplemented with one complete protease inhibitor cocktail EDTA-free minitablet and 100μL of 1 M MgCl2), and centrifuged for 15 min at 13,000 g, 4°C. This supernatant was removed and added to the periplasmic fraction along with 20 mM Imidazole.

The combined Fab containing supernatant was bound onto a 1 mL Ni-NTA that had been pre-washed with 3 column volumes of Fab wash buffer (50 mM Tris pH 8.0, 250 mM NaCl, 20 mM Imidazole). Unbound or non-specifically bound proteins were removed by washing with 10 column volumes of Fab wash buffer. The recombinant Fab was eluted from the column in
a continuous gradient (20 mL) of Imidazole from 0 mM to 500 mM by mixing Fab wash buffer and Fab elution buffer (50 mM Tris pH 8.0, 100 mM NaCl, 500 mM Imidazole) which was done using an AKTA FPLC system. The collected fractions were analysed by SDS-PAGE and the fractions containing Fab were pooled. The Fab was buffer exchanged into PBS and concentrated using a 30-kDa molecular-mass protein concentrator. The purified Fab was aliquoted and stored at -80°C.

2.2.3.7 Extraction of cell wall proteins from \textit{Lactococcus lactis}

\textit{L. lactis} (constitutively expressing the pilus proteins encoded in the pilus operon) were grown in GM17 media (3.725% (w/v) M17 powder, 0.5% (w/v) Glucose) supplemented with 200 μg/mL KAN for 16 hours at 30°C. 1 mL of this culture was taken and diluted 1:100 in GM17 media-KAN-200. 1 mg of the αE3 Fab or B9 Fab (a T1 binding Fab that does not bind to T18.1-used here as a control for the addition of 1 mg of protein to the growth media) were added to the culture and it was then grown for 3 hours at 30°C. The culture was then pelleted by centrifugation for 10 min at 5,000 g, 4°C. The supernatant was discarded, the pellet was washed with 1 mL PBS and centrifuged again for 10 min at 5,000 g, 4°C. The pellet was washed with 1 mL PBS and centrifuged again for 10 min at 5,000 g, 4°C. The pellet was resuspended in 1 mL of cold protoplast buffer (40% Sucrose, 10 mM MgCl2, 0.1 M KPO4 pH 6.2, 2 mg/ml Lysozyme, 400 U Mutanolysin) and incubated for 3 hours at 37°C. The protoplast and cell wall extract were separated by centrifugation for 15 min at 13,000, 4°C. The protoplast was resuspended in PBS and stored at -20°C until use. The cell wall extract was stored separately at -20°C.
2.2.4 X-ray crystallography

2.2.4.1 Protein crystallisation using sitting drop vapour diffusion crystallisation

For T18.1, crystallisation trials were carried out using a 480-condition screen developed at the University of Auckland (Moreland et al., 2005) in addition to the MORPHEUS screen (Gorrec, 2009). For the αE3-T18.1 and αH3-T18.1 complexes, the crystallisation trials were carried out using a 288-condition screen comprising the JCSG+/PACT premier screen (Newman et al., 2005) and the MORPHEUS screen (Gorrec, 2009).

The proteins for all crystallisation experiments were purified by IMAC and SEC (as outlined in section 2.2.3.3 and section 2.2.3.4) and concentrated to between 5 mg/mL and 140 mg/mL. The concentrated proteins were then centrifuged for 30 min at 13,000 g, 4°C to pellet any particulates.

Each of the screen solutions (85 μl) was dispensed into the deep wells of an Intelli-Plate 96 (Hampton Research). The Intelli-Plate 96 was then transferred to the Oryx4 crystallisation robot (Douglas Instruments) that dispensed 0.15 μL of the purified recombinant protein followed by 0.15 μL of the screen precipitant into the shallow drop wells of the Intelli-Plate 96. The plate was then sealed with a clear plastic film and stored at 18°C.

The principle of this method is that, over time, the solvent in the crystallisation drop (in the shallow drop wells) diffuses until equilibrium is reached with the deep well solution. This results in an increase in the concentrations of the screen precipitant and protein, which may lead to protein crystal formation in some cases. In accordance with this, all crystallisation drops were monitored for crystal formation with a light microscope on a regular basis. Selected crystals were then removed from their drops and mounted on cryoloops (Hampton Research) and flash cooled in liquid nitrogen with cryoprotectant (20% Glycerol). The crystals selected
from the MORPHEUS screen were not flash cooled in 20% Glycerol as the precipitant is a cryoprotectant (Gorrec, 2009).

### 2.2.4.2 X-ray diffraction data collection

The diffraction data for this study were collected at the Australian Synchrotron. The Australian Synchrotron has two macromolecular crystallography beamlines, MX1 and MX2. The MX1 beamline uses an ADSC Quantum 210r Detector while the MX2 beamline uses an Eiger x 16M detector. The samples were loaded using a sample mounting robot and data were collected at a constant wavelength (MX1= 0.95468Å, MX2= 0.953724Å) while the crystals were rotating 360° around an axis.

### 2.2.4.3 Integration and Scaling

The dataset was indexed and integrated using XDS (Kabsch, 2010) and scaled using AIMLESS. The AIMLESS log file contains information such as Rmerge, signal-to-noise ratio (mean I/ σ(I)), completeness, and redundancy to monitor data quality. The number of molecules per asymmetric unit was estimated by using Matthews coefficient analysis (Matthews, 1968) as part of the CCP4 software suite (Winn et al., 2011). The Matthews Coefficient is determined by dividing the volume of the unit cell by the molecular weight of the protein in the unit cell.

### 2.2.4.4 Structure solution using Molecular Replacement

Molecular Replacement (MR) can be used to solve the crystallographic phase problem if a known structure with a similar fold to the unknown structure is used as the search model. This generally requires a sequence identity of greater than 30% between the two structures (P. Evans & McCoy, 2008). The phases are calculated from the search model which is placed in the unit cell of the unknown structure and translated and rotated to determine the position of the unknown structure in the unit cell (P. Evans & McCoy, 2008). In this thesis, MR was used to
solve the structures of both T18.1 and the αE3-T18.1 complex. The structure of T18.1 was solved by MR using the previously published T1 structure (PDB ID: 3B2M)(Kang et al., 2007) as a search model. The structure of T18.1 was then used as a search model to solve the structure of the αE3-T18.1 complex. The programme PHASER, part of the CCP4 software suite, was used for MR (McCoy, 2007). PHASER uses maximum likelihood approaches and multivariate statistics to find the correct orientation and position of the search model in the unit cell of the unknown structure. The MR solutions are scored based on Log Likelihood Gain (LLG) and by a Z-score. The LLG measures the probability that the observed data (from the unknown structure) would have been measured, given the predicted model. The Z-score measures the signal-to-noise ratio by comparing the LLG from the model with the LLG for a random set of translations and rotations. The higher the LLG and Z-scores, the better the MR solution is with a Z-score >7 indicating that the structure has ‘probably’ been solved (McCoy, 2007).

2.2.4.5 Model building, refinement and validation

Once the structures were solved using MR, the initial model was manually built into the electron density using COOT (Emsley, Lohkamp, Scott, & Cowtan, 2010). Alternating rounds of building in COOT and refinement using Refmac5 (Murshudov et al., 2011) were performed. The models were inspected and modified using the $2F_{\text{obs}}$-$F_{\text{calc}}$ and $F_{\text{obs}}$-$F_{\text{calc}}$ electron density maps where $F_{\text{obs}}$ is the structure factor observed in the experiment and $F_{\text{calc}}$ is the structure factor calculated from the model. The model building tools in COOT were used to build a model that best fit the electron density. The quality of the model was then assessed using validation tools within COOT such as the Ramachandran plot.

During refinement, the quality of the model was assessed by assessing the R factor calculated by Refmac5. This measures how well the model structure factors ($F_{\text{calc}}$) and the experimental structure factors ($F_{\text{obs}}$) agree. This is calculated using the following formula:
The $R_{\text{free}}$ of the model was also assessed after each round of refinement. This is calculated in the same way as the R factor but only uses a subset (5%) of randomly selected reflections which are excluded from the refinement cycle. This gives a measurement of the model quality that is unbiased by refinement. A large difference between R factor and $R_{\text{free}}$ may indicate that the fit of the model is biased. Alternating model building and refinement steps were continued until the R factor and $R_{\text{free}}$ reached their minimum. The final models were assessed using MOLPROBITY (Davis et al., 2007). The MOLPROBITY score is an aggregated score that combines the analysis of steric clashes, side chain rotamers, geometry and Ramachandran outliers (Davis et al., 2007).

2.2.5 Construction of the Fab-phagemid library

2.2.5.1 Animal vaccinations with T18.1 containing vaccines

Prior to the beginning of this project, Dr Loh immunised 10 female FVBn mice with *Lactococcus lactis* constitutively expressing the T18.1 pilus operon (*L. lactis*-pil18) and ten female FVBn mice with *L. lactis* carrying empty vector (pLZ12-Km2 P23R). Each mouse was immunised intranasally with 25 µL of culture per nostril for a total inoculation of $2 \times 10^7$ CFU/mouse. They received two boosts at day 14 ($8 \times 10^7$ CFU/mouse) and 28 ($2.75 \times 10^8$ CFU/mouse) post immunisation. They were euthanised on day 52. Blood was collected and processed to extract serum (section 2.2.5.2) and the spleens were harvested and snap frozen in liquid nitrogen.

A female New Zealand white rabbit was vaccinated subcutaneously with 100 µg of purified recombinant T18.1 mixed 1:1 with Incomplete Freund’s adjuvant (IFA). It received three
identical boosters at 14, 28 and 42 days post immunisation and was euthanised on day 56. Blood was collected and processed to extract serum (section 2.2.5.2). The spleen and bone marrow (collected from the femurs) were harvested and snap frozen in liquid nitrogen.

2.2.5.2 Processing serum

The animal blood was collected and allowed to clot for 40 min at RT. The serum was separated by centrifugation at 1250 g for 20 min at 4 °C then immediately frozen and stored at -80 °C.

2.2.5.3 RNA extraction from L. lactis-pil18 immunised mouse spleens

The spleens from five of the L. lactis-pil18 immunised mice (mouse B1- mouse B5) were disrupted by grinding them to a powder in liquid nitrogen and homogenised by spinning in Qiashredder spin-columns (Qiagen) at 13,000g for 2 min. Total RNA was purified using an RNeasy Mini Kit (Qiagen) as per the manufacturer’s instructions. The concentration and purity of the extracted RNA were measured using a NanoDrop 2000 (Thermo Fisher Scientific). The RNA from each mouse was stored separately at -80°C.

2.2.5.4 Reverse transcription of RNA to cDNA

The conversion of total RNA to cDNA was performed using the Superscript III First-Strand Synthesis System (Invitrogen) as per the manufacturer’s instructions. Reverse transcription was performed individually for each mouse. The cDNA synthesis mix consisted of 3 μg total RNA, 2.5 μM of oligo(dT)$_{20}$ primer, 0.125 mM of each dNTP, 5 mM MgCl$_2$, 10 mM DTT, 40 U of RNaseOut and 200 U of SuperScript III reverse transcriptase, in RT buffer (total volume 20 μL). The reaction was incubated for 50 min at 50°C then stopped by incubating at 85°C for 5 min. The cDNA was stored at -20°C.
2.2.5.5 Amplification of the mouse variable and human constant heavy and light chains via PCR

The amplification of antibody genes and Fab-phagemid library construction is based on published protocols (Rader, 2009). The heavy (V<sub>H</sub>) and light chain (V<sub>L</sub>) antibody genes were amplified from immunised mouse cDNA using PCR. 100 μL reactions were performed using an optimised protocol which consisted of: 0.75 μL of template cDNA, 10 μM sense primer, 10 μM antisense primer, and 20 μL 5 x PCR buffer, 0.2 mM of each dNTP and 0.5μL of PrimeSTAR HS DNA polymerase (Takara). The primer pairs and number of reactions are shown in Table 2.6. The primer sequences and the make-up of the primer mixes are shown in appendix A. The PCR parameters are shown in Table 2.7. The V<sub>H</sub> and V<sub>L</sub> genes were amplified from each mouse individually.

Table 2.6 Primer pairs for the amplification of mouse V<sub>H</sub> and V<sub>L</sub> genes.

<table>
<thead>
<tr>
<th></th>
<th>Sense primer</th>
<th>Antisense primer</th>
<th>Number of reactions per mouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>V&lt;sub&gt;k&lt;/sub&gt; amplification</td>
<td>MSCVK mix</td>
<td>MHybLJ-B mix</td>
<td>9</td>
</tr>
<tr>
<td>V&lt;sub&gt;λ&lt;/sub&gt; amplification</td>
<td>MSCVL-1</td>
<td>MHybLJ-B</td>
<td>1</td>
</tr>
<tr>
<td>V&lt;sub&gt;H&lt;/sub&gt; amplification</td>
<td>MHyVH mix</td>
<td>MHyIgGCH1-B mix</td>
<td>12</td>
</tr>
</tbody>
</table>

The PCR products were analysed on a 2% agarose gel (section 2.2.1.1). All of the products from the V<sub>k</sub> and V<sub>λ</sub> PCR reactions were pooled (for each mouse). All of the products from the V<sub>H</sub> PCR reactions were also pooled (for each mouse). The pooled PCR products were purified using spin-columns as per the manufacturer’s instructions (Macherey Nagel).
Table 2.7 PCR parameters used for the amplification of mouse V_<sub>H</sub> and V_<sub>L</sub> genes.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>98°C</td>
<td>3 min</td>
</tr>
<tr>
<td>30 cycles of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>98°C</td>
<td>15 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>56°C</td>
<td>30 s</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>90 s</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>10 min</td>
</tr>
</tbody>
</table>

Further PCR reactions were performed to amplify the human heavy (CH1γ) and light (Cκ) constant regions from the pComb3XTT phagemid which contains a cloned human Fab fragment. The primer combinations used in the PCR are detailed in appendix A. Ten 100μL PCR reactions were performed for each of the human constant regions. These consisted of: 1 ng pComb3XTT, 10 μM sense primer, 10 μM antisense primer, 20μL 5x PCR buffer, 0.2 mM of each dNTP and 0.5 μL of PrimeSTAR HS DNA polymerase. The PCR parameters are shown in Table 2.8.

The PCR products were analysed on a 2% agarose gel (section 2.2.1.1). The CH1γ products were pooled together and the Cκ products were combined into another pool. The pooled PCR products were purified using spin-columns as per the manufacturer’s instructions (Macherey Nagel).
Table 2.8 PCR parameters for the amplification of the human C_H1γ and C_κ regions.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>98°C</td>
<td>3 min</td>
</tr>
<tr>
<td>20 cycles of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>98°C</td>
<td>10 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>56°C</td>
<td>30 s</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>60 s</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>10 min</td>
</tr>
</tbody>
</table>

2.2.5.6 Assembly of chimeric mouse/human heavy and light chains via overlap extension PCR

To create chimeric heavy (V_H-linker-C_H1γ) and light (V_L-linker-C_κ) chains, overlap extension PCR was performed with the pooled V_H and V_L products and the corresponding human constant regions. The primer combinations are detailed in appendix A. Ten 100μL PCR reactions were performed for each of the heavy and light chains per mouse. These consisted of: 100 ng of mouse variable region product (V_L or V_H), 100 ng of human constant region product (C_κ or C_H1γ), 10 μM sense primer, 10 μM antisense primer, 20 μL 5X PCR buffer, 0.2 mM of each dNTP and 0.5μL of PrimeSTAR HS DNA polymerase. The PCR parameters are shown in Table 2.9.
Table 2.9 Parameters for overlap extension PCR to generate chimeric mouse/human heavy and light chains.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>98°C</td>
<td>3 min</td>
</tr>
<tr>
<td>15 cycles of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>98°C</td>
<td>15 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>56°C</td>
<td>30 s</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>90 s</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>10 min</td>
</tr>
</tbody>
</table>

The PCR products were analysed on a 1.5% agarose gel (section 2.2.1.1). The light chain products were pooled together (in a separate pool for each mouse) and the heavy chain products were combined into another pool (in a separate pool for each mouse). The pooled products were then purified by ethanol precipitation and gel extraction as described in section 2.2.5.7.

2.2.5.7 DNA purification by ethanol precipitation and gel extraction

In order to concentrate the large volumes of the DNA products, ethanol precipitation was performed. Sodium Acetate (3 M pH 5.2) was added at 1/10 of the DNA volume to give a final concentration of 0.3 M Sodium Acetate. 2.5 volumes of 100% ethanol (Merck) were added and the mixture was incubated for 2 hours at 4°C. The mixture was then centrifuged at 14,000 g for 45 min at 4°C. The supernatant was removed and the DNA pellet was washed with -20°C 70% ethanol, then centrifuged at 14,000g for 15 min at 4°C. The supernatant was removed and the pellet was air-dried. The pellet was resuspended in 20 µL of MQ.
The concentrated DNA products were separated on an agarose gel (1% - 1.5%) and the desired band was cut from the gel (visualised with a Safe Imager Blue Light Transilluminator, Thermo Fisher Scientific). The DNA was extracted from the gel fragments using a NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) as per the manufacturer’s instructions, and eluted in 20 µL MQ.

2.2.5.8 Assembly of the Fab constructs via secondary overlap extension PCR

A final overlap extension PCR was used to combine the previously amplified heavy and light chains to generate chimeric Fab fragments. Ten 100µL PCR reactions were performed for each of the mice. These consisted of: 100 ng of light chain product, 100 ng of heavy chain product, 10 µM RSC-F sense primer, 10 µM dp-EX antisense primer, 20 µL of 5X PCR buffer, 0.2 mM of each dNTP and 0.75 µL of PrimeSTAR HS DNA polymerase. The primer sequences are shown in appendix A. The PCR parameters are shown in Table 2.10.

Table 2.10 Parameters for overlap extension PCR to generate chimeric mouse/human Fab constructs.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>98°C</td>
<td>3 min</td>
</tr>
<tr>
<td>15 cycles of:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>98°C</td>
<td>15 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>56°C</td>
<td>30 s</td>
</tr>
<tr>
<td>Extension</td>
<td>72°C</td>
<td>3 min</td>
</tr>
<tr>
<td>Final extension</td>
<td>72°C</td>
<td>10 min</td>
</tr>
</tbody>
</table>
The PCR products were analysed on a 1% agarose gel (section 2.2.1.1). The chimeric Fab products were pooled together (separate pool for each mouse). The pooled products were purified by ethanol precipitation and gel extraction as described in section 2.2.5.7.

2.2.5.9 Restriction digestion of Fab and pComb3XSS for cloning

At this point, 2μg of the chimeric Fab fragments from all five mice were pooled so that the final phagemid library would contain antibody genes from all five mice equally. The Fab constructs and pComb3XSS phagemid were both digested with Sfi1 (Roche). The total volume of the reaction was 200 μL. The reaction mixture is shown in Table 2.11. The digestions were incubated for five hours at 50°C with periodic mixing. The digested Fab and pComb3X were separated on 1% agarose (section 2.2.1.1). The bands corresponding to the digested Fab and pComb3X (1600 bp and 3400 bp respectively) were cut from the gel and purified as described in section 2.2.5.7.

<table>
<thead>
<tr>
<th>Table 2.11 Restriction digestion of Fab and pComb3XSS with Sfi1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Template (μg)</strong></td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>Sfi1 (U)</td>
</tr>
<tr>
<td>10x buffer M (Roche) (μL)</td>
</tr>
</tbody>
</table>

2.2.5.10 Ligation of the Fab constructs into pComb3X to generate the Fab-phagemid library

Three large-scale ligations were performed to generate a phagemid library with a theoretical size of 1.47x10^7 Fab clones. Each of the three ligation reactions (total volume 200 μL) contained: 1400 ng Sfi1 digested pComb3X, 650 ng Sfi1 digested Fab, 40 μL 5x T4 ligase buffer (Invitrogen) and 1 U T4 DNA ligase. The reaction was incubated for 16 hours at RT.
The ligated Fab-pComb3X (Fab-phagemid) library was ethanol precipitated as described in section 2.2.5.7 and stored at -20°C.

2.2.5.11 Preparation of M13KO7 helper phage

A 5 mL culture of XL1-blue *E. coli* was grown in 2xTY supplemented with 10μg/ml tetracycline (2xTY/TET) overnight at 37°C, 200 rpm. 1mL of the overnight culture was diluted in 100mL 2xTY/TET and grown at 37°C, 200 rpm until the O.D$_{600}$ reached 0.6. 75 μL of M13KO7 helper phage (Invitrogen) was added to the culture which was incubated for one hour at 37°C, 200 rpm. Kanamycin was added to a final concentration of 70 μg/mL and the culture was incubated for 6 hours at 37°C, 200 rpm.

The culture was pelleted at 4000 g for 20 min at 4°C. The helper phage-containing supernatant was collected and incubated at 65°C in a water bath for 15 minutes to lyse the remaining cells. The cell debris was pelleted by centrifugation at 4000 g for 20 min at 4°C. The supernatant was collected and Phage-PEG solution was added to a final concentration of 20%. The phage were precipitated for 2 hours on ice then pelleted at 4,000 g for 45 min at 4°C. The phage were resuspended in PBS and stored at -80°C (with 15% Glycerol).

2.2.5.12 Titering M13KO7 helper phage

A 3 mL culture of TG1 *E. coli* was grown in 2xTY medium (1.6% bacto-tryptone, 1% Yeast extract, 0.5% (w/v) NaCl) at 37°C, 180 rpm until the O.D$_{600}$ reached 1.0. Once this O.D$_{600}$ was reached, 75 μL aliquots of TG1 *E. coli* were added to 4 mL of melted (45°C) 2xTY top agar (2xTY supplemented with 0.7% agar), mixed thoroughly and poured on top of 2xTY agar plates (2xTY supplemented with 1.5% agar).

Once these had air dried, 10 μL drops of phage that had been serially diluted in PBS to 1x10$^{-4}$, 1x10$^{-6}$, 1x10$^{-8}$ and 1x10$^{-10}$ were spotted in triplicate onto the TG1-containing top agar and left
to air dry. The plates were incubated for 16 hours at 37°C. The plaques in each 10 µL spot were counted and averaged and the titer was calculated by multiplying the number of colonies by the dilution factor and expressed as PFU/mL.

2.2.5.13 Transformation of the Fab-phagemid library into XL1-Blue E. coli

The Fab-phagemid library was spin-column purified (Machery Nagel) before being transformed into 900 µL of electrocompetent XL1-Blue E. coli. In each transformation (18 total) 200 ng of the Fab-phagemid library was added to 50 µL XL1-Blue E. coli and transferred to a 0.2 cm–gap electroporation cuvette (Bio-Rad). This mixture was incubated on ice for 1 min and then electroporated (2.5 kV, 25 µF, 200 Ω and τ ~4.0 ms) using a MicroPulser Electroporator (Bio-Rad). The cuvette was immediately flushed with 800 µL of prewarmed SOC medium (37°C). The cells were combined in batches of 6 such that there were 3x 50 mL tubes each containing 5.1 mL of transformed culture.

The 3 tubes were incubated for 1 hour at 37°C, 220 rpm. 10 mL of prewarmed SB medium (37°C), 3 µL of 100 mg/mL ampicillin (AMP-100) and 7.5 µL of 20 mg/mL tetracycline (TET-20) were added to each tube. To titer the library, 2 µL of culture was taken from each tube and pooled. This was diluted and plated as described below.

The 3x 15 mL cultures were incubated for 1 hour at 37°C, 220 rpm. 4.5 µL of AMP was added to each tube and they were incubated for an additional hour at 37°C, 220 rpm. Following this incubation, 500 µL of M13KO7 helper phage (1.4x10^{13} PFU/mL) (made in 2.2.5.11) was added to each of the three 50 mL tubes and then each of the three cultures were transferred into individual 500 mL flasks. 183 mL of prewarmed SB medium (37°C), 92.5 µL AMP-100 and 92.5 µL TET-20 were added to each flask and the cultures were incubated for 2 hours at 37°C, 220 rpm. 280 µL of kanamycin 50 mg/mL (KAN-50) was added to each flask and then they were incubated for 16 hours at 37°C, 220 rpm.
To titer the library, the pooled 6 µL (taken above) was diluted 1:100 in SB medium then 10 µL and 100 µL aliquots were plated onto LB+AMP plates (LB media supplemented with 1.5% agar and 100µg/mL of ampicillin). The plates were incubated for 16 hours at 37°C.

The total number of transformants was calculated using the following formula:

\[
\text{Number of colonies x culture volume (µL)} / \text{plating volume (µL)}
\]

2.2.5.14 Purification of Fab-phagemid particles by PEG precipitation

The 3 overnight cultures were pooled and the Fab-phagemid library was harvested by centrifugation for 15 min at 3,000 g, 4°C. The pellet was saved and stored at -20°C.

To precipitate the Fab-phagemid particles, the phagemid-containing supernatant was transferred to a clean centrifuge bottle and mixed with Phage-PEG solution (20% (w/v) PEG-6000, 2.5M NaCl) to a final volume of 20% then incubated on ice for 2 hours. After the incubation, the precipitated phagemid particles were pelleted by centrifugation for 45 min at 4,000 g, 4°C. The supernatant was removed and the phagemid particles were resuspended in 2 mL PBS. The resuspended phagemid particles were transferred to 1.5 mL tubes and centrifuged at 13,000 g for 15 min to pellet cell debris. The phagemid-containing supernatant was removed and transferred to clean tubes then spun again at 13,000 g for 10 min to remove the last of the cell debris. The supernatant was divided into 0.5 mL aliquots and either stored at -80°C or at 4°C (with 1% EDTA).

2.2.6 Biopanning the phagemid library and analysis of the selected clones

The pil18 mouse phagemid library was panned against T18.1-biotin bound to streptavidin coupled M280 Dynabeads (Invitrogen) in four biopanning cycles. The conditions used in each round of biopanning are summarised in Table 2.12. In the first round, T18.1-biotin was added to 0.5% casein to a final concentration of 1400 nM in 500 µL. This solution was added to 70
μL of M280 Dynabeads (Thermo Fisher Scientific) that had been pre-washed with 1 mL PBS. To allow the binding of the T18.1- biotin to the streptavidin coated M280 Dynabeads, the mixture was incubated on a rotating wheel for 1 hour at RT. Following this incubation, the M280 Dynabeads were washed 3 times in 1 mL PBS to remove any unbound T18.1-biotin.

To block the Fab-phagemid library, 200 μL of the library (9.4E+11 PFU/mL) was diluted to 500 μL in PBS and blocked with 125 μL MPBS (skim milk diluted in PBS) to a final concentration of 2%. This mixture was incubated on a rotating wheel for 1 hour at RT. In rounds 2-4, M280 Dynabeads (5 μg) were added at this step to deselect for Fab-phagemid that bound to streptavidin.

Following incubation, the blocked library was added to the washed T18.1-coated M280 Dynabeads and incubated on a rotating wheel for 2 hours at RT. The beads were then washed in PBST and PBS (Table 2.12) to remove unbound Fab-phagemid particles. The bound Fab-phagemid particles were then eluted with 350 μL of 0.2 M glycine-HCl pH 2.2 on a rotating wheel for 5 min at RT. The pH of the eluted phage was neutralised with 20 μL of 2 M Tris pH 9.0.

The eluted Fab-phagemid particles were used to infect 4 mL of log-phase XL1-blue *E. coli* (O.D<sub>600</sub> = 0.7) for 15 min at RT. To amplify the phagemids, 4 mL of pre-warmed SB media (37°C), 1.6 μL AMP-100 and 3 μL TET-20 were added to the cells. At this point, 2μL of the culture was taken to titer the output phage as described in section 2.2.6.1. The culture was then incubated for 1 hour at 37°C, 180 rpm. Following this incubation, 2.4 μL of AMP-100 was added and then the culture was incubated for a further 1 hour at 37°C, 180 rpm.

The phage were rescued with 80 μL of M13KO7 helper phage (1.9x10<sup>13</sup> PFU/mL) and the culture was transferred to a 500 mL flask. 91 mL of pre-warmed SB media (37°C), 96 μL AMP-100 and 46 μL TET-20 were added. The culture was incubated for 30 min at 37°C, 200
rpm. Following this incubation, 140 µL of KAN-50 was added and the culture was incubated for 16 hours at 37°C, 200 rpm. The phage from the overnight culture were purified by PEG precipitation as described in 2.2.5.14. A final volume of 1 mL purified phage was obtained. 500 µL of this was stored at 4°C (1% EDTA) for use in the next round of biopanning (as the input phage) and phage ELISAs.

Table 2.12 Biopanning conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Panning round</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Antigen (nM)</td>
<td>1400</td>
</tr>
<tr>
<td>Antigen volume (µL)</td>
<td>500</td>
</tr>
<tr>
<td>M280 Beads volume (µL)</td>
<td>70</td>
</tr>
<tr>
<td>Phage library blocking</td>
<td>0.5% Casein</td>
</tr>
<tr>
<td></td>
<td>2% Milk</td>
</tr>
<tr>
<td></td>
<td>5µg M280 Beads</td>
</tr>
<tr>
<td>PBST washes</td>
<td>6</td>
</tr>
<tr>
<td>PBS washes</td>
<td>2</td>
</tr>
</tbody>
</table>
2.2.6.1 Titering input and output phage for biopanning

To titer the input phage for the next round of biopanning, 50 μL of log-phase XL1-blue *E. coli* were infected with phage that had been serially diluted in PBS to $10^{-7}$, $10^{-8}$ and $10^{-9}$. The infected cells were spread onto 2xTYAG-agar plates and incubated for 16 hours at 37ºC. The input titer was calculated using the following formula: Number of colonies x phage dilution factor / plating volume.

To titer the output phage (the phage selected in that round of biopanning), the 2 μL sample that was taken in section 2.2.6 was diluted in 200 μL of SB media and plated on 2xTYAG. The plate was incubated for 16 hours at 37ºC. The output titer was calculated using the following formula: Number of colonies x dilution factor x culture volume/ plating volume.

2.2.6.2 Polyclonal phage ELISA

To ensure that there had been enrichment of T18.1-binding Fab-phage clones, the polyclonal output phage from each round of biopanning was screened for reactivity to T18.1 by ELISA. A Nunc immunoplate (Thermo Fisher Scientific) was coated with recombinant T18.1 (5 μg/mL), recombinant SPy_0136 (5 μg/mL) and BSA (1%) in PBS, pH 7.4 for 1 hour at RT. The plate was washed twice with PBST and then blocked with MT for 16 hours at 4°C. The plate was then washed twice with PBST. The purified Fab-phage from each round of biopanning were diluted 1:4 in MT were added to the plate and incubated for 1 hour at RT. Following this incubation the plate was washed 5 times with PBST. The binding of the Fab-phage to the antigens was detected with an Anti-M13-HRP (GE Healthcare) diluted 1:5000 (in MT) for 1 hour at RT. The plates was washed 4 times with PBST then developed as described in section 2.2.7.1.
2.2.6.3 Small-scale phage rescue of individual clones from the phagemid library

Single XL1-blue *E. coli* colonies were selected from the output titer plates following the fourth round of biopanning and added to the wells of two 96 deep-well plates (Eppendorf) containing 500 µL of 2xTYAG. The plates were incubated for 16 hours at 37°C, 800 rpm using a Thermomixer (Eppendorf).

Following incubation, a 5 µL sample was taken from each of the wells and added to new plates containing 500µL of 2xTYAG/well. The plates were incubated on a Thermomixer at 37°C at 800 rpm for 3 hours until the O.D$_{600}$ reached 0.5. M13KO7 helper phage was added to each well at a multiplicity of infection (MOI) of 20, and the plate was incubated undisturbed for 30 min at 37°C. The cells were pelleted by centrifugation for 5 min at 4000 g and the supernatant was removed. The pellets were resuspended in 500 µL/well of 2xTY media, supplemented with AMP-100 and KAN-50 and incubated on a Thermomixer for 16 hours at 30°C, 800 rpm.

The cells were pelleted by centrifugation for 5 min at 4000 g and the supernatant was removed and stored at 4°C for use in monoclonal phage ELISAs.

2.2.6.4 Monoclonal phage ELISA

Nunc Immunoplates were coated and blocked as described in section 2.2.6.2. The plates were washed twice with PBST. Phage-containing supernatant from the small-scale phage rescue (section 2.2.6.3) was diluted 1:4 in MT and incubated on the plates for 1 hour at RT. The assay was then continued as described in section 2.2.6.2.

2.2.6.5 Assessing the diversity of the Fab library

The 94 clones identified as being the top binders by monoclonal phage ELISA (section 2.2.6.4) were Sanger sequenced to assess their diversity. The clones were miniprepped using a Zippy-96 Plasmid Kit (Zymo Research) as per the manufacturer’s instructions. The purified plasmids
were then Sanger sequenced (Macrogen Inc) using the Ompseq primer to sequence the variable region of the light chain and the Pelseq primer to sequence the variable region of the heavy chain. Primer details are shown in appendix A. The sequences were analysed and assembled into full Fab fragments using Geneious (Biomatters Limited).

2.2.7 Immunoassays

2.2.7.1 Enzyme-linked immunosorbent assay-general method

The protein of interest was coated onto Nunc immunoplates at 5 µg/mL in PBS, pH 7.4 for 16 hours at 4°C. All incubations took place in a humidity chamber. Plates were washed 3 times with PBST and blocked with 100 µL of MT for 1 hour at RT. Plates were washed 3 times with PBST and the primary antibody (e.g. Fab or animal serum) was added, diluted in MT, for 1 hour at RT. Plates were washed 3 times in PBST, and the secondary antibody (e.g. Anti-human IgG-HRP or Anti-mouse IgG-HRP) was added, diluted in MT, and incubated for a further hour at RT. After a final 3 washes in PBST, 50 µL of 3, 3’, 5, 5’-Tetramethylbenzidine (TMB, Sigma-Aldrich) was added as substrate for the HRP. The reaction was stopped with 50 µL 1 M hydrochloric acid (HCl) and the absorbance was read at 450 nm using an EnSpire plate reader (PerkinElmer).

2.2.7.1.1 ELISA to determine serum end point titers

Recombinant T18.1 was coated onto Nunc immunoplates at 5 µg/mL in PBS, pH 7.4 for 16 hours at 4°C. The plates were washed three times with PBST and blocked with 100 µL MT for 1 hour at RT. Sera from the vaccinated mice and rabbit were added and serially diluted 5-fold in MT beginning at 1:100. The plates were incubated for 1 hour at RT then washed 3 times in PBST. To detect the binding of antibodies in the mouse serum to T18.1, an Anti-mouse IgG-HRP (GE Healthcare), diluted 1:5000 in MT, was added for 1 hour (RT). To detect the binding of antibodies in the rabbit serum to T18.1, an Anti-Rabbit IgG-HRP (Abcam), diluted 1:5000,
was added for 1 hour (RT). The plates were then washed and developed as in section 2.2.7.1. The end point titer was defined as the highest dilution that gave an optical density (OD) reading of more than three standard deviations (SD) above the mean OD of control wells containing unvaccinated mouse sera (Batzloff et al., 2004).

2.2.7.1.2 ELISA to estimate affinity

Recombinant T18.1 was coated onto Nunc immunoplates at 5 µg/mL in PBS, pH 7.4 for 16 hours at 4°C. The plates were washed three times with PBST and blocked with 100 µL MT for 1 hour at RT. The Fab and IgG1 were added in a 3-fold dilution series beginning at 1000 nM (diluted in MT) and incubated for one hour at RT. The plates were washed 3 times in PBST and an Anti-human IgG (H+L)-HRP (Jackson Immunoresearch), diluted 1:2500 in MT, was added for 1 hour (RT). The plates were then washed and developed as in section 2.2.7.1.

2.2.7.1.3 ELISA to determine T antigen cross-reactivity

The 14 recombinant T antigens in the panel were coated onto Nunc immunoplates at 5 µg/mL in PBS, pH 7.4 for 16 hours at 4°C. The plates were washed three times with PBST and blocked with 100 µL MT for 1 hour at RT. To determine the cross-reactivity of the Fab the 18 recombinant Fab were diluted to 100 nM in MT and added for 1 hour at RT. To determine the cross-reactivity of the vaccinated animal sera the mouse and rabbit sera were diluted 1:200 in MT and added for 1 hour at RT. The plates were washed 3 times in PBST. The binding of the Fab and animal sera was detected with the species-appropriate secondary antibody which was added for 1 hour (RT). The plates were then washed and developed as in section 2.2.7.1.

2.2.7.1.4 Competition ELISA: Fab vs IgG1

The 4 IgG1 clones were coated onto Nunc immunoplates at 5 µg/mL in PBS, pH 7.4 for 16 hours at 4°C. The plates were washed three times with PBST and blocked with 100 µL MT for
1 hour at RT. Recombinant T18.1 was added at 5 µg/mL (diluted in MT) for 1 hour at RT. The plates were washed three times with PBST. The Fab were diluted to 1 nM (αE3 and αH3), 10 nM (αE1) or 100 nM (βC1) in MT and added for 1 hour at RT. The plates were washed three times with PBST and then the binding of the Fab to T18.1 was detected with an Anti-HA-HRP antibody (Roche), diluted 1:2500 in MT, for 1 hour at RT. The plates were then washed and developed as in section 2.2.7.1.

2.2.7.1.5 Competition ELISA: Fab vs peptide array

Recombinant T18.1 was coated onto Nunc immunoplates at 5 µg/mL in PBS, pH 7.4 for 16 hours at 4°C. The plates were washed three times with PBST and blocked with 100 µL MT for 1 hour at RT. The plates were washed three times with PBST. The 4 Fab and peptides 10-15 of the overlapping peptide array were diluted in MT and incubated with each other at a ratio of 1:5 (Table 2.13) for 30 min at RT. Total competition was shown by pre-incubating the 4 Fab with full-length T18.1 at the same concentrations and ratio. For the ‘no competition’ control, the 4 Fab were not pre-incubated with either T18.1 or the peptides. The Fab-peptide mixtures, positive control and ‘no competition’ control were added to the plates for 1 hour at RT. The plates were washed 3 times with PBST and the binding of the Fab to T18.1 was detected with an Anti-human IgG (H+L)-HRP, diluted in MT, for 1 hour at RT. The plates were then washed and developed as in section 2.2.7.1.

Table 2.13 The concentrations of the Fab and peptides used in the competition ELISA.

<table>
<thead>
<tr>
<th>Fab</th>
<th>Fab concentration (nM)</th>
<th>Peptide concentration (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>αE3</td>
<td>25</td>
<td>125</td>
</tr>
<tr>
<td>αH3</td>
<td>25</td>
<td>125</td>
</tr>
<tr>
<td>αE1</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td>βC1</td>
<td>100</td>
<td>500</td>
</tr>
</tbody>
</table>
2.2.7.1.6 Competition ELISA: Fab vs vaccinated animal sera

Recombinant T18.1 was coated onto Nunc immunoplates at 5 µg/mL in PBS, pH 7.4 for 16 hours at 4°C. The plates were washed three times with PBST and blocked with 100 µL MT for 1 hour at RT. Sera from the vaccinated mice and rabbit were diluted 1:200 in MT and incubated for 1 hour at RT. For the ‘no competition’ control, 1% BSA was added instead as a non-T18.1 binding serum. The plates were washed 3 times with PBST. The Fab were diluted to 1 nM (αE1 and αH3), 10 nM (αE1) or 100 nM (βC1) in MT and added for 1 hour at RT. The plates were then washed three times with PBST and the binding of the Fab to T18.1 was detected with an Anti-human IgG (H+L)-HRP, diluted 1:2500 in MT, for 1 hour at RT. The plates were then washed and developed as in section 2.2.7.1.

2.2.7.2 Western blot to determine if Fab bind to linear epitopes on T18.1

A 200 ng sample of T18.1 was run on reducing SDS-PAGE gels then transferred onto polyvinylidene fluoride (PVDF) membranes using a Bio-Rad wet transfer apparatus (100V, 60 min) in Towbin transfer buffer (192 mM glycine, 0.37% (w/v) SDS, 20% (v/v) methanol, 25 mM Tris HCl pH 8.3). The membranes were blocked in MT for 16 hours at 4°C then probed with Fab diluted to 10 nM (in MT) for 1 hour at RT. The membranes were washed 4 times with PBST for 15 min and probed with an Anti-human IgG (H+L) (Jackson Immunoresearch) diluted 1:2500 (in MT) for 1 hour at RT. The membranes were washed 4 times with PBST for 15 min. Binding to denatured T18.1 was detected using the ECL western blotting analysis system (GE Healthcare) and imaged using a ChemiDoc Touch (Bio-Rad).

2.2.7.3 Dot Blot to determine if Fab bind to conformational epitopes on T18.1

Recombinant T18.1 was spotted onto nitrocellulose membrane (0.45 µm pore size) (GE Healthcare) in 2 µL drops containing 200 ng of protein. The membranes were air dried for 10
min. The membranes were blocked in MT for 16 hours at 4°C then probed, washed and developed as in section 2.2.7.2.

2.2.7.4 Dot Blot to map the epitopes of the Fab using the overlapping peptide library

The 29 peptides in the overlapping peptide array were spotted onto nitrocellulose membrane (0.45 µm pore size) (GE Healthcare) in 2 µL drops containing 200 ng of peptide. The membranes were air dried for 10 min. The membranes were blocked in MT for 16 hours at 4°C then probed with Fab diluted (in MT) to 100 nM (αE3 and αH3) or 500 nM (αE1 and βC1) for 1 hour at RT. The membranes were washed 4 times with PBST for 15 min and probed with an Anti-human IgG (H+L) (Jackson Immunoresearch) diluted 1:2500 (in MT) for 1 hour at RT. The membranes were washed and developed as in section 2.2.7.2.

2.2.8 Biosensor analysis of Fab binding to T18.1

The binding affinities of the αE3, αH3, αE1 and βC1 Fab for T18.1 were analysed using a Biacore T200 (GE Healthcare) at 25°C. Biotinylated T18.1 was immobilised to flow cell 4 of a SA sensor chip (GE Healthcare) as per the manufacturer’s instructions. Flow cell 3 (the reference cell) and flow cell 4 were conditioned with three 1 min injections of 1 M NaCl in 50 mM NaOH. Biotinylated T18.1 was then immobilised onto flow cell 4 of the SA chip at 6.75 nM in PBST (pH 7.4) at a density of 52 RU. The Fab were diluted in PBST to the concentration range 5-150 nM and were injected over the chip at a flow rate of 30 µl/min for 360 s. Dissociation was measured for 120 s. The surface was regenerated with 100 mM glycine pH 2 at 30 µl/ min for 30 s between each injection. Duplicate injections of each sample were performed in each experiment and experiments were repeated three times. The equilibrium binding response was fitted to a single site binding model using the Biacore T200 Evaluation software V 2.0, software, and the equilibrium dissociation constant (KD) was calculated for steady state affinity.
2.2.9 Flow cytometry analysis of IgG1 binding to the polymerised T18.1 pilus

The binding of the IgG1 to the polymerised T18.1 pilus expressed on the surface of *L. lactis*-pil18 was analysed by flow cytometry. *L. lactis*-pil18 were grown in GM17 media (3.725% (w/v) M17 powder, 0.5% (w/v) Glucose) for 16 hours at 30°C. Cells (1.5 mL) were harvested by centrifugation at 5000g for 5 min and resuspended in FACS blocking buffer (PBS/3% FBS/5mM EDTA) at an O.D$_{600}$ of 0.4. The cell suspension was sonicated in a water bath sonicator for 2 min then incubated on ice for 30 min. The blocked cells were harvested by centrifugation and resuspended in FACS buffer (PBS/1% FBS/5mM EDTA) at an O.D$_{600}$ of 0.4. Aliquots of cells (200 µL) were harvested by centrifugation and the supernatant was removed. The cells were resuspended in 100 µL FACS buffer containing 250 nM IgG1 and incubated at 4°C or 37°C for 30 min. The cells were washed with 1 mL FACS buffer and harvested by centrifugation. The cells were stained with 100 µL of Anti-Human IgG (H+L)-Alexa Fluor 488 diluted 1:1000 in FACS buffer for 30 min on ice. After a final wash in 1 mL FACS buffer, the cells were resuspended in 500 µL FACS buffer and transferred to FACS tubes. The binding of the IgG1 clones to polymerised T18.1 was analysed by flow cytometry using a LSR II flow cytometer (BD). Doublets were excluded and B9 Fab (a Fab that binds to T1 and does not bind to T18.1) was used as a negative control to set up the T18.1 positive gate.

2.2.10 In vitro protection assays

2.2.10.1 Opsonophagocytic killing assays

*S. pyogenes* strains were grown in Todd Hewitt broth supplemented with 1% Yeast extract (THY) at 37 °C, 5% CO$_2$ to mid-logarithmic phase (O.D$_{600}$= 0.6-0.7). Once mid-logarithmic phase was reached, the cultures were mixed 1:1 with STGG medium (3% (w/v) bacto-tryptone, 0.5% (w/v) D-Glucose, 10% (v/v) Glycerol) and stored in 0.5 ml aliquots at -80°C until required. Prior to performing the OPKA, the dilution of bacterial working stocks that resulted
in spots containing 100-200 CFU was determined for each strain. On the day of the assay, aliquots were rapidly thawed at 37°C in a water bath. The cells were then harvested by centrifugation and washed 3 times in HBSS with Ca/Mg (HBSS++) then diluted in opsonisation buffer (10% (v/v) FBS, 0.1% (w/v) Gelatin, in HBSS++) (OPS buffer) to the pre-determined dilution.

Serum from the T18.1 vaccinated rabbit was heat inactivated for 30 min at 56°C and 20 µL was added in duplicate to round-bottom microtiter plates (Greiner). The serum was serially diluted 3-fold in OPS buffer for a total of 8 dilutions. 10 µL of diluted S. pyogenes were added to each well. The plate was then incubated for 30 min on a mini-orbital shaker at 700 rpm at RT.

10 µL of BRC (diluted 1:2 in OPS buffer) was added to each well with 40 µL of differentiated HL-60 cells (0.8% DMF for 5/6 days, cultured and provided by Dr Marta Zancolli). The ratio between HL-60 cells and bacteria in the OPKA was approximately 500-2000 HL-60 cells to one bacterial cell. Two complement only controls (control A and control B) were included on each plate to calculate the level of non-specific killing in each assay. Control A contained bacteria, HL-60 cells and heat-inactivated complement only and control B contained bacteria, HL-60 cells and active complement only. The plates were incubated for 45-180 min on a mini-orbital shaker at 700 rpm at 37 °C, 5% CO₂.

After the second incubation, the plates were placed on ice for 30 min to stop the reaction. A 10 µL aliquot from each well was then plated onto THY-agar (THY supplemented with 1.5% (w/v) agar) and air dried. 20 mL of overlay agar (THY supplemented with 0.75% (w/v) agar and 0.0025% 2,3,5-tetraphenyltetrazolium chloride) was poured over each plate and allowed to set. Plates were then incubated for 16 hours at 37°C, 5% CO₂. The number of CFU per spot was enumerated using a ProtoCOL3 (Synbiosis) automated colony counter.
The percentage killing at each dilution of a sample was calculated as: 
\[ \frac{\text{CFU [control B]} - \text{CFU [antisera]}}{\text{CFU [control B]}} \times 100 \]. The dilution of the sample resulting in 50% killing was calculated as the opsonic index (OI).

The percentage of nonspecific killing in each assay was calculated as: 
\[ 1 - \frac{\text{CFU [control B]}}{\text{CFU [control A]}} \times 100 \]. Assays were accepted if levels of nonspecific killing were <35% and the CFU of both complement controls were between 50 and 200.

**2.2.10.1.1 OPKA to determine the specificity of killing**

The specificity of killing observed in the OPKA was determined by measuring the ability of homologous (T18.1) and heterologous (M1 and M6) proteins to inhibit the killing of bacteria by anti-T18.1 serum. Serum from the T18.1 vaccinated rabbit was heat inactivated for 30 min at 56°C. 1.5μg of recombinant T18.1, M1 or M6 (diluted in OPS buffer) was added to the serum and incubated for 30 min at RT. The assay was then performed as described in section 2.2.10.1.

**2.2.10.2 Bactericidal assays**

M217_11574 *S. pyogenes* were grown in THY at 37 °C, 5% CO2 to mid-logarithmic phase (\( \text{O.D}_{600} = 0.6-0.7 \)). Once mid-logarithmic phase was reached, the cultures were mixed 1:1 with STGG medium and stored in 0.5 ml aliquots at -80°C until required. On the day of the assay, aliquots of M217_11574 were rapidly thawed at 37°C in a Thermomixer (Eppendorf). The pellets were harvested by centrifugation (5000g for 5 min) then resuspended and diluted in HBSS++.

Heparinised whole blood was collected from healthy adult volunteers who had been prescreened by ELISA to confirm their nonimmune status, and was used within 2 hours of collection.
50uL of heat-inactivated serum from the vaccinated animals, or 50 µL of heat-inactivated BRC spiked with 1.5 µg of IgG1 (to give a final concentration of 5 µg/mL IgG1 in the assay), was added to 200 µL whole blood and 50uL M217_11574 (containing approximately 50 CFU). The mixture (total volume 300 µL) was incubated in a sealed tube for 3 hours at 37°C with constant rotation. Following incubation, 100 µL of the assay mixture was plated onto THY-agar and grown for 16 hours at 37°C with 5% CO2. The CFU were manually counted. Percentage killing was calculated by comparing the CFU after 3 hours growth in immune sera (or IgG1) with the CFU after 3 hours growth in pre-immune serum as follows: ((CFU [preimmune serum] - CFU [immune serum]) / CFU [preimmune serum]) x 100. Percentage killing greater than 50% was considered a positive result.


3 Structural analysis of T18.1

3.1 Introduction

*S. pyogenes* produce multi-protein pili that form long, thin fibrils and aid in adhesion. They usually consist of three protein components. The pilus shaft makes up the majority of the pilus and consists of repeated copies of a single monomer called the T antigen. The pilus is capped with Ancillary Protein 1 (AP1), which functions as an adhesin, and the whole structure is linked to the bacterial cell wall by the basal protein AP2 in most strains. The genes for these proteins are found within the variable Fibronectin-binding, Collagen-binding, T antigen (FCT) Pathogenicity Island, which also encodes the sortase enzymes required for pilus construction. The components of the pilus are polymerised by sortase C which is a transpeptidase and forms isopeptide bonds between a pilin lysine in the N-domain of one pilus component and the threonine in the QVPTG (LPXTG-like) sortase motif in the C-domain of the next. Pili in *S. pyogenes* can be up to 3 µm in length but are only one protein (2-3 nm) in width (Proft & Baker, 2009).

Due to the nature of the pilus backbone, T antigens are repeated tens or even hundreds of times per pilus. This makes them attractive vaccine candidates as they are highly immunogenic (Manetti et al., 2007; Raynes et al., 2016; Young, Moreland, et al., 2014) and any T antigen epitope will be presented to the immune system multiple times (Raynes et al., 2018). Systemic and mucosal vaccination with whole pili has been shown to be protective in mouse models (Mora et al., 2005) and the protective capacity has been largely attributed to the T antigen (Loh et al., 2017). This indicates that protective epitopes are present on T antigens but these are yet to be identified.
This project aims to use structural knowledge of T antigens to understand how functional antibodies interact with them at the molecular level. This includes the identification and characterisation of protective epitopes and their potential use in vaccine development. There is some variation among T antigens, with 18 tee types, but currently only two solved structures. These are T1 (FCT2) and T6 (FCT1) which are both genetically distinct from the majority of the other T antigens (FCT3/4/7/8) (Figure 1.4) and may not be structurally representative of most T antigens. As a first step in furthering our understanding of T antigen-host interactions, the structure of T18.1 (a representative FCT3/4/7/8 T antigen) has been solved using x-ray crystallography.

3.2 Results

3.2.1 Cloning of tee18.1

The tee18.1 gene was PCR amplified from a plasmid containing tee18.1 (kindly provided by Dr Loh) from the MGAS8232 strain (section 2.2.2.1). This was re-cloned to minimise additional amino acid residues at the C-terminus and so maximise the chance of success in crystal trials. The amplified DNA was cloned into pProEXHTa-Avitag using Kas1 and Xba1 restriction sites. The primers were designed with a C-terminal stop codon in order to prevent expression of the Avitag and ensure that this construct was in its native state for crystallography. Following sequence confirmation, pProEXHTa-tee18-Avitag was transformed using the heat shock method (section 2.2.2.2) into E. coli BL21 (DE3) cells (Invitrogen) for protein expression. This T18.1 construct approximates the mature T antigen before it is polymerised into the pilus. The N-terminal signal sequence has been removed and it ends at a conserved aspartic acid residue that immediately precedes the C-terminal sortase recognition site (RD-TQVPTG).
3.2.2 Expression and purification of T18.1

T18.1 was over-expressed in *E. coli* BL21 (DE3) cells using 2xTY media supplemented with ampicillin (100 µg/mL) (section 2.2.3.1). Cells were grown at 37°C until they reached exponential phase, then induced with 1 mM IPTG for 4 hours. The cell pellets were harvested by centrifugation and stored at -20°C.

T18.1 had an N-terminal polyhistidine (His$_6$) tag and was initially purified from cell lysate by Immobilised Metal Affinity Chromatography (IMAC) using a Nickel (Ni$^{2+}$) charged column (GE Healthcare) (section 2.2.3.3). The elution profile for T18.1 is shown in Figure 3.1.

![Figure 3.1](image)

**Figure 3.1** Immobilised metal affinity chromatography of T18.1 using a HisTrap HP column (GE Healthcare) charged with 100 mM Nickel Chloride. (A) Chromatogram of T18.1 elution. (B) SDS-PAGE analysis of the eluted fractions corresponding to the peak shown in (A). P = Protein mass standards, S = soluble fraction.

The fractions containing T18.1 (31 kDa) were pooled. The His$_6$ tag was removed by cleavage with rTEV-His$_6$ protease at a ratio of 1:50 to the pooled protein and T18.1 was then purified
from the cleaved His6 tags and rTEV-His6 using a second round of IMAC. T18.1 was further purified using Size Exclusion Chromatography (SEC) and was eluted in crystallisation buffer (10 mM Tris pH 8.0, 100 mM NaCl) (section 2.2.3.4). The SEC chromatogram is shown in Figure 3.2. Fractions judged to be >99% pure by SDS-PAGE were pooled.

![Size Exclusion Chromatography of T18.1 using a Superdex 200 10/300GL column (GE Healthcare). (A) Size exclusion chromatogram. (B) SDS-PAGE analysis of the fractions collected from the peak in (A). P = Protein mass standards.](image)

### 3.2.3 Crystallisation of T18.1

Crystallisation trials were carried out using a 480-condition screen developed at the University of Auckland (Moreland et al., 2005) in addition to the MORPHEUS screen (Gorrec, 2009). Vapour diffusion sitting drop trials were performed at 18°C using an Oryx4 crystallisation robot (Douglas Instruments) as described in section 2.2.4.1. T18.1 was concentrated to 140 mg/mL and 0.15 µL of this was mixed with 0.15 µL of the screen precipitant. T18.1 crystals formed in >100 conditions, with most being either plates or fine needles (Figure 3.3). Crystals
from 7 conditions were mounted on cryoloops (Hampton Research) and flash cooled in liquid nitrogen with cryoprotectant (20% glycerol) except for those selected from the MORPHEUS screen as the precipitant is a cryoprotectant (Gorrec, 2009).

![Figure 3.3 Typical crystal morphologies of T18.1 crystals. (A) Large plate-like crystals grown in 20% (w/v) PEG 3350, 0.2 M Potassium/Sodium Tartrate and 0.1 M Bis-Tris Propane pH 6.5. (B) Clusters of long thin needles grown in 10% (w/v) PEG 20000, 20% (v/v) PEG MME 550, 0.3 M Diethylene glycol, 0.3 M Triethylene glycol, 0.3 M Tetaethylene glycol, 0.3 M Pentaethylene glycol and 0.1 M Bicine/Trizma base pH 8.5.](image)

3.2.4 Structural analysis of T18.1

X-ray diffraction data were collected using the MX1 beamline at the Australian Synchrotron (section 2.2.4.2). Most of the samples diffracted to resolutions higher than 3 Å but one crystal (grown in 10% (w/v) PEG 20000, 20% (v/v) PEG MME 550, 0.02 M 1,6-Hexanediol, 0.02 M 1-Butanol, 0.02 M 1,2-Propanediol, 0.02 M 2-Propanol, 0.02 M 1,4-Butanediol, 0.02 M 1,3-Propanediol, 0.1 M Bicine/Tris base pH 8.5) diffracted to approximately 1.8 Å and a full dataset from this crystal was to be collected (Figure 3.4). However, the beam was lost towards
the end of collection, limiting the dataset to images covering the range 0° to 340° rather than 0° to 360°.

Figure 3.4 Data collection of T18.1. (A) The T18.1 crystal used to solve the structure is shown here during collection on the MX1 beamline at the Australian synchrotron. The white arrows indicate the ends of the crystal. (B) Representative diffraction pattern from this T18.1 crystal. There is greater diffraction in one dimension (black line) than the other (red line) indicating anisotropy in the dataset.

The dataset was indexed and integrated using XDS (Kabsch, 2010) and scaled using AIMLESS. The structure of T18.1 was solved by molecular replacement using PHASER-MR (McCoy, 2007) with the previously published T1 structure (PDB ID: 3B2M) as a search model (Kang et al., 2007). The structure underwent multiple rounds of manual building in COOT (Emsley et al., 2010) and refinement using Refmac5 (Murshudov et al., 2011). Final validation of the structure was performed with Molprobity (Davis et al., 2007) which scored the model in the top 100th percentile of structures of similar resolution. T18.1 was refined with data to 1.8 Å resolution, giving final R and R_free values of 0.23 and 0.27 respectively. These are
comparatively high for a structure of this resolution and this is likely due to the mildly anisotropic nature of the data or some undetermined pathology in the crystal lattice (Figure 3.4). In two dimensions the data were robust to 1.75 Å but only robust to 1.9 Å in the third as indicated by CC½ measurement. This phenomenon has been seen previously when solving the structures of *S. pyogenes* T antigens (Kang et al., 2007; Young, Moreland, et al., 2014) and is likely due to the long but thin nature of the T antigen proteins. T18.1 crystallised in space group P1 with two molecules in the asymmetric unit arranged antiparallel to each other. These molecules are fully modelled except for one loop in the N-domain of each monomer where the electron density was too poor to model for: residues 126-129 in molecule A and residues 124-130 in molecule B. This is likely due to flexibility in this loop, which is positioned at the top of the N-domain and would normally be at the interface between the N-domain of one T18.1 monomer and the C-domain of the next T18.1 monomer in the polymerised pilus. Because it is the more complete monomer, monomer A has been presented here and used in all analyses. Full data collection parameters and refinement statistics are shown in Table 3.1.
# Table 3.1 Data collection and refinement statistics for T18.1

<table>
<thead>
<tr>
<th>Collection Statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>0.95468</td>
</tr>
<tr>
<td>No. of images</td>
<td>340</td>
</tr>
<tr>
<td>Oscillation angle (°)</td>
<td>1.0</td>
</tr>
<tr>
<td>*Resolution range (Å)</td>
<td>44.83-1.75 (1.78-1.75)</td>
</tr>
<tr>
<td>*Total no. of observations</td>
<td>189630 (9186)</td>
</tr>
<tr>
<td>*Unique reflections</td>
<td>51303 (2572)</td>
</tr>
<tr>
<td>*Redundancy</td>
<td>3.7 (3.6)</td>
</tr>
<tr>
<td>Space group</td>
<td>P 1</td>
</tr>
<tr>
<td>Unit-cell axial lengths (Å)</td>
<td>a = 103.54 b = 93.53 c = 90.03</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>*Completeness (%)</td>
<td>96.7 (89.1)</td>
</tr>
<tr>
<td>*Mean I/σ(I)</td>
<td>9.4 (1.5)</td>
</tr>
<tr>
<td>*R_{merge} (%) †</td>
<td>0.091 (0.654)</td>
</tr>
<tr>
<td>*CC(1/2) ‡</td>
<td>0.995 (0.636)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>*Resolution range (Å)</td>
<td>40.54-1.75 (1.78-1.75)</td>
</tr>
<tr>
<td>Molecules per A.U.</td>
<td>2</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>40.16</td>
</tr>
<tr>
<td>R_{work}/R_{free} (%) ‡</td>
<td>23.4/27.7</td>
</tr>
<tr>
<td>Protein atoms</td>
<td>4567</td>
</tr>
<tr>
<td>Water molecules</td>
<td>354</td>
</tr>
<tr>
<td>RMS deviation from ideal bond length/angle (Å°)</td>
<td>0.010/1.361</td>
</tr>
<tr>
<td>Mean B-factor (Å²)</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>17.3</td>
</tr>
<tr>
<td>Water</td>
<td>19.8</td>
</tr>
<tr>
<td>Ramachandran Most Favoured (%)</td>
<td>98.53</td>
</tr>
<tr>
<td>Ramachandran Outliers (%)</td>
<td>0</td>
</tr>
</tbody>
</table>

*Numbers in parentheses for outermost shell. \( \phi \) Mn(I) half-set correlation CC(1/2) as calculated by Scala. 
†\( R_{merge} = \sum_{hkl} \sum_i |I_i(hkl)| - ⟨I(hkl)⟩)/\sum_{hkl} \sum_i |I_i(hkl)|. ‡ \( R_{work} \) and \( R_{free} = \sum |F_{obs}|-|F_{calc}|/\sum |F_{obs}| \), where \( R_{free} \) was calculated over 5% of amplitudes that were chosen at random and not used in refinement.
The T18.1 monomer is 90 Å long and 28 Å wide, at its widest. It consists of two domains each possessing an Ig-like fold. This fold comprises two β-sheets made up of antiparallel β-strands. Within each domain is an isopeptide bond (formed between the side chains of K15 and N153 in the N-domain, and K164 and N279 in the C-domain) that covalently links the first and the last β-strand of each domain. Each T18.1 monomer also contains one calcium ion in the C-domain, which coordinates with D253, D269, S271 (sidechain), S271 (backbone), D273 and two water molecules (Figure 3.5).

Figure 3.5 The structure of T18.1 presented as a ribbon diagram. The lysine and asparagine residues that form the isopeptide bonds in each domain are shown in blue and red, respectively. The pilin lysine (residue 146) that is predicted to polymerise with the next T18.1...
monomer is also shown (K). The calcium atom is represented as a grey sphere. The inset shows the coordination of the calcium atom and was generated using CheckMyMetal (Zheng et al, 2014).

Overall, the structure of T18.1 is remarkably similar to the previously published T1 (or Spy_0128) structure. This is particularly notable given that the amino acid identity between T18 and T1 is only 37% (Figure 3.6). Despite this low sequence identity, when the Cα backbone of the two proteins are overlaid the structures appear to have a similar core with the only major structural variation found in surface loops (Figure 3.7). The Root Mean Square Deviation (RMSD) calculates the average distance between the Cα atom positions in overlaid structures. The RMSD for T1 and T18.1 is only 1.11 Å over 258 residues indicating significant structural homology (calculated using PDBeFold (Krissinel & Henrick, 2005)). Furthermore, when the amino acid conservation and variation is mapped onto the structure of T1 it clearly shows that the residues in the hydrophobic cores of T1 and T18.1 are highly conserved (Figure 3.8). This suggests that there is some evolutionary pressure to maintain the core fold across genetically diverse T antigens. In contrast, the residues exposed on the surface and loops tend to be variable which may provide a form of passive immune evasion through antigenic diversity (Figure 3.9).
Figure 3.6 Sequence alignment of mature T18.1 and T1 proteins. The alignment was generated using ClustalW (Larkin et al., 2007) and ESPript (Robert & Gouet, 2014). Black columns indicate identical residues. Framed columns indicate that >70% of the framed residues have similar physico-chemical properties. Secondary structure elements were calculated from the structure of T18.1. Solvent accessibility (acc) is shown as a bar below the sequences. Darker shading indicates more accessible residues.
Figure 3.7 Structural overlay of the polypeptide backbone of T1 and T18.1. T1 (blue) and T18.1 (green) exhibit a conserved overall fold and core structure with significant variation being limited to the loop regions.
Figure 3.8 Identification of sequence conservation between T18.1 and T1 mapped onto the T1 structure. Blue represents identical residues in both T18.1 and T1. White represents conservative substitutions. Red represents non-conservative substitutions.
Figure 3.9 Visualisation of sequence conservation between T18.1 and T1 mapped onto the T1 surface. The T antigen surface consists of patches of sequence variation and conservation scattered along the length of the protein. Blue represents identical residues in both T18.1 and T1. White represents conservative substitutions. Red represents non-conservative substitutions.

3.3 Discussion

The structure of T18.1 is the third T antigen structure to be solved from *S. pyogenes*. T1 (PDB: 3B2M) and T6 (PDB: 4P0D) have been previously solved and shown to consist of two or three tandem IgG-like domains respectively (Kang et al., 2007; Young, Moreland, et al., 2014). Due to its size T18.1 was also predicted to be a two-domain T antigen, but because of the low
sequence identity between T1 (FCT2) and the T antigens from FCT 3, FCT 4, FCT 7 and FCT 8 its structure could not be predicted from sequence alone.

Despite a sequence identity of only 37% (Figure 3.6), the overall structure of T18.1 is almost identical to that of T1. The hydrophobic cores of the two proteins are highly conserved, as are the locations of the intra-domain isopeptide bonds and the pilin lysine involved in polymerisation. This suggests that there is functional pressure to retain this core fold across the genetically diverse T antigen family. Mutation of the residues involved in the formation of the isopeptide bonds has been shown to markedly decrease the mechanical, thermal, and proteolytic stability of pilus components (Alegre-Cebollada et al., 2010; Kang & Baker, 2009; Kang et al., 2007; Walden, Crow, Nelson, & Banfield, 2014) indicating that this fold is required to maintain the strength of the pilus. Given this apparent functional requirement, and the much higher sequence conservation amongst T18.1 and the other FCT3/4/7/8 T antigens, it is probable that all two domain T antigens share a similar overall structure.

Outside of the conserved core, T1 and T18.1 show a large degree of sequence variability. This variability is spread evenly across the surface of the protein (Figure 3.9) and is not confined to a single loop or domain. The development of antigenic diversity between strains is most likely driven by host immune pressure (Gupta, Ferguson, & Anderson, 1998; Gupta et al., 1996; Lipsitch & O'Hagan, 2007) as it protects strains from immunological memory responses directed against heterologous strains. If all T antigens were wholly conserved, then a protective immune response developed to a T antigen expressed by one strain would recognise all T antigens and could be protective against all strains. However, because T antigens display antigenic diversity across their surface, an immune response developed to a T antigen from one T type does not necessarily recognise a T antigen from a heterologous T type (Falugi et al., 2008; Manetti et al., 2007; Raynes et al., 2016; Young, Moreland, et al., 2014). Antigenic
diversity is commonly observed in bacterial surface proteins including the *S. pyogenes* M protein. Unlike the T antigen, however, variability in the M protein is localised to the N-terminal region of the protein. The first 50 residues comprise the hypervariable region (HVR) and this variability is the basis for M and *emm* typing of strains (see section 1.4.3.2). Antigenic diversity in this region is proposed to be driven by immunological pressure as the HVR is the target of protective antibodies (Jones et al., 2018; Lannergård et al., 2011a; Sandin et al., 2006). These antibodies are generally M type (Pandey, Ozberk, et al., 2016; Penfound et al., 2010) or potentially *emm* cluster specific (Dale et al., 2011; Dale et al., 2017; Frost et al., 2017; Sanderson-Smith et al., 2014) so diversification of the HVR prevents antibody binding of heterologous M types. This results in HVR-dependent protection being largely type-, or cluster-, specific. Antigenic diversity in surface proteins is by no means unique to *S. pyogenes* and is observed in many bacterial species (Croucher et al., 2017; Evans et al., 2010; Jones, Gordon, Hewinson, & Vordermeier, 2010; Nuccitelli et al., 2011; Soriani et al., 2010; Thompson, Feavers, & Maiden, 2003).

Type-specific immune responses have also been observed to T antigens and these form the basis of traditional T serotyping (section 1.4.4.4) (Falugi et al., 2008; Manetti et al., 2007; Raynes et al., 2016; Young, Moreland, et al., 2014). This uses commercial sera (which have been characterised as being type-specific or cross-reactive) raised against T antigens as a supplementary strain typing tool. T serotyping exploits the presence of type-specific antibodies in the sera which must bind to type-specific epitopes on T antigens. These are likely to be found in regions of variability which are located in loops and across the surface of T antigens (as shown in red and white in Figure 3.8 and Figure 3.9). Underlying the amino acid variation in these regions is genetic variation within the *tee* gene, and this is the foundation for *tee* typing which has replaced serological T typing just as *emm* typing has superseded M typing (Chochua et al., 2017; Falugi et al., 2008; Steemson et al., 2014).
Conversely, cross-reactive immune responses to T antigens have been observed in animals (Loh et al., 2017) and humans (Falugi et al., 2008; Raynes et al., 2016; Young, Moreland, et al., 2014) and the antibodies causing these must bind to cross-reactive epitopes. These are likely to be found in regions of surface conservation such as those shown in blue in Figure 3.9. Although cross-reactivity between T1 and T18 has not been shown (presumably due to their low sequence identity), T typing sera has revealed a few well characterised cross-reactivity patterns such as T3/T13 and these indicate the presence of cross-reactive epitopes on the surface of distinct T types (Johnson et al., 2006; Padula, Facklam, & Moody, 1969; W. A. Stewart, Lancefield, Wilson, & Swift, 1944). It is also possible that cross-reactive 3-dimensional epitopes may be present due to the conserved nature of the structures - while the linear sequence may be different, the fold could be similar enough for cross-reactivity.

Solving the structure of T18.1 has provided insight into the structural conservation of two domain T antigens and the structural basis for cross-reactive and type-specific immune responses to them. In chapter 4, T18.1 has been used as a vaccine to generate antibody responses to T antigens. In chapter 5, the epitopes of these antibodies on T18.1 are identified.
4 Generation and characterisation of T antigen antibodies from mouse-human phagemid libraries

4.1 Introduction

Infection with *S. pyogenes* activates both cellular and humoral adaptive immune responses in the host. Control of *S. pyogenes* largely relies on antibody-mediated opsonophagocytosis of extracellular bacteria by innate immune cells such as neutrophils (Hidalgo-Grass et al., 2006; Mortensen et al., 2015). During opsonophagocytosis, antibodies bind to their specific epitope (binding site) on the bacteria and induce clearance by stimulating phagocytosis and complement deposition. Early proteins of the complement cascade can also bind to the antibody-antigen complex and promote phagocytosis or the formation of the membrane-attack-complex and cell lysis (Esser, 1994; Raynes et al., 2018).

Antibodies are tetramers made up of two heavy chains and two light chains linked by a number of disulphide bonds. The variable regions in the heavy and light chains contain hypervariable complementarity determining regions (CDRs) which determine an antibody’s specificity and epitope. The portion of the antibody which binds to its antigen is the Fab fragment (Fragment antigen binding) while the portion of the antibody which binds to phagocytes and complement proteins is the Fc fragment (Fragment crystallisable) (Figure 1.6). There are five classes of antibody in humans: IgA, IgD, IgE, IgG and IgM with IgG being most common class in serum. There are four subtypes of IgG (IgG1-IgG4) that each have different effector properties. IgG1 is the most common subtype in serum, is highly opsonic (Brüggemann et al., 1987) and has been implicated in providing protection against *S. pyogenes* (Mortensen et al., 2015).
This project seeks to investigate how individual antibodies interact with T18.1 using phage display technology. Vaccination elicits polyclonal immune responses which can be ‘captured’ by constructing antibody libraries from immune animals. Phage display methodology allows for the isolation of T18.1-specific monoclonal antibodies from these large libraries. The libraries used in this project are chimeric mouse/human Fab libraries which fuse the variable genes from immunised mice to human constant regions. The libraries are cloned into a phagemid vector (pComb3XSS, Scripps research institute) such that the Fab constructs are fused to the gene for the pIII coat protein of M13 filamentous phage (Figure 4.1) (Rader & Barbas, 1997). The library is transformed into F’ E.coli which are then infected with helper phage via the F pilus. The helper phage have defective origins of replication which promotes the assembly of phage containing the Fab-phagemid instead of the helper phage genome. The resulting phagemid particles express 1-5 copies of pIII-Fab on their surface and can be repeatedly affinity purified with T18.1 in a process called biopanning. The Fab-phagemid particles that bind T18.1 are then used to infect F’ E.coli and produce another pool of Fab-phagemid that is enriched for T18.1 binders (Figure 1.7). At the end of this process, individual colonies of infected E. coli contain the genetic information for an individual Fab, allowing the recombinant expression of monoclonal Fab.

This chapter describes the process of constructing Fab-phagemid libraries, from immunised mice to the identification of unique clones, and the pipeline used to characterise and identify the most interesting Fab for further study.
Figure 4.1 Display of a Fab fragment on the pIII coat protein of M13 filamentous phage. The phage particle contains a pComb3X phagemid encoding a single Fab fragment that is fused to the pIII gene. Expression of pIII results in display of the Fab fragment on the surface of the phage(mid) particle. Figure based on (Barbas, Burton, Scott, & Silverman, 2001).

4.2 Results

4.2.1 Animal vaccinations

Prior to the beginning of this project Dr Loh had intranasally (I.N) vaccinated ten female FVBn mice with *Lactococcus lactis* constitutively expressing the T18.1 pilus operon (pil18) described in Loh et al. (2017) (Figure 4.2) and ten female FVBn mice with *L. lactis* carrying empty vector (section 2.2.5.1). They received two booster vaccinations at 14 and 28 days post immunisation and were euthanised on day 52. Blood was collected and processed to extract serum (section 2.2.5.2) and the spleens were harvested and snap frozen in liquid nitrogen.
A female New Zealand white rabbit was vaccinated with purified recombinant T18.1 mixed 1:1 with Incomplete Freund’s adjuvant (IFA) (Figure 4.2). It received three boosters at 14, 28 and 42 days post immunisation and was euthanised on day 56 (section 2.2.5.1). Blood was collected and processed to extract serum. The spleen and bone marrow (collected from the femurs) were harvested and snap frozen in liquid nitrogen. The serum from this rabbit is used extensively in this thesis but, due to the success of the Fab library generated from the pil18 immunised mice, the rabbit antibody library was not pursued further.

Figure 4.2 Schematic showing the T18.1 pilus and the two T18.1-containing vaccines used in this thesis. The New Zealand white rabbit was vaccinated with a vaccine containing monomeric recombinant T18.1 (left). The FVBn mice were immunised with *L. lactis* that constitutively express the full T18.1 pilus on their surface.

To confirm that the animals had developed a humoral immune response to T18.1, serum samples were screened by ELISA to determine their antibody endpoint titers. T18.1 was coated onto immunoplates and the sera were titrated to generate binding curves from which endpoint
titers were calculated (section 2.2.7.1.1). Each of the pil18 vaccinated mice had a raised antibody titer (>45,000) to T18.1 whereas the mice vaccinated with empty vector did not (Figure 4.3 A). Pre-immune sera from these animals was not available for analysis. The T18.1 vaccinated rabbit also had a raised antibody titer (>450,000) to T18.1 which was not present in the pre-immune serum (Figure 4.3 B). These titers indicated that the pil18 and T18.1 vaccinated animals had developed antibody responses to T18.1 and were suitable for further work. As the mice from cage B (mice B1-B5) had higher mean antibody titers than those from cage A, these mice were selected for constructing antibody libraries.

Figure 4.3 Antibody endpoint titres of vaccinated animal sera to T18.1. (A) Comparison of the α-T18.1 titre in serum from the pil18 vaccinated (Cages A and B) and sham vaccinated (Cages C and D) mice. (B) Comparison of the α-T18.1 titre in the T18.1 vaccinated rabbit before and after vaccination. Each point represents a single animal.

### 4.2.2 Preparation of genetic material from mouse spleens

Total RNA was isolated from the frozen spleens of mice B1-B5 (section 2.2.5.3). The spleens were disrupted by grinding them to a powder in liquid nitrogen and homogenised using
QIAshredders (Qiagen). The RNA was purified using an RNeasy Mini Kit (Qiagen) with an average yield of 40.5 μg/spleen. The sample purity was measured using the 260/280 nm absorbance ratio on a Nanodrop 2000 (ThermoFisher) with an average sample 260/280 of 2.07 where a ratio of 2.0 is deemed to be pure RNA.

First-strand cDNA was synthesised from the purified RNA using SuperScript III Reverse Transcriptase (ThermoFisher) primed with oligo(dT)20 primers (section 2.2.5.4).

4.2.3 Generation of chimeric mouse/human Fab libraries

This section describes the construction of a Fab library from each of the five pil18 vaccinated mice by overlap PCR. The variable regions of the Fab (which determine its binding properties) were amplified from vaccinated mouse cDNA then joined to constant regions derived from a human antibody to create chimeric Fab. The genetic material from each of the five mice was kept separate to maximise the potential library diversity and ensure that each mouse was represented in the final combined library. As such, the Fab fragments from the five mice were only pooled at the final stage of library construction when transforming into E. coli.

4.2.3.1 Amplification of mouse immunoglobulin variable genes

The V_L and V_H genes were PCR amplified from pil18 mouse cDNA using mouse V-specific primers. Nine V_κ amplifications, one V_λ amplification and twelve V_H amplifications were performed for each mouse (section 2.2.5.5). Figure 4.4 shows example amplifications of the V_κ and V_H genes from mouse B1 and the amplification of the V_λ genes from all five mice. All of the variable genes amplified equally well in all mice and were pooled by gene family, such that there was a pool of amplified V_κ genes, a pool of V_λ genes and a pool of V_H genes from each mouse.
Figure 4.4 PCR amplification of immunoglobulin variable genes from pil18 mouse cDNA. (A) Amplification of $V_\kappa$ from mouse B1. (B) Amplification of $V_\lambda$ from mouse B1-5 (lanes 1-5 respectively). (C) Amplification of $V_H$ genes from mouse B1. The expected product sizes for the $V$ genes are approximately 350bp. The amplifications shown from B1 are representative of all mice.

4.2.3.2 Amplification of human immunoglobulin constant genes

The constant region of the human $\kappa$ light chain ($C_\kappa$) and the first constant region of the human $\gamma$ heavy chain ($C_{H1\gamma}$) were PCR amplified from a human Fab cloned into the pComb3XTT phagemid (section 2.2.5.5). Ten amplifications of each constant region were performed which yielded sufficient DNA for the following overlap PCR (>5 μg) (Figure 4.5). The ten reactions were pooled and spin column purified.

Figure 4.5 PCR amplification of human constant regions from pComb3XTT. (A) Amplification of $C_\kappa$. (B) Amplification of $C_{H1\gamma}$. The expected product size for both reactions is approximately 400bp.
4.2.3.3  **Overlap extension PCR: round one**

The previously amplified mouse variable regions and human constant regions were combined to generate light-chain and heavy-chain fragments (section 2.2.5.6). $V_k$ and $V_h$ were pooled ($V_L$) and combined with $C_k$ while $V_H$ was combined with $C_H1_\gamma$. The primers in this overlap PCR bind to complementary sequences added to the ends of the amplified genes by the primers used in the initial amplification. Ten amplifications generating chimeric light-chain fragments and ten amplifications generating chimeric heavy-chain fragments were performed for each mouse. The major product in both reactions was a fragment of the expected size, however larger products were also amplified. Figure 4.6 shows the overlap extension PCR performed using genetic material from mouse B1. Similar banding patterns were observed for all mice. The products of the correct size for each fragment were purified by gel extraction and ethanol precipitation then pooled (section 2.2.5.7).

![Image](image_url)

**Figure 4.6** Overlap extension of variable and constant gene regions from mouse B1. (A) Overlap extension of $V_L + C_k$ producing a light-chain fragment of approximately 800bp (marked by arrow). (B) Overlap extension of $V_H + C_H1_\gamma$ producing a heavy-chain fragment of approximately 750bp (marked by arrow). The amplifications shown are representative of all mice.

4.2.3.4  **Overlap extension PCR: round two**

A final overlap extension PCR was used to combine the previously amplified light and heavy chains to generate chimeric Fab fragments (section 2.2.5.8). The sense primer binds to the 5’ end of the light-chain (the $V_L$ region) and the antisense primer binds to the 3’ end of the heavy-
chain (the C_H1γ region). The shared pelB leader sequence at the 3’ end of the light-chain and the 5’ end of the heavy-chain is the site of overlap. Ten amplifications generating chimeric Fab fragments were performed for each mouse. The major product was a fragment of the expected size (1500bp) but a larger product was also amplified. Figure 4.7 shows the overlap extension PCR performed using genetic material from mouse B1. Similar banding patterns were observed for all mice. The Fab fragments were purified by gel extraction and ethanol precipitation and then pooled.

Figure 4.7 Overlap extension of light-chain and heavy-chain fragments from mouse B1. The full chimeric Fab fragment is approximately 1500bp (marked by arrow). The amplifications shown are representative of all mice.

4.2.3.5 Ligation into pComb3X and transformation of the phagemid library into *E. coli*

At this point, the chimeric Fab fragments were pooled so that the final phagemid library would contain antibody genes from all five mice equally. The Fab constructs and pComb3XSS phagemid were digested with Sfi1 (Roche) (Figure 4.8) then purified by gel extraction and ethanol precipitation (section 2.2.5.9).

When the pComb3XSS phagemid is digested with Sfi1, a 1600bp stuffer fragment is removed. Following this, it is referred to as pComb3X. The Sfi1-digested Fab fragments can then be ligated into pComb3X at the Sfi1 restriction sites. When the Fab is cloned into pComb3X, there is an amber stop codon between the heavy chain constant region and geneIII of the phagemid.
This allows for the production of a Fab-geneIII fusion protein in suppressor E. coli and soluble Fab in non-suppressor E. coli. The soluble Fab are tagged with both an HA tag and a His6-tag.

![Figure 4.8 Sfi1 digestion of Fab and pComb3XSS. (Lane 1) Sfi1 Digested Fab fragments from pil18 mice. (Lane 2) Sfi1 digest of pComb3xSS showing the undigested vector (5000bp) the digested vector (3400bp) and the stuffer fragment (1600bp) released when pComb3XSS is digested.](image)

Small-scale test ligations and transformations into electroporation-competent XL1-Blue E. coli (Agilent) were performed to calculate the transformation efficiency and estimate the maximum complexity/library size (the number of unique antibodies) of the library. The transformation efficiency (and therefore the maximum complexity) was calculated as being $4.9 \times 10^6$ CFU/μg of pComb3X-Fab where $1.0 \times 10^7$-$1.0 \times 10^8$ clones is recommended for immune animal libraries (Barbas et al., 2001). In order to reach this library complexity, three large-scale ligations and library transformations were performed (section 2.2.5.13) to reach a theoretical library size of $1.47 \times 10^7$. In total, 4.2 μg of Sfi1 digested and purified pComb3X was ligated with 1.95 μg Sfi1 digested Fab insert using T4 DNA Ligase (Invitrogen). The pComb3X-Fab was spin column purified then electroporated into 900 μL XL1-Blue E. coli. A 6 μL sample of this total culture volume (45 mL) was diluted 1:100 and plated on 2xTY plates supplemented with 100μg/mL.
ampicillin and 2% glucose (2xTYAG) to calculate the final library size. The culture was rescued with M13KO7 (Invitrogen) helper phage and grown for sixteen hours. The phagemid particles that were produced were purified from the culture by PEG precipitation and titered (section 2.2.5.14, section 2.2.6.1).

The final library size was calculated to be 7.79x10^7 transformants which is approximately five times greater than estimated from small-scale testing and well within the range recommended for an immune library (1.0x10^7-1.0x10^8, (Barbas et al., 2001). The total number of phagemid particles in the phagemid library was calculated to be 1.88x10^{12} PFU therefore each transformant is represented approximately 2.4x10^5 times.

4.2.4 Preparation of biotinylated T18.1 for biopanning

In order to best preserve epitope integrity and the accessibility of the antigen, the Fab-phagemid particles were to be biopanned against T18.1 in solution. Solution biopanning allows the phage to be screened against the antigen in its native state. It prevents the partial denaturation of proteins, and therefore the potential loss of epitopes, which can occur when using plastic solid state biopanning and the orientation of the antigen can also be controlled (Schier et al., 1995). This project required the cloning and purification of a T18.1 construct featuring a C-terminal Avitag which could be specifically biotinylated and bound to streptavidin coated beads.

4.2.4.1 Cloning of tee18.1-Avitag

The tee18.1 gene was PCR amplified from a plasmid containing tee18.1 from the MGAS8232 strain and cloned into pProEXHTa-Avitag using Kas1 and Xba1 restriction sites. Following sequence confirmation, pProEXHTa-tee18.1-Avitag was co-transformed into E. coli BL21 (DE3) cells (Invitrogen) with pACYC184-BirA for protein expression and in vivo biotinylation (section 2.2.2.1).
4.2.4.2 Expression and purification of biotinylated T18.1.

T18.1 was over-expressed in *E. coli* BL21 (DE3) cells using 2xTY media supplemented with ampicillin (100 μg/mL) and chloramphenicol (25 μg/mL). Cells were grown at 37°C until they reached exponential phase, then induced with 1 mM IPTG and supplemented with 20 μM D-biotin for 4 hours. The cell pellets were harvested by centrifugation and stored at -20°C.

T18.1-biotin was initially purified from cell lysate by IMAC using a Ni²⁺ charged column (GE healthcare) (section 2.2.3.3). The elution profile is shown in Figure 4.9. The fractions containing T18.1-biotin were pooled. The His₆ tag was removed by cleavage with rTEV-His₆ at a ratio of 1:50, and T18.1-biotin was then purified from the cleaved His₆ tag and rTEV-His₆ using a second round of IMAC. T18.1 was further purified using SEC and eluted into Phosphate-buffered saline (PBS) (section 2.2.3.4). The SEC chromatogram is shown in Figure 4.9. Fractions judged to be >95% pure by SDS-PAGE were concentrated and pooled.
Figure 4.9 Purification of T18.1-biotin. (A) Immobilised metal affinity chromatography of T18.1-biotin using a HisTrap HP column (GE healthcare) charged with 100 mM Nickel chloride. (B) SDS-PAGE analysis of the eluted fractions. (C) Size Exclusion chromatography of T18.1-biotin using a superdex 200 10/300GL column (GE healthcare). (D) SDS-PAGE analysis of the fractions collected in the peak shown in (C). P= protein mass standard.

4.2.4.3 Estimation of the efficiency of biotinylation

Success of the in vivo biotinylation of T18.1-Avitag was confirmed with a pull-down on streptavidin-coupled M-280 Dynabeads (ThermoFisher). T18.1-biotin was added in excess of
the binding capacity of the beads in order to maximise binding. Following incubation, the beads were washed six times in PBS then analysed by SDS-PAGE (section 2.2.1.4) (Figure 4.10).

Figure 4.10 Pull-down of biotinylated T18.1 on streptavidin-coupled beads. Lane 1: The T18.1-biotin sample prior to pull-down. Lane 2: The contents of the first PBS wash following incubation. Lane 3: The contents of the sixth PBS wash following incubation. Lane 4: The T18.1-biotin sample bound to the washed beads.

The fluorescence intensities of the coomassie-stained band known to be biotinylated (Figure 4.10, lane 4) and a standard curve of known T18.1 amounts were measured at 680 nm using an Odyssey® CLx imager (Li-Cor biosciences). From this standard curve, the in vivo biotinylation of T18.1 was calculated to be 13% efficient. This was accounted for in downstream applications.

4.2.5 Biopanning the phagemid library against T18.1

The mouse pil18 phagemid library was panned against T18.1-biotin bound to streptavidin coupled M280 Dynabeads (Invitrogen) in four biopanning cycles (section 2.2.6). In the first round, 3.76x10^{11} PFU from the library were blocked with 2% Milk in PBS then incubated with M280 Dynabeads bound with T18.1 in 0.5% casein. These were washed six times with PBST and twice with PBS to remove the unbound phage. The bound Fab-phagemid particles were eluted with 0.2 M glycine-HCL (pH 2.2) and neutralised in 2M Tris (pH 9.0). These round one
Fab-phagemid particles were used to infect logarithmic phase XL1-blue *E. coli* then rescued with M13KO7 helper phage in order to amplify the number of eluted phage in the pool. In subsequent panning rounds, the concentration of T18.1 was reduced and the number of washing steps were increased to increase the stringency of selection. An additional blocking step was also added where the library was pre-incubated with M280 Dynabeads to deselect any bead-binding phagemid particles prior to biopanning (Table 4.1). Over the four rounds of panning the percentage of eluted phagemid particles ((Titred output/Titred input) x 100) increased >200 times indicating that positive selection for T18.1-binding phagemid particles had occurred.

To ensure that there had been an enrichment of Fab-phagemid clones that were specific for T18.1, the output pools of phage from each of the biopanning rounds were screened by ELISA (section 2.2.6.2). The Fab-phagemid were diluted 1:4 in PBS containing 5% Milk and 0.1% Tween-20 (MT) and incubated with the antigens. Fab-phagemid binding was detected with a secondary antibody recognising the VIII protein of the M13 phage coat (Anti-M13-HRP). The reactivity to T18.1 increased in the polyclonal pools of Fab-phagemid from each round of biopanning, whereas the reactivity to PBS, BSA and SPy-0136 (which control for non-specific binding of phage to plastic or protein) did not increase with each round (Figure 4.11). This indicates that the biopanning process has enriched for Fab-phagemid clones specific for T18.1.
Table 4.1 Summary of solution-phase biopanning against T18.1.

<table>
<thead>
<tr>
<th></th>
<th>Panning round</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>**Blocking and pre-</td>
<td>0.5%</td>
</tr>
<tr>
<td>incubation**</td>
<td>Casein</td>
</tr>
<tr>
<td></td>
<td>2% Milk</td>
</tr>
<tr>
<td><strong>Antigen (nM)</strong></td>
<td>1400</td>
</tr>
<tr>
<td><strong>M280 Beads volume (μL)</strong></td>
<td>70</td>
</tr>
<tr>
<td><strong>PBST washes</strong></td>
<td>6</td>
</tr>
<tr>
<td><strong>PBS washes</strong></td>
<td>2</td>
</tr>
<tr>
<td><strong>Library volume (mL)</strong></td>
<td>0.2</td>
</tr>
<tr>
<td><strong>Input</strong></td>
<td>9.4E+11</td>
</tr>
<tr>
<td><strong>Library concentration</strong></td>
<td>(PFU/mL)</td>
</tr>
<tr>
<td><strong>Titred input</strong></td>
<td>3.76E+11</td>
</tr>
<tr>
<td><strong>Output</strong></td>
<td>8.00E+4</td>
</tr>
<tr>
<td><strong>Titered output</strong></td>
<td>2.13E-5</td>
</tr>
<tr>
<td><strong>T18.1 binders (%)</strong></td>
<td>2.13E-5</td>
</tr>
</tbody>
</table>
4.2.6 Screening individual clones from the pil18 mouse phagemid library.

To identify individual clones that bound to T18.1, 190 colonies were selected from the biopanning round four plates. These were grown in deepwell plates (Eppendorf) in 500 μL 2xTYAG and small-scale phage rescued (section 2.2.6.3). The supernatant (containing phage) was diluted 1:4 in MT and screened by ELISA for reactivity to T18.1 and BSA. Of the 190 clones screened, 173 bound specifically to T18.1 with no binding to BSA (Figure 4.12). The 47 clones with the highest signal to T18.1 from each plate were selected for further analysis. The individual clones were assigned names based on their plate and plate position at this time e.g. αE3, βC1.

Figure 4.11 Polyclonal phage ELISA. Pools of output phage were diluted 1:4 and screened for specificity to T18.1.
Figure 4.12 Monoclonal phage ELISA. Individual phage clones from pan 4 were tested for specific reactivity to T18.1. (A) Reactivity of the clones from the alpha plate to T18.1 (left) and BSA (right). (B) Reactivity of the clones from the beta plate to T18.1 (left) and BSA (right). In both plates A1 was a sterile/negative control during the small scale phage rescue process and ELISA.

4.2.7 Assessing the diversity of the Fab library

The pComb3X phagemids from the 94 selected clones were miniprepped (Zymo research) and their heavy and light chains were Sanger sequenced (Macrogen Inc) in order to identify unique Fab clones (section 2.2.6.5). Of the 192 sequencing reactions, 183 yielded usable sequence data for the whole Fab. The complete Fab sequences were assembled and analysed using Geneious (Biomatters Limited) and IMGT/V-QUEST (Brochet, Lefranc, & Giudicelli, 2008). Sequencing identified 23 unique Fab clones which are summarised in
Table 4.2. Three of these clones (βG1, βB2 and βG8) had >99.7% sequence identity and any differences between them were single conservative substitutions in either framework region two of the light chain or framework region three of the heavy chain which do not contribute directly to Fab binding. Due to this, they were treated as one clone from this point onwards, represented by βG1. Similarly, βF11 and αE1 had >99.5% sequence identity with one conservative substitution in each of framework regions one and three of the heavy chain. They were treated as one clone, represented by αE1. This left twenty unique clones.

All of the clones used kappa light chains which is not unexpected given that >95% of circulating immunoglobulins in mice utilise kappa genes (Haughton et al., 1978; Larijani et al., 2006). Five different Vκ gene families were represented with IGKV17 being used in one clone and IGKV3, IGKV4, IGKV6 and IGKV10 used in multiple clones. Overall, fourteen of the twenty VL CDR3 (complementarity determining region 3) sequences are unique. Only 3 different VH gene families were represented with IGHV2 used in 1 clone, IGHV5 used in 2 clones and IGHV1 used in 17 clones. The IGHV1 gene family is the most frequently used in mice, comprising >60% of the gene usage in the mouse repertoire (Huang et al., 2013). It appears to have been selected for during panning, increasing its gene usage among the T18.1 binding Fab to 91%. The CDR3 of the VH is known to be the major determinant of an antibody’s specificity due to its greater potential for variability (Xu & Davis, 2000). Of the twenty clones, fifteen had unique VH CDR3 (HCDR3) sequences suggesting that certain CDR3 sequences were selected for during panning and therefore represented multiple times. Each HCDR3 is 12-15 residues long which is slightly longer than the average murine CDR3 length of 11 residues (Shi et al., 2014). The amino acid usage and positioning in these Fab is typical of murine CDR3s being dominated by tyrosine (Y), arginine (R), aspartic acid (D), alanine (A) and glycine (G) (Summarised in Figure 4.13) (Shi et al., 2014; Zemlin et al., 2003).
Table 4.2 Sequence analysis of the 23 Fab clones identified by phage display.

<table>
<thead>
<tr>
<th>Fab clone</th>
<th>Nucleic acid identity†</th>
<th>CDR Length‡</th>
<th>CDR sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V&lt;sub&gt;L&lt;/sub&gt; (% / gene&lt;sup&gt;*&lt;/sup&gt;allele)</td>
<td>V&lt;sub&gt;H&lt;/sub&gt; (% / gene&lt;sup&gt;*&lt;/sup&gt;allele)</td>
<td>V&lt;sub&gt;L&lt;/sub&gt;</td>
</tr>
<tr>
<td>αC2</td>
<td>99.25 / IGKV4-59*01</td>
<td>94.79 / IGHV1-26*01</td>
<td>[5.3.9]</td>
</tr>
<tr>
<td>αC3</td>
<td>97.13 / IGKV3-2*01</td>
<td>96.53 / IGHV155*01</td>
<td>[10.3.9]</td>
</tr>
<tr>
<td>αC9</td>
<td>83.15 / IGKV3-2*01</td>
<td>95.49 / IGHV155*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>αD2</td>
<td>96.28 / IGKV17-121*01</td>
<td>94.44 / IGHV155*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>αE1</td>
<td>99.62 / IGKV6-23*01</td>
<td>95.14 / IGHV5-12-1*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>βF11</td>
<td>100 / IGKV6-23*01</td>
<td>94.79 / IGHV5-12-1*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>αE3</td>
<td>96.76 / IGKV10-96*01</td>
<td>93.75 / IGHV1581*02</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>αE4</td>
<td>95.17 / IGKV10-96*01</td>
<td>91.67 / IGHV155*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>αE5</td>
<td>97.58 / IGKV3-2*01</td>
<td>94.79 / IGHV155*01</td>
<td>[10.3.9]</td>
</tr>
<tr>
<td>αf8</td>
<td>96.19 / IGKV3-12*01</td>
<td>96.53 / IGHV155*01</td>
<td>[10.3.9]</td>
</tr>
<tr>
<td>αG3</td>
<td>97.44 / IGKV10-96*01</td>
<td>95.83 / IGHV155*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>αH1</td>
<td>99.64 / IGKV6-17*01</td>
<td>96.18 / IGHV155*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>αH3</td>
<td>98.88 / IGKV4-72*01</td>
<td>94.79 / IGHV1-9*01</td>
<td>[5.3.9]</td>
</tr>
<tr>
<td>αH5</td>
<td>99.62 / IGKV4-72*01</td>
<td>94.44 / IGHV155*01</td>
<td>[5.3.9]</td>
</tr>
<tr>
<td>βB2</td>
<td>94.12 / IGKV6-13*01</td>
<td>93.75 / IGHV15130*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>βG1</td>
<td>94.07 / IGKV6-13*01</td>
<td>93.75 / IGHV15130*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>βG8</td>
<td>94.27 / IGKV6-13*01</td>
<td>93.75 / IGHV15130*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>βC1</td>
<td>97.12 / IGKV6-15*01</td>
<td>97.19 / IGHV2-2*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>βC7</td>
<td>97.04 / IGKV10-96*01</td>
<td>96.53 / IGHV155*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>βD11</td>
<td>98.56 / IGKV10-96*01</td>
<td>92.71 / IGHV155*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>βD7</td>
<td>100 / IGKV4-80*01</td>
<td>92.36 / IGHV1-9*01</td>
<td>[5.3.8]</td>
</tr>
<tr>
<td>βF12</td>
<td>94.81 / IGKV10-96*01</td>
<td>97.57 / IGHV155*01</td>
<td>[6.3.9]</td>
</tr>
<tr>
<td>βG3</td>
<td>88 / IGKV6-23*01</td>
<td>95.83 / IGHV155*01</td>
<td>[6.3.9]</td>
</tr>
</tbody>
</table>

†Percentage identity to the germline gene as calculated by IMTG/V-QUEST. ‡CDR length of [CDR1, CDR2, CDR3] as determined by IMTG/V-QUEST.
Overall, sequence analysis of the Fab clones demonstrated that biopanning had successfully selected for twenty similar, but not identical, Fab that bound to T18.1. The fact that some of the unique clones have identical HCDR3 sequences (for example αC3, αE4, βG3 and βF12) is interesting, and suggests that there is selection for certain HCDR3 sequences that bind to T18.1. It is possible that T18.1 contains epitopes that are targeted by multiple Fab from this library. In order to select the most promising clones for co-crystallography and protection studies, the Fab were further characterised to determine their expression level, affinity and T antigen specificity or cross-reactivity.

4.2.8 Expression of recombinant Fab

The pComb3X-Fab phagemid from each of the 20 clones was transformed into chemically competent TOP10F’ E. coli (Invitrogen) using the heat shock method (section 2.2.2.2). TOP10F’ E. coli is a non-suppressor strain of E. coli that recognises amber stop codons. pComb3X contains an amber stop codon located between the constant region of the heavy
chain of the cloned Fab and geneIII. In this way, protein expression in Top10F’ _E. coli_ allows the expression of soluble Fab and prevents the expression of geneIII.

The Fab were over-expressed in 2xTY media supplemented with 100 μg ampicillin and 0.1% glucose. The cells were grown at 30°C until they reached exponential phase then induced with 1 mM IPTG at 18°C for 16 hours (section 2.2.3.5). The light chain and heavy chain constructs were directed to the periplasm by their OmpA and pelB leader sequences allowing the Fab to be purified from the periplasm by cold osmotic shock (section 2.2.3.6) and then by IMAC. A representative elution profile for a Fab purification is shown in Figure 4.14. The fractions containing the Fab were pooled, buffer exchanged into PBS and concentrated. All Fab were expressed using the same method, and most clones yielded between 200 μg/L and 1 mg/L. However, βG3 and βD7 could not be recombinantly expressed. As their gene families and heavy chain CDR3 sequences were represented by other Fab (Table 4.2) these clones were not pursued further.

**Figure 4.14** Immobilised metal affinity chromatography of αH3 Fab using a HisTrap HP column (GE healthcare). (A) Chromatogram of αH3 elution. (B) SDS-PAGE analysis of the eluted fractions showing the separation of the Fab into heavy (upper band) and light (lower band) chains in reducing conditions. P = Protein mass standards.
4.2.9 Estimating the affinity of T18.1 binding Fab

The affinity of the eighteen recombinant Fab to T18.1 was initially estimated as the Fab concentration needed to reach half-maximum binding (EC50) by ELISA (section 2.2.7.1.2). Each Fab was diluted to generate binding curves to T18.1 which were plotted in Graphpad Prism (Graphpad) (Figure 4.15). The EC50 values were calculated using an asymmetric sigmoidal, five-parameter logistic equation and are shown in Table 4.3.
Figure 4.15 Binding of the eighteen Fab to T18.1. Binding curves were generated by ELISA with serial dilutions of Fab starting at 1000 nM. These curves are representative of three experiments.
Table 4.3 Estimated affinity of the eighteen recombinantly expressed Fab for T18.1.

<table>
<thead>
<tr>
<th>Fab</th>
<th>αC2</th>
<th>αC3</th>
<th>αC9</th>
<th>αD2</th>
<th>αE1</th>
<th>αE3</th>
<th>αE4</th>
<th>αE5</th>
<th>αF8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affinity (nM)</td>
<td>4.2</td>
<td>5.6</td>
<td>19.1</td>
<td>5.6</td>
<td>30.2</td>
<td>3.1</td>
<td>15.1</td>
<td>3.2</td>
<td>2.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fab</th>
<th>αG3</th>
<th>αH1</th>
<th>αH3</th>
<th>αH5</th>
<th>βC1</th>
<th>βC7</th>
<th>βD11</th>
<th>βF12</th>
<th>βG1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affinity (nM)</td>
<td>16.9</td>
<td>4.1</td>
<td>1.8</td>
<td>5.6</td>
<td>63.4</td>
<td>16.4</td>
<td>15.3</td>
<td>6.0</td>
<td>6.8</td>
</tr>
</tbody>
</table>

The solution biopanning process has selected for Fab with high affinity to T18.1. The majority (16/18) of Fab had an estimated affinity <20 nM with 11 having an estimated affinity <10 nM. The αE1 and βC1 Fab have somewhat lower affinity to T18.1 (30 nM and 63 nM respectively) and are the only two Fab to have unique V<sub>H</sub> genes (Table 4.2). It is interesting to note that the two lowest affinity Fab are the only clones to have unique V<sub>H</sub> genes (αE1 uses IGHV5 and βC1 uses IGHV2), while the remaining sixteen Fab, all with affinity <20 nM, use IGHV1 gene. This suggests that the heavy chain is influencing the binding affinity of the Fab for T18.1.

The affinity of the Fab has implications for co-crystallography as higher affinity Fab may be more likely to remain in a complex with T18.1 in a wide range of crystallisation conditions.

4.2.10 Determining the reactivity of the T18.1 binding Fab to a T antigen panel

4.2.10.1 Construction of a T antigen panel

This project required a panel of recombinant T antigens that covers all of the circulating tee types to investigate the presence of cross-reactive epitopes. A number of T antigens had been prepared in whole or in part for previous projects, including my biomedical honours, and by Dr Jace Loh (Faculty of Medical and Health Sciences), Dr John Steemson (School of Biological Sciences) Dr Nikki Moreland (Faculty of Medical and Health Sciences) and Dr
Paul Young (School of Biological Sciences). A summary of the properties of the T antigen panel is given in Table 4.4. The purity and integrity of all proteins was analysed by SDS-PAGE (Figure 4.16).
Table 4.4 Properties of the recombinantly expressed T antigens used in this study.

<table>
<thead>
<tr>
<th>tee type</th>
<th>Strain</th>
<th>Construct length (amino acids)</th>
<th>Vector</th>
<th>Mass (kDa)</th>
<th>Purification Tag</th>
<th>Cloned</th>
<th>Expressed and purified</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reference: SF370</td>
<td>291</td>
<td>pGEX-3c</td>
<td>32.2</td>
<td>GST</td>
<td>Previously (JL)</td>
<td>This project (JR)</td>
</tr>
<tr>
<td>2</td>
<td>Reference: MGAS10270</td>
<td>513</td>
<td>pProEXHTb</td>
<td>56.0</td>
<td>His6</td>
<td>Previously (JL)</td>
<td>Previously (JL)</td>
</tr>
<tr>
<td>3.2</td>
<td>Clinical: ESR 13637</td>
<td>282</td>
<td>pProEX Hta</td>
<td>31.4</td>
<td>His6</td>
<td>Previously (PY)</td>
<td>This project (JR)</td>
</tr>
<tr>
<td>5</td>
<td>Clinical: ESR 10514</td>
<td>285</td>
<td>pProEX Hta</td>
<td>31.8</td>
<td>His6</td>
<td>Previously (JDS)</td>
<td>This project (NM/JR)</td>
</tr>
<tr>
<td>6</td>
<td>Reference: MGAS10394</td>
<td>488</td>
<td>pProEXHTb</td>
<td>52.6</td>
<td>His6</td>
<td>Previously (JL)</td>
<td>Previously (HF)</td>
</tr>
<tr>
<td>9</td>
<td>Clinical: ESR 11262</td>
<td>287</td>
<td>pProEX Hta</td>
<td>32.0</td>
<td>His6</td>
<td>Previously (JDS)</td>
<td>This project (JR)</td>
</tr>
<tr>
<td>10</td>
<td>Clinical: ESR 13347</td>
<td>287</td>
<td>pET101/D-TOPO</td>
<td>31.9</td>
<td>His6</td>
<td>Previously (JR)</td>
<td>This project (JR)</td>
</tr>
<tr>
<td>11</td>
<td>Clinical: ESR 05151</td>
<td>292</td>
<td>pProEX Hta</td>
<td>32.8</td>
<td>His6</td>
<td>Previously (JR)</td>
<td>Previously (JR)</td>
</tr>
<tr>
<td>12</td>
<td>Clinical: ESR 12303</td>
<td>280</td>
<td>pProEX Hta</td>
<td>30.8</td>
<td>His6</td>
<td>Previously (JDS)</td>
<td>This project (NM/JR)</td>
</tr>
<tr>
<td>13</td>
<td>Clinical: ESR 131465</td>
<td>280</td>
<td>pProEX Hta</td>
<td>31.0</td>
<td>His6</td>
<td>Previously (PY)</td>
<td>Previously (PY)</td>
</tr>
<tr>
<td>18.1</td>
<td>Reference: MGAS8232</td>
<td>288</td>
<td>pProEX Hta-Avitag</td>
<td>31.5</td>
<td>His6</td>
<td>This project (JR)</td>
<td>This project (JR)</td>
</tr>
<tr>
<td>18.2</td>
<td>Reference: NZ131</td>
<td>289</td>
<td>pProEX Hta</td>
<td>31.9</td>
<td>His6</td>
<td>Previously (PY)</td>
<td>Previously (PY)</td>
</tr>
<tr>
<td>25</td>
<td>Clinical: LN144</td>
<td>499</td>
<td>pProEX Hta</td>
<td>55.0</td>
<td>His6</td>
<td>Previously (JDS)</td>
<td>This project (JR)</td>
</tr>
<tr>
<td>28.1</td>
<td>Reference: MGAS6180</td>
<td>290</td>
<td>pProEX Hta</td>
<td>31.7</td>
<td>His6</td>
<td>Previously (JL)</td>
<td>This project (JR)</td>
</tr>
</tbody>
</table>

JL=Dr Jace Loh. JR=Jeremy Raynes. PY=Dr Paul Young. JDS=Dr John Steemson. NM= Dr Nikki Moreland. HF= Hannah Frost.
The T antigens all migrated as per their expected molecular weights except for T18.2 (Table 4.4, Figure 4.16). This migrates with the 75 kDa protein standard because it has been expressed as an uncleavable T18.2-Maltose binding protein (MBP) fusion protein. To account for this, whenever T18.2 was used in immunoassays, native MBP was also used to ensure that any detected signals were specific to T18.2 and not to MBP. T6 and T25 appear to have partially degraded into separate domains. However, as the majority of each T antigen was full-length (52.6 kDa and 55.0 kDa respectively) they were used in subsequent immunoassays.

![Figure 4.16 SDS-PAGE analysis of the T antigen panel.](image)

### 4.2.10.2 Determining the reactivity of the T18.1 binding Fab to the T antigen panel

The Fab, pil18 vaccinated mice and T18.1 vaccinated rabbit were screened against the panel of T antigens by ELISA (section 2.2.7.1.3). Each of the animals reacted strongly to T18.1 and T18.2 (89% identity to T18.1), and also showed strong reactivity to T3.2 and weaker reactivity to T13 (Figure 4.17). Some individual animals also had minor cross-reactivity with T5, T11 and T28.1. Based on T antigen reactivity, the Fab appear to cluster into three groups. One group of Fab only react with T18.1 and T18.2 (αC2, αE1, αF8, αH1, αH5, αH3, βC1 and βF12).
second group reacts with T18.1 and T18.2 and cross-reacts with T3.2 (αC3, αC9, αD2, αE3 and βD11), while a third group cross-react strongly with T3.2 and moderately with T13 (αE4, αE5, αG3, βC7 and βG1). This indicates that biopanning the pil18 library has successfully selected Fab that exhibit the major cross-reactivity patterns observed in all of the animals. However, no Fab were identified that exhibit the moderate cross-reactivity with T5, T11 and T28.1 observed in some of the immunised mice.

Figure 4.17 Specificity of the eighteen recombinantly expressed Fab and animal sera to the T antigen panel. Fab and sera were diluted to 100 nM and 1:200 respectively and screened against the T antigen panel in duplicate by ELISA. The absorbance at 450 nm is shown in a colour gradient from white through to red.
4.2.11 Selection of Fab for epitope mapping and protection studies

Characterisation of the T18.1 binding Fab from the pil18 mouse library informed the selection of four Fab for in depth epitope mapping and protection studies. They are summarised in Table 4.5.

Table 4.5 Summary of the Fab selected for epitope mapping and protection studies.

<table>
<thead>
<tr>
<th>Clone</th>
<th>T antigen reactivity</th>
<th>$\kappa$ gene family</th>
<th>$\lambda$ gene family</th>
<th>$\kappa$ CDR3 length</th>
<th>Affinity (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$E3</td>
<td>T3.2, T18.1, T18.2</td>
<td>IGKV10</td>
<td>IGHV1</td>
<td>13</td>
<td>3.1</td>
</tr>
<tr>
<td>$\alpha$H3</td>
<td>T18.1, T18.2</td>
<td>IGKV4</td>
<td>IGHV1</td>
<td>15</td>
<td>1.8</td>
</tr>
<tr>
<td>$\alpha$E1</td>
<td>T18.1, T18.2</td>
<td>IGKV6</td>
<td>IGHV5</td>
<td>15</td>
<td>30.2</td>
</tr>
<tr>
<td>$\beta$C1</td>
<td>T18.1, T18.2</td>
<td>IGKV6</td>
<td>IGHV2</td>
<td>12</td>
<td>63.4</td>
</tr>
</tbody>
</table>

The $\alpha$E1 and $\beta$C1 Fab were selected because they are the only two Fab to have a $\lambda$ gene that is not from the IGHV1 family. As the $\lambda$ CDR3 is the major determinant of antigen specificity, selecting antibodies with the greatest variability in this region is important for understanding the range of antibody binding and function produced by vaccination. They also have the shortest ($\beta$C1=12) and longest ($\alpha$E1=15) HCDR3 regions. Their somewhat lower affinities for T18.1, compared with the clones that derived from the IGHV1 gene family, suggests that they may interact with T18.1 in a different manner to the IGHV1 clones. Both the $\alpha$E1 and $\beta$C1 Fab are considered to be ‘type-specific’ as they only react with T18.1 and T18.2 in T antigen panel.

In contrast, $\alpha$H3 was selected as it is the highest affinity Fab in the library based on the EC50 data (Table 4.3). This is an IGHV1 Fab clone and ELISA screening against the T antigen panel indicates that it only reacts with T18.1 and T18.2. Thus, $\alpha$H3 is also a ‘type-specific’ Fab and
characterisation will enable direct comparison between this higher affinity Fab and the lower affinity αE1 and βC1 Fab that share the same reactivity profile.

The βG1 Fab was initially selected as it is a cross-reactive Fab that reacts with T3.2 and T13 in addition to T18.1 and T18.2. However, due to its behaviour in initial flow cytometry and crystallisation trials (and the time constraints associated with this project) it was not pursued and will not be discussed further in this thesis.

Finally, the αE3 Fab was selected as it is the highest affinity cross-reactive Fab. It has an estimated EC50 of 3 nM and cross-reacts with T3.2 in addition to T18.1 and T18.2. It uses IGHV1 and has a 13-residue HCDR3 which are representative of the most common V\textsubscript{H} gene families and HCDR3 lengths in the library.

### 4.3 Discussion

The Fab library generated from the pil18 immunised mice is a large library (7.79x10\textsuperscript{7} transformants) from which high affinity Fab, with a range of light chain gene families and T antigen specificities, have successfully been selected. However, the IGHV1 gene family clearly dominates the selected Fab with 91\% of heavy chains using this family. The clones do display variation within the IGHV1 genes with seven different alleles present, and all clones showing somatic mutation (mutation from the germline genes) between 3.47\% and 8.33\% (
Table 4.2). This indicates that the high use of IGHV1 is not due to a single gene that amplified better than the others in PCR and consequently dominated the library. There are also differences in the specificity/cross-reactivity profiles of the Fab which use IGHV1 (Figure 4.17) which suggests that a single clone has not dominated the biopanning selection process and the high use of IGHV1 may be representative of the anti-T18.1 antibody repertoire.

Screening the vaccinated animals against the T antigen panel has shown that vaccination with *L. lactis*-pil18 or monomeric T18.1 (Figure 4.2) generates antibodies with some reactivity to other T antigens from the FCT 3/4/7/8 cluster (Figure 1.4, Figure 4.17). In particular there is dominant cross-reactivity observed between T3.2, T18.1 and T18.2. These three T antigens have significant sequence identity (>60%) and, as shown in Figure 4.18, the regions of identity are spread along the length of the proteins. Since T18.1 and the evolutionarily distinct T1 have conserved atomic structures (Figure 3.7), it is likely that T3.2, T18.1 and T18.2, as more closely related T antigens, will also share the same overall structure. The combination of significant sequence identify and structural homology has resulted in some antibodies binding to epitopes on T18.1 that are also present on T3.2 and T18.2. Therefore it is possible that vaccination with T18.1 may cross-protect against some non-T18.1 strains. As biopanning has successfully captured Fab clones that share the same major cross-reactivity profile as the animals, this potential protection can be tested in *in vitro* assays.

Chapter five explores the interactions of the four selected Fab with monomeric and polymeric T18.1, locates their epitopes and investigates their protective capacity.
Figure 4.18 Protein alignment T3.2, T18.1 and T18.2. The sequences in the alignment represent the mature T antigen monomers. The N-terminal signal peptide has been removed and the sequence ends after the threonine in the LPXTG-like sortase C cleavage site. The residues are numbered according to the sequence of T18.1. The alignment was generated using ClustalW and ESPript.
5 Functional characterisation of T18.1 binding antibodies

5.1 Introduction

The functional characterisation and epitope mapping of the four selected antibodies (αE3, αH3, αE1 and βC1) are key steps in understanding how they interact with T18.1. Recombinantly expressed and affinity purified anti-T18.1 Fab were initially screened by ELISA to estimate their binding properties. The true affinity of a Fab for T18.1 can be calculated using biophysical techniques such as surface plasmon resonance (SPR) to measure antibody-antigen binding. This system has several advantages over ELISA methods in that it measures interactions in real time and without labeling or secondary detection steps. A common issue encountered when using SPR to measure antibody-antigen binding is the bivalent binding of IgG antibodies to the antigen. This can confound the determination of affinity, which is a measure of the strength of interaction between a single Fab (binding arm of the antibody) and the antigen (Drake, Myszka, & Klakamp, 2004). This issue can be avoided by using recombinant Fab in SPR experiments. The accurate characterisation of antibody affinity is important as it can be predictive of function, with high affinity generally correlating with better protection (Maynard et al., 2002; Reddy et al., 2012; Safari et al., 2008; Steckbeck et al., 2005; Tkaczyk et al., 2012); however this is not always the case (Kontio, Jokinen, Paunio, Peltola, & Davidkin, 2012).

As T18.1 is present as a polymer in natural infections, it is also important to investigate how antibodies bind to their epitopes in the polymerised pilus as opposed to the recombinant monomer. The importance of investigating antibody binding to antigens in their native, infectious form has been previously illustrated in structural studies of anti-viral antibodies. Antibodies to Poliovirus (Li, Yafal, Lee, Hogle, & Chow, 1994), Hepatitis C virus (Sabo et al., 2012) and Dengue virus (Kudlacek et al., 2018; Lok et al., 2008) show variable levels of
binding to their respective viral envelopes depending on the conformation of the epitope in the context of the virion. Some epitopes are only in the correct conformation for binding on the virion surface (Fibriansah et al., 2014), and not in the recombinant protein, while other epitopes are only available at physiological temperatures (37°C) following temperature-induced conformational changes (Fibriansah et al., 2015; Lok et al., 2008). Therefore it is essential to study the binding of the four antibodies to T18.1 in the polymerised pilus under physiological conditions.

High affinity binding to an antigen does not necessarily indicate that an antibody is functional. Indeed, non-functional antibodies are known to be produced to candidate vaccine antigens for a number of organisms (Ditlev et al., 2012; Dokmetjian, Della Valle, Lavigne, Eriksson, & Manghi, 1998; Drulhe et al., 2005; Weiss et al., 2006). As ELISA-based methods detect total binding and do not discriminate between functional and non-functional antibodies (Lorenz et al., 2017), it cannot be assumed that protective antibodies were selected during the biopanning and initial characterisation stages of this project. Protection against S. pyogenes has largely been attributed to opsonic antibodies. These are antibodies that bind to their epitope on the target antigen and act as an immunological tag to increase recognition of the pathogen by other components of the immune system, such as phagocytes. The measurement of opsonisation in in vitro bactericidal assays is currently the accepted method to assess function for S. pyogenes antibodies (Dale et al., 2013; Tsoi et al., 2015). However, as discussed in section 1.5.1, the current bactericidal assays have low reproducibility and are not standardised between research groups. To address this, a standardised opsonophagocytic killing assay (OPKA) for S. pyogenes has recently been developed (Jones et al., 2018). The OPKA uses the HL-60 cell line as a source of phagocytes and commercial baby rabbit complement (BRC) as a source of complement proteins. HL-60 cells are differentiated with DMF into neutrophil-like granulocytes and these (and the BRC) are used in place of whole human blood. This removes
the main sources of variability that cause the poor reproducibility observed in current bactericidal assays.

This chapter describes the characterisation of binding for the four selected antibodies to both monomeric and polymeric T18.1 and maps the epitopes of these antibodies. Finally, the ability of the monoclonal antibodies to induce killing of *S. pyogenes* is tested in OPKA and bactericidal assays.

5.2 Results

5.2.1 Characterisation of full-length IgG1 produced from the selected Fab

The Fab fragments produced and characterised in Chapter four represent the ‘binding arm’ of a full-length antibody in isolation (Figure 1.6). They contain the components required to bind to their epitope on T18.1 but lack the Fc ‘tail’ fragment, which is necessary for interactions with white blood cells (Indik et al., 1995; Kiyoshi et al., 2015) and complement proteins (Duncan & Winter, 1988; Kaul & Loos, 1997). In order to test the protective capacity of the Fab in bactericidal and opsonophagocytic killing assays (section 5.2.9), full-length antibodies were needed. Four full-length IgG1, corresponding to the αE3, αH3, αE1 and βC1 Fab, were cloned and purified by Genscript (New Jersey) using the sequence from the mouse variable regions fused to a human IgG1 backbone. These were expressed in Expi293F cells (Life Technologies) and affinity purified using protein A columns. The avidity of the IgG1 to T18.1 and their reactivity to the T antigen panel were determined by ELISA and compared to the corresponding Fab.

As the IgG1 are bivalent, the ELISA estimates their avidity (the interaction of the whole antibody with T18.1) rather than affinity (the interaction between one epitope and its paratope). Each IgG1 was diluted to generate binding curves to T18.1, which were plotted in Graphpad
Prism (Graphpad) (Figure 5.1A). The EC50 values were calculated using an asymmetric sigmoidal, five-parameter logistic equation. The binding curves of the Fab over the same dilution series are shown for comparison. The binding curves of the full-length IgG1 are right-shifted compared to the corresponding Fab and each IgG1 has avidity below 0.1 nM (Figure 5.1B). The avidity of all the IgG1 increased between one and two orders of magnitude which is in line with previously observed increases in avidity upon Fab to IgG conversion (Dmitriev et al., 2001; Hofer et al., 2007; Moreland et al., 2012).
Figure 5.1 Comparison of IgG1 and Fab binding to T18.1. Binding curves were generated by ELISA with serial dilutions of Fab and IgG1 starting at 1000 nM. The solid lines represent the IgG1 clones while the dotted lines represent the Fab. (A1) Comparison of the αE3 Fab and IgG1. (A2) Comparison of the αH3 Fab and IgG1. (A3) Comparison of the αE1 Fab and IgG1. (A4) Comparison of the βC1 Fab and IgG1. (B) Estimated avidity of the IgG1 and Fab to T18.1.
The full-length IgG1 were screened against the T antigen panel to compare their reactivity to the corresponding Fab (Figure 5.2). Overall the specificity of each IgG1 matched the specificity of the Fab as expected. The αE3 IgG1 binds strongly to T18.1, T18.2 and T3.2 and binds weakly to T13. The αH3 IgG1, the αE1 IgG1 and the βC1 IgG1 all only bind to T18.1 and T18.2, as do their corresponding Fab. Interestingly, this has revealed that the epitope targeted by αE3 has minor cross-reactivity to T13, which had not been observed with the Fab.

Figure 5.2 Specificity of the IgG1 clones to the T antigen panel. The IgG1 (grey) were diluted to 10 nM in PBS and screened against the T antigen panel by ELISA. Fab (black) binding at 100 nM is shown for comparison. (A) αE3 IgG1 specificity. (B) αH3 IgG1 specificity. (C) αE1 IgG1 specificity. (D) βC1 IgG1 specificity. Each bar represents the mean and standard deviation of duplicate readings.
5.2.2 Analysis of the affinity and kinetics of anti-T18.1 Fab using Surface Plasmon Resonance

In section 4.2.9 the affinity of the monovalent Fab to T18.1 were estimated by ELISA. This information was used to determine which were ‘high’ and ‘low’ affinity Fab and to rank them by affinity. However, the binding of the Fab to surface-bound T18.1 may not reflect the binding in solution. In order to accurately measure the affinity and observe the binding kinetics of the four Fab to T18.1 in solution, a Biacore T200 (GE Healthcare) was used to measure surface plasmon resonance (SPR).

C-terminally biotinylated T18.1 was immobilised onto the streptavidin coating of a SA Biacore sensor chip (GE healthcare) yielding approximately 50 response units (RU) (section 2.2.8). This immobilisation density was chosen to minimise mass transport effects and crowding. For each Fab, a twelve point dilution series was injected in duplicate over the surface at 30 μL/min. The association phase was measured for 360 s and the dissociation phase was measured for 120 s. The sensorgram in Figure 5.3A shows the dose-dependent binding of βC1 to T18.1, with the highest concentration (150 nM) generating a response of approximately 20 RU and the lowest concentration (5 nM) generating a response of 2 RU. The binding response at equilibrium was fitted to a steady state model using the Biacore T200 evaluation software V 2.0 (GE Healthcare) (Figure 5.3B). The $K_D$ for βC1 was calculated as 82.67 ± 4.49 nM (mean of three repeats ± standard deviation of three repeats). This is comparable to the ELISA measurement (63.4 nM).
Figure 5.3 Analysis of βC1 binding to T18.1 using surface plasmon resonance. (A) Sensorgram showing the dilution series of βC1 binding to T18.1. The response at equilibrium was used to fit a curve (B) and calculate the $K_D$. These curves are representative of three experiments.

Attempts were made to measure the affinities of αE3, αH3 and αE1 for T18.1 in a similar manner but were unsuccessful. As shown in Figure 5.4 the Fab-T18.1 binding kinetics are quite different for βC1 compared with αE3, αH3 and αE1. The αE3 and αH3 Fab have the fastest on-rates (or association constant, $k_a$), which are very similar to each other, while the αE1 Fab has the slowest on-rate followed by βC1. At this concentration, all of the Fab reach equilibrium but αE3, αH3 and αE1 do not dissociate. In other experiments, across a range of concentrations,
dissociation was followed for up to ten minutes but only βC1 dissociated (data not shown). Because of this, the dissociation constant \( (k_d) \) could not be calculated; therefore the \( K_D \) of αE3, αH3 and αE1 could not be calculated using kinetic data (where \( K_D = k_d / k_a \)). This was because the software models for independently calculating the \( k_a \) and \( k_d \) were unable to fit curves that had fast association and slow dissociation rates. The affinity of αE3, αH3 and αE1 could also not be calculated using steady state data (as for βC1) because at concentrations below 50 nM equilibrium was not reached. It is recommended that the dilution series of the sample covers orders of magnitude between 0.01 x \( K_D \) to 100 x \( K_D \) (Van Der Merwe, 2001). However, with concentrations <50 nM not reaching equilibrium and concentrations >50 nM saturating the sensor chip and not dissociating, adequate dilution series could not be generated. It is likely that αE3, αH3 and αE1 have off-rates that are too slow to be easily measured by SPR. It is not unusual for antibodies that have undergone affinity maturation to have extremely slow off-rates as a result of the strong antibody-antigen interactions (England, Nageotte, Renard, Page, & Bedouelle, 1999). Some antibody-antigen binding reactions have been observed to take days to reach equilibrium and up to 4 hours to produce minimal measurable dissociation data (Drake et al., 2004; Katsamba et al., 2006). Given the time constraints associated with this project, as well as the practicalities of using the Biacore facility for the lengths of time required to measure interactions with very low off-rates, it was decided not to pursue SPR further. These experiments did highlight important differences in Fab binding kinetics and generally agreed with the ELISA affinity data showing that αE3 and αH3 have similarly high affinities while αE1 and βC1 have lower affinities.
Figure 5.4 Comparison of the kinetics of Fab binding to T18.1. Sensorgram showing the association and dissociation of 100 nM Fab to T18.1. The arrow denotes the beginning of the dissociation phase.

5.2.3 Analysis of binding to the polymerised pilus using flow cytometry

The antibody libraries were constructed from mice vaccinated with *L. lactis* expressing the whole T18.1 pilus. They were biopanned against recombinant T18.1 monomer to isolate T antigen specific antibodies and all assays thus far have used this monomeric antigen. In order to confirm that the monoclonal antibodies bound to T18.1 in the polymerised pilus (as presented on live bacteria), flow cytometry was used to measure their binding to *L. lactis*-pil18 (section 2.2.9). When the IgG1 were incubated with *L. lactis*-pil18 at 4°C, as is standard for flow cytometric assays, αE1 and βC1 bound strongly while αE3 and αH3 only bound weakly (Figure 5.5A). This is the opposite of what might be expected from their respective affinities for monomeric T18.1. When the same experiment was performed at 37°C (representing
physiological temperatures for antibody-T18.1 interactions) the binding of αE1 and βC1 was similar to that at 4°C but αE3 and αH3 showed a >2-fold increase in binding as measured by the geometric mean of fluorescence intensity (MFI) (Figure 5.5B). This shift indicated that the αE3 and αH3 IgG1 were able to bind to their epitopes on polymerised T18.1 better at physiological infection temperatures than at 4°C. The lack of a shift in αE1 and βC1 binding indicates this is not due to a general increase in binding at 37°C, rather that the binding of αE3 and αH3 has specifically increased. This is likely due to the epitopes of αE3 and αH3 being more accessible at 37°C while the epitopes of αE1 and βC1 are equally accessible at either temperature.
Figure 5.5 Binding of anti-T18.1 IgG1 to the polymerised pilus. Binding of the IgG1 clones at 250 nM to \textit{L. lactis}-pil18 at 4°C shown as a histogram (A1) and binding statistics (A2). Binding of the IgG1 clones at 250 nM to \textit{L. lactis}-pil18 at 37°C shown as a histogram (B1) and binding statistics (B2). Cells were gated on the binding of B9 Fab, which is a non-pil18 binding negative control. Data represent two experiments.

### 5.2.4 Epitope mapping of the selected Fab and T18.1 by co-crystallography

Initial epitope mapping was attempted using x-ray crystallographic analysis of Fab-T18.1 co-crystals. X-ray crystallography is the gold standard for epitope mapping as it can result in high
resolution structures of the complex including atomic level detail of the Fab-epitope interactions as well as identification of both linear and conformational epitopes.

5.2.4.1 Analysing the formation and stability of Fab-T18.1 complexes using Size-Exclusion Chromatography

Before crystallisation trials were undertaken, small-scale tests were performed to test the stability of Fab-T18.1 complexes under pressure. If they did not remain bound during size exclusion chromatography (SEC), they would be difficult to purify as a complex in high quantities and less likely to crystallise. Each of the four Fab were incubated with T18.1 at a molar ratio of 2:1 at room temperature for 2 hours. The reactions were then passed over a S200 size-exclusion column (GE Healthcare) (section 2.2.3.4) and complex formation was analysed by comparing the elution profiles of the mixtures containing Fab and T18.1 with Fab-only and T18.1-only controls (Figure 5.6). Molar ratios of 1:1 and 1:2 Fab to T18.1 were also tested but these were significantly less efficient at producing complexes (data not shown). All elution volume measurements were made from protein injection to the centre of the corresponding peak.

The T18.1 monomer has a predicted molecular weight of 31.5 kDa and eluted in a single peak at 17.8 mL. The Fab have a predicted molecular weight of 50kDa and eluted in a single peak at 16.0 mL. The combined Fab-T18.1 complexes have predicted molecular weights of approximately 81 kDa and were expected to elute between 12 and 14 mL. The αE3 reaction eluted as two peaks (Figure 5.6C). The centre of the major peak eluted at 13.4 mL while the minor peak eluted at 15.5 mL. This indicates that the major product is αE3-T18.1 complex and the minor product is excess αE3 Fab. This was confirmed by SDS-PAGE analysis of the eluted peaks (Figure 5.7). The ratio of αE3-T18.1 complex to unbound Fab was approximately 7:1 after SEC. The αH3 reaction eluted similarly, with αH3-T18.1 eluting at 13.3 mL and excess
Fab eluting at 15.6 mL (Figure 5.6D). The ratio of αH3-T18.1 complex to unbound Fab was approximately 4:1. In contrast, the αE1-T18.1 complex (eluting at 14.2 mL) was only the minor product after SEC (Figure 5.6E). The ratio of αE1-T18.1 complex to substrate was approximately 1:2. The βC1 reaction mixture showed the same pattern with the βC1-T18.1 complex eluting as the minor peak at 14.3 mL (Figure 5.6F). The ratio of complex to substrate was also approximately 1:2.

Because αE3-T18.1 and αH3-T18.1 complexes formed at favourable amounts and remained stable during SEC, they were determined to be good candidates for crystallisation trials. As αE1-T18.1 and βC1-T18.1 complexes were not the major product after SEC, crystallography was not attempted. The epitopes of αE1 and βC1 would be determined using an overlapping peptide array of T18.1 (section 5.2.6).
Figure 5.6 Size exclusion chromatography of Fab-T18.1 complex formation and stability. (A) Elution profile of T18.1. (B) Elution profile of αE3. (C) Elution profile of αE3-T18.1. (D) Elution profile of αH3-T18.1. (E) Elution profile of αE1-T18.1. (F) Elution profile of βC1-T18.1.
Figure 5.7 SDS-PAGE analysis of eluted αE3-T18.1 complex after SEC. The fractions collected from peak 1 (Figure 5.6C) contain the αE3-T18.1 complex represented by bands corresponding to T18.1, the αE3 heavy chain and the αE3 light chain. Peak 2 only contains bands corresponding to the αE3 heavy chain and the αE3 light chain.

5.2.4.2 Crystallisation of αE3-T18.1 and αH3-T18.1 complexes

The αE3 Fab was expressed and purified as in sections 2.2.3.5 and 2.2.3.6. Following this, αE3 Fab was incubated with T18.1 at a 2:1 molar ratio for two hours at room temperature. The αH3 Fab was purified and complexed in the same way. The αE3-T18.1 and αH3-T18.1 complexes were purified from any unbound substrate by SEC (Figure 5.6C-D) and eluted in crystallisation buffer (10 mM Tris pH 8.0, 100 mM NaCl).

Crystallisation trials were carried out using a 288-condition screen comprising the JCSG+/PACT premier screen (Newman et al., 2005) and the MORPHEUS screen (Gorrec, 2009). Vapour diffusion sitting drop trials were performed at 18°C using an Oryx4 crystallisation robot (Douglas Instruments) as described in section 2.2.4.1. The αE3-T18.1 and αH3-T18.1 complexes were concentrated to 88 mg/mL and 8 mg/mL respectively and 0.15 µL of complex was mixed with 0.15 µL of the screen precipitant. Crystals of the αE3-T18.1 complex formed in 92 conditions while crystals of the αH3-T18.1 complex formed in >100 conditions. In both cases, the majority of crystals developed in the PACT premier screen. Similarly to T18.1 crystals (Figure 3.3), most of the Fab-T18.1 crystals formed either fine
needles or plates. αE3-T18.1 crystals from 6 conditions and αH3-T18.1 crystals from 12 conditions were mounted in cryoloops (Hampton Research) and flash cooled in liquid nitrogen with cryoprotectant (20% glycerol) except for those from the MORPHEUS screen as the precipitant is a cryoprotectant.

5.2.4.3 Data collection and analysis

X-ray diffraction data were collected for crystals of both Fab-T18.1 complexes using the MX2 beamline at the Australian Synchrotron (section 2.2.4.2). Full 360° data sets were collected for one crystal of the αE3-T18.1 complex (grown in 10% (w/v) PEG 8000, 20% (v/v) ethylene glycol, 0.3 M diethyleneglycol, 0.3 M triethyleneglycol, 0.3 M tetraethyleneglycol, 0.3 M pentaethyleneglycol and 0.1 M MOPS/HEPES-Na pH 7.5) which diffracted to approximately 1.8 Å and for one αH3-T18.1 crystal (grown in 0.2 M potassium thiocyanate, 0.1 M Bis-Tris propane pH 7.5 and 20% (w/v) PEG 3350) which diffracted to approximately 2.2 Å (Figure 5.8).
Figure 5.8 Data collection of the Fab-T18.1 co-crystals. (A1) The αE3-T18.1 co-crystal during collection at the MX2 beamline at the Australian synchrotron. (A2) Representative diffraction pattern from this crystal. (B1) The αH3-T18.1 co-crystal during collection at the MX2 beamline at the Australian synchrotron. (B2) Representative diffraction pattern from this crystal. The white arrows indicate the path of the beam through the crystals.

The dataset from the αE3-T18.1 crystal was indexed and integrated using XDS (Kabsch, 2010) and scaled using AIMLESS. The structure of the αE3-T18.1 complex was solved by molecular replacement (MR) using PHASER-MR (McCoy, 2007) with the T18.1 structure solved in section 3.2.4 as a search model. The structure underwent multiple rounds of manual building in COOT (Emsley et al., 2010) and refinement using Refmac5 (Murshudov et al., 2011). Molprobity (Davis et al., 2007) was used to validate the structure, and scored the model in the top 99th percentile of structures of similar resolution. The αE3-T18.1 complex was solved to 1.9 Å resolution, with final R and R_free values of 0.219 and 0.250 respectively.
The dataset from the αH3-T18.1 crystal was initially analysed in the same way except that the solved αE3-T18.1 structure was used as the search model for MR. However, while the T18.1 component of the complex could be solved, the electron density of the αH3 Fab was not of sufficient quality to be solved by MR. Automated building using ARP/wARP (Perrakis, Harkiolaki, Wilson, & Lamzin, 2001) and Phenix (Adams et al., 2002) as well as manual building in COOT were attempted but the Fab component could not be solved. As such, the epitope of αH3 would also be determined using an overlapping peptide array of T18.1 (section 5.2.6).

5.2.4.4 Identification of the αE3 epitope

The αE3-T18.1 complex crystallised in space group I1 2 1 with one complex in the asymmetric unit. The T18.1 molecule is fully modelled including the N-domain loop comprising residues 124-130 that could not be modelled in the structure of T18.1. The light chain of αE3 is also fully modelled but the heavy chain has two short unmodelled spans comprising residues 134-141 and 195-198. Both of these unmodelled regions are in the framework region of the C\textsubscript{H\textgamma} domain and are not directly involved in the binding of the αE3 Fab to T18.1. Full data collection parameters and refinement statistics are shown in Table 5.1.
Table 5.1 Collection and analysis parameters for αE3-T18.1.

<table>
<thead>
<tr>
<th>Collection Statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wavelength (Å)</strong></td>
<td>0.953724</td>
</tr>
<tr>
<td><strong>No. of images</strong></td>
<td>3600</td>
</tr>
<tr>
<td><strong>Oscillation angle (°)</strong></td>
<td>0.1</td>
</tr>
<tr>
<td><strong>Resolution range (Å)</strong></td>
<td>41.96 -1.90 (1.94 – 1.90)</td>
</tr>
<tr>
<td><strong>Total no. of observations</strong></td>
<td>451654 (29469)</td>
</tr>
<tr>
<td><strong>Unique reflections</strong></td>
<td>67091 (4288)</td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
<td>6.7 (6.9)</td>
</tr>
<tr>
<td><strong>Space group</strong></td>
<td>ıııı</td>
</tr>
<tr>
<td><strong>Unit-cell axial lengths (Å)</strong></td>
<td>a=126.57 b= 41.51 c= 172.35</td>
</tr>
<tr>
<td><strong>α = 90.00, β = 110.10, γ = 90.00</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
<td>100 (100)</td>
</tr>
<tr>
<td><strong>Mean I/σ(I)</strong></td>
<td>8.6 (0.8)</td>
</tr>
<tr>
<td><strong>R_{merge} (%) †</strong></td>
<td>0.095 (2.243)</td>
</tr>
<tr>
<td><strong>CC(1/2) ‡</strong></td>
<td>0.998 (0.432)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Refinement statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resolution range (Å)</strong></td>
<td>41.96 -1.90 (1.94 – 1.90)</td>
</tr>
<tr>
<td><strong>Molecules per A.U.</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Solvent content (%)</strong></td>
<td>52</td>
</tr>
<tr>
<td><strong>R_{work}/R_{free} (%) ‡</strong></td>
<td>21.9/25.0</td>
</tr>
<tr>
<td><strong>Protein atoms</strong></td>
<td>5284</td>
</tr>
<tr>
<td><strong>Water molecules</strong></td>
<td>312</td>
</tr>
<tr>
<td><strong>RMS deviation from ideal bond length/angle (Å/°)</strong></td>
<td>0.003/0.824</td>
</tr>
<tr>
<td><strong>Mean B-factor (Å²)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Protein</strong></td>
<td>53.1</td>
</tr>
<tr>
<td><strong>Water</strong></td>
<td>51.7</td>
</tr>
<tr>
<td><strong>Ramachandran Most Favoured (%)</strong></td>
<td>97.1</td>
</tr>
<tr>
<td><strong>Ramachandran Outliers (%)</strong></td>
<td>0.15</td>
</tr>
</tbody>
</table>

*Numbers in parentheses for outermost shell. † Mn(I) half-set correlation CC(1/2) as calculated by Scala.
‡ R_{work} and R_{free} = \frac{\sum \left| F_{obs} \right| - \left| F_{calc} \right|}{\sum \left| F_{obs} \right|}, where R_{free} was calculated over 5% of amplitudes that were chosen at random and not used in refinement.
The αE3 Fab binds to the top of the N-domain of T18.1 (Figure 5.9). The epitope was identified using ProFunc protein-protein interaction analysis (Laskowski, 2017) and consists of fourteen residues on T18.1 which interact with sixteen residues on the αE3 Fab (Figure 5.10). Of these, twelve are found in the heavy chain CDRs: S31, F32 and Y33 are located in CDR1; N52, N55, A58 and N59 are located in CDR2; and F100, Y101, Y102, G103 and W105 are located in CDR3. In contrast, only four interacting residues are located in the light chain and all of these are found in CDR3 (V91-L94). This indicates that the heavy chain is the main contributor to the bonding interactions and therefore the main determinant of binding. This is particularly interesting given that 18/20 clones isolated from the pil18 mouse library use the same IGHV1 gene family (Table 4.2). This suggests that the majority of the Fab clones, including αH3, may bind to T18.1 in a similar manner.
Figure 5.9 The structure of the αE3-T18.1 complex presented as a ribbon diagram. The heavy chain (blue) of the αE3 Fab forms the majority of the binding interface with only one loop of the light chain (light blue) interacting with T18.1 (green). The isopeptide bonds in T18.1 are shown in blue and red. The pilin lysine (residue 146) that is predicted to polymerise with the next T18.1 monomer is also shown (K). The calcium atom is represented as a grey sphere.
Figure 5.10 ProFunc calculation of the αE3 epitope on T18.1. Hydrogen bonds are shown in blue. Non-bonding interactions are shown in orange where the width of the striped line is proportional to the number of atomic contacts. Residue colours: Blue = positive; Red = negative; Green = neutral; Grey = aliphatic; Purple = aromatic.

In the αE3-T18.1 complex there are nine hydrogen bonds (interactions shown in Table 5.2) and a network of 80 non-bonding interactions. N125 is likely a key determinant of binding as it forms two hydrogen bonds with Y101 in CDR3 of the αE3 heavy chain and one hydrogen bond
with S31 in CDR1 of the heavy chain. K45 is also able to form hydrogen bonds with A58 and N59 in CDR2 of the heavy chain while K47 can hydrogen bond with L94 in CDR3 of the light chain. In this way, all four of the αE3 CDRs that contribute to the paratope are hydrogen bonded to T18.1.

Table 5.2 Hydrogen bonds between the αE3 Fab and T18.1.

<table>
<thead>
<tr>
<th>T18.1</th>
<th>αE3 Heavy Chain</th>
<th>T18.1</th>
<th>αE3 Light Chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>LYS45NZ</td>
<td>ALA58O</td>
<td>LYS47O</td>
<td>LEU94N</td>
</tr>
<tr>
<td>LYS45NZ</td>
<td>ASN59OD1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASP48O</td>
<td>TRP105NE1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLY49O</td>
<td>ASN59ND2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASN125N</td>
<td>TYR101O</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASN125D1</td>
<td>TYR101N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASN125ND2</td>
<td>SER31O</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLY12O</td>
<td>SER31OG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
When the residues which make up the epitope are mapped onto the surface of T18.1 they clearly show that the αE3 Fab binds exclusively across the top of the N-domain (Figure 5.11). The αE3 epitope is spread across three loops (D40-K53, D84-V95 and Y122-Y135) and the binding of αE3 induces movement in all three of these loops (Figure 5.12) This αE3-T18.1 interface is also predicted to form the interface between the N-domain of one T antigen monomer and the C-domain of the next (as well as between the final T antigen and AP1) in the polymerised pilus (Kang et al., 2007). The area of the αE3-T18.1 interface is 803 Å² which is almost the same
size as the predicted interface between the T antigen monomers (850 Å²) (Kang et al., 2007). Therefore, in the polymerised pilus the αE3 Fab will likely be competing with the C-domain of the preceding T18 monomer for binding at this interface. This may reduce accessibility to the epitope in the polymerised pilus, which could explain why the αE3 Fab only shows partial binding to the pilus in flow cytometry (Figure 5.5).

Figure 5.12 Binding of the αE3 Fab induces the movement of three loops in T18.1. The structure of T18.1 without αE3 bound is shown in ribbon diagram in grey. The structure of T18.1 with αE3 bound is shown in green. The remainder of T18.1 is structurally unchanged by the binding of αE3.

5.2.5 Binding of the αE3 Fab does not prevent pilus assembly

As the αE3 Fab binds at the interface between the N-domain of one T18.1 monomer and the C-domain of the next (and binds somewhat weakly to the polymerised pilus (Figure 5.5), it was hypothesised that pilus assembly may be inhibited by the bulk of the Fab. Pilus expression can be visualised by western blotting as a ladder of high molecular weight bands that represent pili incorporating varying numbers of T antigen subunits (Loh et al., 2017). In order to test for
inhibition of pilus assembly, cell wall extracts were taken from *L. lactis*-pil18 grown in the presence of 1 μM αE3 Fab for 3 hours (the earliest time point at which pilus could be detected) (section 2.2.3.7). The banding patterns of the cell wall extracts and the protoplasts were detected using serum from the T18.1 vaccinated rabbit and compared to untreated *L. lactis*-pil18 by western blot (Figure 5.13). If αE3 had an inhibitory effect then there would be a decrease in high molecular weight pili and potentially an increase in low molecular weight bands as αE3 binding to T18.1 monomers or small polymers prevented their incorporation into larger pili. However, pili are still able to assemble in the presence of αE3 as shown by the high molecular weight pili in lanes 3 (cell wall extract) and 7 (protoplast) in Figure 5.13. The patterns of laddering and band intensity are similar to the controls indicating that pili of similar sizes are present in similar amounts. While only qualitative, this suggests that binding of the αE3 Fab to T18.1 does not prevent pilus assembly.

![Figure 5.13 Pilus assembly in the presence of αE3 Fab. Cell wall extracts from: (1) untreated *L. lactis*, (2) *L. lactis* treated with 1 μM B9 Fab, (3) *L. lactis* treated with 1 μM αE3 Fab, (4) Recombinant T18.1. Protoplasts from: untreated *L. lactis* (5), *L. lactis* treated with 1 μM B9 Fab (6) and *L. lactis* treated with 1 μM αE3 Fab (7). (8) Recombinant T18.1. This blot is representative of three experiments.](image)
5.2.6 Epitope mapping of the selected Fab and T18.1 using an overlapping peptide library

5.2.6.1 Determining the linearity of the epitopes

An overlapping peptide library of T18.1 was to be used to map the epitopes of the remaining three Fab. However, a library consisting of short peptides is only likely to contain linear epitopes or the linear components of epitopes. Therefore, if any of the Fab epitopes were entirely conformational, a peptide library would not be able to identify that epitope. In order to determine whether or not the Fab bound to linear epitopes, immunoblots with full-length T18.1 were performed (sections 2.2.7.2 and 2.2.7.3). Each Fab was screened for binding to 200 ng of native T18.1 in a dot blot and to 200 ng of reduced and denatured T18.1 in a western blot. As expected, all four Fab bound to T18.1 on the dot blots, which present their epitopes, whether linear or conformational, in their native state (Figure 5.14A). All four Fab also bound to denatured T18.1 showing that their epitopes contain linear elements (Figure 5.14B). In the case of αE3 this is not unexpected due to its epitope containing flexible loop regions (K45-D51 and Y122-K134). This indicated that an overlapping peptide library of T18.1 was an appropriate tool for mapping the epitopes of the αH3, αE1 and βC1 Fab.
Figure 5.14 Immunoblot analysis of the selected Fab binding to T18.1. (A) Dot blot showing binding of the four Fab to native T18.1 at 10 nM (in triplicate). (B) Western blot showing the binding of the four Fab to denatured T18.1 at 10 nM. T1 was included as a negative control. Fab binding was detected with an Anti-Human IgG (H+L)-HRP (Jackson Immunoresearch) secondary antibody diluted 1:2500.

5.1.1.1 Design of the overlapping peptide library of T18.1

The peptide library was designed to cover the entirety of the mature T18.1 monomer. The N-terminal signal sequence was removed and the library ended at the conserved threonine residue of the QVPTG Sortase C recognition site. The peptide library consisted of 29 peptides with a length of 15 residues and an overlap of 5 residues (Table 5.3). The peptides were synthesised by Genscript (New Jersey) using proprietary technology based on fmoc chemistry and purified to >85%.
<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
<th>Length</th>
<th>Position</th>
<th>Molecular Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ETAGVIDGSTLVVKK</td>
<td>15</td>
<td>1-&gt;15</td>
<td>1516.74</td>
</tr>
<tr>
<td>2</td>
<td>LVVKKTFSYTDDKV</td>
<td>15</td>
<td>11-&gt;25</td>
<td>1740.02</td>
</tr>
<tr>
<td>3</td>
<td>TDKVLMKPADYTFK</td>
<td>15</td>
<td>21-&gt;35</td>
<td>1772.05</td>
</tr>
<tr>
<td>4</td>
<td>DYTTFKVEADNAGK</td>
<td>15</td>
<td>31-&gt;45</td>
<td>1700.81</td>
</tr>
<tr>
<td>5</td>
<td>NAKGKTDGLDIKPG</td>
<td>15</td>
<td>41-&gt;55</td>
<td>1541.74</td>
</tr>
<tr>
<td>6</td>
<td>DIKPGVIDGLENKT</td>
<td>15</td>
<td>51-&gt;65</td>
<td>1599.79</td>
</tr>
<tr>
<td>7</td>
<td>ENTKTIHYGNSDKTT</td>
<td>15</td>
<td>61-&gt;75</td>
<td>1708.80</td>
</tr>
<tr>
<td>8</td>
<td>SDKTTAKEKSVNFDF</td>
<td>15</td>
<td>71-&gt;85</td>
<td>1716.86</td>
</tr>
<tr>
<td>9</td>
<td>VNFDPAVKNFKPGVG</td>
<td>15</td>
<td>81-&gt;95</td>
<td>1609.82</td>
</tr>
<tr>
<td>10</td>
<td>PGVGVRYTSENG</td>
<td>15</td>
<td>91-&gt;105</td>
<td>1596.75</td>
</tr>
<tr>
<td>11</td>
<td>SEVNGNKAGIAYDSQ</td>
<td>15</td>
<td>101-&gt;115</td>
<td>1552.60</td>
</tr>
<tr>
<td>12</td>
<td>AYDSQQWVDVYVVN</td>
<td>15</td>
<td>111-&gt;125</td>
<td>1786.90</td>
</tr>
<tr>
<td>13</td>
<td>VYVVRREDGGFEAKY</td>
<td>15</td>
<td>121-&gt;135</td>
<td>1745.93</td>
</tr>
<tr>
<td>14</td>
<td>FEAKYIVSTEGQSD</td>
<td>15</td>
<td>131-&gt;145</td>
<td>1630.73</td>
</tr>
<tr>
<td>15</td>
<td>GQSDKKPVLFKNFF</td>
<td>15</td>
<td>141-&gt;155</td>
<td>1711.96</td>
</tr>
<tr>
<td>16</td>
<td>FKNFDTSSLKVTKK</td>
<td>15</td>
<td>151-&gt;165</td>
<td>1804.11</td>
</tr>
<tr>
<td>17</td>
<td>KVTKKVTGNTGEHQR</td>
<td>15</td>
<td>161-&gt;175</td>
<td>1682.88</td>
</tr>
<tr>
<td>18</td>
<td>GEHQRSFSFTLLTLP</td>
<td>15</td>
<td>171-&gt;185</td>
<td>1732.97</td>
</tr>
<tr>
<td>19</td>
<td>LLLTPNECFEGQVV</td>
<td>15</td>
<td>181-&gt;195</td>
<td>1689.98</td>
</tr>
<tr>
<td>20</td>
<td>KGQVNLQGETKK</td>
<td>15</td>
<td>191-&gt;205</td>
<td>1598.83</td>
</tr>
<tr>
<td>21</td>
<td>GETKVVIGEEYSFT</td>
<td>15</td>
<td>201-&gt;215</td>
<td>1686.88</td>
</tr>
<tr>
<td>22</td>
<td>EYSFTLKDKESTTLS</td>
<td>15</td>
<td>211-&gt;225</td>
<td>1746.94</td>
</tr>
<tr>
<td>23</td>
<td>SVTSLQLPGIEYKV</td>
<td>15</td>
<td>221-&gt;235</td>
<td>1632.91</td>
</tr>
<tr>
<td>24</td>
<td>IEYKVTEEDVTKDGY</td>
<td>15</td>
<td>231-&gt;245</td>
<td>1788.93</td>
</tr>
<tr>
<td>25</td>
<td>TKDGYKTSATLKDGD</td>
<td>15</td>
<td>241-&gt;255</td>
<td>1599.71</td>
</tr>
<tr>
<td>26</td>
<td>LKDGDVTDGYNLGD</td>
<td>15</td>
<td>251-&gt;265</td>
<td>1568.60</td>
</tr>
<tr>
<td>27</td>
<td>NLGDSKTSDKSTDEI</td>
<td>15</td>
<td>261-&gt;275</td>
<td>1623.69</td>
</tr>
<tr>
<td>28</td>
<td>STDEIVTNKRDTQV</td>
<td>15</td>
<td>271-&gt;285</td>
<td>1704.85</td>
</tr>
<tr>
<td>29</td>
<td>RDTQVPT</td>
<td>7</td>
<td>281-&gt;287</td>
<td>815.88</td>
</tr>
</tbody>
</table>

Residues that make up the αE3 epitope are in bold.
5.1.1.2 Epitope mapping using the T18.1 overlapping peptide library

The epitopes of the four selected Fab were mapped by screening their reactivity to the overlapping peptide library using dot blots (section 2.2.7.4). The individual 15-mers were spotted onto nitrocellulose membrane (GE Healthcare) in 2 μL drops containing 200 ng of peptide. Each membrane was probed with one of the four Fab and distinct peptide spots were detected (Figure 5.15). Interestingly, all four Fab bound to peptide 12, peptide 13 and peptide 14.

In the case of the αE3 Fab, this result generally matches the epitope identified by x-ray crystallography as peptides 12-14 (residues 111-145) contain seven of the fourteen epitope residues identified in the co-crystal (Figure 5.10, Figure 5.11). The binding of the αE3 Fab to peptide 13 is unsurprising given that all seven of these epitope residues are present in the peptide, allowing the formation of four hydrogen bonds (as well as numerous non-bonding interactions) during peptide binding. The binding of the αE3 Fab to peptide 14 is perhaps more surprising given that it only contains two epitope residues (F131 and K134) which provide non-bonded contacts (Figure 5.10). This suggests that these two residues are sufficient for binding of the αE3 Fab. Peptide 12 contains N125, which is able to form three hydrogen bonds with the αE3 Fab as well as Y122 and V124, which form non-bonded contacts. Again, this indicates that these residues are sufficient for binding of the αE3 Fab and, notably, this is independent of F131 and K134 in peptide 14. Unfortunately, the 5-residue overlap of the peptide library does not have high enough resolution to identify if individual residues are sufficient for binding of the αE3 Fab.
Figure 5.15 Mapping the linear epitopes of the four selected Fab using the T18.1 overlapping peptide library. Spots 1-29 are 15 amino acid length peptides covering the length of T18.1 with an overlap of 5 residues. Spot 30 is a full-length T1 negative control. Binding was detected with an anti-Human IgG (H+L)-HRP antibody diluted 1:2500. (A) Membrane probed with 100 nM αE3 Fab. (B) Membrane probed with 100 nM αH3 Fab. (C) Membrane probed with 500 nM αE1 Fab. (D) Membrane probed with 500 nM βC1 Fab. (E) Membrane probed with 100 nM B9 Fab. The B9 Fab is a T1 binding Fab (does not bind to T18.1) that acts as a control for non-specific binding of peptides. (F-H) The location of each peptide bound by the four Fab mapped onto the surface of T18.1 as follows: E= peptide 12; F= peptide 13; G= peptide 14. The pilin Lysine is shown in yellow.
In contrast, the dot blot has not detected binding of the αE3 Fab to peptide 5 which contains six nearly consecutive epitope residues (K45, K47, D48, G49, L50 and D51) that together can form five hydrogen bonds with αE3. This suggests that this region is not essential for the binding of the αE3 Fab to T18.1. It is noteworthy that this indicates that the light chain of the αE3 Fab (which only interacts with K47, D48 and G49) does not appear to be required for the binding of αE3 to T18.1.

The remaining three Fab all bind to peptide 12, peptide 13 and peptide 14. For comparison, the αH3 Fab shows a similar pattern to the αE3 Fab with stronger binding to peptides 12 and 13 and weaker binding to peptide 14. In contrast, the αE1 Fab and βC1 Fab show more even binding to all three peptides although this is purely qualitative and may not accurately reflect the locations of the epitopes. Without higher resolution data the epitopes of the αH3, αE1 and βC1 Fab cannot be defined at the molecular level but it is certainly striking that all four Fab appear to bind to the same 35-residue region of T18.1.

5.1.1.3 Identification of distinct epitopes within the immunogenic region bound by the four selected Fab

While the four Fab all bind to the same three peptides in the overlapping peptide library, previous data suggest that they may not share the same epitope within this 35-residue region. For example, the αE3 Fab is cross-reactive with T3.2, and αH3, αE1 and βC1 are type-specific (Figure 4.17, Figure 5.2). This suggests that the αE3 epitope contains elements that are conserved between T18.1 and T3.2, while the epitopes of αH3, αE1 and βC1 do not. Also, the flow cytometry experiments in section 5.2.3 clearly show that there are differences in epitope accessibility for each of the antibodies on the polymerised pilus.

In order to experimentally determine whether or not they are binding to the same epitopes, competition ELISAs were performed. The IgG1 were coated onto immunoplates and used to
capture recombinant T18.1. Each of the Fab was then bound to T18.1 and Fab binding was specifically detected using an anti-HA-HRP antibody (Roche) (section 2.2.7.1.4). The percentage of binding was calculated by comparing Fab binding to T18.1 in the presence of IgG1 with Fab binding to T18.1 in the absence of IgG1 (no competition control). In this system, Fab binding to T18.1 (which is already bound by an IgG1) indicates that there is no competition and the Fab and IgG1 have different epitopes. If the Fab are unable to bind then there is competition between the IgG1 and the Fab.

The anti-T18.1 antibodies clearly fall into two groups based on their patterns of competition with each other. Binding of the αE3 Fab to T18.1 is blocked by the αE3 and αH3 IgG1 but not the αE1 or βC1 IgG1 (Figure 5.16A). The αH3 Fab shows the same pattern of competition. This indicates that the epitopes of the αH3 and αE3 Fab either overlap or are the same and are distinct from the αE1 and βC1 epitopes. Supporting this, the binding of the αE1 Fab is not blocked by the αE3 or αH3 IgG1 but is blocked by the αE1 and βC1 IgG1 (Figure 5.16C). Again, βC1 shows the same pattern of competition confirming that the epitopes of αE1 and βC1 overlap or are the same and are distinct from the epitopes of αE3 and αH3.
Figure 5.16 Capture ELISA competing the four Fab and the four IgG1 to identify overlapping epitopes. Each IgG1 was coated onto immunoplates (Nunc) at 5 µg/mL. T18.1 (5 µg/mL) was captured by the IgG1. Fab binding to T18.1 was specifically detected using an anti-HA-HRP secondary antibody. The percentage of binding was calculated by comparing Fab binding to the no competition control (Fab binding to T18.1 in the absence of IgG1). (A) Competition with 1 nM αE3 Fab. (B) Competition with 1 nM αH3 Fab. (C) Competition with 10 nM αE1 Fab. (D) Competition with 100 nM βC1 Fab. Data is representative of three independent experiments.
To further investigate this, competition ELISAs were performed with the Fab and the overlapping peptide library to see if the patterns of competition could discriminate between overlapping and identical epitopes. The Fab were pre-incubated with peptide at a ratio of 1:5 and bound to T18.1-coated immunoplates (section 2.2.7.1.5). Total competition was shown by pre-incubating the Fab with full-length T18.1. The percentage of binding was calculated by comparing the binding of Fab in the presence of peptide to the binding of Fab to T18.1 in the absence of peptide (no competition control).

As expected from the co-crystal and dot blot data, the binding of the αE3 Fab to T18.1 is most strongly blocked by peptide 13 (Figure 5.17A). Binding is also reduced slightly by peptides 12 and 14. The pattern of competition for the αH3 Fab is similar with peptide 13 blocking Fab binding most strongly. However, peptides 12 and 14 decrease the binding of the αH3 Fab to T18.1 more than they block αE3 binding. In contrast, peptides 12, 13 and 14 appear to block the binding of both the αE1 and the βC1 Fab almost equally (Figure 5.17C-D).

Once again, the data show that the binding properties of the αE3 and αH3 Fab to T18.1 are similar, but not identical, suggesting that these two Fab have distinct but overlapping epitopes. Similarly, the αE1 Fab shows essentially identical behaviour to the βC1 Fab (as seen previously). These two Fab either target one epitope or have very tightly overlapped epitopes, which these immunoassays cannot distinguish from one another. Therefore, the 35-residue region of T18.1, which is covered by peptides 12-14, contains at least three different epitopes. A protein alignment of this region is presented and discussed in section 6.2.1.
Figure 5.17 Competition ELISA competing the four Fab and the peptides spanning their epitopes. The percentage of binding was calculated by comparing the binding of Fab to T18.1 in the presence of peptide to the binding of Fab to T18.1 in the absence of peptide (no competition control). (A) Competition with 25 nM αE3 Fab. (B) Competition with 25 nM αH3 Fab. (C) Competition with 100 nM αE1 Fab. (D) Competition with 100 nM βC1 Fab. Fab binding was detected with an Anti-Human IgG (H+L)-HRP antibody diluted 1:2500. Data is representative of three independent experiments.
5.2.7 Animals vaccinated with T18.1-containing vaccines generate circulating antibodies that bind to the same immunogenic region of T18.1

Given that all four Fab bind to the same region on T18.1, it was important to confirm that this region is also targeted by anti-T18.1 repertoires in vivo by the vaccinated animals. To do this, each of the four Fab was competed against sera from the five pil18 mice and the T18.1 vaccinated rabbit in competition ELISA. The Fab were competed with animal sera diluted 1:200 (and a 1% BSA ‘no competition’ control) and the reduction in Fab binding to T18.1 was detected (section 2.2.7.1.6). The percent binding was calculated by comparing Fab binding in the presence of vaccinated animal sera to the Fab binding in the no competition control.

Serum from each of the vaccinated animals strongly competes with each of the Fab when binding to T18.1 (Figure 5.18). The pil18 mouse sera reduce the binding of the Fab to T18.1 between 44% (B3 pil18 mouse and the βC1 Fab, Figure 5.18D) and 74% (B1 pil18 mouse and the αE3 Fab, Figure 5.18A) while serum from the T18.1 vaccinated rabbit reduces the binding of the Fab to T18.1 between 50% (αE1 Fab) and 56% (βC1 Fab).

This confirms that the vaccinated animals produce antibodies that bind to the same epitopes as the Fab and that the phage display selection process has isolated Fab which are, at least partially, representative of the animals’ anti-T18.1 responses.
Figure 5.18 Competition ELISA competing the Fab with sera from the vaccinated animals. The no competition control measures binding of Fab to T18.1 blocked with 1% BSA as a non-T18.1 binding serum. The percent binding was calculated by comparing Fab binding in the presence of vaccinated animal sera to the Fab binding in the no competition control. (A) Competition with 1 nM αE3 Fab. (B) Competition with 1 nM αH3 Fab. (C) Competition with 10 nM αE1 Fab. (D) Competition with 100 nM βC1 Fab. Fab binding was detected with an Anti-Human IgG (H+L)-HRP antibody diluted 1:2500. Data is representative of three independent experiments.
5.2.8 Human paediatric patients who generate an immune response to T18.1 produce circulating antibodies which bind to the same immunogenic region of T18.1

To investigate whether these epitopes are also targeted by human antibodies, sera from patients recruited as part of the Rheumatic Fever Risk Factor (RFRF) study were epitope mapped using the T18.1 peptide library. This study recruited paediatric patients diagnosed with a recent *S. pyogenes* infection followed by acute rheumatic fever from across the North Island, New Zealand. Initially, sera from 47 patients were screened by ELISA for reactivity to full-length T18.1. Of these, four patients had a robust response to T18.1 (data not shown). Three of these patients (RFRF0005, RFRF0039 and RFRF0165) had sufficient sera available to enable epitope mapping with the peptide library. The dot blots were performed in the same way as for the Fab (section 5.1.1.2) with the patient sera diluted 1:100 when probing the blots.

Interestingly, all three patients show responses to the same peptides as the anti-T18.1 Fab (Figure 5.19). The serum from patient RFRF0005 binds to peptides 13 and 14 with weaker binding to peptide 12, while the serum from patient RFRF0039 binds weakly to peptides 11-14. The serum from RFRF0165 binds to peptides 12-14. There was no information available on the *tee* type of the *S. pyogenes* strain that these patients were infected with prior to developing rheumatic fever. The overall weakness of the signals, despite the high serum dilution, suggest these dot blots may be measuring antibodies generated to a historical infection that occurred a considerable time prior to recruitment. Nonetheless, it is intriguing that the three patients appear to only target epitopes on these same peptides. While it is possible that other epitopes are targeted, but too weakly to be detected, the blots clearly demonstrate that the location of the dominant antibody response to T18.1 in these patients is contained in peptides 12-14 (A111-D145). Unfortunately, due to the limited amount of patient sera available,
additional immunoassays could not be performed to further analyse human antibody binding to these epitopes. However, these human data do confirm that the same region of T18.1 is immunogenic in mice, rabbits and humans and contains epitopes targeted by all three species.

Figure 5.19 Mapping the linear epitopes of three human patients using the T18.1 overlapping peptide library. Spots 1-29 are 15-residue peptides covering the length of T18.1 with an overlap of 5 residues. Spot 30 is a full-length T1 negative control. Binding was detected with an anti-Human IgG (H+L)-HRP antibody diluted 1:2500. (A) Membrane probed with serum from patient RFRF0005. (B) Membrane probed with serum from patient RFRF0039. (C) Membrane probed with serum from patient RFRF0165. (D) Membrane probed with serum from patient RFRF0169 who reacts to T1 but does not react to T18.1.

5.2.9 Analysing the protective capacity of T18.1 antibodies

The final part of this chapter describes two different methodologies used to test the protective capacity of the vaccinated animal sera and monoclonal IgG1 in vitro. The first section describes how an adapted HL-60 based opsonophagocytic killing assay (OPKA), which is routinely used in testing S. pneumoniae vaccines, can be used to measure opsonophagocytosis of S. pyogenes using anti-T antigen serum. This section of work was undertaken in the Goldblatt laboratory at University College London (UCL). Once again I would like to thank Professor David Goldblatt and Dr Marta Zancolli for hosting me, and the Maurice Wilkins Centre for funding this work.
The second section measures the protective capacity of the monoclonal IgG1 and animal sera using the traditional *S. pyogenes* bactericidal assay.

### 5.2.9.1 Measuring the protective capacity of anti-T18.1 rabbit sera in an adapted OPKA

In the OPKA the testing serum is serially diluted and incubated with the bacteria to allow antibody binding. Baby rabbit complement (BRC) and HL-60 cells are then added to the bacteria and the mixture is further incubated (2.2.10.1). Samples are then plated and colony forming units (CFU) are counted using an automated colony counter (ProtoCOL3; Synbiosis) (Figure 1.5). The assay result is presented as the opsonic index (OI), which is the dilution of the testing serum resulting in 50% killing. An OI >8 indicates killing.

At the time this preliminary work was undertaken, the four monoclonal IgG1 identified had not yet been produced. Due to the Goldblatt laboratory’s prior experience working with rabbit serum in HL-60 assays, the serum from the T18.1 vaccinated rabbit was used in these proof-of-principle experiments to ascertain whether T antigen antibodies induced killing through opsonophagocytosis. The T18.1 rabbit serum was tested against the vaccine strain (MGAS8232), another strain that also expresses the T18.1 protein (*emm*217_007; >99% identity), a strain expressing a closely related T18.2 protein (*emm*49_156; >90% identity) and two strains expressing the distantly related T1 (M1 (43); ~37% identity) and T6 proteins (M6 (2); ~25% identity) (Table 5.4).

In initial experiments the incubation length was 45 min as per the *S. pneumoniae* OPKA (Nahm & Burton, 2013; Romero-Steiner et al., 2006). With this short incubation time, only the M217/T18.1 strain was able to be killed with an OI of 80 (Figure 5.20; Table 5.4).
Figure 5.20 OPKA killing curve for M217/T18.1 after 45 min with rabbit anti-T18.1 serum and matched pre-immune serum. The testing sera were serially diluted 1:3 and incubated with M217/T18.1, BRC and differentiated HL-60 cells for 45 min. The samples were plated and the colony forming units (CFU) counted. The horizontal black line at y=132 represents 0% killing as determined by the growth of M217/T18.1 in the OPKA without testing serum. The horizontal dotted line at y=66 represents 50% killing. The OI is the dilution of the testing serum resulting in 50% killing.

When the incubation time was increased to >75 min, the M49/T18.2 strain was also able to be killed with an OI of 144 though the complement mediated non-specific killing (NSK) also increased (Table 5.4). This is killing due to complement activity independent of any antigen-specific antibodies. NSK was measured in the OPKA by two complement only controls. Control A contained *S. pyogenes*, differentiated HL-60 cells and heat-inactivated complement only and control B contained *S. pyogenes*, differentiated HL-60 cells and active complement only. NSK is calculated as a percentage of growth of the strain in the presence of active complement (control B) compared to the growth of the strain in the presence of heat inactivated complement (control A). While the non-specific killing (NSK) increased with these longer incubation times it was still within the acceptable range (<35%) for these assays.
Table 5.4 Summary of *S. pyogenes* killing achieved with the anti-T18.1 rabbit serum in the HL-60 OPKA. These data are representative of at least two individual experiments.

<table>
<thead>
<tr>
<th>Strain</th>
<th>M type</th>
<th>T type</th>
<th>Incubation</th>
<th>OI*</th>
<th>Killing</th>
<th>Max killing</th>
<th>NSK†</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGAS8232</td>
<td>M18</td>
<td>T18.1</td>
<td>45 min</td>
<td>2</td>
<td>N</td>
<td>13%</td>
<td>0%</td>
</tr>
<tr>
<td><em>emm217_007</em></td>
<td>M217</td>
<td>T18.1</td>
<td>45 min</td>
<td>80</td>
<td>Y</td>
<td>67%</td>
<td>10%</td>
</tr>
<tr>
<td><em>emm49_156</em></td>
<td>M49</td>
<td>T18.2</td>
<td>120 min</td>
<td>144</td>
<td>Y</td>
<td>93%</td>
<td>25%</td>
</tr>
<tr>
<td>M1 (43)</td>
<td>M1</td>
<td>T1</td>
<td>75 min</td>
<td>2</td>
<td>N</td>
<td>22%</td>
<td>8%</td>
</tr>
<tr>
<td>M6 (2)</td>
<td>M6</td>
<td>T6</td>
<td>75 min</td>
<td>2</td>
<td>N</td>
<td>11%</td>
<td>4%</td>
</tr>
<tr>
<td>M18HS66</td>
<td>M18</td>
<td>T18.1</td>
<td>75 min</td>
<td>2</td>
<td>N</td>
<td>22%</td>
<td>3%</td>
</tr>
<tr>
<td>M18HS66 pControl</td>
<td>M18</td>
<td>T18.1</td>
<td>75 min</td>
<td>2</td>
<td>N</td>
<td>29%</td>
<td>3%</td>
</tr>
<tr>
<td>M18HS66 pRocA_M89</td>
<td>M18</td>
<td>T18.1</td>
<td>75 min</td>
<td>960</td>
<td>Y</td>
<td>88%</td>
<td>17%</td>
</tr>
</tbody>
</table>

*Opsonic Index. †Complement-mediated non-specific killing.

Conditions that promoted the killing of the M18/T18.1 strain with the anti-T18.1 serum could not be found, despite this strain expressing T18.1 that is identical in sequence to the antigen used for vaccination. However, M18 strains are often hyper-encapsulated due to a conserved mutation in the *rocA* gene (Lynskey et al., 2013) and it was possible that the excess capsule was inhibiting opsonophagocytosis (Wessels et al., 1991). To test this hypothesis, a hyper-encapsulated M18HS66 strain and the non-hyper-encapsulated isogenic mutant M18HS66 pRocA_M89 (described in (Lynskey et al., 2013)) were kindly provided by Professor Shiranee Sriskandaran. When these strains were used in the OPKA, the M18HS66 pRocA_M89 strain
was successfully killed by the anti-T18.1 serum (OI=960) while the hyper-encapsulated M18HS66 and M18HS66\textsubscript{pControl} strains were not (Table 5.4). This confirmed that the \textit{S. pyogenes} hyaluronic acid capsule was anti-phagocytic and hyper-encapsulation was preventing the killing of MGAS8232. As expected (due to the lack of anti-T1 or anti-T6 antibodies in the anti-T18.1 serum (Figure 4.17)) the M1/T1 and M6/T6 strains could not be killed with anti-T18.1 serum in any conditions.

The specificity of killing observed in the OPKA was determined by measuring the ability of homologous (T18.1) and heterologous (M1 and M6) proteins to inhibit the killing of bacteria by anti-T18.1 serum. The recombinant proteins were pre-incubated with the anti-T18.1 serum for 30 minutes before the bacteria were added (section 2.2.10.1.1). Killing of all the T18 strains was blocked by pre-incubation of the serum with recombinant T18.1 but not M1 or M6 proteins (Table 5.5). This confirmed that the vaccine serum contained T18.1 antibodies able to induce opsonophagocytosis of T18.1 and T18.2 strains of \textit{S. pyogenes}.

Table 5.5 Testing the specificity of killing in the \textit{S. pyogenes} OPKA.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Incubation</th>
<th>OI</th>
<th>Max Killing</th>
<th>NSK</th>
<th>Killing with T18.1 (OI)</th>
<th>Killing with M1 (OI)</th>
<th>Killing with M6 (OI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{emm217_007}</td>
<td>45min</td>
<td>49</td>
<td>61%</td>
<td>30%</td>
<td>N (2)</td>
<td>Y (164)</td>
<td>N/D</td>
</tr>
<tr>
<td>\textit{emm49_156}</td>
<td>100min</td>
<td>144</td>
<td>93%</td>
<td>25%</td>
<td>N (5)</td>
<td>Y (148)</td>
<td>Y (104)</td>
</tr>
<tr>
<td>M18HS66\textsubscript{pRocA}</td>
<td>75min</td>
<td>329</td>
<td>75%</td>
<td>32%</td>
<td>N (2)</td>
<td>Y (403)</td>
<td>Y (OI 333)</td>
</tr>
</tbody>
</table>

N/D Experiment not done
The work to develop and optimise the OPKA for *S. pyogenes* in the Goldblatt laboratory continued after my visit in 2016. Indeed, the first paper describing the OPKA for selected strains has recently been published on which I am a co-author (Jones et al., 2018). Efforts are now underway to establish the *S. pyogenes* OPKA in the Moreland laboratory at the University of Auckland. However, the assay was not available in a timeframe that enabled testing of the four anti-T18.1 IgG1.

5.2.9.2 Measuring the protective capacity of the monoclonal IgG1 and vaccinated animal sera in bactericidal assays

The vaccinated animal sera and monoclonal IgG1 were tested for killing in a modified indirect bactericidal assay (Frost et al., 2017). While these assays have their limitations (such as high variability in donor whole blood, poor reproducibility and difficulty in comparing assays), in the absence of an established OPKA they are still able to provide some indication of protective capacity. Based on the findings in London that the M18/T18.1 strain (expressing T18.1 that is identical to the antigen used for vaccination) could not be killed due to hyper-encapsulation, M217/T18.1 was selected for this work. M217/T18.1 bacteria and the testing samples (sera or IgG1) were added to blood collected from a healthy volunteer who had tested negative for antibodies to T18.1 in an ELISA based pre-screen (data not shown). This assay mixture was incubated for 3 hours at 37°C in a sealed tube then plated and manually counted (section 2.2.10.2). The assay output is presented as percentage killing compared to bacterial growth in rabbit pre-immune serum, where >50% killing is considered a positive result (Dale, 1999; Dale et al., 2011).

Both the pil18 mouse and recombinant T18.1 rabbit sera were able to kill M217/T18.1 *S. pyogenes* (Figure 5.21) with the magnitude of killing being similar for both (88.0% and 81.5% for the mouse and rabbit respectively). The rabbit anti-T18.1 serum appeared to induce
strong killing, as previously seen in the OPKA, confirming that both assay formats can detect bactericidal activity (Figure 5.21, Table 5.4). All four of the monoclonal IgG1 were also able to induce >50% killing of the M217/T18.1 strain. The αH3 and αE1 IgG1 were able to induce killing that was comparable to the polyclonal antisera but the killing caused by the αE3 and βC1 IgG1 was somewhat lower. It is possible that the reason αE3 IgG1 only shows intermediate killing, despite having the second highest avidity for the T18.1 (Figure 5.1), may be due to a lack of epitope accessibility on the polymerised pilus (Figure 5.5). The αE3 IgG1 appeared to have the weakest binding to the polymerised pilus as detected by flow cytometry (Figure 5.5). The βC1 IgG1 had the lowest percentage killing in these assays and this may be a functional consequence of it being the lowest affinity IgG1 clone (Figure 5.4). Due to the time constraints associated with this project, bactericidal assays assessing cross-protection and investigating potential synergism between the IgG1 could not be performed.
Figure 5.21 The vaccinated animals and monoclonal IgG1 have bactericidal activity against M217/T18.1. Polyclonal sera were tested at a 1:6 dilution. The IgG1 clones were tested at 5 µg/mL. The dotted line marks 50% killing.

5.2.10 Discussion

This chapter has identified an immunodominant region on the N-domain of T18.1 that contains epitopes bound by all four of the antibodies isolated from the pil18 mouse library as well as mouse, rabbit and human sera. The epitope of the αE3 antibody was successfully mapped at molecular resolution using x-ray crystallography. This identified a fourteen-residue conformational epitope with nine hydrogen bonding interactions between αE3 and T18.1. This epitope could also be detected using an overlapping peptide library of T18.1. Peptide binding indicated that as few as two residues (F131 and K134) are sufficient for the, albeit weak, binding of αE3 to T18.1, while the light chain interactions are not required for binding.
The epitopes of the αH3, αE1 and βC1 Fab were also mapped to this immunodominant region. In all experiments performed, the αE3 and αH3 clones have shown similar characteristics with the notable exception that they have slightly different T-antigen cross-reactivity patterns (Figure 4.17, Figure 5.2). They have similar affinity, bind to the same peptides in the peptide library and show complete competition with each other when binding to T18.1 in competition ELISAs (Figure 5.16). This indicates that their epitopes are overlapping, but that some of the interacting residues are likely to be different, which alters their specificity (discussed in section 6.2). The epitope of αH3 is therefore predicted to be located across the ‘top’ of the N-domain (Figure 5.22) overlapping with the αE3 epitope (Figure 5.22, black).

In contrast, the epitopes of αE1 and βC1 are entirely distinct from the epitopes of αE3 and αH3. While αE3 and αH3 compete with each other and αE1 and βC1 compete with each other, there is no competition for binding to T18.1 between Fab from the two groups. Again, the αE1 and βC1 antibodies have exhibited similar characteristics throughout this project and the epitope mapping techniques used cannot separate their epitopes. The notable difference between them is that βC1 has markedly lower affinity for T18.1. It is possible that this is a result of the heavy chain gene having undergone less somatic hypermutation than αE1 (2.8% and 4.85% respectively; Table 4.2), which is a sign of antibody maturation. Somatic hypermutation can significantly increase antibody affinity (Methot & Di Noia, 2017; Peled et al., 2008), often through decreasing the rate of dissociation (England et al., 1999). This supports the observation that while the four antibodies have similar rates of association, βC1 has a much higher rate of dissociation (Figure 5.4). Despite this distinction, αE1 and βC1 bind to the same three peptides (peptides 12-14) in the overlapping library as αE3 and αH3. As shown in Figure 5.22, there is a significant region covered by these three peptides that is not bound by αE3. It is likely that the αE1 and βC1 antibodies bind to an epitope, or epitopes, in this region, which runs along the ‘side’ of the N-domain of T18.1 as opposed to across the top.
Figure 5.22 The immunogenic region mapped onto the surface of T18.1. The αE3 epitope residues are shown in black. The residues corresponding to peptides 12 are shown in blue. The residues corresponding to peptide 13 are shown in teal. The residues corresponding to peptide 14 are shown in green. The pilin lysine is shown in yellow.

As the epitope(s) of αE1 and βC1 are located on the side of the N-domain, they are most likely equally accessible in monomeric and polymeric T18.1 and should be insensitive to the conformation of the pilus. This is supported by their equal binding to the polymerised pilus at 4°C and 37°C (Figure 5.5). However, polymerisation of the pilus is predicted to partially obscure the epitopes of αE3 and αH3, which bind at the interface between T antigen monomers (Kang et al., 2007). Indeed, these two antibodies show clear temperature-dependent differences in binding to the polymerised pilus, such that more binding occurs at 37°C than at 4°C. This suggests that their epitopes are most accessible at physiological temperatures (Figure 5.5). Epitope accessibility has been identified as a significant factor determining the ability of antibodies to bind to epitopes on viral proteins in the context of the virion (Pierson & Diamond, 2008, 2012). Landmark studies with Dengue virus have shown that the virion can change the conformation of the envelope protein depending on environmental conditions, including
temperature, which in turn exposes epitopes that were not available at lower temperatures (Lim et al., 2017; Lok et al., 2008; Zhang et al., 2016). While there is a large body of literature describing this phenomenon in viruses, there is a lack of studies investigating structural changes that occur at physiological temperatures for large bacterial assemblies such as pili. Yet the data presented here suggest that marked structural motion does occur in the polymerised pilus. Because proteins are dynamic, the movement of the T18.1 monomers within the polymerised pilus at 37°C is likely to be temporarily exposing epitopes to antibody binding which are partially hidden in the low temperature, less mobile, pilus. Once αE3 and αH3 bind, their extremely slow dissociation rates (Figure 5.4) coupled with steric hindrance between the antibodies and pilus backbone may ‘lock them in place’ and prevent the epitopes being obscured again. This is discussed further in section 6.4.

The identification of a single 35-residue region to which all four of the selected antibodies bind is striking. While it is possible that this region is truly immunodominant, it is also possible the biopanning process was biased towards the selection of antibodies that bind to the N-domain of T18.1. The phagemid library was biopanned in solution against C-terminally biotinylated T18.1 that was coupled to streptavidin coated beads. Biopanning in solution ensured that T18.1 was presented in its native state and not adsorbed to plastic, which can denature the antigen and cause the loss of epitopes or the presentation of cryptic epitopes. (Schier et al., 1995; Zhuang et al., 2001). It also ensured that T18.1 was orientated with the C-domain located towards the bead and the N-domain exposed in solution, which reflects the orientation that the immune system would encounter during natural infection. It is possible that this may have preferentially selected for antibodies with epitopes in the N-domain. In order to truly confirm whether this is the case, the library would need to be re-panned with randomly oriented T18.1 and the domain specificity of the enriched antibodies mapped.
However, the sera from the vaccinated mice and rabbit show significant competition with the antibodies that bind to this region in all experiments. This indicates that the pil18 and T18.1 vaccinated animals developed significant serum antibody responses to this same region of T18.1. The same region is also bound by human patients who have presumably developed an antibody response to this region of T18.1 after natural infection. The patient sera bind exclusively to the same peptides as the antibodies isolated from the vaccinated mice which indicates that these epitopes are also targeted by human antibodies in vivo. It should be noted that due to the linear nature of peptides libraries it is possible that not all epitopes could be identified with this method (Gershoni, Roitburd-Berman, Siman-Tov, Freund, & Weiss, 2007; Van & Pellequer, 1994). The fact that multiple sera show reactivity to the same region of T18.1 confirms that the antibodies selected by phage display bind to epitopes that are representative of the epitopes targeted by circulating antibodies produced in vivo in three different species.

These epitopes have been shown to be protective with all four of the monoclonal antibodies having bactericidal activity against T18.1 strains of S. pyogenes. Sera from the pil18 vaccinated mice and the T18.1 vaccinated rabbit were also shown to be bactericidal. Although combinations of proteins from the S. pyogenes pilus have been shown to be protective in a number of models (Loh et al., 2017; Mora et al., 2005; Tsai et al., 2017), this is the first time that isolated T antigen antibodies have been shown to be bactericidal.

Serum from the T18.1 vaccinated rabbit was also shown to be protective in a newly developed OPKA. Using this method, antibodies to T18.1 were shown to be cross-reactive and induced the killing of T18.1 and T18.2, but not T1 or T6, strains through opsonophagocytosis. However, hyper-encapsulation was shown to inhibit T antigen dependent opsonophagocytosis with hyper-encapsulated strains being completely protected (Table 5.4). Hyper-encapsulation has also been observed to protect against M protein antisera in bactericidal assays (Wessels et al.,
The precise manner in which hyper-encapsulation protects strains from opsonophagocytosis is not known. As the hyaluronic acid (HA) component of the capsule is immunogenically silent, it may simply be a masking effect where the extra bulk of the capsule covers epitopes and prevents the binding of complement proteins (Moses et al., 1997) and antibodies. As both HA (Fraser et al., 1997) and the surface of the neutrophils (Gallin, 1980; Wardle, 1986) are negatively charged, there may also be repulsive forces inhibiting phagocytosis even if protective antibodies are able to bind.

The OPKA provides significantly more information than the bactericidal assay about how *S. pyogenes* are killed. It measures the complement-mediated non-specific killing, which can be significant and is difficult to quantify in the traditional bactericidal assays that use human whole blood as a source of both complement and neutrophils. The testing sera or antibodies are also titrated in the OPKA whereas they are used at a single, usually high, concentration in bactericidal assays. The high throughput nature of the OPKA, as well as the standardisation of phagocytes and complement, also simplifies the comparison of test samples. However, it is apparent that, unlike for *S. pneumoniae*, the OPKA needs to optimised for each strain, as factors such as the incubation time significantly affect opsonophagocytosis (Table 5.4).

The OPKA has the potential to become a standardised, reproducible assay for testing immunity to surface-expressed *S. pyogenes* antigens. In the time since the work in this thesis was completed, the OPKA has been validated as a method for assessing opsonophagocytosis caused by antibodies to the M protein (Jones et al., 2018).
6 Discussion

6.1 Project Summary

The aim of this study was to investigate the interactions between antibodies and T antigens using T18.1 as a representative 2-domain T antigen. The structure of T18.1 has been solved using x-ray crystallography and found to be highly conserved when compared to the evolutionarily distinct T1. Regions of type-specific variation were mapped to the loops and surface of T antigens concealing a conserved core. Ten Mice were immunised with *L. lactis* expressing a full T18.1-containing pilus (pil18), and a rabbit was vaccinated with recombinant T18.1. Immunisation with T18.1 containing vaccines elicited a cross-reactive antibody response to T3.2, T13 and T18.2 in all of the animals. Minor cross-reactivity with other T antigens was also observed in some animals. An antibody library was generated from five of the pil18 vaccinated mice and phage display methods were used to identify twenty unique antibodies which bound to T18.1 with high affinity. Four of these antibodies were biochemically and biophysically characterised and had their epitopes mapped to a 35-residue region on the N-domain of T18.1. This region was confirmed to be immunogenic in mice, rabbits and humans. The binding of two of these antibodies (αE3 and αH3) to the polymerised pilus was impaired at 4°C, with increased binding observed at 37°C, while the binding of two other antibodies (αE1 and βC1) was not impaired at 4°C. This suggested that the epitopes of αE3 and αH3 are partially inaccessible at 4°C but become more accessible at physiological temperatures. Finally, the animal sera and all four of the monoclonal antibodies were shown to be bactericidal *in vitro* while the anti-T18.1 rabbit sera was also shown to be opsonic in a newly adapted opsonophagocytic killing assay (OPKA).
The work presented here adds to our understanding of how antibodies interact with T antigens, provides a basis for the rational design of T antigen based vaccines and adds molecular detail to the model of the assembled pilus.

6.2 Understanding Antibody-T antigen interactions

6.2.1 Insight into T antigen cross-reactivity

This project has identified an immunogenic region on T18.1 (A111-D145) which is bound by four antibodies generated from mice vaccinated with pil18. The αE3 and αH3 antibodies were shown to have overlapping epitopes as they compete for binding to T18.1 but have different specificity patterns when screened against a T antigen panel comprised of fourteen proteins that cover all of the major circulating tee types (Figure 4.17, Figure 5.16). The αE3 antibody is cross-reactive and binds strongly to T3.2, T18.1 and T18.2, with weak cross-reactivity to T13, whereas the αH3 antibody is type-specific and only binds to T18.1 and T18.2 (Figure 4.17, Figure 5.2). To investigate the basis of this difference in specificity, the immunogenic region in the eleven 2-domain T antigens that were part of the panel were aligned. In this protein alignment, the cross-reactive T antigens (T3.2, T13, T18.1 and T18.2) cluster together suggesting that there are conserved elements which contribute to cross-reactivity (Figure 6.1). Overall, the core residues in this region are fairly conserved across all of the proteins (other than the evolutionarily distinct T1) and not just among the cross-reactive T antigens. Indeed, of the residues in the core αE3 epitope (determined from the αE3-T18.1 co-crystal), only R126 is unique to T18.1 with Y122, V124, N125, G129, F131 and K134 all conserved in the majority of the T antigens in the panel. While R126 is unique, the cross-reactive T antigens (as well as T5) are the only other T antigens to have a positively charged residue (lysine) in this position. It is possible that a positive residue at this position within the core epitope enables cross-reactivity between T3.2, T13, T18.1 and T18.2. The lack of cross-reactivity defined for the
αH3 antibody suggests that the specificity of this antibody is not dictated in the same way. However, without molecular detail of the αH3 epitope this cannot be confirmed.

Figure 6.1 Protein alignment of the 35-residue immunogenic region (A111-D145) in all 2-domain T antigens in the panel. The residues are numbered according to the sequence of T18.1. The T antigens that the αE3 antibody cross-reacts with are bracketed. Residues which are part of the αE3 epitope on T18.1 are shown in red boxes. The coloured lines represent peptide 12 (blue), peptide 13 (teal) and peptide 14 (green) from the overlapping peptide library. The alignment was generated using ClustalW and ESPript.

Like αH3, the αE1 and βC1 antibodies are type-specific and only bind to T18.1 and T18.2 (Figure 4.17), but they do not compete with αH3 (or αE3) for binding to T18.1. This indicates that the epitope, or epitopes, of αE1 and βC1 do not overlap with the epitopes of αH3 and αE3. However, epitope mapping with the peptide library indicates that αE1 and βC1 do bind to this 35-residue region (Figure 5.15). Competition ELISAs with the individual peptides revealed that the patterns of competition are different for αE3 and αH3 compared to αE1 and βC1 (Figure 5.17). The binding of the αE3 and αH3 antibodies to T18.1 is most strongly inhibited by peptide 13, which indicates that the residues in the core of their epitopes are located on peptide 13. In the case of αE3, this agrees with the epitope residues identified in the αE3-T18.1 co-crystal and is illustrated in Figure 6.1. In contrast, the binding of αE1 and βC1 to T18.1 is inhibited almost equally by peptide 12, peptide 13 and peptide 14 (Figure 5.17). This indicates that the residues
in the cores of their epitopes are spread across the three peptides. While the residues represented by peptide 13 are relatively conserved across the 2-domain T antigens, the residues in peptides 12 and 14 are less conserved (Figure 6.1). They contain flanking regions (114-116 and 139-142) that display many subtle differences in the charge, polarity and size of the residues. If αE1 and βC1 bind to residues in these variable flanking regions, which the competition ELISA data suggest they may, the complex pattern of variation between T antigens would likely prevent them from cross-reacting with non-T18 T antigens.

The epitope only forms one half of the antibody-T antigen interface with the paratope (the residues on the antibody which interact with the epitope) forming the other. Protein alignments of the CDRs from the four antibodies are shown in Figure 6.2. The residues found in the CDRs of the four antibodies are diverse, particularly in CDR3 of the heavy chain (HCDR3). This is unsurprising since the HCDR3 is the main determinant of an antibody’s specificity (Noel et al., 1996; Xu & Davis, 2000). In the case of αE3 binding to T18.1, the HCDR3 contains 5/16 paratope residues which form 3/9 of the hydrogen bonds between αE3 and T18.1. Given the variability of the CDRs, it is difficult to draw conclusions as to how these produce specificity. Differences in the length of the HCDR3 coupled with even seemingly small amino acid substitutions will alter the complementarity of the interface meaning that each antibody will interact with its epitope in a unique manner. Even if the epitope contains conserved residues found in the core of the 35-residue immunogenic region, the subtle differences in paratope-epitope interactions will ultimately determine T antigen specificity.

Characterisation of antibody-antigen structures deposited in the Protein Data Bank (PDB) has shown that paratopes are often enriched in the aromatic amino acids tyrosine (Y), tryptophan (W) and phenylalanine (F) (Fellouse, Wiesmann, & Sidhu, 2004; Koide & Sidhu, 2009; Ramaraj, Angel, Dratz, Jesaitis, & Mumey, 2012) and that these are energetically important for
antibody binding (Bostrom et al., 2009; Dall’Acqua, Goldman, Eisenstein, & Mariuzza, 1996). Interestingly the lower affinity αE1 and βC1 antibodies only have two of these aromatic residues in their HCDR3 compared to αE3 and αH3 which have six and five respectively (Figure 6.2). This may partially account for the somewhat lower affinities of αE1 and βC1 for T18.1 (Table 4.3, Figure 5.4).

Figure 6.2 Protein alignment of the CDRs of the four antibody clones identified in this project. (A) Protein alignment of the light chain CDRs. (B) Protein alignment of the heavy chain CDRs. The residues which make up the paratope on αE3 are shown in green boxes. The CDRs were identified using IMTG/V-QUEST. The alignment was generated using ClustalW and ESPript.
6.2.2 The heavy chain dominant binding mode of αE3

Canonically, the CDRs of both the heavy and light chains make up an antibody’s paratope and contribute to antigen-specificity. While the HCDR3 is considered to be the main determinant of binding (Xu & Davis, 2000), the magnitude of heavy chain dominance appears to vary significantly between antibody-antigen complexes. Mapping the epitope of the αE3 antibody at the molecular level has revealed that the binding of αE3 to T18.1 is dominated by the heavy chain (Figure 5.9). The heavy chain interacts with thirteen of the fourteen epitope residues while the light chain only interacts with three epitope residues (Figure 5.10). Furthermore, αE3 is able to bind to peptides in the overlapping library that only contain residues bound by the heavy chain (Figure 5.15). This indicates that, unusually, the light chain of αE3 is not necessary for binding to T18.1.

This phenomenon has been observed before by Noel et al. (Noel et al., 1996) when the heavy and light chains of the TG10 antibody were expressed independently and assessed for binding to thyroglobulin. The heavy chain was able to bind to thyroglobulin in isolation but the cognate light chain was not. However, the affinity of the heavy chain for thyroglobulin was ten-fold lower than that of TG10, suggesting that while the light chain does not initiate binding or determine specificity, it may stabilise the TG10-thyroglobulin complex (Noel et al., 1996).

Even more extreme examples of heavy chain dominant antigen-binding have been observed with broadly neutralising antibodies to HIV-1 and Influenza virus. The binding of the b12 antibody to its epitope on HIV-1 gp120 is solely mediated by the heavy chain (Zhou et al., 2007). Unlike αE3 or TG10 where the light chain has a minor role in binding, the closest residue on the b12 light chain is over 10Å from gp120. Similarly, the binding of the F10 (Sui et al., 2009) and CR6261 (Ekiert et al., 2009) antibodies to their epitopes on Influenza hemagglutinin is exclusively heavy chain driven. The light chain of CR6261 is separated from hemagglutinin
by at least 8Å (Ekiert et al., 2009) while the light chain of F10 is oriented into solution away from hemagglutinin (Sui et al., 2009).

This suggests that the influence that the heavy chain has in determining antigen-specificity falls along a spectrum from canonical heavy and light chain mediated antigen-binding to complete heavy chain dominance with the binding of αE3 to T18.1 somewhere in between.

6.2.3 Understanding T antigen-based protection

This project has shown for the first time that anti-T antigen antibodies in isolation are protective. Over the course of this project the four monoclonal antibodies and the sera from the pil18 immunised mice and the T18.1 vaccinated rabbit were shown to be bactericidal in vitro. The T18.1 vaccinated rabbit serum was also shown to be opsonic in a newly adapted OPKA for *S. pyogenes* (Jones et al., 2018). This OPKA is being established in the Moreland laboratory at the University of Auckland but was not available in a timeframe that enabled testing of the four anti-T18.1 IgG1 in this project. When the OPKA is established it will allow experiments to be performed which will further add to the understanding of T antigen based protection.

The T18.1 vaccinated rabbit serum has already been shown to induce the killing of M217/T18.1 and M49/T18.2 strains of *S. pyogenes* (Table 5.4) in the OPKA, and it is potentially more cross-protective than this. When screened against the T antigen panel, the anti-T18.1 rabbit serum also showed cross-reactivity to T3.2, T11, T13 and T28.1 (Figure 4.17). The OPKA will be used to screen this serum against a panel of *S pyogenes*, including those carrying these cross-reactive T antigens, to identify the breadth of protection afforded by the monomeric T18.1 vaccine.
Similarly, serum from each of the pil18 vaccinated mice reacts with T3.2, T18.1 and T18.2 with individual mice also showing various patterns of cross reactivity to T5, T11, T13 and T28.1 (Figure 4.17). The OPKA will first be used to assess whether or not the pil18 mouse sera are opsonic and then to measure the breadth of cross-protection. This will enable the protective capacity of the monomeric T18.1 vaccine and the whole pilus vaccine to be directly compared (with the caveat that different species have been vaccinated). The two vaccine sera have a similar capacity to kill M217/T18.1 at the single 1:6 dilution used in the bactericidal assay (Figure 5.21). As the OPKA uses titrated serum it may reveal differences in protection that aren’t apparent in the bactericidal assay. For example, the pil18 vaccine contains polymerised T18.1 in addition to the two accessory proteins AP1 and AP2 whereas the recombinant T18.1 vaccine only contains monomeric T18.1 (Figure 4.2). If antibodies to the AP proteins contribute significantly to opsonophagocytosis then the Opsonic Index (OI) of the pil18 mouse sera may be greater than the OI of the T18.1 vaccinated rabbit serum.

The protective capacity of the four monoclonal antibodies generated in this project will also be tested in the OPKA. While all four antibodies have been shown to kill M217/T18.1 in a bactericidal assay (Figure 5.21), αH3, αE1 and βC1 only bind to T18.1 and T18.2. Therefore they are only likely to be protective against strains carrying these T antigens. In contrast, the αE3 antibody exhibits the major cross-reactivity pattern observed in this thesis (between T3.2, T13, T18.1 and T18.2) and is potentially cross-protective.

Furthermore, the titration of antibody in the OPKA may aid in the understanding of the biochemical nature of T antigen-based protection. Antibody-mediated protection is a product of both antibody affinity and epitope accessibility (Dowd & Pierson, 2011). The four anti-T18.1 antibodies identified in this project form two groups based on the affinity of their paratope-epitope interactions. The αE3 and αH3 antibodies are higher affinity (3.1 nM and 1.8
nM respectively) and αE1 and βC1 are lower affinity (30 nM and 63 nM respectively) (Table 4.3). However, flow cytometry has shown that αE3 and αH3 bind less well to the polymerised pilus than αE1 and βC1, particularly at 4°C (Figure 5.5). This is likely due to their epitopes (located on the top of the N-domain of T18.1) being partially occluded in the polymerised pilus. This did not appear to inhibit killing in bactericidal assays, but these were only performed at a single concentration of IgG1 (5μg/mL) (Figure 5.21). In the OPKA, the IgG1 clones will be titrated down to low concentrations and the dose-response curves may reveal differences in the protective capacity of the four antibodies that allow the contributions of affinity and epitope accessibility in protection to be teased apart.

The interplay between affinity and epitope accessibility has been shown to be important for anti-viral antibodies, but there is an absence of studies that have explored these relationships for antibodies that bind to large multimeric structures. Notable viral examples include neutralisation of flaviviruses by antibodies that bind to the viral envelope protein. The envelope protein is arranged in 90 anti-parallel dimers across the virion surface, and neutralisation occurs when a threshold number of antibodies bind to the envelope proteins (Pierson & Diamond, 2008). Generally, antibodies that bind to the envelope protein with high affinity are more potent neutralising antibodies than those that bind with low affinity (Dowd & Pierson, 2011). However, this is not always the case as antibodies with similarly high affinities can differ in their capacity to neutralise flaviviruses by several orders of magnitude. This has been attributed to differences in the relative accessibility of their epitopes (Nelson et al., 2008; Pierson et al., 2007). It is possible that a similar phenomenon will be observed with the antibodies identified in this thesis. In this case, αE3 and αH3 are high affinity antibodies that bind to partially occluded epitopes on the polymerised pilus whereas αE1 and βC1 are lower affinity clones that bind to more accessible epitopes. The combination of affinity and epitope accessibility will ultimately determine which of the four are the most protective.
6.3 Implications for Vaccinology

Showing that T antigen binding antibodies are able to kill *S. pyogenes* in isolation confirms that the T antigen has potential as a vaccine candidate. Currently, the most advanced *S. pyogenes* vaccine candidate is the 30-valent vaccine which combines the M protein hyper-variable regions (HVR) from 30 strains into four chimeric proteins (Dale et al., 2011). Due to the extreme antigenic diversity of the HVR there are over 220 *emm* types and the 30-valent vaccine has been predicted to have varying levels of coverage depending on the circulating strains in a given location (Dale et al., 2011; Dale et al., 2013; Engel et al., 2014; Williamson et al., 2015; Williamson et al., 2016). When compared to the M protein, the T antigen displays significantly less antigenic diversity with only eighteen *tee* types and therefore would require fewer antigens to reach similar levels of coverage as the HVR-based vaccines (Steemson et al., 2014). Indeed, it has been estimated that a combination of only twelve T antigens would provide 90% coverage of circulating strains (Falugi et al., 2008). This number might be reduced further if the major T antigen cross-reactivity pattern observed with the pil18 mouse sera, the T18.1 rabbit sera and the αE3 antibody (T3.2, T13, T18.1 and T18.2) results in cross-protection in the OPKA. This would demonstrate that a T18.1 vaccine could protect against non-T18 strains and that T antigens from those cross-protected strains would not be required in a vaccine.

Alternatively, a structural vaccinology approach similar to that taken with Group B Streptococcus (GBS) pili could be considered. The structure of the GBS pilus is similar to the structure of the *S. pyogenes* pilus, with two accessory proteins and a pilus shaft comprised of repeating units of the backbone protein (BP, analogous to the T antigen). Antibodies to the BP are protective but, in GBS, the BPs are encoded by three pilus islands, only two of which are highly conserved across GBS strains (Margarit et al., 2009). The third pilus island encodes six
immunological variants of the BP and antibodies to each of the variants only confer variant-specific protection (Margarit et al., 2009). Nuccitelli et al. mapped the protective epitopes on the six BPs (which are comprised of four domains) to a single variable domain, D3. They then fused the D3 domains from the six BP-variants into a chimeric protein vaccine and showed that vaccinated mice were protected against GBS strains expressing any of the six variants (Nuccitelli et al., 2011). In a similar manner, the protective αE3 epitope identified in this project could be mutated onto the surface of a structurally comparable T antigen, for example T1, in order to make a chimera that could potentially protect against T1 and T18.1 (and potentially T3.2, T13 and T18.2). The identification of protective epitopes on other T antigens and their incorporation into the chimera (assuming that they are not all located in the same place as the αE3 epitope) could potentially widen the breadth of protection even further. This is not dissimilar to the concept of ‘Pilvax’ where peptides from unrelated proteins are inserted into certain loops of T1 from S. pyogenes (Wagachchi et al., 2018). When the whole pilus is expressed, each copy of T1 in the pilus shaft is decorated with the vaccine peptide. This enables the vaccine peptide to be displayed in high copy number due to the repeating nature of the pilus backbone. This increases the immunogenicity of the peptide, and vaccination with Pilvax has been show to elicit the production of antibodies to both the vaccine peptide and the T1 carrier (Wagachchi et al., 2018).

A multivalent vaccine approach drawing inspiration from the 30-valent vaccine could also be considered. This project has identified protective epitopes in the N-domain of T18.1 and it is possible that T antigen based protection is targeted towards one domain. In the case of T18.1, the C-domain may not be necessary to induce vaccine-mediated immunity. Therefore, single protective domains from numerous T antigens could be identified and fused together (with appropriate linkers to prevent unwanted inter-domain interactions) into hybrid proteins analogous to the HVRs in the 30-valent vaccine (Dale et al., 2011). However, further
investigation into T antigen-based protection is needed before this could be done. Using T18.1 as an example, the N-domain would need to be shown to be as protective as the whole T antigen. This could be done by vaccinating with single domain constructs of T18.1 and comparing the protective capacity of the single domains to each other and to the full-length protein. Alternatively, single domain constructs could be used in the OPKA to determine domain-specific inhibition of opsonophagocytosis just as the full-length T18.1 protein was used to show specificity (Section 2.2.10.1.1, Table 5.5). A similar procedure would need to be repeated for the other T antigens in the vaccine construct to confirm that their single domains also confer protection.

6.4 Implications for the Streptococcus pyogenes pilus model

Solving the structure of the αE3-T18.1 complex has shown that the αE3 Fab binds across the top of the N-domain of monomeric T18.1 (Figure 5.9). This is the surface that interfaces with the C-domain of the preceding monomer in the polymerised pilus (Kang et al., 2007). Due to the interactions between T antigen monomers across this surface, polymerisation of the pilus is predicted to mask the epitope of αE3 and inhibit binding. This is supported by ELISA (Figure 5.1), SPR (Figure 5.4) and Flow cytometry data (Figure 5.5) showing that while αE3 binds with high affinity to the T18.1 monomer, it binds somewhat poorly to the polymerised pilus. The inhibitory effect is particularly pronounced at 4°C with the binding of αE3 increasing >2-fold between 4°C and 37°C. This indicates that there is significant temperature induced movement in the polymerised pilus which alters the accessibility of the αE3 epitope.

The current model of pilus assembly (based on a crystal structure from crystals grown at 18°C) proposes that the T antigens stack end-on-end with a 120° rotation along the long axis. This causes the wedge shaped base of the C-domain of one monomer to slot into the groove in the top of the N-domain of the next monomer burying ~850Å² of surface (Kang et al., 2007).
However, rather than being a rigid structure, the work presented in this thesis suggests that there is some flexibility in the pilus fibres. Indeed, electron microscopy (EM) has also shown that the pilus fibres as a whole exhibit some ‘give’ with immunogold labelled pilus shafts displaying some curvature (Mora et al., 2005; Quigley et al., 2010). This flexibility is likely possible due to temperature induced movement between adjacent T antigen monomers. Evidence for this is provided by the structural change required for the αE3 Fab to bind to the polymerised pilus (Figure 6.3). At 4°C the interactions between T antigen monomers may hold the pilus in a fairly rigid structure. In this conformation the epitope of αE3 is mostly inaccessible due to steric hindrance between αE3 and the C-domain of the preceding T antigen monomer (Figure 6.3B) and, the binding of αE3 to the pilus is inhibited (as measured by flow cytometry (Figure 5.5). At 37°C, αE3 is able to bind to the polymerised pilus more effectively (Figure 5.5), indicating that the epitope is being exposed at physiological temperatures. Thus, there must be enough flexibility between the polymerised monomers to firstly expose the epitope, and then to accommodate the bulk of αE3 once bound (Figure 6.3C). While Figure 6.3 shows a simplistic representation of a complex three-dimensional phenomenon, it likely represents the range of motion that the T antigen monomers may have in relation to each other at physiological temperatures. In order to visualise this more accurately, EM could be performed where T18.1 S. pyogenes (or L. lactis expressing the T18.1 pilus) are incubated with the αE3 Fab at 37°C and then the binding of αE3 is detected by immunogold labelling. This may show the ‘kinks’ that would be introduced into the pilus by αE3 binding, and allow the flexibility between T antigen monomers to be quantified.
The binding of the αE3 Fab to the polymerised T18.1 pilus indicates significant flexibility between T antigen monomers. (A) Ribbon diagram showing the polymerisation of T18.1 subunits in the pilus. This is modelled on the proposed structure of the T1 pilus (Kang et al., 2007). The inter-domain TQVPT linker is shown in black. The final threonine in this linker forms an isopeptide bond (catalysed by Sortase C) with the pilin lysine in the N-domain of the bottom T18.1 subunit to covalently join the subunits. (B) The structure of the αE3-T18.1 complex overlaid onto the model polymer shown in (A). The preceding T18.1 subunit (top) is shown in space filling form to highlight the clash with the αE3 Fab. (C) The movement of T18.1 subunits within the pilus to accommodate binding of the αE3 Fab.

6.5 Conclusion

By studying the interactions between antibodies and a model T antigen, insight has been gained into how antibodies bind to complex bacterial structures such as pili. They have been revealed to induce significant movement in the backbone of the pilus, which was previously modelled as a rigid structure. Each of the antibodies studied in this thesis is highly protective and this knowledge will inform the design of current and future S. pyogenes vaccines.
7 Appendix A: Primers for the construction of the chimeric mouse/human Fab library

Primer mixes for the amplification of mouse Vκ and VH genes were made from the following primers:

<table>
<thead>
<tr>
<th>MSCVK primer mix</th>
<th>Primer sequence 5’ to 3’</th>
<th>Volume added to primer mix (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSCVK-1</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TCC AGC TGA CTC AGC C</td>
<td>1</td>
</tr>
<tr>
<td>MSCVK-2</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG TTC TCW CCC AGT C</td>
<td>2</td>
</tr>
<tr>
<td>MSCVK-3</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG TGM TMA CTC AGT C</td>
<td>5</td>
</tr>
<tr>
<td>MSCVK-4</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG TGY TRA CAC AGT C</td>
<td>3.5</td>
</tr>
<tr>
<td>MSCVK-5</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG TRA TGA CMC AGT C</td>
<td>4</td>
</tr>
<tr>
<td>MSCVK-6</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG TGC TCA WCC AGT C</td>
<td>7</td>
</tr>
<tr>
<td>MSCVK-7</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTC AGA TGA YDC AGT C</td>
<td>2</td>
</tr>
<tr>
<td>MSCVK-8</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TYC AGA TGA CAC AGA C</td>
<td>1.5</td>
</tr>
<tr>
<td>MSCVK-9</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG TTC TCA WCC AGT C</td>
<td>2</td>
</tr>
<tr>
<td>MSCVK-10</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG WGC TSA CCC AAT C</td>
<td>3.5</td>
</tr>
<tr>
<td>MSCVK-11</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTS TRA TGA CCC ART C</td>
<td>8</td>
</tr>
<tr>
<td>MSCVK-12</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTK TGA TGA CCC ARA C</td>
<td>8</td>
</tr>
<tr>
<td>MSCVK-13</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG TGA TGA CBC AGK C</td>
<td>6</td>
</tr>
<tr>
<td>MSCVK-14</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG TGA TGA CYC AGG A</td>
<td>2</td>
</tr>
<tr>
<td>MSCVK-15</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG TGA TGA AGW T</td>
<td>2</td>
</tr>
<tr>
<td>MSCVK-16</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTG TGA TGA CAC AAC C</td>
<td>1</td>
</tr>
<tr>
<td>MSCVK-17</td>
<td>GGG CCC AGG CGG CCG AGC TCG AYA TTT TGC TGA CTC AGT C</td>
<td>1</td>
</tr>
<tr>
<td>MHybLJ-B mix</td>
<td>Vκ 3’ antisense primers</td>
<td>Primer sequence 5’ to 3’</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>MHybJK12-B</td>
<td>AGA TGG TGC AGC CAC AGT TCG TTT KAT TTC CAG YTT GGT CCC</td>
<td>1</td>
</tr>
<tr>
<td>MHybJK4-B</td>
<td>AGA TGG TGC AGC CAC AGT TCG TTT TAT TTC CAA CTT TGT CCC</td>
<td>1</td>
</tr>
<tr>
<td>MHybJK5-B</td>
<td>AGA TGG TGC AGC CAC AGT TCG TTT CAG CTC CAG CTT GGT CCC</td>
<td>1</td>
</tr>
<tr>
<td>VH 5’ sense primers</td>
<td>Primer sequence 5’ to 3’</td>
<td>Volume added to primer mix (µL)</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>MHyVH1</td>
<td>GCT GCC CAA CCA GCC ATG GCC CTC GAG GTR MAG CTT CAG GAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>4</td>
</tr>
<tr>
<td>MHyVH2</td>
<td>GAG GTB CAG CTB CAG CAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>4</td>
</tr>
<tr>
<td>MHyVH3</td>
<td>GAG GTG CAG CTG AAG SAS TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>3</td>
</tr>
<tr>
<td>MHyVH4</td>
<td>GAG GTC CAR CTG CAA CAR TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>4</td>
</tr>
<tr>
<td>MHyVH5</td>
<td>GAG GTY CAG CTB CAG CAR TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>7</td>
</tr>
<tr>
<td>MHyVH6</td>
<td>GAG GTY CAR CTG CAG CAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>2</td>
</tr>
<tr>
<td>MHyVH7</td>
<td>GAG GTC CAC GTG AAG CAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>1</td>
</tr>
<tr>
<td>MHyVH8</td>
<td>GAG GTG AAS STG GTG GAA TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>2</td>
</tr>
<tr>
<td>MHyVH9</td>
<td>GAG GTG AWG YTG GTG GAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>5</td>
</tr>
<tr>
<td>MHyVH10</td>
<td>GAG GTG CAG SKG GTG GAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>2</td>
</tr>
<tr>
<td>MHyVH11</td>
<td>GAG GTG CAM CTG GTG GAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>2</td>
</tr>
<tr>
<td>MHyVH12</td>
<td>GAG GTG AAG CTG ATG GAR TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>2</td>
</tr>
<tr>
<td>MHyVH13</td>
<td>GAG GTG CAR CTT GTT GAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>1</td>
</tr>
<tr>
<td>MHyVH14</td>
<td>GAG GTR AAG CTT CTC GAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>2</td>
</tr>
<tr>
<td>MHyVH15</td>
<td>GAG GTG AAR STT GAG GAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>2</td>
</tr>
<tr>
<td>MHyVH16</td>
<td>GAG GTT ACT CTR AAA GWG TST G GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>5</td>
</tr>
<tr>
<td>MHyVH17</td>
<td>GAG GTC CAA CTV CAG CAR CC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>3.5</td>
</tr>
<tr>
<td>MHyVH18</td>
<td>GAG GTG AAC TTG GAA GTG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>0.7</td>
</tr>
<tr>
<td>MHyVH19</td>
<td>GAG GTG AAG GTC ATC GAG TC GCT GCC CAA CCA GCC ATG GCC CTC</td>
<td>0.7</td>
</tr>
</tbody>
</table>
The following primer pair was used to amplify the mouse $V_\lambda$ genes:

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_\lambda$ 5' sense primer</td>
<td>MSCVL-1 GGG CCC AGG CGG CCG AGC TCG ATG CTG TTG TGA CTC AGG AAT C</td>
</tr>
<tr>
<td>$V_\lambda$ 3' antisense primer</td>
<td>MHybLJ-B AGA TGG TGC AGC CAC AGT TCG ACC TAG GAC AGT CAG TTT GG</td>
</tr>
</tbody>
</table>

The following primer pair was used to amplify the human $C_k$ genes from pComb3XTT:

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_k$ sense primer</td>
<td>HKC-F CGA ACT GTG GCT GCA CCA TCT GTC</td>
</tr>
<tr>
<td>$C_k$ antisense primer</td>
<td>Lead-B GGC CAT GGC TGG TTG GGC AGC</td>
</tr>
</tbody>
</table>
The following primer pair was used to amplify the human $C_H1$ genes pComb3XTT:

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_H1$ sense primer</td>
<td>HlgGCH1-F GCC TCC ACC AAG GGC CCA TCG GTC</td>
</tr>
<tr>
<td>$C_H1$ antisense primer</td>
<td>dpseq AGA AGC GTA GTC CGG AAC GTC</td>
</tr>
</tbody>
</table>

Chimeric mouse/human light chains were generated by overlap PCR using the following primers:

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sense primer</td>
<td>RSC-F GAG GAG GAG GAG GAG GAG GCG GGG CCC AGG CGG CCG AGC TC</td>
</tr>
<tr>
<td>Antisense primer</td>
<td>Lead-B GGC CAT GGC TGG TTG GGC AGC</td>
</tr>
</tbody>
</table>

Chimeric mouse/human heavy chains were generated by overlap extension PCR using the following primers:

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sense primer</td>
<td>leadVH GCT GCC CAA CCA GCC ATG GCC</td>
</tr>
<tr>
<td>Antisense primer</td>
<td>dpseq AGA AGC GTA GTC CGG AAC GTC</td>
</tr>
</tbody>
</table>
Chimeric mouse/human Fab were generated by overlap PCR using the following primers:

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sense primer</td>
<td>RSC-F GAG GAG GAG GAG GAG GAG GCG GGG CCC AGG CGG CCG AGC TC</td>
</tr>
<tr>
<td>Antisense primer</td>
<td>dp-EX GAG GAG GAG GAG GAG AGA AGC GTA GTC CGG AAC GTC</td>
</tr>
</tbody>
</table>

The Fab clones were Sanger sequenced with the following primers:

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequencing of the variable region of the light chain</td>
<td>ompseq sense primer AAGACAGCTATCGCGATTGCAG</td>
</tr>
<tr>
<td>Sequencing of the variable region of the heavy chain</td>
<td>pelseq sense primer ACCTATTGCCTACGGCAGGCG</td>
</tr>
</tbody>
</table>

8 Appendix B: Papers published during this PhD

Serological Evidence of Immune Priming by Group A Streptococci in Patients with Acute Rheumatic Fever

Jeremy M. Raynes1,2, Hannah R. C. Frost3, Deborah A. Williamson2,4, Paul G. Young1,2, Edward G. Baker1,2, John D. Steemson1, Jacelyn M. Loh5, Thomas Prifti2,5, P. R. Dunbar2,5, Polly E. Atatoa Carr5, Anita Bell6 and Nicole J. Moreland1,2,4

1 School of Biological Sciences, University of Auckland, Auckland, New Zealand, 2 Maurice Wilkins Centre for Molecular Biodiscovery, University of Auckland, Auckland, New Zealand, 3 Institute of Environmental Science and Research, Wallingford, New Zealand, 4 The Peter Doherty Institute, University of Melbourne, Melbourne, Australia, 5 School of Medical Sciences, University of Auckland, Auckland, New Zealand, 6 Waikato District Health Board, Hamilton, New Zealand

Acute rheumatic fever (ARF) is an autoimmune response to Group A Streptococcus (GAS) infection. Repeated GAS exposures are proposed to ‘prime’ the immune system for autoimmunity. This notion of immune-priming by multiple GAS infections was first postulated in the 1960s, but direct experimental evidence to support the hypothesis has been lacking. Here, we present novel methodology, based on antibody responses to GAS T-antigens, that enables previous GAS exposures to be mapped in patient sera. T-antigens are surface expressed, type specific antigens and GAS strains fall into 18 major clades or T-types. A panel of recombinant T-antigens was generated and immunoassays were performed in parallel with serum depletion experiments allowing type-specific T-antigen antibodies to be distinguished from cross-reactive antibodies. At least two distinct GAS exposures were detected in each of the ARF sera tested. Furthermore, no two sera had the same T-antigen reactivity profile suggesting that each patient was exposed to a unique series of GAS T-types prior to developing ARF. The methods have provided much-needed experimental evidence to substantiate the immune-priming hypothesis, and will facilitate further serological profiling studies that explore the multifaceted interactions between GAS and the host.

Keywords: acute rheumatic fever, T-antigen, group A Streptococcus, tee-type, immune priming, Immunohistochemistry

INTRODUCTION

Acute rheumatic fever (ARF) is an autoimmune condition that can develop after a Group A Streptococcus (GAS) infection. ARF is now rare in high-income countries, but is associated with significant disease burden in low-income countries and some indigenous populations of high-income countries (Carapetis et al., 2016). The rates of ARF in Maori and Pacific children in New Zealand and Aboriginal children in Australia are amongst the highest in the world (Jain et al., 2008; Maguire et al., 2012). The peak incidence for ARF occurs in the 5–14 years old age band, with a mean peak in 9–12 year olds observed in a recent study (Jain et al., 2008). Episodes in children younger than 5 years of age are extremely rare. It has been postulated that repeated infections with GAS are needed to ‘prime’ the immune system before the first episode of ARF occurs (Carapetis et al., 2005, 2016). This may partly explain the lack of disease in pre-school children. Superficial
MINIREVIEW

Protein adhesins as vaccine antigens for Group A Streptococcus

J.M. Raynes1,2,*, P.G. Young2,3, T. Proft1,2, D.A. Williamson4, E.N. Baker2,3 and N.J. Moreland1,2,*

1School of Medical Sciences, The University of Auckland, 85 Park Road, Auckland 1023, New Zealand, 2Maurice Wilkins Centre for Molecular Biodiscovery, The University of Auckland, 3A Symonds Street, Auckland 1010, New Zealand, 3School of Biological Sciences, University of Auckland, 5 Symonds Street, Auckland 1010, New Zealand and 4Microbiological Diagnostic Unit Public Health Laboratory, Doherty Institute for Infection and Immunity, University of Melbourne, Melbourne, Victoria 3000, Australia

*Corresponding author: Department of Molecular Medicine and Pathology, University of Auckland, 85 Park Road, Auckland 1023, New Zealand. Tel: +64 9-923-5394; E-mail: n.moreland@auckland.ac.nz

One sentence summary: An overview of the current state of the Group A Streptococcus vaccine field focusing on the development of adhesin vaccine candidates.

Editor: Jacques Schrenzel

J.M. Raynes, http://orcid.org/0000-0002-2564-7893

ABSTRACT

Group A Streptococcus (GAS) is a globally important human pathogen that causes a broad spectrum of disease ranging from mild superficial infections to severe invasive diseases with high morbidity and mortality. Currently, there is no vaccine available for human use. GAS produces a vast array of virulence factors including multiple adhesin molecules. These mediate binding of the bacteria to host tissues and are essential in the initial phases of infection. Prophylactic vaccination with adhesins is a promising vaccine strategy and many GAS adhesins are currently in development as vaccine candidates. The most advanced candidates, having entered clinical trials, are based on the M protein, while components of the pilus and a number of fibronectin-binding proteins are in pre-clinical development. Adhesin-based vaccines aim to induce protective immunity via two main mechanisms: neutralisation where adhesin-specific antibodies block the ability of the adhesin to bind to host tissue and opsonisation in which adhesin-specific antibodies tag the GAS bacteria for phagocytosis. This review summarises our current knowledge of GAS adhesins and their structural features in the context of vaccine development.

Keywords: Group A Streptococcus; vaccine; adhesin; M protein; T antigen; fibronectin-binding protein

INTRODUCTION

Attachment and adhesion to host tissues are essential steps in microbial pathogenesis. These processes are mediated by virulence factors called adhesins that are, in general, surface-bound protein or polysaccharide molecules which confer tissue-specific binding (Pizarro-Cerdá and Cossart 2006). Bacterial attachment is initially weak, driven by non-specific interactions between bacteria and host tissues. This is followed by higher affinity bacteria-host interactions mediated by adhesins binding to their specific targets. These can be proteins, sugars or lipids exposed on host cell surfaces as well as moieties on components of the extracellular matrix (ECM) such as fibronectin, collagens and laminin (Krishnan and Narayana 2011). Adhesins are
9 References


to Protection. *The Journal of Infectious Diseases*, 187(10), 1598-1608. 10.1086/374800


Duncan, A. R., & Winter, G. (1988). The binding site for C1q on IgG. Nature, 332, 738. 10.1038/332738a0


240


