A Rapid Approach for Linear Epitope Profiling of a Streptococcus A Vaccine Candidate.

A thesis
submitted in partial fulfilment
of the requirements for the degree
of
Master of Science (Research) in Molecular and Cellular Biology
at
The University of Waikato
by
Kirsten Jane Browne-Cole

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2025

Abstract

The creation of effective vaccines is a pinnacle achievement of humankind, with the maintenance and expansion of the current repertoire being crucial in supporting human health and well-being globally. Nevertheless, identification of the mechanisms responsible for protection from infection and disease remains difficult, hindering vaccine candidate progression. These factors illustrate the importance of developing a wide range of techniques available to identify and investigate potential antigens. Current tools for assessing antigen-antibody interactions, although hugely valuable, possess limitations in accessibility and throughput which this research aimed to address.

To implement an accessible rapid protocol that profiles the antibody response to linear epitopes after immunisation, we utilised phage display in conjunction with Nanopore Sequencing. A Group A Streptococcus vaccine candidate, TeeVax3, was selected as a model to explore the binding antibody population produced from immunisation and evaluate the immunogenicity of the antigen. Production of a phage library expressing overlapping TeeVax3 peptides was panned against TeeVax3 specific polyclonal antibodies from rabbits, revealing an unexpected bias for the N terminal epitope tag. Further investigation confirmed the immunogenicity of this region for both linear and conformational antibodies, with an epitope of 10-residues being identified via evaluation of binding with synthetic peptides. Although removal of purification tags is inconsistent, particularly in early stages of vaccine development, these findings illustrate the importance of doing so to ensure relevant immune responses are observed. Other potential areas of immunogenicity on the antigen were able to be observed by excluding the N-terminal region dominating the response, verifying the methods potential for evaluating vaccine candidates. The protocol presented here outlines a valuable addition to the toolkit for development of vaccine candidates, which could be easily adapted to a wide range of antigens and correlating antibody responses.

Acknowledgements

It has taken a village to complete this thesis, but first I would like to express my deepest gratitude to my chief supervisor, Dr William Kelton. I am sincerely grateful for the opportunity to work on this project; your unwavering support and positivity have provided me much needed motivation and confidence in challenging times. I am incredibly thankful for your leadership and drive which has led to the fantastic achievement of publishing this project, and I look forward to continuing to work together in the next step of my career pathway.

A special thank you to Kyrin Hanning, my co first author, for your impeccable technique, desire for elegant solutions and endless patience. My progress would have been greatly hampered without your early morning and evening assistance. To Kevin and Meghan for your efforts helping me achieve publication worthy data and listening to all my life dramas.
I would also like to extend thanks to Dr Jacelyn Loh from Auckland University for always happily accommodating my requests for information and reagents.

To the C2 lab backbone, Dr Judith Burrows and Dr Geetanjali Rai, thank you for being a kind ear and a font of knowledge. I always walk away from our conversations feeling uplifted. To the many members of C2 that I have had the pleasure to work with and get to know, I greatly appreciate the fantastic work environment you all help to create. A special thanks to Marina, my travel buddy and sounding board, I have thoroughly enjoyed our time working together and appreciated having someone to figure out all the administration tasks with.

I greatly appreciate my family for their constant support, particularly to Nana and Grandma for all the childcare that made this work possible. As well as my children, Alex and Holly, for reminding me to keep some perspective and enjoy the small things. Finally, to my husband Myles, thank you for always celebrating my successes, commiserating the challenges, and your endless encouragement to dream big.

Acknowledgement of Funds: This research was supported by The Maurice Wilkins Centre (MWC project grant #4026).

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Preface

The development of vaccines has led human society to successfully protect the global population from substantial impacts of many infectious diseases. This thesis outlines the design and implementation of a high throughput phage display technique used to map linear epitope antibody responses as important determinants of vaccine efficacy. The method is applied to a Group A Streptococcus vaccine candidate being developed by the University of Auckland. The research has been accepted for publication as a peer-reviewed Scientific Reports Article, with this thesis being written independently, following the submission.

Kirsten Browne-Cole, Kyrin R. Hanning, Kevin Beijerling, Meghan Rousseau, Jacelyn Loh, William Kelton. (in press). A rapid approach for linear epitope vaccine profiling reveals unexpected epitope tag immunogenicity. *Scientific reports.*

1.1.1 Author contributions

As a co-first author, I led all the experimental work for this research (including antigen production, antibody purification, library creation, phage panning, sequencing and enzyme-linked immunosorbent assays), conducted data analysis, and contributed to the manuscript drafting and revision process. Co-first author Kyrin Hanning contributed to the experimental design, completed the bioinformatics, and was involved in writing the manuscript from draft to submission. Kevin Beijerling assisted with enzyme-linked immunosorbent assays and reviewed the draft manuscript. Meghan Rosseau assisted with purification of the TeeVax3 antigen both before and after epitope tag removal. Dr Jacelyn Loh contributed to the conceptualisation of the research, while also providing the TeeVax plasmid and both serum from unimmunised and TeeVax immunised rabbits. Dr William Kelton conceptualised and supervised the research, assisted with experiment work, gave recommendations for data interpretation and led the manuscript creation process from draft through to submission and review. All

contributions to generating experimental data have been further described in the figure captions.

1.2 Drivers of Vaccine development

Vaccines are one of the most impactful scientific developments made for the benefit of humankind. They are used to prepare the body for exposure to pathogens, ideally to prevent infection altogether, or to limit the infectiousness and severity of disease if it still occurs (Cunningham et al., 2016). This results in a lower disease burden on society, by minimising deaths, hospitalisations and economic cost. The success of our current repertoire of vaccines is illustrated by the eradication of smallpox infection, alongside the World Health organisation's estimation that immunisation prevents 2-3 million deaths annually (Gebre et al., 2021). These outcomes illustrate how vaccines contribute to billions of dollars saved globally each year (C. M. C. Rodrigues & Plotkin, 2020).

The importance of having a robust and timely vaccination development pipeline became apparent during the COVID-19 global pandemic, where both the economic and societal impacts of infection control without a vaccine available had substantial repercussions. A major public health strategy during this time was to close nonessential businesses and limit unnecessary contact between individuals. The economic costs of these closures are estimated to be in the trillions (Kolahchi et al., 2021) with continued concerns about the societal impacts, particularly in the areas of child development and education (Abufhele et al., 2024; Mazrekaj & De Witte, 2024). The risk of future global pandemics caused by zoonotic events is increasing due to habitat destruction, climate change and human development bringing wildlife in close proximity to human populations (L. Zhang et al., 2022). The potential to avoid or minimise these scenarios, in addition to the achievements of the current vaccine collection, makes the arduous process of vaccine development well worth the investment for the global community.

A particularly challenging pathogen to create a vaccination for is *Streptococcus pyogenes,* often referred to as Group A Streptococcus (GAS). While the diseases caused by this bacterium are currently treatable with antibiotics, they continue to pose a significant health challenge, particularly for Māori and Pacific populations

in Aotearoa, New Zealand (Ammar et al., 2024). The focus of this thesis is to assist in the development of TeeVax, a Group A Streptococcus vaccine, by creating a phage display method which profiles the linear epitope portion of the immune response resulting from immunisation. This work aims to establish a rapid method for evaluating vaccine candidates, offering an accessible tool to enhance the vaccine development pipeline and, in turn, improve health outcomes for groups disproportionately affected by Group A Streptococcus.

1.3 Group A Streptococcus

TeeVax is a recombinant protein vaccine designed to target the pilus of GAS, a Gram-positive bacterium that exclusively infects humans. Historically, GAS was classified based on serotype-specific antiserum recognizing the T or M proteins on its cell surface. Currently the primary typing strategy is to sequence the *emm* gene, which codes for the highly variable M protein allowing the bacteria to be classified into one of over 250 serotypes. In the past the T protein, which comprises the main component of the pilus structure, has also been used to complement the above strategy. However, the accessibility of full genome sequencing which allows comprehensive understanding of strain development and identification of potential vaccine targets is superseding both of these tools as the preferred option (Ajay Castro & Dorfmueller, 2023; Loh et al., 2021; Smeesters et al., 2024).

GAS causes a wide range of diseases, from the reasonably benign pharyngitis and impetigo to the more severe rheumatic fever and necrotising fasciitis. There is a strong link between M type and the manifestation of disease, with superficial pharyngitis being associated with M types 1, 3, 5, 6, 12, 14, 17, 19, and 24, while M types 1, 3, 5, 6, 11, 12, 14, 17, 18, 19, 24, 27, 29, 30, 32, and 41 have been connected to acute rheumatic fever (Walker et al., 2014). However, it is interesting to note the most prevalent type for each disease frequently changes from year to year as well as geographically.

Although easily treated in most western nations using antibiotics, GAS remains responsible for over 500,000 deaths worldwide per year and persists as one of the World Health Organisation's diseases of global relevance (Fan et al., 2024). The burden of GAS infections is disproportionately experienced by middle- and low-

income countries, along with the indigenous populations of some more wealthy nations including Australia and Aotearoa, New Zealand. The source of this discrepancy is often contributed to socioeconomic factors such as crowded housing and inadequate access to medical care, although how much these factors influence the increased strain diversity observed in disadvantaged populations is yet to be confirmed (Smeesters et al., 2024).

A substantial increase in paediatric cases of GAS and scarlet fever have been recorded globally since the Covid-19 pandemic. New Zealand's data mirrored this trend with cases increasing by 120% after lifting of Covid-19 restrictions. Māori and Pacific populations continue to be affected at higher rates than other ethnic groups (Ammar et al., 2024) emphasising the persistent need for an effective Group A Streptococcus vaccine.

1.3.1 Difficulties in creating a Strep A vaccine

Despite concerted effort from the scientific and medical communities there is still no licensed vaccine for Group A Streptococcus. One factor contributing to this is the considerable strain diversity, which appears to be achieved through homologous recombination and driven by competition between lineages as well as selective forces from host immune responses (Davies et al., 2019). Another substantial obstacle is the likelihood of autoimmune events occurring, particularly in M protein vaccine candidates. This response was initially reported by Massell et al during a clinical trial in 1969, where three healthy participants developed rheumatic fever like symptoms after being immunised with a type 3 M protein vaccine (Massell, 1969). Although the study was not well controlled, the concern was supported by Dale and Beachey with their identification of protein cross reactivity with cardiac myosin (Dale & Beachey, 1985) causing the US Food and Drug Administration to place a ban on GAS organisms and their derivatives in vaccines which stood until 2006 (Walkinshaw et al., 2023).

With the lifting of the FDA ban and rejuvenation of research around vaccines for Group A Streptococcus*,* the range of potential candidates targeting a variety of antigens (Figure 1.1) has improved the outlook for an effective solution being achieved. However, when compared to other infectious diseases with a high global

impact, the development of GAS vaccines remains relatively neglected. This gap was further highlighted by the World Health Organisation releasing a research and development technology roadmap in 2019, confirming GAS as a priority worldwide (Vekemans et al., 2019).

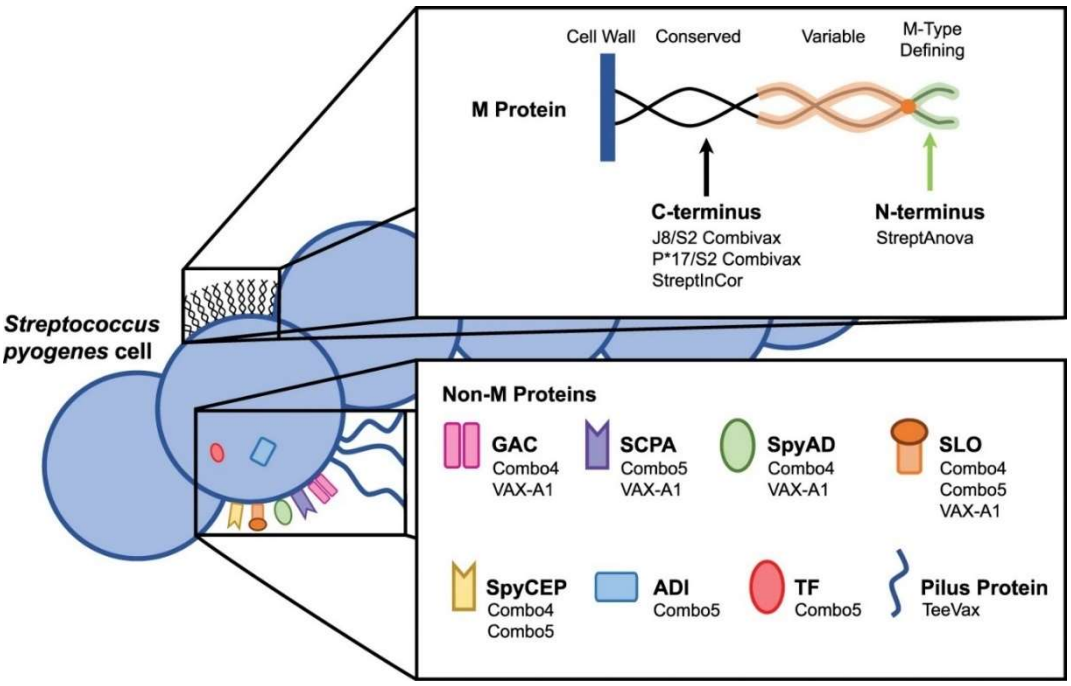


Figure 1.1. Schematic of vaccine candidates in development and their corresponding antigens from Walkinshaw et al, 2023.

1.4 The development of TeeVax candidate vaccines.

The ongoing challenge of Group A Streptococcus disease in Aotearoa has led to a local interest in the development of an effective vaccine solution. This focus has produced the TeeVax series of vaccine candidates. These protein subunit vaccines target the pilus structure on the surface of Group A Streptococcus. Pili are hairlike structures around 1 μm in length and are known virulence factors which play a role in adhesion, colonisation, and avoidance of the host immune system. Targeting the pilus offers the advantage of having less variability than the historically favoured M protein target, with 18 of the potential T antigens covering over 95% of circulating GAS strains. These projections are made up of anywhere between 10 and 100 T antigens polymerised together with adaptor sequences at both ends.

(Figure 1.2). The majority of these antigens are two domain structures with a select few containing three or even four domains (Loh et al., 2021).

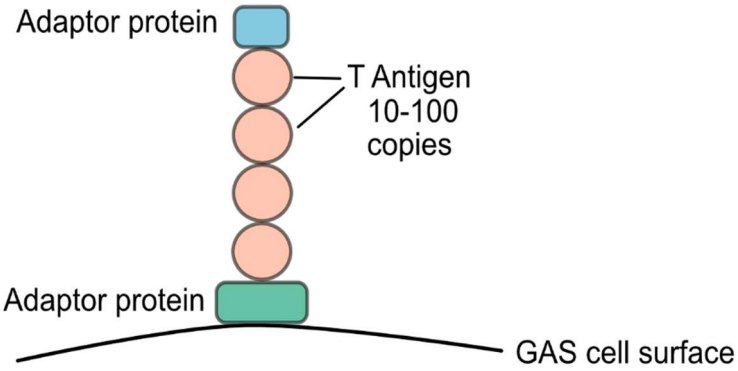


Figure 1.2. Illustration of pilus structure made up of multiple T antigens flanked by Adaptors. Figure made using Affinity Designer 2 https://affinity.serif.com/en-us/designer/

TeeVax antigen development involved the creation of three fusion proteins named TeeVax1, 2 and 3, with each antigen containing 6, 7, and 5 concatenated domains respectively. The antigen we investigated, TeeVax3, was composed of the middle domain from four three-domain T antigens and the N-terminal domain from the only known four-domain T antigen (Figure 1.3). These sections were linked together with a GSGSG spacer sequence and cloned into an *E.coli* expression vector for production.

To determine the antibody response to TeeVax antigens, three strategies were employed: enzyme-linked immunosorbent assay (ELISA) using serum from immunised rabbits, opsonophagocytic killing assays (OPKA), and a humanised plasminogen transgenic mouse model. Initially all three approaches were completed with TeeVax1 to investigate the validity of the approach. ELISA analysis of TeeVax1 immunised rabbit sera showed IgG responses to the T antigens included in the TeeVax 1 structure, with some cross reactivity to similar T antigens. OPKA assays against GAS SF370 identified vaccination with the T1 antigen resulted in much higher killing rates than both M1 protein and TeeVax1. Finally, the humanised plasminogen transgenic mouse model was used to determine vaccine efficacy. Before the immunised mice were challenged with an invasive disease GAS strain, ELISA analysis showed a strong antibody response to the antigens used

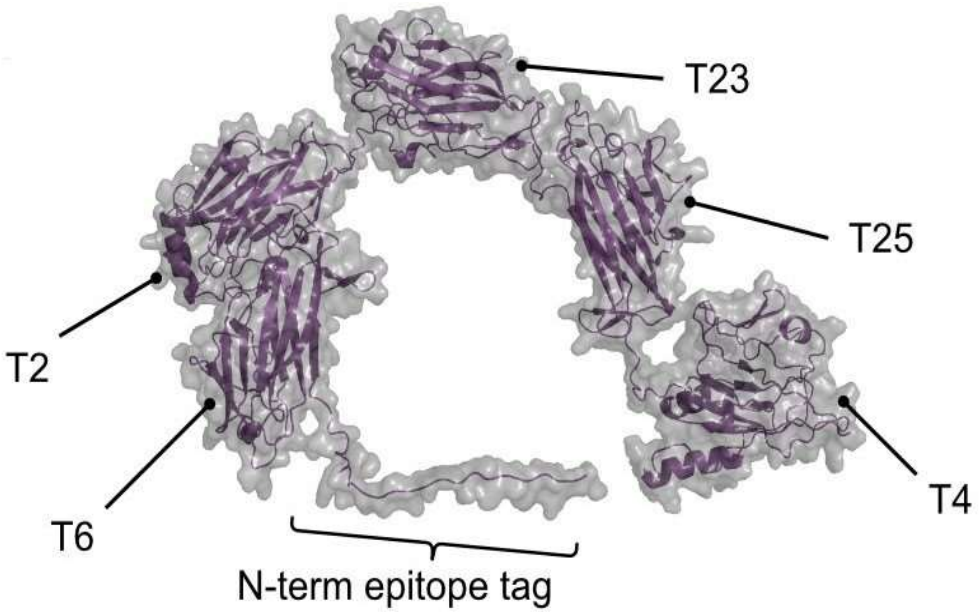


Figure 1.3. Model of TeeVax3 antigen showing the domains it is comprised of and the N terminal epitope tag containing repeated Histidine residues and TEV cleavage site. Structure from Browne-Cole et al.

for immunisation when compared to phosphate buffered saline immunised control mice. TeeVax1 immunised mice had higher IgG titres for the six T-antigens comprising the recombinant protein, with no significant difference in response when comparing TeeVax1 immunisation to full length T1 protein. The survival rates for mice challenged with GAS were highest in the M1 protein immunised group, followed by T1 and then TeeVax1 with 33% survival after 10 days. These results provided sufficient evidence of protection conferred by TeeVax1 to prompt construction of two more recombinant proteins, TeeVax2 and TeeVax3. Subsequent experiments involved rabbits being immunised with the vaccine candidates either individually or as a mixture. TeeVax3 antiserum showed reactivity to the constituent antigens when exposed to 21 full length recombinant T antigens during ELISA. However, this did not translate to binding with two domain T antigens present in TeeVax1 and TeeVax2. Moreover, these results do not provide detailed information on the exact nature or location of the epitopes that are being recognised.

Given this lack of molecular-level epitope data, we designed and implemented a protocol to map the antibody response from immunisation to the recombinant antigen structure. Our approach sought to provide further information about regions

of immunodominance and identify any biases in binding on the antigen surface. Such data provides guidance to the continued development of a broadly protective vaccine against GAS, while also being easily applicable to other vaccine candidates.

1.5 Correlates of vaccine-induced protection

To determine a vaccine candidates' ability to protect against the specific disease being targeted, evidence of elicited immune protection is required. Correlates of protection are immune markers which indicate the vaccine will provide immunity when exposure to a pathogen occurs. One of the most common markers used by the current vaccines is to measure the number of neutralising antibodies produced, which prevent extracellular pathogens from interacting with host cells and establishing infection. This mechanism is used for the current vaccines for Human papilloma virus and Small pox (Z. Chen et al., 2022; Plotkin, 2020). However, it is important to consider other mechanisms, particularly for bacterial pathogens such as *Streptococcus pyogenes*, where generating neutralising antibodies does not appear to be a predominant factor in establishing protection. The T cell produced interferon response, is one such example of this, which provides protection from tuberculosis in the Bacillus Calmette–Guérin vaccine. (Z. Chen et al., 2022). Even with the ability to measure a range of potential correlates of protection for any given vaccine candidate, challenges can still arise when there is no clear mechanism which is the most significant. For instance, malaria vaccine candidates trigger a range of responses from antibody production, natural killer cell responses and secretion of interferon from T cells, none of which can be identified as the primary method of protection (Plotkin, 2020).

Another challenge when determining vaccine efficacy is the difficulty of measuring antibody-antigen interactions in high throughput. Only a small fraction of an individual's B cell population will be specific to the antigen in question, requiring substantial effort to be applied to purify the relevant sample and the potential to lose antibodies involved in lower affinity or temporary interactions. These barriers, in combination with the resource intensive nature of many of the high throughput protocols, could result in the investigation strategies applied being less than optimal. (De Leon et al., 2024; D. Hu & Irving, 2023).

1.5.1 The antibody response

Antibodies are a core strategy in the adaptive immune systems response to external threats. These 'Y' shaped proteins are produced in a type of B cell called a Plasma cell, after which they are located free in sera or bound to B cells as receptors. The primary purpose of these proteins is to either destroy or tag a pathogen for destruction, which they achieve through a range of mechanisms including neutralisation, antibody-dependent cellular cytotoxicity, phagocytosis and complement mediated lysis. As mentioned earlier, neutralising antibodies are frequently used as the primary indicator of vaccine efficacy due to their ability to prevent infection through a variety of strategies. Pre-attachment neutralisation occurs when antibodies bind to pathogens before interaction with host cells has occurred, this results in immune complexes that effectively trap the pathogen preventing infection and marking it for destruction. Another neutralisation process is to block the pathogens' ability to attach to the host cells. This can be achieved directly with antibodies binding the surface ligands used to interact with other cells or indirectly resulting in a structural change or barrier that prevents binding from occurring. If attachment is achieved there are antibodies with the function of inhibiting the pathogens ability to fuse with or enter the cells, as well as those that can act intracellularly to disrupt the invading microbes lifecycle (Forthal, 2014; Kapingidza, A. B., Kowal, K., & Chruszcz, M, 2020). Phagocytosis and complement mediated lysis are both integral processes for the protection against bacterial invasion, with GAS possessing a multitude of adaptations to resist attack.(Walker et al., 2014) The role of antibodies in these mechanisms is to bind the microbe and activate the complement cascade. This can initiate a range of responses including sequestering the pathogen and recruitment of phagocytotic cells to implement destruction (Forthal, 2014).

Although these pathways all differ in some respects, identification and interaction with the pathogenic cell is always required. The specific sequence of amino acid residues recognised by an antibody is called an epitope. If the sequence is made up of consecutive amino acids the epitope is linear, whereas conformational epitopes result from protein folding bringing the amino acids into close proximity to form the sequence recognised by the antibody (Figure 1.4).

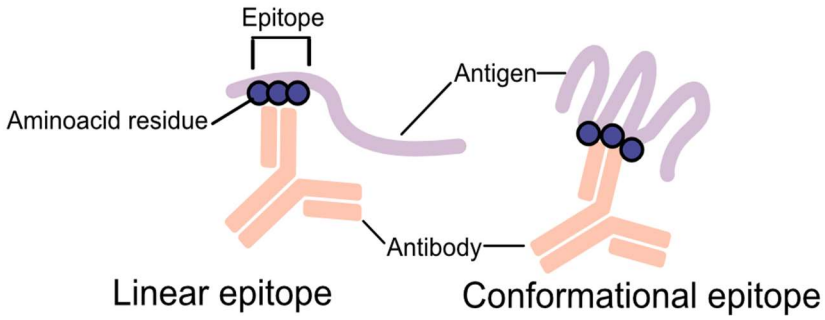


Figure 1.4. Figure 1.4. Schematic showing how linear and conformational epitopes are recognised by antibodies. Figure made using Affinity Designer 2 https://affinity.serif.com/en-us/designer/

For the immune response to be as effective as possible, both linear and conformational epitopes need to be targeted. The ability to identify the pathogen in its native form is determined by the binding of conformational epitopes, however, changes in this structure can easily occur due to the host environment, interactions between pathogen and host immune system or mutations that occur in the surface proteins over time. The recognition of both linear and conformational epitopes increases the likelihood of a potential pathogen being detected by the host and preventing infection from occurring (Forsström et al., 2015; Francino-Urdaniz & Whitehead, 2021).

1.5.2 B cell generation for linear and conformational epitopes

The mechanisms used to generate antibodies against linear and conformational epitopes differ due to the requirement to display the antigen processed for linear epitopes and intact for conformational ones. The antigen presenting cells which result in antibodies that recognise linear epitopes include dendritic cells, B cells, and macrophages. These cells internalise the antigen and dismantle it into segments which are then presented on major histocompatibility complex class II to activate helper T cells. These helper T cells then interact with B cells to trigger the process of plasma cell formation and antibody production. On the other hand, generating antibodies against conformational epitopes requires the antigen to remain complete and unmodified. Both B cells and follicular dendritic cells are capable of binding native antigen, with follicular dendritic cells displaying the native antigen bound

with antibody or complement proteins to ensure continued B cell maturation and antibody production (Abd El-Aleem et al., 2022; Adler et al., 2017; Rastogi et al., 2022).

1.6 Methods for high-throughput vaccine epitope profiling

The profiling of both linear and conformational B cell epitopes produced in response to immunisation can be used to identify immunogenic sites that have potential to be effective vaccines or to investigate the immune response to increase the effectiveness of a vaccine candidate. Currently there are a range of techniques available to investigate the epitope profile of an immune response (Table 1.1). Historically, favoured methods for mapping epitopes have involved X ray Crystallography and more recently Cryo-electron microscopy. While the accuracy of these methods is a substantial advantage, the inability to capture temporary or weaker interactions and the difficulty in analysing a polyclonal antibody repertoire highlight the need for an alternative tool. Moreover, the considerable resources required makes these techniques less desirable to implement for vaccine candidates in early development stages (Antanasijevic et al., 2022; De Leon et al., 2024; Hamed et al., 2023; D. Hu & Irving, 2023).

In order to address some of these hurdles, an array of Mass spectrometry-based approaches have been established. With the benefits of small sample requirements, less time-consuming preparation and more instrument availability, alongside the capacity to resolve epitopes at an amino acid level, these methods provide alternatives for continued progress of the vaccine development pipeline. However, the challenges of identifying interactions in less stable antigen-antibody complexes and the requirement to validate the results produced by these protocols still remain (Jethva & Gross, 2023).

Another approach is to use protein display systems which can be broken into broad categories based on the type of surface used for display. Peptide microarrays involve the immobilisation of synthesised peptide segments to a membrane, plate or slide (Wang et al., 2020). Whereas, cell-based techniques allow antigenic peptides to be expressed on living cells such as bacteria, yeast or mammal, as well as the firmly established phage display technique. These systems require

cloning of the peptide library into expression vectors for display on the surface of the cell or phage (Mahdavi et al., 2022; Robertson et al., 2020). All peptide display approaches involve exposure of the immobilised peptides to the antibody pool, with epitopes that are bound being identified using a range of techniques including colorimetric secondary antibodies and flow cytometry. The substantial advantage of these techniques is the ability to create epitope maps in high throughput, with a flexibility and speed not achievable in structural based methods. The possibility of identifying more transient interactions is potentially obtainable with peptide display systems, however length of exposure and washing protocols need to be carefully considered when trying to achieve this. (D. Hu & Irving, 2023). Moreover, both cell-based and phage display methods require DNA sequencing to link the bound peptide with the antigen reference library, which requires careful contemplation of error rates and creates a barrier of accessibility for some projects (Mohan et al., 2018).

Table 1.1. Summary of Epitope mapping techniques.

Technique	Linear	Conformational	Resolution	High through-put	Low Avidity detection	Time	Cost	References
Xray crystallography	Yes	Yes	High	No	No	High	High	De Leon et al Hu & Irving Hamad et al
Cryo-electron microscopy	Yes	Yes	High	No	No	High	High	De Leon et al Hu & Irving Hamad et al
Nuclear Magnetic resonance	Yes	Yes	High	No	Yes	High	High	De Leon et al Hu & Irving Hamad et al
Mass spectrometry	Yes	Yes	High	Yes	Yes	Mid	Mid	Hu & Irving Hamad et al
Peptide microarray	Yes	No	Low	Yes	Yes	Low	Low	De Leon et al Hu & Irving Hamad et al Antanasijevic et al
Bacteriophage display	Yes	Some	Low	Yes	Yes	Low	Low	De Leon et al Hu & Irving
Deep mutational scanning	Yes	Yes	High	Yes	Yes	Mid	Mid	Hu & Irving Hanning et al
Yeast display	Yes	Some	Low	Yes	Yes	Low	Low	Hu & Irving

1.6.1 Phage display systems for epitope mapping

Bacteriophages are viruses that specifically infect prokaryotic cells and are the most abundant biological entities on Earth. Like other viruses, they replicate by hijacking the molecular machinery of their bacterial hosts to produce new phage particles (Ackermann, 2003). Traditionally, the phage lifestyle has been classified as either lytic, which results in destruction of the host cell to release phage progeny, or lysogenic where the bacterial cell remains intact and functional while producing phage (Mäntynen et al., 2021). The lysogenic mechanism is utilised in the phage display technique applied in this work, with M13 bacteriophage being engineered to express peptides on their surface to allow investigation of protein- protein interactions.

M13 bacteriophage are a filamentous phage with a long, thin structure that infect a range of Gram-negative bacteria via the F pilus (Figure 1.5). The genome consists of an extended single-stranded DNA, containing non-essential regions that can be readily modified to incorporate exogenous sequences for expression (Deng et al., 2017). In phage display protocols, the coding sequence for the display peptide is frequently inserted into the pIII coat protein (Figure 1.5A), which results in expression on the surface of one end of the phage. However, peptides of eight or less amino acids can be attached to the pVIII protein resulting in larger quantities of displayed peptide (Figure 1.5B) (Fadaie et al., 2023).

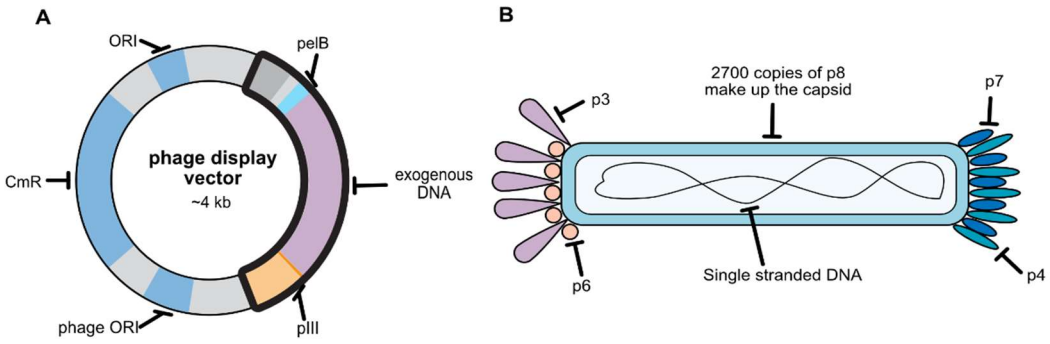


Figure 1.5. Phage display components. (A) Phage display vector with labelled origin of replication (ORI), Chloramphenicol resistance (CmR), and DNA insert between pelB sequence to ensure transport to the periplasm and p3 protein gene for insert display. (B) M13 bacteriophage structure identifying the five capsid proteins and single stranded DNA genome.

A phage display library is a collection of bacteriophages, each displaying a unique peptide due to the inclusion of exogenous DNA into the genome. Libraries can be designed using randomly generated peptides, or by breaking a known protein sequence into small peptides (Jaroszewicz et al., 2022), as was the strategy used in this research .

Once a library is created and expressed in phage, bio panning can be completed. This is the process of collecting and identifying peptides that interact with the antibody population in question. The phage library is exposed to a target, either immobilised on a surface or in solution, with unbound phage being washed away and the bound fraction released from the target by altering pH or cleaving with an

enzyme. This positive fraction is then used to reinfect bacteria producing more binding phage, with the process being repeated for three to five rounds depending on the application. The final step of this technique is to sequence the population of phage produced, whether this is completed after each panning round or just at the final stage will depend on the aims of the project. The resultant data identifies which peptides were the most successful binders to the target in question and provides valuable information about the protein-protein interactions occurring (Jaroszewicz et al., 2022; Kong et al., 2020).

1.6.1.1 Linear epitope mapping by phage display

The way phage display can be utilised for investigating the binding affinity and distribution of linear epitopes has been illustrated by various groups in recent years. One such example is the work of Xu et al., which used the VirScan platform to profile over 1 x 10^8 binding interactions between human antibodies and viral peptides. This protocol involved synthesising a bacteriophage library which displayed peptides 56 amino acids in length with a 28-residue overlap. These tiles were identified from reference sequences of all human viruses and screened against serum samples from 569 humans worldwide. Once the phage library was screened with the serum, unbound phages were washed away and the bound fraction collected and amplified for sequencing with Illumina. Their results identified a few specific peptides that were recognised by 95% of the sampled population, showing the conserved nature of epitopes and identifying potential delivery peptides for vaccine development (Xu et al., 2015).

Another related technique is phage deep mutational scanning, which allows the influence of single amino acid residues within an epitope to be investigated in high throughput. Garrett et al utilised this method to explore known epitopes of HIV antibodies, confirming potential residues implicit in viral escape and identifying unique epitopes. The primary distinction between this technique and that used in VirScan experiments is the creation of a phage library which includes modifying the amino acid at each position to encompass all 20 possibilities while the other amino acid positions remain unchanged (Garrett et al., 2020). Both examples from Xu et al and Garrett et al utilised Illumina for their sequencing needs. Although highly accurate which is particularly important for deep mutational scanning, this

method does have drawbacks including the possible biases introduced from PCR amplification to the slow turnaround and accessibility issues. We aimed to address these problems by utilising Oxford Nanopore technologies for our sequencing needs.

1.6.1.2 Conformational epitope mapping by phage display

A large proportion of any antibody response will consist of immunoglobulins that recognise conformational epitopes (Gershoni et al., 2007), however, the ability to identify these using phage display techniques alone is limited due to the short length of the peptides expressed resulting in a lack of native structure. Despite this, phage display can be used to identify areas of immunogenicity for further investigation in the hope of discovering conformational epitopes present on the native antigen. One technique available is using computational tools where the data from phage display experiments can be mapped against the antigen structure. EpiSearch is an example of such a tool, described by Negi and Braun, which automatically searches and predicts possible conformational epitopes in high throughput. This method has the flexibility of using both single or groups of peptide sequences, and is capable of rapid analysis. Although it appears to identify epitope locations correctly, finding consensus on sequences remains difficult (Negi & Braun, 2009).

Another strategy involves mutagenesis, as mentioned previously, this technique allows fine mapping of antibody binding sites by altering amino acid residues to observe the effects on interaction. Careful consideration needs to be given to experimental design for this technique, as altering multiple sites at once can make it difficult to identify crucial residues, and changes in affinity could result not just from altering the binding site itself, but also amino acids vital in protein folding structure (Rojas et al., 2014). Even with these tools, resource intensive structural techniques are generally required to verify conformational epitopes in native antigens due to the large number of factors involved in antigen folding and interactions between epitope and paratope (Negi & Braun, 2009).

Considering the above, there remains a substantial need for rapid identification of regions of immunogenicity on an antigen, which can then be further defined by high resolution techniques. The work of this thesis aims to provide an easily accessible

high throughput method to analyse linear epitopes using a synthetic oligo pool and Nanopore sequencing. These elements allow easy implementation, greater flexibility and improved resolution when compared to other phage display protocols, with the data produced identifying areas of immunodominance and providing guidance to assist conformational epitope discovery.

1.7 Project aims

The importance of continual progress in the vaccine development space is illustrated by the recent challenges from emerging pathogens and the global SARS-CoV2 pandemic. Great achievements have occurred in response to these events including the first widespread administration of mRNA vaccines to the population, further refinement of Cryo-EM technology, and advancement in machine learning for prediction tools (Antanasijevic et al., 2022; Szabó et al., 2022; Taft et al., 2022) Despite this, there remains the requirement to experimentally assess the ability of a vaccine candidate to offer protection in an easily accessible manner.

This project aims to use a high throughput approach to map the antibody response from TeeVax immunisation to the antigen structure. Establishing this technique will provide an attainable option for epitope mapping potential vaccine candidates and provide valuable information for the future development of TeeVax antigens.

Objective 1:

Establish a method for rapid profiling of linear epitopes in high throughput.

Objective 2:

Characterise the antibody response elicited from TeeVax antigens.

Chapter 2

Materials and Methods

2.1 Creation of phage display libraries

To establish the protocol for creating phage libraries displaying TeeVax3 peptides, we first confirmed our ability to display two control peptides successfully. These steps where then repeated with minor adjustments in order to produce the TeeVax3 library. The methods outlined in this chapter are taken directly from the published paper with some supplementary information for enhanced understanding.

2.1.1 Control peptide phage display

2.1.1.1 Cloning of control peptide sequences

To confirm suitability of the phage display system, we designed control oligo nucleotides flanked by 5' and 3' adapters containing a coding sequence, peptide transport signal and restriction sites (Appendix A1). These were added to the control oligonucleotides encoding the PEP1 cyclic peptide and 6-His epitopes (Appendix A2).

2.1.1.1.1 PCR amplification of peptides

The two control peptides were PCR amplified in separate reactions using HOT FIREPol® Blend Master Mix (Solis Biodyne), with primers (designed in house and ordered from Integrated DNA Technologies, Appendix A3) included at 0.2 μM in the reaction alongside 5 ng of control peptide template. Amplification was achieved under the conditions outlined in table 2.1.

Table 2.1. PCR conditions

Step	Temperature (°C)	Time	Number of rounds
Initital activation	95	15 min	1
Denaturation	95	10 sec	35
Annealing	55	30 sec	35
Elongation	72	6 sec	35
Final extension	72	5 min	1

2.1.1.1.2 Restriction digest of peptides and fragments

One hundred ng of the amplified control peptide sample oligod in rCutSmart™ Buffer (NEB) and ultrapure water were digested with 0.5 µL Sfil restriction enzyme (NEB) for 2 hrs at 50°C. Digested fragments were cleaned using a QIAquick nucleotide removal kit (Qiagen) and DNA concentration was quantified using a Nanodrop™ 2000 (Thermofisher).

2.1.1.1.3 Ligation of vector and peptides

Both peptides were then ligated into the digested pAK200 vector (sourced from Prof George Georgiou at the University of Texas) at a ratio of 1:5, with a final vector concentration of 2 nM. The 10 µL reactions were composed of 0.5 µl T4 ligase, 1 µL of 10x T4 DNA ligase reaction buffer (NEB) with approximately 31.7 ng and 30.3 ng of ligated vector for PEP-1 and 6-His respectively. Reactions were incubated at room temperature for 15 min, followed by 65 °C for 10 min in a T100 Thermal Cycler (Bio-Rad). The ligated vector was then bound to Monarch columns (NEB) after the addition of five volumes of 5 M Guanidine Hydrochloride containing 30% isopropanol. The sample was then washed with 10 mM Tris-HCl pH 7.5, containing 80% ethanol, and eluted with ultra-pure water.

2.1.1.1.4 Transformation into Escherichia coli SS320 cells

One µL of each ligation solution was added to 25 µl of electrocompetent SS320 cells and electroporated at 1.8 mV, with 1 mL of pre warmed Super optimal medium with catabolite repression (0.5 % (w/v) yeast extract, 2 % (w/v) tryptone, 10 mM

NaCl, 2.5 mM KCL, 20 mM $MgSO_4$, 20 mM glucose) added immediately after. These cells were then plated on Luria Bertani (LB) agar in the presence of 20 µg/mL Chloramphenicol and incubated overnight at 37°C.

2.1.1.2 Colony PCR and sequencing preparation

Eight colonies of the PEP1 cyclic peptide and 6-His epitope were amplified using the PCR protocol outlined in section 2.1.1.1.1. Primers designed to amplify inserts in the pAK200 vector by Kyrin Hanning (Appendix A4) were added at a concentration of 0.3 µM. The resulting reaction mixtures were run on a 3% agarose gel in Tris-borate-EDTA (TBE) buffer (0.13 M Tris, pH 7.6, 45 mM boric acid, 2.5 mM EDTA) at 130 V for 30 mins. Sanger sequencing of both controls was completed by the Massey Genome Service (Massey University, New Zealand) with samples being prepared by growing up two colonies of each control peptide at 37 °C overnight in LB media accompanied by 20 µg/mL chloramphenicol. DNA was extracted using a QIAprep® Spin Miniprep Kit (Qiagen) and eluted with ultra-pure water before being sent away, with resulting data pairwise aligned to expected sequences in Geneious Prime®.

2.1.1.3 Phage production

Control peptide glycerol stocks were grown at 37 °C from OD 0.1 to 0.4 - 0.6 in LB media supplemented with 20 µg/mL chloramphenicol. M13K07 helper phage were added at a multiplicity of infection (MOI) of 20x and the culture incubated without shaking for 1 hour to allow infection. Kanamycin antibiotic was added at a concentration of 35 µg/mL and peptide expression was induced with 1 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG). Cultures were then incubated overnight at 37 °C with shaking at 200 rpm to produce phage. Culture supernatants were filtered through 0.22 µm filters and phage were precipitated by adding 20% of the total volume of 20% PEG/2.5 M NaCl and incubating on ice for at least 30 min. Phage were centrifuged at 16,000 g, resuspended in 1x PBS and spun again twice more at 16,000 g changing the tube between spins. Titers were quantified by spectrophotometry (DeNovix DS11) using the formula virions/mL = (A369 – A420) $\times 6 \times 10^{16}$/(phage ssDNA base) (Smith & Scott, 1993).

2.1.1.4 ELISA analysis of control peptide phage

ELISA plates were coated with 50 µL volumes of Sodium Carbonate buffer pH 9.6 (34 mM $NaHCO_3$, 16 mM Na_2CO_3) containing 4 µg/mL anti PEP-1 (mAb 2C7) (produced in house) or 2 µg/mL anti-His (Thermo Fisher Scientific MA1-21315) overnight at 4 °C. The following day the plates were blocked with 360 µL of PBSMT (PBS containing 2% skim milk powder and 0.05% Tween20) for 30 min at room temperature. Control phage expressing PEP-1 or 6-His peptides were serially diluted in PBSM (PBS containing 2% skim milk powder) and incubated for 30 min at room temperature. The plates were washed 3x with PBST (PBS containing 0.05% Tween20) and 50 µL of anti-M13 HRP diluted 1:20,000 (Sino Biological MM05T) was added in PBSM. After 30 min of incubation the plates were washed a further 3 times before the addition of 50 µL 1-Step™ Ultra TMB-ELISA (ThermoFisher) for less than 10 min. The reaction was quenched with 50 µL 1 M H_2SO_4 and the plates read at 450 nm on a Spectramax M4. All ELISAs were repeated in triplicate.

2.1.2 TeeVax3 phage library construction

2.1.2.1 Tiling of TeeVax 3 antigen

The tile library was created computationally by splitting the 941 amino acid sequence of the TeeVax3 antigen (Appendix A5) into 308 segments, each 20 amino acids long. These peptides were offset by three amino acids to ensure complete coverage of potential binding sites. Bioinformatics software (Geneious Prime®) was then used to back translate the tiles into nucleotide sequences, followed by the addition of adaptors identical to those used in section 2.1.1.1 (Appendix A1) to both 5' and 3' ends. The resulting 308 tiles of 105 nucleotides each were then ordered as a single oligo pool (Twist Bioscience).

2.1.2.2 Cloning and production of TeeVax3 phage

The primary difference between the protocol followed to produce the control peptide phage (section 2.1.1) and the TeeVax3 library was the amplification step. Oligos were duplexed by a single cycle of PCR using OneTaq® Quick-Load® 2X Master Mix with Standard Buffer (NEB). A reverse primer (Appendix A3) was included at 0.2 µM in the reaction alongside 12 ng of oligo template and the

following conditions were used: denaturation at 94 °C for 1 min, 1 cycle of elongation at 60 °C for 1 min, followed by final extension for 5 min at 68 °C. The double-stranded oligo pool was digested with SfiI restriction enzyme and then cleaned as per the control peptide method, with DNA concentration being quantified using a DeNovix double-stranded Ultra sensitivity assay. The ligation, transformation, (section 2.1.1.1) colony PCR (section 2.1.1.2) and phage production (section 2.1.1.3) steps were all carried out as for the control peptides.

2.2 Nanopore sequencing of TeeVax3 libraries.

2.2.1 Preparation for sequencing

Library plasmid stocks from pre and post phage expression were directly purified from corresponding *E. coli* glycerol stocks using QIAprep Spin Miniprep Kit (Qiagen).

2.2.1.1 Restriction enzyme double digestion

Linear 723 bp DNA fragments representing the tile library coding sequence were extracted by restriction digest of each library plasmid stock with BamHI-HF and XbaI (NEB). Reaction mixtures were composed of 1 µg of DNA, 1µL of each enzyme, 5 µL of 10X rCutSmart Buffer and ultrapure water, which were incubated overnight at room temperature.

2.2.1.2 Gel extraction and purification

The resulting samples were run on a 2% Agarose gel with TAE buffer at 110 V for 45 mins, before being excised and purified with a QIAquick Gel Extraction Kit (Qiagen). Each purified library fragment was then prepared according to the SQK-NBD114.24 Oxford Nanopore Technologies (ONT, UK) ligation sequencing amplicons protocol and kit (200 fmol input).

2.2.2 Sequencing and data analysis

The resulting pooled sample was loaded onto a R10.4.1 flow cell (ONT). Sequencing was stopped once the estimated raw read count reached ≥100x coverage per tile per sample. Base-calling and demultiplexing of raw read data was

performed using Dorado (v0.5.3; model dna_r10.4.1_e8.2_400bps_sup@4.3.0). Read data were filtered for quality control using fastp (v0.23.438) and binned into tile counts via local alignment against the 308 segments using minimap2 (v2.2439) and samtools (v1.19.240).

2.3 Production of TeeVax3 antigen

2.3.1 TeeVax 3 antigen production

TeeVax3 antigen (provided by J Loh, Auckland University) in a pET-32a vector was transformed into BL21 (DE3) cells using a heat shock protocol. This glycerol stock was then stored at -80 °C. To produce the antigen, the glycerol stock was streaked onto a small agar plate of LB agar with 100 µg/mL Ampicillin and grown overnight at 37 °C. An overnight culture was grown from a single colony from the original plate in the same conditions using LB broth. In order to express the TeeVax3 antigen the 400 µL of overnight culture was inoculated into 400 mL of LB media with 100 µg/mL Carbenicillin. The culture was then grown to mid-log phase (optical density at A600nm = 0.4-0.6) at 37 °C and shaking at 110 rpm. Expression was then induced using 400µl 0.1 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG) for 18 hrs at 18 °C before centrifugation at 5000 g for 15 mins.

2.3.2 TeeVax 3 antigen purification

TeeVax3 antigen was isolated and refined using both nickel-nitrilotriacetic acid and size exclusion chromatography. The resulting sample was then analysed using SDS-PAGE to ensure accurate size and purity.

2.3.2.1 Nickel-nitrilotriacetic acid chromatography

Cell pellets containing the antigen were thawed on ice before resuspension at a ratio of three millilitres of lysis buffer (50 mM Tris-Cl pH8, 150 mM NaCl, 20 mM imidazole) for every milligram of wet pellet, with the addition of a protease inhibitor cocktail tablet (Roche). The cells were then sonicated using a Qsonica Q700 with a 417-A microtip for a total time of 3 mins (pulse-on time: 1 s, pulse-off time: 1s) and the supernatant recovered via centrifugation at 14000 g for 30 mins at

4°C. The resulting supernatant was sequentially filtered through 1.2 µm, 0.45 µm and 0.2 µm filters before Ni-NTA chromatography using 5 mL HisTrap HP columns (Cytiva) on a BioRad NGC FPLC allowed the His-tagged antigen to be captured. The unbound fraction was removed with a wash buffer of 50 mM Tris-Cl pH 8, 150 mM NaCl, 20 mM imidazole, and the purified TeeVax3 antigen was eluted using a 50 mM Tris-Cl pH 8, 150 mM NaCl, 1 M imidazole buffer and concentrated in Amicon Ultra-15 10 kDa centrifugal filters (Merck Millipore).

2.3.2.2 Size exclusion chromatography

The concentrated protein was syringe loaded onto a S75 16/60 column (GE Healthcare) using 50 mM Tris pH 8, 150 mM NaCl buffer at a rate of 1 mL/1min. Eluted sample was collected in 2 mL fractions which were then buffer exchanged into 1x Phosphate Buffered Saline (PBS) at a concentration of 0.6 mg/mL.

2.3.2.3 SDS-PAGE

SDS-PAGE gels (12%) manufactured by BioRad or made in house were used to analyse purity of the TeeVax antigen. In house gels consisted of a 12% resolving and 5% stacking layer made using a five gel multi-gel caster (Hoefer). The precise combination of components is listed in Table 2.1, with the resolving gel solution being prepared first and added to within three centimetres of the top of the caster. Isopropanol is then layered over the resolving gel solution and left to polymerise at room temperature. The Stacking gel solution is then prepared and added to the top of the caster after careful removal of the isopropanol layer. A 10 well comb is inserted in each of the five gels before leaving them to polymerise at room temperature for around 30 mins. Gels were then stored for up to four weeks at 4°C. Two - six µg of TeeVax antigen in PBS was made up to 15 µL with ultrapure water before the addition of 5µL of SDS-loading dye (250 mM Tris pH 6.8, 20% (w/v) glycerol, 4% (w/v) SDS, 10% (w/v) beta-2-mercaptoethanol, 0.025% (w/v) bromophenol blue). Samples were then incubated at 95 °C for five mins before loading into the gel. 15 µL of antigen samples were loaded alongside 10 µL of Precision Plus Protein^{TM} (Bio-Rad) and run in 1x Tris-Glycine SDS Running Buffer at 110 V until the dye line reached the bottom of the gel.

Table 2.2. SDS-PAGE gel components. Makes 5 gels with APS prepared weekly.

Component	12% Resolving gel (mL)	5% Stacking gel (mL)
Resolving Buffer (1.5 M Tris pH 8.5)	7.5	-
Stacking Buffer (1 M Tris pH 6.8)	-	1.6
30% Acrylamide	12	2.125
10% SDS (w/v)	0.3	0.125
10% APS (w/v)	0.15	0.063
TEMED	0.015	0.0063
Milli Q water	10.05	8.5

2.4 Isolation of TeeVax3 specific polyclonal antibodies

2.4.1 Animal experiments

All animal experiments were approved by The University of Auckland Animal Ethics Committee (Approval #R1663) and were performed in the Vernon Jansen Unit at The University of Auckland. All methods and procedures were performed according to the regulations and standards of the University of Auckland. Experimental methods are reported in accordance with the ARRIVE guidelines. TeeVax3 immunised rabbit serum was obtained from a New Zealand white rabbit immunised subcutaneously with 100 µg of recombinant protein emulsified 1:1 with incomplete Freund's adjuvant and boosted at 2 and 4 weeks. Control serum was obtained in parallel from an unimmunised individual. All animals were deeply anaesthetized at day 42 with 90 mg/kg of intravenous sodium pentobarbitone, and antiserum was collected by terminal blood draw.

2.4.2 Custom resin production

The resin was created by coupling TeeVax3 antigen to AminoLink Plus Coupling Resin (ThermoFisher). Eight mg of TeeVax3 antigen in 0.1 M sodium phosphate, 0.15 M NaCl, pH 7.2 coupling buffer were incubated with 2 mL of resin at 4 °C overnight with gentle rotation. The following day, 5 mL of pH 7.2 PBS and 100 µL of Cyanoborohydride Solution were added and the mixture rotated for 4 hrs at room

temperature. Resin flowthrough was collected for analysed using a Nanodrop™ 2000 (Thermofisher, USA) and verified by SDS-PAGE.

2.4.3 Purification of TeeVax specific antibodies

Two mL of conjugated resin was added to a gravity flow column, and sera diluted five-fold in PBS was passed through the column. The bound fraction was washed with PBS, eluted with 0.1 M glycine pH 3, and neutralized with 10% the solution volume of 1 M Tris pH 9.0. Amicon Ultra-15 10 kDa centrifugal filters (Merck Millipore) were used for buffer exchange into PBS prior to purity analysis by SDS-PAGE. In parallel, Negative control antibodies from unimmunised rabbits were purified by Protein G chromatography. Sera from these animals was again diluted more than five-fold in PBS and passed through a 1 mL Protein G column (Cytiva). After washing with >10 column volumes of PBS, the bound antibody pool was eluted with Glycine HCl pH 3.0, neutralized with 1 M Tris pH 9.0 and buffer exchanged into PBS following the method used for anti-TeeVax3 antibodies.

2.5 Analysis of background binding by ELISA

To perform ELISAs to measure background antibody binding, the plates were coated with 4 µg/mL of polyclonal antibodies isolated from rabbit serum (section 2.4.3) alongside 2 µg/mL anti-His as a positive control. The remaining steps were as outlined in section 2.1.1.4 with the TeeVax3 phage being applied in place of the control peptide phage for rabbit serum wells and 6-His control phage for the positive control.

2.6 Linear epitope profiling of the TeeVax3 antibody response

2.6.1 Phage panning

Phage panning was carried out in ELISA 8 well strips (Corning). The wells were coated overnight at 4 °C with 100 µL of purified anti-TeeVax3 antibodies at 4 µg/mL in 50 mM Na_2CO_3 buffer at pH 9.5 or incubated with Na_2CO_3 buffer alone. The next day, the solutions were removed, and the plates were blocked with 300 µL PBSMT. After 1 hour of shaking at room temperature, the strips were washed 3x with PBST. Aliquots of 1 x 10^{11} phage per well were then added for 1 hour to the

strips coated with buffer alone for non-specific phage binding depletion before transfer to the wells coated with antigen. The wells were washed again after 1 hour, although the number of washes was increased by three each round to increase selection pressure. Phages were eluted with a 10 min incubation using 100 µL of 0.1 M glycine/HCl pH 2.7, transferred to a 2 mL tube previously blocked with PBS containing 2% skim milk powder, and neutralized with 1 M Tris pH 8.0 at a 1:5 ratio (vol Tris/vol phage). Twenty mL SS320 E. coli cells were grown in 2xYT media and 2% glucose medium to an OD 0.4-0.6 at 37 °C without shaking, and incubated with the eluted phages for 30 min at 30°C. The shaking speed was then increased to 250 rpm for another 30 min. The phage library was recovered by plating infected cells on three 145 mm round LB agar plates containing 2% glucose and 20 µg/mL chloramphenicol and overnight growth at 37°C.

2.6.2 Nanopore sequencing of panned TeeVax3 libraries

Sequencing of the panned libraries was carried out as per the protocol outline in section 2.2 with proportions of tiles being calculated for each library before comparison.

2.7 Exploration of the high affinity peptide tag.

2.7.1 Confirmation of epitope tag binding

2.7.1.1 Histidine tag removal

To remove the His tag, 1 mg of TeeVax3 antigen was buffer exchanged into 20 mM Tris pH 7.5 using Amicon Ultra-15 10 kDa spin columns to a concentration of 3.5 mg/mL. 300 µg of antigen was incubated with 10 µL of TEV Protease (NEB) and 50 µl of reaction buffer in a total volume of 500 µL at 4 °C overnight. Both digested His tags and undigested antigen were isolated by Nickel-nitrilotriacetic acid chromatography as outlined above, with the flowthrough fraction collected as untagged antigen.

2.7.1.2 ELISA analysis of tagged and untagged antigen

ELISA analysis was used to compare antibody binding to TeeVax3 with and without the His epitope tag. Antigens were coated overnight at 4 °C on ELISA

plates (Corning 3590) at 4 µg/mL in Coating Buffer (50 mM $NaHCO_3$, pH 9.5), washed the next day three times with PBST, and blocked with PBSMT. Blocked plates were washed a further three times with PBST. Serial dilutions of TeeVax3-specific antibodies isolated from rabbit sera were added to the plate, and a goat anti-rabbit IgG HRP antibody (Abcam) was used at 1:50,000 for detection of binding. Control samples confirming His tag removal from the TeeVax3 antigen used a HRP conjugated anti-polyHistidine antibody [HIS-1, Abcam] diluted 1:20,000 in PBSM. 50 µL 1-Step™ Ultra TMB-ELISA was used for detection of HRP signal in each well and the reaction quenched after 10-15 min with an equal volume of 1 M H_2SO_4 before reading at 450 nm. All ELISA steps were undertaken at room temperature with incubation periods of 1 hour unless otherwise indicated.

2.7.2 Profiling epitope tag via ELISA

Fine epitope mapping was performed by ordering synthetic peptides with N-terminal biotin residues (Genscript, Appendix C1). ELISA analysis followed the protocol above except the ELISA plates were coated with biotinylated peptides that had been precomplexed with neutravidin at a 4:1 molar ratio for 30 min at room temperature in 50 mM $NaHCO_3$, pH 9.5. ELISA data was fit by non-linear regression of a reparametrized 5 parameter logistic model in JupyterLab (Liao & Liu, 2009). All ELISAs were repeated in triplicate. Pairwise t-tests were used to compare data points at each antibody concentration and a Bonferroni correction was applied to account for multiple comparisons.

Chapter 3

Results

3.1 Construction of tiled TeeVax peptide library for phage display

To create a rapid protocol for investigating the linear epitope profile of the TeeVax3 antibody response, we first had to confirm the suitability of our phage display system for peptide expression. The two control peptides (Appendix A2) - a cyclic peptide mimetic of lipooligosaccharide derived from the pathogen *N. gonorrhoeae* (PEP-1) (Ngampasutadol et al., 2006) and the ubiquitous 6-His epitope tag, were selected due to the availability of high-affinity antibodies to target these peptides.

3.1.1 Validation of library design

The initial step in this MSc research was the cloning of control peptides into pAK200 phage display vectors (Krebber et al., 1997), which were sourced before this work began from Prof George Georgiou at the University of Texas. Both peptides were amplified using primers previously created by Kyrin Hanning (Appendix A3), before digestion with the restriction enzyme SfiI. The phage display vector was also prepared with SfiI digestion and separated via agarose gel electrophoresis (Figure 3.1) before purification with gel excision. Digested peptide and vector were then ligated and transformed into electrocompetent SS320 *E. coli* cells prior to plating. Eight colonies of each peptide were selected for PCR amplification (Appendix A4 for primers) with inserts of the expected sizes, 253 bp and 244 bp for the PEP-1 and 6-His peptides respectively, being observed (Figure 3.2). Sanger sequencing results from the Massey Genome Service (Massey University, New Zealand) further confirmed the cloning had been successful (Appendix A6).

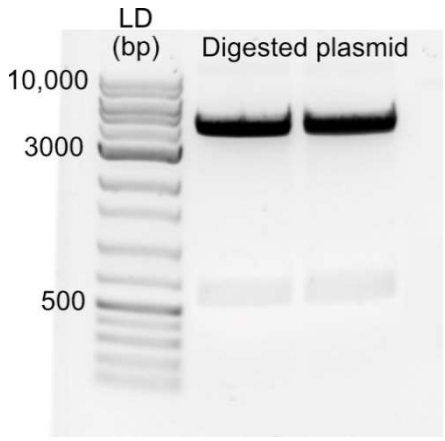


Figure 3.1. 1% Agarose gel of SfiI digested pAK200 phage display vector.

Ladder used was SiZerTM 1000 plus DNA Marker Solution, which indicates plasmid backbone at expected size of 4287 bp.

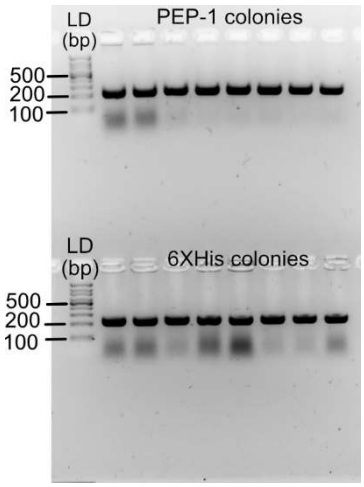


Figure 3.2. 3% Agarose gel of single colony PCR's for both the PEP-1 and 6-His peptides. 100 bp Ladder (NEB) used. Bands correspond to insert size of 253 bp for PEP-1 and 244 bp for 6-His.

Phage were then produced from these transformed stocks and ELISA analysis was used to confirm the expression and display of both control peptides by binding with anti-PEP-1 (mAb 2C7) and anti-His (HIS.H8) monoclonal antibodies (Figure 3.3).

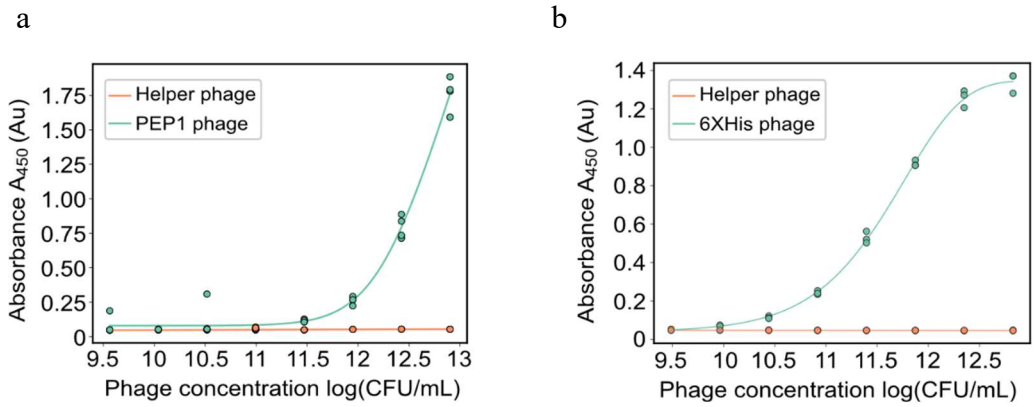


Figure 3.3. ELISA analysis illustrating the binding of control phage compared to helper phage, confirming peptides are displayed on phage surface. (A) Phage expressing PEP1 cyclic peptides bound by 2C7 monoclonal antibodies. (B) 6-His displaying phage binding anti-His antibodies. The 6-His data was collected by Kevin Beijerling during revisions of the published manuscript.

3.1.2 TeeVax3 phage library construction

After confirming the expression of short peptides on the phage surface, we produced a library of peptide tiles from the TeeVax3 sequence. The 940 amino acid protein (Appendix A5) was computationally broken into 20 amino acid tiles with each tile offset by 3 amino acids, except for the 2 amino acid offset in the final tile. This process resulted in a final library diversity of 308 individual tiles. Each peptide sequence was then back translated and adaptor sequences identical to the control peptides were added using Geneious Prime®. The resulting library was ordered as an oligo pool from Twist Biosciences and amplified by a single round of PCR, before cloning into pAK200. The initial transformation into E *.coli* (SS320) in accordance with the control peptide protocol was performed, and colony PCR was carried out for 16 colonies from the TeeVax library plate together with a single colony that appeared on the negative control plate (Figure 3.4). The subsequent agarose gel showed bands near the expected size of 266 bp for all 16 colonies, with no band visible for the negative control sample. The presence of the TeeVax3 tiles was confirmed with sequencing completed by Massey Genome Service (Massey University, New Zealand) (Appendix A7).

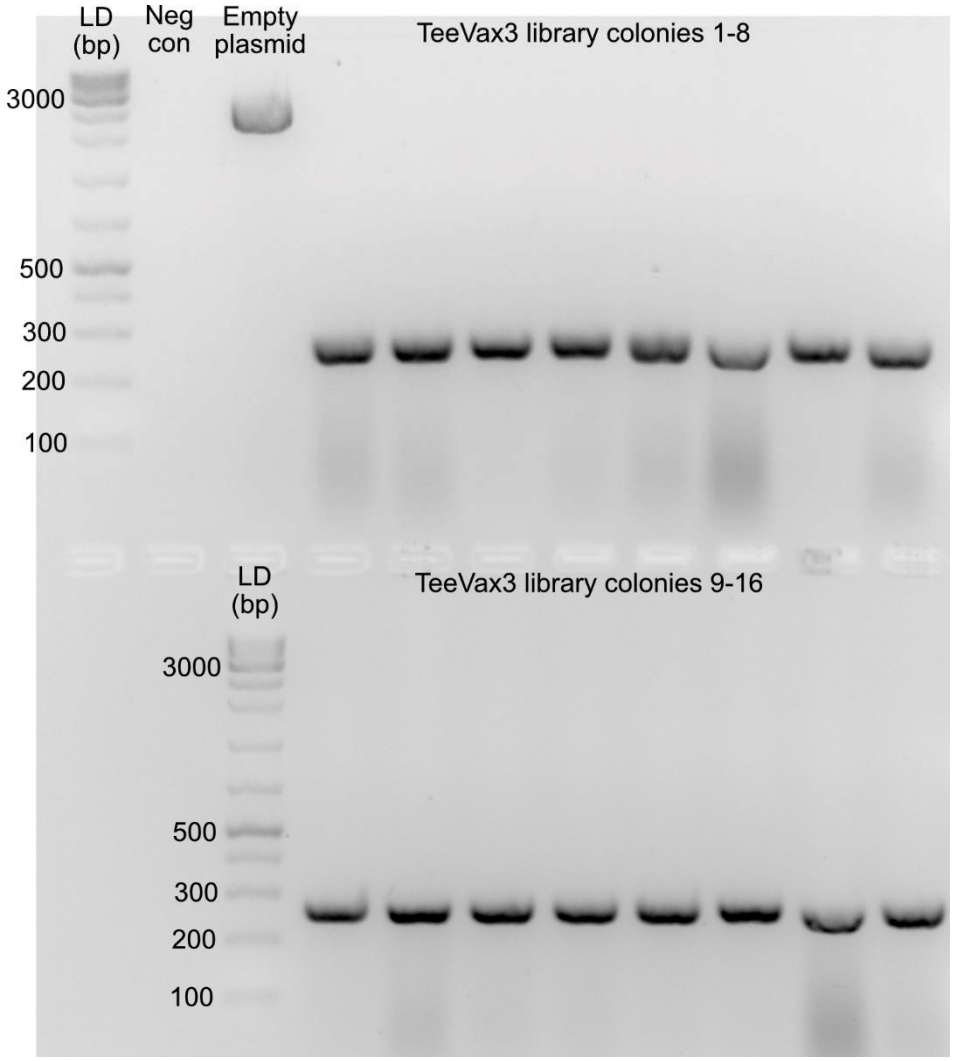


Figure 3.4. Colony PCR of TeeVax3 library run on 3% agarose gel. The ladder used was SiZer^{TM} 1000 plus DNA Marker Solution, with the negative control (Neg con) being a colony from the control plate and empty pAK200 plasmid employed as a positive control. All 16 colonies from the TeeVax3 library plate produced bands near the expected 266 bp, with the empty plasmid being near 2254 bp as predicted.

After verifying the successful cloning of the TeeVax3 tile library, the transformation was repeated on a larger scale using fresh *E. coli* (SS320) cells. This transformation resulted in 26,260 colonies being collected across all plates, giving a coverage of over 85 times the predicted library diversity. Phage were then created and purified, before being used to reinfect more *E. coli* (SS320) cells, with the

resulting population of phage infected bacteria being collected for nanopore sequencing.

3.1.3 Sequencing of unpanned TeeVax libraries

To first monitor library expression bias caused by the display of peptides on the phage surface, samples of both the initial bacterial stock (before phage production) and expressed stock (post phage production) of the TeeVax3 library were extracted for Nanopore sequencing. To avoid the introduction of inadvertent amplification errors, we used a double digest of BamHI and Xbal restriction enzymes. The resulting fragments were then separated on a 2% Agarose gel. The expected bands of size 727 bp were observed (Figure 3.5), confirming adequate digestion of the library.

These two samples were then excised and purified in preparation for Nanopore sequencing using a Ligation sequencing amplicons protocol from Oxford Nanopore technologies. Sequencing was carried out for both samples simultaneously, obtaining over 1.4 million reads with a Q score cutoff of 15. I then bioinformatically binned the reads into the corresponding 308 tiles using Minimap2, with support from Kyrin Hanning who refined the scripts for publication (Appendix B1).

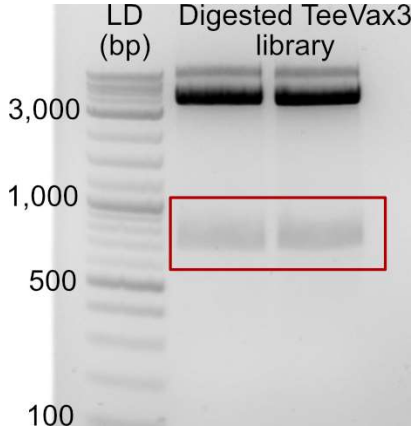


Figure 3.5. 2% Agarose gel of digested TeeVax3 library plasmid. Lane 1 contains NEB 1kb plus ladder and Lanes 2 and 3 the digested plasmid. The tile fragments for each lane are highlighted and the remaining 3651 bp of the plasmid is visible in the upper portion of the gel.

All 308 tiles were observed in both the initial library and the phage expressed library (Figure 3.6a), with the pre library individual tile frequencies ranging from 0.07%

to 0.57% indicating a reasonably even distribution within the sample (Figure 3.6b). The expressed library tile frequencies had a wider range of 0.036% to 1.82%, however, when mapped against location in the antigen (Figure 3.6c) there did not appear to be any regions of concentrated bias present.

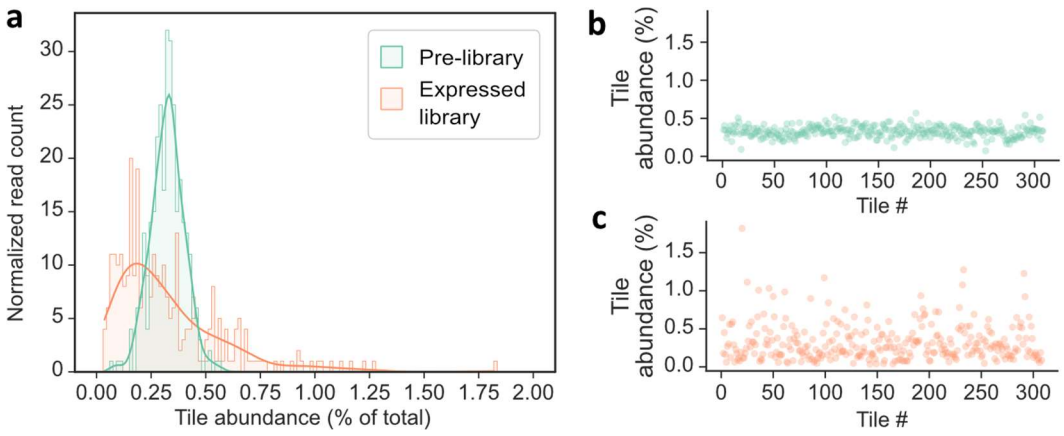


Figure 3.6. Nanopore sequencing evaluation of TeeVax3 library before panning. (a) Normalised read count distribution for initial library (green) and post phage expression (orange) with kernel density estimations shown by the solid lines. (b) Indicates tile abundance of the initial library according to location in the antigen with (c) illustrating this for the phage expressed library.

3.2 Production of TeeVax3 antigen and isolation of vaccine specific antibodies

Investigating the antibody response to immunisation with the TeeVax3 vaccine candidate required adequate samples of TeeVax specific polyclonal antibodies. These were purified from immunised rabbit serum using a custom resin bound with TeeVax3 antigen, which was expressed and purified in advance.

3.2.1 TeeVax3 expression and purification

Cell pellets from expression cultures of TeeVax3 antigen were resuspended, sonicated and centrifuged, before the resulting supernatant was collected for purification by immobilized metal affinity chromatography. The successful isolation of the antigen was confirmed by the single 3000 mAu peak observed on

the chromatogram (Figure 3.7), with corresponding fractions being collected for further purification by size exclusion chromatography.

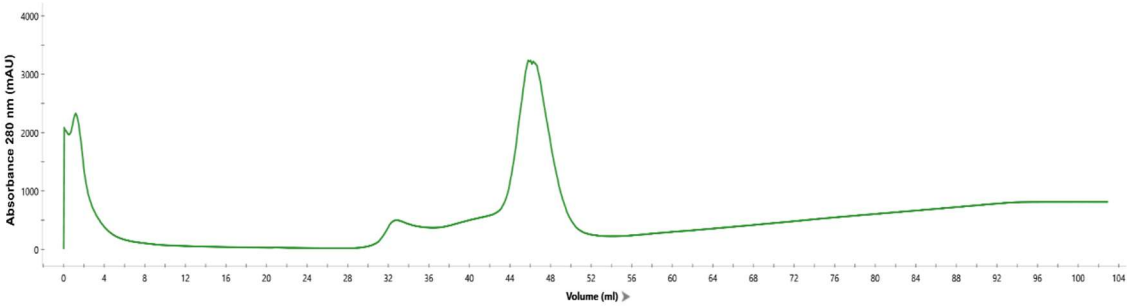


Figure 3.7. Chromatogram of Ni-NTA purification using 5 mL HisTrap HP columns (Cytiva) on a BioRad NGC FPLC. A clear increase in absorbance at 280 nm is observed as protein was eluted from the column.

The isolated sample was then analysed on a 12% SDS-PAGE gel where 3.6 µg of the 105 kDa TeeVax3 antigen was visible as a distinct band (Figure 3.8). We also observed some faint banding below the primary band; however, these were not investigated in greater depth given the apparent high purity of the preparation.

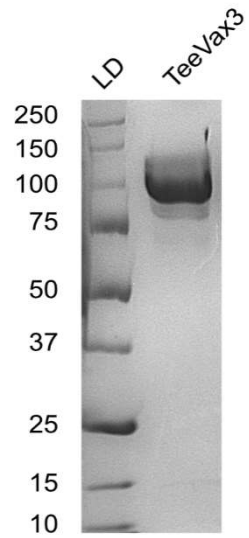


Figure 3.8. SDS-PAGE gel of purified TeeVax3 antigen at expected size of 105 kDa. Precision Plus Protein™ Unstained Standards (Bio-Rad) ladder used with units in kDa.

3.2.2 Pulldown of TeeVax3 specific polyclonal antibodies

3.2.2.1 Creation of TeeVax specific chromatographic resin

The purified TeeVax3 antigen sample was then utilised to create the custom resin for TeeVax3 specific polyclonal antibody collection (Figure 3.9). Eight mg of antigen was conjugated to AminoLink Plus Coupling Resin™, with samples of the loading solution and resin flowthrough being collected to determine the conjugation efficiency.

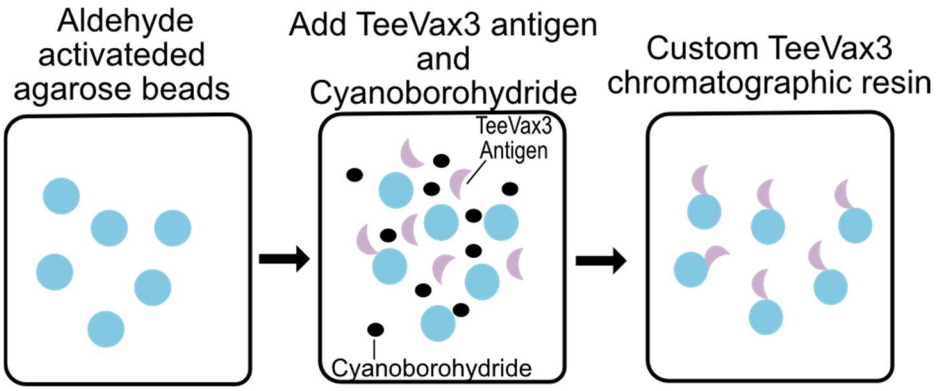


Figure 3.9. Schematic of how custom TeeVax resin was produced by binding TeeVax 3 antigen to AminoLink Plus Coupling Resin (ThermoFisher). Aldehyde-activated agarose beads were exposed to TeeVax3 antigen in the presence of Cyanoborohydride solution to bind the antigen to the resin by reductive amination.

The coupling efficiency of the resin production was calculated to be 97.6 %, found by comparing the concentrations of the TeeVax3 loading solution and the resin flowthrough as measured by the absorbance at 280 nm. The conjugation efficiency was independently confirmed with SDS-PAGE analysis showing no visible TeeVax3 protein in any of the flowthrough samples loaded (Figure 3.10).

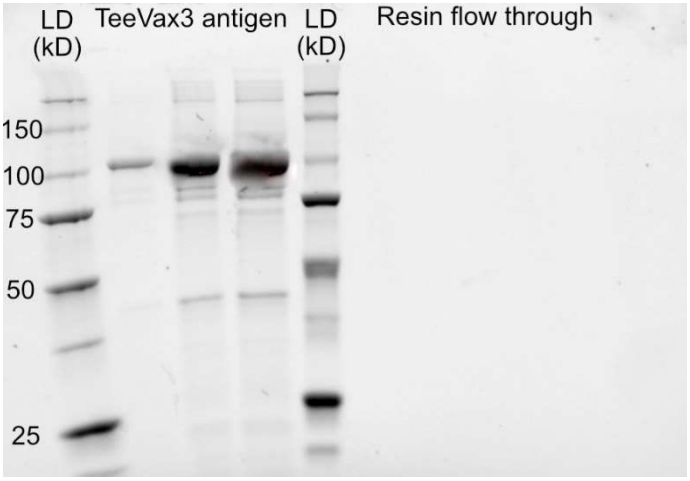


Figure 3.10. SDS-PAGE gel of TeeVax3 antigen (2.67µg/µL) and resin flowthrough samples. The ladder in lanes one and four is Precision Plus Protein™ Standards (Bio-Rad), lanes two, three and four contain 1 µL, 2 µL and 3 µL of TeeVax3 antigen, while lanes five, six and seven contain column flowthrough at the same volumes illustrating the vast majority of antigen is bound to the resin.

3.2.2.2 Purification of TeeVax3 binding antibodies

Serum from rabbits immunised with TeeVax3 was loaded onto the resin to purify antibodies recognising both linear and conformational epitopes on the TeeVax3 antigen. Approximately 1.3 mg of polyclonal antibodies were obtained for every mL of rabbit sera. SDS-PAGE was used to verify the purity of the sample (Figure 3.11). While multiple bands were observed within the heavy and light chain fractions, this is expected given the polyclonal nature of the isolated antibodies, which comprise multiple isotypes with variable glycan states and diverse complementarity-determining regions.

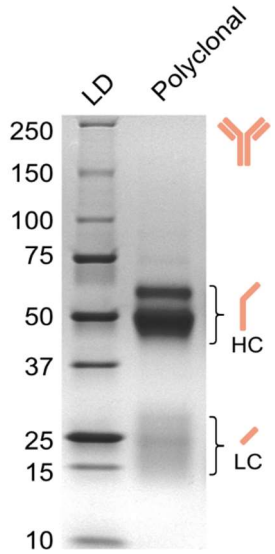


Figure 3.11. 12% SDS-PAGE gel of isolated TeeVax3 specific polyclonal antibodies from immunised rabbit sera. An assortment of heavy chain (HC) and light chain (LC) sizes are visible as labelled.

3.3 Analysis of background binding of antibody repertoires with TeeVax3 libraries.

To determine the background binding of the antibody pool to the TeeVax3 phage library, ELISA analysis was undertaken by Kevin Beijerling. This compared the TeeVax3 immunised antibody repertoire to that of unimmunised rabbits (Figure 3.12). As a negative control, antibodies from the sera of unimmunised rabbits were purified using a protein G pulldown, which specifically isolates IgG isotypes. The unimmunised antibody pool showed a 1.3-fold increase in binding to TeeVax phage compared to the PBSM background negative control, indicating low but detectable levels of non-specific binding. In contrast, the TeeVax3 specific antibody pool exhibited the highest binding signal to the TeeVax3 phage library, with a 4.7-fold increase compared to the unimmunised samples. The positive control, anti-His antibodies (His.H8), was included to confirm detection of the 6-His control phage by the anti-helper phage secondary antibody.

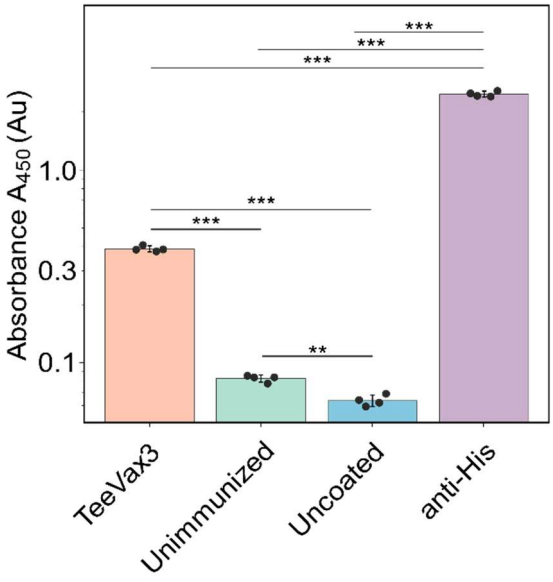


Figure 3.12. Background binding of polyclonal antibodies to the TeeVax3 phage library via ELISA. Plates were coated with anti-TeeVax3 antibodies, antibodies purified from unimmunised sera with protein G chromatography, and a positive control of anti-His antibodies binding control His-phage. PBSM was used to calculate background signal, and all points are a single well (n=4). Statistical significance was calculated using a pairwise t-tests with a Bonferroni correction applied for multiple comparisons (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$). This data was collected by Kevin Beijerling during revisions for publication with data analysis completed by Dr William Kelton.

3.4 Linear epitope profiling of the TeeVax3 antibody response

Once the TeeVax3 phage library was created and the TeeVax specific antibodies isolated, investigation of the epitope interactions could begin.

3.4.1 Phage panning to isolate and enrich binding epitopes

Phage panning involves exposure of a phage library to a population of antibodies to allow collection and amplification of the binding phage with the corresponding displayed peptide. In this protocol, the polyclonal TeeVax3 antibody repertoire was bound to a plate which the TeeVax3 phage library were then exposed to. Unbound phages were washed away, leaving behind those that interacted with the antibodies.

The positive phage fraction was then eluted by lowering the pH and collected (Figure 3.13).

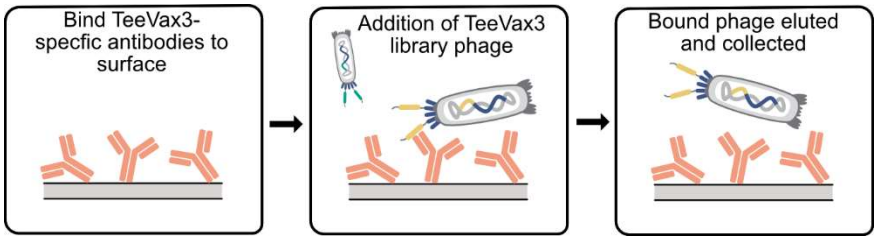


Figure 3.13. Illustration of the phage panning process. TeeVax3 specific polyclonal antibodies are bound to the plate, the TeeVax3 peptide library phages are then exposed to the antibodies with unbound phage being washed away. Bound phages are then eluted and collected for amplification and subsequent panning.

The panning process was completed in three sequential rounds with the number of wash steps increasing by three each round to ensure high-affinity binders were being selected. The collected binding fractions from each stage were used to infect cultures of *E.coli* (SS320) and produce more phage to be purified for the next panning step. A sample of the purified phage was also collected at each round for Nanopore sequencing. In addition, colonies were selected from the final round of panning to be grown for Sanger sequencing at Massey University. Of the nine colonies sent for sequencing, all were aligned to tile one indicating that three rounds of panning had been sufficient for enrichment.

3.4.2 Quality control of Nanopore sequencing reads.

The DNA samples from round one, two, and three of panning were extracted and prepared for sequencing as outlined for the pre-library (Round 0) and expressed library (Round 0 reinfected) (section 2.2). Sequencing was completed and base calling was performed by Kyrin Hanning using Dorado-0.5.3. The resulting density plots illustrate the average q-score for each library at different read lengths (Figure 3.14a), with the dark blue colour indicating the values possessed by the majority of reads. These plots convey the high quality of the data collected as the densest areas correspond to the expected tile length of 723 bp with Q scores of greater than 15. The histograms in Figure 3.14b depict the distribution of read lengths, supporting

the findings from Figure 3.14a. In addition to exhibiting the total read count for each library, with over 64,000 high quality reads selected for further analysis- providing more than 200-fold coverage of the TeeVax3 library diversity.

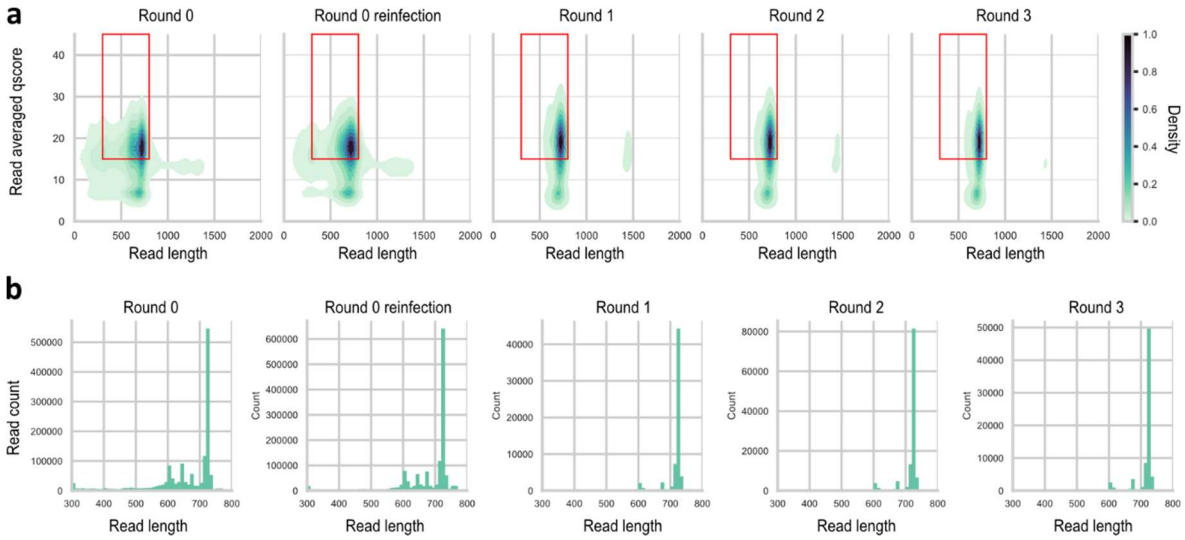


Figure 3.14. Quality control data from Nanopore sequencing. (a) Density plots of average Q score for all reads compared to read length for libraries at each stage of the method. Red boxes indicate reads selected for further analysis. (b) Histograms of the number of reads at each length for libraries pre panning and following each panning round. Data analysis was supported by Kyrin Hanning who refined the scripts used for publication.

3.4.3 Clonal enrichment of TeeVax3 tiles during panning.

Selected reads, that met the quality control criteria for read length and Q score, were then binned via local alignment into the 308 tiles. The substantial increase in tile 1 and tile 2 in the first round of panning (Figure 3.15), confirmed the enrichment indicated by the earlier Sanger sequencing (section 3.4.1). These two tiles, which included the Histidine tag and TEV cleavage site sequences from the TeeVax3 purification tag, represented 81.1% and 9.5% of the round 2 panning population, with no change being observed in round 3.

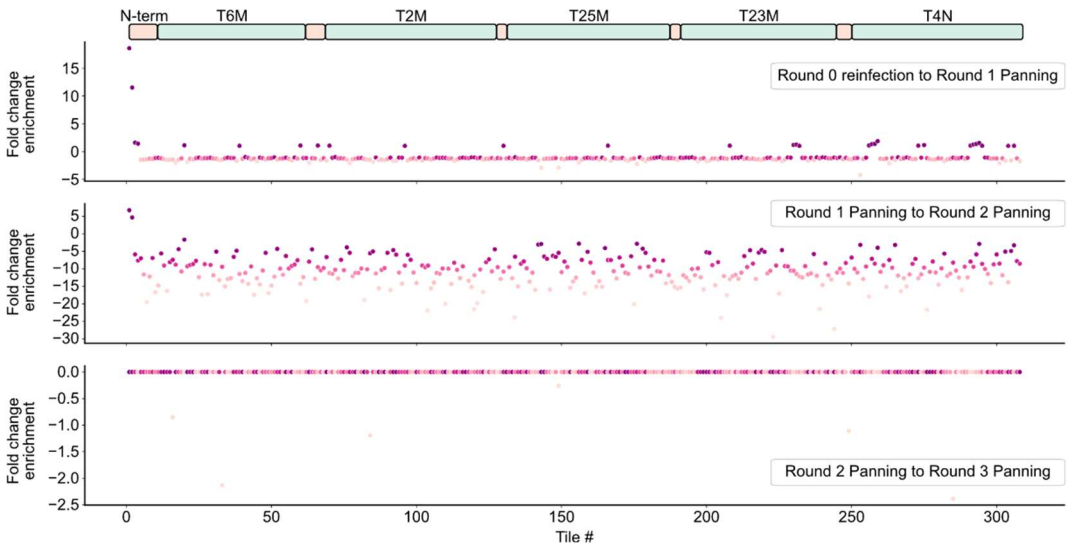


Figure 3.15. Enrichment of peptide tiles during the panning process. Each dot represents a single tile mapped against location in the TeeVax3 antigen. Enrichment values were calculated by comparing the proportion of each tile present in successive panning steps. I completed the initial data analysis which was further refined by Kyrin Hanning for publication.

To determine if this result was due to expression bias, the proportion of tile 1 and tile 2 in the pre library and expression library were compared. Tile 1 was found to increase 1.9-fold, from 0.347% to 0.645%, with tile 2 decreasing 0.44-fold from 0.404% to 0.176%. Although tile 1 did show a slight expression advantage, the fact that tile 2 did not, and the substantial 10-fold enrichment seen for both tiles in round 1, would suggest the affinity for the Histidine tag and TEV cleavage site to be a real phenomenon.

To explore regions of immunogenicity beyond the epitope tag, tiles 1-3 were excluded from further analysis. The remaining tiles with enrichment in the 95th percentile for round 1 panning were mapped onto an AlphaFold model of the TeeVax3 antigen. This identified regions of increased binding in the T4N, T6M and T25M domains in addition to the linking region between T6M and T2M domains (Figure 3.16), while T2M and T23M domains appear to have fewer immunogenic regions.

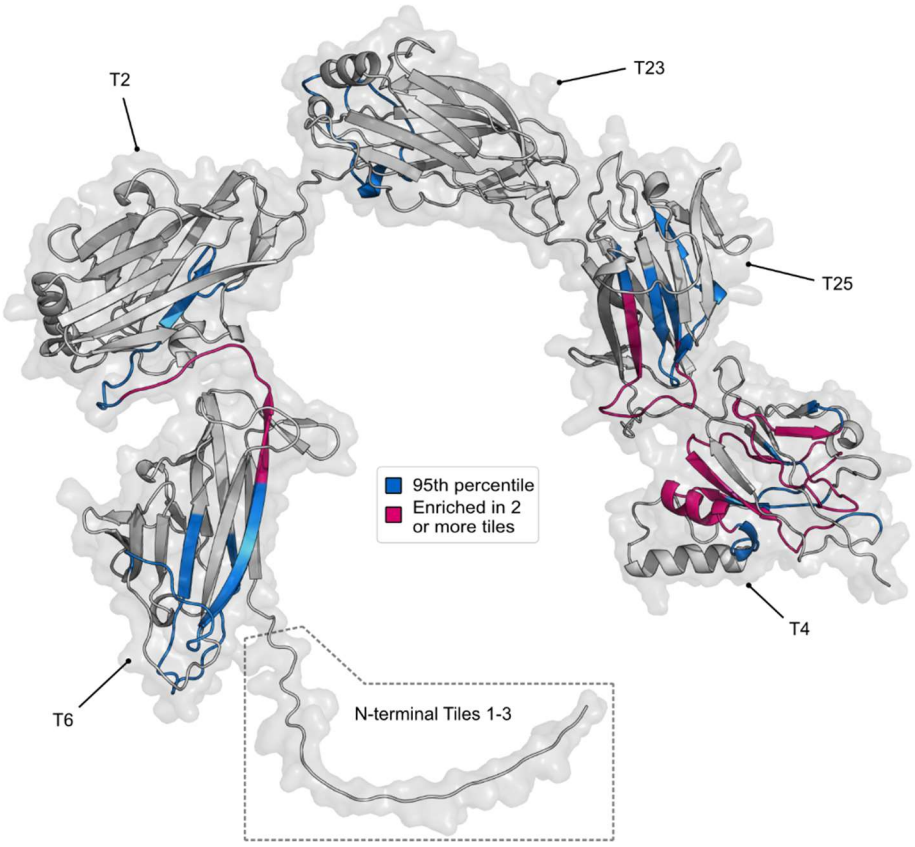


Figure 3.16. Model of TeeVax3 antigen with 95th percentile enriched tiles from the first round of panning, excluding tiles 1-3. Sequences represented by two or more tiles are coloured magenta. TeeVax3 model made in AlphaFold2, with the model and data analysis completed by Dr William Kelton.

3.4.4 Exploration of a high-affinity peptide within the epitope tag

Given the unexpected immunogenicity of the epitope tag, we wanted to confirm that this result was significant when both linear and conformational epitopes were considered. This was accomplished by evaluating the TeeVax3 antibody repertoire response to removal of the histidine tag from the TeeVax3 antigen. Furthermore, profiling the specific amino acid residues involved in the response was explored via the binding of a series of short synthetic peptides created from tile 1 and 2 sequences.

3.4.4.1 Histidine epitope tag removal

TeeVax3 antigen was incubated with TEV protease (NEB) overnight to cleave the epitope tag. The sample was then loaded onto a HisTrap HP column (Cytiva) to capture the removed tag and any undigested sample, with the flowthrough collected as untagged antigen. SDS-PAGE confirmed the presence of TeeVax3 antigen in the flowthrough sample (Figure 3.17), which showed a visible band at the expected size.

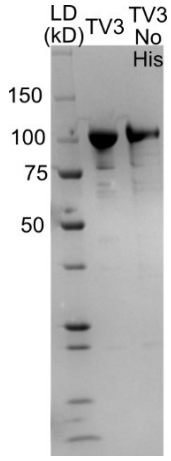


Figure 3.17. SDS-PAGE gel of TeeVax3 antigen before and after removal of the Histidine tag with TEV protease. Bands show the presence of protein at expected the size of 105 kDa.

To verify the successful removal of the histidine tag, we carried out a series of single well ELISAs (Figure 3.18). Histidine tagged TeeVax3 antigen exhibited substantially higher signal when exposed to anti-histidine antibodies, in comparison to His tag removed antigen, confirming successful cleavage of the epitope tag.

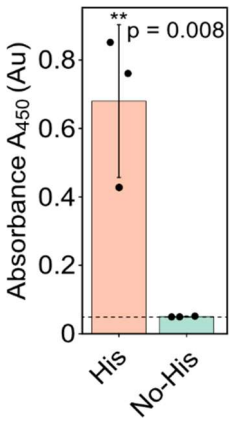


Figure 3.18. Confirmation of Histidine tag removal by single well ELISA. Data are presented as the mean ± SD (n = 3) and a pairwise t-test was performed to calculate the p-value (p = 0.008). The dashed line represents background binding. I optimised the protocol and collected initial data. The data presented here was collected and analysed by Dr William Kelton for publication.

3.4.4.2 Confirmation of epitope tag binding preference in isolated TeeVax3 specific antibodies.

A series of ELISAs were performed to determine what proportion of the TeeVax3 specific antibody population was interacting with the Histidine tag (Figure 3.19). Removal of the tag resulted in a 6.3-fold decrease in signal, illustrating that a large fraction of the purified TeeVax3 antibodies were in fact linear binders of the epitope tag even with the inclusion of antibodies recognising conformational epitopes.

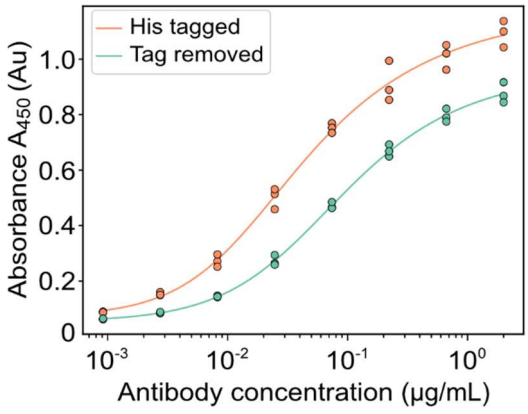


Figure 3.19. ELISA plot of the TeeVax3 polyclonal antibody response to TeeVax3 antigen with and without the histidine tag. Data points are graphed individually (n=3). Data analysed for publication by Dr William Kelton.

3.4.4.3 Comprehensive profile of epitope tag immunogenicity.

To determine which regions within the epitope tag were presenting with the highest immunogenicity, five synthetic peptides were designed from the sequences of tile 1 and 2. The longest 17 amino acid chain encompassed the whole overlapping region of the two tiles and the shortest 7 amino acid segment was composed of the histidine tag with an additional tyrosine residue (Figure 3.20).

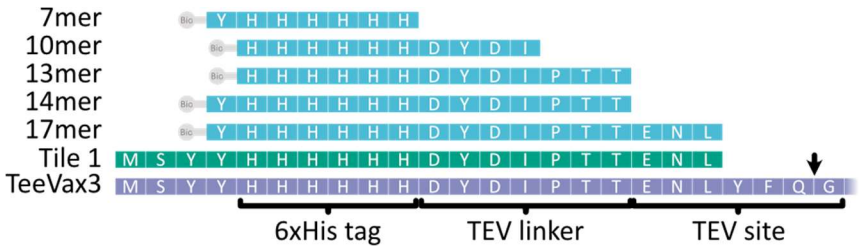


Figure 3.20. Series of biotinylated synthetic peptides generated from the overlapping sequence of tile 1 and 2.

These peptides were bound to plates with neutravidin and exposed to TeeVax3 specific antibodies to assess their binding ability (Figure 3.21). The longest 17 amino acid residue peptide displayed the strongest signal with the shortest 7-mer exhibiting considerably less. The three middle length peptides, 14-mer, 13-mer and 10-mer, all displayed similar results that fell in between the two extremities. This evidence strongly implies the sequence of immunogenicity includes both the histidine tag and the TEV linker regions.

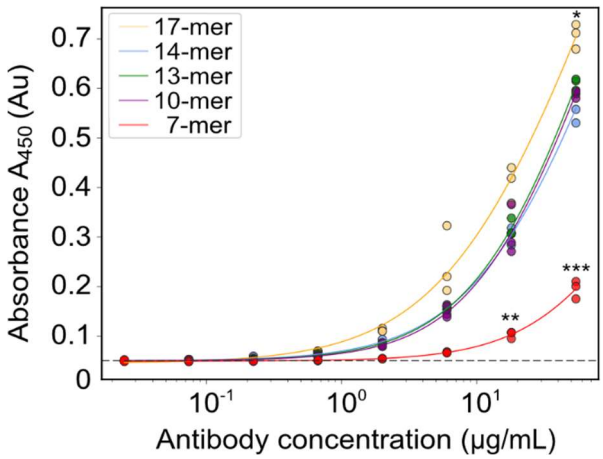


Figure 3.21. ELISA analysis of TeeVax3 polyclonal antibodies binding the peptide series. I optimised this method and prepared the plates and solutions, with Dr William Kelton collecting and analysing the data for publication.

Chapter 4

Discussion

Techniques that allow rapid, high throughput exploration of the antibody response to antigens, whether from vaccine candidates or pathogens, are crucial for the efficiency of the vaccine development pipeline. This thesis endeavoured to implement a protocol to rapidly map the linear epitope binding antibodies produced in response to immunisation with the Streptococcus A vaccine candidate TeeVax3. An oligo pool of overlapping tiles created from the TeeVax3 sequence was cloned into phage display vectors to create a phage library, which was subsequently panned against TeeVax specific polyclonal antibodies. The interactions were rapidly analysed with Nanopore sequencing to observe tile enrichment. Upon analyses of the enrichment data from three rounds of panning, there was a clear bias towards the first two tiles in the antigen sequence, which correspond to the histidine tag and TEV cleavage sequence on the N terminus. To determine the proportion of the TeeVax3 specific antibody population that was causing this effect, we compared the binding ability of the TeeVax3 antigen with and without the epitope tag. These results confirmed the dominance of linear epitopes recognising the N terminal tag in the TeeVax3 antibody response and led us to investigate which amino acid residues were contributing to this effect. A series of five biotinylated peptides were designed, and all were observed to bind with the TeeVax3-specific antibody pool with a 10 amino acid sequence encompassing the 6-His tag and TEV linker sequence identified as the most immunogenic segment.

Previous linear epitope mapping studies have used high accuracy sequencing techniques, such as Illumina (Mohan et al., 2018; Qi et al., 2021; Xu et al., 2015), which can be restrictive in terms of accessibility and consequently result in a slow turnaround. Illumina also requires PCR amplification of samples before sequencing which has the potential to introduce errors and bias to the resulting data. To circumvent this issue, we employed Nanopore sequencing, which allowed for the in-house analysis of unamplified samples. The tile library coding sequences were extracted from the plasmid DNA of each library specimen by restriction enzyme digest. This ensured that all DNA sequences originated directly from the library

sample, eliminating any amplification bias or misassembled sequences from incomplete extension and re-annealing.

The length of tiles used in phage display experiments is influenced by the objectives of the work, with short peptides (less than 10 amino acids) permitting deep profiling to pinpoint regions of interest and longer peptides (over 15 amino acids) providing the ability to investigate more complex interactions due to structures being more physiologically relevant (Fadaie et al., 2023; Jaroszewicz et al., 2022). The decision to design the TeeVax3 library tiles 20 amino acids in length was driven by the aim to characterise the antibody response to immunisation with the TeeVax3 antigen. Despite tiles of this length limiting the diversity of a library, the prospect of structures mimicking native epitopes results in data collected being of greater biological relevance. In addition to length, tile overlap can be manipulated to alter the resolution and increase the coverage of results in phage display investigations. Here, we used an offset of three amino acid residues, with each tile overlapping the previous by 17 amino acids. This could have been extended to a one or two residue offset, however the additional clarity generated from this adjustment would have been of little value given our aim.

It is well established that Nanopore sequencing can have higher rates of error when compared to other next generation approaches (J. Chen & Xu, 2023; Delahaye & Nicolas, 2021). Tile design is one way to mitigate this limitation, which influenced our decision to use tiles 20 amino acids in length with a three-residue offset. Although real time sequencing of long reads is one of the substantial advantages of Nanopore, it is unclear how significant the impact of read length is on accuracy (Frascarelli et al., 2023; T. Hu et al., 2021; T. Zhang et al., 2024). Given this uncertainty, we decided to focus on the offset aspect of our tile design to avoid error rates affecting the correct identification of tiles. Sequencing errors tend to be randomly distributed throughout a read, and with a 17 residue overlap between tiles the probability of reads accumulating enough mutations to prevent correct recognition is low. This reasoning was verified by the correct matching of tiles observed in this work. However, it is important to recognise that our library of 308 tiles is limited in comparison to others in the literature (Kong et al., 2020; Zantow et al., 2016). Large libraries where sequences have a greater likelihood to include

repeated regions or stretches of homopolymer are more susceptible to base calling errors resulting in identical tile sequences at differing parts of an antigen. This can be problematic as there is no way to determine which part of the antigen the two identical sequences are derived from, generating inaccurate binning of tiles and potentially misidentifying epitopes.

The expected affinity of binders and the possibility of background signal needs to be considered when developing a phage panning protocol (Smith & Petrenko, 1997). Here, we increased the number of washing steps per round of panning in order to isolate the epitopes with the strongest affinity and decrease background binding. This may have led to the loss of those more transient interactions and the failure to identify the corresponding epitopes. An alternative strategy to identify non-specific binding would be to include a control trial which investigates the interactions of the antibody pool from unimmunised sera, or an IgG control antibody with the TeeVax3 phage library. Furthermore, carrying out a depletion step during the antibody purification process would be beneficial considering the slightly elevated signal from the ELISA analysis of unimmunised sera with the TeeVax3 library. This would involve exposing TeeVax3 specific polyclonal antibodies to helper phage and collecting the unbound fraction, subsequently removing any antibody clones that are interacting with other phage components as opposed to the displayed peptide.

In addition to the above approaches, the number of panning rounds can be adjusted depending on the selection stringency required for the experiment. Our results showed rapid enrichment in the first round of panning, with little change in subsequent rounds. This differs from other studies which often employ three rounds before sequencing the bound fraction (Zantow et al., 2016; Zhou et al., 2014). Whether this effect is influenced by the small size of our library is to be considered, as other protocols contain libraries of up to 10^{10} peptides (Kong et al., 2020; Zantow et al., 2016). Moreover, it is essential to confirm that the dominating peptides are not due to an expression bias, where the ease of production and display is the cause of an increase in frequency not the peptide's ability to bind with the antibody population. This factor was alleviated in our work by analysing the tile frequencies

in the library before and after expression, alongside the validation of binding affinity using a synthetic peptide series. Given these considerations we would advise two rounds of panning to be sufficient when using Nanopore sequencing, as additional rounds may be unwarranted due to peptides with strong binding affinity dominating the library.

The rapid enrichment we observed after the first round of panning was caused by the strong affinity tiles covering the N terminal epitope tag. This outcome was surprising given the general sentiment that Histidine tags are not immunogenic (López-Laguna et al., 2022; Zhao et al., 2013), despite some instances of them affecting an antigen's ability to generate an immune response (Lin et al., 2022; Singh et al., 2020). We verified the outcome by evaluating the antibody binding ability of the TeeVax3 antigen with and without the Histidine tag, and on further investigation of the sequence causing this result, we identified a 10-residue segment which included the 6-His tag and TEV cleavage site. This indicates the importance of removing Histidine tags, and any other superfluous residues, which presently occurs inconsistently in early stages of vaccine development (Otsyula et al., 2013; M. X. Rodrigues et al., 2020). Exclusion of these high affinity tiles did uncover some regions of immunogenicity on the TeeVax3 antigen, which supports previous work (Raynes et al., 2023) and indicates further investigation of T-antigens as vaccine candidates to be worthwhile.

The production of binding antibodies in response to an antigen is one of the primary measurements in determining a vaccine candidate's ability to protect from infection and disease (Pulendran & Ahmed, 2011; Rueckert & Guzmán, 2012). In this thesis, one of the objectives was to identify where on the TeeVax3 antigen surface the antibodies produced from immunisation were binding. Although this identifies regions that are immunogenic, it does not reveal the mechanisms that lead to the desired goal of protection. While some of the important functions of binding antibodies include neutralisation and tagging for destruction, it could be argued that the T cell responses are just as influential, if not more so, in the development of a successful immune response and resulting protection from disease (King et al., 2024).

Despite our success at implementing a high throughput method for the rapid epitope profiling of TeeVax3, this technique remains limited to evaluating the response of antibodies recognising linear epitopes. The ability to expand phage display techniques to incorporate detection of conformational epitopes is attempted in some studies by displaying peptides of up to 90 amino acids in length (Mohan et al., 2018). However, the applicability of the identified epitopes to the native scenario may be underwhelming when posttranslational modifications and other physiological conditions are considered. Nevertheless, the ability to rapidly map the linear epitope response to antigens remains a valuable tool to establish any bias in the immune response, illustrated by our discovery of epitope tag immunogenicity in this example. This method provides an accessible rapid high throughput option to assist vaccine candidate progress through the development pipeline.

Chapter 5

Conclusions and Future Directions

The continual demand for effective vaccines in the future will require a development pipeline that includes a range of assessable protocols (Cunningham et al., 2016; Gebre et al., 2021; C. M. C. Rodrigues & Plotkin, 2020; Rueckert & Guzmán, 2012). The current repertoire of techniques is lacking in tools that provide both rapid and assessable high throughput evaluation of antigen-antibody interactions (De Leon et al., 2024; Gershoni et al., 2007). The successful implementation of the method described in this thesis, which rapidly profiles linear epitopes in the antibody response to TeeVax3 immunisation, provides a highly efficient approach that can assist in the identification of epitopes and inform future antigen design.

The challenges involved in identifying the correlates of protection for any individual vaccine candidate remain a substantial barrier to the development of new effective vaccines (Cunningham et al., 2016; Plotkin, 2020). Measurement of binding antibodies continues to be one of the primary methods for evaluating this characteristic (Pulendran & Ahmed, 2011), therefore techniques that provide the ability to collect this information remain in high demand. Currently, methods can be broadly categorised into two groups, the high resolution and resource intensive structural tools, such as X-ray crystallography and Cryo-electron microscopy, or the high-throughput lower resolution methods such as peptide microarray and cell display systems (De Leon et al., 2024; Hamed et al., 2023; D. Hu & Irving, 2023). The protocol outlined in this thesis falls into the latter group, with the additional advantage of improved resolution due to considered tile design. Furthermore, the use of Oxford Nanopore sequencing greatly improves the accessibility and speed at which the antibody response to a protein antigen can be evaluated.

The use of phage display for profiling linear epitopes is well established with many reports of large libraries being successfully utilised to investigate protein-protein interactions (Kong et al., 2020; Mohan et al., 2018). Although the size of the library used in this work was limited due to the antigen sequence length and tile design,

application of the protocol with larger libraries should not pose significant difficulty. Both the display of tiled candidate vaccine antigens and proteins from native pathogens ought to result in valuable data for subsequent vaccine development. Further expansion into the monitoring of population immunity and longevity of vaccine efficacy could be achieved by establishing phage display libraries for the antigens of interest and panning them against the antibody repertoire of recovered or vaccinated individuals.

Continued assessment of the T antigens suitability as a vaccine candidate remains a notable focus in the aim to create a Group A Streptococcus vaccine. Extending the work covered in this thesis by applying the technique to the remaining TeeVax candidates, TeeVax1 and TeeVax2, would be the next logical step to confirm the viability of these recombinant proteins as potential vaccine candidates. Ideally, the epitope tag bias observed would be addressed by ensuring the tag is cleaved before animal studies, however, removing the tag sequence from the phage library design may still allow for valuable insight into immunogenic regions on the antigen surfaces. Moreover, utilisation of the native T antigen library created by Raynes et al. could provide an insight into the interactions between the antibody response initiated by TeeVax immunisation and the native Strep A pathogen.

Antibodies recognising conformational epitopes comprise the majority of the antibody population produced in an immune response. Despite this, the investigation of conformational epitopes remains relatively difficult given the tools currently available (Gershoni et al., 2007; Rojas et al., 2014), and the rapid epitope profiling method we have implemented is limited in addressing this obstacle. However, utilising our data with computational tools, such as EpiSearch, it is feasible that prediction of areas on the antigen surface where conformational epitopes exist ought to be possible. As both methods are rapid and high throughput the ability to acquire this information promptly is a substantial advantage when investigating potential vaccine candidates.

Overall, this work has developed a functional tool to rapidly assess the binding antibody response to a protein antigen, while calling attention to the potential ramifications of not removing purification tags during early development of vaccine

candidates. The technique presented here can be applied to numerous antigen-antibody investigations, including antigen identification and evaluation of vaccine candidates, making it a valuable addition to the vaccine development tool kit.

References

Abd El-Aleem, S. A., Saber, E. A., Aziz, N. M., El-Sherif, H., Abdelraof, A. M., & Djouhri, L. (2022). Follicular dendritic cells. *Journal of Cellular Physiology*, *237*(4), 2019–2033. https://doi.org/10.1002/jcp.30662

Abufhele, A., Bravo, D., Lopez-Boo, F., & Soto-Ramirez, P. (2024). Developmental Losses in Young Children from Preprimary Program Closures during the COVID-19 Pandemic. *Comparative Education Review*, *68*(3), 423–442. https://doi.org/10.1086/731588

Ackermann, H.-W. (2003). Bacteriophage observations and evolution. *Research in Microbiology*, *154*(4), 245–251. https://doi.org/10.1016/S0923-2508(03)00067-6

Adler, L. N., Jiang, W., Bhamidipati, K., Millican, M., Macaubas, C., Hung, S., & Mellins, E. D. (2017). The Other Function: Class II-Restricted Antigen Presentation by B Cells. *Frontiers in Immunology*, *8*. https://doi.org/10.3389/fimmu.2017.00319

Ajay Castro, S., & Dorfmueller, H. C. (2023). Update on the development of Group A Streptococcus vaccines. *Npj Vaccines*, *8*(1), 135. https://doi.org/10.1038/s41541-023-00730-x

Ammar, S., Anglemyer, A., Bennett, J., Lees, J., Addidle, M., Morgan, J., DuBray, K., Galloway, Y., Grey, C., & Duff, P. (2024). Post-pandemic increase in invasive group A strep infections in New Zealand. *Journal of Infection and Public Health*, *17*(11), 102545. https://doi.org/10.1016/j.jiph.2024.102545

Antanasijevic, A., Bowman, C. A., Kirchdoerfer, R. N., Cottrell, C. A., Ozorowski, G., Upadhyay, A. A., Cirelli, K. M., Carnathan, D. G., Enemuo, C. A., Sewall, L. M., Nogal, B., Zhao, F., Groschel, B., Schief, W. R., Sok, D., Silvestri, G., Crotty, S., Bosinger, S. E., & Ward, A. B. (2022). From structure to sequence: Antibody discovery using cryoEM. *Science Advances*, *8*(3), eabk2039. https://doi.org/10.1126/sciadv.abk2039

Chen, J., & Xu, F. (2023). Application of Nanopore Sequencing in the Diagnosis and Treatment of Pulmonary Infections. *Molecular Diagnosis & Therapy*, *27*(6), 685–701. https://doi.org/10.1007/s40291-023-00669-8

Chen, Z., Gao, X., & Yu, D. (2022). Longevity of vaccine protection: Immunological mechanism, assessment methods, and improving strategy. *VIEW*, *3*(1), 20200103. https://doi.org/10.1002/VIW.20200103

Cunningham, A. L., Garçon, N., Leo, O., Friedland, L. R., Strugnell, R., Laupèze, B., Doherty, M., & Stern, P. (2016). Vaccine development: From concept to early clinical testing. *Vaccine*, *34*(52), 6655–6664. https://doi.org/10.1016/j.vaccine.2016.10.016

Dale, J. B., & Beachey, E. H. (1985). Epitopes of streptococcal M proteins shared with cardiac myosin. *The Journal of Experimental Medicine*, *162*(2), 583–591. https://doi.org/10.1084/jem.162.2.583

Davies, M. R., McIntyre, L., Mutreja, A., Lacey, J. A., Lees, J. A., Towers, R. J., Duchêne, S., Smeesters, P. R., Frost, H. R., Price, D. J., Holden, M. T. G., David, S., Giffard, P. M., Worthing, K. A., Seale, A. C., Berkley, J. A., Harris, S. R., Rivera-Hernandez, T., Berking, O., … Walker, M. J. (2019). Atlas of group A streptococcal vaccine candidates compiled using large-scale comparative genomics. *Nature Genetics*, *51*(6), 1035–1043. https://doi.org/10.1038/s41588-019-0417-8

De Leon, A. J., Tjiam, M. C., & Yu, Y. (2024). B cell epitope mapping: The journey to better vaccines and therapeutic antibodies. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1868*(10), 130674. https://doi.org/10.1016/j.bbagen.2024.130674

Delahaye, C., & Nicolas, J. (2021). Sequencing DNA with nanopores: Troubles and biases. *PLOS ONE*, *16*(10), e0257521. https://doi.org/10.1371/journal.pone.0257521

Deng, X., Wang, L., You, X., Dai, P., & Zeng, Y. (2017). Advances in the T7 phage display system (Review). *Molecular Medicine Reports*. https://doi.org/10.3892/mmr.2017.7994

Fadaie, M., Dianat-Moghadam, H., Ghafouri, E., Naderi, S., Darvishali, M. H., Ghovvati, M., Khanahmad, H., Boshtam, M., & Makvandi, P. (2023). Unraveling the potential of M13 phages in biomedicine: Advancing drug nanodelivery and gene therapy. *Environmental Research*, *238*, 117132. https://doi.org/10.1016/j.envres.2023.117132

Fan, J., Toth, I., & Stephenson, R. J. (2024). Recent Scientific Advancements towards a Vaccine against Group A Streptococcus. *Vaccines*, *12*(3), Article 3. https://doi.org/10.3390/vaccines12030272

Forsström, B., Bisławska Axnäs, B., Rockberg, J., Danielsson, H., Bohlin, A., & Uhlen, M. (2015). Dissecting Antibodies with Regards to Linear and Conformational Epitopes. *PLOS ONE*, *10*(3), e0121673. https://doi.org/10.1371/journal.pone.0121673

Forthal, D. N. (2014). Functions of Antibodies. *Microbiology Spectrum*, *2*(4), 2.4.21. https://doi.org/10.1128/microbiolspec.AID-0019-2014

Francino-Urdaniz, I. M., & Whitehead, T. A. (2021). An overview of methods for the structural and functional mapping of epitopes recognized by anti-SARS-CoV-2 antibodies. *RSC Chemical Biology*, *2*(6), 1580–1589. https://doi.org/10.1039/D1CB00169H

Frascarelli, C., Zanetti, N., Nasca, A., Izzo, R., Lamperti, C., Lamantea, E., Legati, A., & Ghezzi, D. (2023). Nanopore long-read next-generation sequencing for detection of mitochondrial DNA large-scale deletions. *Frontiers in Genetics*, *14*, 1089956. https://doi.org/10.3389/fgene.2023.1089956

Garrett, M. E., Itell, H. L., Crawford, K. H. D., Basom, R., Bloom, J. D., & Overbaugh, J. (2020). Phage-DMS: A Comprehensive Method for Fine Mapping of Antibody Epitopes. *iScience*, *23*(10), 101622. https://doi.org/10.1016/j.isci.2020.101622

Gebre, M. S., Brito, L. A., Tostanoski, L. H., Edwards, D. K., Carfi, A., & Barouch, D. H. (2021). Novel approaches for vaccine development. *Cell*, *184*(6), 1589–1603. https://doi.org/10.1016/j.cell.2021.02.030

Gershoni, J. M., Roitburd-Berman, A., Siman-Tov, D. D., Tarnovitski Freund, N., & Weiss, Y. (2007). Epitope Mapping: The First Step in Developing Epitope-Based Vaccines. *BioDrugs*, *21*(3), 145–156. https://doi.org/10.2165/00063030-200721030-00002

Hamed, S. M., Sakr, M. M., El-Housseiny, G. S., Wasfi, R., & Aboshanab, K. M. (2023). State of the Art in Epitope Mapping and Opportunities in COVID-19. *Future Science OA*, *9*(1), FSO832. https://doi.org/10.2144/fsoa-2022-0048

Hu, D., & Irving, A. T. (2023). Massively-multiplexed epitope mapping techniques for viral antigen discovery. *Frontiers in Immunology*, *14*, 1192385. https://doi.org/10.3389/fimmu.2023.1192385

Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human Immunology*, *82*(11), 801–811. https://doi.org/10.1016/j.humimm.2021.02.012

Jaroszewicz, W., Morcinek-Orłowska, J., Pierzynowska, K., Gaffke, L., & Węgrzyn, G. (2022). Phage display and other peptide display technologies. *FEMS Microbiology Reviews*, *46*(2), fuab052. https://doi.org/10.1093/femsre/fuab052

Jethva, P. N., & Gross, M. L. (2023). Hydrogen deuterium exchange and other mass spectrometry- based approaches for epitope mapping. *Frontiers in Analytical Science*, *3*, 1118749. https://doi.org/10.3389/frans.2023.1118749

Kapingidza, A. B., Kowal, K., & Chruszcz, M. (2020). Antigen–antibody complexes. In *Vertebrate and Invertebrate Respiratory Proteins, Lipoproteins and other Body Fluid Proteins* (pp. 465–497).

King, D. F., Groves, H., Weller, C., Jones, I., Cramer, J. P., Gilbert, P. B., Goldblatt, D., Gruber, M. F., Kampmann, B., Maïga, D., Pasetti, M. F., Plotkin, S. A., Precioso, A., Wassie, L., Wittke, F., & Kaslow, D. C. (2024). Realising the potential of correlates of protection for vaccine development, licensure and use: Short summary. *Npj Vaccines*, *9*(1), 82. https://doi.org/10.1038/s41541-024-00872-6

Kolahchi, Z., De Domenico, M., Uddin, L. Q., Cauda, V., Grossmann, I., Lacasa, L., Grancini, G., Mahmoudi, M., & Rezaei, N. (2021). COVID-19 and Its Global Economic Impact. In N. Rezaei (Ed.), *Coronavirus Disease—COVID-19* (Vol. 1318, pp. 825–837). Springer International Publishing. https://doi.org/10.1007/978-3-030-63761-3_46

Kong, X.-D., Carle, V., Díaz-Perlas, C., Butler, K., & Heinis, C. (2020). Generation of a Large Peptide Phage Display Library by Self-Ligation of Whole-Plasmid PCR Product. *ACS Chemical Biology*, *15*(11), 2907–2915. https://doi.org/10.1021/acschembio.0c00497

Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H. R., & Plückthun, A. (1997). Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *Journal of Immunological Methods*, *201*(1), 35–55. https://doi.org/10.1016/S0022-1759(96)00208-6

Liao, J. J. Z., & Liu, R. (2009). Re-parameterization of five-parameter logistic function. *Journal of Chemometrics*, *23*(5), 248–253. https://doi.org/10.1002/cem.1218

Lin, T.-W., Huang, P.-H., Liao, B.-H., Chao, T.-L., Tsai, Y.-M., Chang, S.-C., Chang, S.-Y., & Chen, H.-W. (2022). Tag-Free SARS-CoV-2 Receptor Binding Domain (RBD), but Not C-Terminal Tagged SARS-CoV-2 RBD, Induces a Rapid and Potent Neutralizing Antibody Response. *Vaccines*, *10*(11), Article 11. https://doi.org/10.3390/vaccines10111839

Loh, J. M. S., Rivera-Hernandez, T., McGregor, R., Khemlani, A. H. J., Tay, M. L., Cork, A. J., M. Raynes, J., Moreland, N. J., Walker, M. J., & Proft, T. (2021). A multivalent T-antigen-based vaccine for Group A Streptococcus. *Scientific Reports*, *11*(1), 4353. https://doi.org/10.1038/s41598-021-83673-4

López-Laguna, H., Voltà-Durán, E., Parladé, E., Villaverde, A., Vázquez, E., & Unzueta, U. (2022). Insights on the emerging biotechnology of histidine-rich peptides. *Biotechnology Advances*, *54*, 107817. https://doi.org/10.1016/j.biotechadv.2021.107817

Mahdavi, S. Z. B., Oroojalian, F., Eyvazi, S., Hejazi, M., Baradaran, B., Pouladi, N., Tohidkia, M. R., Mokhtarzadeh, A., & Muyldermans, S. (2022). An overview on display systems (phage, bacterial, and yeast display) for production of anticancer antibodies; advantages and disadvantages. *International Journal of Biological Macromolecules*, *208*, 421–442. https://doi.org/10.1016/j.ijbiomac.2022.03.113

Mäntynen, S., Laanto, E., Oksanen, H. M., Poranen, M. M., & Díaz-Muñoz, S. L. (2021). Black box of phage–bacterium interactions: Exploring alternative phage infection strategies. *Open Biology*, *11*(9), 210188. https://doi.org/10.1098/rsob.210188

Massell, B. F. (1969). Rheumatic Fever Following Streptococcal Vaccination: Report of Three Cases. *JAMA*, *207*(6), 1115. https://doi.org/10.1001/jama.1969.03150190037007

Mazrekaj, D., & De Witte, K. (2024). The Impact of School Closures on Learning and Mental Health of Children: Lessons From the COVID-19 Pandemic. *Perspectives on Psychological Science*, *19*(4), 686–693. https://doi.org/10.1177/17456916231181108

Mohan, D., Wansley, D. L., Sie, B. M., Noon, M. S., Baer, A. N., Laserson, U., & Larman, H. B. (2018). PhIP-Seq characterization of serum antibodies using oligonucleotide-

encoded peptidomes. *Nature Protocols*, *13*(9), 1958–1978. https://doi.org/10.1038/s41596-018-0025-6

Negi, S. S., & Braun, W. (2009). Automated Detection of Conformational Epitopes Using Phage Display Peptide Sequences. *Bioinformatics and Biology Insights*, *3*, BBI.S2745. https://doi.org/10.4137/BBI.S2745

Ngampasutadol, J., Rice, P. A., Walsh, M. T., & Gulati, S. (2006). Characterization of a peptide vaccine candidate mimicking an oligosaccharide epitope of *Neisseria gonorrhoeae* and resultant immune responses and function. *Vaccine*, *24*(2), 157–170. https://doi.org/10.1016/j.vaccine.2005.07.065

Otsyula, N., Angov, E., Bergmann-Leitner, E., Koech, M., Khan, F., Bennett, J., Otieno, L., Cummings, J., Andagalu, B., Tosh, D., Waitumbi, J., Richie, N., Shi, M., Miller, L., Otieno, W., Otieno, G. A., Ware, L., House, B., Godeaux, O., … Spring, M. D. (2013). Results from tandem Phase 1 studies evaluating the safety, reactogenicity and immunogenicity of the vaccine candidate antigen Plasmodium falciparum FVO merozoite surface protein-1 (MSP142) administered intramuscularly with adjuvant system AS01. *Malaria Journal*, *12*(1), 29. https://doi.org/10.1186/1475-2875-12-29

Plotkin, S. A. (2020). Updates on immunologic correlates of vaccine-induced protection. *Vaccine*, *38*(9), 2250–2257. https://doi.org/10.1016/j.vaccine.2019.10.046

Pulendran, B., & Ahmed, R. (2011). Immunological mechanisms of vaccination. *Nature Immunology*, *12*(6), 509–517. https://doi.org/10.1038/ni.2039

Qi, H., Ma, M., Hu, C., Xu, Z., Wu, F., Wang, N., Lai, D., Li, Y., Zhang, H., Jiang, H., Meng, Q., Guo, S., Kang, Y., Zhao, X., Li, H., & Tao, S. (2021). Antibody Binding Epitope Mapping (AbMap) of Hundred Antibodies in a Single Run. *Molecular & Cellular Proteomics*, *20*, 100059. https://doi.org/10.1074/mcp.RA120.002314

Rastogi, I., Jeon, D., Moseman, J. E., Muralidhar, A., Potluri, H. K., & McNeel, D. G. (2022). Role of B cells as antigen presenting cells. *Frontiers in Immunology*, *13*, 954936. https://doi.org/10.3389/fimmu.2022.954936

Raynes, J. M., Young, P. G., Lorenz, N., Loh, J. M. S., McGregor, R., Baker, E. N., Proft, T., & Moreland, N. J. (2023). Identification of an immunodominant region on a group A Streptococcus T-antigen reveals temperature-dependent motion in pili. *Virulence*, *14*(1), 2180228. https://doi.org/10.1080/21505594.2023.2180228

Robertson, N., Lopez-Anton, N., Gurjar, S. A., Khalique, H., Khalaf, Z., Clerkin, S., Leydon, V. R., Parker-Manuel, R., Raeside, A., Payne, T., Jones, T. D., Seymour, L., & Cawood, R. (2020). Development of a novel mammalian display system for selection of antibodies against membrane proteins. *Journal of Biological Chemistry*, *295*(52), 18436–18448. https://doi.org/10.1074/jbc.RA120.015053

Rodrigues, C. M. C., & Plotkin, S. A. (2020). Impact of Vaccines; Health, Economic and Social Perspectives. *Frontiers in Microbiology*, *11*, 1526. https://doi.org/10.3389/fmicb.2020.01526

Rodrigues, M. X., Yang, Y., de Souza Meira, E. B., do Carmo Silva, J., & Bicalho, R. C. (2020). Development and evaluation of a new recombinant protein vaccine (YidR) against *Klebsiella pneumoniae* infection. *Vaccine*, *38*(29), 4640–4648. https://doi.org/10.1016/j.vaccine.2020.03.057

Rojas, G., Tundidor, Y., & Infante, Y. C. (2014). High throughput functional epitope mapping: Revisiting phage display platform to scan target antigen surface. *mAbs*, *6*(6), 1368–1376. https://doi.org/10.4161/mabs.36144

Rueckert, C., & Guzmán, C. A. (2012). Vaccines: From Empirical Development to Rational Design. *PLoS Pathogens*, *8*(11), e1003001. https://doi.org/10.1371/journal.ppat.1003001

Singh, M., Sori, H., Ahuja, R., Meena, J., Sehgal, D., & Panda, A. K. (2020). Effect of N-terminal poly histidine-tag on immunogenicity of *Streptococcus pneumoniae* surface protein SP0845. *International Journal of Biological Macromolecules*, *163*, 1240–1248. https://doi.org/10.1016/j.ijbiomac.2020.07.056

Smeesters, P. R., De Crombrugghe, G., Tsoi, S. K., Leclercq, C., Baker, C., Osowicki, J., Verhoeven, C., Botteaux, A., & Steer, A. C. (2024). Global Streptococcus pyogenes strain diversity, disease associations, and implications for vaccine development: A systematic review. *The Lancet Microbe*, *5*(2), e181–e193. https://doi.org/10.1016/S2666-5247(23)00318-X

Smith, G. P., & Petrenko, V. A. (1997). Phage Display. *Chemical Reviews*, *97*(2), 391–410. https://doi.org/10.1021/cr960065d

Smith, G. P., & Scott, J. K. (1993). Libraries of peptides and proteins displayed on filamentous phage. In *Methods in Enzymology* (Vol. 217, pp. 228–257). Academic Press. https://doi.org/10.1016/0076-6879(93)17065-D

Szabó, G. T., Mahiny, A. J., & Vlatkovic, I. (2022). COVID-19 mRNA vaccines: Platforms and current developments. *Molecular Therapy*, *30*(5), 1850–1868. https://doi.org/10.1016/j.ymthe.2022.02.016

Taft, J. M., Weber, C. R., Gao, B., Ehling, R. A., Han, J., Frei, L., Metcalfe, S. W., Overath, M. D., Yermanos, A., Kelton, W., & Reddy, S. T. (2022). Deep mutational learning predicts ACE2 binding and antibody escape to combinatorial mutations in the SARS-CoV-2 receptor-binding domain. *Cell*, *185*(21), 4008-4022.e14. https://doi.org/10.1016/j.cell.2022.08.024

Vekemans, J., Gouvea-Reis, F., Kim, J. H., Excler, J.-L., Smeesters, P. R., O'Brien, K. L., Van Beneden, C. A., Steer, A. C., Carapetis, J. R., & Kaslow, D. C. (2019). The Path to Group A Streptococcus Vaccines: World Health Organization Research and Development Technology Roadmap and Preferred Product Characteristics. *Clinical Infectious Diseases*, *69*(5), 877–883. https://doi.org/10.1093/cid/ciy1143

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

Walkinshaw, D. R., Wright, M. E. E., Mullin, A. E., Excler, J.-L., Kim, J. H., & Steer, A. C. (2023). The Streptococcus pyogenes vaccine landscape. *Npj Vaccines*, *8*(1), Article 1. https://doi.org/10.1038/s41541-023-00609-x

Wang, H., Wu, X., Zhang, X., Hou, X., Liang, T., Wang, D., Teng, F., Dai, J., Duan, H., Guo, S., Li, Y., & Yu, X. (2020). SARS-CoV-2 Proteome Microarray for Mapping COVID-19 Antibody Interactions at Amino Acid Resolution. *ACS Central Science*, *6*(12), 2238–2249. https://doi.org/10.1021/acscentsci.0c00742

Xu, G. J., Kula, T., Xu, Q., Li, M. Z., Vernon, S. D., Ndung'u, T., Ruxrungtham, K., Sanchez, J., Brander, C., Chung, R. T., O'Connor, K. C., Walker, B., Larman, H. B., & Elledge, S. J. (2015). Viral immunology. Comprehensive serological profiling of human populations using a synthetic human virome. *Science (New York, N.Y.)*, *348*(6239), aaa0698. https://doi.org/10.1126/science.aaa0698

Zantow, J., Just, S., Lagkouvardos, I., Kisling, S., Dübel, S., Lepage, P., Clavel, T., & Hust, M. (2016). Mining gut microbiome oligopeptides by functional metaproteome display. *Scientific Reports*, *6*(1), 34337. https://doi.org/10.1038/srep34337

Zhang, L., Rohr, J., Cui, R., Xin, Y., Han, L., Yang, X., Gu, S., Du, Y., Liang, J., Wang, X., Wu, Z., Hao, Q., & Liu, X. (2022). Biological invasions facilitate zoonotic disease emergences. *Nature Communications*, *13*(1), 1762. https://doi.org/10.1038/s41467-022-29378-2

Zhang, T., Li, H., Jiang, M., Hou, H., Gao, Y., Li, Y., Wang, F., Wang, J., Peng, K., & Liu, Y.-X. (2024). Nanopore sequencing: Flourishing in its teenage years. *Journal of Genetics and Genomics*, *51*(12), 1361–1374. https://doi.org/10.1016/j.jgg.2024.09.007

Zhao, X., Li, G., & Liang, S. (2013). Several Affinity Tags Commonly Used in Chromatographic Purification. *Journal of Analytical Methods in Chemistry*, *2013*, 1–8. https://doi.org/10.1155/2013/581093

Zhou, W., Lu, S., Su, Y., Xue, D., Yu, X., Wang, S., Zhang, H., Xu, P., Xie, X., & Liu, R. (2014). Decreasing oxidative stress and neuroinflammation with a multifunctional peptide rescues memory deficits in mice with Alzheimer disease. *Free Radical Biology and Medicine*, *74*, 50–63. https://doi.org/10.1016/j.freeradbiomed.2014.06.013

Appendices

Appendix A: Creation of phage libraries

A1: Adaptor sequences

5’ adaptor CGAGGGCCCAGCCGGCCATGGCCGAGGGT

3’ adaptor TCGGCCTCGGGGGCCA

A2: Control Peptide sequences

PEP1 cyclic peptide	TGCGGTCCGATCCCAGTGCTGGATGAAAACGGCCTGTTTGCTCCGGGTCCGTGC
6-His epitopes	CATCACCATCACCATCACGACTACAAAGACGATGACGACAAG

A3: PCR primers for peptide amplification

Reverse primer for amplifying peptides for phage display	TGGCCCCCGAGGCC
Forward primer for amplifying peptides for phage display	CGAGGGCCCAGCCGG

A4: Primers for colony PCR

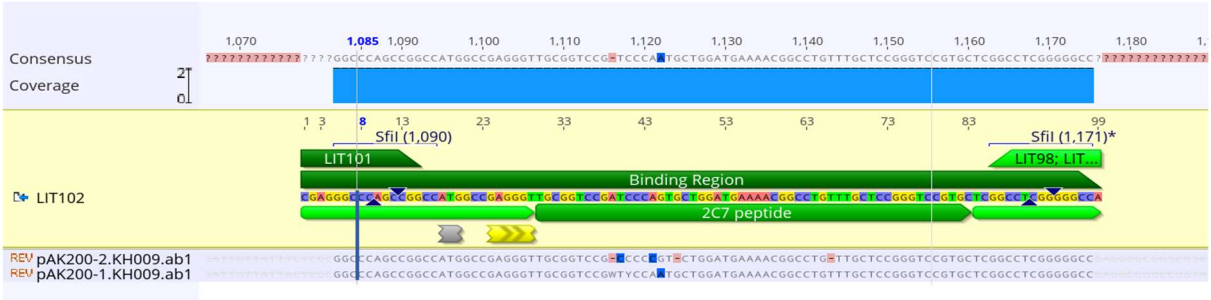
pAK200-pIII, reverse	TGCCATCTTTTCATAATCAAAATCACC
M13pUCF (Lac promoter), forward	AGCGGATAACAATTTCACACAGG

A5: Amino acid sequence of the TeeVax3 antigen. The underlined sequence is the N-terminal epitope tag

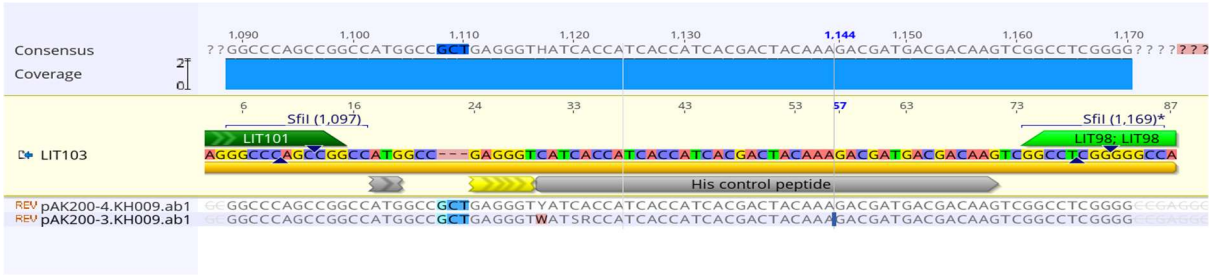
>TeeVax3	MSYYHHHHHHDYDIPTTENLYFQGAMGSPSITKKVTGTIDDVNKKTTSLGSVLSYSLTFEL PSYTKEAVNKTVYVSDNMSEGLTFNFNSLTVEWKGKMANITEDGSVMVENTKIGIAKEVN NGFNLSFIYDSLESISPNISYKAVVNNKAIVGEEGNPNKAEFFYSNNPTKGNTYDNLDKKPD KGNGITSKEDSKIVYTYRSGSGSGKNTTVNEPKVDKDVTKLGKDDDTYQIGDKITWFLKST VPSNIKTLDKFGFTDTLNKGLSFIGDKTQTVTKVQFGTTVLSPDTDYTVEILDSKLTVSLTS AGIEKVSGLVASKQLITEAEKLYKAEDNTDEAAFLSVEVNAKLNADAVMGSRIENDVELD YGHESDIYKSKVPTNEVPEVHTRSGSGSGPKPGKDVKELGLNHSSYNIGERFSWFLKGTVP KNMLDYEKYSFTDTLDSQLDFISVKSVKYGSQILEKNNDYTLSEPTAQNRTLKVELTEAGI KKVAGLYPDRQEVLDTEIEAIKENTDQKPFLEVEFETNINSTVILGKPVTNEVKIEFDNKPD KIAKPVTTPPSDNPEVHTRSGSGSGPTISKSITKSTKDGDKDTASVGEKVDYKLTVQLPSYS KDAINKTVFITDKLSQGLTFLPKSLKIIWNGQTLTKVNEEFKAGDKVIAQLKVENNGFNLN FNYDNLDNHAPEVNYSALLNENAVVGKGGNDNNVDYYYSNNPNKGETHKTTEKPKEGE GTGITKKTDKKTVYTYRSGSGSGESSHKTDVVIHKIKMTSLKGWPKEKNPDGTYTGLGDK NYNGEKIDTITSYFGEGAEELDGVSFTYWSVDKEKYKKLTKNPQNYDTVPKMKAFLQGTE KNKALENSSETIDGKTTGHTADKGGVKVKDLADGYYWFVENSGSNIANGETLSSSAAVPF GLELPVYKADGSTITELHVYPKNTTTKRS*

A6: Alignment of control peptides Sanger sequencing to expected sequence.

PEP-1 control peptide sequence alignment.



6XHis control peptide sequence alignment.



A7: Alignment of TeeVax3 tile Sanger sequencing to expected sequence.

Appendix B: Bioinformatics and data analysis

B1: Processed data and scripts

https://github.com/ImmunoLab/2025-peptide_tiling

Appendix C: High affinity peptide tag exploration

C1: Synthetic biotinylated peptide sequences

Name	Sequence
17-mer	YHHHHHHDYDIPTTENL
14-mer	YHHHHHHDYDIPTT
13-mer	HHHHHHDYDIPT
10-mer	HHHHHHDYD
7-mer	YHHHHHH