



**MONASH** University

# **New malaria vaccines based on merozoite surface protein 2**

**Jeffrey Tien-Jie Seow**

*Bachelor of Pharmaceutical Science (Honours) Monash University, Australia*

A thesis submitted for the degree of Doctor of Philosophy at  
Monash University in 2018

Medicinal Chemistry  
Monash Institute of Pharmaceutical Sciences

## **Copyright notice**

© The author (2018).

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

## Table of contents

<b>Thesis abstract</b> .....	4
<b>Publications</b> .....	6
<b>General declaration</b> .....	7
<b>Acknowledgments</b> .....	10
<b>Abbreviations</b> .....	11
<b>Chapter 1: Introduction</b> .....	13
1.1 Malaria .....	14
1.2 Merozoite surface protein 2 .....	16
1.3 Merozoite surface protein 2 as a vaccine candidate.....	18
1.4 Anti-MSP2 mouse monoclonal antibodies .....	19
1.5 N-terminal epitopes.....	19
1.6 C-terminal epitopes .....	20
1.7 Antibody fragments .....	21
1.8 Peptide vaccines.....	22
1.9 Structure-based vaccine design.....	24
1.10 Project scope and aims.....	26
<b>Chapter 2: Cloning, expression and structural characterisation of antibody fragments against MSP2</b> .....	28
2.1 Introduction.....	29
2.2 Materials and methods .....	31
2.3 Results.....	39
2.4 Discussion.....	50
2.5 Conclusions.....	51
<b>Chapter 3: Structural basis of epitope masking an strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate</b> .....	52
3.1 Chapter introduction .....	53
3.2 Journal article.....	54
3.3 Supplementary material .....	64
<b>Chapter 4: Structure and characterisation of a key epitope in the conserved C-terminal domain of the malaria vaccine candidate MSP2</b> .....	81
4.1 Chapter introduction .....	82
4.2 Journal article.....	83
4.3 Supplementary material .....	94

<b>Chapter 5: Guiding the immune response of MSP2, and intrinsically disordered vaccine candidate.....</b>	<b>105</b>
5.1 Chapter introduction .....	106
5.2 Abstract.....	108
5.3 Introduction.....	109
5.4 Results.....	110
5.5 Discussion.....	119
5.6 Materials and methods .....	121
5.7 Supplementary material .....	125
<b>Chapter 6: Conclusions and future directions .....</b>	<b>132</b>
<b>References.....</b>	<b>138</b>
<b>Appendix I: Disordered epitopes as peptide vaccines .....</b>	<b>146</b>
<b>Appendix II: Lipid interactions modulate the structural and antigenic properties of the C-terminal domain of the malaria antigen merozoite surface protein 2 .....</b>	<b>158</b>

## Thesis abstract

Malaria is a deadly disease responsible for over 430,000 deaths per year. Highly effective therapeutics and increased efforts in vector control have seen a steady decrease in disease prevalence over the past decades. However, to reach the ultimate goal of global eradication of malaria, a robust and broadly protective vaccine is sorely needed.

The target antigen and focus of this thesis is merozoite surface protein 2 (MSP2), an intrinsically disordered protein that is abundant on the surface of the parasite. The polymorphisms in MSP2 are confined to a central variable region that characterises the protein into two allelic families, 3D7 and FC27. Flanking this region are a highly conserved N- and C-terminal regions. Vaccine trials with recombinant MSP2 resulted in a reduction in parasite density, although this was strain-specific. If the immune response can be biased towards conserved regions of MSP2, problems with strain-specific protection may be circumvented.

In this work, the emerging approach of structure-based vaccine design is applied to MSP2. The rational design of antigens based on their native conformation may provide a path forward where traditional vaccine methods have failed. To achieve this, a panel of monoclonal antibodies (mAb) generated in mice immunised with recombinant MSP2 was utilised. Acquiring structural information on antibody-bound MSP2 should provide insights into the conformation of native MSP2 on the merozoite surface and inform the design of a better malaria vaccine.

The N-terminal region of MSP2 is known to adopt an  $\alpha$ -helical conformation in the presence of lipid and it is likely that this conformation is present on the parasite surface. The mAb 6D8 recognises an epitope in this region, but is unable to bind to native MSP2 on merozoites. A crystal structure of 6D8 Fv bound to its minimal binding epitope showed that the structure of the antibody-bound epitope was incompatible with the lipid-bound conformation. Additional crystal structures and NMR spectroscopy also revealed that transient interactions mediated by residues beyond the conserved region of the 6D8 epitope were able to subtly modulate epitope binding.

The C-terminal region of MSP2 is recognised by mAbs that are able to bind to parasite MSP2. One of these mAbs, 4D11, was crystallised with its minimal binding epitope. During the process of crystallisation, this peptide was found to form a homo-dimer with both peptides able to bind 4D11. The structure was stabilised by multiple intramolecular hydrogen bonds and presented a template for further optimisation as a peptide vaccine. This was achieved by designing a series of dimeric, linear and backbone-cyclised peptides, conjugating them to a

carrier protein, keyhole limpet haemocyanin (KLH), and immunising mice. Each peptide was able to induce high titres of antibody specific to the 4D11 epitope. Moreover, the specificity of the native peptide sequence could be biased towards the more accessible 4D11 epitope with a single amino acid point mutation.

In summary, high-resolution crystal structures of antibody-bound epitopes within the conserved regions of MSP2 were solved and described in this study. These structures gave us a better understanding of possible mechanisms of immune escape employed by the parasite and the effects of disorder on antibody recognition. Additionally, the 4D11-bound epitope served as an excellent template for the rational structure-based design of several peptide vaccine candidates. The findings of this study represent an exciting and novel approach towards a better MSP2-based malaria vaccine and may have broader implications for the design of other disordered vaccine candidates.

## Publications

Morales, R. A. V., MacRaid, C. A., **Seow, J.**, Krishnarjuna, B., Drinkwater, N., Rouet, R., Anders, R. F., Christ, D., McGowan, S., Norton, R. S. Structural basis for epitope masking and strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate. *Scientific Reports*. 5, 10103 (2015)

**Seow, J.**, Morales, R. A. V., MacRaid, C. A., Krishnarjuna, B., McGowan, S., Dingjan, T., Jaipuria, G., Rouet, R., Wilde, K. L., Atreya, H. S., Richards, J. S., Anders, R. F., Christ, D., Drinkwater, N., Norton, R. S. Structure and characterisation of a key epitope in the conserved C-terminal domain of the malaria vaccine candidate MSP2. *Journal of Molecular Biology*. 429, 836-846 (2017)

Das, S. C., Morales, R. A. V., **Seow, J.**, Krishnarjuna, B., Dissanayake, R., Anders, R. F., MacRaid, C. A., Norton, R. S. Lipid interactions modulate the structural and antigenic properties of the C-terminal domain of the malaria antigen merozoite surface protein 2. *FEBS J*. 284, 2649–2662 (2017).

MacRaid, C. A., **Seow, J.**, Das, S. C., Norton, R. S. Disordered epitopes as peptide vaccines. *Peptide Science*. In press

Krishnarjuna, B., Suguki, T., Morales R. A. V., **Seow, J.**, Fujiwara, T., Wilde, K. L., Norton, R. S., MacRaid, C. A. Finding specificity in unexpected places: how transient interactions mediate the strain-specific recognition of a conserved malaria epitope. *Communications Biology*. In press

## **General declaration**

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes two original papers published in peer reviewed journals and one prepared manuscript. The core theme of the thesis is the design of new malaria vaccines based on merozoite surface protein 2. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the Medicinal Chemistry Theme of the Monash Institute of Pharmaceutical Science under the supervision of Prof. Raymond S. Norton and Dr. Chris A. MacRaid.

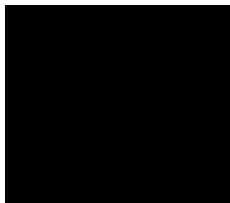
In the case of Chapters 3, 4 and 5 my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution*	Monash student
3	Structural basis for epitope masking and strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate	Published	25%. Designed experiments, prepared samples (synthesis and expression), Solved crystal structures, analysed data and manuscript preparation	1) Rodrigo Morales; Designed experiments, prepared samples, manuscript preparation, 35% 2) Chris MacRaild; Designed experiments, NMR analyses, manuscript preparation, 25% 3) Krishna Bankala; NMR analyses, manuscript preparation, 2% 4) Nyssa Drinkwater; X-ray crystallography 2.5% 5) Romain Rouet ; prepared samples 1% 6) Robin Anders ; manuscript preparation 1% 7) Daniel Christ ; prepared samples, manuscript preparation, 1% 8) Sheena McGowan; X-ray crystallography, 2.5% 9) Ray Norton ; manuscript preparation, 5%	No No No No No No No No No
4	Structure and characterisation of a key epitope in the conserved C-terminal domain of the malaria vaccine candidate MSP2	Published	70% Designed experiments, prepared samples(synthesis and expression), solved crystal structures, analysed data and manuscript preparation	1) Rodrigo Morales; Designed experiments, prepared samples, manuscript preparation, 5% 2) Chris MacRaild; Designed experiments, NMR analyses, manuscript preparation, 5% 3) Krishna Bankala; NMR analyses, manuscript preparation, 2.5% 4) Sheena McGowan; X-ray crystallography, 1% 5) Tamir Dingjan; MD simulations, 2% 6) Garima Jaipuria; NMR analyses, 1% 7) Romain Rouet ; prepared samples 1% 8) Karyn Wilde; sample preparation 1% 9) Hanudatta Atreya; NMR analyses, 1% 10) Jack Richards; manuscript preparation 1% 11) Robin Anders ; manuscript preparation 1% 12) Daniel Christ ; prepared samples, manuscript preparation, 1% 13) Nyssa Drinkwater; X-ray crystallography 2.5% 14) Ray Norton ; manuscript preparation, 5%	No No No No Yes No No No No No No No No No

5	Guiding the immune response of MSP2, an intrinsically disordered malaria vaccine candidate	Prepared Manuscript	75% Designed experiments, prepared samples (synthesis), analysed data and manuscript preparation	1) Chris MacRaid; Designed experiments, manuscript preparation, 5%	No
				2) Sreedam Das; ELISA experiments, manuscript preparation, 2%	Yes
				3) Rodrigo Morales; Designed experiments, prepared samples, manuscript preparation, 4%	No
				4) Vashti Irani; Merozoite ELISA experiments 2%	No
				5) Bankala Krishnarjuna; manuscript preparation 1%	No
				6) Mitchell Silk; MD simulations, 2%	Yes
				7) David Chalmers; MD simulations, 1%	No
				8) Kaye Whycherley ; prepared samples 1%	No
				9) Jack Richards; manuscript preparation 1%	No
				10) Robin Anders ; manuscript preparation 1%	No
				11) Ray Norton ; manuscript preparation, 5%	No

I have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

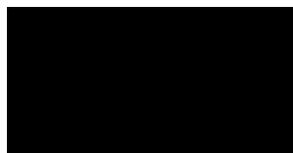
**Student signature:**



**Date:** 26/3/18

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

**Main Supervisor signature:**



**Date:** 26/3/18

## Acknowledgements

Firstly I would like to thank my supervisor Prof. Ray Norton for his guidance and support throughout my candidature. He has been an excellent leader and role model with advice that has always been constructive and valuable. I can't thank Ray enough for giving me the opportunity to work in his lab, from when he first offered me a summer scholarship that led to an Honours project, and ultimately a PhD. His excellent taste in football club is also an added bonus.

I am indebted to Dr. Chris MacRaid for his help with experimental design and expertise in NMR. His suggestions and constant encouragement have been invaluable and pushed me through many difficulties. I would also like to express my gratitude to Dr. Rodrigo Morales, for his friendship and patience in teaching me everything I needed to know in the lab.

I am extremely grateful to Drs. Nyssa Drinkwater and Sheena McGowan for their mentorship in X-ray crystallography, as well as allowing me to utilise their time and resources. I would also like to thank everyone on the MSP2 project including Prof. Robin Anders for his huge wealth of knowledge in everything malaria and Drs. Jack Richards, Vashti Irani and Kaye Wycherley for their advice and support with animal immunisations.

Thanks to my fellow lab members and colleagues, Krishna Bankala, Sreedam Das, Eleanor Leung, Tamir Dingjan, Trayder Thomas, Andrew Tang, Anthony Lai, Mitchell Silk, Michela Mitchell, Maiada Hassan, Mansura Akter, Cael Debono, Shane Devine, Stephen Drane, Steven Yap and Tony Wang for keeping me sane and sharing the highs and lows of research with me. A special thanks to Joanne Du for always being around to help me vent my problems and endlessly try and explain my project to.

Lastly, I have to thank my family, for their endless support and most importantly, for keeping me alive throughout my many years of study. Without them, this whole thing would not have been possible.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

## Abbreviations

ADCI	Antibody-dependent cellular inhibition
ABTS	2,2-azinobis(3-ethylbenzthiazolinesulfonic acid
Fmoc	9-fluorenylmethoxycarbonyl
ACN	Acetonitrile
AMP	Ampicillin
Fab	Antigen-binding fragment
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
cDNA	Complementary deoxyribonucleic acid
cryo-EM	Cryo-electron microscopy
DCM	Dichloromethane
DMB	Dimethylbenzene
DMF	Dimethylformamide
DTT	Dithiothreitol
ELISA	Enzyme-linked immunosorbent assay
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
fHbp	Factor H-binding protein
FPLC	Fast protein liquid chromatography
GPI	Glycophosphatidylinositol
GIA	Growth inhibition assay
HPLC	High-performance liquid chromatography
HRP	Horseradish peroxidase
IFA	Immunofluorescence assays
IgG	Immunoglobulin G
IDP	Intrinsically disordered protein
IPTG	Isopropyl- $\beta$ -D-thiogalactopyranoside
KLH	Keyhole limpet hemocyanin
LB	Luria broth
MSP1	Merozoite surface protein 1
MSP2	Merozoite surface protein 2
Mmt	Methoxytrityl
MD	Molecular dynamics
mAbs	Monoclonal antibodies
DIPEA	N,N-diisopropylethylamine
NCS	N-chlorosuccinimide
NTA	Ni-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
HCTU	O-(1H-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium Hexafluorophosphate
OD	Optical density
PNG	Papua New Guinea
PS3	Peptide synthesiser 3

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein database
RSV	Respiratory syncytial virus
RES	Ring-infected erythrocyte surface antigen
RT	Room temperature
RSV F	RSV fusion
scFv	Single-chain variable fragment
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
s.d.	Standard deviations
SPR	Surface plasmon resonance
TET	Tetracycline
THDX	Thioredoxin
TEV	Tobacco etch virus
TEA	Triethylamine
TFA	Trifluoroacetic acid
TIPS	Triisopropylsilane
Fv	Variable fragment
V <sub>H</sub>	Variable heavy chain
V <sub>L</sub>	Variable light chain

# Chapter 1

## Introduction

## 1.1 Malaria

Malaria is a mosquito-borne infectious disease caused by protozoa of the *Plasmodium* genus, which is responsible for over 200 million cases and 430,000 deaths per year (1, 2). Most of these deaths occur in Sub-Saharan African and South East Asian countries, where younger children and pregnant women are at the highest risk. The social and economic burdens of the disease are exacerbated by ineffective vector control and limited access to treatments and prevention, with the disease often associated with poverty (3, 4). Although the prevalence of the disease has been decreasing over the last few decades, the goal of worldwide eradication of malaria still requires considerable work.

Of the many species of *Plasmodium*, five are known to commonly infect humans: *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale* and *P. knowlesi* (5, 6). The disease varies in virulence and lethality among these species, with the most prevalent and deadly, *P. falciparum*, accounting for more than 90% of human mortalities (7). The clinical symptoms of malaria typically occur between 10 and 15 days after a bite from an infected mosquito (8). These symptoms can be debilitating and may include anaemia, fever, joint pains, nausea and headaches. More severe cases of malarial infection can be fatal, such as in cerebral malaria, where parasite infected red blood cells sequester in the brain microcirculation, causing swelling in the brain that can lead to seizures, coma and death (9).

Early strategies to control malaria, such as vector control and drug therapies, have seen great success in reducing the burden of malaria in some regions (2, 10). In more developed countries these measures have been sufficient to eliminate the disease altogether. Unfortunately, in areas of intense malaria transmission, relying on these strategies alone will not be sufficient to achieve global eradication of malaria. The successful global eradication of smallpox, and significant reduction in occurrences of polio and measles have shown that vaccines are immensely powerful tools against infection. However, the complex biology of the *Plasmodium* parasite, exemplified by its extensive antigenic polymorphism and diverse immune evasion strategies, has complicated the development of an effective malaria vaccine (5, 8, 10–12). A robust malaria vaccine, able to generate antibodies that block invasion or induce phagocytosis of infected cells, would limit parasite replication and greatly reduce the severity of clinical illness, giving patients time to build up their own natural immunity. The most advanced malaria vaccine so far, RTS,S/AS01, has progressed through Phase III clinical trials and is currently undergoing further evaluation in Phase IV trials before wider roll-out. However, these trials showed that the vaccine had only modest efficacy, with a 30% reduction

in severe malaria cases in children aged 5-17 months (13). Additionally, protection conferred by the vaccine diminished rapidly over time, with no efficacy observed four years after immunisations (14). These results show there is still much room for improvement and highlight the urgent need for more effective second generation malaria vaccines.

The *Plasmodium* parasite has a complex life cycle involving an insect vector and vertebrate host, both of which are essential for parasite survival (15). Infection begins with the blood meal of an infected female *Anopheles* mosquito, whereupon sporozoites in the mosquito's salivary glands are injected into the host circulation (**Figure 1**). These sporozoites localise in the liver and invade hepatocytes, where they mature into merozoites before release into the bloodstream. This initiates a cycle of erythrocyte invasion by the merozoite known as the blood stage. It is at this point where the symptoms of malaria arise (6). The blood stage is marked by the rapid invasion of erythrocytes by merozoites, followed by the asexual replication of merozoites. Upon rupture of infected erythrocytes, a new load of merozoites is released into circulation, ready to infect more healthy red blood cells and continue the cycle (11). Some merozoites will leave this cycle and develop into sexual forms of the parasite known as gametocytes (8). When taken up by the mosquito during a blood meal, gametocytes undergo sexual recombination and develop into sporozoites, continuing the cycle. Disruption of the *Plasmodium* life cycle by drug therapies or vaccination is crucial, not only for limiting the damage to the host, but also for controlling the dissemination of parasites in the population.

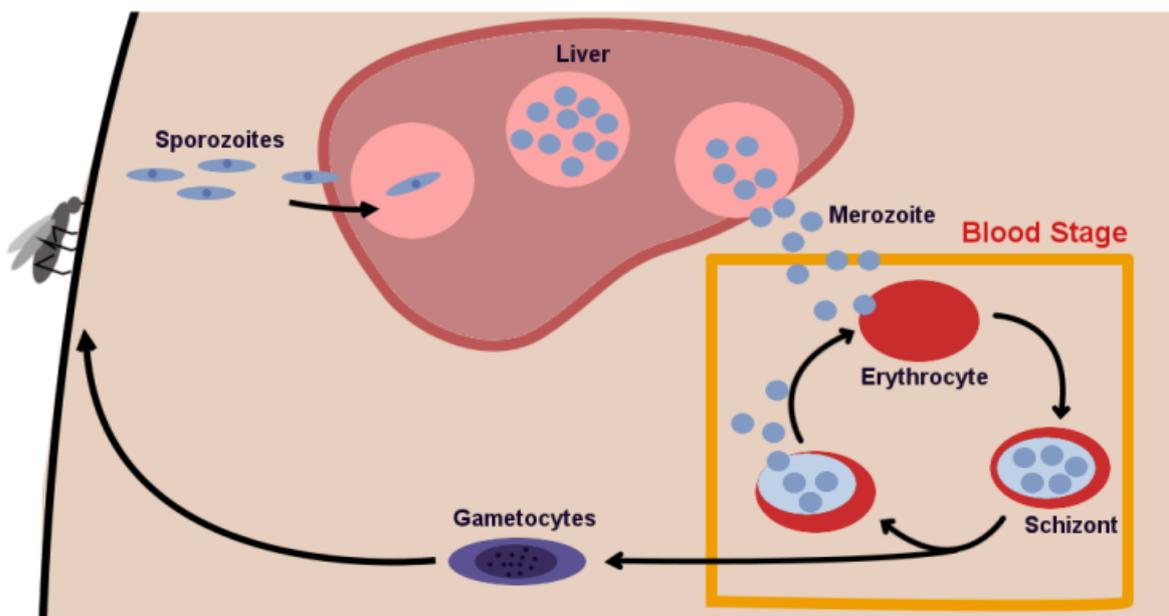


Figure 1: Life cycle of *Plasmodium falciparum* in the human host

## 1.2 Merozoite surface protein 2

The complexity of the *Plasmodium* life cycle, coupled with its ability to evade the host immune system, makes it likely that a successful malaria vaccine will require multiple components, targeting different life stages of the parasite. Pre-erythrocyte vaccines will reduce the chances of infection by clearing the parasite before they can infect erythrocytes. Transmission blocking vaccines, that target the sexual, sporogonic, or mosquito stages of the parasite, will stop the spread of malaria by hindering the parasite's ability to reproduce. Vaccines successfully targeting the asexual blood stage will inhibit parasite replication, ideally clearing infection, or at the very least reduce the severity of symptoms and the risk of death. In recent years, interest in blood-stage vaccines has waned, in part owing to the focus on development of pre-erythrocyte vaccines such as RTS,S. Despite this, the ability of blood-stage vaccines to prevent or reduce the severity of clinical illness dictates that they will be invaluable in reducing the burden of malaria. This thesis details work on merozoite surface protein 2 (MSP2), a blood-stage antigen unique to *Plasmodium falciparum* and found in high abundance on the merozoite surface (16).

This ~23 kDa protein is almost entirely intrinsically disordered (17) and anchored C-terminally to the surface of the merozoite by a glycosylphosphatidylinositol (GPI) moiety (18–20). The precise role of MSP2 in *P. falciparum* is unknown, but knockout of the MSP2 gene results in non-viable parasites (18), suggesting that MSP2 is essential for merozoite infection of the erythrocyte. The protein is present on the merozoite surface coat during invasion and quickly degrades after entering the erythrocyte (21). All MSP2 proteins can be grouped into two allelic families, 3D7 and FC27 (22–24), that differ in a central variable region, flanked by highly conserved N- and C-terminal regions (**Figure 2**). The central variable region constitutes up to 60% of the protein, with the 3D7 allele characterised by short GSA repeats (25), whilst the FC27 allele carries longer tandem repeats of 32- and 12-residue sequences. The protein contains a single disulfide bond within the conserved C-terminal region.

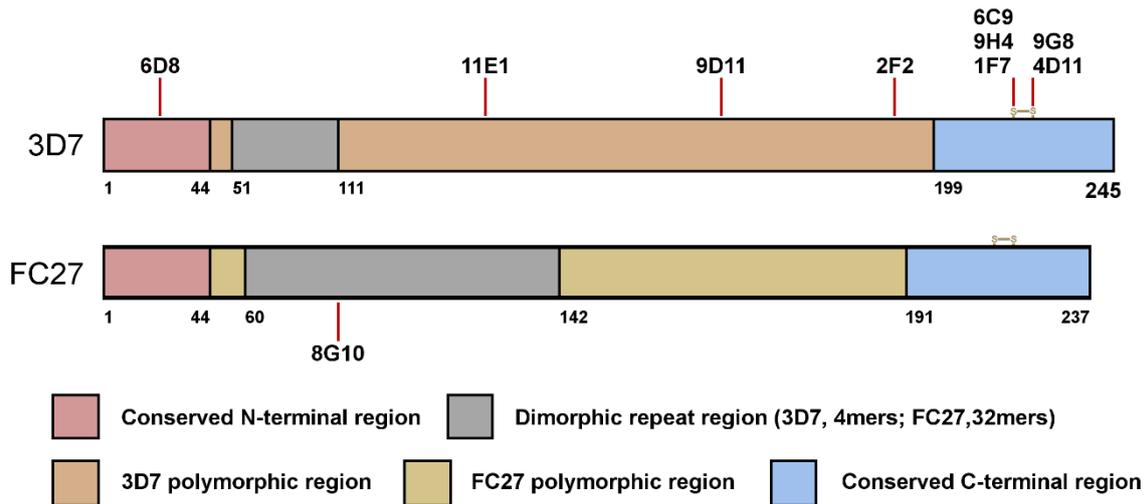


Figure 2: Schematic diagram of 3D7 and FC27 allelic families of MSP2, depicting conserved, dimorphic and polymorphic domains. Anti-MSP2 mouse mAb epitopes are indicated with red lines.

MSP2 is well characterised as a vaccine candidate. Several human vaccine trials involving the antigen have been being carried out, with one trial demonstrating that an immune response against MSP2 was efficacious (22, 26–30), Phase I/IIb trials of the “Combination B” vaccine, consisting of three recombinant antigen components, RESA, MSP1 and 3D7-MSP2, were undertaken on 120 Papua New Guinean (PNG) infants (27, 28). A 62% reduction in parasite densities was observed, with participants having a lower prevalence of parasite carrying the 3D7 allelic form of MSP2. Furthermore, over a 12-month period, a higher incidence of morbidity was associated with FC27-type parasites, suggesting that the protection observed was due at least in part to the MSP2 component of the vaccine. In an effort to circumvent the problem of strain specific protection, a Phase I trial was carried out using a mixture of full-length 3D7 and FC27 MSP2 adjuvanted in Montanide ISA720 (26). Although the trial was terminated midway due to adjuvant reactogenicity and concerns regarding the vaccine stability, antibodies induced by the vaccine exhibited functional activity by antibody-dependent cellular inhibition (ADCI) assays.

Unfortunately, the specific mechanisms behind antibody-mediated protection are poorly understood. Growth inhibition assays (GIAs) have been the most widely used functional assay for anti-merozoite antibodies, but there is a poor correlation between growth inhibitory antibodies and protective immunity (31, 32). Anti-MSP2 antibodies show limited inhibitory activity in standard GIAs (26, 30) despite showing efficacy in vaccine trials (22). Even antibodies from immune adults are often unable to inhibit parasite replication in these standard assays. More reliable functional correlates of protection, including the ADCI (26) and the

opsonic phagocytosis assay (31), have shown that anti-MSP2 antibodies are able to both inhibit *in vitro* growth of parasite in the presence of monocytes and induce opsonic phagocytosis.

A recent study identified antibody-mediated complement-dependent inhibition as a dominant mechanism in the targeting of merozoites by both vaccine-induced and naturally-acquired immune responses (33). Antibodies able to activate the classical complement pathway were found to be strongly correlated with protection from disease, with MSP2 identified as a major target. While antibodies to MSP2 can be internalised into the erythrocyte when bound to the merozoite surface in complement-free systems (31), in the presence of complement these antibodies effectively inhibited invasion. These studies suggest that mechanisms we are only now beginning to understand are mediating the inhibitory activity of antibodies that, by themselves, are not directly inhibitory. These results further strengthen the case for MSP2 as a promising malaria vaccine candidate.

## **1.2 Merozoite surface protein 2 as a vaccine candidate**

MSP2, in common with many other merozoite surface proteins (34), is intrinsically disordered in solution (35). The sequence of MSP2 is remarkably hydrophilic and shows low sequence complexity, both of which are hallmarks of intrinsically disordered proteins (IDPs). IDPs are a class of proteins that do not have a well-defined structure under native, functional conditions (36, 37). The discovery of these proteins has challenged the long-held paradigm that a defined structure is a pre-requisite for protein function (38, 39) with an increasing number of examples of functional IDPs able to assume a structure when complexed with their binding partner. Moreover, IDPs are promiscuous in their protein-protein interactions and are able to adopt differing conformations dependent on their binding partners (40). It is possible that the abundance of IDPs in the proteome of *P. falciparum* may provide a mechanism of immune escape for the parasite by acting as a “smokescreen”, biasing the host immune response towards non-protective epitopes. Little is known of the implications of antigen disorder for immune recognition. However, recent analysis of antibody recognition of disordered antigens has challenged the long-held view that disorder is detrimental to immune recognition (41). Instead, it was shown that the affinity of antibodies to their cognate epitopes was only weakly dependent on the degree of disorder. Furthermore, disordered epitopes were generally smaller and more efficient in their binding interactions than ordered epitopes, reinforcing disordered antigens as important targets of immune recognition. MSP2 is an excellent candidate to better understand this largely understudied area and may have broader implications in the optimisation of other disordered vaccine candidates.

## 1.4 Anti-MSP2 mouse monoclonal antibodies

A significant challenge in the development of vaccines targeting blood-stage antigens is the extensive diversity of these antigens (19, 42, 43). These polymorphisms have evolved under the pressure of immune selection to assist the parasite in immune evasion. As this diversity can be detrimental to the development of broadly protective antibodies, vaccines need to cover the majority of strains able to cause infection (11). One way to avoid these problems is to skew the immune response towards conserved regions. In the case of MSP2, the highly conserved N- and C- terminal regions are promising targets for the design of broadly protective MSP2-based vaccines, and are the main focus of this thesis.

This project builds on previous work by Adda *et al* (44), where a panel of monoclonal antibodies (mAbs) was generated from mice immunised with recombinant forms of both 3D7 and FC27 MSP2. To locate the binding region of each antibody, epitope mapping was carried out using a peptide array of overlapping 13-residue peptides spanning both alleles of MSP2. Overall, ten mAbs recognising epitopes in conserved and variable regions of MSP2 were identified (**Figure 2**). Five antibodies; 9H4, IF7, 6C9, 4D11 and 9G8, recognised an overlapping epitope within the conserved C-terminal region of MSP2, while a single mAb, 6D8, recognised an epitope within the conserved N-terminal region. Within the variable regions of MSP2 three antibodies, 2F2, 9D11 and 11E1, recognised 3D7-type MSP2 and mAb 8G10 recognised the FC27 allelic form of MSP2.

Intriguingly, these epitopes showed differing abilities to recognise parasite MSP2. These differences are discussed in detail below. Understanding the structural determinants of the binding of these antibodies to their cognate epitopes may provide valuable insight into how MSP2 is presented in its native environment.

## 1.5 N-terminal epitopes

Despite the intrinsically disordered nature of MSP2, some regions of local structure are present, notably within the conserved regions (45, 46). In particular, the 25-residue N-terminus of MSP2 has been found to adopt an  $\alpha$ -helical conformation upon lipid interaction (35, 46–48). The N-terminus of MSP2 could thus interact with the merozoite surface membrane in a similar way. A single antibody, 6D8, was able to recognise an epitope within the conserved N-terminal of recombinant MSP2 (**Figure 3**) (44). However, immunofluorescence assays (IFA) showed that 6D8 was unable to recognise native MSP2 on the parasite. Understanding the structural determinants of the interaction between 6D8 and its cognate epitope, and comparing the epitope

structure with its  $\alpha$ -helical lipid-bound conformation, should shed light on possible mechanisms of immune escape.

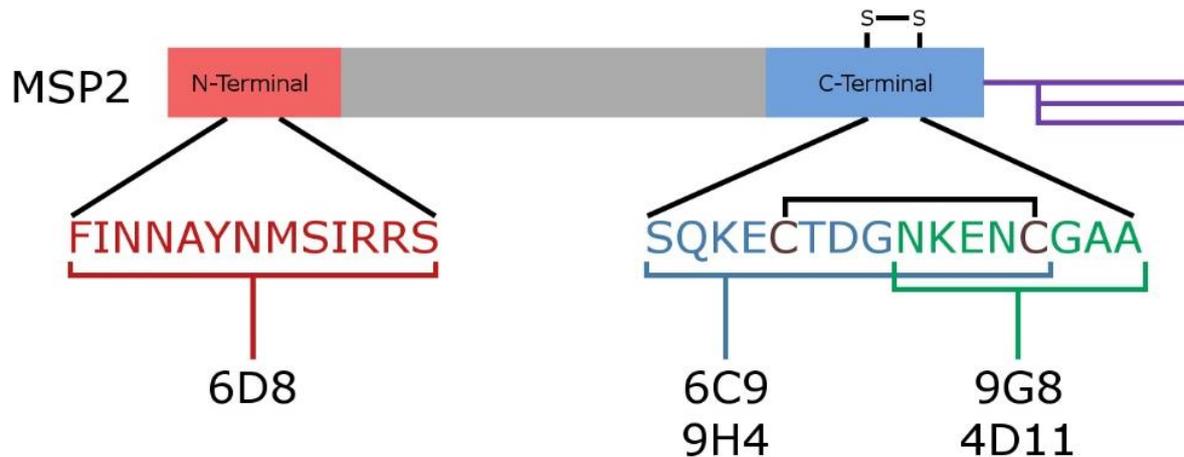


Figure 3: Amino acid sequence of key epitopes within the conserved regions of MSP2

### 1.6 C-terminal epitopes

The conserved C-terminal region of MSP2 spans 50 amino acid residues and contains the only disulfide bond in the protein (19, 43). Five antibodies (9H4, 1F7, 6C9, 4D11 and 9G8) are able to recognise overlapping epitopes within an 18-residue stretch encompassing the disulfide bond of the C-terminal region (**Figure 4**) (44). As previously alluded to (section 1.4), these antibodies showed differing reactivity towards native (parasite) MSP2 by both IFA and western blot. Of the five antibodies recognising conserved C-terminal epitopes, only 4D11 and 9G8 gave strong IFA signal to parasite MSP2. The other C-terminal epitope antibodies, 9H4, 6C9 and 1F7, recognise an adjacent overlapping epitope with 10-fold weaker IFA signal to parasite MSP2 than 4D11 and 9G8, suggesting that this epitope is less accessible on the merozoite surface. There may also be slight conformational differences between recombinant and native MSP2. As the protein is C-terminally anchored by a GPI moiety (18, 44), the close proximity of C-terminal epitopes to the merozoite membrane surface may contribute to this difference. Furthermore, it is important to acknowledge that the conformation of MSP2 on live parasite may differ from that of native MSP2 in IFA and western blot. The process of parasite fixation in IFA and the use of lysed parasite in western blot can change the protein conformation and in turn influence antibody recognition (49).

## 1.7 Antibody fragments

The above studies utilised antibody isotype immunoglobulin G (IgG), a key player in the humoral immune response. IgGs are large molecules, usually around 150 kDa in size (**Figure 4**), which are composed of four polypeptide chains, two identical heavy and two identical light chains arranged in a Y-shape, typical of monomeric antibodies. The use of full-size IgGs or even antigen-binding fragments (Fabs) is not ideal for nuclear magnetic resonance (NMR) spectroscopy or other structural studies. Furthermore, their production is time-consuming and expensive. Smaller antibody domains, variable fragments (Fv) and single-chain variable fragments (scFv) make ideal substitutes for full size IgGs and have been used in a number of applications, including as immunotherapeutics (50–52). These fragments consist of just the variable domains ( $V_H$  and  $V_L$  chains) of the IgG and retain the parent molecule's binding properties. A scFv differs from an Fv in having a short peptide linker between the  $V_H$  and  $V_L$  (53–56). These antibody fragments can be expressed recombinantly in *Escherichia coli*, significantly reducing both the time and costs of production. Their smaller size (25-30 kDa) also means they are amenable to NMR spectroscopy (57). In previous work during my honours project in collaboration with the Garvan institute (unpublished) scFv and Fv antibody fragments were designed for all antibodies targeting conserved regions of MSP2 (58). The expression of these constructs was optimised with 6D8, 9H4 and 4D11 scFv/Fv antibody fragments giving suitable yields for structural analyses.

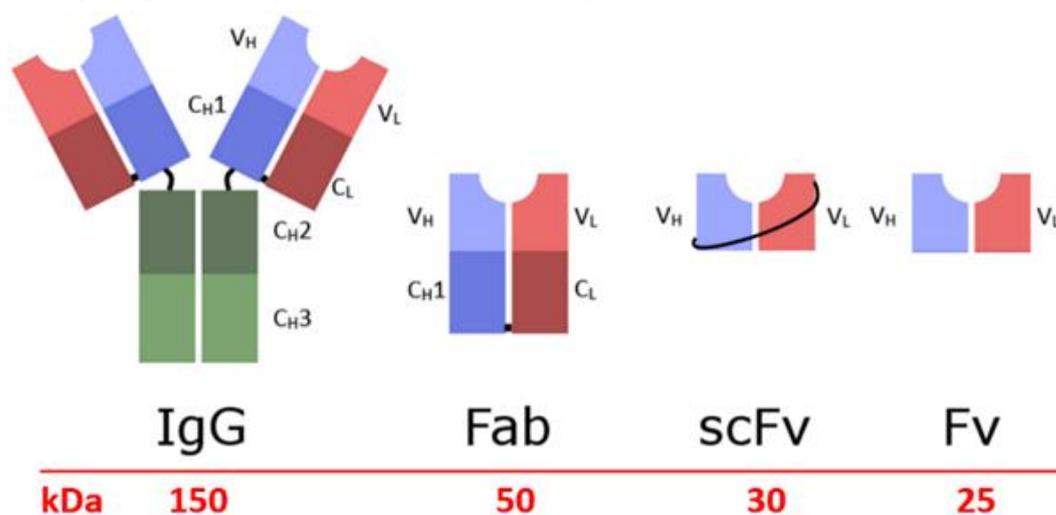


Figure 4: Schematic diagram of IgG, Fab, scFv and Fv antibody domains

## 1.8 Peptide vaccines

One of the focuses of the work presented in this thesis are peptide vaccines and their application to MSP2. The following section is based on a published review article by MacRaild, C. A. *et.al.* in *Peptide Science*, 2018 (**Appendix I**).

Vaccines are an indispensable tool in the fight against disease and have had a significant impact on public health over several decades (59). Traditionally, vaccines have relied on the use of live-attenuated or inactivated forms of the pathogen to induce a protective immune response. However, in some cases, such as malaria, the pathogen is difficult to culture *in vitro* at large scale. Moreover, the use of whole organisms as vaccines introduces a high antigenic load when often only a small subset of antigens are driving protection (60). There is also the possibility of adverse allergic reactions to certain antigens in these preparations. These obstacles have led to an increased interest in subunit vaccines, in which single, or a select few, proteins are used in a vaccine formulation to induce protective immunity (61).

The use of peptides as vaccines takes this rationale further, as even a single protein antigen such as MSP2 may have many epitopes, not all of which contribute to protective immunity. Peptide vaccines offer a means to formulate vaccines containing only epitopes that are capable of inducing a positive and efficient immune response (52, 60). The ease of peptide synthesis makes large-scale production feasible, whilst also offering a cleaner vaccine preparation lacking biological contaminations commonly associated with recombinant expression or whole organism vaccines. The specificity of peptides also allows for improved customisability, facilitating, for example, a multi-epitope approach to target different strains or stages in the life cycle of the pathogen.

For these reasons, there has been long-standing interest in this area of vaccine design, with over 500 peptide vaccines reaching clinical trials, targeting a wide range of indications (**Table 1**). Seven of these vaccines have reached Phase III trial (**Table 2**). Much of this recent interest has focussed on T-cell epitopes, with all of the vaccines to reach Phase III being T-cell based. For many diseases, however, B-cell responses also play an important role in immune protection, although peptide-based vaccines based on B-cell epitopes have proved more challenging for a number of reasons. Owing to the small size of peptides, they are generally poorly immunogenic. In addition, T-cell epitopes required for the establishment of a robust immune response may also be absent in smaller peptides. To address these issues, carrier proteins or adjuvants are typically included in the vaccine formulation. As peptides are also prone to enzymatic degradation, epitopes can be modified by conformational stabilisation or

cyclisation, or peptide mimetics that are resistant to proteolysis can be developed (62). Finally, and perhaps most importantly, peptide-based approaches are limited because the antibody response to many protein antigens is dominated by conformational epitopes, which are difficult or impossible to capture effectively in a peptide-based design. Intrinsically disordered antigens such as MSP2 are uniquely suited for implementation as peptide vaccines because disordered epitopes are invariably linear, eliminating the challenge of capturing conformational epitopes in a peptide vaccine.

Table 1: Peptide vaccines in clinical trials

	<b>Number of active or completed clinical trials</b>	<b>Conditions being treated with peptide vaccines</b>
<b>Phase III</b>	7	Cancer immunotherapies, Multiple sclerosis, Type 1 Diabetes
<b>Phase II</b>	203	Cancer immunotherapies, Myelodysplastic syndrome, HIV, HBV, HCV, Cytomegalovirus, Myasthenia gravis, Influenza, Malaria ( <i>falciparum</i> )
<b>Phase I and Early Phase 1</b>	307	Cancer immunotherapies, HIV, HPV, HBV, HCV, Age related macular degeneration, Respiratory syncytial virus, Malaria ( <i>vivax</i> ), Malaria ( <i>falciparum</i> ), Hand foot and mouth disease, Influenza, Multiple sclerosis, Alzheimer's disease, Listeria, Cat allergy, Ragweed allergy, House dust mite allergy, Grass allergy

Table 2: Phase III clinical trials currently active or completed

Candidate	Construct	Condition	Clinical status	Clinical Trials Identifier
<b>MDX-1379</b>	Two peptides from gp100 melanocyte protein	Metastatic melanoma	Phase III	NCT00094653
<b>PR1 leukaemia peptide vaccine</b>	Derived from proteinase 3 and neutrophil elastase	Acute myeloid leukaemia	Phase III	NCT00454168
<b>Telomerase peptide vaccine GV1001</b>	Derived from reverse transcriptase subunit of telomerase (hTERT)	Pancreatic cancer	Phase III	NCT00425360
<b>NeuVax</b>	Derived from human leukocyte antigen HER2	Breast cancer	Phase III	NCT01479244
<b>NeuroVax</b>	Two peptides from T-cell receptor	Multiple sclerosis	Phase III	NCT02057159
<b>MAGE-A3 and NY-ESO-1 Immunotherapy</b>	Peptides from MAGE-A3 and NY-ESO-1 proteins	Multiple myeloma	Phase III	NCT00090493
<b>Diaprep277</b>	T-cell epitope of heat shock protein 60	Type 1 Diabetes	Phase III	NCT01281072

## 1.9 Structure-based vaccine design

Structure-based vaccine design, or structural vaccinology, is an emerging strategy for the rational design of vaccine candidates (63, 64). The recent improvement in high-resolution structural analyses such as X-ray crystallography, NMR spectroscopy, and cryo-EM have enabled us to move away from traditional methods of vaccine design that rely heavily on trial-and-error testing of antigens. Instead, by using high-resolution structures, antigens can be designed to better mimic their native conformations and yield a more efficient immune response, offering a way forward for pathogens where traditional methods have failed (65). Whilst the structures of therapeutic targets have been integral in the design of many small molecule drugs on the market today, the approach has only recently been implemented in vaccine design. Although the field is still in its infancy, structure-based vaccine design has already shown promise in a number of disease conditions, including *meningococcus* B (66), respiratory syncytial virus (RSV) (67, 68), Group B *Streptococcus* (69), and HIV (70–72). The first two will be discussed in detail below.

Similar to malaria, *Neisseria meningitidis* employs a high degree of polymorphism in its genome to evade the immune system (73, 74). In the case of the surface-exposed lipoprotein, factor H-binding protein (fHbp), 500 known sequences could be categorised into three distinct variant groups. Although immune responses against fHbp have proven to be bactericidal, there

are difficulties with strain specificity as immunisation with each group alone is unable to induce a cross-reactive response (75). Using a NMR and X-ray crystallography to determine the structure of fHbp, a large-scale design effort was initiated to overcome this sequence variability. Key conformational epitopes were first identified from each variant, then these epitopes were categorised depending on their location on the protein into 11 discrete regions. A total of 54 chimeric antigens was designed and expressed by grafting these key epitopes in varying combinations. Two of these fHbp chimeras were able to induce a broad bactericidal immune response against strains carrying any of the three variant groups of fHbp (66).

Another example of structural vaccinology being applied successfully can be seen in a respiratory syncytial virus vaccine based on the RSV fusion glycoprotein (RSV F) (68, 76). Efforts to generate an effective RSV F-based vaccine are complicated by large conformational changes during fusion to the host-cell membrane. The work focused on an antigenic site specific to the pre-fusion state of the glycoprotein, which is meta-stable, but functionally relevant and known to elicit potent neutralising antibodies (67, 77). Using six high-resolution crystal structures of RSV F in complex with these antibodies, an analogue was designed with the key antigenic site stabilised. This site was accessible even when the protein was exposed to extreme pH, osmolality and temperature. Ultimately, the stabilised RSV F glycoprotein was able to elicit an RSV-specific immune response with high neutralising activity in mice and macaques (68).

The above examples involve the application of structure-based vaccine design towards highly structured proteins. In these examples, extensive effort was needed to preserve conformational epitopes and correct protein folding throughout the design process. Because MSP2 is intrinsically disordered, these considerations do not apply. Moreover, as a consequence of the focus on conformational epitopes, the field of structural vaccinology has not yet ventured into the domain of peptide vaccines. MSP2 provides an exciting platform to combine these two strategies for the design of novel malaria vaccines. Our results may have wider implications in contributing to a general understanding of the immune response towards disordered protein antigens.

## 1.10 Project Scope and Aims

MSP2 is a promising blood-stage vaccine candidate that has shown evidence that it can induce an efficacious immune response (22). The protein is highly abundant on the merozoite surface and has proven to be essential for parasite invasion of the erythrocyte (16, 18). In previous work, a panel of anti-MSP2 mouse mAbs was identified and their binding epitopes were determined by peptide array (44). The structural determinants underlying the differing abilities of these antibodies to recognise parasite MSP2 are unknown. Determining the antibody-bound conformations of epitopes that are accessible on the parasite surface will provide valuable insight into possible conformations of native MSP2. The main aim of this study is to merge the growing field of structure-based vaccine design with the advantages of peptide vaccines to design novel malaria vaccine candidates based on MSP2. By studying an intrinsically disordered protein such as MSP2, we also aim to add to the knowledge of the immune response towards disordered antigens.

Naturally-acquired protective immunity to MSP2 in malaria is dominated by the central variable region. Strain-specific reduction in parasite density induced by the ‘Combination B’ vaccine also suggested that the variable region was playing a role in vaccine-induced protection. However, the high degree of polymorphism in this area poses a problem for vaccine design. Chapter 2 describes the analysis of key epitopes in the central variable region of MSP2 and the design and expression of antibody fragments for the 3D7-specific 2F2, 9D11 and 11E1 mAbs. This Chapter also describes efforts to crystallise 9H4 antibody fragments with its cognate epitope.

From Chapter 3 onwards the focus shifts to the conserved regions of MSP2. This Chapter describes work on the N-terminal region specific mAb 6D8, an antibody unable to recognise native MSP2 by western-blot and IFA. The N-terminal region of MSP2 is known to adopt an  $\alpha$ -helical conformation when interacting with lipid. It is possible that this helical conformation is present on the parasite surface. X-ray crystallography and NMR were used to understand why the 6D8 epitope is masked on the parasite surface. Additionally, the peculiar phenomenon of strain-specific binding within a highly conserved region is explored.

Although our studies of the N-terminal region provided valuable insight into the effects of disorder on antibody recognition, in Chapter 4, the focus shifts towards the conserved C-terminal region. As antibodies targeting the C-terminal region of MSP2 are able to recognise parasite MSP2, this region is appealing as a target for vaccine design. This Chapter describes

a high-resolution crystal structure of 4D11 Fv in complex with its minimal binding epitope, and analyses the structural determinants of the binding of this key epitope.

Based on this structural information, Chapter 5 describes the design of a series of constrained dimeric, linear and backbone cyclised peptides, for use as possible MSP2-based peptide vaccines. MD simulations were performed to assess if the additional constraints would add too much strain on the peptides. Promising candidate peptides were synthesized, then surface plasmon resonance (SPR) was used to determine their binding to 4D11 before conjugation to KLH and immunisation in mice. The resulting sera were analysed by enzyme-linked immunosorbent assay (ELISA) to determine antibody titre and specificity.

Finally, Chapter 6 summarises the findings of this study and discusses potential future directions of this work.

# Chapter 2

Cloning, expression and structural characterisation of antibody fragments against MSP2

## 2.1 Introduction

The polymorphisms of MSP2 are confined to the central variable region of the protein, which consists of highly polymorphic tandem repeats, flanked by non-repetitive dimorphic sequences that define the two allelic families of MSP2, 3D7 and FC27 (19, 78). Although polymorphisms in the central variable regions of MSP2 may make the region less appealing as a vaccine target, vaccine trials and sero-epidemiological studies suggest that antibodies targeting these regions are associated with a reduction in parasite density in children. Naturally-acquired immune responses in PNG children were found to be directed against the allele-specific central variable region (79, 80). Furthermore, the multicomponent “Combination B” vaccine containing 3D7-type MSP2 reduced parasite density in PNG children by 62% (22). This reduction was only seen in parasites of the 3D7-MSP2 genotype and was accompanied by a higher incidence of morbidity associated with FC27-type parasite. A subsequent Phase I vaccine trial including both allelic forms of MSP2 as components was able to induce antibodies that exhibited cellular inhibition of parasite growth (26). Moreover, a recent immunisation trial in mice used chimeric MSP2 analogues consisting of conserved and variable regions of both FC27 and 3D7 MSP2 was also shown to elicit a robust strain-transcending response across both families (81).

All four antibodies that recognised epitopes within the central variable region of MSP2 (8G10, 2F2, 9D11 and 11E1) were strongly reactive to native parasite MSP2 by IFA, indicating that the central variable region is accessible to the immune response (44). The epitopes recognised by 2F2, 9D11 and 8G10 are all located in dimorphic regions of MSP2 that are fairly conserved within their respective allelic families. In contrast, the 11E1 epitope resides in a more polymorphic region of MSP2 (**Figure 1**). Thus, the 11E1 mAb would be an excellent candidate with which to investigate the effects of polymorphisms on antibody affinity and their role in possible mechanisms of immune escape. A recent study characterising the naturally-acquired immune response from four malaria endemic regions using epitope mapping identified antibodies specific to both variable and conserved regions were present (82). Among others, variable region epitopes recognised by 2F2 and 9D11, and the conserved C-terminal 4D11/9H4 epitopes, were found to be immunodominant. The lack of antibodies able to bind the 11E1 epitope may be attributed to its high degree of polymorphism, as peptides used to probe antibodies would be unable to capture all MSP2 variants. Although an immune response towards conserved regions of MSP2 would be ideal for broad protection against both allelic families of MSP2, inducing production of antibodies able to recognise variable regions of MSP2 may be necessary for effective protection.

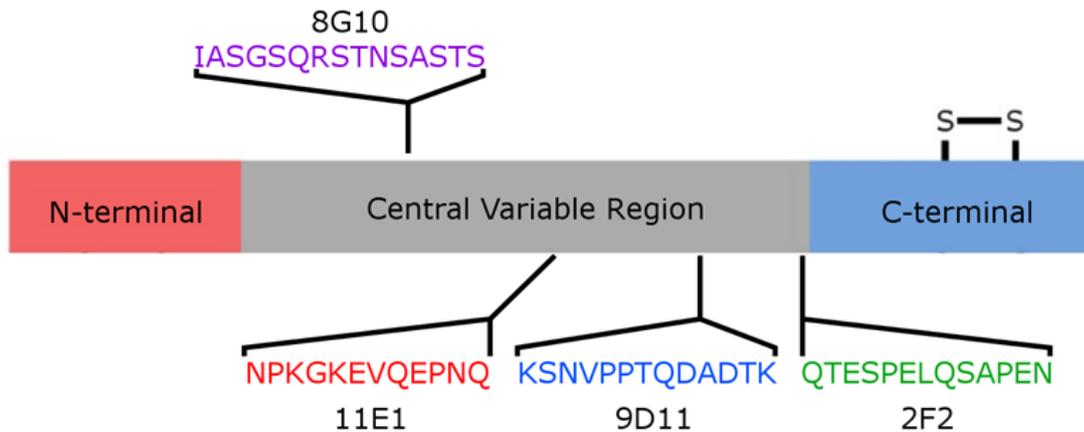


Figure 1: Amino acid sequence of key epitopes within the variable regions of MSP2. 8G10 is FC27-specific, whilst 11E1, 9D11 and 2F2 are 3D7-specific.

In this Chapter, known MSP2 sequences of variable region epitopes recognised by mAbs 2F2, 9D11 and 11E1 are analysed. Understanding mutations that are present in the population, and their frequency, may assist in the design of MSP2 analogues able to induce a broader protective response against a wide range of MSP2 variants. Constructs for antibody fragments derived from 2F2, 9D11 and 11E1 were designed and expressed by first acquiring their sequences from their corresponding mouse hybridoma cells lines. These sequences were cloned into scFv and Fv scaffolds using available 4D11, 9G8, 6C9, 9H4 and 6D8 antibody fragments as a template.

The C-terminal region of MSP2 is recognised by five mAbs: 4D11, 9G8, 6C9, 1F7 and 9H4. Each of these shows different abilities to recognise parasite MSP2, suggesting that some epitopes are more accessible than others on the merozoite surface. Also described in this Chapter are efforts to determine crystal structures of 9H4 bound to its cognate epitope. Although the 9H4 epitope is poorly accessible on the parasite surface by IFA and western blot, understanding the conformation of the antibody-bound epitope may account for this epitope masking. Furthermore the structure can inform the design of better MSP2-based antigens able to steer the immune response towards more accessible epitopes.

## 2.2 Materials and methods

### *Expression and purification of 3D7-MSP2 thioredoxin fusion protein*

Thioredoxin (THDX) fusion protein of 3D7-MSP2 was expressed in *E. coli* BL21 (DE3) overnight at 37°C in 50 mL of Luria broth (LB) with 100 µg/mL ampicillin (AMP) and 15 µg/mL tetracycline (TET) as a seed culture. Overnight culture was diluted 1000-fold with fresh LB medium with AMP and grown at 37°C to an optical density (OD<sub>600</sub>) of 0.8-1.0. At this point isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce expression. After a further 6 h at 37°C, the cells were harvested by centrifugation (8000 g for 10 min at 4°C), lysed by resuspending in 50 mM phosphate buffer, pH 7.0 and incubated in a boiling water bath for 15 min. The lysis of cells by boiling is unconventional and would usually be detrimental to a protein, but the intrinsically disordered nature of MSP2 allows it to withstand such high temperatures (35). Lysate was cleared by centrifugation (16000 g for 20 min at 4°C) and the His-tagged protein was purified by fast protein liquid chromatography (FPLC) using a high-capacity Ni-nitrilotriacetic acid (NTA) resin. A gradient of 0-500 mM imidazole in 250 mM NaCl / 50mM phosphate buffer pH 7.0 over 6 min was run at a flow rate of 2.5 mL/min. Eluted samples were concentrated and buffer-exchanged into 50 mM phosphate buffer using an Amicon Ultra-15 10 kDa centrifugal filters. The concentrated sample was then treated with tobacco etch virus (TEV) protease for 48 h at 24°C with the addition of 1 mM dithiothreitol (DTT) for removal of THDX. Cleaved protein was filtered and further purified by high-performance liquid chromatography (HPLC) through an Altima 5 µM 250 mm C8 column. A gradient of 5% - 80% acetonitrile (ACN) 0.1% trifluoroacetic acid (TFA) was used over 80 min at 5 ml/min flow rate.

### *Assessing polymorphisms in known sequences of MSP2*

Sequences of MSP2 were found by a nucleotide and protein Basic Local Alignment Search Tool (BLAST) (83) search of the swissprot and pdb database using both gene sequences of 3D7 and FC27 MSP2. Sequences corresponding to short fragments or not containing the MSP2 variable region were removed. Multiple sequence alignment was then carried out with Clustal Omega (84). The sequences were translated to their amino acid sequence and each variable region epitope was assessed separately. Geneious version R7 (<http://www.geneious.com>) (85) was used to compile these data and create the sequence logos. The sequences were sorted manually by their mutations and the frequency of these mutations within the total sample was calculated.

### Isolation and PCR amplification of Fv domains

Mouse hybridoma cells lines producing anti-MSP2 mAbs 2F2, 9D11 and 11E1 were provided by the WEHI antibody facility (Bundoora, Victoria). Cells were grown to confluence, harvested, washed in phosphate buffered saline (PBS) and provided as frozen cell pellets. The cells were lysed and homogenised in TRIzol solution (Life Technologies). Total RNA was purified using the PureLink RNA Mini Kit (Life Technologies) as per supplier instructions. First strand cDNA was synthesised using the commercially available SMART complementary deoxyribonucleic acid (cDNA) library construction kit (Clontech Laboratories). Amplification of cDNA libraries was performed with Platinum Pfx DNA polymerase (Life Technologies) following the supplier's protocol. Polymerase chain reaction (PCR) amplification of mouse variable-region domains were performed with the same degenerate primers used by Fields *et.al.* (86) (**Table 1**). As the yields from most of these PCR amplifications were low, the PCR product was re-amplified using the same primers to acquire suitable amounts of DNA. The DNA was gel purified and cleaned with the PureLink quick gel extraction kit (Life technologies) and stored at -20 °C. Each insert was restriction digested with Nco1 and Not1 and cloned into pET32a vector for sequencing. The functionality of each V<sub>H</sub> and V<sub>L</sub> sequence was evaluated by sequence alignment with the mouse immunoglobulin set stored at the IMGT reference directory (<http://www.imgt.org>) (87). Following validation of functional sequences, codon-optimised genes for each chain were ordered from Genscript with appropriate restriction sites at the 3' and 5' ends.

Table 1: Degenerate primers used for amplification of mouse variable heavy (VH) and light (VLL and VLK) chains, highlighted sequence indicate restriction sites

Name	Sequence (5'-3')
VH-FOR	CTTCCGCCATGGSARGTBMAGCTGSAGSAGTCWGG
VH-REV	TACAGG <b>GCGGCCGC</b> GGACAGTGGATARACMGATGG
VKAPPA-FOR	CTTCCGCCATGGGAYATTGTGMTSACMCARWCTMCA
VKAPPA-REV	TACAGG <b>GCGGCCGC</b> GGGATACAGTTGGTGCAGCATC
VLAMBDA-FOR	CTTCCCATGGCAGGCTGTTGTGACTCAGGAA
VLAMBDA-REV	TACAGG <b>GCGGCCGC</b> GCTTGGGCTGACCTAGGACAGT
Degeneracy code:	R (AG) , B (GCT) , M (AC) , Y (CT) , W (AT)

### Construction of Fv constructs

The new anti-MSP2 Fv chains were constructed using a previously sequenced, codon-optimised framework from 9H4 Fv in pET12a. The novel V<sub>H</sub> and V<sub>L</sub> regions required two consecutive rounds of restriction digestion and ligation to replace the existing V<sub>H</sub> and V<sub>L</sub> in the 9H4 Fv frame (**Figure 2**). Specific primers (**Table 2**) were designed incorporating suitable restriction sites for each V<sub>H</sub> (SalI and NsiI) and V<sub>L</sub> chain (NcoI and XhoI) including stop codons at the reverse anchors. Each chain was amplified with Platinum Pfx DNA polymerase (Life Technologies) by PCR and the product was gel-purified before restriction digestion. The V<sub>H</sub> was first cloned into the 9H4 Fv frame using T4 ligase (Life Technologies) at room temperature for 20 min, followed by insertion of the V<sub>L</sub>. The ligated product was transformed into ultracompetent XL10 cells and colonies were chosen for plasmid isolation. Completed Fv constructs were sequenced and stored at -20°C, ready for transformation and expression.

Table 2: Specific primers used for construction of Fv and scFv constructs, highlighted sequence indicate restriction sites

Name	Sequence (5'-3')
2F2VH_SalI_For	TTTGC <b>GTCGAC</b> CAGGTCAAGCTGGAGGAGTC
9D11VH_SalI_For	TTTGC <b>GTCGAC</b> GAAAGTTCAGCTGGAGCAGTC
11E1VH_SalI_For	TTTGC <b>GTCGAC</b> GAAAGTCAAGCTGCAGCAGTC
VH_FV_NsiI_Rev	TCCGCT <b>ATGCAT</b> <b>TTATTA</b> AGGGTGTTCGTTTTGGCTG
2F2VLKA_NcoI_For	AGCCGG <b>CCATGG</b> <b>CC</b> GACATTGTGCTGACCCAGAC
9D11VLKA_NcoI_For	AGCCGG <b>CCATGG</b> <b>CC</b> GATATTGTGCTGACACAATCTCC
11E1VLLA_NcoI_For	AGCCGG <b>CCATGG</b> <b>CC</b> GATATTGTGATGACCCAGACTACATC
VL_Fv_XhoI_Rev	TCCGCT <b>CTCGAG</b> <b>TTATTA</b> TGCAGCATCAGCCCG
OmpT_scFv_NdeI_For	GATATA <b>CATATG</b> CGGGCGAAACTCCTAGGA
11E1_Fv_NsiI_BamHI_Rev	CCAGGA <b>ATGCAT</b> TTATTA <b>GGATCC</b> GGAGACGGTGACCGTGGTCCC
2F2VH_BamHI_Rev	T <b>GGATCC</b> GGGTGTTCGTTTTGGCTGCA
9D11VH_BamHI_Rev	T <b>GGATCC</b> GGGTGTTCGTTTTGGCTGAG
2F2_9D11VH_GGGGS_Bam_NcoI_Rev	GTCGG <b>CCATGG</b> AACCACCGCCACCAGAACCAGCCACCACCGCTACCACC
ACCACC <b>GGATCC</b> GGGTGTTCGT	
11E1_GGGGS_Bam_NcoI_Rev	GTCGG <b>CCATGG</b> AACCACCGCCACCAGAACCAGCCACCACCGCTACCACC
ACCACC <b>GGATCC</b> GGAGACGGT	

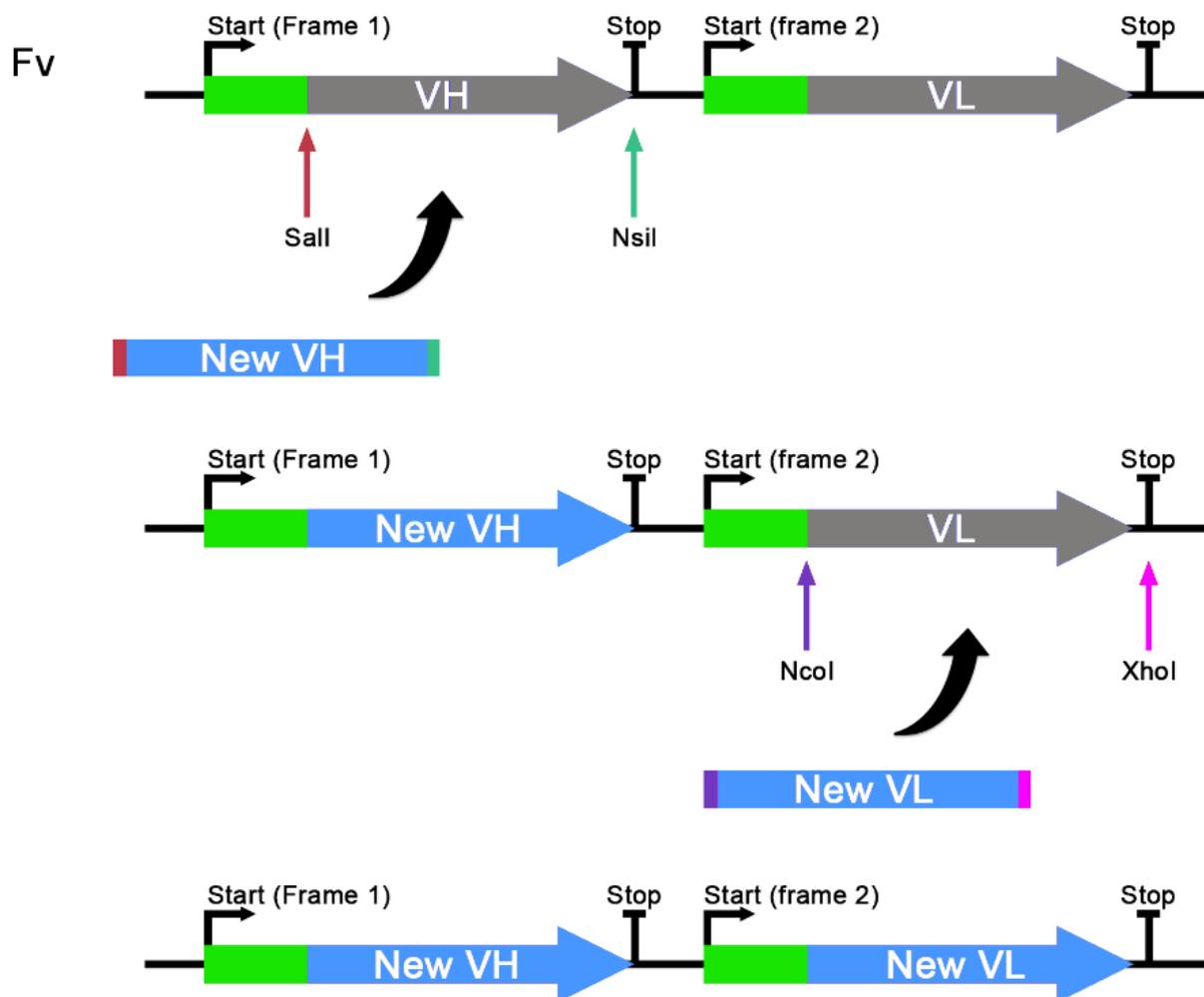


Figure 2: Schematic diagram of construction of novel Fv antibody fragments utilising 9H4 Fv as a base

### *Construction of scFv constructs*

The initial method for generation of scFv constructs utilised previously-sequenced Fv fragment constructs as a template. This involved replacing the silent linker region between the V<sub>H</sub> and V<sub>L</sub> chains with a flexible spacer corresponding to three GGGGS repeats (**Figure 3**). Specific primers were designed to incorporate suitable restriction sites along with the Gly-Ser linker (**Table 2**). The Fv constructs were first restriction digested with Nde1, Nsi1 and Nco1 to release two fragments corresponding to the V<sub>H</sub> (450 bp) and the silent linker (90 bp), the V<sub>H</sub> fragment and linear plasmid were gel-purified. The V<sub>H</sub> was used as a base for the installation of the Gly-Ser linker before ligation into the linear pET12a-V<sub>L</sub> plasmid. As the Fv constructs contained a frameshift between the V<sub>H</sub> and V<sub>L</sub>, a single nucleotide was added into the linker region to correct the reading frame. Two PCR amplifications with Platinum Pfx DNA polymerase (Life

Technologies) were used to install the scFv 3XGGGGS linker to the V<sub>H</sub>. The first set of primers was used to remove the stop codon downstream of the V<sub>H</sub> and also to extend the consensus 3' terminal sufficiently for efficient primer pairing of the second set of primers necessary for installation of the GGGGS linker. The V<sub>H</sub>-linker fragment was restriction-digested with NdeI and NcoI and ligated into the linear pET12a-V<sub>L</sub> plasmid with T4 ligase (Life Technologies). The ligated product was transformed into ultracompetent XL10 cells and colonies were chosen for plasmid isolation. Completed scFv constructs were sequenced and stored at -20°C, ready for transformation and expression. This method was used to create scFv constructs for 2F2 and 9D11, although, owing to poor expression, codon-optimised scFv genes were subsequently purchased from Genscript.

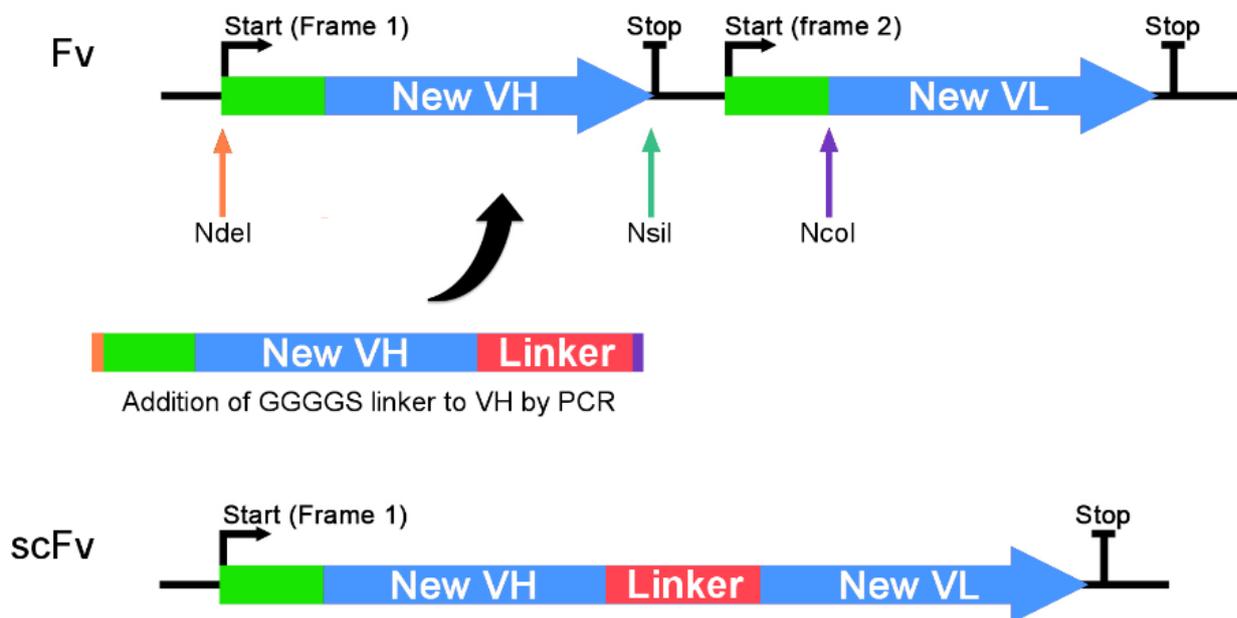


Figure 3: Schematic diagram detailing the conversion of Fv constructs into their corresponding scFv constructs by insertion of a Gly-Ser linker

#### *IPTG induction and expression of antibody fragments*

All scFv or Fv antibody fragments were expressed in *E. coli* strain BL21 (DE3). Overnight cultures were grown at 37°C in 50 mL of LB with 100 µg/mL AMP, 15 µg/mL TET and 4% glucose and diluted 1000-fold with fresh LB medium containing AMP, TET and 4% glucose at 37°C to an OD<sub>600</sub> of 0.8-1.0. The cells were then pelleted (5000 g for 15 min at 4°C) and re-suspended in fresh LB medium containing AMP and TET with no glucose. After the cells were equilibrated in the fresh media for 2 h, IPTG was added to a final concentration of 1 mM to induce expression. The induced cells were incubated at room temperature (RT) overnight and harvested by centrifugation.

### *Autoinduction of antibody fragments*

To improve the yield of poorly-expressed antibody fragments and increase the simplicity and consistency of induction, autoinduction was trialled. (88) ZYM-5052 autoinduction media was prepared with (per litre): tryptone 10 g, yeast extract 5 g, Na<sub>2</sub>HPO<sub>4</sub> 3.56 g, KH<sub>2</sub>PO<sub>4</sub> 3.4 g, NH<sub>4</sub>Cl 2.68 g, Na<sub>2</sub>SO<sub>4</sub> 0.71 g, glycerol 4 ml, glucose 0.5 g, lactose 2 g, 1 x BME vitamins solution (Sigma-aldrich), and MgSO<sub>4</sub> 1 mM. Overnight cultures were grown at 37°C in 50 mL of LB broth with 100 µg/mL AMP, 15 µg/mL TET and 4% glucose. The overnight culture was diluted 1000-fold with fresh ZYM-5052 medium and incubated for 41 h at 30°C. Cells were harvested by centrifugation and periplasmic extraction (see below) was performed.

### *Periplasmic extraction*

Each antibody fragment construct contains a periplasmic signal sequence that allows the expressed protein to be secreted into the periplasm. The oxidising environment of the periplasm allows for the formation of disulfide bonds and correct folding. Extraction of protein from the periplasm is achieved by lysis of the bacterial outer membrane by osmotic shock. The cell pellet (6 g) was resuspended in 35 mL of sucrose buffer (30 mM Tris 2 mM Ethylenediaminetetraacetic acid (EDTA)/30% sucrose/pH 8.0) prior to centrifugation at 10000xg for 20 min at 4°C. The supernatant was discarded and the pellet resuspended in 30 mL ice-cold MQ water with 2 mM MgCl<sub>2</sub> and 10 µg/ml DNase I. The resuspended cells were incubated in ice for 15 min for periplasmic rupture. Cells were pelleted at 15000xg for 20 min at 4°C and the supernatant periplasmic extract collected. The periplasmic extract was filtered and subjected to affinity purification for removal of antibody fragments.

### *Affinity purification*

The MSP2 affinity column was prepared by amide conjugation of full-length 3D7 MSP2 to hydroxysuccinimidyl-sepharose resin. This was performed following the manufacturer's protocol with 5-10 mL of resin for each column. Purification was performed in a gravity column. The periplasmic extract of scFv and Fv fragments were flowed through a MSP2 affinity column, washed with 5 column volumes of wash buffer (50 mM phosphate buffer at pH 7.0), then eluted in 15 mL of 200 mM glycine pH 2.7 and neutralised immediately with 3 mL of 1M Tris base. Eluted samples were then concentrated using an Amicon Ultra-15 10K Centrifugal Filter Unit to 1/10<sup>th</sup> of the original volume. The concentrated sample was dialysed twice against 50 mM ammonium bicarbonate at 4°C, over 24 h then lyophilised.

### *Refolding of novel variable region antibody fragments*

Owing to the possibility of incorrect folding in the periplasm or denaturation during the process of extraction, some antibody fragments were unable to be purified by affinity purification. The first method of refolding involved placing the periplasmic extract under redox conditions. Reduced and oxidised forms of glutathione were added to 10 mL of periplasmic extract and kept at 4°C for 3 h. The next method of refolding was by dilution in GdnHCl, with periplasmic extract buffer exchanged into 3.5 M GdnHCl, 5 mM DTT, pH 7.4 to fully denature the protein. Four subsequent dilutions were performed 4 h apart, gradually reducing the GdnHCl concentration. The first dilution contained 2M GdnHCl, 50 mM Tris, 0.2 M NaCl, 1mM EDTA, pH 7.4, and the two subsequent dilutions contained 1 M and 0.5 M GdnHCl with 0.4 M ArgHCl, 1mM GSSG, 2mM GSH, 50 mM Tris, 0.2 M NaCl, 1mM EDTA, pH 7.4. Before affinity purification, a final dilution was carried out in PBS.

### *Peptide synthesis*

Peptides including 3D7-MSP2<sub>207-222</sub> and 3D7-MSP2<sub>215-222</sub> were synthesised in-house by standard 9-fluorenylmethoxycarbonyl (Fmoc) solid-phase chemistry using an automated peptide synthesiser 3 (PS3). The peptides were assembled via coupling of 0.3 mmol (3 equiv) of Fmoc-protected amino acids to 0.1 mmol rink amide AM resin (0.53 mmol/g loading). Coupling reactions were carried out for 50 min under the activation of 0.3 mmol (3 equiv) O-(1H-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) and 0.6 mmol (6 equiv) N,N-diisopropylethylamine (DIPEA). A double coupling was performed on the first residue of each peptide. Chain deprotection was carried out with 20% piperidine in dimethylformamide (DMF) for 2 min. The peptides were N-terminally capped with an acetyl moiety using 0.5 mmol (5 equiv) of acetic anhydride in 0.5 mmol (5 equiv) DIPEA. Cleavage of the complete peptides was performed with TFA:triisopropylsilane (TIPS):water (95:2.5:2.5 v/v).

The cleaved material was precipitated in cold diethyl ether overnight at -20 °C. The insoluble peptide material was spun down at 4,000 rpm for 30 min at 0 °C and the pellet washed twice in cold diethyl ether prior to removal of the organic phase. The crude peptide mixture was resuspended in 50% acetonitrile/0.1% TFA, filtered and freeze-dried prior to further purification. Disulfide bond-containing peptides were oxidised by air in 100 mM ammonium bicarbonate for 2 h before HPLC purification. MSP2 peptides were purified on a reverse-phase C18 column (Zorbax, 10 x 300 mm) using a linear gradient of 5 to 60% of solvent B (80%

acetonitrile / 9.9% water / 0.1% TFA) against solvent A (0.1% TFA in water) over 1 h at 5 ml/min flow rate.

#### *Removal of purification tag from scFv constructs*

To make constructs of 4D11 and 9H4 scFv more amenable to crystallisation, analogues were designed to remove the flexible His-tag and Myc-tag at their N-terminus. Conveniently, a BamHI restriction site was located directly after the V<sub>L</sub> chain, upstream of the His-tag. A simple restriction double digest with NdeI and BamHI was used to remove the 800 bp insert corresponding to the V<sub>H</sub>-linker-V<sub>L</sub> insert. This insert was then ligated into a pET12a plasmid without the purification tag completing the construct leaving the desired construct.

#### *Antibody crystallisation*

Freeze-dried antibody fragments were dissolved in 1 ml of 20 mM Tris buffer with 100 mM NaCl at pH 7.0 and mixed with synthetic peptide epitopes at a scFv/Fv:peptide molar ratio of 1:1.5. Size-exclusion chromatography was then performed with a Superdex 10/300 GL column (0.5 mL/min) and fractions containing complexed monomeric antibody fragment were collected and concentrated. Once the desired concentration (15-30 mg/mL) was reached, the complexed antibody-peptide was filtered through a 0.22 µm filter. Promising conditions from the initial trial (JCSG+ screen kit) were optimised in a subsequent narrow screen with variation in buffer pH and precipitant concentrations.

Crystals were grown using the hanging-drop vapour diffusion method, with 1:1 (vol/vol) ratio of protein to reservoir solution, using well volume of 0.5 mL. Crystals were cryo-protected by the addition of 10 % glycerol prior to data collection. Each dataset was collected at 100 K using the Australian Synchrotron macro crystallography MX1 beamline 3BM1. Diffraction images were processed using iMosfilm and Scala from the CCP4 suite.(89) 5% of each dataset was flagged for calculation of R<sub>Free</sub>(90) with neither a sigma nor a low-resolution cut-off applied to the data.

Structure determination proceeded using the Molecular Replacement method and the program PHASER (91). 6D8 Fv (PDB ID: 4QYO) was used as an initial model, the peptide epitope was removed and sequence alignment with the 9H4 Fv sequence was performed. The aligned sequences were input into CHAINSAW (92) to mutate the 6D8 Fv residues into those of 9H4 Fv. Long stretches of residues that were not in consensus between the two sequences were automatically removed from the model. PHASER MR (91, 93) was used to solve the structure by molecular replacement.

## 2.3 Results

### *Polymorphisms in variable regions of MSP2*

Of the panel of mAbs available, four are able to recognise epitopes within the central variable regions of MSP2. There is one FC27-specific antibody, 8G10, and three 3D7-specific antibodies, 2F2, 9D11 and 11E1. 11E1 is particularly interesting as it recognises an epitope within a highly polymorphic region of MSP2. BLAST searches of the MSP2 gene sequence for both 3D7 and FC27 were performed to gather all known MSP2 sequences in the population. A total of 420 sequences was found for 3D7-type MSP2 and 203 for FC27-type MSP2. When aligned and translated to their amino acid sequences, the 13-residue epitopes of each antibody were analysed individually for mutations, deletions, amino acid frequency, and the frequency of each mutation among the total sample analysed. Sequence logos were created to visualise the sequence conservation of each epitope (**Figure 4**).

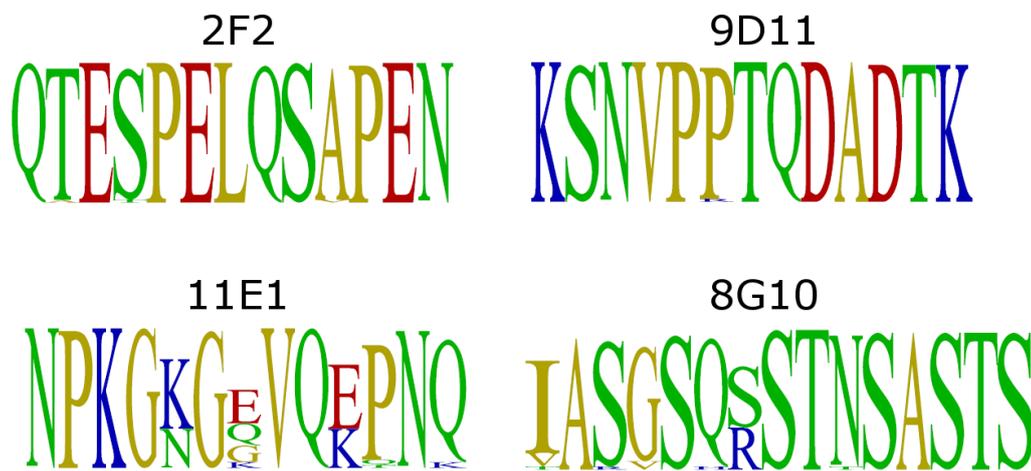


Figure 4: Sequence logos of variable-region epitopes 2F2 (3D7-MSP2<sub>172-184</sub>), 9D11 (3D7-MSP2<sub>142-154</sub>) 11E1 (3D7-MSP2<sub>112-124</sub>) and 8G10(FC27-MSP2<sub>77-89</sub>). Amino acid frequency for each position is indicated by the height of their single letter code, overall height represents mean pairwise identity over all pairs in the column, and colour represents amino acid charge or hydrophobicity.

The results showed high sequence conservation in both 2F2 and 9D11 epitopes within the 3D7 allelic family. This is unsurprising as these epitopes are within dimorphic regions of MSP2. The 8G10 epitope also showed moderate sequence conservation across the FC27 allelic family, with only one frequent mutation of Arg in place of a Ser. The 11E1 epitope had a high degree of polymorphisms, with some mutations corresponded to fairly drastic changes in amino acid properties. One of the most common mutations involved a negatively-charged Glu swapping to a positively-charged Lys. Other mutations corresponded to loss of charge (K216N) and increased flexibility (E218G). Such extreme changes would be expected to substantially

change the antigenic surface of the epitope and may be an effective mechanism of immune evasion.

The overall frequency of each 11E1 mutation was calculated; the top 8 most frequent sequences are shown in **Table 3**. Surprisingly, the reference sequence present in our form of 3D7-MSP2 was only seen in 3.3% of the sample, instead, single point mutations corresponding to Q118E (26.4%) and E121K (25.7%) were the most common in the sampled sequences.

Table 3: 11E1 epitope sequence frequency

<b>Freq.</b>	<b>%</b>	<b>Sequence</b>
14	3.3	NPKGKGEVQEPNQ (ref)
111	26.4	NPKGKGQVQEPNQ
108	25.7	NPKGKGEVQKPNQ
50	11.9	NPKGNGGVQEPNQ
34	8.1	NPKGNGGVQKPNQ
22	5.2	NPKGNGKVQEPNQ
12	2.9	NPKGKGEVQEQNQ
10	2.4	NPKGNGEVQEPNQ

#### *CDNA library sequencing of variable region monoclonal antibodies*

In previous work by the Norton lab, scFv and Fv constructs for mAbs recognising conserved regions of MSP2 (4D11, 9G8, 6C9, 9H4 and 6D8) were designed and made available. mAbs targeting the variable region have not been sequenced and their corresponding scFv and Fv antibody fragments constructs are yet to be created. The following work describes the design and expression of these constructs.

Hybridoma cells for 2F2, 9D11 and 11E1 were acquired from Professor Robin Anders at La Trobe University. To isolate the antibody DNA required for the design of scFv and Fv antibody fragments, total RNA was extracted from the hybridomas. mRNA was then isolated and a cDNA library was synthesised (**Figure 5A**). The genetic information for the variable heavy and light chains ( $V_H$  and  $V_L$ ) was amplified by polymerase chain reaction (PCR) from the cDNA libraries using degenerate primer sets designed for amplification of mouse variable domains (94–96). As there are two classes of  $V_L$  domains,  $\lambda$  and  $\kappa$ , a separate primer set was used for each. As initial PCR reactions were unsuccessful or had low yields, reaction conditions for each cDNA library and primer set required individual optimisation (**Figure 5B**).

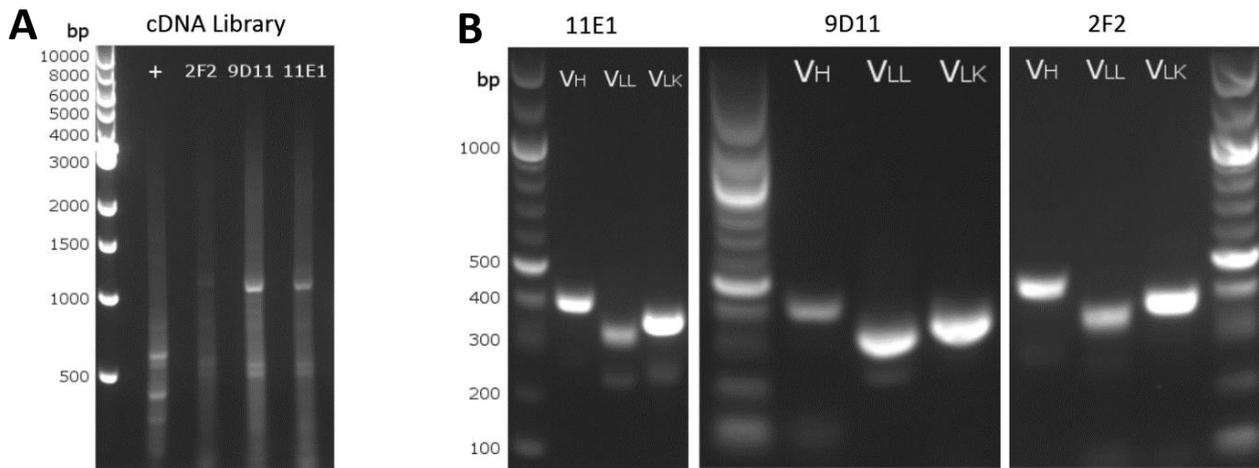


Figure 5: (A) DNA agarose gel of cDNA library amplified from variable region hybridomas, (B) DNA agarose gel showing antibody domain PCR amplification products.

Intriguingly, two  $V_L$  chains were amplified ( $\kappa$  and  $\lambda$ ) from the cDNA of all hybridomas. The ability of a single cell line to express more than one heavy or light chain is not uncommon and has been observed by others in the literature. (97, 98) The fragments for each chain were cloned into pET32a for sequencing. Sequencing and subsequent evaluation with the IMGT reference directory (<http://www.imgt.org>) revealed that there was a functional  $V_H$  and  $V_L$  for each mAb (**Table 4**). This involved checking the sequence for unwanted stop codons, and frame shift errors. The sequences were also matched with known V-gene and J-gene alleles in the database. The sequences were also The  $V_{LL}$  sequences of each mAb were not functional. For 11E1, both  $V_L$  sequences were identical, although this was probably because the  $V_{LL}$  primer was able to amplify the  $V_{LK}$  sequence. The full  $V_H$  and  $V_L$  sequences are shown in **Table 5**.

Table 4. Summary of gene products sequenced from mouse hybridoma 2F2, 9D11 and 11E1

Hybridoma	Primer set	V-gene	J-gene	Status
2F2	$V_H$	IGHV1-52*01 IGHV1-61*01 F	IGHJ3*01 F	Functional
	$V_{LL}$	IGLV1*01	IGLJ1*01	Non-functional
	$V_{LK}$	IGKV3-7*01	IGKJ2*01	Functional
9D11	$V_H$	IGHV1-53*01 IGHV1S16*01	IGHJ2*01 IGHJ2*02	Functional
	$V_{LL}$	IGLV1*01	IGLJ1*01	Non-functional
	$V_{LK}$	IGKV8-30*01	IGKJ1*01	Functional
11E1	$V_H$	IGHV14-3*02 IGHV14-4*01	IGHJ1*01	Functional
	$V_{LL}$	IGKV8-28*01	IGKJ5*01	Functional (same as VLK)
	$V_{LK}$	IGKV8-28*01	IGKJ5*01	Functional

Table 5: Variable region amino acid V<sub>H</sub> and V<sub>L</sub> sequences, Grey and yellow highlights indicate CDRs

Name	Translated sequence
2F2 V <sub>H</sub>	QVKLEESGAELVKPGASVKLSCKASGYTFTSYW <sup>W</sup> MHWVKQRPGQG <sup>LEWIGNIDPSD</sup> SETHYNQKFKDKATLTV <sup>DKSSSTAYMQLSSLTSEDSAVYY</sup> CAR <sup>VGLLYYGDYDWF</sup> AYWGQGT <sup>TLVTVS</sup>
2F2 V <sub>LK</sub>	DIVLTQTPASLAVSLGQRATISCRASQSVSTSSYSYMH <sup>W</sup> YQQKPGQPPKLLIK <sup>YA</sup> SNLESGV <sup>PARFSGSGSTDFTLNIHPVEEEDTATYY</sup> CQHSWEIPYT <sup>FGGGTKLEI</sup> K
9D11 V <sub>H</sub>	EVKLQQSGAEVVKPGASVKLSCKASGYTFTDYWIH <sup>W</sup> VKQGPQG <sup>LEWIGEINPTN</sup> DDTNYNEKFKNRATLTV <sup>DESSTAYLQLSSLTSEDSAVYY</sup> CTV <sup>RGLIRWVDYWGQ</sup> GTALTVS
9D11 V <sub>LK</sub>	DIVLTQSPSSSLAVSVGEKVTMSC <sup>KSSQSVLYSSNQKNYLAW</sup> YQQKPGQSPKVL <sup>LIY</sup> WASTRES <sup>GVPDRFTGSGSGTDFTLT</sup> ISSVKAEDLAVYY <sup>CQYYTYRT</sup> FGGG <sup>TKLEI</sup> IK
11E1 V <sub>H</sub>	EVKLQQSGAELMRPGTSVKLSCTASGFNIKDDYIH <sup>W</sup> VKQRPEQG <sup>LEWIGRIDPAN</sup> GNTKYVPKFQDRAT <sup>ITADTSSNIAYLQLSSLTSEDTAVYY</sup> CARRDGNYGWYFDVW <sup>GAGTTVTVS</sup>
11E1 V <sub>LK</sub>	DIVMTQTTSSLSVSAGERVTMSC <sup>KSSQSLNLSGNQKNYLAW</sup> YQQKPGQPPKVL <sup>LIY</sup> GAFTRES <sup>GVPDRFTGSGSGTDFTLT</sup> ISSVQAEDLAVYY <sup>CQNDHSYPLT</sup> FGAG <sup>TKLEI</sup> ELK

### Cloning of variable-region Fvs

After acquiring V<sub>H</sub> and V<sub>L</sub> chain sequences for each variable-region mAb and validating them as functional, the process of antibody fragment construction could move forward. First, the V<sub>H</sub> sequence was amplified with specific primers and digested with SalI and NsiI restriction enzymes. The 9H4 Fv vector that would serve as a scaffold for these new antibody fragments was also digested (**Figure 6A**). After ligation of the new V<sub>H</sub> into the vector, the V<sub>L</sub> was inserted using NcoI and XhoI restriction enzymes (**Figure 6B**). The completed constructs were then sequenced to verify that cloning was successful.

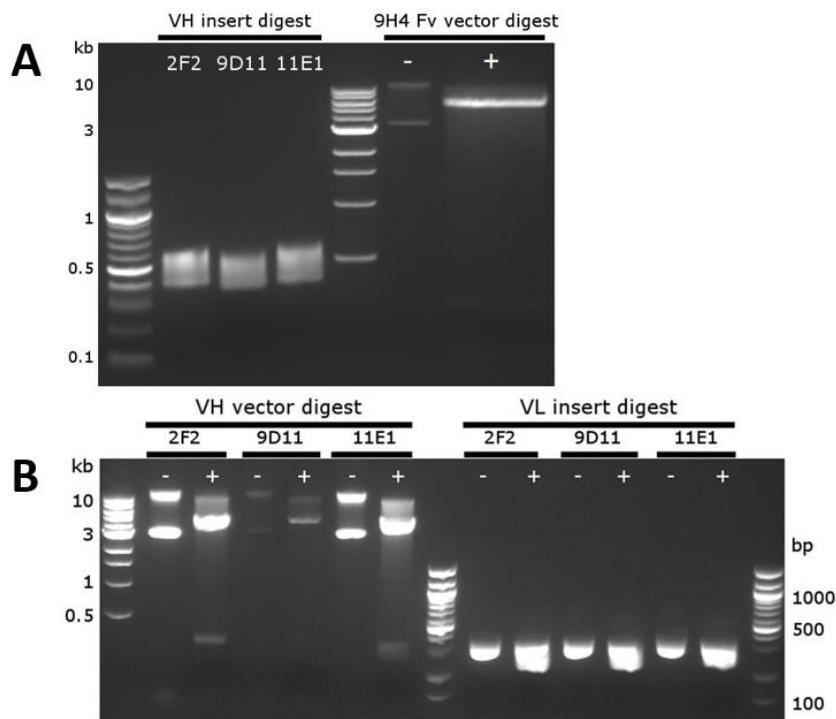


Figure 6: (A) Digestion of 2F2, 9D11 and 11E1 V<sub>H</sub> inserts with SalI and NsiI in preparation for cloning into 9H4 Fv vector (B) Digestion of V<sub>L</sub> inserts with NcoI and XhoI for ligation into their cognate V<sub>H</sub> containing vector. – indicates undigested and + indicate digested.

### Variable-region Fv expression

The plasmids were transformed into *E. coli* BL21 (DE3) for expression. Initial test expressions revealed poor expression for all variable region Fv fragments. Subsequent large-scale expression revealed similar results (**Figure 7**), with 11E1 displaying no noticeable expression whilst 2F2 and 9D11 Fv had some protein in the expected 10-15 kDa range. However, leaky expression was observed in these constructs as protein was present in the uninduced sample. The poor expression may be attributed to the use of mammalian sequences that have not been optimised for recombinant expression.

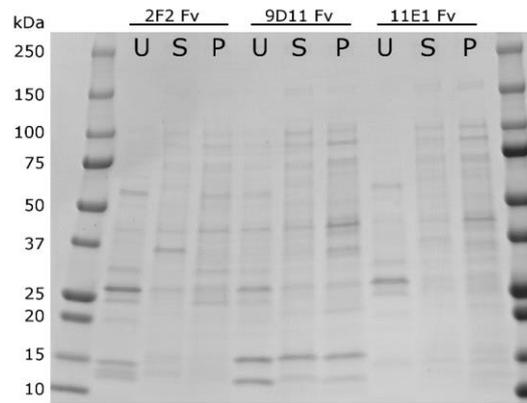


Figure 7: Variable region Fv expression. U - uninduced, S - sucrose extract, P - periplasmic extract

#### *Variable-region scFv expression*

Using the Fv fragments above as templates, scFv antibody fragment constructs were successfully cloned and sequenced for 2F2 and 9D11. However, owing to the poor expression of Fv fragments due to the lack of codon optimisation, work on these constructs was halted. Instead, codon-optimised genes for 2F2, 9D11 and 11E1 scFv fragments were purchased. Test expression of these new constructs yielded some promising results (**Figure 8**), with three colonies picked for each construct. Both 2F2 and 9D11 scFv had good expression, with a large band in the expected 25-30 kDa range by Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Leaky expression was again seen in the uninduced lanes, but the inclusion of 4% glucose used in larger-scale expression should reduce the undesired expression. 11E1 scFv showed no expression, with no clear band in the expected range. Considerably fewer colonies were observed after transformation of the 11E1 construct, suggesting that this antibody fragment may be in some way toxic to the cells.

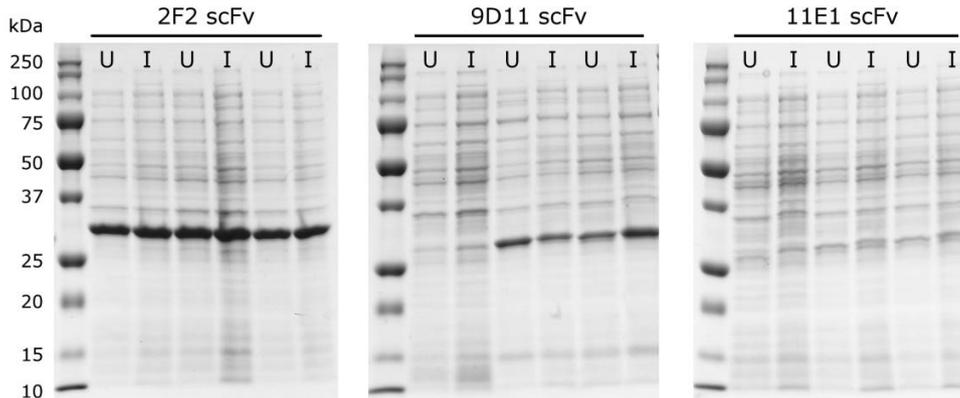


Figure 8: Test expression of variable region scFv antibody fragments; three colonies were picked for each construct. U - uninduced, I - induced

Larger-scale expression was carried out with the 2F2 and 9D11 scFv constructs. The periplasmic extracts showed relatively clean samples with good scFv expression (**Figure 9A**). Unfortunately, affinity purification with 3D7-MSP2 was unable to separate the scFv, suggesting that the antibody fragments were unable to bind MSP2. It is possible that the antibody fragments were misfolded in the periplasm or denatured during the process of periplasmic extraction. In an effort to recover the scFv, first, a simple redox refolding strategy was employed. Two redox refolding conditions using reduced and oxidised forms of glutathione were tested on 2F2, but affinity purification yielded no correctly-folded scFv (**Figure 9B**). Dilution refolding of 2F2 scFv using decreasing concentrations of GdnHCl was also unable to correctly refold the antibody fragment (**Figure 9C**). Further work will be required to express functional variable-region antibody fragments.

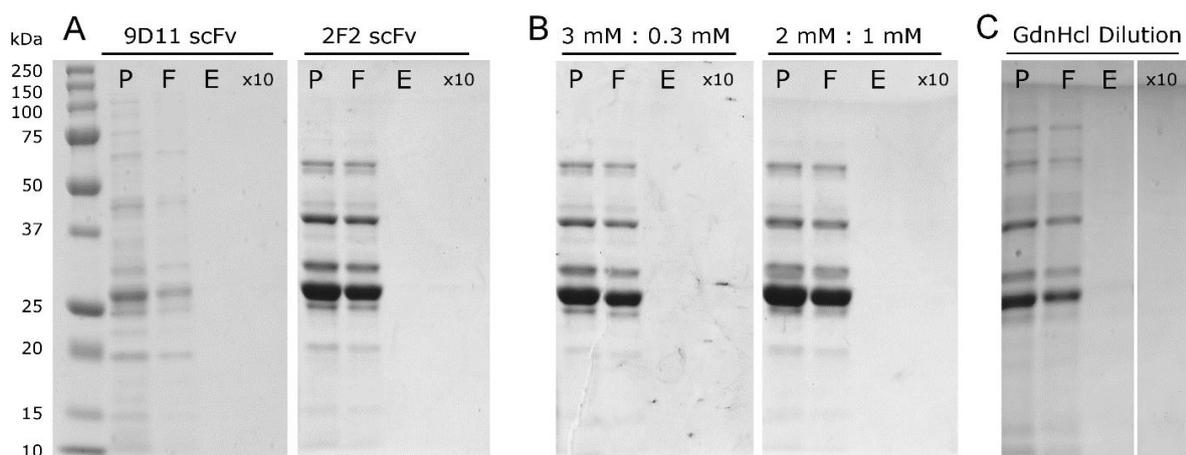


Figure 9: (A) Affinity purification of 9D11 and 2F2 scFv, (B) redox refolding of 2F2 scFv at 3 mM:0.3 mM and 2 mM:1 mM reduced:oxidised glutathione, (C) Dilution refolding in GdnHCl. P - periplasmic extract, F - flow-through, E - elution, x10 - ten-fold concentrated elution

#### *4D11 Fv/9H4 Fv + MSP2<sub>207-224</sub> crystallisation trials*

Prior to my candidature, crystallisation trials involving 4D11 scFv and 9H4 scFv in complex with their shared epitope MSP2<sub>207-224</sub> were carried out. Wide screens using three commercially available screening kits yielded no suitable crystals. Since the flexible linker present in the scFv antibody fragments may be detrimental to crystal formation, the Fv constructs were evaluated instead. The Fv forms of both of 4D11 and 9H4 Fv were complexed with MSP2<sub>207-224</sub> for initial wide screening of crystallisation conditions. Although the 4D11 Fv complex did not yield any promising conditions, encouraging conditions were found for the 9H4 Fv complex. Crystallisation conditions for the 9H4 Fv complex were further optimised in a narrow screen and several large needle-like crystals grew within 4 days (**Figure 10**). X-ray diffraction data for these crystals were collected at the Australian Synchrotron at 1.3 Å resolution.

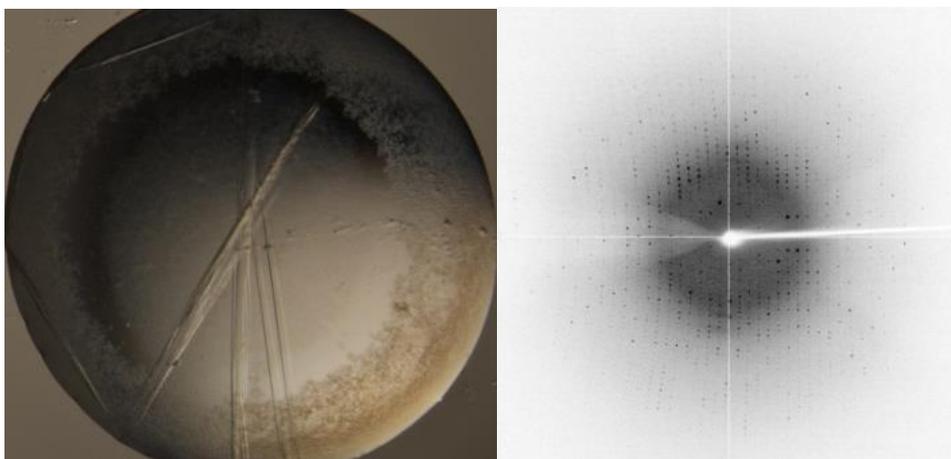


Figure 10: Image of thin needle-like 9H4 Fv crystals, and X-ray diffraction pattern of 9H4 Fv complexed with MSP2<sub>207-224</sub>

Despite the excellent data, attempts to solve the 9H4 Fv + MSP2<sub>207-224</sub> structure were not fruitful. Molecular replacement was used to solve the phases of the diffraction data using the structure of 6D8 Fv (PDB ID: 4QYO) mutated to the 9H4 Fv sequence as an initial model. As the structures of Fv antibody fragments are well characterised and known to not differ dramatically in fold or orientation, the use of a mutated 6D8 Fv as a model should be sufficient to solve the 9H4 structure. However, initial attempts to solve a structure were unsuccessful.

Subsequent attempts included; forcing different space groups, probing the data with V<sub>H</sub> and V<sub>L</sub> chains as separate ensembles, changing solvent content and varying the number of copies to search for. Some combinations of these parameters were able to solve structures, although their statistics were far from ideal ( $R_{\text{Free}}$  and  $R_{\text{cryst}} > 0.5$ ), suggesting that the solved structures were not correct. Manual inspection of these structures also revealed poor electron density occupancy and unusual orientation of Fv chains.

A potential explanation for the inability to solve a 9H4 Fv structure by molecular replacement may be its low binding affinity to its cognate epitope ( $K_d = 2.7 \mu\text{M}$ ). The peptide may have dissociated from 9H4 Fv during the crystallisation process. The lack of bound peptide may have destabilised the Fv resulting in a change in fold or orientation. To test this, a sample of the crystals used for data collection was picked and run by LCMS to confirm the presence of peptide (**Figure 11**). No peptide was detected in the sample, supporting the hypothesis that crystallisation had effectively removed the peptide from the complex.

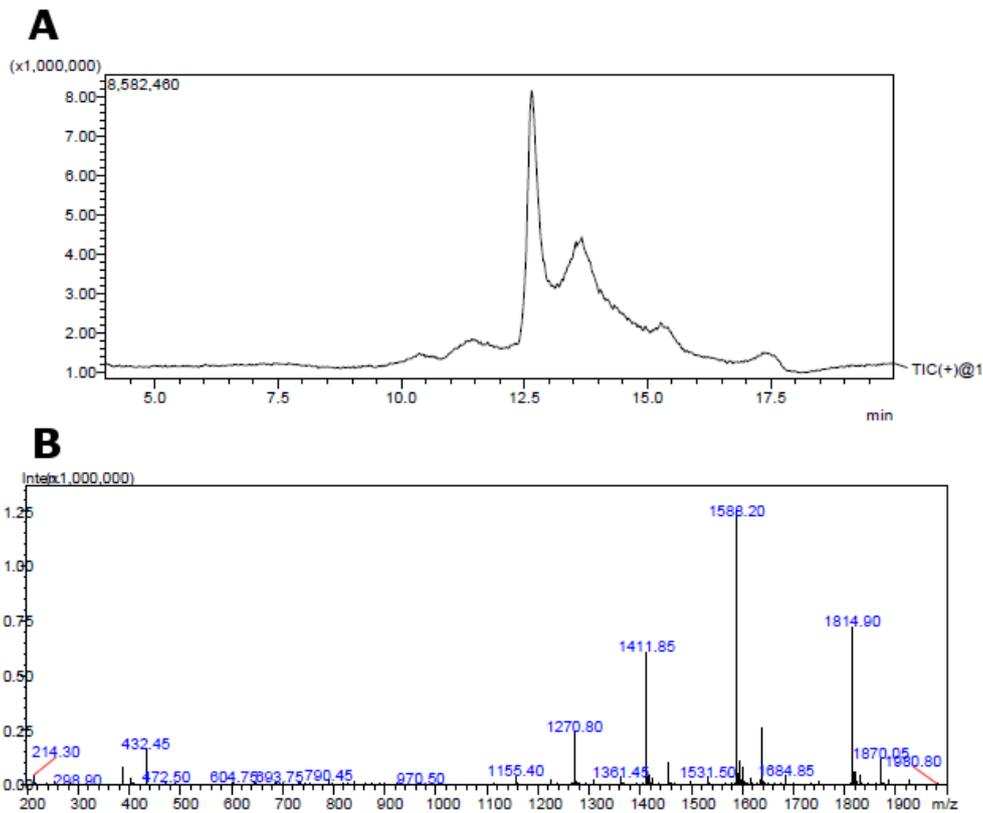


Figure 11: (A) LC trace of 9H4 crystals used in data collection (B) MS spectra of largest peak corresponding to 9H4 VH. The peptide mass (848 Da) was not observed throughout the spectra

### Removal of His-tag from scFv constructs

An alternative strategy for crystallisation of 9H4 and 4D11 with their cognate epitopes involved using their scFv equivalents. The linker between the heavy and light chain would greatly reduce the possibility of the chains dissociating from each other, as seen with 9H4 Fv. Previous attempts to crystallise scFv antibody fragments of 4D11 and 9H4 in a number of buffers and commercial screens had not yielded any protein crystals. It was thought that the flexible His-tag and Myc-tag at the N-terminus of these constructs may have had a detrimental effect on crystal packing. Two new constructs of 4D11 and 9H4 scFv were therefore generated with these tags removed. After sequencing of these constructs to confirm that the His-tag had been removed, large-scale expression was carried out in preparation for crystallisation screens. The overall yields for the no-tag constructs were similar to the original constructs (**Figure 12**).

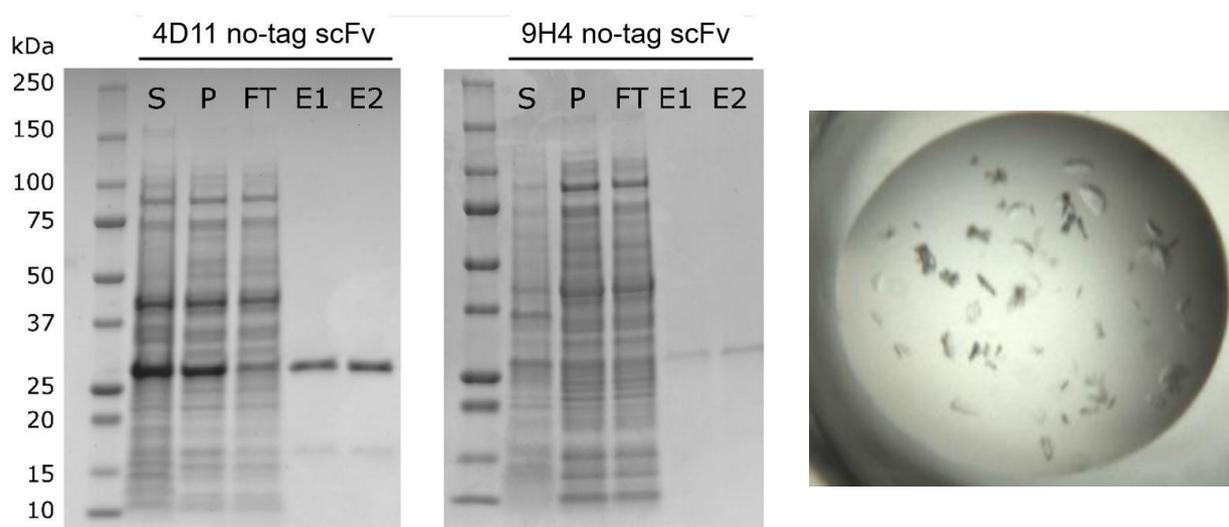


Figure 12. (left) Affinity purification of no-tag analogues of 4D11 and 9H4 scFv, S - sucrose extract, P - periplasmic extract, FT - flow through, E1/E2 - elutions and (right) image of crystals of 9H4 no-tag scFv in complex with MSP2<sub>207-224</sub>.

Both 9H4 and 4D11 no-tag scFv were complexed with 16-residue epitope MSP2<sub>207-224</sub> for crystallisation wide screens. Promising hits were found for the 9H4 constructs and taken further for optimisation in narrow screens. When taken to the synchrotron for data collection these crystals showed poor diffraction. This may be due to poor lattice formation as decreased attenuation and exposure to the X-ray source showed that protein was present. Conditions for this complex will require further optimisation to determine a structure.

### *Autoinduction of 9H4 scFv no-tag construct*

A bottleneck in progress towards crystallisation of 9H4 antibody fragments was its poor expression. Multiple large-scale batches were required to obtain suitable yields for screening of crystallographic conditions. Autoinduction provides a simple approach to protein expression, requiring minimal involvement after the inoculation of culture. This would save time from the standard IPTG expression of antibody fragments, which requires change of media to remove glucose and initiate expression. The 9H4 no-tag scFv construct was used to compare the autoinduction methods with standard IPTG induction (**Figure 13**). Although autoinduction was found to have superior expression even after 1 day of growth, periplasmic extraction was shown to be considerably more efficient in the IPTG-induced cells. Hence, overall yields between the two methods were similar. If periplasmic extraction were to be optimised for autoinduced cells, the yield of not just 9H4 but other antibody fragments may be improved upon significantly.

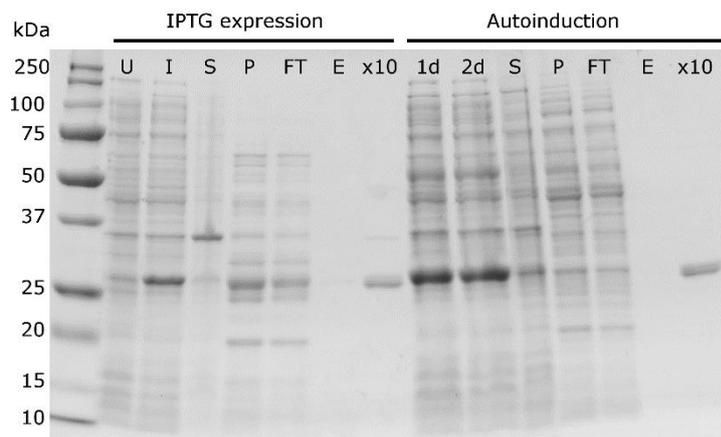


Figure 13: Comparison of IPTG expression and autoinduction of 9H4 no-tag scFv. U- uninduced, I - induced, S - sucrose extract, P - periplasmic extract, FT - flow through, E - elution, x10 – 10-fold concentrated elution

## 2.4 Discussion

The work in this chapter describes considerable efforts to generate scFv and Fv antibody fragments that target variable regions of MSP2.  $V_H$  and  $V_L$  sequences were successfully amplified and sequenced from cDNA libraries generated from mouse hybridomas cell lines. These sequences were first used to create Fv fragments, however, the low expression and poor yields of these constructs led to the purchase of codon-optimised scFv fragments. Although these constructs showed improved expression, they were unable to bind MSP2, suggesting misfolding in the periplasm or denaturation during extraction from the cell. Efforts to refold 2F2 scFv using redox conditions and dilution in GdnHCl were unsuccessful. Different methods of refolding and optimisation of the purification process may be able to remedy this (55, 56, 99). Alternatively, different expression hosts such as *Pichia* or mammalian cells may also improve the yield of functional antibody fragments (100, 101). The method of autoinduction, with some optimisation, showed potential to improve yields of 9H4 no-tag scFv. This approach could be applied to other antibody fragments, including those in the variable region.

The analysis of variable-region epitopes revealed mutations that correspond to quite significant changes in amino acid properties. These mutations were prominent in the 11E1 epitope, located within a highly polymorphic region of 3D7-MSP2. Surprisingly, among the samples examined, the 11E1 sequence present in the reference 3D7-MSP2 variant was relatively infrequent, with the Q118E and E121K variants being most common. As the 11E1 mAb is able to recognise MSP2 on the surface of the parasite, it possible that antibodies targeting the 11E1 epitope may contribute to a protective response. Indeed, the selective pressure of protective antibodies targeting the 11E1 epitope may explain its high degree of diversity. There is strong evidence in literature suggesting that these polymorphisms are employed by pathogens as a mechanism of immune evasion (11, 43, 102). Determining key residues for 11E1 binding and the effect polymorphisms have on recognition may assist in the design of malaria vaccines able to protect against all variants of MSP2.

The second aspect of this chapter involves antibody fragments targeting the conserved C-terminal region of MSP2. Conserved epitopes in antigens are often desirable as vaccine targets as immune response towards them can elicit broad protection, bypassing the problems associated with polymorphic regions. Despite sharing an overlapping epitope, 4D11 mAb is able to recognise native MSP2 on the parasite, whereas 9H4 cannot. Understanding the structural determinants of this difference may provide insight into the conformation of native parasite MSP2. Numerous crystallisation screens were conducted, including hundreds of

crystallisation conditions, in efforts to obtain a structure of 4D11 and 9H4 in complex with their cognate epitope. Furthermore, a range of antibody fragments and their analogues was tested, the latest of which, a scFv without the flexible N-terminal purification tag, showed some promising hits and warrants optimisation in narrow screens. An alternative strategy to obtain crystallographic data for these complexes may entail the use of Fab fragments that are known to crystallise well. However, the cost and time associated with their production are not as favourable as for their smaller scFv and Fv counterparts. In the case of 4D11, the work here ultimately culminated in a crystal structure of 4D11 Fv in complex with its minimal binding epitope MSP2<sub>215-222</sub>, as described in Chapter 4 of this thesis.

## **2.5 Conclusions**

Six constructs corresponding to antibody fragments targeting variable-region epitopes were cloned and expressed. Although these constructs were unable to bind to MSP2, further optimisation of refolding or purification methods may restore functional binding. The analysis of polymorphisms in the variable region of MSP2 offers an alternative path to elicit a broadly protective immune response. Additionally, acquiring crystal structures of key MSP2 epitopes bound to their cognate antibodies is a major focus of this thesis. The work in this chapter has contributed to the progression of this project towards more efficient MSP2-based vaccine candidates.

# Chapter 3

Structural basis of epitope masking and strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate

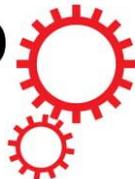
### 3.1 Chapter introduction

The previous Chapter described the production of variable region antibody fragments and our efforts to acquire structural information on the binding of 4D11 and 9H4 to their cognate conserved C-terminal epitope.

The work described here focuses on the conserved N-terminal region of MSP2, which is recognised by the mouse mAb, 6D8. The N-terminal region is well characterised and known to adopt an  $\alpha$ -helical conformation when interacting with lipid. It is likely that this conformation is present in its native form on the merozoite surface. 6D8 is unable to recognise parasite by IFA and western blot, suggesting that the epitope is inaccessible or masked in the native antigen. Here, we determine the minimal binding epitope of 6D8 by SPR by using a series of peptides covering MSP2<sub>11-23</sub>. A high-resolution crystal structure was obtained with 6D8 Fv in complex with the 9-residue peptide (NAYNMSIRR). This structure revealed why the 6D8-bound conformation was incompatible with the lipid-bound  $\alpha$ -helical structure. Also investigated in this chapter, is the strain-specific difference in binding of 6D8 to 3D7 and FC27 MSP2. FC27-MSP2 binds to 6D8 IgG with 5-fold higher affinity than 3D7-MSP2 even though the epitope is fully conserved. Two peptides were synthesised, extending the 6D8 epitope by five residues into the variable region of MSP2. These peptides were able to replicate the difference in binding of the full-length protein and were crystallised in complex with 6D8 Fv. Although these structures were unable to resolve the difference in binding affinity, NMR spectroscopy revealed subtle transient interactions between the variable region residues and 6D8.

A portion of the work in this chapter was carried out prior to my candidature, including the determination of the minimal binding epitope of 6D8. Furthermore, contributions involving NMR and SPR were performed by my colleagues Drs. Chris MacRaild and Rodrigo Morales. However, all crystal structures were solved and further analyses were completed by me within the PhD timeframe. This work provides a proof-of-principle that crystallographic data of anti-MSP2 antibody fragments, in complex with their cognate epitopes, can be acquired and that they can offer important insights into the basis of epitope recognition and masking on the parasite surface. Moreover, the experience obtained working with 6D8 greatly assisted in the work described in later chapters of this thesis. As the results of this Chapter have been published in *Scientific Reports*, they are presented in the format of a published article in the immediate section below.

# SCIENTIFIC REPORTS



OPEN

## Structural basis for epitope masking and strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate

Received: 23 December 2014

Accepted: 30 March 2015

Published: 12 May 2015

Rodrigo A. V. Morales<sup>1</sup>, Christopher A. MacRaild<sup>1</sup>, Jeffrey Seow<sup>1</sup>, Bankala Krishnarjuna<sup>1</sup>, Nyssa Drinkwater<sup>2</sup>, Romain Rouet<sup>3</sup>, Robin F. Anders<sup>4</sup>, Daniel Christ<sup>3</sup>, Sheena McGowan<sup>2</sup> & Raymond S. Norton<sup>1</sup>

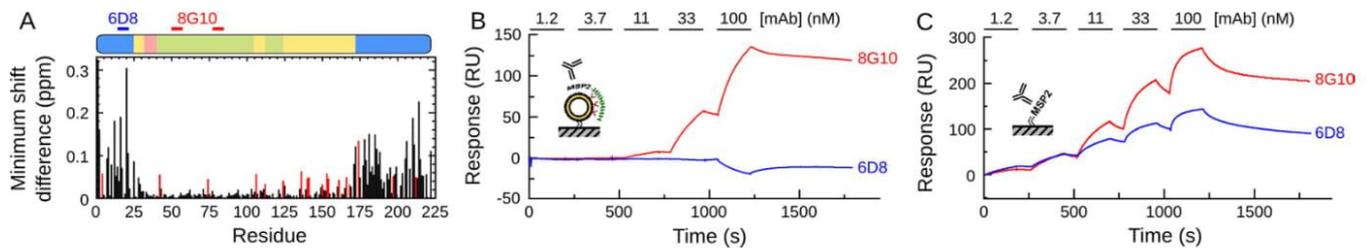
Merozoite surface protein 2 (MSP2) is an intrinsically disordered, membrane-anchored antigen of the malaria parasite *Plasmodium falciparum*. MSP2 can elicit a protective, albeit strain-specific, antibody response in humans. Antibodies are generated to the conserved N- and C-terminal regions but many of these react poorly with the native antigen on the parasite surface. Here we demonstrate that recognition of a conserved N-terminal epitope by mAb 6D8 is incompatible with the membrane-bound conformation of that region, suggesting a mechanism by which native MSP2 escapes antibody recognition. Furthermore, crystal structures and NMR spectroscopy identify transient, strain-specific interactions between the 6D8 antibody and regions of MSP2 beyond the conserved epitope. These interactions account for the differential affinity of 6D8 for the two allelic families of MSP2, even though 6D8 binds to a fully conserved epitope. These results highlight unappreciated mechanisms that may modulate the specificity and efficacy of immune responses towards disordered antigens.

Intrinsically disordered proteins are highly abundant in *Plasmodium* and related pathogenic genera<sup>1</sup>, and several have been identified as potential vaccine candidates for malaria<sup>2–7</sup> and other diseases<sup>8,9</sup>. For example, the protective effects of RTS,S, the most advanced malaria vaccine in clinical development, appear to be mediated by antibodies to the unstructured repeats of the *Plasmodium falciparum* circumsporozoite protein<sup>6,10</sup>. The interactions of intrinsically disordered proteins with their binding partners differ from those of structured proteins in terms of specificity and kinetics<sup>11</sup>, as well as the physicochemical properties of the interacting residues<sup>12</sup>, and are functionally important for a variety of cellular processes<sup>13</sup>. Nonetheless, the role of protein disorder in modulating the immune response against unstructured antigens remains poorly understood.

MSP2 is one of the most abundant and polymorphic glycosylphosphatidylinositol (GPI)-anchored proteins on the surface of the *P. falciparum* merozoite, the invasive blood-stage form of the malaria parasite<sup>14,15</sup>. All variants of MSP2 share conserved N- and C-terminal regions but fall into two allelic families, 3D7 and FC27, distinguished by tandem repeats and dimorphic flanking sequences within the central region of the protein<sup>14,16</sup>. Human vaccine trial subjects immunized with recombinant 3D7 MSP2

<sup>1</sup>Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC 3052, Australia. <sup>2</sup>Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia.

<sup>3</sup>Garvan Institute of Medical Research, Darlinghurst, Sydney, NSW 2010, Australia. <sup>4</sup>Department of Biochemistry, La Trobe Institute for Molecular Science, Melbourne, VIC 3086, Australia. Correspondence and requests for materials should be addressed to R.S.N. (email: Ray.Norton@monash.edu)



**Figure 1.** Membrane anchoring recapitulates the solvent exposure of the central variable region and the masking of an N-terminal epitope in recombinant MSP2. **(A)** NMR chemical shift perturbations of His-tagged MSP2 in DOGS-NiNTA/DPC micelles indicate marked differences at the conserved N and C-terminal portions of the protein. Red bars indicate chemical shift perturbations of Asp/Glu residues due to local pH at the micelle surface. Sequence features of MSP2 are shown above, with conserved regions in blue, repeats in green, allelic-specific dimorphic regions in yellow and polymorphic regions in pink. **(B)** SPR sensorgrams demonstrating the complete loss of N-terminal recognition of lipid-bound FC27 MSP2 by mAb 6D8 despite the dose-dependent binding of the strain-specific mouse antibody 8G10. The locations of the 6D8 and 8G10 epitopes are indicated in **(A)**. **(C)** The binding of mAb 8G10 and mAb 6D8 to MSP2 is concentration dependent in the absence of lipids.

mounted IgG responses capable of recognizing the parasite and significantly reducing parasitemia<sup>17</sup>. However, this vaccine preferentially targeted parasites expressing a 3D7-type MSP2 sequence, indicating that vaccine efficacy was mediated by strain-specific responses to MSP2<sup>18,19</sup>. Consistent with this result, the polymorphic region appears to be immunodominant in the natural immune response to MSP2<sup>20,21</sup> and some conserved region epitopes are cryptic on the parasite surface<sup>22,23</sup>. Understanding the mechanisms by which these epitopes are masked on the parasite surface should facilitate the design of MSP2-based antigens that direct the human immune response towards conserved epitopes and thus achieve strain-transcending protection.

Here, we use the mouse monoclonal antibody (mAb) 6D8, which recognizes a conserved N-terminal epitope on recombinant MSP2<sup>24</sup> but does not recognize the parasite surface,<sup>22</sup> to gain insights into epitope masking and strain specificity of the antibody response to MSP2. Using surface plasmon resonance (SPR) and NMR experiments, we show that recombinant MSP2, when C-terminally anchored to membrane mimetics, adopts a conformation that precludes the binding of mAb 6D8. X-ray crystal structures reveal the structural basis for this epitope masking. In addition, although the 6D8 epitope is fully conserved, its affinity for the antibody is modulated by transient interactions with flanking variable sequences. The ability of a variable region to confer strain specificity on a neighboring conserved epitope has important implications for our understanding of the immunogenic response to disordered vaccine candidates such as MSP2.

## Results

**Lipid interactions block recognition by 6D8.** The N-terminal conserved region of MSP2 was shown previously to undergo disorder-to-order transitions in the presence of dodecylphosphocholine (DPC) micelles<sup>25,26</sup>. These interactions, although weak, were sufficient to stabilize the 25-residue N-terminal peptide as an  $\alpha$ -helix, spanning at least residues 10–22. The possibility that this helical structure may contribute to epitope masking was explored with full-length MSP2 using a novel proxy of GPI anchoring in which a nickel-chelating lipid was used to bind the C-terminally His-tagged MSP2, mimicking the association of the MSP2 C-terminus with the lipid surface (Fig. 1). A comparison of <sup>1</sup>H-<sup>15</sup>N HSQC spectra of C-terminally His-tagged FC27 MSP2 in the presence and absence of dodecylphosphocholine (DPC) micelles containing 1 mol % of the nickel-chelating lipid 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-[[N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl] (DOGS-NiNTA) revealed substantial changes, involving both line-broadening and chemical shift changes, indicative of extensive interactions between MSP2 and the DPC micelle (Fig. S1). In our previous studies of the N-terminal region of MSP2<sup>25</sup> we have shown that extensive line-broadening in the presence of high DPC concentrations, as seen here, is preceded at lower DPC concentrations by weaker broadening together with chemical shift changes that are consistent with the adoption of helical structure. In contrast to those studies of the MSP2-DPC interaction in the absence of the DOGS-NiNTA tether, the interactions seen here are not restricted to the N-terminal region of MSP2. In the absence of assignments for the DPC-bound state, we analyzed these changes as minimum chemical shift changes from the DPC-free state (Fig. 1A), calculated from the distance between each assigned peak in the free MSP2 spectrum and the closest peak in the unassigned spectrum of MSP2 in complex with DOGS-NiNTA/DPC as  $[\Delta\delta H_N^2 + (\Delta\delta N/5)^2]^{1/2}$ . Changes to residues in the variable region of MSP2 are uniformly small and without significant line-broadening, and are restricted to Asp and Glu residues (red bars in Fig. 1A). Because these experiments are performed at pH 4.7, near the expected pK<sub>a</sub> of these residues, and because the directions of the chemical shift changes

Peptide	Sequence <sup>A</sup>	K <sub>d</sub> (nM) <sup>B</sup>
MSP2 <sub>11-23</sub> <sup>C</sup>	FINNAYNMSIRRS	6±3
MSP2 <sub>11-18</sub>	FINNAYNM	>1000
MSP2 <sub>12-19</sub>	INNAYNMS	>1000
MSP2 <sub>13-20</sub>	NNAYNMSI	>1000
MSP2 <sub>14-21</sub>	NAYNMSIR	>1000
MSP2 <sub>15-22</sub>	AYNMSIRR	87±21
MSP2 <sub>16-23</sub>	YNMSIRRS	>1000
MSP2 <sub>15-23</sub>	AYNMSIRRS	34±12
MSP2 <sub>14-22</sub>	NAYNMSIRR	6±2
MSP2 <sub>13-22</sub>	NNAYNMSIRR	14±5
MSP2 <sub>14-23</sub>	NAYNMSIRRS	16±7

**Table 1.** Mapping of the optimal binding region of the mAb 6D8 on the MSP2<sub>11-23</sub> sequence using a panel of overlapping synthetic peptides. <sup>B</sup>K<sub>d</sub> determined by SPR ± standard deviation of three replicate experiments. <sup>A</sup>Synthetic peptides were N-terminally acetylated and C-terminally amidated; <sup>C</sup>Broad epitope mapped previously<sup>22</sup>.

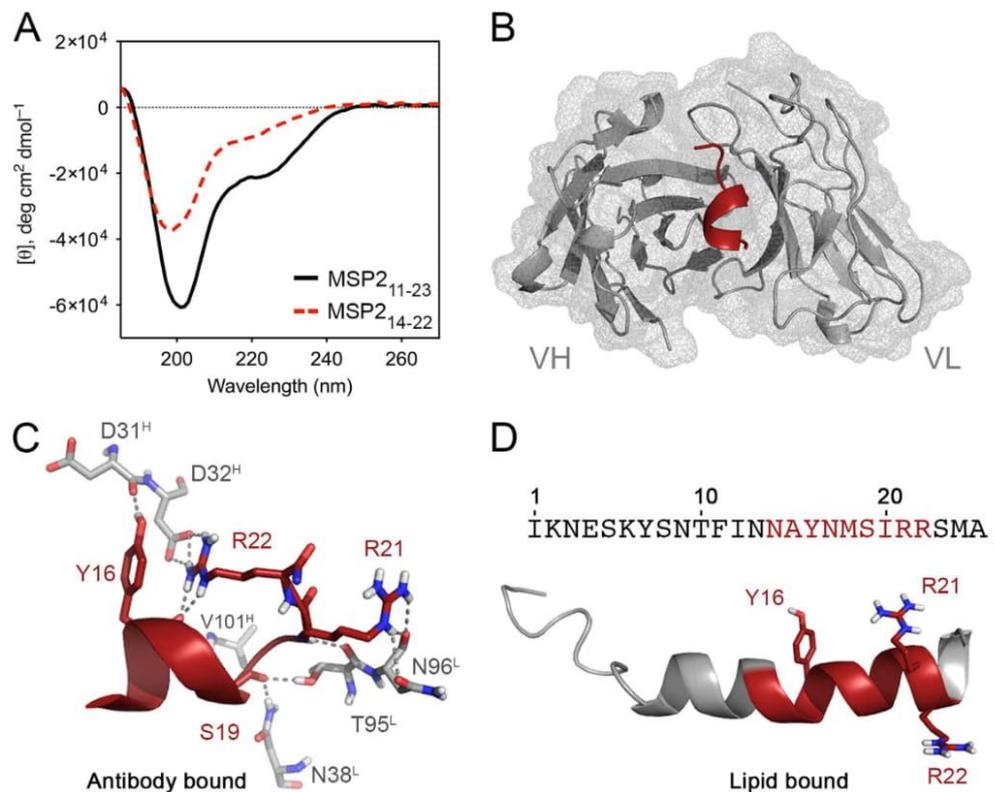
are consistent with those we have observed for lipid-free MSP2 upon small changes to the solution pH around this value (data not shown), it is likely that these changes are due to a small change in local pH at the micelle surface. In contrast, perturbations in both the N- and C-terminal conserved regions are larger, with chemical shift changes and extensive line-broadening affecting all residue types. These data indicate that both the N- and C-terminal conserved regions are involved in lipid interactions in this system, consistent with a possible role for these interactions in masking conserved region epitopes.

To test this possibility, we examined antibody binding to lipid-tethered MSP2 by SPR. Vesicles of POPC containing 1 mol% DOGS-NTA were immobilized on an L1 Biacore chip and loaded with Ni<sup>2+</sup> then C-terminally His-tagged MSP2, and the resulting surface was interrogated for antibody binding. In this assay, mAb 8G10, which recognizes the disordered variable region of FC27 MSP2 on the parasite surface<sup>22,27</sup>, bound strongly to vesicles decorated with C-terminally His-tagged FC27 MSP2. In contrast, mAb 6D8, which recognizes an N-terminal conserved epitope on recombinant MSP2<sup>24</sup> but fails to bind native MSP2 on the parasite surface<sup>22</sup>, did not interact with MSP2 on this artificial lipid surface (Fig. 1B). When MSP2 was immobilized using conventional amide coupling to the SPR chip in the absence of lipid, 6D8 binding was retained (Fig. 1C).

**Determination of the minimal 6D8 epitope.** The mAb 6D8 was first demonstrated to bind the conserved N-terminus of MSP2<sup>24</sup>, with the epitope localized to residues 11 to 23 (MSP2<sub>11-23</sub>)<sup>22</sup>. This region of full-length MSP2 undergoes structural changes in the presence of lipids as demonstrated previously<sup>25,26</sup> and confirmed here (Fig. 1A). To explore the structural basis of epitope masking, it was necessary to first define the minimal epitope of mAb 6D8. Six 8-residue peptides were prepared corresponding to a single amino acid frame shift over MSP2<sub>11-23</sub> (Table. 1) and tested against immobilized mAb 6D8 by SPR. MSP2<sub>15-22</sub> was the only 8-residue peptide of this initial panel to show detectable binding.

The affinity of MSP2<sub>15-22</sub> for mAb 6D8 was found to be 87 nM, which is more than ten-fold weaker than for MSP2<sub>11-23</sub> (6 nM), suggesting that residues beyond the mapped 8-residue region also contribute to antibody recognition. We tested this hypothesis with a second panel of synthetic peptides corresponding to single or double amino acid extensions at the N- and/or C-terminus of MSP2<sub>15-22</sub> (Table 1). Although a single C-terminal amino acid extension had no effect on binding, a single N-terminal extension was sufficient to restore binding to control levels (6 nM) (Table 1, Fig. S3). Consistent with the results from the 8-residue peptide series, these peptides demonstrate that removal of either Ala15 or Arg22 abolishes binding of MSP2 to 6D8, at up to 1 μM peptide concentration (Table. 1; compare MSP2<sub>15-23</sub> with MSP2<sub>16-23</sub>, and MSP2<sub>14-21</sub> with MSP2<sub>14-22</sub>).

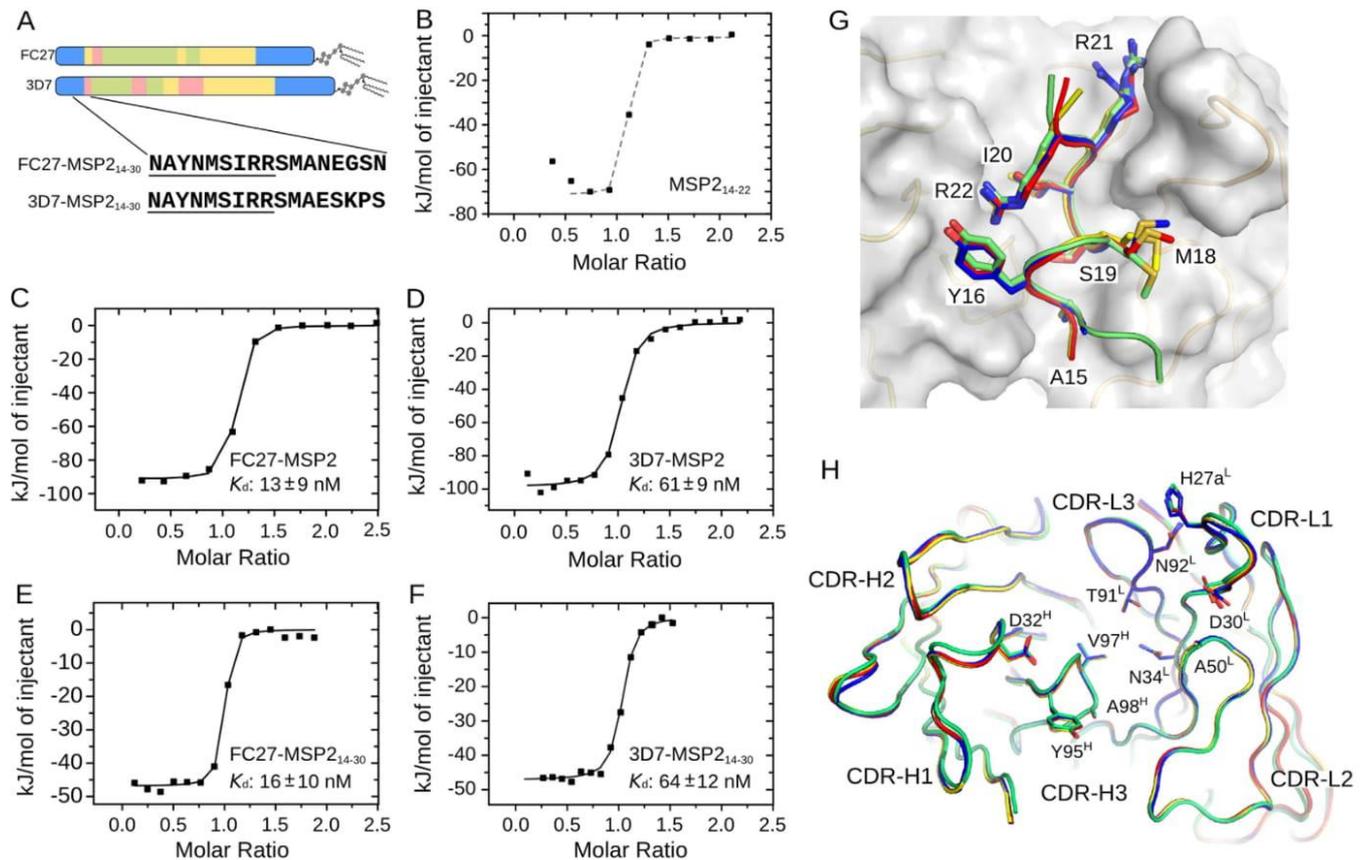
**Helical propensity of peptide epitopes.** We have demonstrated previously that the conserved N-terminal portion of MSP2 is unstructured in solution but adopts an α-helical conformation in the presence of lipids or organic solvents such as TFE<sup>25,28</sup>. This region of MSP2 encompasses the entire 6D8 epitope, suggesting that secondary structure formation could play a role in antibody binding<sup>22</sup>. The α-helical propensities of MSP2<sub>11-23</sub> and the equipotent minimal epitope, MSP2<sub>14-22</sub>, were investigated by circular dichroism. The minimal peptide epitope was substantially less helical than the 13-residue peptide (MSP2<sub>11-23</sub>) in aqueous solution (Fig. 2A) and over a range of TFE concentrations (Fig. S2). However, 6D8 binds these peptides with identical affinity, suggesting that recognition by 6D8 does not depend critically on the extent of helical conformation in the unbound epitope.



**Figure 2.** Crystal structure to 1.2 Å resolution of 6D8 Fv bound to the 9-mer peptide MSP2<sub>14–22</sub>. (A) Helical propensity of synthetic epitope-bearing peptides MSP2<sub>11–23</sub> and MSP2<sub>14–22</sub>. Both peptides showed ellipticity at 190, 208 and 220 nm, consistent with some helical content in water, with the longer peptide having significantly greater tendency for helical conformation. (B) Crystal structure of MSP2<sub>14–22</sub> at the paratope cleft formed by the antibody CDRs. (C) MSP2<sub>14–22</sub> adopts a single turn of helix stabilized by hydrogen bonds with residues Asp31<sup>H</sup>, Asp32<sup>H</sup>, Val101<sup>H</sup> and Asn38<sup>L</sup>. In solution, Arg22 contacts residues Thr91<sup>L</sup> and Asn92<sup>L</sup> while the guanidine group of Arg23 of the peptide bends towards the N-terminus of the peptide, establishing critical hydrogen bonds with Asp32 in the VH chain and the carbonyl of Tyr16. (D) The structure of the lipid-bound form of MSP2<sub>1–25</sub> indicating the relative orientation of residues Tyr16, Arg21 and Arg22. The  $\alpha$ -helical configuration of the lipid-bound MSP2<sub>1–25</sub> removes the backbone flexibility required for Arg22 to access Tyr16, which provides a structural rationale for the lack of binding of mAb 6D8 to MSP2 at the parasite membrane.

**The structure of lipid-bound MSP2 is incompatible with recognition by mAb 6D8.** To determine the structural basis for 6D8 epitope masking by lipid interactions, we have solved the X-ray crystal structure of the N-terminal epitope MSP2<sub>14–22</sub> in complex with the Fv region of mAb 6D8 (6D8 Fv) to 1.2 Å resolution (Fig. 2B). Our results indicate that, despite the moderate  $\alpha$ -helical propensity of the 6D8 epitope in MSP2 and the stabilizing effect of lipid interactions<sup>25</sup>, the antibody stabilizes only a single turn of helix in the peptide. This turn of helix spans residues Asn14–Met18, and breaks at Ser19, allowing the carbonyl of this residue to make hydrogen bonds to 6D8 through the side chains of Asn38<sup>L</sup> and Thr95<sup>L</sup> (6D8 is numbered according to Kabat<sup>29</sup>, with the chain indicated by a superscript)(Fig 2C). The helix is terminated by an unusually long capping motif involving two side-chain to main-chain hydrogen bonds, one between N<sup>ε</sup> of Arg22 and the carbonyl of Tyr16 and the other between the hydroxyl of Ser19 and the carbonyl of Ala14. This unusual motif positions Arg22 for simultaneous intra-molecular cation- $\pi$  interactions with the Tyr16 side chain and an inter-molecular salt bridge with Asp32<sup>H</sup> (Fig. 2C). The complex is further stabilized by hydrogen bonds between peptide residues Tyr16, Ser19, Arg22 and 6D8 residues Asp31<sup>H</sup>, Val101<sup>H</sup> and Asn96<sup>L</sup>, respectively, as well as van der Waals interactions involving the hydrophobic MSP2 residues Ala15, Met18 and Ile20.

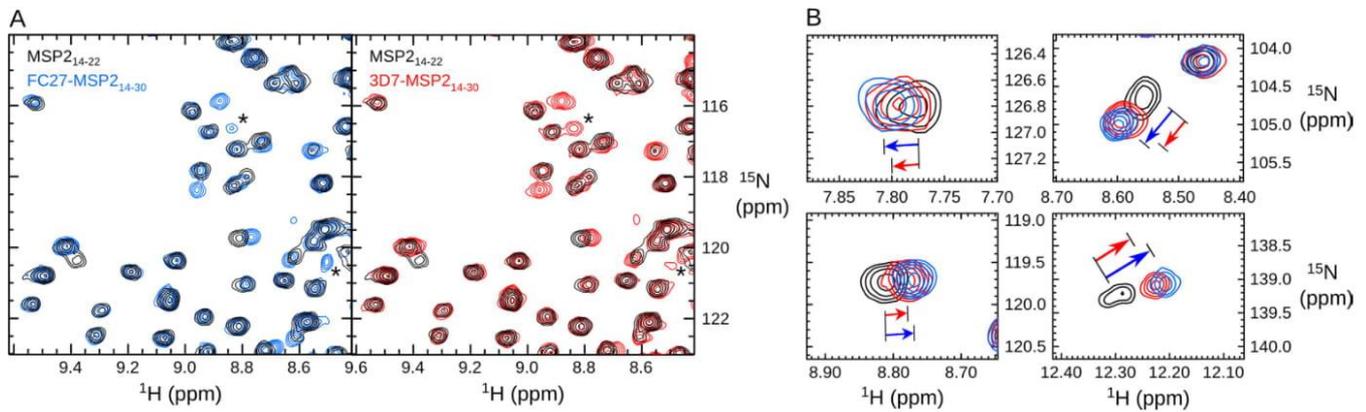
In contrast to the mostly unstructured MSP2 observed in solution<sup>26</sup>, the N-terminal region of MSP2 is predominantly helical in lipid membranes, with well-defined helical structure extending at least as far as Arg22 (Fig. 2D)<sup>25,28</sup>. In the lipid-bound conformation, the side chains of Arg22 and Tyr16 are separated by ~15 Å (Fig. 2D) and are clearly precluded from the interactions jointly made with CDR H1 in the 6D8 complex (Fig. 2C). Similarly, this helical conformation would also prevent the carbonyl of Ser19 from engaging CDRs L1 and L3. Furthermore, lipid interactions are expected to bury the hydrophobic face of this amphipathic helix, preventing access to the key hydrophobic residues that mediate recognition by



**Figure 3.** 6D8 displays strain-specific binding, even though it recognizes a conserved epitope sequence. (A) Schematic of the sequence features of MSP2, showing conserved regions (blue), repeats (green), allelic-specific dimorphic regions (yellow) and polymorphic regions (pink), together with the sequences of the two MSP2<sub>14–30</sub> peptides **b–e**: ITC titrations of MSP2<sub>14–22</sub> (B), full-length FC27 (C) and 3D7 (D) MSP2, FC27-MSP2<sub>14–30</sub> (E) and 3D7-MSP2<sub>14–30</sub> (F) into 6D8 scFv. Solid black lines in panels C–F are lines of best-fit, with corresponding  $K_d$  values shown with error estimates based on replicate titrations against 6D8 scFv and mAb ( $n = 3–5$ ). The dashed grey line in panel B illustrates the titration at the limit of determinable affinities (approx. 5 nM under these conditions). (G) Superposition of MSP2<sub>14–22</sub> (blue), MSP2<sub>11–23</sub> (green), 3D7 MSP2<sub>14–30</sub> (yellow) and FC27 MSP2<sub>14–30</sub> (red) in complex with the 6D8 Fv, with a representative molecular surface of the antibody fragment shown. (H) Backbone traces of 6D8 Fv bound to MSP2<sub>14–22</sub> (blue), MSP2<sub>11–23</sub> (green), 3D7 MSP2<sub>14–30</sub> (yellow) and FC27 MSP2<sub>14–30</sub> (red). CDRs and side chains making significant contacts with the antigen are shown.

6D8, including Tyr16, Met18 and Ile20. The structure of MSP2<sub>14–22</sub> bound to 6D8 Fv is therefore incompatible with the structure adopted by MSP2 on a lipid surface, which offers a clear explanation for the failure of 6D8 to recognize lipid-bound MSP2. This mechanism of binding was also observed in our 1.7 Å resolution crystal structure of MSP2<sub>11–23</sub> bound to 6D8 Fv (Fig. 3), with the three-residue N-terminal extension in this peptide making no additional contacts with the antibody. This finding is consistent with the SPR results, which showed that mAb 6D8 binds both peptides with similar affinities (Table 1).

**6D8 recognizes MSP2 in a strain-specific manner.** The interaction between 6D8 and MSP2, was examined further using isothermal titration calorimetry (ITC). The titration of 6D8 scFv with the minimal epitope, MSP2<sub>14–22</sub>, resulted in data at the upper limit of accessible affinities in this system (approximately 5 nM), consistent with the 6 nM  $K_d$  determined by SPR (Fig. 3B). Full-length recombinant MSP2 binds 6D8 scFv with lower affinity, permitting quantification by ITC. Even though 6D8 recognizes a fully conserved epitope in MSP2, the affinities of 6D8 scFv for full-length 3D7 and FC27 MSP2 differ by as much as 5-fold (Fig. 3C,D). To explore the basis of this strain-specificity, we synthesized peptides encompassing the minimal 6D8 epitope and extending C-terminally into the variable region. We found that peptides containing just the first five residues of the variable region of each allele (3D7 and FC27 MSP2<sub>14–30</sub>; Fig. 3A) were sufficient to fully replicate the weaker, strain-specific binding observed for full-length MSP2 (Fig. 3E,F). Identical affinities were observed when 6D8 IgG was used in place of the scFv (Fig. S4).



**Figure 4.**  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectra of 6D8 scFv reveal strain-specific transient interactions. (A) Comparison of the spectra of 6D8 in the MSP2<sub>14–22</sub> complex (black) with the FC27-MSP2<sub>14–30</sub> complex (blue) and with the 3D7-MSP2<sub>14–30</sub> complex (red) shows extensive spectral changes. A small number of peaks (\*) differ in intensity between the 3D7 and FC27 peptide complex, while others (B) undergo chemical shift changes that differ in magnitude, but not direction, between the 3D7 and FC27 peptides as indicated by the arrows.

**Strain-specific recognition is not explained by crystal structures of complexes.** In an attempt to determine the structural basis for this strain-specific recognition, we solved the X-ray structure of 6D8 Fv co-crystallized with the strain-specific peptides corresponding to residues 14 to 30 of 3D7 and FC27 (MSP2<sub>14–30</sub>) at resolutions of 1.4 Å and 1.6 Å, respectively. Surprisingly, the basic mode of binding observed in the MSP2<sub>14–22</sub>-6D8 complex was replicated in all complexes crystallized (Fig. 3G,H). The conformation of the MSP2 epitope is identical in all four structures, with backbone RMSD over residues 15–22 of 0.13 Å (heavy atom RMSD 0.54 Å) (Fig. 3G). Similarly, there were no differences in antibody conformation at complementarity determining regions (CDRs) that might offer an explanation for the 10-fold variation in affinity across the peptide-Fv complexes (Figs. 3H & S5).

In the case of FC27-MSP2<sub>14–30</sub>, no interpretable electron density was available to resolve residues C-terminal to Ser24, suggesting that this region fails to make stable interactions with 6D8 Fv, consistent with our determination of MSP2<sub>14–22</sub> as the minimal 6D8 epitope. The unresolved region includes all of the residues that differ between the FC27 and 3D7 MSP2<sub>14–30</sub> peptides (Fig. 3A). In contrast, the peptide is resolved until Ser28 in the 3D7-MSP2<sub>14–30</sub>-6D8 Fv complex, although with substantially elevated B-factor values for residues 24–28 (Fig. S6A). In fact, this C-terminal region appears to make only a single significant interaction with the 6D8 Fv, by way of hydrogen bonds between the side chain of Glu27 of 3D7-MSP2<sub>14–30</sub> and the side chains of Asp52<sup>H</sup> and Asn55<sup>H</sup> of 6D8 Fv. Plainly, this interaction depends on one of the carboxylate moieties of Glu27 or Asp52<sup>H</sup> being protonated. Although this is plausible at the crystallographic pH of 4.2, given the highly solvent-exposed nature of this site it is highly unlikely that the pK<sub>a</sub> values of these residues are sufficiently perturbed to allow protonation at physiological pH, where our affinity measurements were made. Moreover, this interaction is stabilized by crystal contacts with Arg18<sup>L</sup> from a crystallographically adjacent 6D8 molecule (Fig. S6B). In NMR spectra recorded at pH 6.5, we see no differences in peaks arising from sidechain NH<sub>2</sub> groups of 6D8 scFv bound to the peptides 3D7-MSP2<sub>14–30</sub> and FC7-MSP2<sub>14–30</sub> (Fig. S6C), suggesting the absence of strain-specific interactions involving Asn55<sup>H</sup> or other NH<sub>2</sub> groups under these conditions. Thus, we conclude that the interaction between Glu27 of 3D7-MSP2<sub>14–30</sub> and Asp52<sup>H</sup> in 6D8 is an artifact of crystallization. As this is the only crystallographically resolved interaction involving residues that differs between 3D7 and FC27 MSP2, the crystal structures provide no explanation for the observed five-fold difference in affinity.

**Strain-specific recognition is mediated by transient interactions.** We therefore turned to NMR spectroscopy to elucidate the basis of the differences in the affinity of 6D8 for 3D7 and FC27 MSP2 (Fig. 4). 6D8 scFv yielded high-quality  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectra, despite the tendency of scFv to dimerise at high concentrations<sup>30</sup>. Comparisons of the spectrum of 6D8 scFv bound to MSP2<sub>14–22</sub> with those of 6D8 scFv bound to FC27-MSP2<sub>14–30</sub> (Fig. 4A & S7) and 3D7-MSP2<sub>14–30</sub> (Fig. 4B & S7) indicated that the C-terminal extension of the peptide induced significant differences in the spectrum of 6D8, suggesting that MSP2 residues 24 to 30 interact with 6D8 scFv in solution, even though such interactions were not observed crystallographically. Moreover, more subtle differences were observed when the spectra of 6D8 scFv bound to FC27-MSP2<sub>14–30</sub> and 3D7-MSP2<sub>14–30</sub> were compared. For some peaks, these differences manifest as differences in intensity (compare peaks marked \* in Fig. 4A), while for other peaks small chemical shift differences are observed (Fig. 4B).

Despite the excellent quality of two-dimensional HSQC spectra for these samples, triple-resonance spectroscopy has proven more challenging, presumably due to scFv self-association. For this reason,

reliable scFv resonance assignments are not available for any of the 6D8-peptide complexes. While this precludes a direct structural interpretation of the spectral changes observed, analysis of these changes reveals rich variation in the dynamics of these interactions. On the one hand, peaks with variation in intensity without change in chemical shift (such as those marked \* in Fig. 4A) are indicative of interactions that are in exchange on a timescale of ms or slower. In this slow exchange case, the variation in peak intensity reflects the population of the exchanging state reported on by that peak.

In contrast, for a subset of peaks (Fig. 4B and Fig. S7), chemical shift changes are observed between the 6D8–MSP2<sub>14–22</sub> complex and those of the longer peptides. These changes vary in magnitude but not in direction, with changes seen for the FC27 complex being co-linear with, but systematically larger than those seen for the 3D7 complex by approximately 25%. This behavior is most simply explained by fast exchange between two states. In this fast exchange regime the observed chemical shift is a population-weighted average of the chemical shifts of the exchanging states. The interactions between 6D8 scFv and MSP2 residues C-terminal to Arg22 therefore include interactions that are exchanging on at least two distinct time-scales, and in which the interacting populations differ between the FC27 and 3D7 peptides. Importantly, all of the chemical shift changes seen here are small, suggesting that these interactions are populated to a relatively small extent, consistent with our inability to resolve these interactions in crystal structures.

## Discussion

Intrinsically disordered proteins, such as MSP2, are abundant components of pathogenic apicomplexan parasites as well as some viruses and other pathogens<sup>1,31</sup>. It has been argued that vaccine formulations containing disordered antigens may be advantageous because the effective antibody response against these proteins is less dependent on their native structure<sup>5</sup>. However, disordered proteins often undergo disorder-to-order transitions that are critical for their function and long-range transient interactions that modulate their interactions with binding partners<sup>11,13,32</sup>. The implications of the conformational dynamics of disordered antigens for the specificity and function of antibodies directed against them are largely unknown, despite posing important challenges for the use of these proteins in vaccine formulations.

During erythrocyte invasion, MSP2 is carried into the invaded cell and rapidly degraded<sup>33</sup>, unlike many merozoite surface antigens, which are released into the blood-stream<sup>34</sup>. As such, interactions between MSP2 and host antibodies occur exclusively at the parasite surface to which MSP2 is attached by a GPI anchor. It has been suggested that local structural changes observed at the N-terminus of MSP2 following lipid interactions<sup>25</sup> or oligomerization<sup>24</sup> at the parasite surface may modulate the accessibility of conserved N-terminal epitopes of MSP2. Our results show that a simple membrane-anchoring system employing C-terminally His-tagged MSP2 reproduces the masking of the distant 6D8 epitope on the parasite surface. We demonstrated previously that the N terminal region of MSP2 (MSP2<sub>1–25</sub>) binds lipid, adopting an extended helical conformation spanning at least residues 10–22<sup>25</sup>. The structure of the antibody-bound epitope differs in several important respects from the lipid-bound structure and offers a clear structural explanation for the loss of antibody recognition at the parasite surface. Taken together, our results indicate that lipid interactions alone or in conjunction with the formation of MSP2 oligomers<sup>24,35</sup> play an important role in determining the antigenic characteristics of MSP2 on the merozoite surface<sup>22</sup>.

Coupled folding and binding is often important in the functional interactions of disordered proteins<sup>36</sup> and has also been demonstrated for disordered antigens<sup>37–39</sup>. Indeed, functional properties of antibodies targeting disordered epitopes have been shown to depend on the specific conformation recognized by these antibodies<sup>38</sup>. The failure of 6D8 to recognize the parasite implies that MSP2 on the merozoite surface, in contrast to recombinant MSP2 in solution, lacks the conformational flexibility required to adopt the conformation bound by the 6D8 antibody. Human monoclonal antibodies targeting the conserved regions of MSP2 have not been characterized to date and polyclonal responses to those regions are significantly underrepresented in patients when compared to variable region antibodies<sup>40,41</sup>. Nonetheless, antibodies able to recognize the N-terminal region of MSP2 on the merozoite surface have been characterized<sup>23,33</sup>. These antibodies appear to recognize epitopes that overlap the 6D8 epitope characterized here, and as such it is likely that these antibodies recognise this region in a distinct conformation, more compatible with the structure of MSP2 on the parasite surface. Engineering the N-terminal region to favor this mode of antibody recognition could provide a route for the development of a strain-transcending MSP2 vaccine.

While disordered proteins often adopt an ordered conformation when interacting with binding partners, significant disorder can persist in these complexes<sup>13,42</sup>. Transient or ‘fuzzy’ interactions allow disordered epitopes to interact with multiple partners in biological roles otherwise not accessible to structurally rigid proteins<sup>13,32,43,44</sup>. However, these interactions have not, to our knowledge, been implicated in the modulation of antibody recognition. Here we have shown that transient interactions involving residues outside the structurally defined epitope of mAb 6D8 confer strain-specificity on the recognition of MSP2 by this antibody. ITC data demonstrate that a C-terminal extension from the minimal 6D8 epitope reduces the affinity of the complex, presumably as a consequence of the entropic costs associated with the restriction of conformational freedom that accompanies antibody binding. This entropic cost is partially offset by transient, fuzzy interactions<sup>13</sup>, shown by NMR to occur over a range of time-scales, between

the Fv domain and regions of MSP2 immediately C-terminal to the defined epitope. These interactions differ in extent between 3D7 and FC27 MSP2, thus accounting for the difference in affinity of mAb 6D8 for the two allelic forms.

The retention of proteins of low sequence complexity in apicomplexan genomes has long been postulated as an adaptive strategy against the host's immune system<sup>45–47</sup>. This hypothesis has been challenged because the majority of disordered apicomplexan proteins are not antigenic<sup>48</sup>. However, recent work has recognized distinct types of low-complexity sequence within the *P. falciparum* genome, one class of which includes several important antigens including MSP2<sup>49</sup>. Our finding that antibody recognition of a conserved epitope can be strain-specific as a consequence of transient interactions between antibody and antigen highlights a potentially important but largely unappreciated mechanism of immune evasion in this class of proteins. In disordered antigens, such as MSP2, conserved epitopes are flanked by disordered polymorphic regions<sup>1</sup>. Transient interactions between these regions and host antibodies may explain the difficulty in establishing broad neutralizing responses to conserved antigens and shed light on the possible benefit of these regions for the survival of *P. falciparum* in the human host.

## Materials and methods

**Recombinant protein expression.** The two allelic forms of MSP2 were prepared as described previously<sup>26,50</sup>. Soluble antibody fragments assembled as heterodimeric (Fv) and monomeric (scFv) polypeptide chains using 6D8 VH and VL sequences (Fig. S8) were produced in *Escherichia coli* and affinity purified using 3D7-MSP2-coated beads (Fig. S9) as described in the SI Materials and methods.

**Affinity measurements.** The affinity of 6D8 IgG and scFv for synthetic peptides and recombinant MSP2 was determined by ITC (Microcal ITC-200, GE Healthcare) and SPR (Biacore T200, GE Healthcare) using a Mouse Antibody Capture kit (GE Healthcare) (SI Materials and methods). Antibody binding to lipid-tethered MSP2 was assessed SPR using an L1 chip (GE Healthcare). Vesicles of POPC containing 1 mol % DOGS-NTA were adsorbed onto the chip, and loaded with Ni<sup>2+</sup> and then FC27 MSP2-6His.

**X-ray data collection and structure refinement.** Complexes of 6D8 Fv with synthetic peptides corresponding to the conserved N-terminal regions MSP2<sub>11–23</sub> and MSP2<sub>14–22</sub> and the allele-specific N-terminal peptides 3D7 MSP2<sub>14–30</sub> and FC27 MSP2<sub>14–30</sub> were crystallized by the hanging-drop method under the conditions listed in Table S1 (SI Materials and methods). Structures were solved by molecular replacement to resolutions ranging from 1.2 Å to 1.7 Å for the different complexes. Refinement statistics are provided in Table S2. The presence and identity of the peptide in each 6D8 crystal was confirmed by LC-MS.

**NMR spectroscopy.** <sup>13</sup>C, <sup>15</sup>N-labelled 6D8 scFv was dissolved in 20 mM sodium citrate, pH 6.5, 7% <sup>2</sup>H<sub>2</sub>O, in the absence of ligand to a concentration of approximately 100 μM. Peptides were added to 20% molar excess. <sup>15</sup>N-labelled FC27-MSP2-6His was dissolved to 200 μM concentration in 20 mM NaAcOH, pH 4.7 containing 100 mM DPC, 1 mM DOGS-NTA and 1 mM NiCl<sub>2</sub>. <sup>1</sup>H-<sup>15</sup>N HSQC spectra were recorded at 35 °C on a 600 MHz Bruker Avance III spectrometer equipped with a TCI triple-resonance cryoprobe. Spectra were processed using Topspin 3.2 (Bruker).

## References

- Feng, Z. P. *et al.* Abundance of intrinsically unstructured proteins in *P. falciparum* and other apicomplexan parasite proteomes. *Mol. Biochem. Parasitol.* **150**, 256–67 (2006).
- Raj, D.K. *et al.* Antibodies to PfSEA-1 block parasite egress from RBCs and protect against malaria infection. *Science* **344**, 871–877 (2014).
- Richards, J. S. *et al.* Association between naturally acquired antibodies to erythrocyte-binding antigens of *Plasmodium falciparum* and protection from malaria and high-density parasitemia. *Clin. Infect. Dis.* **51**, e50–60 (2010).
- Olugbile, S. *et al.* Vaccine potentials of an intrinsically unstructured fragment derived from the blood stage-associated *Plasmodium falciparum* protein PFF0165c. *Infect. Immun.* **77**, 5701–9 (2009).
- Yagi, M. *et al.* Protective epitopes of the *Plasmodium falciparum* SERA5 malaria vaccine reside in intrinsically unstructured N-terminal repetitive sequences. *PLoS One* **9**, e98460 (2014).
- Foquet, L. *et al.* Vaccine-induced monoclonal antibodies targeting circumsporozoite protein prevent *Plasmodium falciparum* infection. *J. Clin. Invest.* **124**, 140–4 (2014).
- Jepsen, M. P. *et al.* The malaria vaccine candidate GMZ2 elicits functional antibodies in individuals from malaria endemic and non-endemic areas. *J. Infect. Dis.* **208**, 479–88 (2013).
- Muster, T. *et al.* A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. *J. Virol.* **67**, 6642–7 (1993).
- Foucault, M. *et al.* UV and X-ray structural studies of a 101-residue long Tat protein from a HIV-1 primary isolate and of its mutated, detoxified, vaccine candidate. *Proteins* **78**, 1441–56 (2010).
- Dyson, H. J., Satterthwait, A. C., Lerner, R. A. & Wright, P. E. Conformational preferences of synthetic peptides derived from the immunodominant site of the circumsporozoite protein of *Plasmodium falciparum* by <sup>1</sup>H NMR. *Biochemistry* **29**, 7828–37 (1990).
- Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197–208 (2005).
- Uversky, V. N. Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta* **1834**, 932–51 (2013).
- Tompa, P. & Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **33**, 2–8 (2008).
- Smythe, J. A. *et al.* Structural diversity in the 45-kilodalton merozoite surface antigen of *Plasmodium falciparum*. *Mol. Biochem. Parasitol* **39**, 227–34 (1990).

15. Gilson, P. R. *et al.* Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite *Plasmodium falciparum*. *Mol. Cell Proteomics* **5**, 1286–99 (2006).
16. Fenton, B. *et al.* Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite surface antigen (MSA-2) of the malaria parasite *Plasmodium falciparum*. *Mol Cell Biol.* **11**, 963–71 (1991).
17. Genton, B. *et al.* A recombinant blood-stage malaria vaccine reduces *Plasmodium falciparum* density and exerts selective pressure on parasite populations in a phase 1-2b trial in Papua New Guinea. *J. Infect. Dis.* **185**, 820–7 (2002).
18. Flück, C. *et al.* Strain-specific humoral response to a polymorphic malaria vaccine. *Infect Immun.* **72**, 6300–5 (2004).
19. Flück, C. *et al.* Effect of the malaria vaccine Combination B on merozoite surface antigen 2 diversity. *Infect Genet Evol.* **7**, 44–51 (2007).
20. Taylor, R. R., Smith, D. B., Robinson, V. J., McBride, J. S. & Riley, E. M. Human antibody response to *Plasmodium falciparum* merozoite surface protein 2 is serogroup specific and predominantly of the immunoglobulin G3 subclass. *Infect Immun.* **63**, 4382–8 (1995).
21. Lawrence, N., Stowers, A., Mann, V., Taylor, D. & Saul, A. Recombinant chimeric proteins generated from conserved regions of *Plasmodium falciparum* merozoite surface protein 2 generate antiparasite humoral responses in mice. *Parasite Immunol.* **22**, 211–221 (2000).
22. Adda, C. G. *et al.* Antigenic characterization of an intrinsically unstructured protein, *Plasmodium falciparum* merozoite surface protein 2. *Infect Immun.* **80**, 4177–85 (2012).
23. Jones, G. L. *et al.* Immunological fine structure of the variable and constant regions of a polymorphic malarial surface antigen from *Plasmodium falciparum*. *Mol Biochem. Parasitol* **48**, 1–9 (1991).
24. Adda, C.G. *et al.* *Plasmodium falciparum* merozoite surface protein 2 is unstructured and forms amyloid-like fibrils. *Mol Biochem. Parasitol* **166**, 159–171 (2009).
25. MacRaid, C.A., Pedersen, M.Ø., Anders, R.F. & Norton, R.S. Lipid interactions of the malaria antigen merozoite surface protein 2. *Biochim. Biophys. Acta.* **1818**, 2572–2578 (2012.).
26. Zhang, X. *et al.* Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2. *J. Mol. Biol.* **379**, 105–121 (2008).
27. Epping, R. J., Goldstone, S. D. & Ingram, L. T. An epitope recognised by inhibitory monoclonal antibodies that react with a 51 kilodalton merozoite surface antigen in *Plasmodium falciparum*. *Mol. Biochem. Parasitol* **28**, 1–10 (1988).
28. Zhang, X. *et al.* Role of the helical structure of the N-terminal region of *Plasmodium falciparum* merozoite surface protein 2 in fibril formation and membrane interaction. *Biochemistry* **51**, 1380–7 (2012).
29. Kabat, E.A., Wu, T.T., Bilofsky, H., Reid-Miller & Perry, H. *Sequences of Proteins of Immunological Interest*, (National Institutes of Health, Bethesda, 1983).
30. Wilkinson, I.C. *et al.* High resolution NMR-based model for the structure of a scFv-IL-1beta complex: potential for NMR as a key tool in therapeutic antibody design and development. *J. Biol. Chem.* **284**, 31928–35 (2009).
31. Xue, B. *et al.* Structural disorder in viral proteins. *Chem. Rev.* **114**, 6880–911 (2014).
32. Ferreon, A. C., Ferreon, J.C., Wright, P. E. & Deniz, A. A. Modulation of allostery by protein intrinsic disorder. *Nature* **498**, 390–4 (2013).
33. Boyle, M.J. *et al.* Sequential processing of merozoite surface proteins during and after erythrocyte invasion by *Plasmodium falciparum*. *Infect Immun* **82**, 924–36 (2014).
34. Blackman, M. J. Proteases in host cell invasion by the malaria parasite. *Cell Microbiol.* **6**, 893–903 (2004).
35. Anders, R. F., Adda, C. G., Foley, M. & Norton, R. S. Recombinant protein vaccines against the asexual blood-stages of *Plasmodium falciparum*. *Hum. Vaccin.* **6**, 1–15 (2010).
36. Uversky, V. N., Oldfield, C. J. & Dunker, A. K. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit* **18**, 343–84 (2005).
37. Chu, H.M. *et al.* Two potential therapeutic antibodies bind to a peptide segment of membrane-bound IgE in different conformations. *Nat. Commun.* **5**, 3139 (2014).
38. Deng, L. *et al.* Discrete conformations of epitope II on the hepatitis C virus E2 protein for antibody-mediated neutralization and nonneutralization. *Proc. Natl. Acad. Sci. USA* **111**, 10690–5 (2014).
39. Ofek, G. *et al.* Elicitation of structure-specific antibodies by epitope scaffolds. *Proc. Natl. Acad. Sci. USA* **107**, 17880–7 (2010).
40. Felger, I., Steiger, S., Hatz, C., Smith, T. & Beck, H. P. Antigenic cross-reactivity between different alleles of the *Plasmodium falciparum* merozoite surface protein 2. *Parasite Immunol* **25**, 531–43 (2003).
41. Weisman, S. *et al.* Antibody responses to infections with strains of *Plasmodium falciparum* expressing diverse forms of merozoite surface protein 2. *Infect Immun* **69**, 959–67 (2001).
42. Fuxreiter, M. & Tompa, P. Fuzzy complexes: a more stochastic view of protein function. *Adv. Exp. Med. Biol.* **725**, 1–14 (2012).
43. Choi, U. B., McCann, J. J., Weninger, K. R. & Bowen, M. E. Beyond the random coil: stochastic conformational switching in intrinsically disordered proteins. *Structure* **19**, 566–76 (2011).
44. Garcia-Pino, A. *et al.* Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell* **142**, 101–11 (2010).
45. Kemp, D. J., Coppel, R. L. & Anders, R. F. Repetitive proteins and genes of malaria. *Annu. Rev. Microbiol.* **41**, 181–208 (1987).
46. Hughes, A. L. The evolution of amino acid repeat arrays in Plasmodium and other organisms. *J. Mol. Evol.* **59**, 528–35 (2004).
47. Ferreira, M. U., Ribeiro, W. L., Tonon, A. P., Kawamoto, F. & Rich, S. M. Sequence diversity and evolution of the malaria vaccine candidate merozoite surface protein-1 (MSP-1) of *Plasmodium falciparum*. *Gene.* **304**, 65–75 (2003).
48. DePristo, M. A., Zilversmit, M. M. & Hartl, D. L. On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins. *Gene.* **378**, 19–30 (2006).
49. Zilversmit, M. M. *et al.* Low-complexity regions in *Plasmodium falciparum*: missing links in the evolution of an extreme genome. *Mol. Biol. Evol.* **27**, 2198–209 (2010).
50. MacRaid, C. A. *et al.* Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2. *PLoS One* **10**, e0119899 (2015).

## Acknowledgements

We thank Xiaopeng Ge for sequencing 6D8 and John Gehman for assisting with the acquisition of CD spectra of MSP2 peptides. This work was partially supported by the Indo-Australian Biotechnology fund (BF050053) and the National Health and Medical Research Council of Australia (APP1042520). RSN acknowledges fellowship support from the National Health and Medical Research Council of Australia. SM is an Australian Research Council Future Fellow (FT100100690). We thank the Monash Macromolecular Crystallization Facility for technical assistance and the Australian Synchrotron (MX-1 & MX-2) and the beam-line scientists for beam time and technical assistance.

### Author Contributions

R.A.V.M, C.A.M, R.F.A, D.C, S.M and R.S.N designed experiments; R.A.V.M, C.A.M, J.S, K.B, N.D and R.R performed experiments; R.A.V.M, C.A.M, J.S, N.D, S.M and RSN analyzed data; R.A.V.M, C.A.M and R.S.N wrote the paper with input from all authors.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Morales, R. A. V. *et al.* Structural basis for epitope masking and strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate. *Sci. Rep.* 5, 10103; doi: 10.1038/srep10103 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

## Supplementary Information

### Structural basis for epitope masking and strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate

Rodrigo A. V. Morales<sup>1</sup>, Christopher A. MacRaid<sup>1</sup>, Jeffrey Seow<sup>1</sup>, Bankala Krishnarjuna<sup>1</sup>,  
Nyssa Drinkwater<sup>2</sup>, Romain Rouet<sup>3</sup>, Robin F. Anders<sup>4</sup>, Daniel Christ<sup>3</sup>, Sheena McGowan<sup>2</sup> and  
Raymond S. Norton\*<sup>1</sup>

1. Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University,  
Parkville, VIC 3052, Australia

2. Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC  
3800, Australia

3. Garvan Institute of Medical Research, Darlinghurst, Sydney, NSW 2010, Australia

4. Department of Biochemistry, Institute for Molecular Science, La Trobe University  
Melbourne, VIC 3086, Australia

\* Corresponding author (Ray.Norton@monash.edu)

## Supplementary Methods

*Recombinant expression and purification of 3D7 and FC27 MSP2.* Two allelic forms of 3D7 and FC27 MSP2 (Genebank accession numbers JN248383 and JN248384, (<http://www.ncbi.nlm.nih.gov>)) were produced recombinantly in *Escherichia coli* BL21(DE3) Gold (Stratagene) cells under different strategies. FC27 MSP2 was expressed in high yield without tags or carrier proteins, as described,<sup>1,2</sup> but this strategy was not useful for 3D7 MSP2 owing to poor expression yields. Optimal expression of 3D7 MSP2 was achieved using a codon-optimised construct (Life Technologies) in a thioredoxin-6xHis expression system (pET32a, Millipore Pty) as described in (MacRaild *et al.* in press).

*Recombinant expression and purification of anti-MSP2 antibody fragments.* Variable heavy and light chain sequences corresponding to the anti-MSP2 6D8 mAb were PCR-amplified directly from the 6D8 mouse hybridoma cell line<sup>3,4</sup> using sequence-specific primers.<sup>5</sup> Primer-derived errors found at the first conserved N-terminal residues of the VH sequence were corrected based on the closely-related V<sub>H</sub> mouse germline, IGHV1-78 (<http://www.imgt.org/>) (Fig. S8). Variable heavy (V<sub>H</sub>) and light (V<sub>L</sub>) chains are available under accession codes KM393285 and KM393286, respectively.

Soluble antibody fragments were assembled as heterodimeric (Fv) and monomeric (scFv) polypeptide chains using 6D8 VH and VL sequences optimised for expression in *E. coli* BL21-Gold as described previously.<sup>6</sup> Synthetic genes were purchased from Genscript and cloned into pET12a (Merck Millipore) for periplasmic expression in *E. coli* were grown in presence of tetracycline (15 µg/mL) and ampicillin (100 µg/mL). Protein expression was carried out under IPTG induction for 20 h at room temperature in LB or M9 minimal media supplemented with <sup>15</sup>N-ammonium chloride and <sup>13</sup>C-glucose at 1 g/L and 4 g/L respectively for isotopic labeling.<sup>7</sup> Periplasm contents were extracted with gentle treatment of the cell mass with sucrose buffer (100 mM Tris-HCl, 1 mM EDTA and 20% sucrose (w/v) pH 8.0) followed by 5 mM MgCl<sub>2</sub> at 4 °C.

The 6D8 antibody fragments were affinity purified from the periplasmic fraction using 3D7 MSP2-bound beads prepared by incubation of full-length recombinant 3D7 MSP2 with hydroxylsuccinimidyl-coated Sepharose beads (Sigma-Aldrich) according to the manufacturer's instructions. Bound antibody fragments were washed extensively with 50 mM phosphate buffer pH 7 and eluted from the MSP2 beads with 100 mM glycine, pH 2.7. Samples were neutralized with 1M Tris, concentrated in a 10 kDa MWCO centrifuge filter (AMICON Ultra-0.5, Merck-Millipore) and dialyzed against 50 mM ammonium bicarbonate prior to lyophilization and storage at -80 °C. The purity of each antibody fragment was assessed by SDS-PAGE and LC-MS (Fig. S9).

*Peptide synthesis.* Standard Fmoc-protected amino acids (Phe, Ile, Asn(Trt), Ala, Tyr(tBu), Met, Ser(tBu) and Arg(Pbf)), Rink amide polystyrene resin and O-(1H-6-chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HCTU) were obtained from Chem-Impex. Organic solvents dimethylformamide (DMF), acetonitrile and diethyl ether were obtained from Merck Pty. Trifluoroacetic acid (TFA) was obtained from Peptides International. Piperidine, diisopropylethylamine (DIPEA), triisopropylsilane (TIPS) and acetic anhydride were obtained from Sigma-Aldrich.

A panel of synthetic peptides corresponding to the conserved N-terminal regions MSP2<sub>11-23</sub>, MSP2<sub>11-18</sub>, MSP2<sub>12-19</sub>, MSP2<sub>13-20</sub>, MSP2<sub>14-21</sub>, MSP2<sub>15-22</sub>, MSP2<sub>16-23</sub>, MSP2<sub>14-23</sub>, MSP2<sub>14-22</sub>, MSP2<sub>13-23</sub>, MSP2<sub>14-23</sub> and the allele-specific regions 3D7 MSP2<sub>14-30</sub> and FC27 MSP2<sub>14-30</sub> were prepared in-house using standard 9-fluorenylmethoxycarbonyl (Fmoc) solid-phase chemistry.<sup>8</sup> Briefly, MSP2 peptides were assembled over Rink amide polystyrene resin (0.1

mM scale) using a three-fold equivalent of Fmoc-protected amino acid, HCTU and DIPEA in DMF per cycle. Each amino acid coupling was carried out for 30 min at room temperature with continuous shaking. Chain deprotection was carried out in 50% piperidine in DMF for 2 min. Each step in the cycle was terminated with an extensive DMF wash. Peptides were N-terminally acetylated after the last coupling and deprotected, using a three-fold equivalent of acetic anhydride and DIPEA in DMF.

The fully assembled peptides were dried under vacuum and cleaved for 3 h in a mixture of TFA:TIPS:water (95:2.5:2.5 v/v). The cleaved material was precipitated in cold diethyl ether overnight at -20 °C. The insoluble peptide material was spun down at 4,000 rpm for 30 min at 4 °C and the pellet washed twice in cold diethyl ether prior to removal of the organic phase. The crude peptide mixture was resuspended in 50% acetonitrile/0.1% TFA, filtered and freeze-dried prior to HPLC purification. All MSP2 peptides with the exception of MSP2<sub>12-19</sub> and MSP2<sub>13-20</sub> were resuspended in solvent A (0.1% TFA in water) to a final concentration of 2 mg/mL for HPLC purification. Peptides MSP2<sub>12-19</sub> and MSP2<sub>13-20</sub> showed a high tendency to aggregate and had to be resuspended in 10 mM acetic acid at concentrations of 1 mg/mL. MSP2 peptides were purified on a reverse-phase C18 column (Zorbax, 10 x 300 mm) using a linear gradient of 5 to 40% of solvent B (90% acetonitrile / 9.9% water / 0.1% TFA) against solvent A (0.1% TFA in water) over 40 min. The purity of MSP2 peptides was assessed by mass spectrometry (LC-MS).

*Affinity measurements.* Binding affinity constants of 6D8 antibodies (IgG and scFv) for synthetic peptides and recombinant allelic forms of MSP2 were determined by isothermal titration calorimetry (Microcal ITC-200, GE Healthcare) using either murine mAb 6D8 produced and affinity-purified by the Walter and Eliza Hall Institute antibody facility<sup>3</sup> or 6D8 scFv. Titrations were performed in 20 mM sodium phosphate, 150 mM NaCl, pH 7.4, at 25°C. Typical antibody concentrations were 10 μM (IgG) or 20 μM (scFv), with MSP2 or peptide titrated from 200 μM stocks. Control titrations of MSP2 into buffer were performed, and the resulting heats of dilution subtracted from the corresponding titration into antibody. Where low peptide solubility precluded ITC measurements, affinity was estimated by surface plasmon resonance (SPR) (Biacore T200, GE Healthcare) using a Mouse Antibody Capture kit (GE Healthcare). Synthetic peptides were resuspended in running buffer (20 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05 % Tween20, pH 7.4) and injected over the 6D8-captured surface. Binding affinity was estimated from the concentration dependence of steady-state responses observed.

*Antibody binding to lipid-bound MSP2.* For NMR samples, micelles of dodecylphosphocholine (DPC) were prepared at 100 mM in 20 mM NaAcOH, pH 4.7 and doped with 1 mol % 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-[(N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl] (DOGS-NTA) (Avanti Polar Lipids). NiCl<sub>2</sub> was added to 1 mM and 7 % <sup>2</sup>H<sub>2</sub>O added, and this solution was used to dissolve <sup>15</sup>N-FC27-MSP2-6His to a final concentration of 0.2 mM. <sup>1</sup>H-<sup>15</sup>N SOFAST HMQC spectra<sup>9</sup> were recorded at 25°C on a 600 MHz Bruker Avance III spectrometer equipped with a TCI triple-resonance cryoprobe. Spectra were processed using Topspin (Bruker). Minimum chemical shift differences are calculated from the distance between each assigned peak in the free MSP2 spectrum and the closest peak in the unassigned spectrum of MSP2 in complex with DOGS-NiNTA/DPC as  $[\Delta\delta H_N^2 + (\Delta\delta N/5)^2]^{1/2}$ .

For SPR assays of antibody binding, lipid vesicles were prepared from 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and DOGS-NTA at a 100:1 molar ratio. The lipids were mixed in chloroform:methanol (1:1), dried, and resuspended to approximately 5 mM lipid in buffer (20 mM HEPES, 150 mM NaCl, pH 7.4), then sonicated until a stable,

optically clear solution was obtained. Freshly-prepared vesicles were coated onto reference and active cells of a Biacore L1 chip (GE Healthcare) by injection of a 0.5 mM stock at 2  $\mu\text{L}/\text{min}$  for 15 min, washed with a 30 s pulse of 50 mM NaOH (30  $\mu\text{L}/\text{min}$ ) and loaded with  $\text{Ni}^{2+}$  with a 3 min injection of 50 mM  $\text{NiSO}_4$  (2  $\mu\text{L}/\text{min}$ ). The prepared surface was rinsed for 2 min (30  $\mu\text{L}/\text{min}$  running buffer) before MSP2 was loaded onto the lipid surface of the active cell with a 5 min injection of 1.5  $\mu\text{M}$  FC27 MSP2-6His at 2  $\mu\text{L}/\text{min}$ . Finally, the surface was allowed to stabilize under flow (30  $\mu\text{L}/\text{min}$  running buffer) for 10 min before antibody binding was assessed, typically with five 3-min injections of increasing antibody concentration. The chip surface was regenerated with a 30 s injection of 30 mM NaOH in 40 % (v/v) isopropanol between cycles. In control experiments, identical injections of mAb were performed over MSP2 immobilized to a comparable level on a CM5 chip by amide coupling following the manufacturer's instructions.

*Antibody-peptide co-crystallization, X-ray data collection and structure refinement.* Crystals of antibody-peptide complexes were obtained using 6D8 Fv bound to synthetic MSP2 peptides corresponding to the conserved N-terminal region MSP2<sub>14-22</sub> and MSP2<sub>11-23</sub> and the allele-specific N-terminal regions 3D7 MSP2<sub>14-30</sub> and FC27 MSP2<sub>14-30</sub>. Antibodies and synthetic peptides were conjugated at 1:1.5 (mol/mol) ratio in crystallization buffer (20 mM Tris-HCl, 100 mM NaCl, pH 7.0) for 1 h at room temperature. The unbound peptide and high molecular weight aggregates were removed by gel filtration (Superdex 75 10/300 GL, GE Healthcare). The complex was concentrated to 20-30 mg/mL by gentle centrifugation (Amicon Ultra 3 kDa, Merck-Millipore Pty) and cleared of precipitants by centrifugation at 10,000 rpm for 10 min, then filtered through a 0.22  $\mu\text{m}$  spin filter at 4°C. Antibody concentrations were determined based on the absorbance at 280 nm ( $A^{1\%}_{280\text{nm}} = 1.6$ ) by NanoDrop (Thermo-Fisher Pty Ltd).

The crystal complexes were grown using the hanging drop vapour diffusion method, with 1:1 (v/v) ratio of protein to mother liquor (0.5 mL well volume). Large, plate-shaped crystals appeared overnight in the presence of polyethylene glycol 8000, sodium acetate buffer and sodium chloride as shown in Supplementary Table S2. Crystals were cryo-protected by the addition of 10 % glycerol prior to data collection. Datasets were collected at 100 K at the Australian Synchrotron Macro crystallography MX1 beamline 3BM1 and Micro crystallography MX2 beamline 3ID1 at resolutions ranging from 1.2 Å to 1.7 Å for the different complexes. Diffraction images were processed using XDS<sup>10</sup> and Aimless from the CCP4 suite.<sup>11</sup> 5% of each dataset was flagged for calculation of  $R_{\text{free}}$ <sup>12</sup> with neither a sigma nor a low-resolution cut-off applied to the data. A summary of statistics is provided in Supplementary Table S3.

Structure determination proceeded using the Molecular Replacement (MR) method and the program PHASER.<sup>13</sup> An initial search model was constructed from the crystal structure of a mouse single chain Fv (PDB ID 3GM0) by removing the ligand from the search model. The 6D8 Fv + MSP2<sub>14-22</sub> dataset showed that a single clear peak was evident in both the rotation and translation functions and packed well within the asymmetric unit. Together with the unbiased features in the initial electron density maps, the correctness of the MR solution was confirmed. Initial electron density maps also clearly showed unbiased features of the MSP2<sub>14-22</sub> peptide ligand between the two protein chains. Automated model building was performed using the program ARP/warp.<sup>14</sup> All subsequent model building and structural validation was undertaken using Phenix<sup>15,16</sup> and COOT.<sup>17</sup> Solvent molecules were added only if they had acceptable hydrogen-bonding geometry contacts of 2.5 to 3.5 Å with protein atoms or with existing solvent and were in good  $2F_o - F_c$  and  $F_o - F_c$  electron density.

The remaining three structures (6D8 Fv + MSP2<sub>11-23</sub>, 6D8 Fv + 3D7 MSP2<sub>14-30</sub> and 6D8 Fv + FC27 MSP2<sub>14-30</sub>) were solved using the model of 6D8 Fv alone (excluding ligand and

solvent) as the MR probe. Hydrogen bonds (excluding water-mediated bonds) and salt bridges were calculated using PDBePISA.<sup>18</sup>

*Circular dichroism.* The secondary structure of the MSP2 peptides MSP2<sub>11-23</sub> and MSP2<sub>14-22</sub> was determined by CD spectroscopy. CD spectra were acquired between 185 and 280 nm at 25 °C on a Jasco J-815 spectropolarimeter (Jasco, ATA Scientific, Japan) using a 1-mm pathlength quartz cell (Starna, Hainault, United Kingdom). Peptide stock solutions were made to a concentration of 100 µM in 10 mM acetic acid. Peptides were tested at a final concentration of 50 µM with increasing amounts of trifluoroethanol (Sigma) (0, 10, 20 and 30% v/v). Spectra were acquired at a rate of 50 nm/min with 1 nm data intervals, 4 s integration time and a 1-nm slit width. Three accumulations were made and averaged to reduce the noise. Signal was recorded in millidegrees and later converted to ellipticity values  $[\theta]$  using the formula  $[\theta] = (\theta \times R_w / L \times C)$ , where  $\theta$  is the recorded ellipticity in millidegrees,  $R_w$  is the mean residue weight,  $L$  is the pathlength in millimeters and  $C$  is the concentration in mg/mL.

**Table S1.** Conditions used for the crystallization of antibody-peptide complexes.

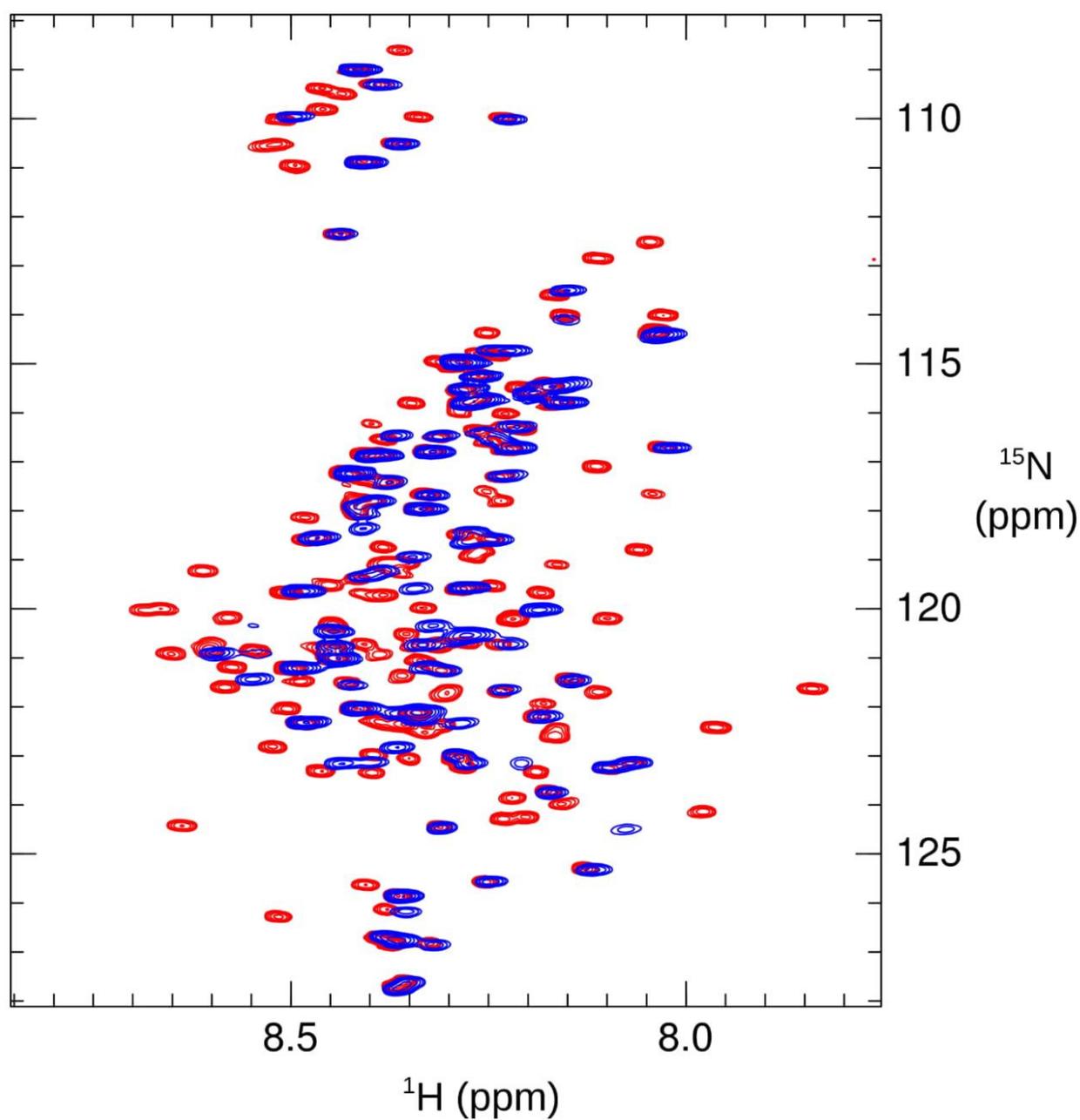
Crystal	Crystallisation Conditions	Dataset Resolution (Å)
6D8 Fv + MSP2 <sub>14-22</sub>	22 % (w/v) PEG8000 0.1 M sodium acetate (pH 4.7) 0.2 M sodium chloride	1.2
6D8 Fv + 3D7 MSP2 <sub>14-30</sub>	30 % (w/v) PEG8000 0.1 M sodium acetate (pH 4.2) 0.2 M sodium chloride	1.4
6D8 Fv + FC27 MSP2 <sub>14-30</sub>	30 % (w/v) PEG8000 0.1 M sodium acetate (pH 4.4) 0.2 M sodium chloride	1.6
6D8 Fv + MSP2 <sub>11-23</sub>	20 % (w/v) PEG8000 0.1 M sodium acetate (pH 4.2) 0.2 M sodium chloride	1.7

**Supplementary Table S2.** Data collection and refinement statistics.

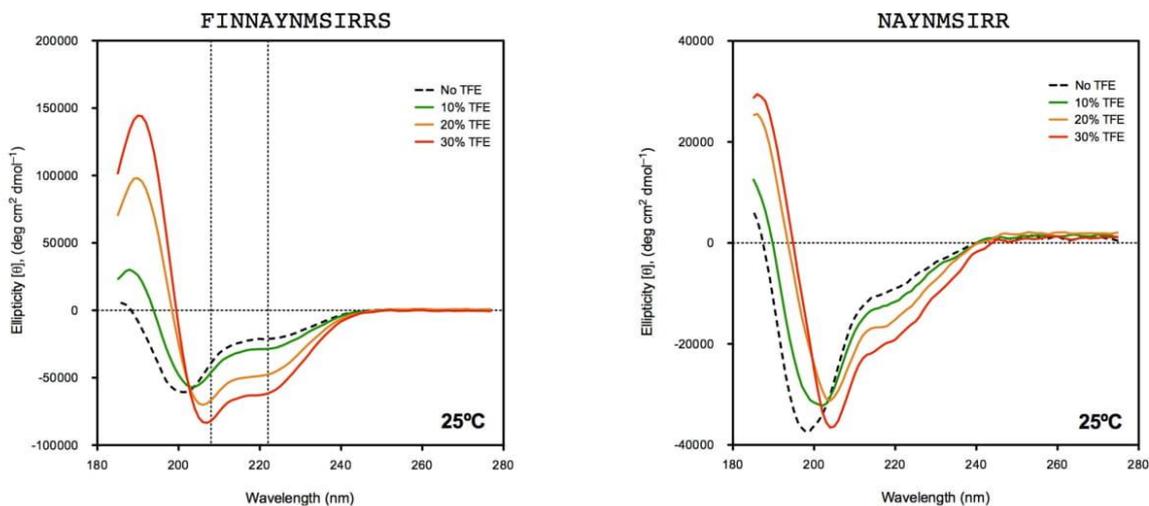
Data collection	6D8 Fv + MSP2 <sub>14-22</sub>	6D8 Fv + 3D7 MSP2 <sub>14-30</sub>	6D8 Fv + FC27 MSP2 <sub>14-30</sub>	6D8 Fv + MSP2 <sub>11-23</sub>
Space Group	P21	P2	P2	P212121
Cell dimensions (Å)	a=41.9 b=60.7 c=43.7, $\alpha = 90^\circ$ ; $\beta = 106.7^\circ$ ; $\gamma = 90^\circ$	a=42.1 b=60.3 c=43.5 $\alpha = 90^\circ$ ; $\beta = 107.0^\circ$ ; $\gamma = 90^\circ$	a=42.1 b=60.2 c=43.4 $\alpha = 90^\circ$ ; $\beta = 107.1^\circ$ ; $\gamma = 90^\circ$	a=35.2 b=64.3 c=89.9 $\alpha = \beta = \gamma = 90^\circ$
Resolution (Å)	41.83 – 1.21 (1.27 – 1.21)	34.39 – 1.35 (1.43 – 1.35)	34.40 – 1.58 (1.67 – 1.58)	52.3 – 1.7 (1.76 – 1.7)
Total reflections	1560904	258993	213302	315156
Unique reflections	62158	44204	28505	23236
Multiplicity	25.1 (3.6)	5.9 (3.5)	7.5 (3.6)	13.6 (12.3)
Data Completeness (%)	96.8 (83.3)	97.4 (95.8)	99.9 (99.7)	100.0 (100.0)
$\langle I/\sigma_I \rangle$	32.5 (1.7)	16.3 (2.1)	21.1 (6.8)	9.4 (3.0)
CC(1/2)	0.904 (0.578)	0.999 (0.707)	0.998 (0.971)	0.991 (0.680)
R <sub>p</sub> (%) <sup>b</sup>	6.4 (51.9)	22 (366)	21 (96)	17.0 (148.0)
PDB	4QYO	4QY8	4QXT	4R3S
<b>Structure refinement</b>				
Non hydrogen atoms				
Protein	1821	1829	1808	1847
Solvent (HOH)	377	369	326	370
R <sub>free</sub> (%)	17.0	16.5	17.8	20.0
R <sub>cryst</sub> (%)	15.1	15.2	15.6	15.2
CC*	0.974	0.906	0.916	0.735
Bond lengths (Å)	0.008	0.005	0.008	0.017
Bond angles (°)	1.25	1.10	1.18	1.61
Ramachandran plot				
Favoured (%)	98	97	98	98
Outliers (%)	-	-	-	-
B factors (Å <sup>2</sup> )				
Mean protein	13.5	15.7	17.0	9.10
Mean water molecule	30.6	29.3	32.0	23.8
Molprobit Score <sup>c</sup>	1.12	0.93	0.76	1.24
	97th percentile (N=1422, 1.208 Å ± 0.25 Å)	100th percentile (N=2978, 1.353 Å ± 0.25 Å)	100th percentile (N=6761, 1.580 Å ± 0.25 Å)	97th percentile (N=9248, 1.700 Å ± 0.25 Å)

a Values in parentheses refer to the highest resolution shell.

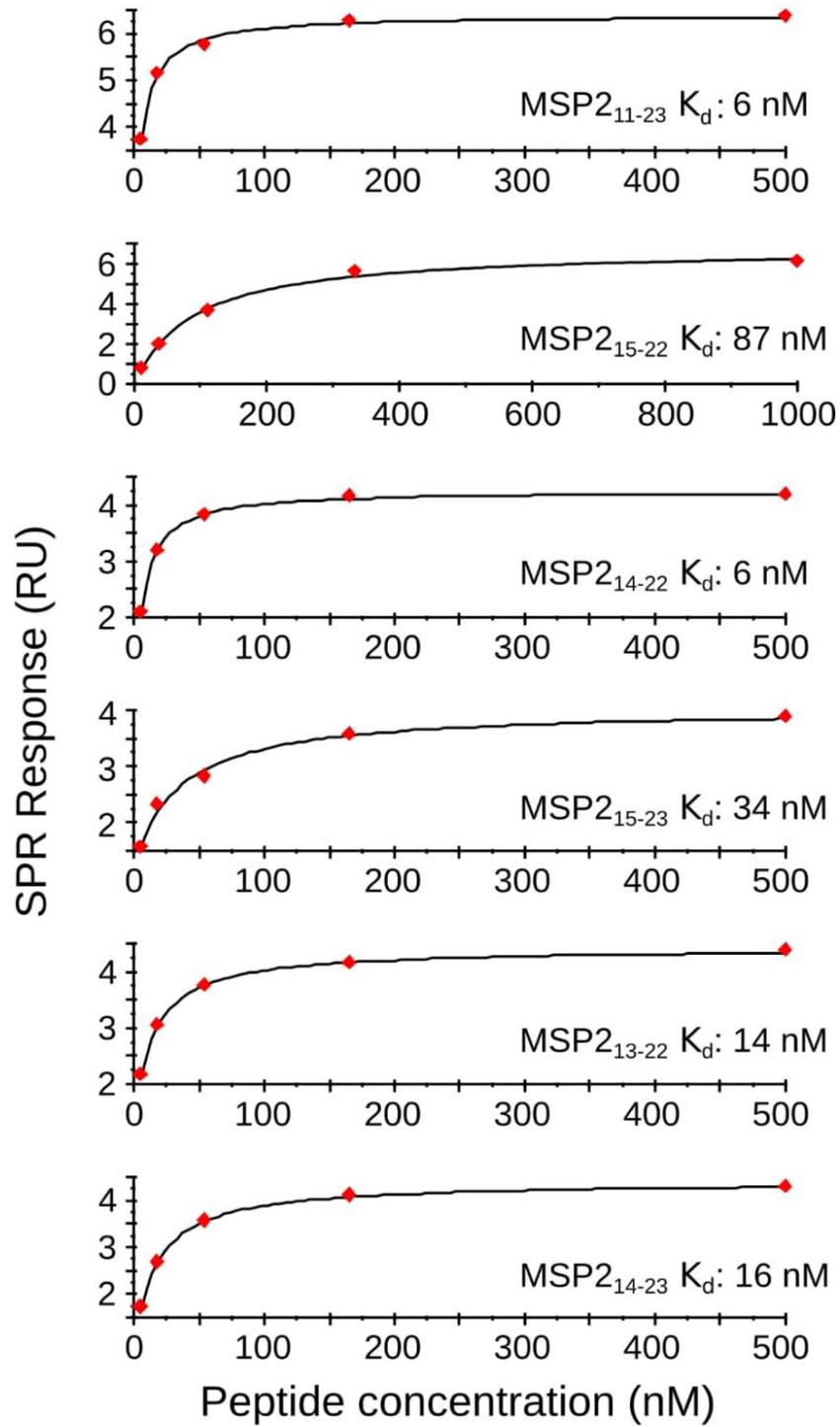
b Agreement between intensities of repeated measurements of the same reflections and can be defined as:  $\sum(I_{h,i} - \langle I_h \rangle) / \sum I_{h,i}$ , where  $I_{h,i}$  are individual values and  $\langle I_h \rangle$  is the mean value of the intensity of reflection h.



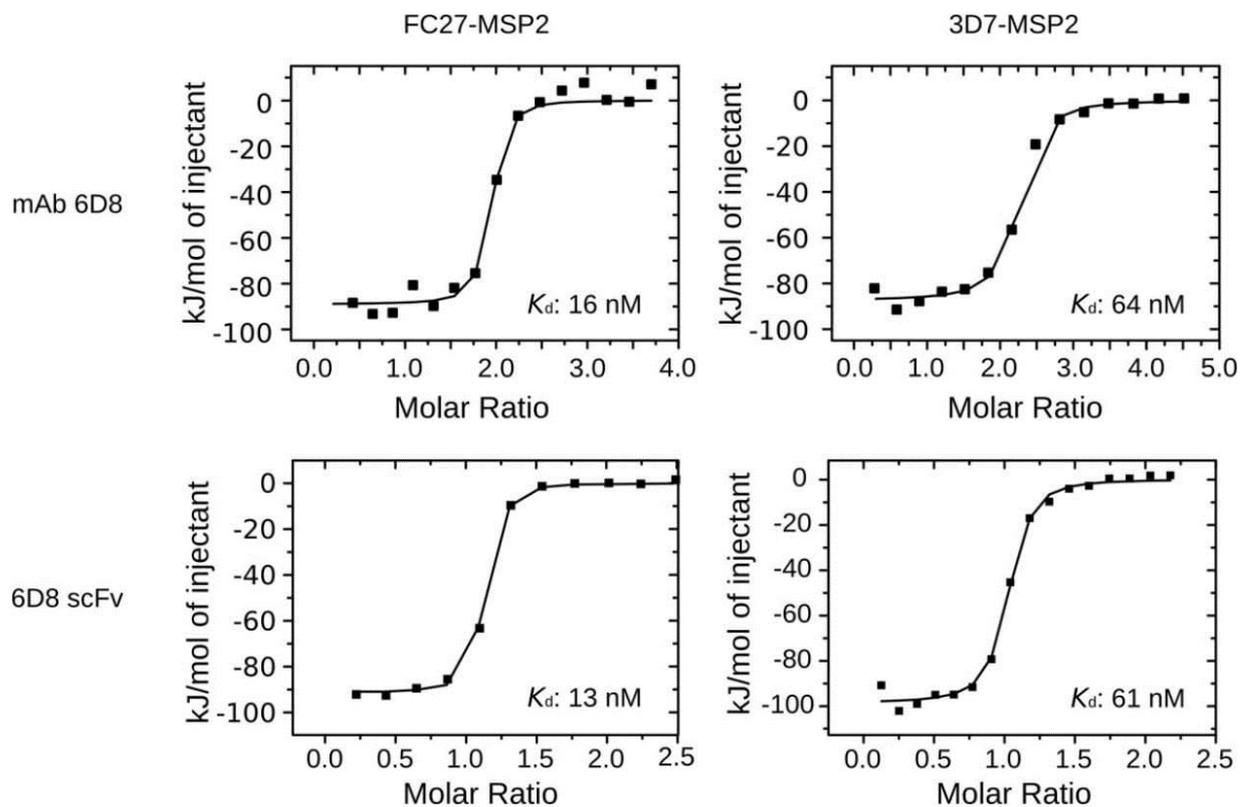
**Figure S1.**  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC spectra of FC27-MSP2-6His in the presence (blue) and absence (red) of dodecylphosphocholine (DPC) micelles containing 1 mol % of the  $\text{Ni}^{2+}$ -charged chelating lipid 1,2-di-(9Z-octadecenoyl)-*sn*-glycero-3-[(N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl] (DOGS-NTA).



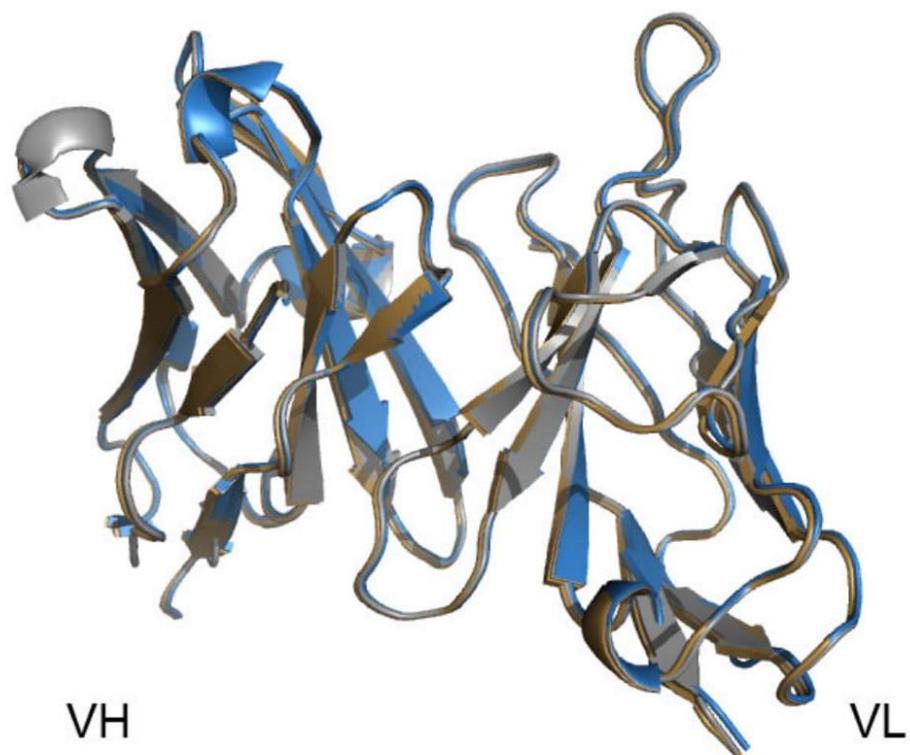
**Figure S2.** Helical propensity of synthetic epitope-bearing peptides MSP2<sub>11-23</sub> and MSP2<sub>14-22</sub>. Both peptides showed decreases in ellipticity at 208 and 220 nm and an increase in peak height at 190 nm, characteristic of  $\alpha$ -helical motifs adopted upon trifluoroethanol addition (0 to 30%). The minimal peptide epitope MSP2<sub>14-22</sub> was less helical than the 13-mer peptide (MSP2<sub>11-23</sub>) as expected for such short peptides. However, a helical component present in both peptides is small but still prone to stabilisation in lipid-like environments.



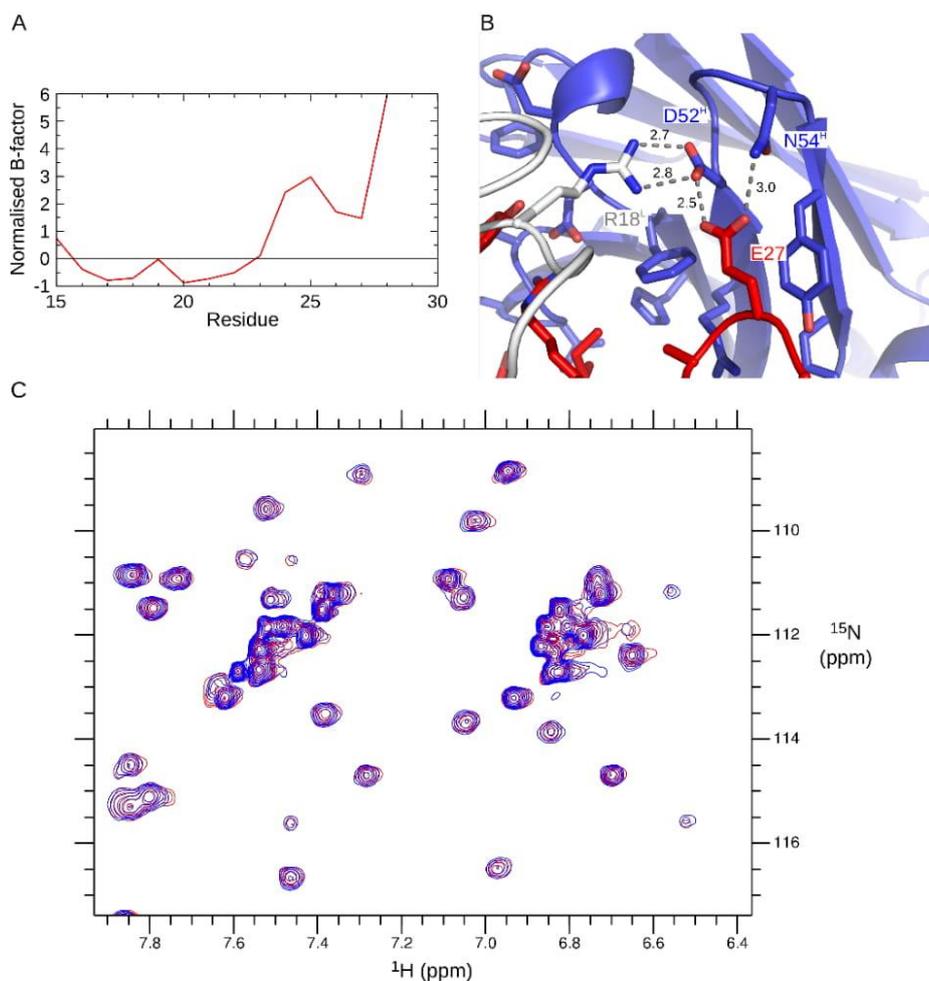
**Figure S3.** Representative SPR data for the determination of peptide binding affinities for immobilised 6D8.



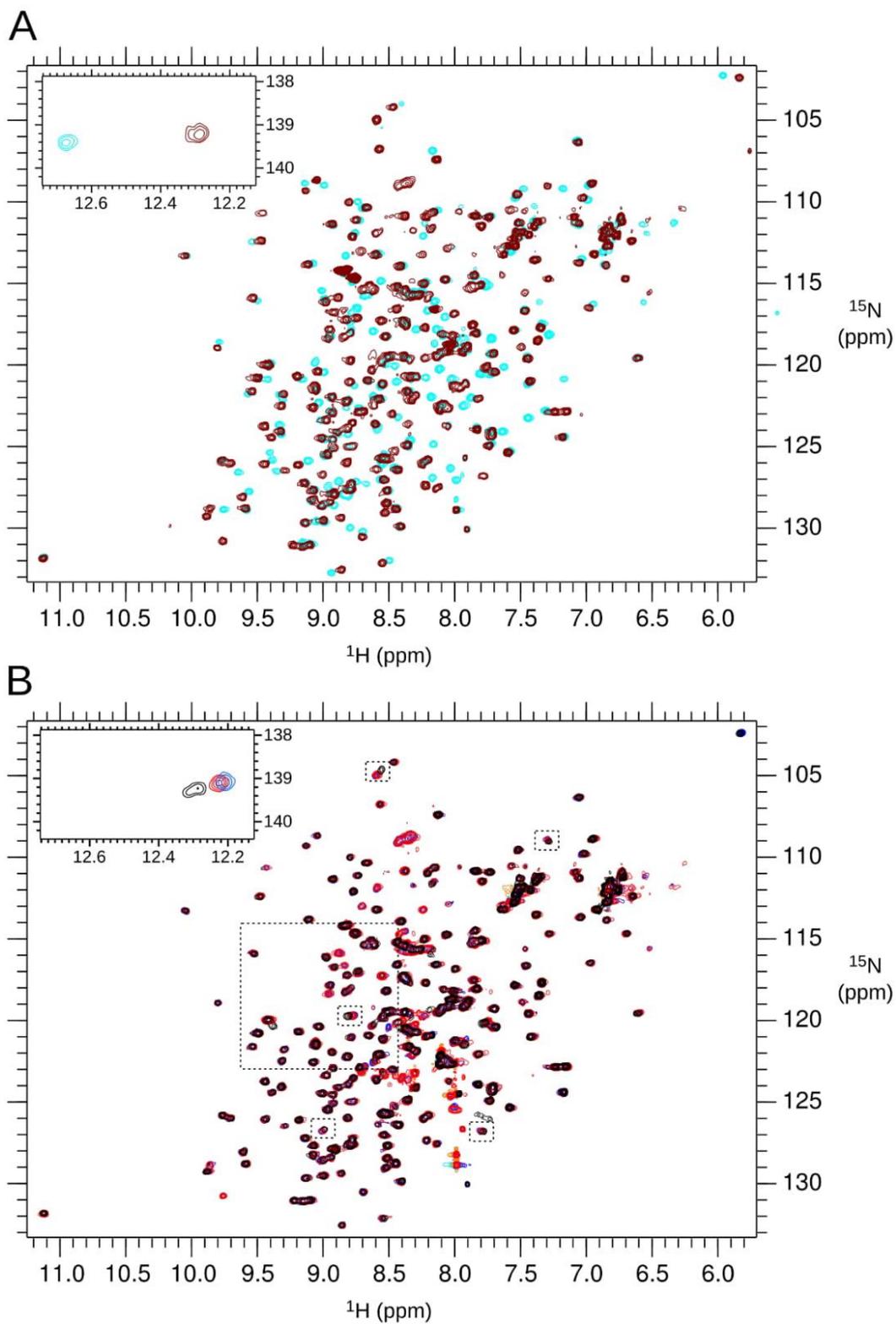
**Figure S4.** ITC titration of MSP2 to the mAb 6D8 (IgG) and 6D8 scFv. Binding data indicate that recognition of full-length 3D7 and FC27 MSP2 alleles is unaffected by the conversion of the full-length IgG antibody into a single-chain fragment. The strain-selectivity observed in 6D8 IgG is also fully reproduced in the scFv.



**Figure S5.** Overlay of the bound antibody structure of crystal complexes 6D8 Fv-MSP2<sub>14-22</sub>, 6D8 Fv-3D7 MSP2<sub>14-30</sub> and 6D8 Fv-FC27 MSP2<sub>14-30</sub> shown in grey, blue and brown respectively. Despite marked differences in affinity, side-chain deviations across the three crystal structures were minimal and unlikely to explain affinity differences (C $\alpha$  RMSD of 1.55 Å between 6D8 Fv-MSP2<sub>14-22</sub> and 6D8 Fv-3D7 MSP2<sub>14-30</sub> complexes and C $\alpha$  RMSD of 1.53 Å between 6D8 Fv-MSP2<sub>14-22</sub> and 6D8 Fv-FC27 MSP2<sub>14-30</sub>).

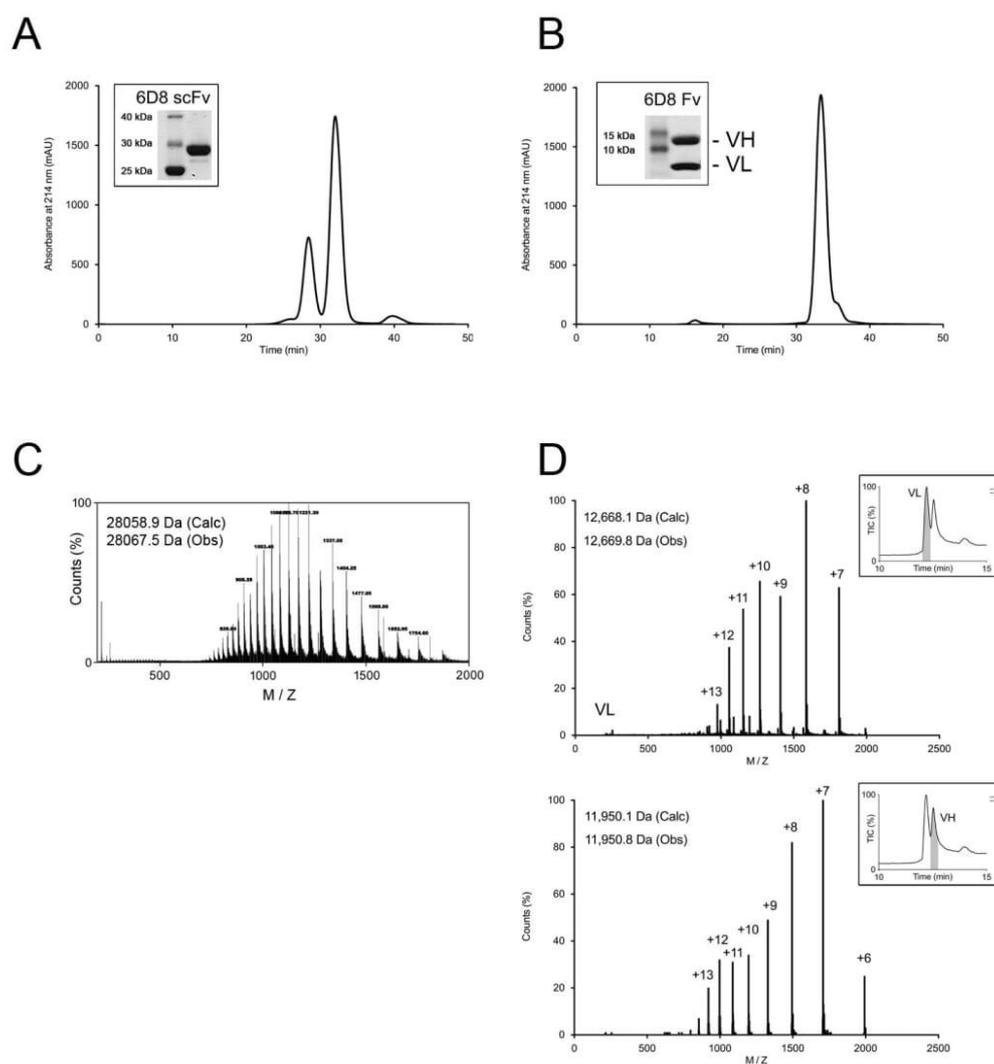


**Figure S6.** Crystal-contact-mediated interactions of 3D7-MSP2<sub>14-30</sub> with 6D8. **(A)** Normalised B-factors of C $\alpha$  atoms of 3D7-MSP2<sub>14-30</sub> in complex with 6D8 Fv. Isotropic B-factors are normalized as  $(B - \langle B \rangle) / \sigma$  where  $\langle B \rangle$  and  $\sigma$  are the average and standard deviation, respectively, of all C $\alpha$  B-factors in the complex. **(B)** H-bonds between 3D7-MSP2 residue E27 (red) and 6D8 V<sub>H</sub> (D52<sup>H</sup> and N54<sup>H</sup>; blue) are stabilized by crystal contacts involving R18<sup>L</sup> of a crystallographically adjacent Fv complex (grey). Donor-acceptor distances (Å) for H-bonds and salt bridges are labelled. **(C)** Sidechain NH<sub>2</sub> region of <sup>1</sup>H-<sup>15</sup>N HSQC spectra of 6D8 scFv in the presence of 3D7-MSP2<sub>14-30</sub> (blue) and FC27-MSP2<sub>14-30</sub> (red) showing the absence of chemical shift perturbations that are expected if E27 of 3D7 MSP2 interacts with the sidechain of N54<sup>H</sup> in solution.



**Figure S7.** (A) Overlay of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of 6D8 scFv in the absence (cyan) and presence (brown) of MSP2<sub>14-22</sub> (B) Overlay of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of 6D8 scFv in the presence of MSP2<sub>14-22</sub> (black), FC27-MSP2<sub>14-30</sub> (blue) and 3D7-MSP2<sub>14-30</sub> (red). Spectral regions shown in Fig 4 (main text) are boxed, and an outlying Trp indole NH peak is inset.





**Figure S9.** Purification profile of recombinant antibody fragments 6D8 scFv (A) and 6D8 Fv (B) used in this work. Antibody fragments were affinity purified, concentrated and freed from aggregates by gel filtration prior to structural and binding studies. The molecular masses observed for scFv (C) and Fv (D) were consistent with their calculated average masses detected by LCMS. The SDS-PAGE profile of each construct is presented in A and B.

## References

1. Zhang, X. et al. Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2. *J Mol Biol* **379**, 105-121 (2008).
2. MacRaid, C.A., et al. Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2. *PLoS ONE* (**in press**)(2015).
3. Adda, C.G. et al. Antigenic characterization of an intrinsically unstructured protein, *Plasmodium falciparum* merozoite surface protein 2. *Infect Immun* **80**, 4177-85 (2012).
4. Adda, C.G. et al. *Plasmodium falciparum* merozoite surface protein 2 is unstructured and forms amyloid-like fibrils. *Mol Biochem Parasitol* **166**, 159-171 (2009).
5. Orlandi, R. & Güssow, D.H. Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. *Proc Natl Acad Sci U S A* **24**, 527-31 (1989).
6. Rouet, R. et al. Expression of high-affinity human antibody fragments in bacteria. *Nat Protoc* **7**, 364-73 (2012).
7. Marley, J., Lu, M. & Bracken, C. A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* **20**, 71-5 (2001).
8. Fields, G.B. & Noble, R.L. Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int J Pept Protein Res* **35**, 161-214 (1990).
9. Schanda, P. & Brutscher, B. Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J Am Chem Soc* **127**, 8014-5 (2005).
10. Kabsch, W. XDS. *Acta Crystallogr* **D66**, 125-132 (2010).
11. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr* **D50**, 760-763 (1994).
12. Brunger, A.T. Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallogr D Biol Crystallogr* **49**, 24-36 (1993).
13. McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C. & Read, R.J. Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* **61**, 458-64 (2005).
14. Cohen, S.X. et al. ARP/wARP and molecular replacement: the next generation. *Acta Crystallogr D Biol Crystallogr* **64**, 49-60 (2008).
15. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr* **D66**, 213-221 (2010).
16. Afonine PV et al. Towards automated crystallographic structure refinement with phenix.refine. . *Acta Cryst. D* **D68**, 352-367 (2012).
17. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32 (2004).
18. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97 (2007).

# Chapter 4

Structure and characterisation of a key epitope in the conserved C-terminal domain of the malaria vaccine candidate MSP2

## 4.1 Chapter introduction

In the previous Chapter, the structural determinants of 6D8 binding to an epitope in the conserved N-terminal region of MSP2 are determined. A high-resolution crystal structure of 6D8 Fv in complex with its minimal binding epitope showed that the antibody-bound conformation was incompatible with its lipid-bound counterpart, which is known to adopt an  $\alpha$ -helical conformation. When bound to 6D8, the  $\alpha$ -helix of the epitope was disrupted at Ser19 owing to interactions of Arg21 and Arg22 with the 6D8 paratope, revealing why 6D8 mAb is unable to recognise native MSP2 on the merozoite surface. Strain-specific binding of 6D8 was also investigated with crystallographic data and NMR. These results showed that residues in the variable region beyond the fully conserved epitope could form transient interactions with 6D8 and modulate binding affinity.

In this Chapter, the crystallographic methods established with 6D8 are applied to epitopes in the conserved C-terminal region of MSP2, specifically the epitope recognised by 4D11 mAb. In contrast to 6D8 mAb, 4D11 mAb is able to recognise parasite MSP2, suggesting that its epitope is accessible to the host immune system. A crystal structure was solved with 4D11 Fv in complex with its cognate 8-residue epitope (NKENCGAA). The bound peptide was found to form a homodimer, during the process of crystallisation, mediated by a disulphide bond between the free cysteines in the monomers. Although this dimer is not present in native MSP2, MD simulations and SPR competition binding assays showed that the bound conformation of the dimer is consistent with the native disulphide-bonded sequence. These results support the relevance of the crystal structure to native MSP2 on the parasite surface. This structure underpinned further work on the design and evaluation of MSP2-based peptide vaccine candidates, as described in the subsequent chapter. As the results of this Chapter have been published in *Journal of Molecular Biology*, they are presented in the format of a published article in the immediate section below.



# Structure and Characterisation of a Key Epitope in the Conserved C-Terminal Domain of the Malaria Vaccine Candidate MSP2

Jeffrey Seow<sup>1</sup>, Rodrigo A.V. Morales<sup>1</sup>, Christopher A. MacRaid<sup>1</sup>, Bankala Krishnarjuna<sup>1</sup>, Sheena McGowan<sup>2</sup>, Tamir Dingjan<sup>1</sup>, Garima Jaipuria<sup>3</sup>, Romain Rouet<sup>4</sup>, Karyn L. Wilde<sup>5</sup>, Hanudatta S. Atreya<sup>3</sup>, Jack S. Richards<sup>6</sup>, Robin F. Anders<sup>7</sup>, Daniel Christ<sup>4</sup>, Nyssa Drinkwater<sup>2</sup> and Raymond S. Norton<sup>1</sup>

<sup>1</sup> - Monash Institute of Pharmaceutical Sciences, Monash University, Parkville 3052, Australia

<sup>2</sup> - Department of Microbiology, Monash University, Clayton 3168, Australia

<sup>3</sup> - NMR Research Centre, Indian Institute of Science, Bangalore, 560012, India

<sup>4</sup> - Garvan Institute of Medical Research, Darlinghurst 2010, Australia

<sup>5</sup> - National Deuteration Facility, Australian Nuclear Science and Technology Organisation, Lucas Heights 2234, Australia

<sup>6</sup> - Centre for Biomedical Research, The Burnet Institute, Melbourne 3004, Australia

<sup>7</sup> - Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne 3086, Australia

Correspondence to Raymond S. Norton: [ray.norton@monash.edu](mailto:ray.norton@monash.edu)

<http://dx.doi.org/10.1016/j.jmb.2017.02.003>

Edited by Dr Arne Skerra

## Abstract

Merozoite surface protein 2 (MSP2) is an intrinsically disordered antigen that is abundant on the surface of the malaria parasite *Plasmodium falciparum*. The two allelic families of MSP2, 3D7 and FC27, differ in their central variable regions, which are flanked by highly conserved C-terminal and N-terminal regions. In a vaccine trial, full-length 3D7 MSP2 induced a strain-specific protective immune response despite the detectable presence of conserved region antibodies. This work focuses on the conserved C-terminal region of MSP2, which includes the only disulphide bond in the protein and encompasses key epitopes recognised by the mouse monoclonal antibodies 4D11 and 9H4. Although the 4D11 and 9H4 epitopes are overlapping, immunofluorescence assays have shown that the mouse monoclonal antibody 4D11 binds to MSP2 on the merozoite surface with a much stronger signal than 9H4. Understanding the structural basis for this antigenic difference between these antibodies will help direct the design of a broad-spectrum and MSP2-based malaria vaccine. 4D11 and 9H4 were reengineered into antibody fragments [variable region fragment (Fv) and single-chain Fv (scFv)] and were validated as suitable models for their full-sized IgG counterparts by surface plasmon resonance and isothermal titration calorimetry. An alanine scan of the 13-residue epitope 3D7-MSP2<sub>207–222</sub> identified the minimal binding epitope of 4D11 and the key residues involved in binding. A 2.2-Å crystal structure of 4D11 Fv bound to the eight-residue epitope NKENCGAA provided valuable insight into the possible conformation of the C-terminal region of MSP2 on the parasite. This work underpins continued efforts to optimise recombinant MSP2 constructs for evaluation as potential vaccine candidates.

© 2017 Elsevier Ltd. All rights reserved.

## Introduction

Development of a malaria vaccine is a high priority. The mosquito-borne disease, caused by protozoa of the *Plasmodium* genus, is responsible for 200 million cases and over 430,000 deaths per year [1]. Although vector control and drug interventions have significantly

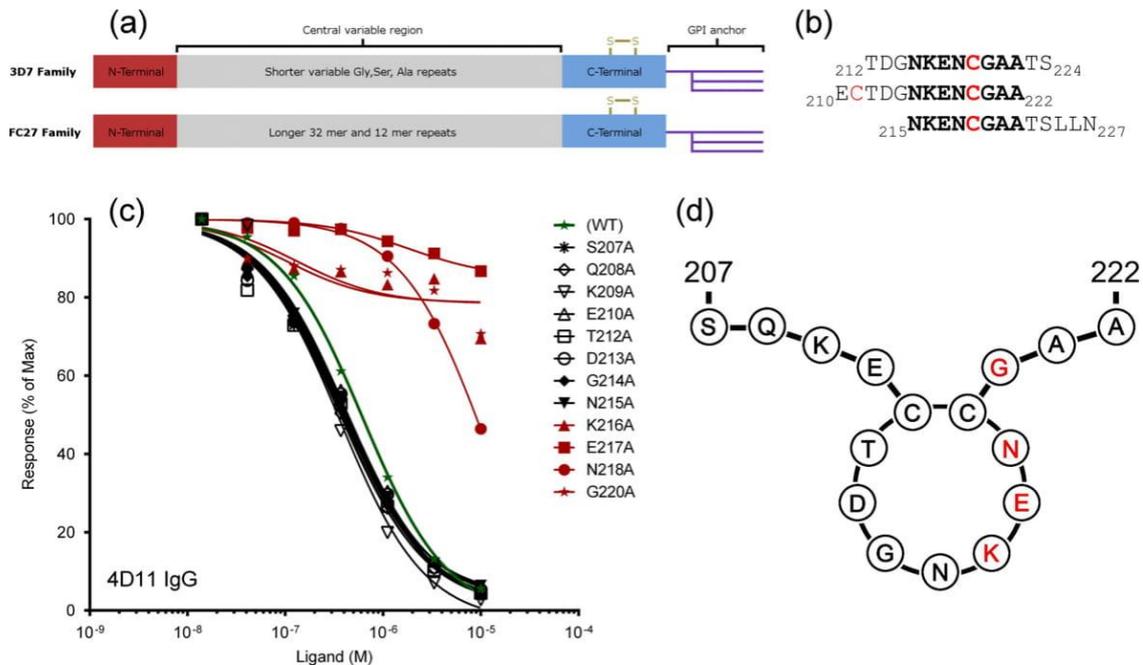
decreased mortality over the past decade, the global eradication of malaria calls for a robust and effective vaccine [2]. The most advanced vaccine candidate is RTS,S/AS01, a pre-erythrocytic vaccine that has shown only modest efficacy in young children and infants in Phase III trials [3]. Blood-stage vaccines are aimed at inhibiting parasite growth and proliferation,

thereby reducing the severity of symptoms and in turn mortality [4,5]. The eradication of malaria will require the development of multistage vaccines able to prevent blood-stage infection and block transmission.

Merozoite surface protein 2 (MSP2) is a ~23-kDa glycosylphosphatidyl inositol (GPI)-anchored protein present in the asexual blood stages of *Plasmodium falciparum* and is one of the most abundant proteins on the surface of the merozoite [6,7]. The protein is highly polymorphic, particularly in the central variable region, which consists of tandemly arrayed repeat sequences and dimorphic sequences that differentiate the protein into two allelic families, 3D7 and FC27 (Fig. 1a) [8,9]. The central variable region is flanked by a highly conserved 25-residue N-terminal region and ~50-residue C-terminal region, the latter containing the only disulphide bond of the protein. Similar to many other blood-stage antigens [10–12], MSP2 is an intrinsically disordered protein lacking a well-defined structure in solution [13], although the protein has been found to have a propensity for amyloid-like fibril formation [14,15]. The structural characteristics of MSP2 when interacting with the merozoite surface and those mediated by the GPI-anchor are still being explored. NMR studies have shown that the conserved N-terminal region adopts an  $\alpha$ -helical conformation when in the presence of lipid [16] and identified

some motional restriction in the conserved C-terminal region caused by the single disulphide bond [13].

MSP2 is well characterised as a vaccine candidate and has been included in several human vaccine trials [17–21]. In Phase I-IIb trials, recombinant 3D7-MSP2 in combination with MSP1 and RESA was tested in Papua New Guinean children [18,22]. A 62% reduction in parasite densities was observed, although this was biased towards parasites expressing the 3D7 allele of MSP2 that was used in the vaccine [19]. This implies that the response to MSP2 dominated the protective effect in this trial, but the MSP2 response was highly strain-specific. A vaccine containing recombinant forms of both 3D7 and FC27 MSP2 was tested in Phase I trials and induced antibodies against both alleles that were able to inhibit parasite growth by antibody-dependent cellular inhibition [20] and complement-mediated inhibition [23]. A recent approach to circumvent the problems of strain-specific immune responses involved the use of MSP2 chimeras that incorporated conserved and variable regions of the 3D7 and FC27 alleles [24]. In mice, these chimeras were able to induce a robust anti-MSP2 antibody response across both alleles. Nonetheless, the success of an MSP2 vaccine is likely to be significantly enhanced by a protective immune response targeting conserved epitopes.



**Fig. 1.** (a) Schematic of the primary structures of 3D7 and FC27 MSP2. (b) Aligned 13-residue peptides from peptide array able to bind with mAb 4D11. 3D7-MSP2<sub>212–224</sub> and 3D7-MSP2<sub>210–222</sub> reacted strongly with mAb 4D11 by ELISA [25]. Overlapping residues are shown in bold, cysteine residues are shown in red. (c) SPR competition assay of alanine scan 16-residue peptides of 3D7-MSP2<sub>207–222</sub>. Black lines show alanine mutants with no effect on binding, the green line indicates the wild-type binding of 3D7-MSP2<sub>207–222</sub>, and red lines indicate the alanine mutations that decreased the binding to 4D11 IgG. (d) Schematic representation of disulphide-bonded peptide sequence indicating the location of key residues in red.

A panel of monoclonal antibodies (mAbs) has been generated from mice immunised with recombinant forms of both 3D7 and FC27 MSP2 [25]. Epitope mapping was performed using an array of overlapping 13-residue peptides spanning both alleles of MSP2. Overall, 10 mAbs were able to recognise linear epitopes in both conserved and variable regions of MSP2. Here, we focus on mAbs 4D11, 9G8, and 9H4, which recognise overlapping epitopes within an 18-residue stretch of the conserved C-terminal region encompassing the single disulphide bond. These antibodies are all able to recognise recombinant MSP2, but they show differing reactivity by immunofluorescence assays (IFAs) and western blot towards MSP2 derived from parasite material. 4D11 and 9G8 exhibited a 10-fold stronger signal by IFA in comparison to 9H4, suggesting that the 9H4 epitope is less accessible on the merozoite surface. In this work, we use the 4D11 mAb to better understand the structural determinants of binding to parasite MSP2.

## Results

### Alanine scan of 4D11 epitope

In previous work, the epitope of mAb 4D11 was mapped with an array of 13-residue peptides covering full-length 3D7-MSP2 and FC27-MSP2, with an overlap of 8 residues. mAb 4D11 was able to bind three of these overlapping peptides (Fig. 1b). As the 8-residue sequence NKENCGAA was common to these three peptides, this 8-mer appears to be the minimal epitope for mAb 4D11. Binding of mAb 4D11 to MSP2 was independent of the presence of the disulphide bond, as the 3D7-MSP2<sub>212–224</sub> and 3D7-MSP2<sub>215–227</sub> peptides each contained only a single cysteine, and this result was confirmed in the context of full-length MSP2 [25].

To determine the contributions to binding of each residue in the minimal epitope sequence, we performed an alanine scan of MSP2<sub>207–222</sub> (Table 1). All peptides were oxidised to form the disulphide bond before HPLC purification. The N- and C termini were capped by acetyl and amide moieties, respectively, to avoid electrostatic interactions with the epitope that would not be present in the full-length protein. The affinities of these peptides for 4D11 IgG were determined by surface plasmon resonance (SPR) competition assays using immobilised 3D7 MSP2 and varying concentrations of peptide with 4D11 IgG. Direct binding assays were not suitable, as the small peptide size gave a poor response. Mutations to Lys216, Glu217, Asn218, and Gly220 caused a marked decrease in binding affinity, suggesting that they make key interactions with 4D11 (Fig. 1c and d, and Supplementary Fig. S1). Mutation of Glu217 to alanine resulted in a complete loss of binding. In

**Table 1.** Alanine scan of the 13-residue mAb 4D11 epitope MSP2<sub>207–222</sub> to determine the key residues for binding

Peptide	16-mer sequence*	K <sub>d</sub> against 4D11 IgG (μM)
3D7-MSP2 <sub>207–222</sub> (WT)	SQKECTDGNKENC <del>G</del> AA	0.9 ± 0.4
S207A	AQKECTDGNKENC <del>G</del> AA	0.5 ± 0.2
Q208A	SAKECTDGNKENC <del>G</del> AA	0.4 ± 0.1
K209A	SQAECTDGNKENC <del>G</del> AA	0.3 ± 0.1
E210A	SQKACTDGNKENC <del>G</del> AA	0.5 ± 0.2
T212A	SQKECADGNKENC <del>G</del> AA	0.5 ± 0.2
D213A	SQKECTAGNKENC <del>G</del> AA	0.5 ± 0.1
G214A	SQKECTDANKENC <del>G</del> AA	0.4 ± 0.1
N215A	SQKECTDGA <del>K</del> ENC <del>G</del> AA	0.5 ± 0.2
K216A	SQKECTDGN <del>A</del> ENC <del>G</del> AA	21.0 ± 6.0
E217A	SQKECTDGNK <del>A</del> NC <del>G</del> AA	219 ± 64
N218A	SQKECTDGNK <del>E</del> AC <del>G</del> AA	6.0 ± 2.3
G220A	SQKECTDGNKENC <del>A</del> AA	42 ± 10

K<sub>d</sub> was determined by SPR (± standard deviation of three replicate experiments). Synthetic peptides were N-terminally acetylated and C-terminally amidated.

contrast, mutation of residues 207–215 did not have a significant effect on binding. The decreased binding caused by the mutation of Gly220 was unexpected but may indicate that the flexibility and/or small size of this residue are important for binding.

### 4D11 antibody fragments are suitable models for their full IgG counterparts

To better understand the antibody-bound conformation of the key 4D11 epitope, we designed variable region (Fv) and single-chain Fv (scFv) antibody fragments based on the 4D11 IgG sequence, which were expressed at high yield in *Escherichia coli* [26,27]. To confirm that these antibody fragments were suitable models for 4D11 IgG, we used isothermal titration calorimetry (ITC) to compare their affinities using the 3D7-MSP2<sub>207–222</sub> peptide (Supplementary Fig. S2). 4D11 Fv and scFv were found to have K<sub>d</sub> values of 2.7 and 3.0 μM, respectively, very close to the binding affinity of 4D11 IgG at 2.2 μM.

### NMR spectroscopy of MSP2<sub>207–224</sub> bound to 4D11 scFv

To further explore the interaction between 4D11 and MSP2, we employed NMR. To enable efficient isotope labelling, we used a recombinant form of 18-residue MSP2<sub>207–224</sub>. The NMR spectra of MSP2<sub>207–224</sub> free in solution reveal limited chemical shift dispersion, consistent with the expected lack of well-ordered structure in the free epitope (Supplementary Fig. S3). Upon the addition of a small excess of unlabelled 4D11 scFv to <sup>15</sup>N-labelled 3D7-MSP2<sub>207–224</sub>, a more dispersed set of peaks appeared, indicating that much of the 18-residue peptide adopts a more defined structure in the presence of the antibody. Indeed, each peak in the [<sup>1</sup>H, <sup>15</sup>N]-heteronuclear single quantum coherence spectrum of the free peptide was significantly

**Table 2.** Data collection and refinement statistics

4D11 Fv + MSP2 <sub>215–222</sub> (PDB ID: 5TBD)	
<i>Data collection</i>	
Space group	P1
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	41.97, 72.97, 81.97
$\alpha$ , $\beta$ , $\gamma$ (°)	68.2, 78.8, 76.8
Resolution (Å)	65.99–2.2 (2.27–2.20)
<i>R</i> <sub>pim</sub> (%)	15.5 (66.4)
<i>I</i> / $\sigma$ <sub>1</sub>	7.5 (2.6)
Completeness (%)	100 (100)
Redundancy	7.3 (7.3)
Unique reflections	43,777
<i>Refinement</i>	
Resolution	2.2
Number of reflections	321,569
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> (%)	20.66/24.70
Number of atoms	
Protein	7064
Water	365
<i>B</i> -factors	
Protein	27.1
Fv	27.3
Epitope	19.9
Water	26.9
RMSDs	
Bond length (Å)	0.003
Bond angles (°)	0.68
Ramachandran Plot	
Outliers (%)	0
Favoured (%)	98.1
Rotamer outliers (%)	0.3
Clashscore [46]	2.65

Values in parentheses refer to the highest-resolution shell.

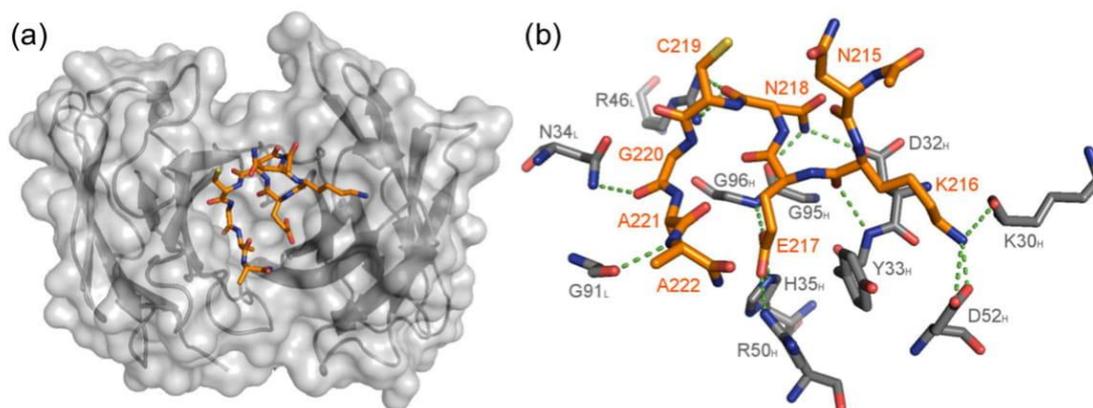
Agreement between intensities of repeated measurements of the same reflections can be defined as  $\sum(|h_{i,j} - \langle h \rangle|) / \sum |h_{i,j}|$ , where  $|h_{i,j}|$  are individual values and  $\langle h \rangle$  is the mean value of the intensity of reflection *h*.

perturbed by antibody binding, despite the inference from the alanine scan that the N-terminal half of the peptide does not make significant interactions with the

antibody. There is, however, significant variation in the peak intensity and linewidth in the antibody-bound spectra, suggesting that some parts of the peptide retain significant flexibility when bound. The relaxation properties of this complex were such that obtaining good quality triple-resonance data from much of the bound peptide proved impossible. Accordingly, we were unable to fully assign the spectra of the bound peptide.

### Crystal structure of 3D7 MSP2<sub>215–222</sub> bound to 4D11 Fv

Initial crystallisation trials involving 4D11 Fv in complex with the 16-residue disulphide-cyclised 3D7-MSP2<sub>207–222</sub> were unsuccessful. This was probably a consequence of significant flexibility in a large portion of the peptide, as inferred from the NMR data. Smaller peptides were therefore tested, including the eight-residue 3D7-MSP2<sub>215–222</sub>, the minimal binding epitope of 4D11. The 4D11 3D7-MSP2<sub>215–222</sub> complex was crystallised and its structure was determined to 2.2-Å resolution (Table 2 and Fig. 2). The crystals contained four Fv–peptide complexes in the asymmetric unit, with the peptide forming two homodimers via disulphide bonds formed between the free cysteines in the monomer sequence. The bound peptide is stabilised as a  $\beta$ -bend ribbon, a subtype of the  $3_{10}$   $\alpha$ -helical structure, which is characterised by consecutive overlapping  $\beta$ -turns. The first turn encompasses Asn215 (i)–Asn218 (i + 3), with dihedral angles typical of a type I  $\beta$ -turn. The second overlapping turn encompasses Glu217 (i)–Gly220 (i + 3) and, although slightly different from the canonical dihedral angles, most closely resembles a type I'  $\beta$ -turn [28]. The crystal structure supports the results from the alanine scan, with Lys216, Glu217, and Asn218 forming key interactions with the 4D11 variable region heavy chain (V<sub>H</sub>) complementarity-determining regions.



**Fig. 2.** The 2.2-Å crystal structure of 4D11 Fv bound to the 8-residue peptide 3D7-MSP2<sub>215–222</sub>. (a) Crystal structure of 3D7-MSP2<sub>215–222</sub> bound within the paratope of 4D11 Fv. (b) The conformation of 3D7-MSP2<sub>215–222</sub> (orange) is stabilised by a number of hydrogen bond interactions (shown as green dashed lines) with the complementarity-determining regions of 4D11 Fv (grey) involving residues Lys216, Glu217, Asn218, Gly220, and Ala221.

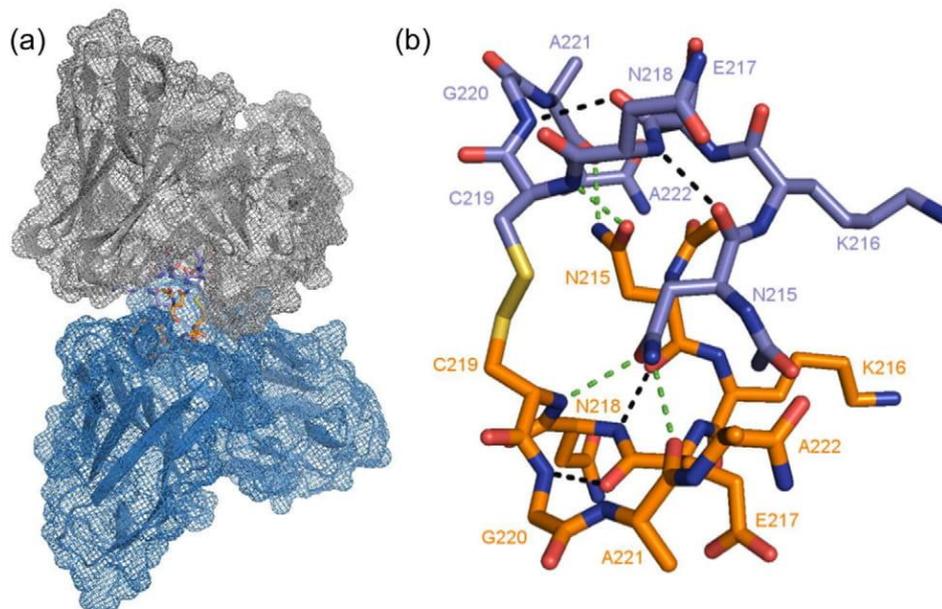
Gly220 formed a hydrogen bond with a light chain Asn residue (Asn34<sub>L</sub>; the Kabat numbering system was used for 4D11 residues [29] with the chain indicated by a subscript). The lack of space surrounding the Gly  $\alpha$ -carbon explains why the mutation to an alanine in this position had a detrimental effect on binding. An additional interaction was also present between the backbones of Ala221 and Gly91<sub>L</sub>.

Although the peptide was purified as a monomer by HPLC, and size-exclusion chromatography confirmed that the 4D11 Fv-peptide complex was monomeric prior to crystallisation (Supplementary Fig. S4), the 8-residue peptide 3D7-MSP2<sub>215-222</sub> formed a disulphide-mediated homodimer in the crystal. The four Fvs and peptides in the asymmetric unit of the crystal were oriented in such a way that two Fvs had their antigen-binding sites facing each other, forming a small pocket in which the peptide homodimer was bound (Fig. 3). Attempts to crystallise 4D11 Fv with eight-residue monomeric peptide analogues with serine in place of the cysteine failed. The tight packing of the Fv-peptide dimer in the crystal suggests that the formation of the dimer may have assisted crystal formation. The peptide dimer itself contains both intra- and intermolecular hydrogen bonds that stabilise its conformation. Asn215 in particular forms intermolecular hydrogen bonds with the backbone of Cys219 and Ala221 and an intramolecular hydrogen bond with Asn218. 4D11 binds recombinant, full-length MSP2, in which Cys219 is engaged in an intramolecular disulphide with Cys211, as a monomer and binds equally well to a cysteine mutant C219S [25]. There is

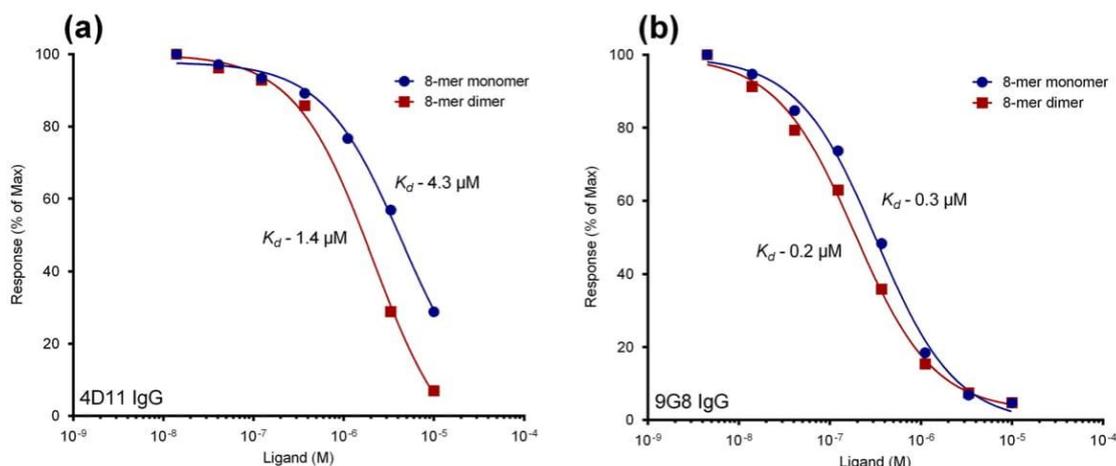
no evidence that native MSP2 exists as a covalent dimer on the parasite surface [14], implying that 4D11 recognises monomeric MSP2 on the parasite surface. Accordingly, it was necessary to establish the relevance of the dimer structure to that interaction.

### Dimeric 8-mer is able to bind 4D11 and 9G8 IgG

The binding affinity of dimeric 3D7-MSP2<sub>215-222</sub> to 4D11 IgG was determined in order to assess the relevance of the bound dimeric conformation revealed in the crystal structure. Both air oxidation and iodine oxidation of the monomeric 8-residue peptide in solution failed to yield dimeric peptide. Instead, the dimer was synthesised on resin with cystine as a building block. The affinities of the monomeric and dimeric peptides were measured against both 4D11 and 9G8 IgG by an SPR competition binding assay. The affinity ( $K_d$ ) of the monomeric 8-mer was 4.3  $\mu$ M, slightly weaker than the dimer at 1.4  $\mu$ M. These affinities are comparable to the 0.9  $\mu$ M  $K_d$  of the longer 3D7-MSP2<sub>207-222</sub> peptide, which contains the native disulphide bond. In addition, mAb 9G8, another antibody able to recognise parasite MSP2 by IFA and western blotting, was able to bind the monomeric and dimeric peptide with  $K_d$  values of 0.3 and 0.2  $\mu$ M, respectively (Fig. 4 and Supplementary Fig. S5). The ability of the dimer to bind both 4D11 IgG and 9G8 IgG at affinities comparable to those of the monomer indicates that the conformation adopted by the bound dimer is accessible to recombinant and parasite MSP2. The slightly higher affinity binding of the dimer



**Fig. 3.** Crystal structure of 4D11 Fv in complex with 3D7-MSP2<sub>215-222</sub> homodimer. (a) Both peptides in the homodimer are able to bind to 4D11 Fv, hence the unusual orientation of two Fvs facing towards each other. (b) Hydrogen bond interactions within the homodimer include intramolecular (black dashed lines) and intermolecular (green dashed lines).



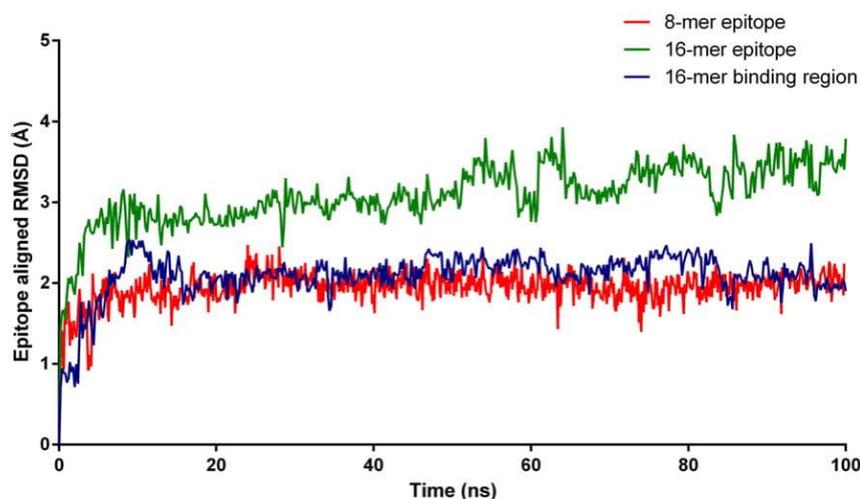
**Fig. 4.** SPR competition assay of dimeric and monomeric 8-residue peptide 3D7-MSP<sub>215–222</sub> using immobilised C-terminal region peptide 3D7-MSP<sub>217–221</sub>, (a) 4D11 IgG, and (b) 9G8 IgG.

to both antibodies suggests that the constraints imposed by the additional intermolecular hydrogen bonds in the dimer structure are favourable for 4D11 binding, probably by reducing the entropic penalty associated with the disordered-to-ordered conformational transition upon binding.

#### Molecular dynamics (MD) simulations with bound 8-mer and modelled 16-mer epitopes

MD simulations were employed to gauge the stability of the 8-residue and 16-residue MSP2 epitopes in complex with 4D11 Fv. This enabled us to assess whether additional adjacent residues and/or the disulphide bond in the native epitope have an effect on the epitope binding conformation. MD simulations over 100 ns were performed to sample the conforma-

tional space of two peptides, MSP<sub>215–222</sub> and MSP<sub>207–222</sub>, bound to 4D11 Fv. For the monomeric 8-residue epitope MSP<sub>215–222</sub>, the starting structure was extracted directly from the crystal structure. For the 16-residue epitope, the crystallographic conformation was extended N-terminally to residue 207 (MSP<sub>207–222</sub>), with the native disulphide modelled in. When only 4D11 Fv was aligned, the epitope RMSD for both peptides remained below 3 Å, suggesting that they remained in the binding pocket during the entire simulation (Supplementary Fig. S6). When the 8-residue epitope was aligned over this period, the epitope RMSD remained at 2 Å throughout the simulation, suggesting that the overall epitope conformation did not change significantly during simulation (Fig. 5). As expected from the NMR data, the flexible N-terminal extensions of the modelled 16-residue



**Fig. 5.** RMSD of the epitope shape over 100 ns with respect to the equilibrated MD simulation starting conformation. All residues were aligned for the 8-mer (red) and 16-mer (green); blue indicates the RMSD of the 16-mer with only the binding residues 217–222 aligned. The conformation of the binding region of the modelled 16-residue epitope does not change significantly.

epitope had a higher epitope-aligned RMSD (3 Å) than the 8-residue epitope (2 Å). However, when only the core binding residues (NKENCGAA) were used for alignment in the modelled 16-residue epitope, they were found to be stable, with an RMSD of 2 Å, identical to that of the 8-residue epitope alone. These simulations showed that the 8-residue epitope and the extended disulphide-bridged 16-residue epitope share identical conformational behaviour and stability when bound to 4D11 Fv, further supporting the relevance of the bound dimeric conformation in the crystal structure.

## Discussion

Intrinsically disordered proteins are abundant in eukaryotic pathogens, with apicomplexans, including the malaria parasite *Plasmodium*, being particularly enriched [10,30]. Moreover, these disordered proteins are frequently targets of host immune recognition [12,31]. The effects of conformational disorder on antigenicity, and the broader implications for vaccine development, are largely unknown and are still being explored. Entropic costs associated with antibody binding to disordered antigens have been thought to be detrimental to the development of high affinity and specific antibodies. However, a comparison of antibody binding to disordered and ordered antigens has shown that disordered epitopes, although usually shorter in length than their ordered counterparts, are highly efficient in their binding [31].

The recombinant form of MSP2 used in clinical trials is highly disordered [13], and the conformations of some regions of the protein differ from that of the native GPI-anchored antigen. Many disordered proteins are known to undergo a disorder-to-order transition when binding to their functional targets [32–34], and given the abundance of proteins in the merozoite surface coat, it is likely that such protein–protein interactions may influence the conformation of MSP2. Additionally, it is possible that lipid interactions or the C-terminal GPI-anchor may impose conformational constraints on the protein, favouring a more ordered conformation for the parasite protein. For example, the N-terminal region of MSP2 has been found to adopt an  $\alpha$ -helical conformation upon interaction with lipid micelles [16]. The mAb 6D8, which recognises an epitope within the conserved N-terminal region of MSP2, is unable to recognise parasite MSP2. The structural basis for this significant difference in binding was revealed by a high-resolution crystal structure of the 6D8–epitope complex, which showed that the antibody-bound and lipid-bound conformations were incompatible [35].

The strain specificity of the vaccine response towards MSP2 implies that epitopes within the variable region play a role in protection [19]. For the development of a strain-transcending MSP2-based vaccine, however, directing the immune response towards the conserved regions of the protein would be beneficial, provided that

epitopes in these conserved regions are recognised by antibodies elicited by the vaccine. Thus, mAb 4D11 and its corresponding scFv and Fv antibody fragments, which recognise an epitope that is accessible in parasite MSP2, present a valuable opportunity to understand how the C-terminal conserved region behaves on the merozoite surface. The crystal structure obtained here revealed that the bound 4D11 epitope adopted a  $\beta$ -bend ribbon conformation with two overlapping hydrogen bonded  $\beta$ -turns. MD simulations showed that this conformation was stable in the context of the extended MSP2<sub>207–222</sub> epitope, which includes the native intramolecular disulphide bond.

The accessibility of the 4D11 epitope on the parasite surface contrasts with poor accessibility of the epitope recognised by mAb 9H4 [25]. Although 9H4 recognises an epitope that overlaps with that of 4D11, 9H4 only weakly recognises parasite MSP2 by IFA and western blotting. The ability to direct the immune response towards epitopes that are accessible on the parasite surface, such as the 4D11 epitope investigated here, and are away from potentially distracting epitopes such as 9H4, suggests a strategy for engineering a more effective MSP2-based vaccine.

The alanine scan data presented in this paper, together with the crystal structure of the bound conformation of the minimal epitope, provide a rational basis for such antigen engineering. A possible approach to optimising the immunogenicity of the C-terminal region of MSP2 involves constraining the conformation towards that observed for this key epitope in the crystal structure. Although this has not been attempted with a disordered antigen such as MSP2, there has been success in the rational design of vaccines based on structure [36–38]. In the case of MSP2, the ability of an 8-residue dimer to bind to two mAbs able to recognise the parasite surface, 4D11 and 9G8, may make it a promising candidate for further study when coupled to a suitable carrier. In addition, the elimination of epitopes unlikely to contribute to protection, such as the 9H4 epitope, may improve vaccine efficacy. Currently, MSP2 analogues without the 9H4 epitope have been designed for immunisation trials in mice. The results obtained here increase our understanding of the conserved C-terminal region of MSP2 and will assist in the design of MSP2-based vaccine components that more closely match the conformational and antigenic properties of parasite MSP2.

## Materials and Methods

### Recombinant expression and purification of 3D7-MSP2 and MSP2<sub>207–224</sub>

3D7-MSP2 and the C-terminal region peptide MSP2<sub>207–224</sub> were produced recombinantly in *E. coli* BL21(DE3) Gold (Stratagene) cells using a codon-

optimised construct (Genscript) cloned into a thioredoxin-His<sub>6</sub> expression system, pET32a (Novagen). The protein was expressed and purified using methods specific for recombinantly expressed disordered proteins, as described previously [30].

### Antibody fragment preparation and purification

4D11 scFv and Fv antibody fragments were assembled using 4D11 V<sub>H</sub> and variable region light chain sequences optimised for expression in *E. coli* BL21-Gold, as described previously [26]. Synthetic genes (Genscript) were cloned into pET12a (Merck Millipore) for periplasmic expression in *E. coli*. Seed culture was grown overnight in the presence of tetracycline (15 µg/mL), ampicillin (100 µg/mL), and 4% glucose. The overnight culture was diluted 1000-fold with fresh LB medium containing ampicillin, tetracycline, and 4% glucose at 37 °C to an OD<sub>600</sub> of 0.8–1.0. The cells were then pelleted (5000g for 15 min at 4 °C), resuspended in fresh LB medium containing ampicillin and tetracycline with no glucose and grown for 2 h. IPTG was then added to a final concentration of 1 mM to induce expression. The induced cells were incubated at room temperature overnight and then harvested by centrifugation. The protein was extracted from the periplasm by osmotic shock by treatment with sucrose buffer [30 mM Tris, 2 mM EDTA, and 30% sucrose (pH 8.0)], followed by lysis of the periplasm with 5 mM MgCl<sub>2</sub> and 10 µg/mL DNase I at 4 °C.

The 4D11 antibody fragment was purified from the periplasmic extract by affinity purification using 3D7 MSP2-bound beads. The MSP2 affinity column was prepared by amide conjugation of full-length 3D7 MSP2 to hydroxysuccinimidyl-Sepharose resin (Sigma-Aldrich) following the manufacturer's protocol. Bound antibody fragments were washed extensively with 50 mM phosphate buffer (pH 7) and eluted from the MSP2 beads with 100 mM glycine (pH 2.7). Samples were neutralised immediately with 1 M Tris, concentrated in a 10-kDa MWCO centrifuge filter (Amicon Ultra-0.5, Merck-Millipore), and dialysed against 20 mM ammonium bicarbonate prior to lyophilisation and storage at –80 °C. The purity of each antibody fragment was assessed by SDS-PAGE and liquid chromatography-mass spectroscopy (LC–MS).

### Peptide synthesis

Peptides including the alanine scan of 3D7-MSP2<sub>207–222</sub> and MSP2<sub>215–222</sub>, along with the corresponding 8-residue monomeric analogues MSP2<sub>215–222</sub>(C219S) and MSP2<sub>215–222</sub>(C219α-aminobutyric acid), were synthesised in-house by standard 9-fluorenylmethoxycarbonyl (Fmoc) solid-phase chemistry using an automated peptide synthesiser 3 (PS3, Pti Instruments). The peptides were assembled by coupling 0.3 mmol (3 equiv) of Fmoc-protected amino acids to 0.1 mmol rink amide

AM resin (0.53 mmol/g loading). Coupling reactions were carried out for 50 min under the activation of 0.3 mmol (3 equiv) *O*-(1H-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate and 0.6 mmol (6 equiv) *N,N*-diisopropylethylamine. A double coupling was performed on the first residue of each peptide. Chain deprotection was carried out with 20% piperidine in dimethylformamide (DMF) for 2 min. The peptides were N-terminally capped with an acetyl moiety using 0.5 mmol (5 equiv) of acetic anhydride in 0.5 mmol (5 equiv) *N,N*-diisopropylethylamine. Dimerisation of the MSP2<sub>215–222</sub> peptide proved difficult, as conventional air oxidation over 5 days and iodine oxidation over 24 h were unsuccessful. Instead, the dimeric peptide was synthesised on resin using bis-Fmoc-L-cystine and constructed by fragment condensation.

Cleavage of the complete peptides was performed with trifluoroacetic acid (TFA):triisopropylsilane (TIPS):water [95:2.5:2.5 (vol/vol)]. The cleaved material was precipitated in cold diethyl ether overnight at –20 °C. The insoluble peptide material was spun down at 4000 rpm for 30 min at 0 °C, and the pellet washed twice in cold diethyl ether prior to removal of the organic phase. The crude peptide mixture was resuspended in 50% acetonitrile/0.1% TFA, filtered, and freeze-dried prior to further purification. Disulphide-bond-containing peptides were oxidised by air in 100 mM ammonium bicarbonate for 2 h before HPLC purification. MSP2 peptides were purified on a reverse-phase C18 column (Zorbax; 10 × 300 mm) using a linear gradient of 5 to 60% of solvent B (80% acetonitrile/9.9% water/0.1% TFA) against solvent A (0.1% TFA in water) over 1 h. The purity of MSP2 peptides was assessed by mass spectrometry (LC–MS; Supplementary Figs. S7, S8, and S9).

### Affinity measurements

The affinities of 4D11 IgG and Fv for synthetic peptides and recombinant MSP2 were determined by ITC (Microcal ITC-200, GE Healthcare) using murine mAb 6D8, 6D8 scFv, or 6D8 Fv. Titrations were performed in 10 mM Hepes and 100 mM NaCl (pH 7.4) at 25 °C. Antibody concentrations were 10 µM (IgG) or 20 µM (scFv/Fv), with peptide titrated from 200-µM stocks. Control titrations of peptide into buffer were performed, and the resulting heats of dilution were subtracted from the corresponding titration into antibody. Affinities for peptide epitopes and full-length 3D7-MSP2 to 4D11 IgG were measured by SPR (Biacore T200, GE Healthcare) using a Mouse Antibody Capture kit (GE Healthcare). Initial direct binding SPR assays with the alanine scan peptides were not feasible due to sensitivity issues caused by the small size of the peptides. Instead, affinities were measured by an SPR competition assay with 3D7-MSP2 immobilised on a CM5 chip. A standard curve was established by flowing eight

twofold dilutions of 4D11 IgG from 100 nM stock. Varying threefold dilutions of peptide were used from 10- $\mu$ M stock to compete with the immobilised antigen. We used 50 nM 4D11 IgG throughout the assay.

### X-ray data collection and structure refinement

4D11 Fv and synthetic peptide 3D7-MSP<sub>215–222</sub> were mixed at 1:1.5 (mol/mol) ratio in crystallisation buffer [20 mM Tris–HCl and 100 mM NaCl (pH 7.0)] for 1 h at room temperature. Unbound peptide and high molecular weight aggregates were removed by gel filtration (Superdex 75 10/300 GL, GE Healthcare). The complex was concentrated to 20–30 mg/mL by gentle centrifugation (Amicon Ultra 3 kDa, Merck-Millipore Pty) and cleared of precipitates by centrifugation at 10,000 rpm for 10 min, then filtered through a 0.22- $\mu$ m spin filter at 4 °C. Antibody concentrations were determined based on the absorbance at 280 nm ( $A^{1\%}_{280\text{nm}} = 1.6$ ) by NanoDrop (Thermo-Fisher Pty Ltd).

Crystals were grown using the hanging-drop vapour-diffusion method, with 1:1 (vol/vol) ratio of protein to reservoir solution and 0.5 mL well volume. Crystals grew in 23% (wt/vol) polyethylene glycol 3350, 0.1 M BIS Tris (pH 6.0), and 0.2 M MgCl<sub>2</sub> and were cryo-protected by the addition of 10% glycerol prior to data collection. Data were collected at 100 K using the Australian synchrotron macro crystallography MX1 beamline 3BM1 [39]. Data were processed using the CCP4 suite [40]. Diffraction images were indexed and integrated using iMosfilm, and Aimless was used to merge data sets and perform scale and average. The final structure was solved with merged data from two crystals collected under identical conditions. Then, 5% of each data set was flagged for calculation of  $R_{\text{Free}}$  [41]. Structure determination proceeded using the Molecular Replacement method and the program PHASER [42]. A search model was constructed from the crystal structure of mouse Fv 6D8 (PDB ID 4QYO) by removing the ligand from the search model and by mutation of residues to the 4D11 sequence using CHAINSAW from the CCP4 suite [40]. Initial electron density maps clearly showed unbiased features of the 3D7-MSP<sub>215–222</sub> peptide ligand between the two protein chains. All subsequent model building and structural validation were undertaken using Phenix [43,44] and COOT [45]. The refined  $2F_o - F_c$  electron density of the 3D7-MSP<sub>215–222</sub> peptide ligand is shown in Supplementary Fig. S10. Solvent molecules were added only if they had acceptable hydrogen-bonding geometry contacts of 2.5 to 3.5 Å with protein atoms or with existing solvent and were in good  $2F_o - F_c$  and  $F_o - F_c$  electron density. MolProbity was used to assess the quality of the refined structure [46].

### MD simulations

MD simulation was employed to gauge the stability of the 8-residue and 16-residue MSP2 epitopes in

complex with 4D11 Fv. The initial models of the 16-residue disulphide-bonded epitope MSP<sub>207–222</sub> were created using Modeller version 9.17 [47], with the 8-residue epitope from the crystal structure as a template. A total of 30 models were prepared, those with disfavoured rotamers and dihedral angles were removed, and a single model was chosen as a starting point for further MD studies. The simulations and analysis of 4D11 Fv in complex with both the monomeric 8-residue epitope and the modelled 16-residue disulphide-bonded peptide were performed using GROMACS version 5.1.2 software [48,49] and Amber 99SB-ILDN force field [50]. The protonation state of ionisable amino acids was set for a pH of 7.0, with Na<sup>+</sup> or Cl<sup>-</sup> counterions added to neutralise the system as required. The complex was placed in a rhombic dodecahedron box with a minimal distance between protein and the wall of the unit cell set to 10 Å and was solvated using the TIP3P water model. The solvated system was minimised using the steepest descent algorithm for 5000 steps. The system was equilibrated in three stages; first, a 100-ps MD simulation at 10 K with positional restraints on the protein (1000 kJ/mol-nm<sup>2</sup>) in an NVT ensemble. The V-rescale-modified Berendsen thermostat with a time coupling constant of 0.1 ps was used for temperature regulation [51]. This simulation was then repeated with no restraints. Finally, the system was equilibrated at 300 K for 100 ps in an NPT ensemble. The Parrinello–Rahman barostat with a pressure coupling constant of 2 ps was used to control the system pressure [52]. The LINCS algorithm was used to constrain covalent bonds, allowing a simulation time step of 2 fs [48]. A non-bonded interaction cutoff of 9 Å was used. Long-range electrostatics were calculated with the particle mesh Ewald method [53]. The production simulations were performed in an NPT ensemble at 300 K and 1 bar for 100 ns. Snapshots were stored every 2 ps, resulting in a total of 50,000 conformations, which were used in the MD trajectory analysis. Post-processing of the MD simulations was performed using the GROMACS utility *rmsdist*.

### Accession numbers

Coordinates and structure factors have been deposited in the Protein Data Bank with accession number 5TBD. Nucleotide sequences of the 4D11 V<sub>H</sub> and variable region light chain have been deposited in the Genbank database (KY038039 and KY03040, respectively).

### Acknowledgements

We thank David Zahra for his assistance in sequencing mAbs 4D11 and 9H4 and in the design

of antibody fragments. R.S.N. acknowledges fellowship support from the National Health and Medical Research Council of Australia. This work was supported in part by the Australian Government Research Training Program Scholarship and the National Health and Medical Research Council (project grant 1042520). We thank the Australian Synchrotron for beam time (MXCAP 8208) and technical assistance. The National Deuteration Facility is partially funded by the National Collaborative Research Infrastructure Strategy, an initiative of the Australian Federal Government.

## Appendix A. Supplementary Data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jmb.2017.02.003>.

Received 16 November 2016;

Received in revised form 22 January 2017;

Accepted 5 February 2017

Available online 8 February 2017

### Keywords:

malaria;  
antibody;  
merozoite surface protein 2;  
disordered protein;  
structure

### Abbreviations used:

MSP2, merozoite surface protein 2; GPI, glycosylphosphatidyl inositol; mAb, monoclonal antibody; IFA, immunofluorescence assay; SPR, surface plasmon resonance; Fv, variable fragment; scFv, single-chain Fv; ITC, isothermal titration calorimetry; V<sub>H</sub>, variable heavy chain; MD, molecular dynamics; FMOC, 9-fluorenylmethoxycarbonyl.

## References

- [1] WHO, World Malaria Report 2015, WHO Press, Geneva, Switzerland, 2015.
- [2] C.J.L. Murray, L.C. Rosenfeld, S.S. Lim, K.G. Andrews, K.J. Foreman, D. Haring, et al., Global malaria mortality between 1980 and 2010: a systematic analysis, *Lancet* 379 (2012) 413–431.
- [3] B.M. Greenwood, Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial, *Lancet* 386 (2015) 31–45.
- [4] J.S. Richards, J.G. Beeson, The future for blood-stage vaccines against malaria, *Immunol. Cell Biol.* 87 (2009) 377–390.
- [5] L. von Seidlein, P. Bejon, Malaria vaccines: past, present and future, *Arch. Dis. Child.* 98 (2013) 981–985.
- [6] P.R. Gilson, T. Nebl, D. Vukcevic, R.L. Moritz, T. Sargeant, T.P. Speed, et al., Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite *Plasmodium falciparum*, *Mol. Cell. Proteomics* 5 (2006) 1286–1299.
- [7] P.R. Sanders, L.M. Kats, D.R. Drew, R.A. O'Donnell, M. O'Neill, A.G. Maier, et al., A set of glycosylphosphatidyl inositol-anchored membrane proteins of *Plasmodium falciparum* is refractory to genetic deletion, *Infect. Immun.* 74 (2006) 4330–4338.
- [8] J.A. Smythe, R.L. Coppel, K.P. Day, R.K. Martin, A.M. Oduola, D.J. Kemp, et al., Structural diversity in the *Plasmodium falciparum* merozoite surface antigen 2, *Proc. Natl. Acad. Sci. U. S. A.* 88 (1991) 1751–1755.
- [9] B. Fenton, J.T. Clark, C.M. Khan, J.V. Robinson, D. Walliker, R. Ridley, et al., Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite surface antigen (MSA-2) of the malaria parasite *Plasmodium falciparum*, *Mol. Cell. Biol.* 11 (1991) 963–971.
- [10] Z. Feng, X. Zhang, P. Han, N. Arora, R.F. Anders, R.S. Norton, Abundance of intrinsically unstructured proteins in *P. falciparum* and other apicomplexan parasite proteomes, *Mol. Biochem. Parasitol.* 150 (2006) 256–267.
- [11] S. Olugbile, C. Kulangara, G. Bang, S. Bertholet, E. Suzarte, V. Villard, et al., Vaccine potentials of an intrinsically unstructured fragment derived from the blood stage-associated *Plasmodium falciparum* protein PFF0165c, *Infect. Immun.* 77 (2009) 5701–5709.
- [12] A.J. Guy, V. Irani, C.A. MacRaild, R.F. Anders, R.S. Norton, J.G. Beeson, et al., Insights into the immunological properties of intrinsically disordered malaria proteins using proteome scale predictions, *PLoS One* 10 (2015) 1–22.
- [13] X. Zhang, M.A. Perugini, S. Yao, C.G. Adda, V.J. Murphy, A. Low, et al., Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2, *J. Mol. Biol.* 379 (2008) 105–121.
- [14] C.G. Adda, V.J. Murphy, M. Sunde, L.J. Waddington, J. Schloegel, G.H. Talbo, et al., *Plasmodium falciparum* merozoite surface protein 2 is unstructured and forms amyloid-like fibrils, *Mol. Biochem. Parasitol.* 166 (2009) 159–171.
- [15] X. Yang, C.G. Adda, C.A. MacRaild, A. Low, X. Zhang, W. Zeng, et al., Identification of key residues involved in fibril formation by the conserved N-terminal region of *Plasmodium falciparum* merozoite surface protein 2 (MSP2), *Biochimie* 92 (2010) 1287–1295.
- [16] C.A. MacRaild, M.Ø. Pedersen, R.F. Anders, R.S. Norton, Lipid interactions of the malaria antigen merozoite surface protein 2, *Biochim. Biophys. Acta* 1818 (2012) 2572–2578.
- [17] R.F. Anders, C.G. Adda, M. Foley, R.S. Norton, Recombinant protein vaccines against the asexual blood stages of *Plasmodium falciparum*, *Hum. Vaccin.* 6 (2010) 39–53.
- [18] B. Genton, F. Al-Yaman, R. Anders, A. Saul, G. Brown, D. Pye, et al., Safety and immunogenicity of a three-component blood-stage malaria vaccine in adults living in an endemic area of Papua New Guinea, *Vaccine* 18 (2000) 2504–2511.
- [19] B. Genton, I. Betuela, I. Felger, F. Al-Yaman, R.F. Anders, A. Saul, et al., A recombinant blood-stage malaria vaccine reduces *Plasmodium falciparum* density and exerts selective pressure on parasite populations in a phase 1-2b trial in Papua New Guinea, *J. Infect. Dis.* 185 (2002) 820–827.
- [20] J.S. McCarthy, J. Marjason, S. Elliott, P. Fahey, G. Bang, E. Malkin, et al., A phase 1 trial of MSP2-C1, a blood-stage malaria vaccine containing 2 isoforms of MSP2 formulated with Montanide® ISA 720, *PLoS One* 6 (2011) e24413.

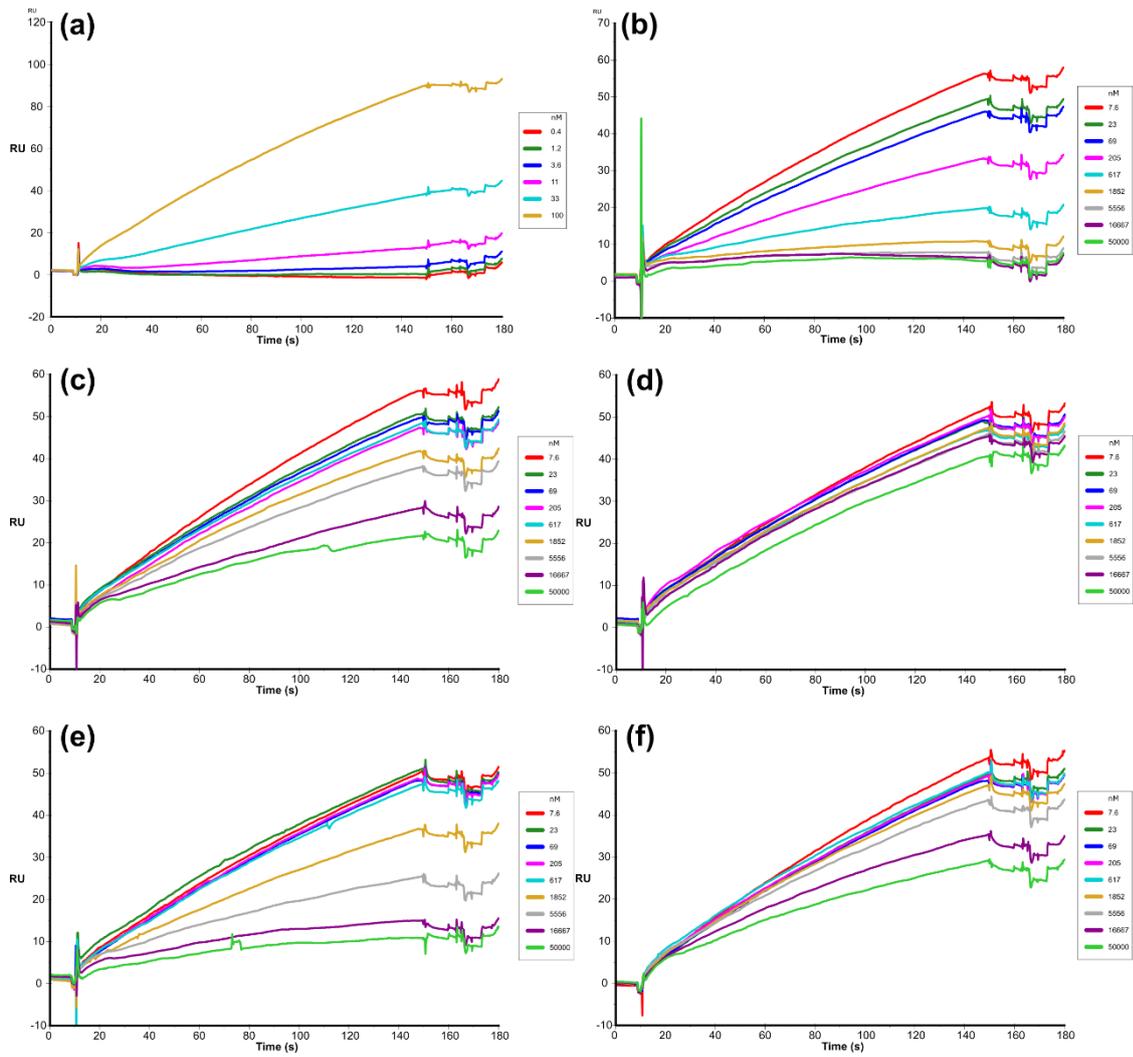
- [21] S.D. Polley, D.J. Conway, D.R. Cavanagh, J.S. McBride, B.S. Lowe, T.N. Williams, et al., High levels of serum antibodies to merozoite surface protein 2 of *Plasmodium falciparum* are associated with reduced risk of clinical malaria in coastal Kenya, *Vaccine* 24 (2006) 4233–4246.
- [22] B. Genton, F. Al-Yaman, I. Betuela, R.F. Anders, A. Saul, K. Baea, et al., Safety and immunogenicity of a three-component blood-stage malaria vaccine (MSP1, MSP2, RESA) against *Plasmodium falciparum* in Papua New Guinean children, *Vaccine* 22 (2003) 30–41.
- [23] M.J. Boyle, L. Reiling, R.F. Anders, J.G. Beeson, M.J. Boyle, L. Reiling, et al., Human antibodies fix complement to inhibit *Plasmodium falciparum* invasion of erythrocytes and are associated with protection against malaria, *Immunity* 42 (2015) 580–590.
- [24] B. Krishnarajuna, D. Andrew, C.A. MacRaild, R.A.V. Morales, J.G. Beeson, R.F. Anders, et al., Strain-transcending immune response generated by chimeras of the malaria vaccine candidate merozoite surface protein 2, *Sci. Rep.* 6 (2016) 20,613.
- [25] C.G. Adda, C.A. MacRaild, L. Reiling, K. Wycherley, M.J. Boyle, V. Kienzle, et al., Antigenic characterization of an intrinsically unstructured protein, *Plasmodium falciparum* merozoite surface protein 2, *Infect. Immun.* 80 (2012) 4177–4185.
- [26] R. Rouet, D. Lowe, K. Dudgeon, B. Roome, P. Schofield, D. Langley, et al., Expression of high-affinity human antibody fragments in bacteria, *Nat. Protoc.* 7 (2012) 364–373.
- [27] C.M. Lee, N. Iorno, F. Sierro, D. Christ, Selection of human antibody fragments by phage display, *Nat. Protoc.* 2 (2007) 3001–3008.
- [28] E.G. Hutchinson, J.M. Thornton, A revised set of potentials for beta-turn formation in proteins, *Protein Sci.* 3 (1994) 2207–2216.
- [29] E.A. Kabat, T.T. Wu, H. Bilofsky, Reid-Miller, H. Perry, Sequences of Proteins of Immunological Interest, National Institutes of Health, Bethesda, 1983.
- [30] C.A. MacRaild, M. Zachrdla, D. Andrew, B. Krishnarajuna, J. Novacek, L. Zidek, et al., Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2, *PLoS One* 10 (2015) 1–19.
- [31] C.A. Macraild, J.S. Richards, R.F. Anders, R.S. Norton, Antibody recognition of disordered antigens, *Structure* 24 (2016) 148–157.
- [32] V.N. Uversky, A.K. Dunker, Understanding protein non-folding, *Biochim. Biophys. Acta* 1804 (2010) 1231–1264.
- [33] H.J. Dyson, P.E. Wright, Intrinsically unstructured proteins and their functions, *Nat. Rev. Mol. Cell Biol.* 6 (2005) 197–208.
- [34] P.E. Wright, H.J. Dyson, Intrinsically disordered proteins in cellular signalling and regulation, *Nat. Rev. Mol. Cell Biol.* 16 (2014) 18–29.
- [35] R.A.V. Morales, C.A. MacRaild, J. Seow, B. Krishnarajuna, N. Drinkwater, R. Rouet, et al., Structural basis for epitope masking and strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate, *Sci. Rep.* 5 (2015) 10,103.
- [36] M. Scarselli, B. Aricò, B. Brunelli, S. Savino, F. Di Marcello, E. Palumbo, et al., Rational design of a meningococcal antigen inducing broad protective immunity, *Sci. Transl. Med.* 3 (2011) (91ra62).
- [37] J.S. McLellan, B.E. Correia, M. Chen, Y. Yang, B.S. Graham, W.R. Schief, et al., Design and characterization of epitope-scaffold immunogens that present the motavizumab epitope from respiratory syncytial virus, *J. Mol. Biol.* 409 (2011) 853–866.
- [38] J.S. McLellan, M. Chen, M.G. Joyce, M. Sastry, G.B.E. Stewart-Jones, Y. Yang, et al., Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus, *Science* 342 (2013) 592–598.
- [39] N.P. Cowieson, D. Aragao, M. Clift, D.J. Ericsson, C. Gee, S.J. Harrop, et al., MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron, *J. Synchrotron Radiat.* 22 (2015) 187–190.
- [40] The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D. Biol. Crystallogr.* 50 (1994) 760–763.
- [41] A.T. Brünger, Assessment of phase accuracy by cross validation: the free R value. Methods and applications, *Acta Crystallogr. D. Biol. Crystallogr.* 49 (1993) 24–36.
- [42] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, *J. Appl. Crystallogr.* 40 (2007) 658–674.
- [43] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, et al., PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D. Biol. Crystallogr.* 66 (2010) 213–221.
- [44] P.V. Afonine, R.W. Grosse-Kunstleve, N. Echols, J.J. Headd, N.W. Moriarty, M. Mustyakimov, et al., Towards automated crystallographic structure refinement with phenix.refine, *Acta Crystallogr. D Biol. Crystallogr.* 68 (2012) 352–367.
- [45] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr. D Biol. Crystallogr.* 60 (2004) 2126–2132.
- [46] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, et al., MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr.* 66 (2010) 12–21.
- [47] M.A. Marti-Renom, A. Stuart, A. Fiser, R. Sánchez, F. Melo, A. Sali, Comparative protein structure modeling of genes and genomes, *Annu. Rev. Biophys. Biomol. Struct.* 29 (2000) 291–325.
- [48] B. Hess, C. Kutzner, D. Van Der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, *J. Chem. Theory Comput.* 4 (2008) 435–447.
- [49] S. Pronk, Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, et al., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, *Bioinformatics* 29 (2013) 845–854.
- [50] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, et al., Improved side-chain torsion potentials for the Amber ff99SB protein force field, *Proteins Struct. Funct. Bioinforma.* 78 (2010) 1950–1958.
- [51] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.* 126 (2007).
- [52] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, *J. Appl. Phys.* 52 (1981) 7182–7190.
- [53] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, *J. Chem. Phys.* 103 (1995) 8577–8593.

## Supplementary Data

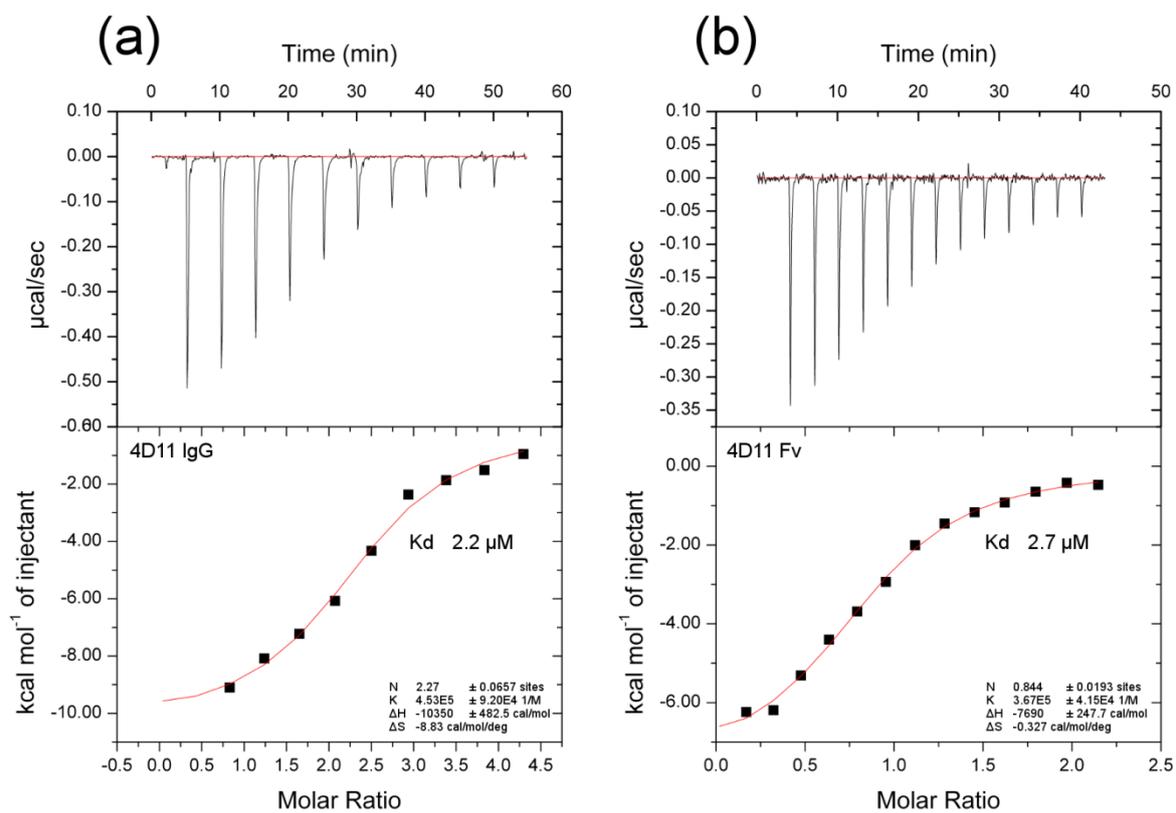
### Structure and characterisation of a key epitope in the conserved C-terminal domain of the malaria vaccine candidate MSP2

Jeffrey Seow <sup>1</sup>, Rodrigo A. V. Morales <sup>1</sup>, Christopher A. MacRaild <sup>1</sup>, Bankala Krishnarjuna <sup>1</sup>, Sheena McGowan <sup>2</sup>, Tamir Dingjan <sup>1</sup>, Garima Jaipuria <sup>3</sup>, Romain Rouet <sup>4</sup>, Karyn L. Wilde <sup>5</sup>, Hanudatta S. Atreya <sup>3</sup>, Jack S Richards <sup>6</sup>, Robin F. Anders <sup>7</sup>, Daniel Christ <sup>4</sup>, Nyssa Drinkwater <sup>2</sup> and Raymond S. Norton<sup>1</sup>

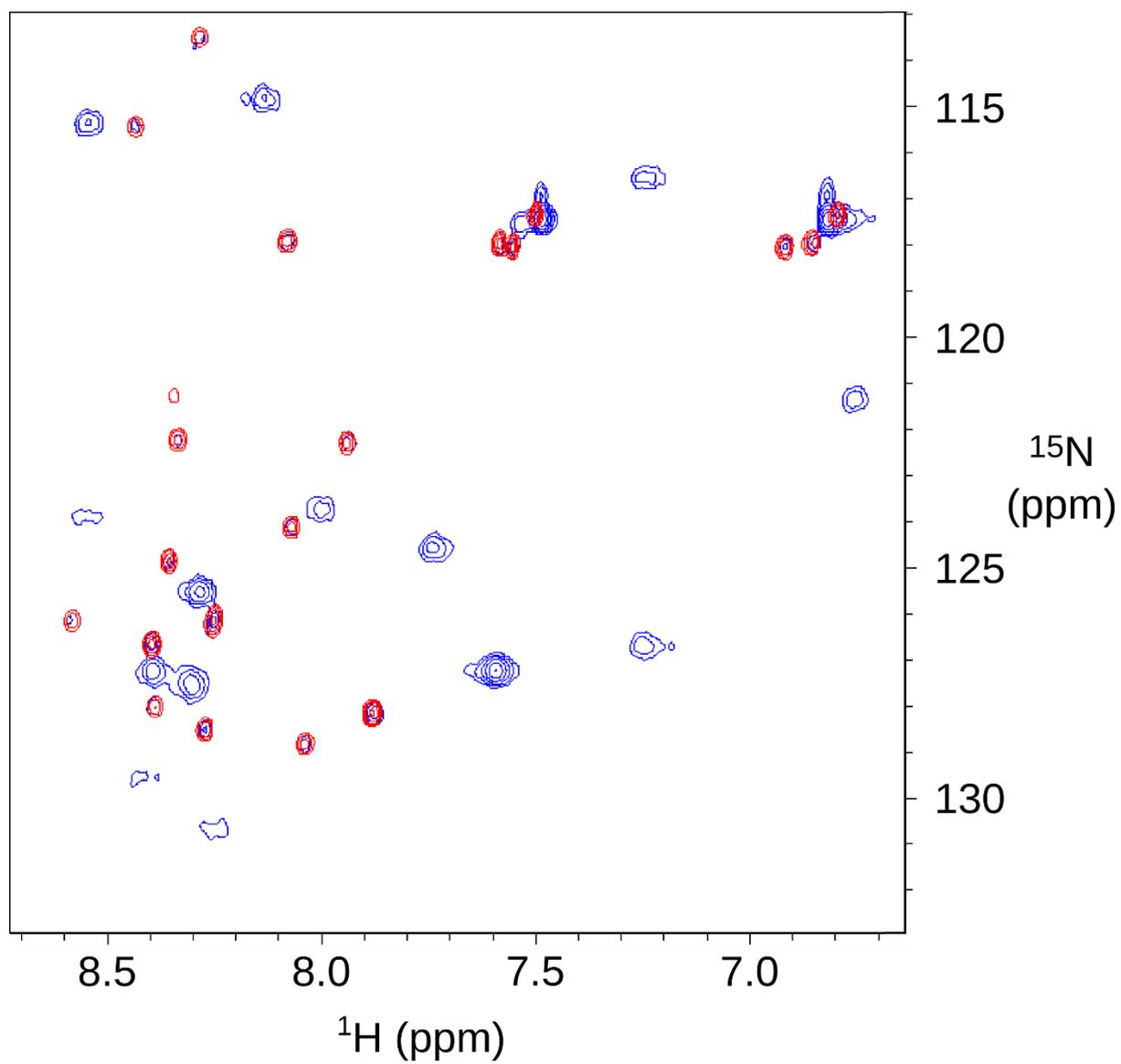
1. Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, 3010, Australia
2. Department of Microbiology, Monash University, Clayton, 3168, Australia
3. NMR Research Centre, Indian Institute of Science, Bangalore, India
4. Garvan Institute of Medical Research, Darlinghurst, 2010, Australia
5. National Deuteration Facility, Australian Nuclear Science and Technology Organisation, Lucas Heights, 2234, Australia
6. Centre for Biomedical Research, The Burnet Institute, Melbourne, 3004, Australia
7. Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, 3086, Australia



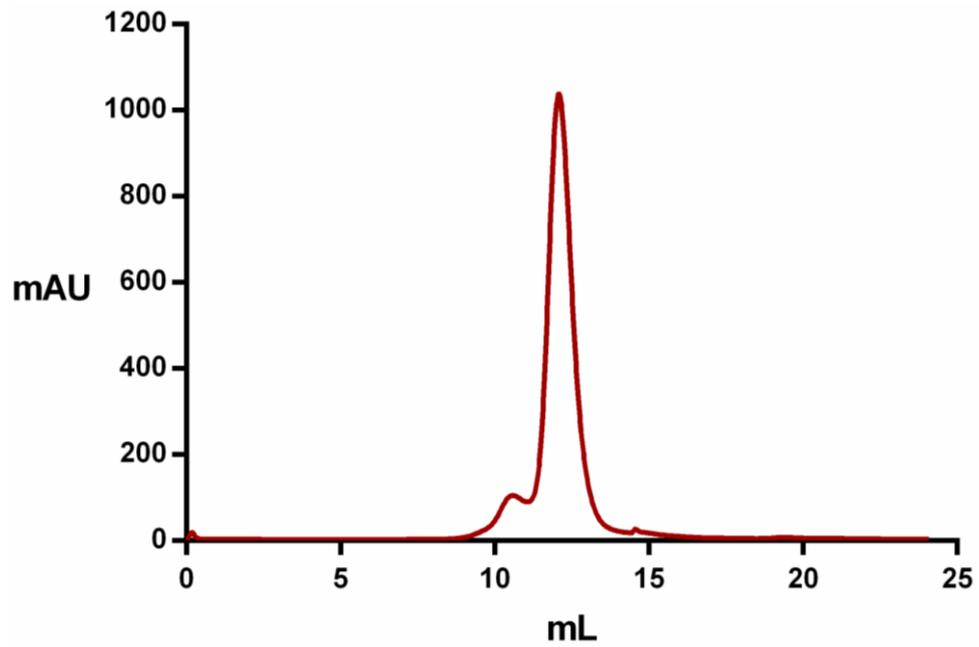
**Figure S1.** SPR sensograms of competition binding assay using alanine scan peptides and 4D11 IgG (a) 4D11 IgG calibration curve (b) Wild type 3D7-MSP2<sub>207-222</sub> (c) K216A (d) E217A (e) N218A (f) G220A. These results are summarised in **Figure 1c** and **Table 1** in the main text.



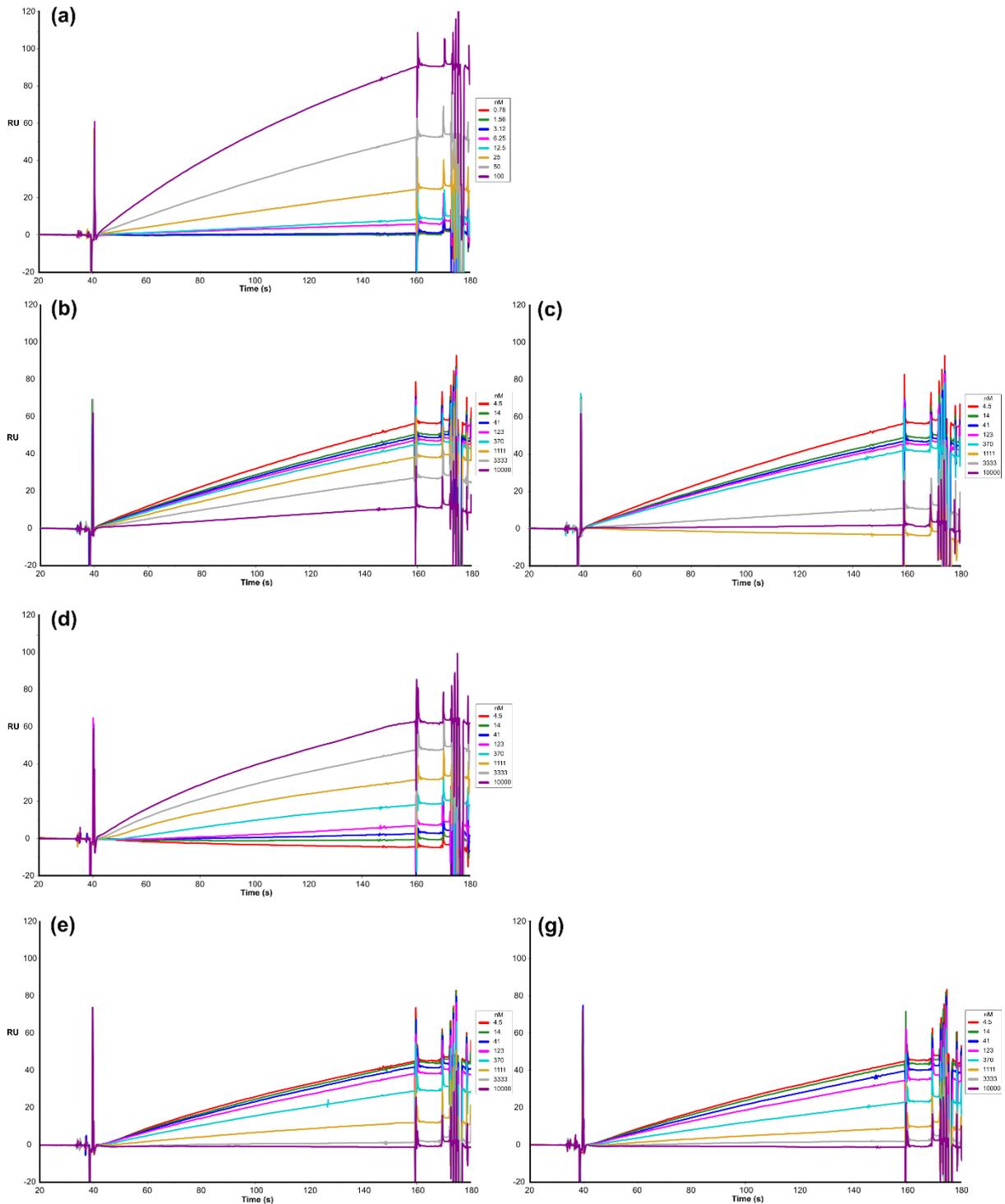
**Figure S2.** ITC titrations of 3D7-MSP<sub>207-222</sub> into **(a)** 4D11 IgG and **(b)** 4D11 Fv. Solid red lines are lines of best fit with corresponding  $K_d$  values shown.



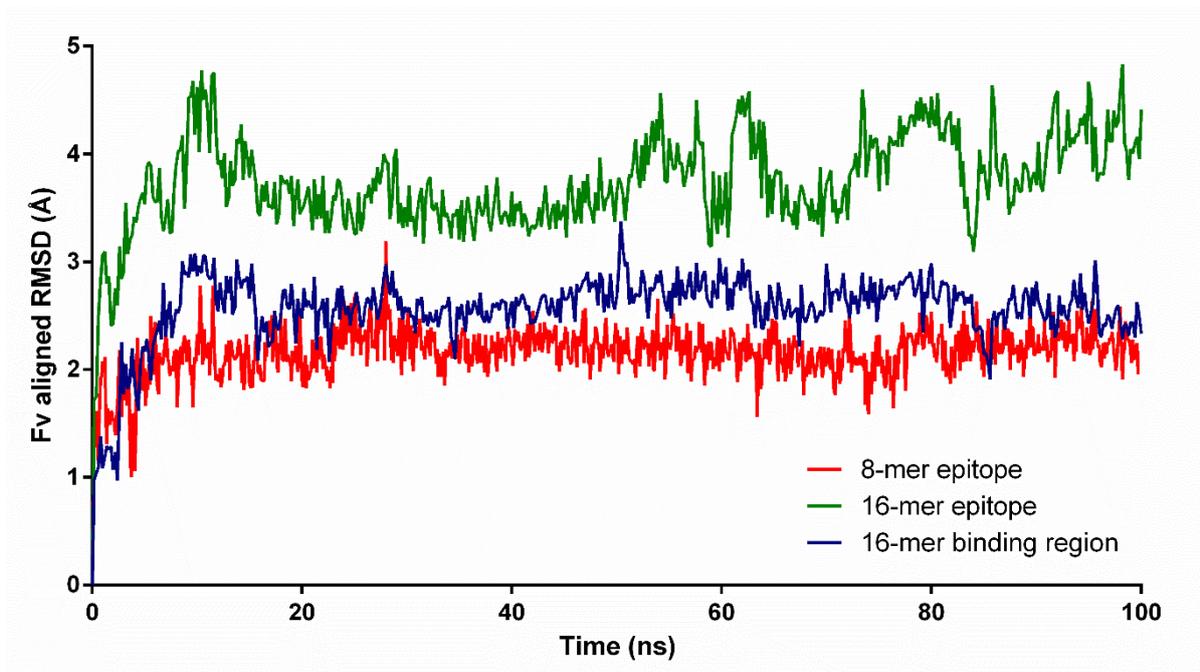
**Figure S3.** [ $^1\text{H}$ ,  $^{15}\text{N}$ ]-HSQC spectrum of  $^{15}\text{N}$ -labelled 3D7-MSP<sub>207-224</sub> in the absence (red) and presence (blue) of 4D11 scFv.



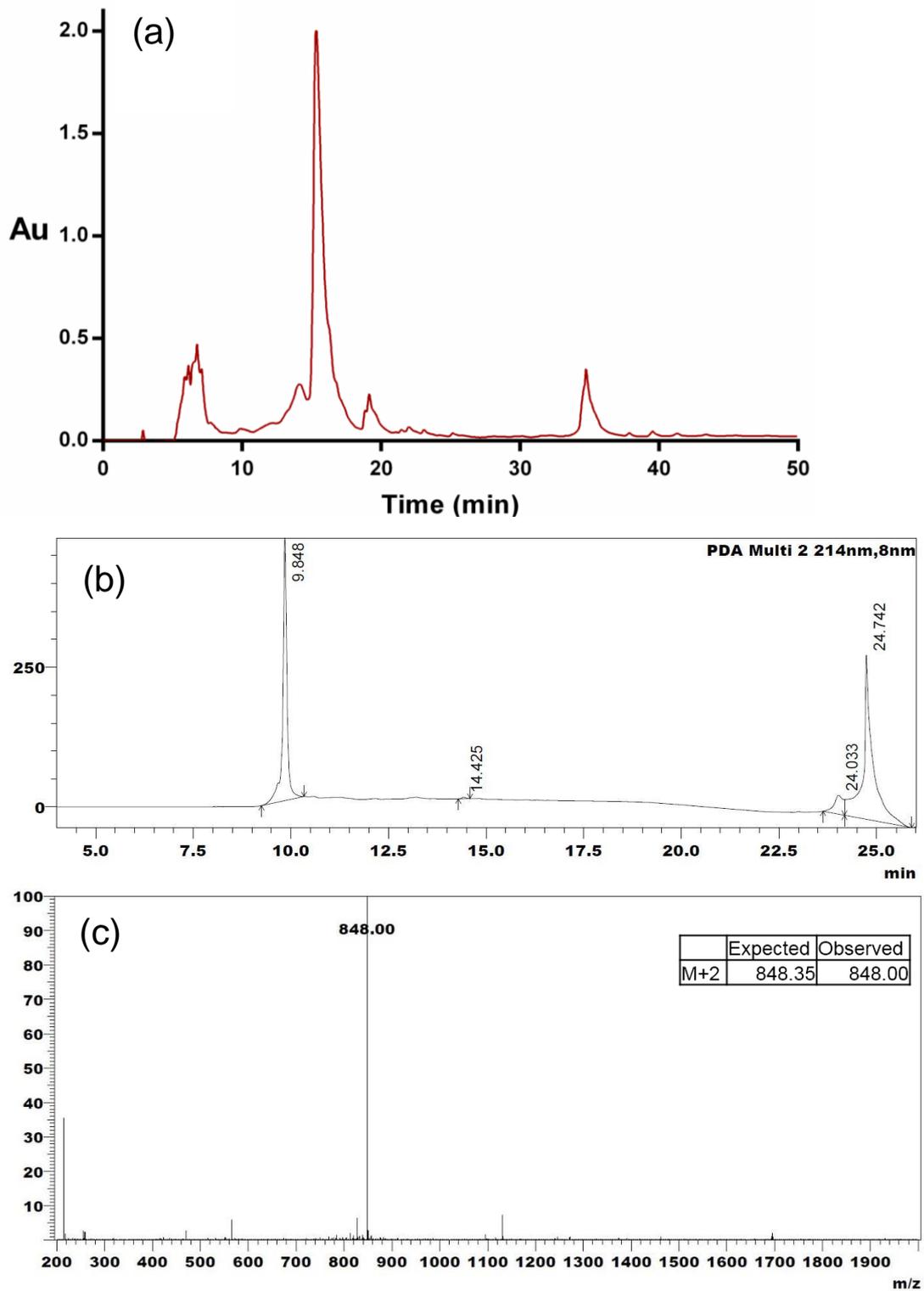
**Figure S4.** Size-exclusion chromatography of 4D11 Fv complexed with 3D7-MSP<sub>215-222</sub> using a Superdex 75 10/300 column (GE healthcare) at 0.5 mL/min shows that the sample is predominantly monomeric.



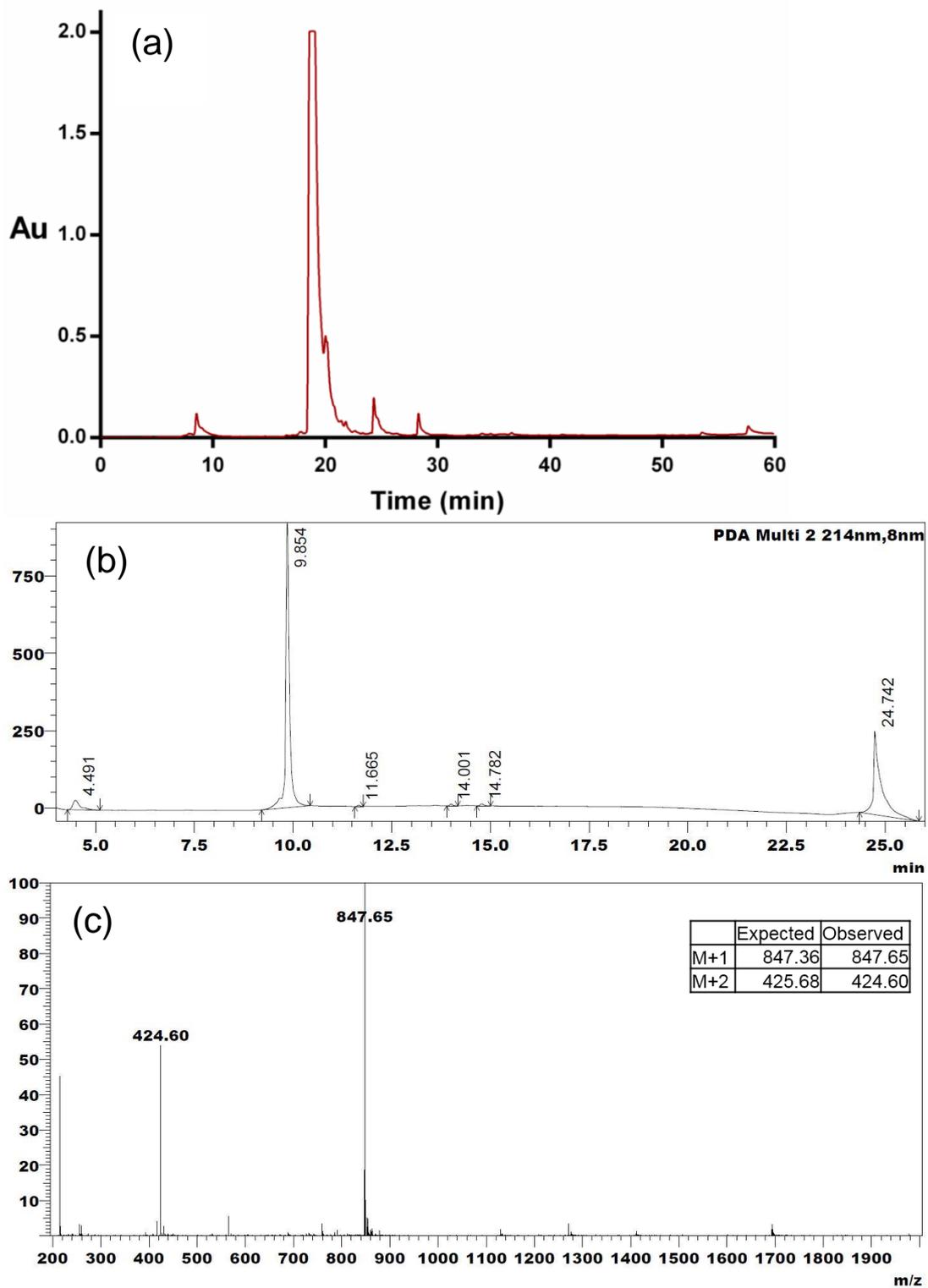
**Figure S5.** SPR sensograms of competition binding assay using monomeric and dimeric peptides MSP2<sub>215-222</sub> 8-residue peptides with both 4D11 IgG and 9G8 IgG **(a)** 4D11 IgG calibration curve **(b)** 4D11 IgG + monomeric MSP2<sub>215-222</sub> **(c)** 4D11 IgG + dimeric MSP2<sub>215-222</sub> **(d)** 9G8 IgG calibration curve **(e)** 9G8 IgG + monomeric MSP2<sub>215-222</sub> **(f)** 9G8 IgG + dimeric MSP2<sub>215-222</sub>. These results are summarised in **Figure 4** in the main text.



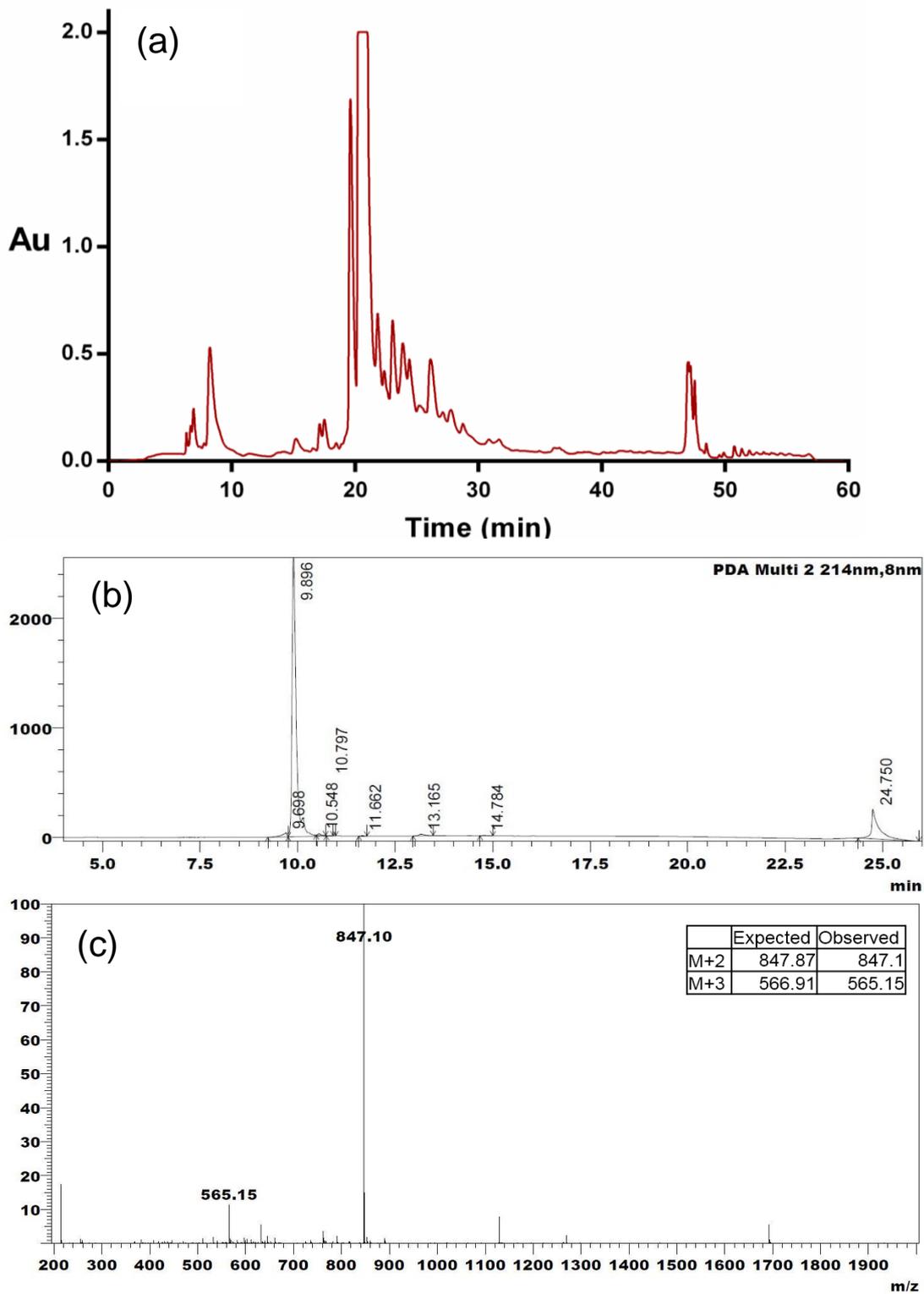
**Figure S6.** RMSD of each epitope after alignment of 4D11 Fv over 100 ns with respect to equilibrated MD simulation starting conformation. All residues of 4D11 Fv were aligned before calculation of each epitope RMSD. Neither the conformation nor the position of the bound epitope change significantly throughout the simulation. The binding region (NKENCGAA) of the 16-residue modelled epitope has an identical RMSD to the 8-residue peptide throughout the simulation suggesting that the flexible N-terminal extensions are responsible for the higher RMSD observed from the full 16-residue peptide.



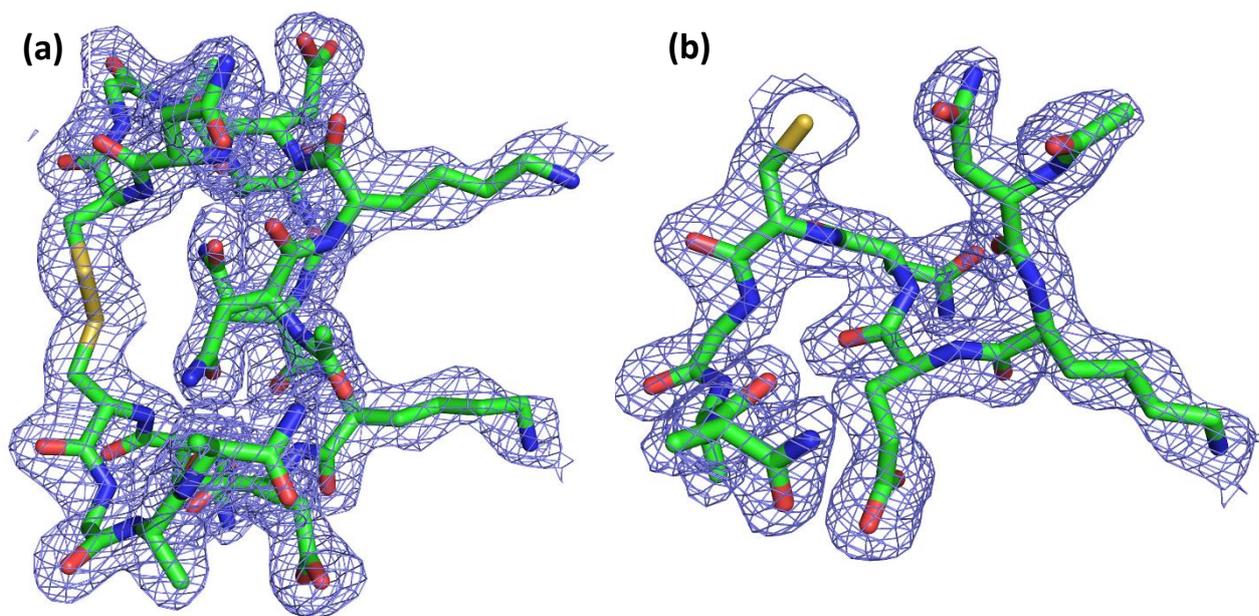
**Figure S7.** (a) HPLC purification of crude synthetic 3D7-MSP<sub>207-222</sub> peptide, (b) LC of collected fraction and (c) mass-spectrometry profile of purified peptide



**Figure S8.** (a) HPLC purification of crude synthetic monomeric 3D7-MSP<sub>215-222</sub>, (b) LC of collected fraction and (c) mass-spectrometry profile of purified peptide



**Figure S9.** (a) HPLC purification of crude synthetic dimeric 3D7-MSP<sub>215-222</sub> peptide, (b) LC of collected fraction and (c) mass-spectrometry profile of purified peptide



**Figure S10.**  $2F_o - F_c$  electron density map of MSP<sub>215-222</sub> shown at  $1.0 \sigma$  contour level of the **(a)** dimer and **(b)** monomer.

# Chapter 5

Guiding the immune response of MSP2,  
and intrinsically disordered vaccine  
candidate

## 5.1 Chapter introduction

In the previous Chapter, a high-resolution crystal structure of 4D11 Fv in complex with its cognate minimal binding epitope (NKENCGAA) was solved. 4D11 mAb recognises an epitope in the highly conserved C-terminal region of MSP2 that is accessible on native MSP2 by IFA and western blot, and hence presents a promising target for immune protection. The crystal structure revealed that the peptide epitope adopted a  $\beta$ -bend ribbon conformation when bound to 4D11. Key interactions between the epitope and the 4D11 paratope in the structure agreed with alanine scan data, which showed that residues Lys216, Glu217, Asn218 and Gly220 were essential for binding. The bound peptide was found to form an asymmetrical homo-dimer in the process of crystallisation via formation of a disulphide bond between the free cysteines in the sequence. MD simulations and SPR competition binding assays were utilised to show that this dimeric conformation was relevant to the native wild-type disulphide-bonded epitope.

In this Chapter, the structure-based vaccine design approach is applied to MSP2 to develop peptide immunogens capable of eliciting protective antibodies with broad strain specificity. The 4D11-bound dimeric peptide was used as a template for the design of a series of conformationally-constrained peptides to be conjugated to carrier protein KLH for immunisation in mice. The dimer peptides were stabilised by linking the N- and C- termini of each peptide to form linear or backbone-cyclised analogues. MD simulations, SPR, and direct ELISA were used to inform the design of these peptides and prioritise those that would be moved forward to animal immunisations. Additionally, key residues for binding of 9H4 were determined by alanine scanning. The 9H4 mAb binds to an epitope in close proximity to the 4D11 epitope, but is unable to recognise parasite MSP2. In an effort to bias the antibody response towards 4D11-like antibodies and away from 9H4-like antibodies, two peptides were designed. The first was a 17-residue peptide with the wild-type sequence (A1), the second was identical, but with a single point mutation (K209A) that removed a key residue for 9H4 binding. Antibody responses for each peptide group were analysed by ELISA and a peptide array was used to determine specificities.

This Chapter is presented in the form of a manuscript in preparation. Some experimental aspects are incomplete and require further work, specifically, merozoite ELISAs that are currently being undertaken by collaborators. However, the results presented are largely complete and should be ready for publication soon after thesis submission.

## **Guiding the immune response of MSP2, an intrinsically disordered malaria vaccine candidate**

Jeffrey Seow<sup>1</sup>, Christopher A. MacRaid<sup>1</sup>, Sreedam C. Das<sup>1</sup>, Rodrigo A. V. Morales<sup>1</sup>, Racardo Ataide<sup>2</sup>, Bankala Krishnarjuna<sup>1</sup>, Mitchell Silk<sup>1</sup>, David Chalmers<sup>1</sup>, Jack Richards<sup>2</sup>, Robin F. Anders<sup>3</sup> and Raymond S. Norton<sup>1</sup>

1. Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, 3010, Australia
2. Centre for Biomedical Research, The Burnet Institute, Melbourne, 3004, Australia
3. Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, 3086, Australia

Key words: malaria; merozoite surface protein 2; disordered protein; peptide vaccines; structural vaccinology

## Abstract

Vaccine trials involving the intrinsically disordered vaccine candidate, merozoite surface protein 2, have shown that the antigen is in part responsible for a reduction in parasite density. However, this reduction was strain-specific, suggesting that polymorphic regions of MSP2 are immuno-dominant. These polymorphisms are localised in a central variable region that is flanked by highly conserved N- and C-terminal regions. One strategy to bypass the hurdle of strain-specificity is to bias the immune response towards conserved regions. Two mouse monoclonal antibodies, 4D11 and 9H4, recognise the conserved C-terminal region of MSP2. Although both epitopes overlap, they each show different antigenic properties, with 4D11 able to recognise parasite MSP2 whilst 9H4 cannot, suggesting that the 4D11 epitope is accessible on the parasite surface. In this work, the emerging field of structure-based vaccine design is applied to MSP2 using a crystal structure of 4D11 Fv in complex with its cognate minimal binding epitope. Molecular dynamics simulations and surface plasmon resonance have informed the design a series of constrained peptides based on the 4D11-bound epitope structure. These peptides were conjugated to keyhole limpet hemocyanin and immunised in mice. High to moderate peptide-specific antibody titres were generated in all groups. The specificities of antibody responses revealed that single point mutations can focus vaccine responses towards more favourable epitopes. This rational approach to vaccine design may be useful not only to MSP2-based malaria vaccines, but also other intrinsically disordered antigens.

## Introduction

The global health burden of malaria is still significant today, with over 200 million cases and 430,000 deaths per year (1). Despite the recent regulatory approval of the pre-erythrocytic RTS,S/A01 vaccine, the modest efficacy in young infants in Phase III trials calls for further research towards more effective and robust malaria vaccines (13). An ideal malaria vaccine will probably need to be multi-valent, targeting multiple stages of the *Plasmodium* life cycle. Inhibition of pre-erythrocytic stages would minimise or prevent symptomatic infection whilst vaccines targeting the blood-stage would combat breakthrough infection. This work focuses on a merozoite surface protein 2 (MSP2), a blood-stage antigen found in abundance on the parasite surface. All MSP2 proteins can be characterised into two allelic families, 3D7 and FC27, that are defined by a central variable region. Flanking this polymorphic region are highly conserved N-terminal and C-terminal regions.

Phase I-IIb immunisation trials in Papua New Guinean children with the Combination B vaccine, of which 3D7-MSP2 was a component, showed a 62% reduction in parasite densities (22, 30). However, this was skewed towards parasites expressing 3D7-MSP2. This suggested that, although immunisation with MSP2 is efficacious, the response is highly strain-specific. Subsequently, a vaccine containing both allelic forms of MSP2 was tested in Phase I trials and was able to induce antibodies that could inhibit parasite growth and invasion (26). Recent efforts to address the problem of strain-specificity involve production of MSP2 chimaeras composed of the central variable region of both 3D7 and FC27 and the conserved N- and C-terminal regions (81). Animal immunisations with these constructs yielded a robust immune response towards both MSP2 alleles.

An alternative method to circumvent the problem of strain-specificity is targeting epitopes in the conserved regions an antigen. The C-terminal conserved region of MSP2 is recognised by five mouse monoclonal antibodies (mAb), 4D11, 9G8, 9H4, 6C9 and 1F7, which bind to overlapping epitopes (44). Despite the close proximity of binding sites, these antibodies have different antigenic properties. 4D11 and 9G8 mAbs show strong recognition of parasite MSP2 by western blot and immunofluorescence assay (IFA), whilst 9H4, 6C9 and 1F7 mAbs are unable to bind parasites, suggesting that these epitopes are less accessible in native MSP2 on parasite surface. Guiding affinity maturation towards antibodies able to recognise native MSP2 may yield a more efficient immune response.

Peptide vaccines may be a useful tool in guiding the immune response towards these key epitopes (60). They offer a cleaner antigen preparation with minimal allergic and

autoimmune response owing to their synthetic origin, whilst avoiding redundant or detrimental epitopes not associated with protection. Moreover, with the ability to include multiple epitopes in the vaccine formulation, different life stages of the parasite can be targeted for improved efficacy (103). The disordered nature of MSP2 also lends itself to a peptide vaccine owing to a lack of discontinuous or conformational epitopes in the antigen (44). Although this is advantageous, with an increased flexibility, a large variety of conformations can be sampled by recombinant MSP2 in solution, not all of which will guide affinity maturation of antibodies able to recognise parasite MSP2 efficiently. Structural vaccinology, an emerging field in rational vaccine design, involves the use of antigen structure to inform the design of better vaccine candidates. This approach has shown promise in a variety of disease conditions, including meningococcus B (66), respiratory syncytial virus (67, 68), Group B Streptococcus, (69) and HIV (70–72), although the applications to disordered protein antigens and peptide vaccines remain largely unexplored.

Recently, the crystal structure of a key epitope in the C-terminal region of MSP2, bound to the 4D11 variable fragment (Fv) was solved at 2.2 Å resolution (104). The structure revealed that the bound epitope adopts a  $\beta$ -bend ribbon conformation that was stabilised by two intramolecular hydrogen bonds. The peptide crystallised as a homo-dimer via the free cysteine in the peptide sequence, with both peptides able to bind their separate 4D11 Fv antibody fragment. In this work, we use the dimeric 4D11-bound epitope structure as a template for the rational design of a better MSP2-based vaccine.

## Results

### *Alanine scan of 9H4 and 4D11 epitope*

In our previous work, the minimal binding epitope and key residues involved in 4D11 recognition was confined to the 8-residue peptide 3D7-MSP2<sub>215-222</sub> with Lys216, Glu217 and Asn218 crucial for binding (44, 104). To understand the epitope of 9H4 and compare it with 4D11, the same alanine scan peptides were used to probe 9H4 IgG binding (**Figure 1A and 1B**). Affinities for and 9H4 IgGs were measured by surface plasmon resonance (SPR), using a competition binding assay described previously (104). Mutation of Lys209, Thr212 or Asp213 to Ala resulted in a significant loss of 9H4 binding, suggesting that they are crucial for 9H4 recognition (**Table 1**). Intriguingly, mutations to residues Gly214-Glu217 resulted in a 10-fold increase in binding to 9H4 relative to the wild-type sequence. Attempts to crystallise 9H4 and 6C9 Fv with their cognate epitopes and the tighter binding mutants were unsuccessful, so the

structural determinants of this binding interaction are still unknown. The epitopes of 6C9 and 9H4 were located closer to the N-terminus of the peptide and no residues involved in binding were shared with the 4D11 epitope (**Figure 1C and 1D**).

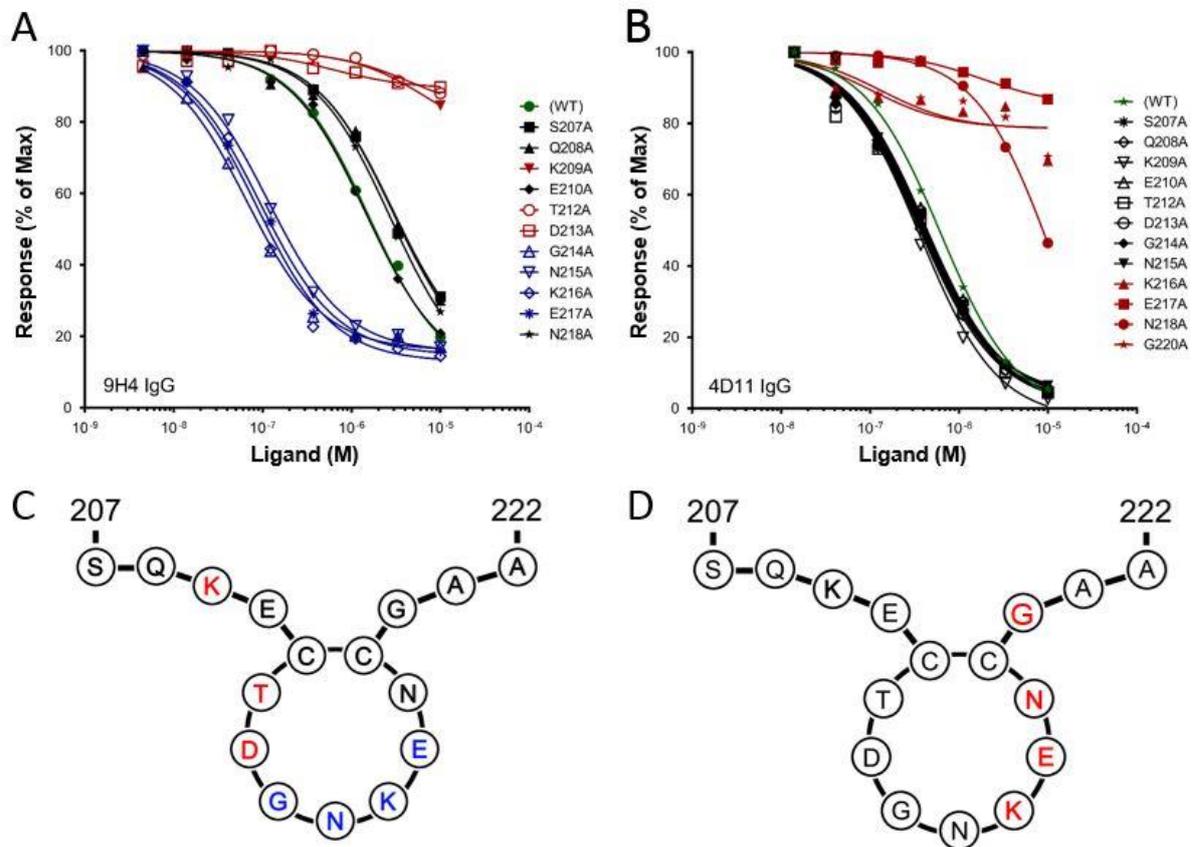


Figure 1: SPR competition assay using alanine scan 16-residue peptides of 3D7-MSP2<sub>207-222</sub> with (A) 9H4 and (B) 4D11 IgG. Schematic representation of disulphide bonded epitope sequence for both (C) 9H4 and (D) 4D11 IgG indicating the location of key residues. Black indicates alanine mutants with no effect on binding, the green shows the wild-type binding of 3D7-MSP2<sub>207-222</sub>, red indicate the alanine mutations that decreased the binding and blue indicate mutations that increased binding.

Table 1. Alanine scan of 16-residue epitope MSP2<sub>207-222</sub> to determine key residues for binding of 9H4 and 4D11 IgG

Peptide	Sequence	$K_a$ against 9H4 IgG ( $\mu$ M)	$K_a$ against 4D11 IgG ( $\mu$ M)
<b>3D7-MSP2<sub>207-222</sub> (WT)</b>	SQKECTDGNKENC $C$ GAA	1.5	0.9
<b>S207A</b>	<b>A</b> SQKECTDGNKENC $C$ GAA	2.9	0.5
<b>Q208A</b>	S <b>A</b> QKECTDGNKENC $C$ GAA	2.9	0.4
<b>K209A</b>	SQ <b>A</b> ECTDGNKENC $C$ GAA	7.3	0.3
<b>E210A</b>	SQK <b>A</b> CTDGNKENC $C$ GAA	1.4	0.5
<b>T212A</b>	SQKE <b>C</b> ADGNKENC $C$ GAA	4.8	0.5
<b>D213A</b>	SQKECT <b>A</b> GNKENC $C$ GAA	0.6	0.5
<b>G214A</b>	SQKECTD <b>A</b> NKENC $C$ GAA	0.07	0.4
<b>N215A</b>	SQKECTDG <b>A</b> KENC $C$ GAA	0.1	0.5
<b>K216A</b>	SQKECTDGN <b>A</b> ENC $C$ GAA	0.09	21.0
<b>E217A</b>	SQKECTDGNK <b>A</b> NC $C$ GAA	0.09	219.4
<b>N218A</b>	SQKECTDGNKE <b>A</b> CGAA	2.7	6.0
<b>G220A</b>	SQKECTDGNKENC <b>A</b> AA	2.9	42.2

#### Design of peptides for immunisation

Two peptides were designed following results from the alanine scan, a 16-residue peptide with the wild-type sequence MSP2<sub>207-222</sub> (A1) and an identical peptide with the single point mutation (K209A) (**Figure 2**). These peptides were designed to determine if removal of residues key to 6C9/9H4 binding would skew the immune response away from the less accessible 6C9/9H4 epitope and towards 4D11-like antibodies. For immunisations, the commonly-used carrier protein, KLH was chosen. As conjugation to KLH required a free thiol, a cysteine residue was included in all of the peptides.

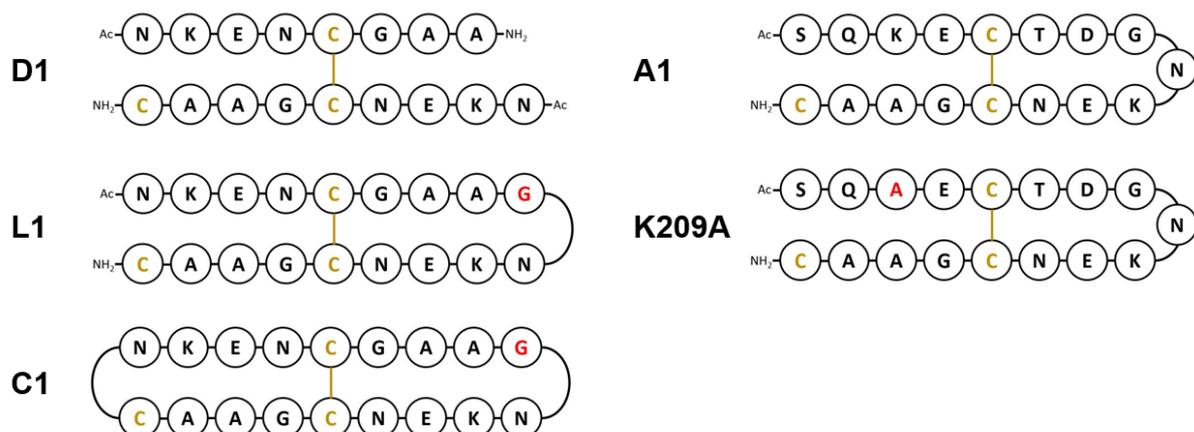


Figure 2: Schematic representation of dimeric 4D11 epitope peptides dimer (D1), linear (L1), backbone cyclised (C1), wildtype sequence 17-residue peptide (A1) and its single point mutant K209A, which has had the key residue Lys209 that is important for 9H4 binding removed.

Another series of peptides was designed using the homo-dimeric, 4D11-bound structure as a template. The symmetry and close proximity of the N- and C-termini (5.0 Å) of peptides in the disulphide-mediated dimer (**Figure 3**) allowed for linkage at one or both termini, making linear (L1-3) or cyclic peptides (C1-3), respectively. Glycine was chosen as a linker owing to its structural flexibility and easy integration by peptide synthesis. It was possible that the additional constraints introduced to the peptides could obstruct the epitope conformation or impede synthesis. Hence, molecular dynamics (MD) simulations were employed to further inform peptide epitope design and to determine the ideal linker lengths in the linear and cyclic peptides.

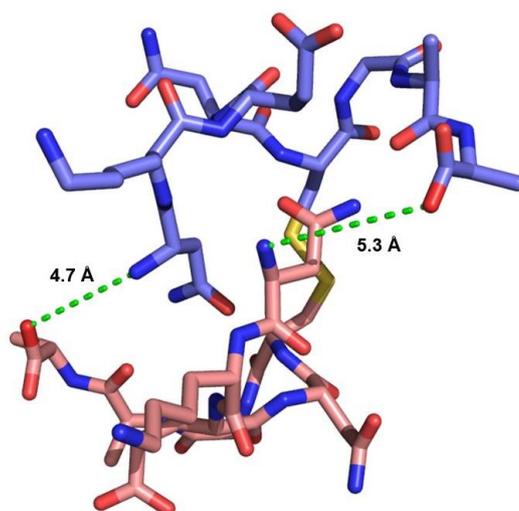


Figure 3: Crystal structure of 4D11-bound homodimer peptide shows close proximity between N- and C-termini of peptides, green dashed lines indicate distance between the peptide termini.

With linker lengths ranging from 1-3 Gly residues, all linear and cyclic peptides proved to be stable during a 100 ns MD simulation (**Supplementary Figure S1**). Each peptide had two repeats of the 4D11 epitope, hence, both epitopes were aligned, individually, with the 4D11-bound peptide from the crystal structure. RMSD values were calculated using the backbone atoms of residues involved in 4D11 binding (MSP2<sub>15-19</sub>). As expected, owing to the additional constraints, the cyclic peptides were more stable, with less variation and flexibility than the linear peptides. In most cases, the conformation of both epitopes in the peptide were within 2.5 Å RMSD of the bound conformation. However, there were instances of one side of the peptide straying from the bound conformation more than the other. This was most pronounced in C3, with one face of the peptide having an RMSD of 2 Å and the other face an RMSD of 3 Å.

Following MD simulations, each peptide was synthesised by standard Fmoc synthesis. Orthogonal protection of cysteines was used to ensure that the desired disulphide bond was formed, leaving the C-terminal cysteine for KLH conjugation. This complicated the synthesis of backbone-cyclised peptides as cysteine de-protection conditions also cleaved the peptide from the resin. Consequently, disulphide bond formation and cyclisation coupling were performed in solution, increasing the time and difficulty of synthesis. To determine if the additions incorporated into the dimer sequence hindered binding to 4D11, SPR was used to measure their binding affinity to 4D11 IgG (**Table 2**). All peptides were able to bind 4D11, with the tightest binding peptides, A1 and K209A having a  $K_d$  of 0.2  $\mu$ M and the weakest, D1, 2.38  $\mu$ M. All linear and cyclic peptides showed tight binding to 4D11, and in both cases the shorter linker length of a single Gly residue had the strongest binding. Moving forward, L1 and C1 were chosen for conjugation to KLH and immunisations along with D1, A1 and K209A.

Table 2. Binding affinities of peptides for KLH conjugation were determined by a SPR competition binding assay. Cysteines required for conjugation to maleimide-activated KLH are indicated in yellow, additional linker residues and mutations are shown in red.

Name	Sequence	$K_d$ against 4D11 IgG ( $\mu$ M)
8mer	NKENCGAA	0.24
D1	NKENCGAA NKENCGAAC	2.38
L1	NKENCGAAGNKENCGAAC	0.47
L2	NKENCGAAGGNKENCGAAC	0.67
L3	NKENCGAAGGGNKENCGAAC	0.58
C1	c [NKENCGAAGNKENCGAAC]	0.44
C2	c [NKENCGAAGGNKENCGAACG]	0.68
C3	c [NKENCGAAGGGNKENCGAAGCG]	1.04
A1	SQKECTDGNKENCGAAC	0.20
K209A	SQAECTDGNKENCGAAC	0.20

*Epitopes in the peptide-KLH conjugates were able to bind to 4D11 and 9H4 IgG*

ELISA was used to assess if 4D11 and 9H4 IgG could still recognise the KLH-conjugated peptide (**Figure 4A**). Each peptide-KLH conjugate was coated on the plate and probed with serial dilutions of 4D11 or 9H4 IgG. All peptide conjugates were able to bind 4D11, with C1-KLH and full-length 3D7+FC27 MSP2 the tightest binders, followed by L1-KLH and D1-KLH. The A1-KLH and K209A-KLH conjugates had weaker binding when compared to other peptide conjugates. This is probably due to the bivalent nature of the dimeric peptides, which presents two faces for 4D11 recognition. The close proximity of the C-terminal cysteine conjugation to the 4D11 epitope may also hinder binding. The K209A-KLH conjugate was unable to bind 9H4 IgG, confirming that the mutation of key residue Lys209 to Ala had successfully inhibited the 9H4 recognition (**Figure 4B**).

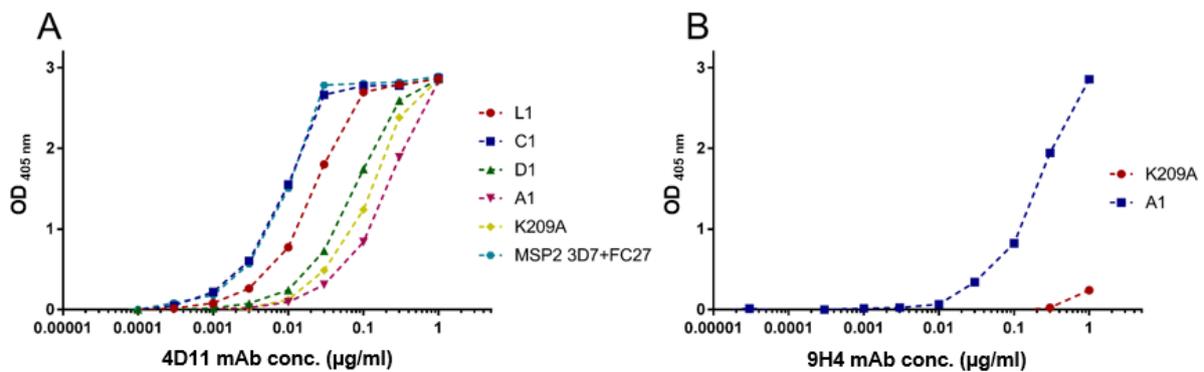


Figure 4: (A) Direct ELISA indicates that all peptide-KLH conjugates are recognised by 4D11 IgG, (B) K209A-KLH is unable to bind to 9H4, indicating that the 9H4 epitope had been removed successfully.

*Peptide-KLH conjugates were able to induce epitope-specific immune response*

Specific anti-peptide titres were determined by coating ELISA plates with their corresponding peptide-bovine serum albumin (BSA) conjugates. High antibody titres were seen for peptides L1-KLH, C1-KLH and the full-length MSP2 mixture (endpoint titre  $> 3.5 \times 10^6$ ) (**Figure 5**). A1-KLH had moderate antibody titre (endpoint titre  $> 1.5 \times 10^6$ ) followed by D1-KLH and K209A-KLH, with an endpoint titre of  $> 7 \times 10^5$ . To confirm that the maleimide linker present in each conjugate was not eliciting a response, a non-related peptide was conjugated to BSA and coated on the plate. None of the pooled sera showed any response to this conjugate at 1000-fold dilution (**Supplementary Figure 2**). These results indicate that immunisation with peptide-conjugates can induce a peptide-specific response.

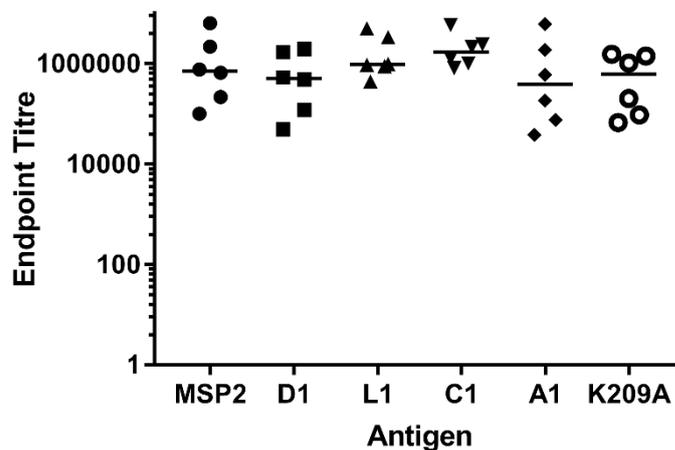


Figure 5: Mouse sera, taken two weeks after final immunisation with antigen, contain antibodies that recognise their corresponding peptide-BSA conjugate or in the case of the MSP2 mix, recombinant 3D7 and FC27 MSP2 coated on the ELISA plate. Each point is the mean of two replicate wells for an individual mouse. Lines indicates group median. Endpoint titres were calculated using a cut-off value three standard deviations greater than  $OD_{405}$  for naïve mouse serum ( $\sim 0.015$ ).

### *Peptide-KLH conjugates directed the immune response towards the 4D11 epitope*

To further characterise the response against these dimeric peptides, specificity was determined by indirect ELISA. Individual mouse sera from each group was probed using a biotinylated peptide array corresponding to nine peptides spanning the epitope recognised by 4D11 and 9H4 mAbs (peptides A-I) (**Table 3**). Peptides C and D contained the 9H4 epitope whilst peptides E to G encompass the 4D11 epitope. The 3D7 + FC27 MSP2 mixture elicited a response towards peptides C-F, encompassing both 9H4 and 4D11 epitopes (**Figure 6**). However, the response against MSP2 was largely dominated by a single mouse in the group. In contrast, responses against each of the dimeric peptide-KLH conjugates (D1, C1, and L1) were more consistent throughout the sample size and were to the 4D11 epitope. Only one instance of cross reactivity was observed, with one mouse in the L1 cohort generating antibodies that recognised peptide D. The lack of binding to the adjacent peptide C suggests that these antibodies are recognising residues closer to the C-terminal, probably the GNKENC residues that are present in the immunising peptide.

### *K209A mutation caused bias in immune response against 4D11 epitope*

The removal of the key residue Lys209 in K209A peptide resulted in a drastic change in antibody specificity when compared to the wild-type A1 peptide. The A1 peptide generated antibodies primarily against peptides C and D, the 9H4 epitope, suggesting that the epitope is immunodominant over the 4D11 epitope. The single mutation in the K209A peptide successfully removed the 9H4 epitope and shifted antibody recognition towards the 4D11 epitope represented by peptides E and F.

Table 3. Peptide array used to probe antibody specificities, peptides C-D encompass the 9H4 epitope (red) and peptides E-G contain the 4D 11 epitope (green).

Peptide name	Sequence
A	HPQNTSDSQKECT
B	QNTSDSQKECTDG
C	SDSQKESTDGNKE
D	SQKECTDGNKENC
E	ECTDGNKENCGAA
F	TDGNKENCGAATS
G	NKENCGAATSLLN
H	ENCGAATSLLNNS
I	CGAATSLLNSSN

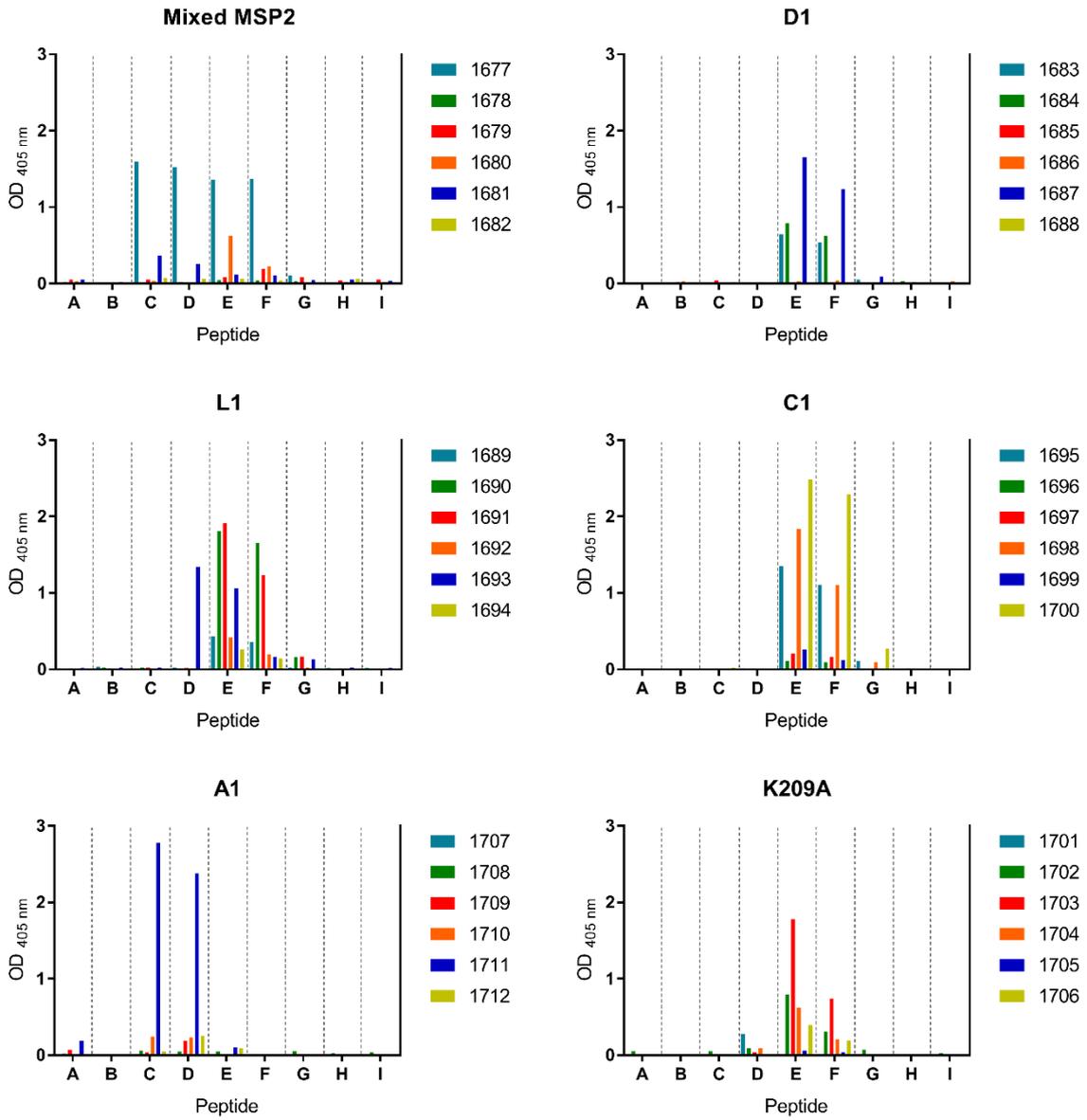


Figure 6: Antibody specificities of individual mice from each group was determined by peptide array (peptides A-I)

## Discussion

Despite the many decades of research, an effective vaccine for some pathogens, such as malaria, remains elusive. The reverse vaccinology approach has brought about a new generation of vaccine candidates (105, 106). Advances in multi-strain genome sequencing have enabled not only the identification of novel surface antigens, but also the ability to discern polymorphic and conserved regions. However, vaccines designed on protein sequence alone are not adequate to induce antibodies that can recognise native antigen. The direct evolution of the reverse vaccinology approach, structural vaccinology, may be the next step in rational vaccine design. Similar to structure-based drug design (107, 108), although largely unexplored, the structural vaccinology strategy employs structural data from X-ray crystallography, NMR spectroscopy and cryo-EM to inform the rational design of novel vaccine antigens.

Early applications of the approach involved grafting of key epitopes or electrostatic surfaces onto proteins (66) and scaffolds (71, 109). Crystal structures have also been used to assist in the design of stabilised analogues of otherwise unstable antigens (68). Significant effort is often required to mimic the conformational epitopes in these structured antigens. In contrast, the intrinsically disordered nature of MSP2 presents a unique challenge that may benefit from the structure-based strategy. Although conformational disorder has commonly been suggested to impede the affinity maturation of specific and high-affinity antibody responses, in-depth analysis and comparison of ordered and disordered epitope-antibody interactions have shown otherwise (41). Indeed, antibody affinity was found to be only weakly dependent on disorder, with similar antibody binding affinity seen in both disordered epitopes and their structured counterparts. Furthermore, disordered epitopes were found to be shorter in length than ordered epitopes, making more efficient interactions with the antibody paratope. This challenges the long-held belief that the entropic costs associated with the transition from disorder to order, are detrimental to antibody binding, and reinforces the notion that disordered antigens are bona fide targets of antibody recognition.

An important consideration when constraining disordered epitopes into their antibody-bound or native conformations is the effects on antibody maturation. Introducing rigidity in epitopes may be favourable to reduce the entropic costs of binding, but it may also lead to activation of fewer germline antibodies. There are examples in literature of multiple neutralising antibodies recognising the same disordered epitope in noticeably different conformations (110–112). Presumably, the flexibility of these epitopes allows for affinity maturation down multiple paths. In this study, although 4D11-like antibodies may be desirable,

confining the immune response to a narrow repertoire of B-cell precursors may be less efficient and some antigen flexibility may be required to elicit a robust immune response.

The utilisation of peptides in structure-based vaccine design is surprisingly unexplored, perhaps due to their inability to present conformational epitopes. However, this is not an obstacle for disordered antigens, such as MSP2, with epitopes that are invariably linear and amenable to peptide-based strategies. The high customisability of peptides, coupled with detailed structural analyses of epitope-antibody interactions, means that peptides can be constrained into their antibody-bound conformation without introducing potentially distracting epitopes. The well-established field of peptide synthesis offers a variety of strategies for introduction of constraints and stabilisation of secondary structure(62, 113–115). Here we use the crystal structure of 4D11 Fv in complex with its minimal binding epitope to inform the design of a series of peptide epitope vaccines. The homo-dimeric peptide present in this structure led to the design and synthesis of a series of dimeric peptides constrained as linear or backbone cyclised analogues. These peptides were all capable of inducing a peptide-specific response specific to the 4D11 epitope.

The stark change in antibody specificities induced by the mutant peptide K209A when compared to the wildtype sequence peptide A1, presents a simple strategy to bias the immune response to favourable epitopes. Recently, chimeric MSP2 antigens that include variable and conserved regions of 3D7 and FC27 alleles were able to induce a broad immune response to both strains in mice (81). Mutations such as K209A can be implemented into recombinant analogues of MSP2 such as these chimeras and may enhance antibody production to the more accessible 4D11 epitope.

These results provide a promising platform for further work on MSP2-based peptide vaccine candidates. Despite the wide use of the growth inhibition assay in evaluation of anti-merozoite antibodies, there is a poor correlation with protective immunity (26, 31). To further characterise the immune response to these peptide vaccines, more reliable functional correlates of protection such as antibody-mediated complement-dependant inhibition (33), antibody-dependent cellular inhibition (31), and opsonic phagocytosis (116) assays will be required. Collectively, these results suggest that, with structural knowledge of antibody-bound epitopes, structural vaccinology can be applied to customisable peptide vaccines to induce a highly specific antibody response.

## Materials and methods

### *SPR*

Affinities for peptide epitopes and dimer peptides to 4D11 and 9H4 IgG were measured by SPR (Biacore T200, GEHealthcare) by a competition assay method developed previously (104). Briefly, 3D7-MSP2 was immobilised on a CM5 chip and a standard curve was established by flowing eight twofold dilutions of 4D11/9H4 IgG from 100 nM stock. To determine binding, eight three-fold dilutions of peptide, from 10  $\mu$ M stock were added with 50 nM 4D11/9H4 IgG to compete with the immobilised antigen.

### *MD simulations*

MD simulation was used to assist in the design of dimer peptides before synthesis. To evaluate if the strain introduced via disulphide bonds and backbone cyclisation would preclude the peptide from adopting a conformation capable of binding 4D11, each peptide was constructed in Maestro (Schrödinger 2016-4) using the 4D11 Fv bound homo-dimeric peptide in the crystal structure (PDB ID: 5TBD) as a template. Each model was checked for favourable rotamers and dihedral angles. The simulations and analysis of each peptide were performed with GROMACS version 5.1.2 software and GROMOS 54A7 forcefield (117, 118). The complex was placed in a cubic box with a minimal distance between protein and the wall of the unit cell set to 10 Å and was solvated using the TIP3P water model. The solvated system was minimised using the steepest descent algorithm for 5000 steps. The system was equilibrated in three stages; first, a 100-ps MD simulation at 10 K with positional restraints on the protein (1000 kJ/mol/nm<sup>2</sup>) in an NVT ensemble. The V-rescale-modified Berendsen thermostat with a time coupling constant of 0.1 ps was then used for temperature regulation (119). This simulation was then repeated with no restraints. Finally, the system was equilibrated at 300 K for 100 ps in an NPT ensemble. The Parrinello–Rahman barostat with a pressure coupling constant of 2 ps was used to control the system pressure (120). The LINCS algorithm was used to constrain covalent bonds, allowing a simulation time step of 2 fs (117). A non-bonded interaction cutoff of 9 Å was used. Long-range electrostatics were calculated with the particle mesh Ewald method (121). The production simulations were performed in an NPT ensemble at 300 K and 1 bar for 100 ns. Post-processing of the MD simulations was performed using the GROMACS utility rmsdist.

### *Peptide synthesis*

All peptides were synthesised in-house by standard 9-fluorenylmethoxycarbonyl (Fmoc) solid-phase chemistry using an automated peptide synthesiser 3 (PS3, Pti Instruments). All linear peptides (L1-L3, A1, K209A and Alanine scan peptides) were assembled by coupling 0.3 mmol (3 equiv.) of Fmoc-protected amino acids to 0.1 mmol rink amide AM resin (0.53 mmol/g loading). Coupling reactions were carried out for 50 min under the activation of 0.3 mmol (3 equiv.) O-(1H-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate and 0.6 mmol (6 equiv.) N,N-diisopropylethylamine (DIPEA). A double coupling was performed on the first residue of each peptide. Chain deprotection was carried out with 20% piperidine in dimethylformamide (DMF) for 2 min. The peptides were N-terminally capped with an acetyl moiety using 0.5 mmol (5 equiv.) of acetic anhydride in 0.5 mmol (5 equiv.) DIPEA. Orthogonal protection of cysteines were employed to ensure that the correct disulphide connectivity was present, with the C-terminal cysteine free for conjugation to BSA or KLH. 4-methoxytrityl (mmt) was used as a thiol protecting group for cysteines taking part in disulphide bond formation. Selective removal of mmt was performed with trifluoroacetic acid (TFA):triisopropylsilane (TIPS): dichloromethane (DCM) [1:2:97 (vol/vol)] for 2 x 30 min. The disulphide bond for linear peptides were formed using 0.2 mmol (2 equiv.) N-chlorosuccinimide (NCS) in DMF for 2 h (122, 123).

Cyclic peptides (C1-C3) were assembled on 2-chlorotriyl chloride resin (1.4 mmol/g loading). Deprotection of mmt and cleavage from the resin was performed concurrently with TFA:TIPS:DCM [1:2:97 (vol/vol)] for 2 x 1 h. Disulphide bonds were formed with air oxidation in 0.175 mM triethylamine (TEA), DMF for 2 days. The linear peptides were then cyclised in solution with PyClock (3 equiv.) and DIPEA (10 equiv.) in DMF for 16 h.

The dimer peptide D1 was synthesised by first assembling the short CGAA and CGAAC peptides separately on rink amide AM resin. The CGAA peptide was fully cleaved from the resin whilst the mmt on the N-terminal cysteine of CGAAC was removed by the methods discussed previously. To form the disulphide bond between these peptides, cleaved CGAA peptide was mixed with resin-bound CGAAC in 0.2 mmol (2 equiv.) NCS, DMF for 2 h. The remaining amino acids were assembled by standard Fmoc solid-phase chemistry and the N-terminals were acetylated. As peptide conjugation with KLH was to be established via activated maleimide on the carrier protein, a free cysteine was included in the design of each peptide.

Cleavage of the complete peptides was performed with TFA:TIPS:dimethylbenzene (DMB) [92.5:2.5:5 (vol/vol)]. The cleaved material was precipitated in cold diethyl ether and

insoluble peptide material was spun down at 4000 rpm for 15 min at 0 °C, and the pellet washed twice in cold diethyl ether prior to removal of the organic phase. The crude peptide mixture was resuspended in 50% acetonitrile/0.1% TFA and freeze-dried prior to further purification. All peptides were purified on a reverse-phase C18 column (Vydac; 10 × 300 mm) using a linear gradient of 5 to 60% of solvent B (80% acetonitrile/9.9% water/0.1% TFA) against solvent A (0.1% TFA in water) over 1 h. The purity of peptides was assessed by mass spectrometry (LCMS; **Supplementary Figures S3, S4, S5, S6 and S7**)

#### *Protein preparation and mice immunisation experiments*

Peptide was conjugated to maleimide-activated KLH (Sigma-Aldrich) following the manufacturer's protocol. Briefly, 1 mg of each peptide was mixed with 1 mg KLH in the provided buffer for a 200-fold excess of peptide to carrier protein. The reaction mixture was degassed under nitrogen and mixed at room temperature for 2 h. Any unreacted peptide was removed by dialysis against PBS using a 10 kDa cutoff membrane. The extent of conjugation was determined using Ellman's reagent. Peptide-KLH conjugate stock solution was stored at 4°C until further use. Prior to immunisation, peptide-KLH conjugates were diluted to 1 mg/mL in PBS and formulated with Montanide ISA720 at 3:7 ratio (antigen:adjuvant) to a final concentration of 0.3 mg/mL. The 3D7 + FC27 MSP2 mix was formulated using 0.15 mg/mL of each allelic form. Female C57BL/6 mice (n=6 per group) were inoculated subcutaneously with 100 µL containing 30 µg of antigen at weeks 0, 4 and 8, then euthanized at week 10. Sera was collected and stored at -80°C.

#### *ELISA*

To determine peptide-KLH binding to 4D11 and 9H4, Maxisorp 96-well microtitre plates (Nunc) were coated overnight at 4°C with 2 µg/mL peptide KLH in PBS. The plates were blocked with 1% BSA in PBS for 1 h before adding 100 µL of 4D11 and 9H4 IgG in eleven half log<sub>10</sub> serial dilutions starting from stock at 1 µg/mL. After 1 h incubation at 4°C and washing, antigen-bound antibodies were detected with goat anti-mouse IgG (1:2000 dilution) and freshly prepared 2,2-azinobis(3-ethylbenzthiazolinesulfonic acid (ABTS) substrate (1 mM) in citric acid buffer pH 4.2 containing horseradish peroxidase (HRP). The absorbance was read at 405 nm using a microplate reader. For determination of peptide-specific antibody titres, peptide-BSA conjugate was coated on the plate instead and sera were added at 1000-fold dilution. End-point titres were taken as the x-axis intercept of the dilution curve at an absorbance value three standard deviations (s.d.) greater than OD 405 for naïve mouse serum.

To generate the ELISA data for peptide array specificity, nine biotinylated 13-residue peptides (A-I), that overlap by 2 or 3 residues, spanning the C-terminal region of MSP2 were used (44, 81). Maxisorp 96-well microtitre plates (Nunc) were coated overnight at 4°C with 1 µg/mL streptavidin in PBS. The plates were blocked with 1% BSA in PBS for 2 h before adding 1:500 biotinylated peptides. Individual sera samples were diluted 1:1000 in blocking solution before addition to the peptide array. Antibody was detected with the same protocol as mentioned above.

## Supplementary

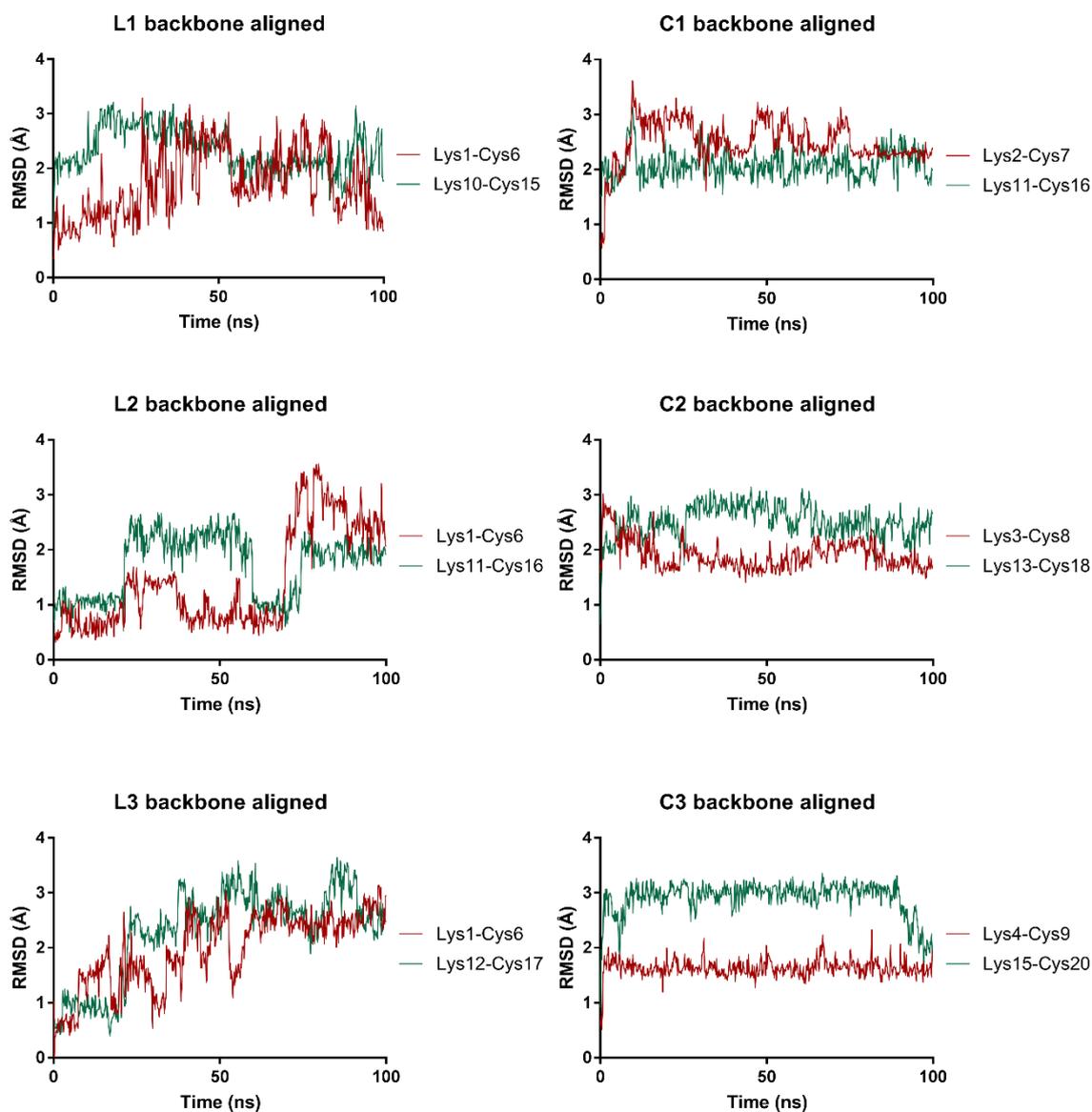


Figure S1. RMSD of the 4D11 epitope shape over 100 ns MD simulation with respect to the conformation of the 4D11-bound epitope from the crystal structure (PDB ID: 5TBD). Only residues involved in 4D11 binding were backbone aligned (MSP<sub>2217-220</sub>). As each peptide was dimeric, RMSD for each 4D11 epitopes was aligned separately, shown in green and red.

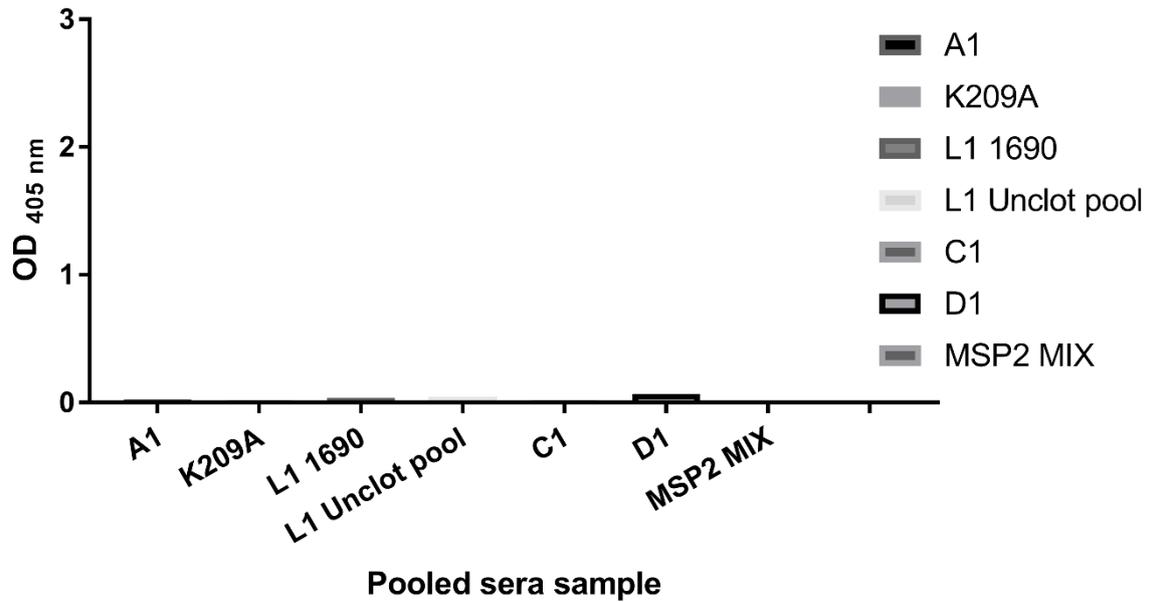


Figure S2. To determine if there was an immune response towards the maleimide linker moiety in KLH, direct ELISA was used to show if pooled sera had any binding to a non-relevant peptide (SKWICANRSVCPI) maleimide-conjugated to BSA. None of the pooled sera had significant response towards this conjugate, indicating that the maleimide linker was not immunogenic.

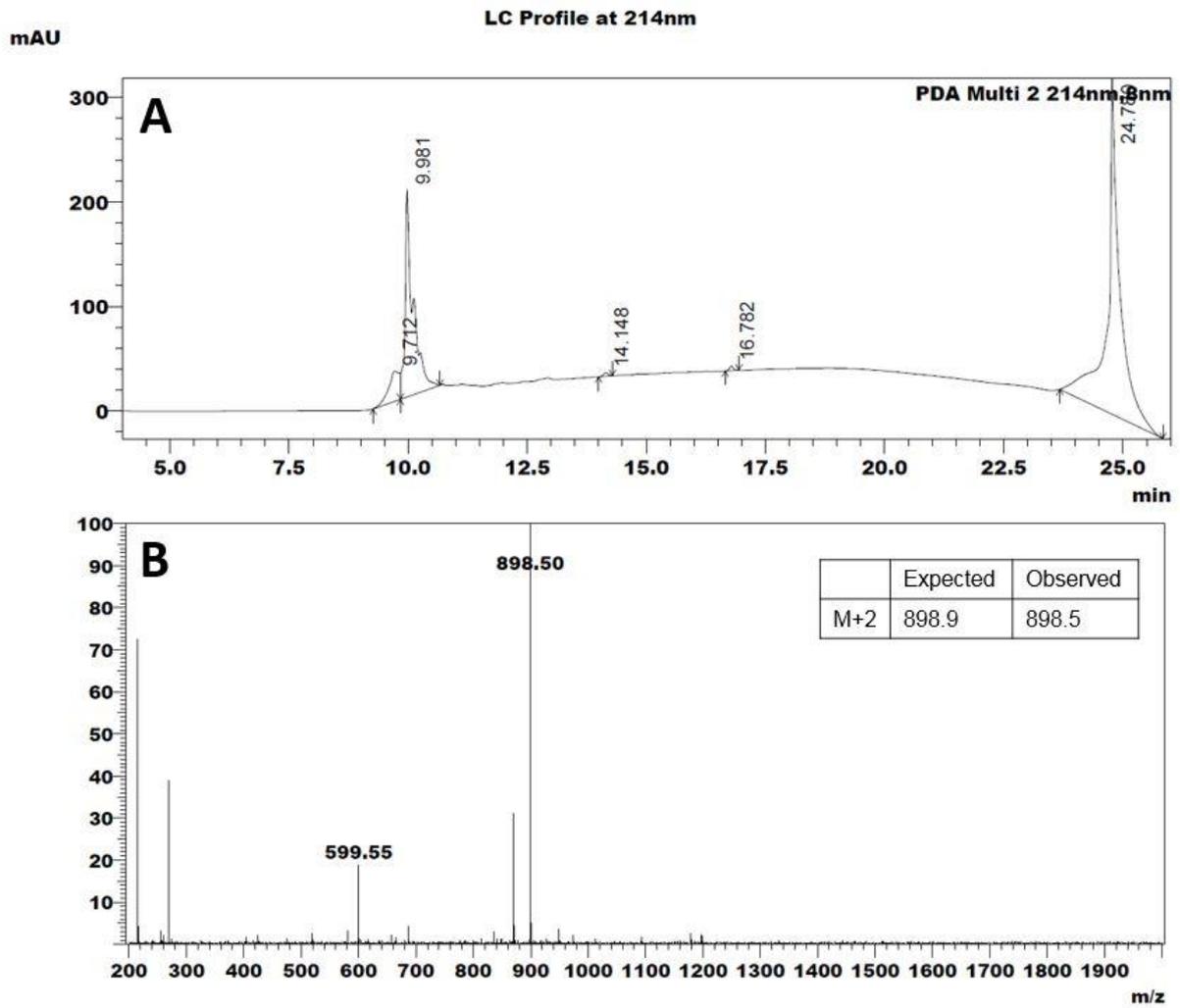


Figure S3. (A) Liquid chromatography trace of purified D1 peptide and (B) mass-spectroscopy profile

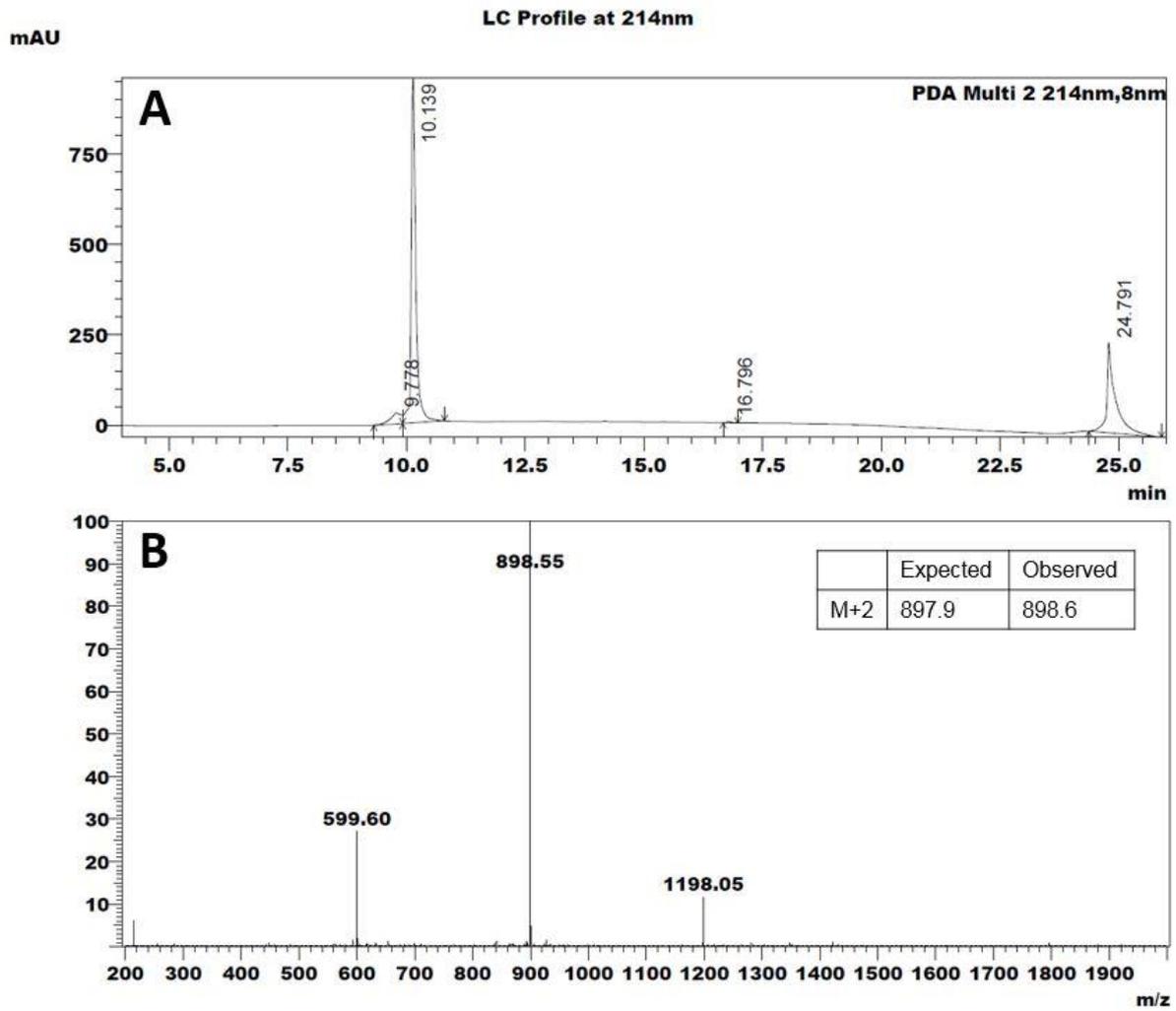


Figure S4. (A) Liquid chromatography trace of purified L1 peptide and (B) mass-spectroscopy profile

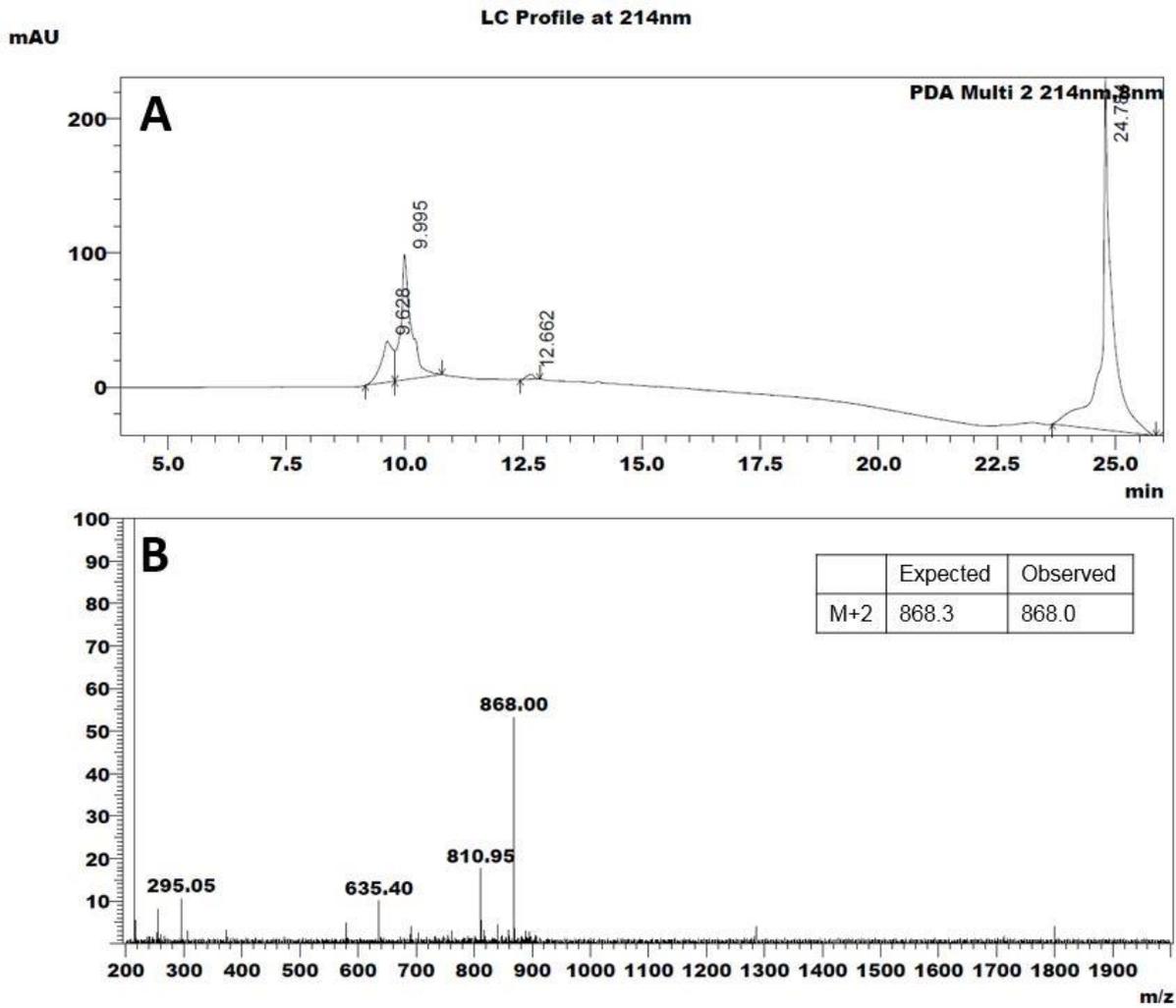


Figure S5. (A) Liquid chromatography trace of purified C1 peptide and (B) mass-spectrometry profile

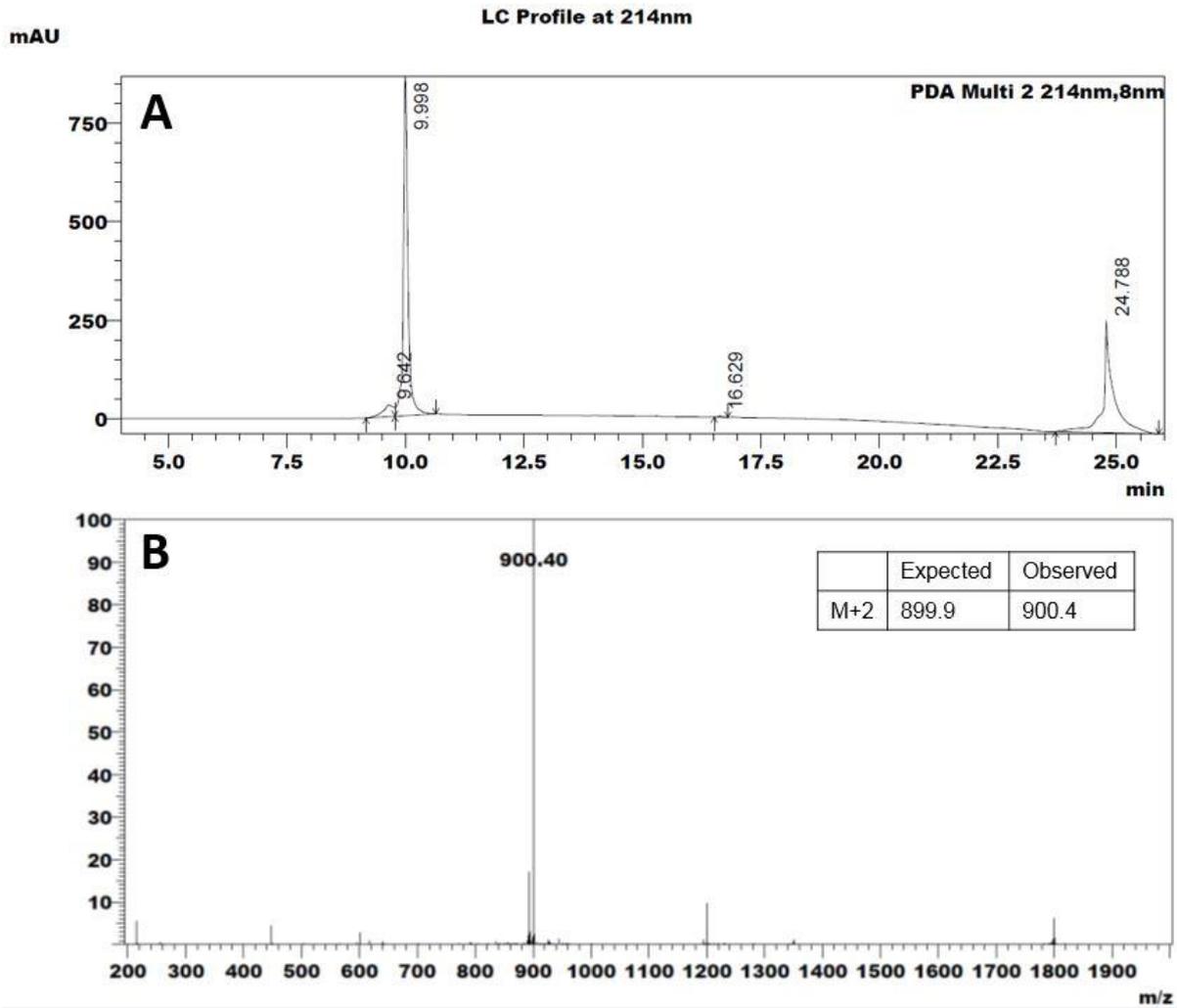


Figure S6. (A) Liquid chromatography trace of purified A1 peptide and (B) mass-spectrometry profile

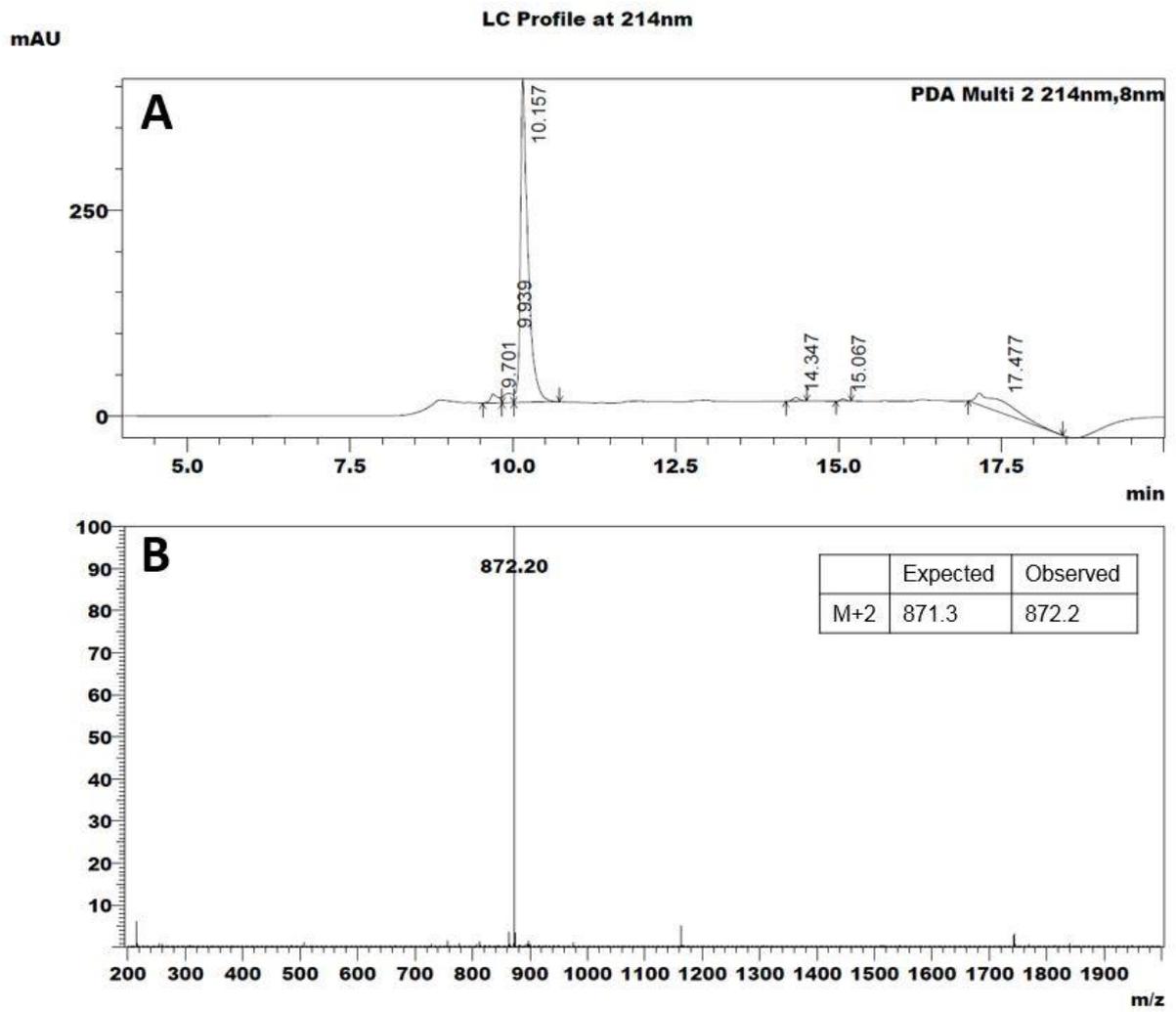


Figure S7. (A) Liquid chromatography trace of purified K209A peptide and (B) mass-spectroscopy profile

# Chapter 6

Conclusions and future directions

The burden of malaria is still significant today despite steady progress over the last decades. In order to achieve the ultimate goal of global eradication, a robust and effective malaria vaccine is an absolute requirement. The work described in this thesis is relevant to this goal.

A major focus of my project was the application of structural biology to vaccine design, in an emerging field commonly known as structural vaccinology or structure-based vaccine design. A large portion of the work builds on previous work in which a panel of monoclonal antibodies was generated in mice against recombinant MSP2 (44). These antibodies were able to recognise native MSP2 on the parasite surface to varying degrees; here, the structural determinants of antibody recognition were explored and that structural information was utilised to inform the design of better MSP2-based vaccine candidates. The organisation of this thesis reflects that of the protein of interest, in that it can be separated into the investigation of three distinct domains, a central variable region, a conserved N-terminal region and a conserved C-terminal region (19, 78).

The first portion of this thesis described the design and expression of antibody fragments corresponding to mAbs that recognise epitopes in the central variable region. The four antibodies specific to this region of MSP2 (2F2, 9D11, 11E1 and 8G10) were strongly reactive to native parasite MSP2 by IFA, suggesting that their cognate epitopes would be accessible to the host immune system. Moreover, antibodies from naturally-acquired and vaccine-induced immunity that target the variable region have been shown to be protective and effective in reducing parasite densities respectively (30, 79, 80). cDNA libraries were generated from hybridoma cell lines expressing variable region mAbs 2F2, 9D11 and 11E1. From these libraries, mouse variable heavy ( $V_H$ ) and light ( $V_{LL}$  and  $V_{LK}$ ) chains were amplified and sequenced. The CDRs of each mAb were cloned onto the previously sequenced, codon-optimised framework from 9H4 Fv to create new anti-MSP2 Fv and scFv antibody fragments that were able to recognise epitopes in the variable region of MSP2. While these antibody fragments were expressed at suitable yields, none was able to bind to recombinant MSP2, and therefore could not be affinity purified using MSP2. This suggested that the antibody fragments had been misfolded in the periplasm or denatured during the process of purification. Further work will be required to correctly fold these variable region specific antibody fragments; this may include the use of different methods of refolding and optimisation of the purification process (55, 56, 99). Alternatively, different expression hosts such as *Pichia* or mammalian cells may improve the yield of functional antibody fragments (100, 101). Expression of the larger antigen-binding fragment (Fab) is also a possibility (124–126).

To complement the work with antibody fragments specific for the variable region and to better understand possible mechanisms of immune evasion employed by the parasite, polymorphisms in the central variable region were analysed in detail. A total of 420 sequences of 3D7-MSP2 and 203 of FC27-MSP2 was aligned and translated. Polymorphisms in the immunodominant epitopes recognised by 3D7-specific mAbs (2F2, 9D11 and 11E1) and FC27-specific antibody 8G10 were evaluated individually for mutations, deletions, amino acid frequency and the frequency of each mutation within the population. These results revealed high sequence conservation in both 2F2 and 9D11 epitopes across the 3D7 allelic family, with 8G10 showing moderate conservation within the FC27 allele. Owing to their location within dimorphic regions of MSP2 this was not unexpected. The 11E1 epitope had a high degree of polymorphisms, with some corresponding to drastic changes in amino acid properties, such as the reversal of side-chain charge and increased flexibility. In terms of mutation frequency, the reference sequence present in our form of 3D7-MSP2, was present in only 3.3% of the sampled sequences. The most frequent sequence within this sample, corresponding to the Q118E mutation in the 3D7-MSP2 sequence, was seen in 26.4% of sequences. There is strong evidence that these polymorphisms, driven by immune pressure, serve as a mechanism of immune evasion (11, 43, 102). Future direction of this work may include synthesis of the most frequent 11E1 epitope sequences to determine their effect on binding. Understanding which residues are important for 11E1 binding may inform the design of vaccine candidates able to elicit a broad protection against all MSP2 variants. Conversely, excluding highly polymorphic epitopes such as 11E1 altogether, in favour of conserved epitopes may achieve the same goal, and was the focus of the remainder of the work presented in this thesis.

The conserved N-terminal region of MSP2 is known to adopt an  $\alpha$ -helical conformation when interacting with lipid (35). 6D8 mAb is able to recognise an epitope within this region, but is unable to bind to native MSP2 on the parasite surface. To better understand the structural determinants of 6D8 binding, the crystal structure of 6D8 Fv in complex with its minimal binding epitope (NAYNMSIRR) were solved. This structure revealed that when bound to 6D8, the epitope did adopt a helical structure, although this was disrupted at Ser19, explaining at least in part why the 6D8-bound conformation was incompatible with the lipid-bound  $\alpha$ -helical configuration, and in turn, why 6D8 was unable to recognise parasite MSP2. Although the antibody-bound structure of the 6D8 epitope does not directly contribute to vaccine development, it may be possible to induce antibodies against the N-terminal region using the lipid-bound structure. The amphipathic nature of the helix suggests that only one face may interact with the merozoite surface, leaving the other side accessible to antibody binding. With

peptide synthesis, stabilised helical analogues of the region could be created by installing paired lactam bridge residues across the sequence. Alternatively, flanking the wild-type sequence that is already known to have high helical propensity, with constrained helical turns may be enough to drive the rest of the sequence into a helical conformation (127, 128). These peptides could be used to affinity-purify N-terminal region specific antibodies from polyclonal sera that could recognise helical epitopes. If present, these antibodies may uncover new conserved epitopes that may contribute to the design of novel MSP2-based vaccines. Even the constrained peptides themselves may serve as good immunogens.

The third domain of MSP2, the conserved C-terminal region, is recognised by two key mAbs, 4D11 and 9H4. Although they recognise overlapping epitopes, these antibodies exhibited distinct antigenic properties (44). Whilst 4D11 mAb was able to recognise native parasite MSP2 by IFA and western blot, 9H4 mAb was not. To determine key residues for antibody binding, an alanine scan was performed along the full 16-residue epitope, which showed that Lys216, Glu217, Asn218 and Gly220 were essential for 4D11 binding, whilst 9H4 binding was dependent on residues closer to the N-terminus, Lys209, Thr212 and Asp213. Intriguingly, removal of residues Gly214 to Glu217 significantly improved binding of 9H4. Crystallisation attempts involving 9H4 antibody fragments in complex with either wildtype or mutated epitopes were unfortunately unsuccessful. However, new 9H4 antibody fragment constructs have shown promise and require further optimisation.

My experience with 4D11 was more favourable, and a crystal structure of 4D11 Fv in complex with its minimal binding epitope (NKENCGAA) was determined at high resolution (2.2 Å). The structure revealed that this epitope adopted a  $\beta$ -bend ribbon conformation when bound to 4D11, characterised by two overlapping  $\beta$ -turns. Furthermore, key interactions observed between 4D11 and the epitope supported the alanine scan results, with additional interactions detected between Ala221 and the V<sub>L</sub> CDRs. Although the peptide was purified as a monomer prior to crystallisation, the bound state of the peptide was found to be homodimeric, mediated by a disulphide bond between the free cysteines of each monomer in the sequence. Although both faces of the peptide were able to bind 4D11 Fv, there is no evidence that native MSP2 exists as a covalent dimer. In order to confirm the relevance of the dimeric bound conformation to native MSP2, MD simulations of the dimer and extended disulphide-bridged 16-residue epitope were performed. These simulations showed that both peptides, when bound to 4D11 Fv, exhibited identical conformational behaviour and stability. Furthermore, SPR binding assays showed that both the monomer and dimer had similar binding affinity to 4D11, supporting the relevance of the dimeric conformation.

Multiple intramolecular hydrogen bonds were found to stabilise the 4D11-bound dimer structure, presenting it as a promising template for application of structure-based vaccine design. Moreover, the close proximity of the N- and C-termini of both peptides suggested the possibility of linking the peptides together, further stabilising the epitope. A series of peptide immunogens was then designed with one or both termini linked to create constrained linear and cyclic analogues of the dimeric peptide. MD simulations were performed to determine ideal linker length and whether the additional constraints would impede the peptide from adopting a conformation that can bind to 4D11. Linker lengths of 1-3 Gly residues were found to be tolerated and peptides were synthesised by standard Fmoc peptide synthesis. Another two peptides were also designed to evaluate the effects of removing the 9H4 epitope. These corresponded to the wildtype 16-residue peptide MSP2<sub>207-222</sub> and an identical peptide with the K209A mutation. SPR competition binding assay showed that all peptides were able to bind 4D11, with the single Gly linker length exhibiting the tightest binding. The dimeric peptide (D1), linear peptide (L1), backbone-cyclised peptide (C1), MSP2<sub>207-222</sub> (A1) and the mutated MSP2<sub>207-222</sub> (K209A) were conjugated to KLH carrier protein and used to immunise mice alongside full-length MSP2 as a control.

Analysis of terminal bleed sera showed that removal of the 9H4 epitope via the K209A mutation had a drastic effect on antibody specificity. The wildtype A1 sequence predominantly elicited antibodies targeting the 9H4 epitope, suggesting its immunodominance over the 4D11 epitope. However, when the 9H4 epitope was removed, the response was biased towards the more desirable 4D11 epitope. This simple mutation could be easily introduced into recombinant forms of MSP2 to induce a more efficient immune response towards the conserved C-terminal region.

All dimeric peptide-KLH conjugates were able to induce a response specific to the 4D11 epitope. When compared to sera from mice immunised with full-length 3D7 and FC27 MSP2, the response was more focused on the 4D11 epitope and more consistent across mice in the group. Future directions of this work involve firstly determining, by using a combination of merozoite ELISA, IFA and western blot, whether the antibodies induced can recognise parasite MSP2. Promising peptide immunogens will then be used to immunise rabbits to generate suitable volumes of sera for functional assays such as ADCI (26), complement fixation (33) and opsonic phagocytosis assays(116) to determine their mechanism of protection.

My work on the intrinsically disordered vaccine candidate MSP2 underpins ongoing efforts to develop next-generation, broadly-protective MSP2-based immunogens. The

application of structural vaccinology to disordered antigens shows significant promise, and may be expected to contribute to the development of vaccines targeting other disordered antigens from a range of pathogens. The results presented in this thesis also enhance our understanding of the effects of disorder on antibody recognition.

# References

1. WHO (2017) *World Malaria Report*.
2. Murray CJL, et al. (2012) Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet* 379(9814):413–31.
3. Thera MA, Plowe C (2012) Vaccines for malaria: how close are we? *Annu Rev Med* 63(2012):345–57.
4. Hemingway J, et al. (2016) Tools and Strategies for Malaria Control and Elimination: What Do We Need to Achieve a Grand Convergence in Malaria? *PLOS Biol* 14(3):e1002380.
5. Good MF, Doolan DL (2010) Malaria vaccine design: immunological considerations. *Immunity* 33(4):555–66.
6. Richards JS, Beeson JG (2009) The future for blood-stage vaccines against malaria. *Immunol Cell Biol* 87(5):377–90.
7. Gardner MJ, et al. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419(6906):498–511.
8. Miller LH, et al. (2002) The pathogenic basis of malaria. *Nature* 415(6872):673–679.
9. Idro R, et al. (2010) Cerebral malaria: mechanisms of brain injury and strategies for improved neurocognitive outcome. *Pediatr Res* 68(4):267–74.
10. Von Seidlein L, Bejon P (2013) Malaria vaccines: past, present and future. *Arch Dis Child* 98(12):981–5.
11. Ferreira M, et al. (2004) Antigenic Diversity and Immune Evasion by Malaria Parasites. *Clinical Diagnostic Lab Immunol* 11(6):987–995.
12. Persson KEM, et al. (2013) Erythrocyte-binding antigens of *Plasmodium falciparum* are targets of human inhibitory antibodies and function to evade naturally acquired immunity. *J Immunol* 191(2):785–94.
13. Greenwood BM (2015) Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa : final results of a phase 3, individually randomised, controlled trial. *Lancet* 386(9988):31–45.
14. Aide P, et al. (2011) Four year immunogenicity of the RTS,S/AS02A malaria vaccine in Mozambican children during a phase IIb trial. *Vaccine* 29(35):6059–6067.
15. Michalakis Y, Renaud F (2009) Malaria: Evolution in vector control. *Nature* 462(7271):298–300.
16. Gilson PR, et al. (2006) Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics* 5(7):1286–99.
17. MacRaild CA, et al. (2015) Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2. *PLoS One* 10(3):1–19.
18. Sanders PR, et al. (2006) A set of glycosylphosphatidyl inositol-anchored membrane proteins of *Plasmodium falciparum* is refractory to genetic deletion. *Infect Immun* 74(7):4330–8.
19. Smythe JA, et al. (1991) Structural diversity in the *Plasmodium falciparum* merozoite surface antigen 2. *Proc Natl Acad Sci U S A* 88(5):1751–5.
20. Bütikofer P, et al. (2001) GPI-anchored proteins: now you see 'em, now you don't. *FASEB J* 15(2):545–8.
21. Boyle MJ, et al. (2014) Sequential Processing of Merozoite Surface Proteins during and after Erythrocyte Invasion by *Plasmodium falciparum*. *Infect Immun* 82(3):924–36.
22. Genton B, et al. (2002) A recombinant blood-stage malaria vaccine reduces *Plasmodium falciparum* density and exerts selective pressure on parasite populations in a phase 1–2b trial in Papua New Guinea. *J Infect Dis* 185(6):820–7.
23. Reddy SB, et al. (2012) High affinity antibodies to *Plasmodium falciparum* merozoite

- antigens are associated with protection from malaria. *PLoS One* 7(2):e32242.
24. Cortés A, et al. (2004) *Plasmodium falciparum*: distribution of msp2 genotypes among symptomatic and asymptomatic individuals from the Wosera region of Papua New Guinea. *Exp Parasitol* 106(1–2):22–9.
  25. Adda CG, et al. (2009) *Plasmodium falciparum* merozoite surface protein 2 is unstructured and forms amyloid-like fibrils. *Mol Biochem Parasitol* 166(2):159–71.
  26. McCarthy JS, et al. (2011) A phase 1 trial of MSP2-C1, a blood-stage malaria vaccine containing 2 isoforms of MSP2 formulated with Montanide® ISA 720. *PLoS One* 6(9):e24413.
  27. Anders RF, et al. (2010) Recombinant protein vaccines against the asexual blood stages of *Plasmodium falciparum*. *Hum Vaccin* 6(1):39–53.
  28. Genton B, et al. (2000) Safety and immunogenicity of a three-component blood-stage malaria vaccine in adults living in an endemic area of Papua New Guinea. *Vaccine* 18(23):2504–11.
  29. Polley SD, et al. (2006) High levels of serum antibodies to merozoite surface protein 2 of *Plasmodium falciparum* are associated with reduced risk of clinical malaria in coastal Kenya. *Vaccine* 24(19):4233–46.
  30. Genton B, et al. (2003) Safety and immunogenicity of a three-component blood-stage malaria vaccine (MSP1, MSP2, RESA) against *Plasmodium falciparum* in Papua New Guinean children. *Vaccine* 22(1):30–41.
  31. Duncan CJA, et al. (2012) Can growth inhibition assays (GIA) predict blood-stage malaria vaccine efficacy? *Hum Vaccin Immunother* 8(6):706–14.
  32. Bergmann-Leitner ES, et al. (2006) Critical evaluation of different methods for measuring the functional activity of antibodies against malaria blood stage antigens. *Am J Trop Med Hyg* 75(3):437–42.
  33. Boyle MJ, et al. (2015) Human Antibodies Fix Complement to Inhibit *Plasmodium falciparum* Invasion of Erythrocytes and Are Associated with Protection against Malaria. *Immunity* 42(3):580–590.
  34. Feng Z-P, et al. (2006) Abundance of intrinsically unstructured proteins in *P. falciparum* and other apicomplexan parasite proteomes. *Mol Biochem Parasitol* 150(2):256–67.
  35. Zhang X, et al. (2008) Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2. *J Mol Biol* 379(1):105–21.
  36. Uversky VN, Dunker AK (2010) Understanding protein non-folding. *Biochim Biophys Acta* 1804(6):1231–64.
  37. Uversky V (2002) Natively unfolded proteins: a point where biology waits for physics. *Protein Sci*:739–756.
  38. Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* 27(10):527–33.
  39. Tompa P (2012) Intrinsically disordered proteins : a 10-year recap. *Trends Biochem Sci* 37(12):509–516.
  40. Chouard T (2011) Structural biology: Breaking the protein rules. *Nature* 471(7337):151–3.
  41. Macrauld CA, et al. (2016) Antibody Recognition of Disordered Antigens. *Structure* 24(1):148–157.
  42. Drew DR, et al. (2012) Defining the antigenic diversity of *Plasmodium falciparum* apical membrane antigen 1 and the requirements for a multi-allele vaccine against malaria. *PLoS One* 7(12):e51023.
  43. Ferreira MU, et al. (2007) Origins and evolution of antigenic diversity in malaria

- parasites. *Curr Mol Med* 7(6):588–602.
44. Adda CG, et al. (2012) Antigenic characterization of an intrinsically unstructured protein, *Plasmodium falciparum* merozoite surface protein 2. *Infect Immun* 80(12):4177–85.
  45. Das SC, et al. (2017) Lipid interactions modulate the structural and antigenic properties of the C-terminal domain of the malaria antigen merozoite surface protein 2. *FEBS J* 284(16):2649–2662.
  46. Yang X, et al. (2010) Identification of key residues involved in fibril formation by the conserved N-terminal region of *Plasmodium falciparum* merozoite surface protein 2 (MSP2). *Biochimie* 92(10):1287–1295.
  47. Yang X, et al. (2007) A partially structured region of a largely unstructured protein, *Plasmodium falciparum* merozoite surface protein 2 (MSP2), forms amyloid-like fibrils. *J Pept Sci* 13(12):839–848.
  48. Zhang X, et al. (2012) Role of the helical structure of the N-terminal region of *Plasmodium falciparum* merozoite surface protein 2 in fibril formation and membrane interaction. *Biochemistry* 51(7):1380–7.
  49. Herrera R, et al. (2015) Reversible conformational change in the *Plasmodium falciparum* circumsporozoite protein masks its adhesion domains. *Infect Immun* 83(10):3771–3780.
  50. Beckman RA, et al. (2007) Antibody constructs in cancer therapy: Protein engineering strategies to improve exposure in solid tumors. *Cancer* 109(2):170–179.
  51. Holliger P, Hudson PJ (2005) Engineered antibody fragments and the rise of single domains. *Nat Biotechnol* 23(9):1126–36.
  52. Kumai T, et al. (2017) Cancer immunotherapy: moving forward with peptide T cell vaccines. *Curr Opin Immunol* 47:57–63.
  53. Rouet R, et al. (2014) Stability engineering of the human antibody repertoire. *FEBS Lett* 588(2):269–77.
  54. Arakawa T, Ejima D (2014) Refolding Technologies for Antibody Fragments. *Antibodies* 3(2):232–241.
  55. Jager M, Pluckthun A (1997) The rate-limiting steps for the folding of an antibody scFv fragment. *FEBS Lett* 418(1–2):106–110.
  56. Jung S, Pluckthun A (1997) Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting. *Protein Eng* 10(8):959–966.
  57. Bachmann MF, et al. (1997) The role of antibody concentration and avidity in antiviral protection. *Science* 276(5321):2024–2027.
  58. Rouet R, et al. (2012) Expression of high-affinity human antibody fragments in bacteria. *Nat Protoc* 7(2):364–73.
  59. Stern AM, Markel H (2005) The history of vaccines and immunization: Familiar patterns, new challenges - If we could match the enormous scientific strides of the twentieth century with the political and economic investments of the nineteenth, the world's citizens might be much healed. *Health Aff* 24(3):611–621.
  60. Skwarczynski M, Toth I (2016) Peptide-based synthetic vaccines. *Chem Sci* 7(2):842–854.
  61. Anders RF (2011) The case for a subunit vaccine against malaria. *Trends Parasitol* 27(8):330–334.
  62. Robinson JA (2013) Max Bergmann lecture protein epitope mimetics in the age of structural vaccinology. *J Pept Sci* 19(3):127–40.
  63. Dormitzer PR, et al. (2008) Structure-based antigen design: a strategy for next generation vaccines. *Trends Biotechnol* 26(12):659–67.
  64. Dormitzer PR, et al. (2012) Structural vaccinology starts to deliver. *Nat Rev Microbiol*

- 10(12):807–13.
65. Grimm SK, Ackerman ME (2013) Vaccine design : emerging concepts and renewed optimism. *Curr Opin Biotechnol*:1–11.
  66. Scarselli M, et al. (2011) Rational design of a meningococcal antigen inducing broad protective immunity. *Sci Transl Med* 3(91):91ra62.
  67. Swanson KA, et al. (2011) Structural basis for immunization with postfusion respiratory syncytial virus fusion F glycoprotein (RSV F) to elicit high neutralizing antibody titers. *Proc Natl Acad Sci U S A* 108(23):9619–24.
  68. McLellan JS, et al. (2013) Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science* 342(6158):592–8.
  69. Nuccitelli A, et al. (2011) Structure-based approach to rationally design a chimeric protein for an effective vaccine against Group B Streptococcus infections. *Proc Natl Acad Sci U S A* 108(25):10278–10283.
  70. Jardine J, et al. (2013) Rational HIV immunogen design to target specific germline B cell receptors. *Science* 340(6133):711–6.
  71. Ofek G, et al. (2010) Elicitation of structure-specific antibodies by epitope scaffolds. *Proc Natl Acad Sci* 107(42):17880–17887.
  72. Azoitei ML, et al. (2012) Computational design of high-affinity epitope scaffolds by backbone grafting of a linear epitope. *J Mol Biol* 415(1):175–192.
  73. Murphy E, et al. (2009) Sequence Diversity of the Factor H Binding Protein Vaccine Candidate in Epidemiologically Relevant Strains of Serogroup B *Neisseria meningitidis*. *J Infect Dis* 200(3):379–389.
  74. Bambini S, et al. (2009) Distribution and genetic variability of three vaccine components in a panel of strains representative of the diversity of serogroup B meningococcus. *Vaccine* 27(21):2794–2803.
  75. Masignani V, et al. (2003) Vaccination against *Neisseria meningitidis* Using Three Variants of the Lipoprotein GNA1870. *J Exp Med* 197(6):789–799.
  76. McLellan JS, et al. (2011) Design and characterization of epitope-scaffold immunogens that present the motavizumab epitope from respiratory syncytial virus. *J Mol Biol* 409(5):853–66.
  77. Magro M, et al. (2012) Neutralizing antibodies against the preactive form of respiratory syncytial virus fusion protein offer unique possibilities for clinical intervention. *Proc Natl Acad Sci* 109(8):3089–3094.
  78. Fenton B, et al. (1991) Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite surface antigen (MSA-2) of the malaria parasite *Plasmodium falciparum*. *Mol Cell Biol* 11(2):963–971.
  79. Franks S, et al. (2003) Genetic Diversity and Antigenic Polymorphism in *Plasmodium falciparum* : Extensive Serological Cross-Reactivity between Allelic Variants of Merozoite Surface Protein 2. *Infect Immun* 71(6):3485–3495.
  80. Taylor RR, et al. (1995) Human antibody response to *Plasmodium falciparum* merozoite surface protein 2 is serogroup specific and predominantly of the immunoglobulin G3 subclass . *Infect Immun* 63(11):4382–4388.
  81. Krishnarjuna B, et al. (2016) Strain-transcending immune response generated by chimeras of the malaria vaccine candidate merozoite surface protein 2. *Sci Rep* 6:20613.
  82. Balam S, et al. (2014) *Plasmodium falciparum* merozoite surface protein 2: epitope mapping and fine specificity of human antibody response against non-polymorphic domains. *Malar J* 13:510.
  83. Altschul SF, et al. (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–10.
  84. Sievers F, Higgins DG (2014) Clustal Omega, accurate alignment of very large

- numbers of sequences. *Methods Mol Biol* 1079:105–16.
85. Kearse M, et al. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12):1647–9.
  86. Pronk S, et al. (2013) GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 29(7):845–854.
  87. Brochet X, et al. (2008) IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res* 36(Web Server issue):503–508.
  88. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41(1):207–234.
  89. The CCP4 suite: programs for protein crystallography. (1994) *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760–3.
  90. Brünger AT (1993) Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallogr D Biol Crystallogr* 49(Pt 1):24–36.
  91. Adams PD, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–21.
  92. Stein N (2008) CHAINSAW : a program for mutating pdb files used as templates in molecular replacement. *J Appl Crystallogr* 41(3):641–643.
  93. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658–674.
  94. Zhou H, et al. (1994) Optimization of primer sequences for mouse scFv repertoire display library construction. *Nucleic Acids Res* 22(5):888–9.
  95. Wang Z, et al. (2000) Universal PCR amplification of mouse immunoglobulin gene variable regions: the design of degenerate primers and an assessment of the effect of DNA polymerase 3' to 5' exonuclease activity. *J Immunol Methods* 233(1–2):167–77.
  96. Fields C, et al. (2013) Creation of recombinant antigen-binding molecules derived from hybridomas secreting specific antibodies. *Nat Protoc* 8(6):1125–48.
  97. Shimizu A, et al. (1991) Trans-splicing as a possible molecular mechanism for the multiple isotype expression of the immunoglobulin gene. *J Exp Med* 173(6):1385–93.
  98. Welschof M, et al. (1995) Amino acid sequence based PCR primers for amplification of rearranged human heavy and light chain immunoglobulin variable region genes. *J Immunol Methods* 179(2):203–214.
  99. Miller BR, et al. (2010) Stability engineering of scFvs for the development of bispecific and multivalent antibodies. *Protein Eng Des Sel* 23(7):549–57.
  100. Cupit PM, et al. (1999) Cloning and expression of single chain antibody fragments in *Escherichia coli* and *Pichia pastoris*. *Lett Appl Microbiol* 29(5):273–277.
  101. Freyre FM, et al. (2000) Very high expression of an anti-carcinoembryonic antigen single chain Fv antibody fragment in the yeast *Pichia pastoris*. *J Biotechnol* 76(2–3):157–163.
  102. Hughes AL (1991) Circumsporozoite protein genes of malaria parasites (*Plasmodium spp.*): Evidence for positive selection on immunogenic regions. *Genetics* 127(2):345–353.
  103. Patarroyo ME, Patarroyo MA (2008) Emerging rules for subunit-based, multiantigenic, multistage chemically synthesized vaccines. *Acc Chem Res* 41(3):377–86.
  104. Seow J, et al. (2017) Structure and characterisation of a key epitope in the conserved C-terminal domain of the malaria vaccine candidate MSP2. *J Mol Biol* 429(6):836–846.
  105. Rappuoli R, et al. (2016) Reverse vaccinology 2.0: Human immunology instructs

- vaccine antigen design. *J Exp Med* 213(4):469–481.
106. Kaufmann SHE, et al. (2014) Challenges and responses in human vaccine development. *Curr Opin Immunol* 28:18–26.
  107. Cui JJ, et al. (2011) Structure based drug design of crizotinib (PF-02341066), a potent and selective dual inhibitor of mesenchymal-epithelial transition factor (c-MET) kinase and anaplastic lymphoma kinase (ALK). *J Med Chem* 54(18):6342–6363.
  108. Lounnas V, et al. (2013) Current Progress in Structure-Based Rational Drug Design Marks a New Mindset in Drug Discovery. *Comput Struct Biotechnol J* 5(6):e201302011.
  109. Burton DR (2010) Scaffolding to build a rational vaccine design strategy. *Proc Natl Acad Sci U S A* 107(42):17859–60.
  110. Triller G, et al. (2017) Natural Parasite Exposure Induces Protective Human Anti-Malarial Antibodies. *Immunity* 47(6):1197–1209.e10.
  111. Meola A, et al. (2015) Structural Flexibility of a Conserved Antigenic Region in Hepatitis C Virus Glycoprotein E2 Recognized by Broadly Neutralizing Antibodies. *J Virol* 89(4):2170–2181.
  112. Oyen D, et al. (2017) Structural basis for antibody recognition of the NANP repeats in *Plasmodium falciparum* circumsporozoite protein. *Proc Natl Acad Sci* 114(48):E10438–E10445.
  113. Henchey LK, et al. (2008) Contemporary strategies for the stabilization of peptides in the  $\alpha$ -helical conformation. *Curr Opin Chem Biol* 12(6):692–697.
  114. Harrison RS, et al. (2010) Downsizing human, bacterial, and viral proteins to short water-stable alpha helices that maintain biological potency. *Proc Natl Acad Sci U S A* 107(26):11686–11691.
  115. Taylor JW (2002) The synthesis and study of side-chain lactam-bridged peptides. *Biopolym - Pept Sci Sect* 66(1):49–75.
  116. Osier FH, et al. (2014) Opsonic phagocytosis of *Plasmodium falciparum* merozoites: mechanism in human immunity and a correlate of protection against malaria. *BMC Med* 12(1):108.
  117. Abraham MJ, et al. (2015) Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2:19–25.
  118. Schmid N, et al. (2011) Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur Biophys J* 40(7):843–856.
  119. Bussi G, et al. (2007) Canonical sampling through velocity rescaling. *J Chem Phys* 126(1). doi:10.1063/1.2408420.
  120. Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys* 52(12):7182–7190.
  121. Essmann U, et al. (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103(1995):8577–8593.
  122. Postma TM, Albericio F (2013) N-Chlorosuccinimide, an efficient reagent for on-resin disulfide formation in solid-phase peptide synthesis. *Org Lett* 15(3):616–619.
  123. Postma TM, Albericio F (2014) Disulfide formation strategies in peptide synthesis. *European J Org Chem* 2014(17):3519–3530.
  124. Roguska MA, et al. (1994) Humanization of murine monoclonal antibodies through variable domain resurfacing. *Proc Natl Acad Sci* 91(3):969–973.
  125. Orum H, et al. (1993) Efficient method for constructing comprehensive Fab antibody libraries displayed on phage murine. *Nucleic Acids Res* 21(19):4491–4498.
  126. Frenzel A, et al. (2013) Expression of recombinant antibodies. *Front Immunol* 4(JUL):1–20.
  127. Forood B, et al. (1993) Stabilization of alpha-helical structures in short peptides via

- end capping. *Proc Natl Acad Sci U S A* 90(3):838–842.
128. Aurora R, Rose GD (1997) Helix capping. *Protein Sci* 7:1648–1652.

# Appendix I

Disordered epitopes as peptide vaccines

## REVIEW

## 1 Disordered epitopes as peptide vaccines

AQ5 2 Christopher A. MacRaild  | Jeffrey Seow | Sreedam C. Das | Raymond S. Norton4 Medicinal Chemistry, Monash Institute of  
5 Pharmaceutical Sciences, Monash University,  
6 399 Royal Parade, Parkville 3052, Australia

## 7 Correspondence

8 Christopher A. MacRaild, Medicinal  
9 Chemistry, Monash Institute of  
10 Pharmaceutical Sciences, Monash  
11 University, 381 Royal Parade, Parkville  
12 3052, Australia

13 Email: chris.macraild@monash.edu

14 And

15 Raymond S. Norton, Medicinal Chemistry,  
16 Monash Institute of Pharmaceutical  
17 Sciences, Monash University, 381 Royal  
18 Parade, Parkville 3052, Australia.

19 Email: ray.norton@monash.edu

## 20 Funding information

21 Australian National Health and Medical  
22 Research Council, Grants Numbers:  
23 1042520, 1125788; NHMRC

## Abstract

The development of clinically useful peptide-based vaccines remains a long-standing goal. This review highlights that intrinsically disordered protein antigens, which lack an ordered three-dimensional structure, represent excellent starting points for the development of peptide vaccines. Disordered proteins represent an important class of antigen in a wide range of human pathogens, and, contrary to widespread belief, they are frequently targets of protective antibody responses. Importantly, disordered epitopes appear invariably to be linear epitopes, rendering them ideally suited to incorporation into a peptide vaccine. Nonetheless, the conformational properties of disordered antigens, and hence their recognition by antibodies, frequently depend on the interactions they make and the context in which they are presented to the immune system. These effects must be considered in the design of an effective vaccine. Here we discuss these issues and propose design principles that may facilitate the development of peptide vaccines targeting disordered antigens.

## KEYWORDS

design, intrinsically disordered antigen, malaria, membrane interactions, peptide epitope, structure

## 22 1 | INTRODUCTION

23 Vaccines are an indispensable tool in the fight against disease and have  
24 had a significant impact on public health over several decades.<sup>[1]</sup> Traditionally,  
25 vaccines have relied on the use of live-attenuated or inactivated forms of the  
26 pathogen to induce a protective immune response.  
27 However, in some cases, such as malaria, the pathogen is difficult to  
28 culture in vitro at large scale. Moreover, the use of whole organisms as  
29 vaccines introduces a high antigenic load when often only a small subset  
30 of antigens are driving protection.<sup>[2]</sup> There is also the possibility of  
31 adverse allergenic reactions to certain antigens in these preparations.  
32 These obstacles have led to an increased interest in subunit vaccines,  
33 in which single, or a select few, proteins are used in a vaccine formulation  
34 to induce protective immunity.<sup>[3]</sup>

35 The use of peptides as vaccines takes this rationale further, as even a  
36 single protein antigen may have many epitopes, not all of which contribute  
37 to protective immunity. Peptide vaccines offer a means to formulate  
38 vaccines containing only epitopes that are capable of inducing a positive  
39 and efficient immune response.<sup>[2,4]</sup> The ease of peptide synthesis makes  
40 large scale production feasible whilst also offering a cleaner vaccine preparation  
41 lacking biological contaminations commonly associated with  
42 recombinant expression or whole organism vaccines. The specificity of  
43 peptides also allows for improved customisability, facilitating, for example,

a multiepitope approach to target different strains or stages in the life  
cycle of the pathogen. 44 45

46 For these reasons, there has been long-standing interest in this area  
47 of vaccine design, with over 500 peptide vaccines reaching clinical trials,  
48 targeting a wide range of indications (Table 1). Seven of these vaccines  
49 have reached Phase III trial (Table 2). Much of this recent interest has  
50 focussed on T-cell epitopes, with all of the vaccines to reach Phase III  
51 being T-cell based. For many diseases, however, B-cell responses also  
52 play an important role in immune protection, but peptide-based vaccines  
53 based on B-cell epitopes have proved more challenging for a number of  
54 reasons. Owing to the small size of peptides, they are generally poorly  
55 immunogenic. In addition, T-cell epitopes required for the establishment  
56 of a robust immune response may also be absent in smaller peptides. To  
57 address these issues, carrier proteins or adjuvants are typically included in  
58 the vaccine formulation. Peptides are also prone to enzymatic degradation;  
59 to combat this, the epitopes can be modified by conformational stabilization  
60 or cyclisation, or peptide mimetics that are resistant to proteolysis can be used.<sup>[5]</sup>  
61 Finally, and perhaps most importantly, peptide-based approaches are limited  
62 because the antibody response to many protein antigens is dominated by  
63 conformational epitopes, which are difficult or impossible to capture effectively  
64 in a peptide-based design.

65 In this review, we argue that intrinsically disordered protein antigens,  
66 which lack an ordered three-dimensional structure, represent

TABLE 1 Peptide vaccines in clinical trials

	Number of active or completed clinical trials <sup>a</sup>	Conditions being treated with peptide vaccines
Phase III	7	Cancer immunotherapies, Multiple Sclerosis, Type 1 Diabetes
Phase II	203	Cancer immunotherapies, Myelodysplastic Syndrome, HIV, HBV, HCV, Cytomegalovirus, Myasthenia Gravis, Influenza, Malaria ( <i>falciparum</i> )
Phase I and Early Phase 1	307	Cancer immunotherapies, HIV, HPV, HBV, HCV, Age Related Macular Degeneration, Respiratory Syncytial Virus, Malaria ( <i>vivax</i> ), Malaria ( <i>falciparum</i> ), Hand foot and mouth disease, Influenza, Multiple Sclerosis, Alzheimer's Disease, Listeria, Cat allergy, Ragweed allergy, House dust mite allergy, Grass allergy

<sup>a</sup>Clinical studies were found using "peptide vaccine" as a search term on ClinicalTrials.gov, withdrawn studies were excluded.

67 excellent starting points for the development of peptide vaccines. Our  
68 argument is structured as follows: first, we establish that disordered  
69 proteins represent an important class of antigen in a wide range of  
70 human pathogens. Second, we describe the unique features of these  
71 antigens and their interactions with antibodies, highlighting those prop-  
72 erties that render them particularly suited to the development of pep-  
73 tide vaccines. In particular, we note that disordered epitopes appear  
74 invariably to be linear epitopes, so the challenge of capturing confor-  
75 mational epitopes in a peptide vaccine may not apply. We then outline  
76 some of the opportunities and challenges presented by the peptide-  
77 based approach to vaccine development, as exemplified by efforts to  
78 develop a vaccine based on the disordered malaria antigen MSP2.  
79 Finally, we speculate on future developments, within peptide science  
80 and beyond, that will more fully enable the promise of this approach to  
81 be realized.

## 82 2 | DISORDERED PROTEINS ARE 83 IMPORTANT ANTIGENS

### 84 2.1 | Disordered proteins are abundant in a range of 85 pathogens

86 Intrinsically disordered proteins are found throughout biology,<sup>[6]</sup> and  
87 are particularly enriched in the context of molecular signaling and regu-  
88 latory functions.<sup>[7]</sup> Consistent with this regulatory role, the extent of  
89 disorder within the proteome of individual species tends to increase  
90 with the environmental complexity of the organisms life-cycle.<sup>[8]</sup> Thus,

91 organisms that live in diverse or highly variable environments tend to  
92 have highly disordered proteomes, reflecting the need for sophisticated  
93 regulatory mechanisms to survive such environmental diversity. As a  
94 result, disordered proteins are particularly abundant in a range of path-  
95 ogenic organisms, and in some viruses,<sup>[9-13]</sup> which must survive diverse  
96 and often hostile host environments. In these contexts, disordered pro-  
97 teins play important roles in host-pathogen interactions,<sup>[11]</sup> and accord-  
98 ingly, disordered proteins are often exposed to the host immune  
99 system.<sup>[12]</sup> Even amongst bacteria, which have significantly more  
100 ordered proteomes than eukaryotic species, many proteins involved in  
101 pathogenic effector mechanisms appear to be disordered.<sup>[11]</sup>

102 It is often assumed that disordered antigens impede the develop-  
103 ment of an effective immune response, and hence that their abundance  
104 represents an adaptive mechanism enabling the pathogen to evade the  
105 immune response.<sup>[14-18]</sup> Mechanisms proposed to account for this  
106 immune evasion are diverse and contested, and we critically assess sev-  
107 eral of these mechanisms in section 2 below. Despite the fact that  
108 these hypotheses have received limited direct experimental support,  
109 the underlying assumption that repetitive or disordered antigens will  
110 fail to elicit an effective antibody response continues to enjoy signifi-  
111 cant currency.<sup>[19-21]</sup> However, recent sequencing of the genomes of  
112 several species of parasite has demonstrated that low-complexity and  
113 disordered sequences are unusually abundant in a wide range of genes,  
114 including many that do not appear to be antigenically important.<sup>[22]</sup>  
115 This suggests that the presence of such sequences in these parasites is  
116 not primarily a strategy to evade the host immune response. This has  
117 prompted significant debate surrounding the mechanisms responsible

TABLE 2 Phase III clinical trials currently active or completed

Candidate	Construct	Condition	Clinical Trials Identifier
MDX-1379	Two peptides from gp100 melanocyte protein	Metastatic melanoma	NCT00094653
PR1 leukaemia peptide vaccine	Derived from proteinase 3 and neutrophil elastase	Acute Myeloid leukemia	NCT00454168
Telomerase peptide vaccine GV1001	Derived from reverse transcriptase subunit of telomerase (hTERT)	Pancreatic cancer	NCT00425360
NeuVax	Derived from human leukocyte antigen HER2	Breast cancer	NCT01479244
NeuroVax	Two peptides from T-cell receptor	Multiple sclerosis	NCT02057159
MAGE-A3 and NY-ESO-1 Immunotherapy	Peptides from MAGE-A3 and NY-ESO-1 proteins	Multiple myeloma	NCT00090493
Diaprep277	T-cell epitope of heat shock protein 60	Type 1 Diabetes	NCT01281072

TABLE 3 Selected disordered antigens under development as vaccines

Disordered antigen	Pathogen	Vaccine type	Stage of development	Refs
Neisserial Heparin Binding Antigen	<i>Neisseria meningitidis</i> Serogroup B	Subunit combination	Licensed (Bexsero)	85,86
CSP	<i>Plasmodium falciparum</i>	Virus-like particle	Phase IV (RTS,S)	29,87
MSP2	<i>Plasmodium falciparum</i>	Subunit combination	Phase IIb	36,37
KMP-11	<i>Leishmania amazonensis</i>	Subunit combination	Phase I	88
P27A	<i>Plasmodium falciparum</i>	Peptide	Phase I	83
preS Antigen	Hepatitis B	Virus-like particle	Preclinical	89
HSP90	<i>Candida albicans</i>	Peptide	Preclinical	90
SAPA	<i>Trypanosoma cruzi</i>	Peptide/fusion-protein	Preclinical	91
Nucleocapsid protein	SARS coronavirus Urbani	Protein	Preclinical	92
Protease precursor	<i>Porphyromonas gingivalis</i>	Peptide	Preclinical	93
Protective recombinant antigen	<i>Taenia crassiceps</i>	Peptide/protein	Preclinical	94,95
Glycoprotein G	Herpes simplex virus	Peptide	Preclinical	96
Glycoprotein D	Herpes simplex virus	Peptide/Fusion protein	Preclinical	97,98
Phosphoprotein 150	Cytomegalovirus	Peptide	Preclinical	99
VP1	Human Parvovirus B19	Virus-like particle	Preclinical	100,101
Glycoprotein G	RSV	Virus-like particle	Preclinical	102,103
Protein P	RSV	Peptide	Preclinical	104,105
Gag region encoded protein	Human T-lymphotropic Virus Type-1	Protein subunit/peptide	Preclinical	106,107

118 for generating and maintaining this abundance of low-complexity  
 119 sequence.<sup>[23-25]</sup> In fact, it seems likely that distinct classes of low-  
 120 complexity sequences exist within parasite genomes, suggesting a  
 121 diversity of contributory mechanisms.<sup>[25]</sup>

## 122 2.2 | Disordered antigens are important targets in 123 vaccine development

124 As noted above, the abundance of disordered and low-complexity sequen-  
 125 ces in *Plasmodium* species that are responsible for malaria and in related  
 126 parasites was first identified through analysis of antigenically important  
 127 proteins. It is not surprising, then, that a substantial fraction of malaria  
 128 vaccine candidates contain regions of disorder, and that these regions are  
 129 often important targets of antibodies. The most prominent example is the  
 130 circumsporozoite protein (CSP), the dominant antigen on the surface of  
 131 the sporozoite, the parasite form that is injected into the host by the mos-  
 132 quito and which must replicate within host hepatocytes prior to the estab-  
 133 lishment of the symptomatic blood-stage infection.<sup>[26]</sup> CSP has been  
 134 extensively studied as an antigen for inclusion in a malaria vaccine, and the  
 135 recently-approved vaccine RTS,S includes the C-terminal half of CSP,  
 136 fused to the hepatitis B S-antigen and assembled as a virus-like particle.<sup>[27]</sup>  
 137 The RTS,S antigen includes the thrombospondin repeat domain at the C-  
 138 terminus of CSP, although it is not clear that this domain is correctly folded  
 139 in the vaccine formulation.<sup>[28]</sup> The rest of CSP is predicted to be disor-  
 140 dered, and these predictions have been confirmed experimentally for the  
 141 tetra-peptide repeats of CSP that appear to be immunodominant and are  
 142 targets of protective antibodies induced by RTS,S.<sup>[29-32]</sup>

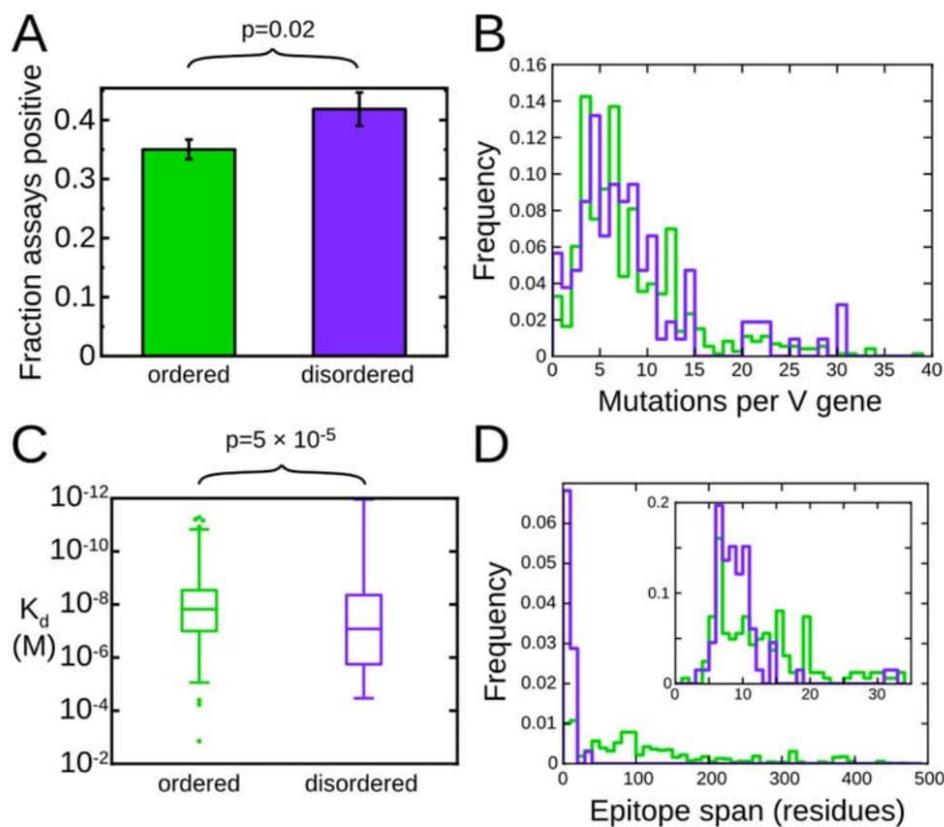
A second important malaria vaccine candidate that is entirely  
 disordered is merozoite surface protein 2 (MSP2). Antibodies target-  
 ing MSP2 arise in the course of malaria infection, and the presence  
 of these antibodies correlates with protection from subsequent  
 infection.<sup>[33-35]</sup> MSP2 also induces robust antibody responses when  
 used as a vaccine, and these responses were shown to mediate pro-  
 tection in a Phase IIb trial in Papua New Guinea.<sup>[36,37]</sup> The confor-  
 mational and antigenic properties of MSP2 have been the subject of  
 extensive study,<sup>[38-41]</sup> providing important insights into the interplay  
 between antigenicity and disorder, to which we will return in section  
 2.2 below.

Many additional antigens of significant interest as vaccine candi-  
 dates against malaria and other diseases are known or predicted to be  
 disordered, and we highlight a number of examples in Table 3. It is  
 apparent, however, that in many cases the disordered nature of these  
 antigens is not widely appreciated, raising the possibility that many  
 important antigens are yet to be recognized as disordered.

## 3 | ANTIBODY INTERACTIONS OF DISORDERED ANTIGENS

### 3.1 | Disordered antigens are targets of antibody recognition

The preceding discussion has established that disordered proteins are  
 important antigens in a broad range of immunological contexts. It  
 would be natural to suppose, then, that antibodies are able to



**FIGURE 1** Disordered antigens are *bona fide* targets of antibody recognition. A, Epitopes within disordered protein regions are more likely to be targets of positive antibody binding assays. B, Antibodies to disordered epitopes (purple) and ordered epitopes (green) are subject to similar levels of somatic hypermutation. C, Disorder accounts for only a small fraction of the variability in antibody affinity. D, Disordered epitopes (purple) are exclusively short linear epitopes, while ordered epitopes (green) are predominantly conformational epitopes spanning many residues in primary sequence. Modified with permission from Ref. 43

167 effectively recognize disordered antigens. On the other hand, we have  
 168 also alluded to a number of arguments that suggest the antibody  
 169 response to disordered antigens will be dysfunctional or even  
 170 absent.<sup>[14,15,19–21,42]</sup> To address this apparent contradiction, we recently  
 171 assembled a large dataset<sup>[43]</sup> of well-characterized protein antigens from  
 172 the Immune Epitope Database (IEDB).<sup>[44]</sup> Epitopes arising from protein  
 173 regions predicted to be disordered are well represented in this dataset.  
 174 In fact, we found a positive correlation between predicted disorder and  
 175 the likelihood that any putative epitope will test positive in an antibody  
 176 binding assay reported in the IEDB, establishing that disorder is no  
 F1 177 impediment to effective antibody recognition (Figure 1A). Others have  
 178 recently made the same observation, and extended it to show that pre-  
 179 dictors of protein disorder out-perform dedicated sequence-based pre-  
 180 dictors for the identification of B-cell epitopes.<sup>[45]</sup> This may reflect the  
 181 fact that residues that score highly in disorder predictors are more likely  
 182 to be solvent exposed,<sup>[43]</sup> as solvent exposure is both a prerequisite for,  
 183 and established predictor of, antigenicity.<sup>[46]</sup>

by which B-cells mature and differentiate into memory or antibody  
 188 producing cells.<sup>[14,15,40,42]</sup> In the process of this maturation, antibodies  
 189 undergo somatic hypermutation and selection for high-affinity antigen  
 190 binding. This affinity maturation is essential to the development of a  
 191 high-affinity and long-lived antibody response. If the affinity maturation  
 192 of antibodies to disordered antigens were significantly impeded, we  
 193 would expect such antibodies to have fewer mutations with respect to  
 194 germline antibody sequences than do antibodies to ordered antigens.  
 195 In fact, when we compared the available antibody sequences to the  
 196 corresponding germline sequences, we detected no difference in the  
 197 number of V-gene mutations between the two classes of antibodies  
 198 (Figure 1B), implying that affinity maturation proceeds with equivalent  
 199 efficacy for antibodies to ordered and disordered antigens.<sup>[43]</sup> 200

Highly repetitive antigens are capable of eliciting B-cell responses  
 201 that are independent of the T-cell help normally required for B-cell  
 202 maturation. Such responses are characterized by defective affinity mat-  
 203 uration and a failure to establish a conventional memory response.<sup>[47]</sup> 204  
 As such, this has been proposed as one reason that B-cells to repetitive  
 205 and disordered antigens may fail to mature.<sup>[15]</sup> Indeed, the repetitive  
 206 malaria antigen CSP can induce a B-cell response in the absence of T-  
 207 cell help, but this response is much weaker and more transient than  
 208 that in the presence of T-cells. Thus, the normal response to this pro-  
 209 typical repetitive antigen is T-cell dependent.<sup>[48]</sup> 210

### 184 3.2 | Affinity maturation of antibodies to disordered 185 antigens

186 It has been argued that disordered antigens may impede the develop-  
 187 ment of an effective immune response by interfering with the process

211 Additional specific evidence for affinity maturation in antibodies  
 212 targeting this epitope comes from recent studies of the antibody  
 213 responses induced in clinical trials of the RTS,S vaccine,<sup>[49]</sup> and in the  
 214 context of natural infection.<sup>[50]</sup> These studies have established the  
 215 presence of extensive somatic hyper-mutation in both infection- and  
 216 vaccine-derived antibody responses. Hyper-mutation in the vaccine-  
 217 induced response was influenced strongly by the vaccination schedule  
 218 and correlated with antibody avidity, confirming that affinity maturation  
 219 is effective in this context. As such, it is clear that the presence of  
 220 a transient T-cell independent response to repetitive and disordered  
 221 antigens such as CSP does not preclude the development of a classical  
 222 T-cell dependent response and accompanying affinity maturation.

### 223 3.3 | Effect of disorder on the affinity of 224 antigen-antibody interactions

225 A further basis for the misconception that disordered proteins will be  
 226 poorly recognized by antibodies arises from the fact that disordered  
 227 epitopes, by definition, fail to adopt a single well-defined conformation  
 228 in the context of the native protein that might be recognized by an  
 229 antibody. It is clear, however, that disordered proteins are capable of  
 230 being recognized by protein binding partners, antibodies or otherwise.  
 231 To achieve this, disordered proteins often undergo a process of  
 232 coupled folding and binding in which the binding partner acts effectively  
 233 as a template into which the disordered protein can fold into a  
 234 relatively well-defined bound conformation.<sup>[51]</sup> Nonetheless, coupled  
 235 folding and binding places both kinetic and thermodynamic constraints  
 236 on the interaction.<sup>[52]</sup> In particular, the significant entropic cost of folding  
 237 reduces the overall binding affinity of these interactions, unless this  
 238 cost can be offset by other means.<sup>[53]</sup> Thus, the interactions of disordered  
 239 proteins are often of relatively low affinity, despite maintaining  
 240 high specificity. These properties are ideally suited to mediators of signals  
 241 that must be switched on or off rapidly, accounting for the enrichment  
 242 of disorder in proteins involved in signal transduction and  
 243 regulation.<sup>[7]</sup> These interactions are quite distinct, however, from typical  
 244 antibody-antigen interactions, which are selected for high affinity  
 245 and slow dissociation.

246 These considerations raise the possibility that antibodies that recognize  
 247 disordered antigens may do so with reduced affinity, reflecting  
 248 the substantial entropic cost that must accompany the disordered antigen  
 249 adopting a single antibody-bound conformation.<sup>[54-56]</sup> In fact, using  
 250 our dataset of ordered and disordered epitopes from the IEDB, we  
 251 have shown that disorder accounts for very little of the wide variation  
 252 observed in the affinities of antibodies for protein antigens.<sup>[43]</sup> Specifically,  
 253 we observed an approximately seven fold difference in median  
 254 affinity between antibodies that recognize ordered epitopes and those  
 255 that recognize disordered epitopes, a loss of affinity that is very much  
 256 smaller than might be expected on the basis of the expected entropic  
 257 cost of coupled folding and binding (Figure 1C).<sup>[54,55]</sup>

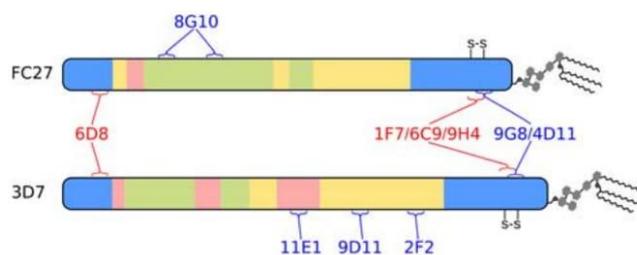
258 To investigate the structural basis for the relatively high affinity  
 259 with which antibodies recognize disordered antigens, we cross-  
 260 referenced our IEDB-derived dataset of ordered and disordered antigens  
 261 with the Protein Databank (PDB),<sup>[57]</sup> resulting in a set of 872

structures of antibody-protein complexes, of which 69 represent disordered  
 epitopes.<sup>[43]</sup> Comparison of these structures revealed that disordered  
 epitopes involve substantially fewer residues than ordered  
 epitopes, perhaps as a means to minimize the entropic cost of binding  
 (Figure 1D). Thus, disordered epitopes rarely span more than 12  
 residues of the antigen primary sequence, and the majority of these  
 residues make direct interactions with antibody. These structural  
 observations are consistent with epitope mapping experiments with a  
 large number of antibodies against several disordered antigens,<sup>[38]</sup>  
 which imply that disordered epitopes are overwhelmingly linear epitopes,  
 and that antibody recognition is determined largely by the local  
 antigen sequence. This stands in contrast to more conventional ordered  
 antigens, in which a large fraction of the antibody response is typically  
 directed to conformational epitopes.

A further implication of the reduction in the size of disordered epitopes  
 is that disordered epitopes are much more efficient than ordered  
 epitopes in their interactions with antibody, in the sense that they  
 achieve more binding free energy per unit of interaction surface area,  
 or per epitope residue.<sup>[43]</sup> Comparison of the structures of antibody-bound  
 ordered and disordered antigens reveals a number of features of the  
 antibody-antigen interaction that may contribute to this efficiency,  
 including a more concave antigen binding site with better complementarity  
 to the shape of the antigen, as well as a two-fold increase in the number  
 of inter-molecular H-bonds per epitope residue. These structural features  
 imply that antibodies to disordered epitopes are better matched to the  
 bound conformation of that epitope than are antibodies to ordered  
 epitopes. It is likely that this reflects to some extent the ability of the  
 disordered epitope to adapt to its binding site on the antibody. Indeed,  
 this flexibility permits some disordered epitopes to adopt distinct  
 conformations bound to different antibodies.<sup>[50,58,59]</sup> Nonetheless, this  
 exquisite complementarity is also likely to require adaptation on the part  
 of the antibody.<sup>[50]</sup> As such, it may also be seen as evidence to support  
 the conclusion that affinity maturation of antibodies to disordered  
 epitopes is at least as effective as is affinity maturation targeting ordered  
 epitopes.

## 297 4 | THE IMPLICATIONS OF DISORDER FOR 298 VACCINE DEVELOPMENT: EXAMPLES 299 FROM THE MALARIA ANTIGEN MSP2

We now place the preceding general considerations on a more specific  
 footing, by outlining their implications in the context of vaccine development  
 efforts against the disordered malaria antigen MSP2.<sup>[39,60]</sup> MSP2 is a highly  
 abundant, C-terminally glycosylphosphatidylinositol (GPI)-anchored blood-stage  
 surface antigen of the malaria parasite *P. falciparum* and a potential component  
 of a malaria vaccine.<sup>[36,61-63]</sup> MSP2 has conserved N- and C-terminal regions,  
 flanking a central region comprising polymorphic repetitive sequences and  
 non-repetitive sequences that are predominantly dimorphic (Figure 2). These  
 dimorphic sequences within the central variable region differentiate MSP2  
 variants into two allelic families, 3D7 and FC27 MSP2.<sup>[64,65]</sup> The Combination  
 B vaccine, which included 3D7 MSP2 as one of three protein antigens,  
 yielded evidence for protection in a clinical trial in Papua New



**FIGURE 2** Schematic of the primary structure of the two allelic families of MSP2. Regions of conserved (blue), repetitive (green), dimorphic (yellow), and polymorphic (pink) sequence are shown. The epitopes of a panel of monoclonal antibodies are also shown, with antibodies that strongly recognize the parasite antigen in blue text, and those that do so only weakly or not at all in red

development is the ability to omit ineffective regions such as these polymorphic regions entirely, ensuring that the response is focused on conserved epitopes that are more likely to elicit broadly protective responses. This is exemplified by the long synthetic peptides derived from the C-terminal half of MSP2, which have been considered potential components of a malaria vaccine.<sup>[41,67]</sup> These peptides include both conserved and dimorphic family-specific sequences, but exclude the highly polymorphic regions of MSP2.

## 4.2 | Lipid interactions modulate the conformation and antigenicity of MSP2

Beyond the specificity of the immediate epitope-antibody interaction, the broader context in which the antibody encounters the antigen is likely to influence the way in which antibodies are able to recognize disordered antigens. Such effects are well-known in the case of ordered antigens, and can arise when, for example, molecular interactions lead to conformational changes that mask or disrupt epitopes or otherwise render them cryptic.<sup>[68,69]</sup> Similar effects are likely in disordered antigens; indeed, the conformational properties of disordered proteins are uniquely sensitive to modulation by interactions, post-translational modifications and other mechanisms. As such, antibody recognition of disordered epitopes may be expected to be particularly sensitive to the context in which the antibody is recognized.

An example of this is seen in the masking of conserved epitopes in MSP2. Amongst a panel of monoclonal mouse antibodies raised against recombinant MSP2, all of those targeting variable epitopes also recognized MSP2 on the parasite surface, whereas many antibodies targeting conserved epitopes recognized the parasites weakly or not at all (Figure 2).<sup>[38]</sup> Together, these data imply that at least some conserved epitopes within MSP2 may be masked by the conformations or interactions adopted by the antigen on the parasite surface.

The structural basis of this masking has recently been characterized in the case of one such epitope. 6D8, a murine monoclonal antibody recognizing an epitope in the conserved N-terminal region binds recombinant MSP2 and epitope-derived peptides with high affinity, yet it fails to recognize the parasite surface.<sup>[38]</sup> The N-terminal region of MSP2, including the 6D8 epitope, binds lipids, adopting an amphipathic helical conformation that is expected to be heavily populated in parasite MSP2 given that it is GPI-anchored to the parasite membrane.<sup>[70]</sup> This helical conformation is incompatible with the conformation that is recognized by 6D8, as revealed by X-ray crystallography (Figure 3A, B).<sup>[71]</sup> Moreover, the 6D8 epitope is lost in recombinant MSP2 when the protein is C-terminally tethered to lipid vesicles,<sup>[71]</sup> indicating that the N-terminal interaction with lipid and the accompanying conformational change are sufficient to account for the observed masking of the 6D8 epitope on the parasite surface.

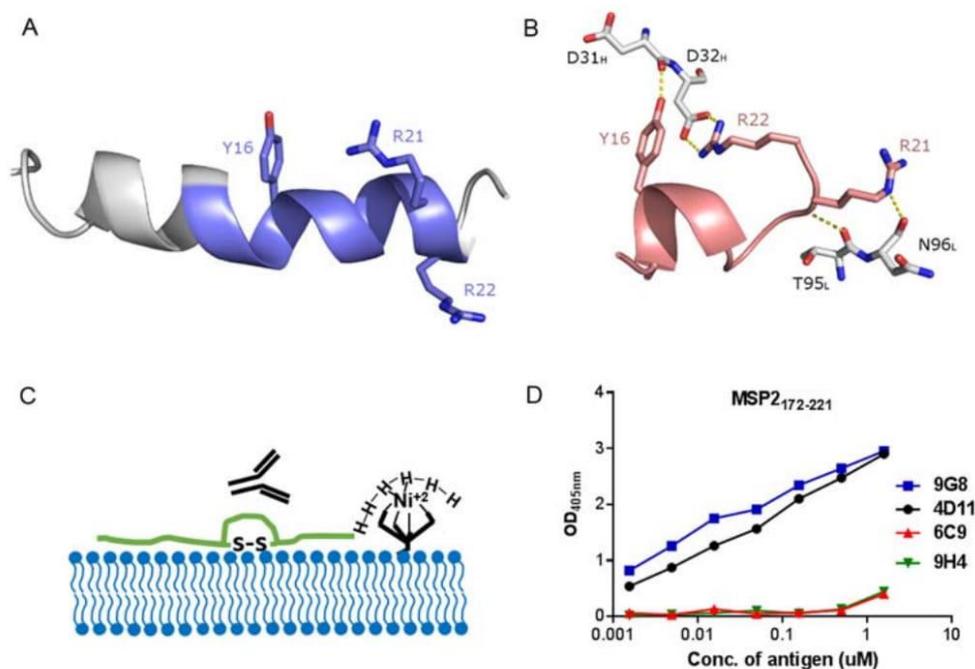
To explore the impact of lipid conjugation on the C-terminal domain of MSP2, a shorter construct of 50 residues, MSP2<sub>172-221</sub>, consisting of only the conserved C-terminal region, was generated. Our NMR data suggest that MSP2<sub>172-221</sub> is unstructured in solution and that the conformational properties of the C-terminal region are identical in MSP2<sub>172-221</sub>, 3D7 and FC27 MSP2. NMR studies also indicate

313 Guinea, but this protection was strain-specific, with vaccines not being  
314 protected from malarial strains expressing FC27-family MSP2 sequen-  
315 ces.<sup>[37]</sup> Thus, a primary concern for vaccine development against  
316 MSP2, like many other malaria antigens, has been to address the poly-  
317 morphism in the antigen.<sup>[66]</sup>

## 4.1 | Disorder and antigenicity in MSP2

319 Although the studies described above have established that disorder is  
320 no impediment to the development of a robust antibody response, the  
321 interplay between disorder and antigenicity may be more complicated  
322 at the level of a single antigen. In our studies of MSP2, we have com-  
323 pared experimental measures of conformational disorder derived from  
324 NMR relaxation rates with observed patterns of antigenicity inferred  
325 from experimental immunizations of animals and humans.<sup>[40]</sup> The result  
326 of that study, in contrast to the bioinformatic analysis,<sup>[43,45]</sup> was a neg-  
327 ative correlation between antigenicity and disorder, with the most flexi-  
328 ble regions of MSP2 being the least likely to be recognized by  
329 antibodies in MSP2-immune sera, and with regions that show slight  
330 conformational restriction apparently dominating the antibody  
331 response in these immunizations. It is important to note that for MSP2  
332 the experimental measures of disorder used in this study correlate  
333 poorly with the disorder prediction scores used in the bioinformatics  
334 analysis, reflecting the uniformly high levels of disorder across MSP2  
335 (>85% of MSP2 residues have IUPred scores >0.8). Indeed, the most  
336 flexible regions of MSP2 are extremely disordered, being composed  
337 exclusively of the small amino acids Gly, Ser, and Ala. Thus, it may be  
338 that this extreme level of disorder and attendant low sequence com-  
339 plexity is sufficient to significantly impede the development of a strong  
340 antibody response, whereas the levels of disorder more typically  
341 encountered in disordered antigens are compatible with a robust  
342 response, as outlined above.

343 Importantly, these most flexible and dynamic regions in MSP2 are  
344 also the most polymorphic regions in the antigen, and this, together  
345 with their low complexity, renders them unattractive targets for  
346 vaccine development. As such, the greater immunogenicity of other  
347 disordered but less flexible regions of the protein is likely to be  
348 advantageous in the context of vaccine development. Moreover,  
349 a key advantage afforded by peptide-based approaches to vaccine



**FIGURE 3** Interaction of different MSP2-based conserved regions specific peptides with antibody and lipid. Comparison between A, lipid bound N-terminal MSP2<sub>1-25</sub><sup>[70]</sup> and B, 6D8 mAb bound MSP2<sub>14-22</sub> (PDB ID 4QYO),<sup>[71]</sup> key 6D8 paratope residues involved in binding are shown in white. The  $\alpha$ -helical configuration of the lipid-bound peptide removes the backbone flexibility required for Arg22 to access Tyr16, which provides a structural rationale for 6D8 epitope masking at the parasite membrane. C, Schematic of lipid tethering of C-terminal region of MSP2 (MSP2<sub>172-221</sub>) where MSP2<sub>172-221</sub> was synthesized with a C-terminal His<sub>6</sub>-tag to immobilize on nickel bound to nickel-chelating lipid. D, ELISA showing the effects of lipid tethering on the binding of four C-terminal region-specific mouse mAbs for MSP2<sub>172-221</sub>.<sup>[72]</sup>

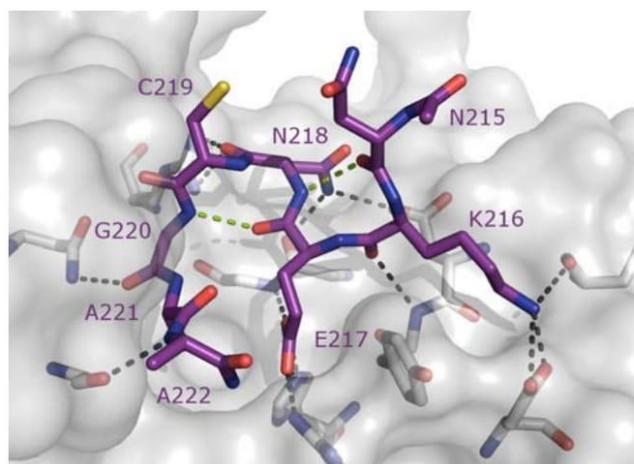
401 that many residues in MSP2<sub>172-221</sub> interact with lipid, including some in  
 402 epitopes recognized by C-terminal region specific monoclonal antibodies  
 403 (mAbs). In contrast to the N-terminal region however, there is no indica-  
 404 tion of stable helical conformation within lipid-bound MSP2<sub>177-221</sub>.<sup>[71,72]</sup>  
 405 Nonetheless, the antigenicity of liposome-conjugated MSP2<sub>172-221</sub> and  
 406 full-length MSP2 are altered in comparison to the lipid-free recombinant  
 407 peptides (Figure 3C). Importantly, it also appears that the binding pattern  
 408 of C-terminal specific mAbs for liposome-conjugated antigens (Figure 3D)  
 409 is consistent with that of the mAbs for parasite lysate. Thus, lipid teth-  
 410 ering appears to alter the antigenic state of these MSP2 constructs, render-  
 411 ing it more similar to that of the antigen on the parasite surface. Lipid-  
 412 conjugation of MSP2 based vaccines may therefore provide a more effec-  
 413 tive antibody response.

414 Similar masking of potentially protective epitopes is seen in the pre-  
 415 erythrocytic malaria antigen CSP. Although the repetitive sequences in  
 416 CSP have long been the focus of vaccine development efforts, recent  
 417 studies have identified epitopes within the conserved N-terminal region  
 418 that appear to contribute to the protective response to CSP in the con-  
 419 text of natural immunity and vaccination.<sup>[73-75]</sup> These antibodies act  
 420 synergistically with antibodies targeting the CSP repeats.<sup>[75]</sup> In the  
 421 course of parasite development within the mosquito and subsequent  
 422 migration through the human host to the liver, CSP undergoes proteo-  
 423 lytic processing, which is essential for the establishment of a liver-stage  
 424 infection. This proteolysis exposes otherwise cryptic epitopes in both  
 425 the N- and C-terminal regions of CSP, indicating conformational  
 426 changes that appear to involve long-range interactions within CSP,

427 which in turn determine the accessibility of some of these N-terminal  
 428 epitopes.<sup>[74,76,77]</sup> The nature of these interactions and the associated  
 429 conformational changes remain largely obscure. In light of the efficacy  
 430 of antibodies targeting N-terminal and possibly also C-terminal epitopes,  
 431 it is likely that a better understanding of these interactions will facilitate  
 432 the design of improved CSP-based vaccine constructs.

### 4.3 | Interactions outside the defined epitope may modulate the specificity of antibodies targeting disordered antigens

436 Disordered proteins typically adopt a well-defined structure when  
 437 interacting with binding partners, as described above. Nonetheless, in  
 438 some cases significant disorder persists in these interactions, giving rise  
 439 to what are known as “fuzzy interactions.”<sup>[78]</sup> It is becoming increas-  
 440 ingly clear that these fuzzy interactions modulate the function of disor-  
 441 dered proteins in several different ways. In particular, transient  
 442 interactions involving the disordered regions flanking the structurally  
 443 defined binding site can modulate the affinity with which disordered  
 444 proteins are recognized by their binding partners.<sup>[79]</sup> The extent to  
 445 which such interactions modulate the antibody recognition of disor-  
 446 dered antigens remains largely unexplored, but we have recently char-  
 447 acterized one example which suggests that such modulation is feasible,  
 448 and possibly quite general. The mAb 6D8, introduced above, recognizes  
 449 a conserved epitope near the N-terminus of MSP2, yet 6D8 binds the  
 450 two allelic forms of MSP2 with affinities that differ by approximately



**FIGURE 4** Murine mAb 4D11 Fv in complex with its cognate 8-residue epitope (PDB ID 5TBD).<sup>[81]</sup> Intramolecular hydrogen bonds are indicated by green dashed lines. Interactions with 4D11 Fv paratope are shown in black dashed lines

found to adopt a  $\beta$ -bend ribbon conformation, characterized by consecutive overlapping  $\beta$ -turns (Figure 4). The two  $\beta$ -turns in the epitope were stabilized by hydrogen bonds involving Asn215-Asn218 and Glu217-Gly220, and allowed key interactions of residues Lys216, Glu217, and Asn218 with the 4D11 paratope. This structural information suggests strategies by which stabilized analogues of the epitope could be optimized for use as a peptide vaccine, with a view to guiding the specificity of the antibody response towards 4D11-like epitopes that are both conserved and accessible on the parasite surface, and thus are more likely to contribute to a broadly protective immune response.

## 5 | CONCLUSIONS AND FUTURE DIRECTIONS

Disordered proteins are highly abundant in infectious organisms across the kingdoms of life, and these proteins are *bona fide* targets of antibody responses in the context of natural infection and vaccination.<sup>[43]</sup> In contrast to widely-held views that disordered antigens will fail to elicit an effective antibody response,<sup>[14–21]</sup> it is clear that they play an important role as protective antigens in a range of diseases (Table 3). Epitopes from disordered antigens are typically short linear epitopes, not the more complex conformational epitopes dominant in ordered antigens. As such, disordered antigens are well suited as targets of peptide vaccines.

Despite the potential promise of this approach, careful consideration must be paid to the context in which target epitopes are recognized by the immune system. The conformational and antigenic properties of disordered antigens can be strongly modulated by their interactions, both in their native parasite context and in any vaccine formulation. Mismatch between the parasite and vaccine context has probably contributed to the failure of a number of peptide-based vaccine strategies.<sup>[82]</sup> The use of long peptides in a vaccine formulation, potentially incorporating several protective epitopes within their native sequence, may offer one strategy to address this issue.<sup>[67,83]</sup> Potentially a more effective approach will be to define the structures of disordered peptide epitopes bound to their cognate antibodies, and to use this structural information as a basis for design of peptides in which these structures are stabilized. To this end, the range of available strategies for stabilizing the conformation of synthetic peptides represents a clear advantage to the peptide vaccine approach. Future developments in this area, enabling the stabilization of a wider range of peptide conformations while limiting perturbation of residues essential for antibody binding, will serve to broaden the scope of the strategy further.

It is clear that the targeted vaccine development strategy described here requires a detailed understanding of the functional determinants of immune protection, and of the specific antigens and epitopes responsible for mediating these functions. While this is well understood for some pathogens, for other less studied or immunologically complex pathogens, such information may still be emerging. In this context, new strategies for high-throughput profiling of immunological responses and identification of correlates of immune protection<sup>[84]</sup> represent a welcome advance that promises to allow peptide-based vaccine development against diseases for which no effective vaccine currently exists.

51 five fold.<sup>[71]</sup> NMR experiments provide evidence of transient interactions that are not resolved in the crystal structure and which involve 52 variable residues that are distal to the structurally defined epitope. 53 These transient or fuzzy interactions thus appear to modulate the specificity of 6D8, rendering the antibody somewhat strain-specific in spite of the conserved nature of its epitope.<sup>[71,80]</sup> Although the generality of 54 this effect remains to be established, it appears to originate from largely non-specific interactions that occur, almost inevitably, as a result 55 of the high effective concentration of the disordered regions that immediately flank the structurally defined epitope.<sup>[80]</sup> As such, it is 56 expected that similar fuzzy interactions may contribute to antibody recognition of many disordered antigens.

### 4.4 | Toward a structural vaccinology of disordered antigens

57 The conserved C-terminal region of MSP2 is particularly immunodominant, with a total of five mouse mAbs able to recognize an overlapping 58 epitope spanning MSP2<sub>207–224</sub>.<sup>[38]</sup> As discussed above, antibodies targeting the conserved regions have shown differing ability to recognize parasite 59 MSP2 (Figure 2). The mAbs 4D11 and 9G8 are able to bind to parasite MSP2 by western blot and immunofluorescence assay (IFA) 60 whilst the mAbs 9H4, 6C9 and 1F7, show significantly weaker signal by IFA and no binding by western blot, despite the five antibodies recognizing 61 overlapping linear epitopes. The close proximity of the C-terminal region to the parasite surface via its GPI tether may influence the local 62 conformation, making parts of the epitope inaccessible to antibody binding. Understanding the conformation of epitopes bound to antibodies 63 such as 4D11 and 9G8 affords a good platform for further peptide vaccine design. By presenting the immune system with epitopes that are 64 accessible on the parasite surface whilst removing distracting sequences, a more efficient and specific immune response may be elicited.

65 Recently, the crystal structure of 4D11 Fv bound to its minimal binding epitope MSP2<sub>215–222</sub>, which it also shares with 9G8, was solved at 2.2 Å.<sup>[81]</sup> Upon antibody binding, the 8-residue peptide epitope was

535 **ACKNOWLEDGMENTS**

536 This work was supported in part by the Australian National Health and  
537 Medical Research Council (project grants 1042520 and 1125788 to  
538 CAM and RSN). RSN acknowledges fellowship support from the  
539 NHMRC. The authors thank Robin Anders and Jack Richards for many  
540 helpful discussions in the course of their studies on MSP2.

541 **ORCID**

542 Christopher A. MacRaid  <http://orcid.org/0000-0002-3694-3989>

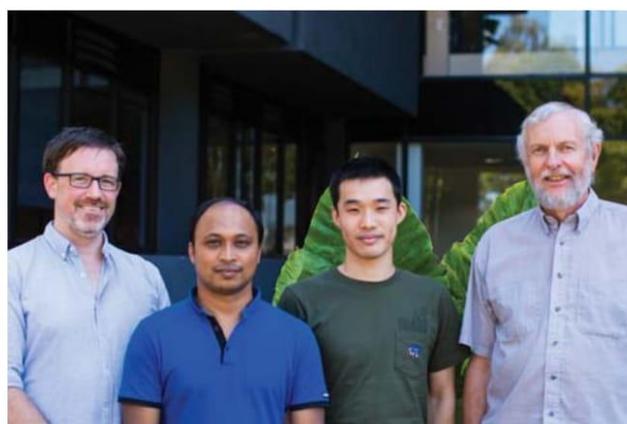
543 **REFERENCES**

- 544 [1] A. M. Stern, H. Markel, *Health Aff (Millwood)* **2005**, *24*, 611.
- 545 [2] M. Skwarczynski, I. Toth, *Chem. Sci.* **2016**, *7*, 842.
- 546 [3] R. F. Anders, *Trends Parasitol.* **2011**, *27*, 330.
- 547 [4] T. Kumai, A. Fan, Y. Harabuchi, E. Celis, *Curr. Opin. Immunol.* **2017**,  
548 *47*, 57.
- 549 [5] J. A. Robinson, *J. Pept. Sci.* **2013**, *19*, 127.
- 550 [6] A. K. Dunker, Z. Obradovic, P. Romero, E. C. Garner, C. J. Brown,  
551 *Genome Inform. Ser. Workshop Genome Inform.* **2000**, *11*, 161.
- 552 [7] P. E. Wright, H. J. Dyson, *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18.
- 553 [8] R. Pancsa, P. Tompa, *PLoS One* **2012**, *7*, e34687.
- 554 [9] B. Xue, D. Blocquel, J. Habchi, A. V. Uversky, L. Kurgan, V. N. Uver-  
555 sky, S. Longhi, *Chem. Rev.* **2014**, *114*, 6880.
- 556 [10] A. Mohan, W. J. Sullivan, Jr., P. Radivojac, A. K. Dunker, V. N. Uver-  
557 sky, *Mol. Biosyst.* **2008**, *4*, 328.
- 558 [11] M. Marin, V. N. Uversky, T. Ott, *Plant Cell* **2013**, *25*, 3153.
- 559 [12] A. J. Guy, V. Irani, C. A. MacRaid, R. F. Anders, R. S. Norton, J. G. Beeson,  
560 J. S. Richards, P. A. Ramsland, *PLoS One* **2015**, *10*, e0141729.
- 561 [13] Z. P. Feng, X. Zhang, P. Han, N. Arora, R. F. Anders, R. S. Norton,  
562 *Mol. Biochem. Parasitol.* **2006**, *150*, 256.
- 563 [14] R. F. Anders, *Parasite Immunol.* **1986**, *8*, 529.
- 564 [15] L. Schofield, *Parasitol Today* **1991**, *7*, 99.
- 565 [16] T. A. Mendes, F. P. Lobo, T. S. Rodrigues, G. F. Rodrigues-Luiz, W.  
566 D. daRocha, R. T. Fujiwara, S. M. Teixeira, D. C. Bartholomeu, *Mol.*  
567 *Biol. Evol.* **2013**, *30*, 951.
- 568 [17] M. U. Ferreira, M. da Silva Nunes, G. Wunderlich, *Clin. Diagn. Lab.*  
569 *Immunol.* **2004**, *11*, 987.
- 570 [18] A. L. Hughes, *J. Mol. Evol.* **2004**, *59*, 528.
- 571 [19] C. J. Penkett, C. Redfield, J. A. Jones, I. Dodd, J. Hubbard, R. A.  
572 Smith, L. J. Smith, C. M. Dobson, *Biochemistry* **1998**, *37*, 17054.
- 573 [20] A. K. Dunker, C. J. Brown, J. D. Lawson, L. M. Iakoucheva, Z. Oba-  
574 radovic, *Biochemistry* **2002**, *41*, 6573.
- 575 [21] P. D. Kwong, M. L. Doyle, D. J. Casper, C. Cicala, S. A. Leavitt, S.  
576 Majeed, T. D. Steenbeke, J. Arthos, *Nature* **2002**, *420*, 678.
- 577 [22] M. J. Gardner, N. Hall, E. Fung, O. White, M. Berriman, R. W.  
578 Hyman, J. M. Carlton, A. Pain, K. E. Nelson, S. Bowman, I. T. Paul-  
579 sen, K. James, J. A. Eisen, K. Rutherford, S. L. Salzberg, A. Craig, S.  
580 Kyes, M. S. Chan, V. Nene, S. J. Shallom, B. Suh, J. Peterson, S.  
581 Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A.  
582 B. Vaidya, D. M. Martin, A. H. Fairlamb, M. J. Fraunholz, D. S. Roos,  
583 S. A. Ralph, G. I. McFadden, L. M. Cummings, G. M. Subramanian,  
584 C. Mungall, J. C. Venter, D. J. Carucci, S. L. Hoffman, C. Newbold,  
585 R. W. Davis, C. M. Fraser, B. Barrell, *Nature* **2002**, *419*, 498.
- 586 [23] M. A. DePristo, M. M. Zilversmit, D. L. Hartl, *Gene* **2006**, *378*, 19.
- 587 [24] H. Y. Xue, D. R. Forsdyke, *Mol. Biochem. Parasitol.* **2003**, *128*, 21.
- [25] M. M. Zilversmit, S. K. Volkman, M. A. DePristo, D. F. Wirth, P. 588  
Awadalla, D. L. Hartl, *Mol. Biol. Evol.* **2010**, *27*, 2198. 589
- [26] A. S. Aly, A. M. Vaughan, S. H. Kappe, *Annu. Rev. Microbiol.* **2009**, 590  
*63*, 195. 591
- [27] P. D. Crompton, S. K. Pierce, L. H. Miller, *J. Clin. Invest.* **2010**, *120*, 592  
*4168*. 593
- [28] M. B. Doud, A. C. Koksai, L.-Z. Mi, G. Song, C. Lu, T. A. Springer, 594  
*Proc. Natl. Acad. Sci. USA* **2012**. 595
- [29] L. Foquet, C. C. Hermsen, G. J. van Gemert, E. Van Braeckel, K. E. 596  
Weening, R. Sauerwein, P. Meuleman, G. Leroux-Roels, *J Clin Invest* 597  
**2014**, *124*, 140. 598
- [30] A. Olotu, J. Lusingu, A. Leach, M. Lievens, J. Vekemans, S. Msham, 599  
T. Lang, J. Gould, M. C. Dubois, E. Jongert, P. Vansadia, T. Carter, P. 600  
Njuguna, K. O. Awuondo, A. Malabeja, O. Abdul, S. Gesase, N. 601  
Mturi, C. J. Drakeley, B. Savarese, T. Villafana, D. Lapiere, W. R. 602  
Ballou, J. Cohen, M. M. Lemnge, N. Peshu, K. Marsh, E. M. Riley, L. 603  
von Seidlein, P. Bejon, *Lancet Infect Dis.* **2011**, *11*, 102. 604
- [31] V. S. Moorthy, W. R. Ballou, *Malar. J.* **2009**, *8*, 312. 605
- [32] H. J. Dyson, A. C. Satterthwait, R. A. Lerner, P. E. Wright, *Biochemis- 606  
try* **1990**, *29*, 7828. 607
- [33] R. R. Taylor, S. J. Allen, B. M. Greenwood, E. M. Riley, *Am. J. Trop. 608  
Med. Hyg.* **1998**, *58*, 406. 609
- [34] F. Al-Yaman, B. Genton, R. F. Anders, M. Falk, T. Triglia, D. Lewis, J. 610  
Hii, H. P. Beck, M. P. Alpers, *Am. J. Trop. Med. Hyg.* **1994**, *51*, 593. 611
- [35] W. G. Metzger, D. M. Okenu, D. R. Cavanagh, J. V. Robinson, K. A. 612  
Bojang, H. A. Weiss, J. S. McBride, B. M. Greenwood, D. J. Conway, 613  
*Parasite Immunol.* **2003**, *25*, 307. 614
- [36] B. Genton, I. Betuela, I. Felger, F. Al-Yaman, R. F. Anders, A. Saul, L. 615  
Rare, M. Baisor, K. Lorry, G. V. Brown, D. Pye, D. O. Irving, T. A. 616  
Smith, H. P. Beck, M. P. Alpers, *J. Infect. Dis.* **2002**, *185*, 820. 617
- [37] C. Flück, T. Smith, H. P. Beck, A. Irion, I. Betuela, M. P. Alpers, R. 618  
Anders, A. Saul, B. Genton, I. Felger, *Infect. Immun.* **2004**, *72*, 6300. 619
- [38] C. G. Adda, C. A. MacRaid, L. Reiling, K. Wycherley, M. J. Boyle, V. 620  
Kienzle, P. Masendycz, M. Foley, J. G. Beeson, R. S. Norton, R. F. 621  
Anders, *Infect. Immun.* **2012**, *80*, 4177. 622
- [39] X. Zhang, M. A. Perugini, S. Yao, C. G. Adda, V. J. Murphy, A. Low, 623  
R. F. Anders, R. S. Norton, *J. Mol. Biol.* **2008**, *379*, 105. 624
- [40] C. A. MacRaid, M. Zachrdla, D. Andrew, K. Bankala, J. Nováček, L. 625  
Židek, V. Sklenář, J. S. Richards, J. G. Beeson, R. F. Anders, R. S. 626  
Norton, *PLoS One* **2015**, *10*, e0119899. 627
- [41] S. Balam, S. Olugbile, C. Servis, M. Diakitè, A. D'Alessandro, G. 628  
Frank, R. Moret, G. Corradin, *Malar. J.* **2014**, *13*, 510. 629
- [42] P. L. Carl, B. R. Temple, P. L. Cohen, *Arthritis Res. Ther.* **2005**, *7*, R1360. 630
- [43] C. A. MacRaid, J. S. Richards, R. F. Anders, R. S. Norton, *Structure* 631  
**2016**, *24*, 148. 632
- [44] R. Vita, J. A. Overton, J. A. Greenbaum, J. Ponomarenko, J. D. Clark, 633  
J. R. Cantrell, D. K. Wheeler, J. L. Gabbard, D. Hix, A. Sette, B. 634  
Peters, *Nucleic Acids Res.* **2015**, *43*, D405. 635
- [45] K. S. Rahman, E. U. Chowdhury, K. Sachse, B. Kaltenboeck, *J. Biol.* 636  
*Chem.* **2016**, *291*, 14585. 637
- [46] J. Novotny, M. Handschumacher, E. Haber, R. E. Bruccoleri, W. B. 638  
Carlson, D. W. Fanning, J. A. Smith, G. D. Rose, *Proc. Natl. Acad.* 639  
*Sci. USA* **1986**, *83*, 226. 640
- [47] T. Defrance, M. Taillardet, L. Genestier, *Curr. Opin. Immunol.* **2011**, *23*, 330. 641
- [48] C. R. Fisher, H. J. Sutton, J. A. Kaczmarek, H. A. McNamara, B. Clif- 642  
ton, J. Mitchell, Y. Cai, J. N. Dups, N. J. D'Arcy, M. Singh, A. Chuah, 643  
T. S. Peat, C. J. Jackson, I. A. Cockburn, *PLoS Pathog.* **2017**, *13*, 644  
e1006469. 645

- 646 [49] J. A. Regules, S. B. Cicatelli, J. W. Bennett, K. M. Paolino, P. S.  
647 Twomey, J. E. Moon, A. K. Kathcart, K. D. Hauns, J. L. Komisar, A. N.  
648 Qabar, S. A. Davidson, S. Dutta, M. E. Griffith, C. D. Magee, M.  
649 Wojnarski, J. R. Livezey, A. T. Kress, P. E. Waterman, E. Jongert, U.  
650 Wille-Reece, W. Volkmoth, D. Emerling, W. H. Robinson, M. Lievens,  
651 D. Morelle, C. K. Lee, B. Yassin-Rajkumar, R. Weltzin, J. Cohen, R. M.  
652 Paris, N. C. Waters, A. J. Birkett, D. C. Kaslow, W. R. Ballou, C. F.  
653 Ockenhouse, J. Vekemans, *J. Infect. Dis.* **2016**, 214, 762.
- 654 [50] G. Triller, S. W. Scally, G. Costa, M. Pissarev, C. Kreschel, A. Bosch,  
655 E. Marois, B. K. Sack, R. Murugan, A. M. Salman, C. J. Janse, S. M.  
656 Khan, S. H. I. Kappe, A. A. Adegnika, B. Mordmuller, E. A. Levashina,  
657 J. P. Julien, H. Wardemann, *Immunity* **2017**, 47, 1197.
- 658 [51] K. Sugase, H. J. Dyson, P. E. Wright, *Nature* **2007**, 447, 1021.
- 659 [52] V. N. Uversky, *Biochim. Biophys. Acta* **2013**, 1834, 932.
- 660 [53] R. Pancsa, M. Fuxreiter, *IUBMB Life* **2012**, 64, 513.
- 661 [54] C. Bracken, P. A. Carr, J. Cavanagh, A. G. Palmer, III, *J. Mol. Biol.*  
662 **1999**, 285, 2133.
- 663 [55] M. C. Baxa, E. J. Haddadian, J. M. Jumper, K. F. Freed, T. R. Sosnick,  
664 *Proc. Natl. Acad. Sci. USA* **2014**, 111, 15396.
- 665 [56] G. Papadakos, A. Sharma, L. E. Lancaster, R. Bowen, R. Kaminska, A.  
666 P. Leech, D. Walker, C. Redfield, C. Kleanthous, *J. Am. Chem. Soc.*  
667 **2015**, 137, 5252.
- 668 [57] H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne,  
669 K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J.  
670 Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H.  
671 Weissig, J. D. Westbrook, C. Zardecki, *Acta Crystallogr. D Biol. Cryst-*  
672 *tallogr.* **2002**, 58, 899.
- 673 [58] L. Deng, L. Ma, M. L. Virata-Theimer, L. Zhong, H. Yan, Z. Zhao, E.  
674 Struble, S. Feinstone, H. Alter, P. Zhang, *Proc. Natl. Acad. Sci. USA*  
675 **2014**, 111, 10690.
- 676 [59] H. M. Chu, J. Wright, Y. H. Chan, C. J. Lin, T. W. Chang, C. Lim,  
677 *Nat. Commun.* **2014**, 5, 3139.
- 678 [60] C. G. Adda, V. J. Murphy, M. Sunde, L. J. Waddington, J. Schloegel, G.  
679 H. Talbo, K. Vingas, V. Kienzle, R. Masciantonio, G. J. Howlett, A. N.  
680 Hodder, M. Foley, R. F. Anders, *Mol. Biochem. Parasitol.* **2009**, 166, 159.
- 681 [61] J. A. Smythe, M. G. Peterson, R. L. Coppel, A. J. Saul, D. J. Kemp, R.  
682 F. Anders, *Mol. Biochem. Parasitol.* **1990**, 39, 227.
- 683 [62] P. R. Gilson, T. Nebel, D. Vukcevic, R. L. Moritz, T. Sargeant, T. P.  
684 Speed, L. Schofield, B. S. Crabb, *Mol. Cell Proteom.* **2006**, 5, 1286.
- 685 [63] J. S. McCarthy, J. Marjason, S. Elliott, P. Fahey, G. Bang, E. Malkin,  
686 E. Tierney, H. Aked-Hurditch, C. Adda, N. Cross, J. S. Richards, F. J.  
687 I. Fowkes, M. J. Boyle, C. Long, P. Druilhe, J. G. Beeson, R. F.  
688 Anders, *PLoS One* **2011**, 6, e24413.
- 689 [64] J. A. Smythe, R. L. Coppel, K. P. Day, R. K. Martin, A. M. Oduola, D.  
690 J. Kemp, R. F. Anders, *Proc. Natl. Acad. Sci. USA* **1991**, 88, 1751.
- 691 [65] B. Fenton, J. T. Clark, C. M. Khan, J. V. Robinson, D. Walliker, R.  
692 Ridley, J. G. Scaife, J. S. McBride, *Mol. Cell Biol.* **1991**, 11, 963.
- 693 [66] R. F. Anders, C. G. Adda, M. Foley, R. S. Norton, *Hum. Vaccin.* **2010**, 6, 39.
- 694 [67] C. Flueck, G. Frank, T. Smith, A. Jafarshad, I. Nebie, S. B. Sirima, S.  
695 Olugbile, G. Corradin, *Vaccine* **2009**, 27, 2653.
- 696 [68] V. I. Zarnitsyna, A. H. Ellebedy, C. Davis, J. Jacob, R. Ahmed, R.  
697 Antia, *Philos. Trans. R Soc. Lond. B Biol. Sci.* **2015**, 370, 20140248.
- 698 [69] A. P. Mould, J. A. Askari, A. Byron, Y. Takada, T. A. Jowitt, M. J.  
699 Humphries, *J. Biol. Chem.* **2016**, 291, 20993.
- 700 [70] C. A. MacRaidl, M. Ø. Pedersen, R. F. Anders, R. S. Norton, *Biochim.*  
701 *Biophys. Acta* **2012**, 1818, 2572.
- 702 [71] R. A. V. Morales, C. A. MacRaidl, J. Seow, B. Krishnarjuna, N. Drink-  
703 water, R. Rouet, R. F. Anders, D. Christ, S. McGowan, R. S. Norton,  
704 *Sci. Rep.* **2015**, 5, 10103.
- [72] S. C. Das, R. A. V. Morales, J. Seow, B. Krishnarjuna, R. Dissanayake,  
705 R. F. Anders, C. A. MacRaidl, R. S. Norton, *FEBS J* **2017**, 284, 2649. 706
- [73] S. E. Bongfen, P. M. Ntsama, S. Offner, T. Smith, I. Felger, M. Tan-  
707 ner, P. Alonso, I. Nebie, J. F. Romero, O. Silvie, R. Torgler, G. Corra-  
708 din, *Vaccine* **2009**, 27, 328. 709
- [74] D. Rathore, R. Nagarkatti, D. Jani, R. Chattopadhyay, P. de la Vega,  
710 S. Kumar, T. F. McCutchan, *J. Biol. Chem.* **2005**, 280, 20524. 711
- [75] D. A. Espinosa, G. M. Gutierrez, M. Rojas-Lopez, A. R. Noe, L. Shi,  
712 S. W. Tse, P. Sinnis, F. Zavala, *J. Infect. Dis.* **2015**, 212, 1111. 713
- [76] R. Herrera, C. Anderson, K. Kumar, A. Molina-Cruz, V. Nguyen, M.  
714 Burkhardt, K. Reiter, R. Shimp, Jr., R. F. Howard, P. Srinivasan, M. J.  
715 Nold, D. Ragheb, L. Shi, M. DeCotiis, J. Aebig, L. Lambert, K. M.  
716 Rausch, O. Muratova, A. Jin, S. G. Reed, P. Sinnis, C. Barillas-Mury, P.  
717 E. Duffy, N. J. MacDonald, D. L. Narum, *Infect. Immun.* **2015**, 83,  
718 3771. 719
- [77] A. Coppi, R. Natarajan, G. Pradel, B. L. Bennett, E. R. James, M. A.  
720 Roggero, G. Corradin, C. Persson, R. Tewari, P. Sinnis, *J. Exp. Med.*  
721 **2011**, 208, 341. 722
- [78] P. Tompa, M. Fuxreiter, *Trends Biochem. Sci.* **2008**, 33, 2. 723
- [79] M. Fuxreiter, P. Tompa, *Adv. Exp. Med. Biol.* **2012**, 725, 1. 724
- [80] B. Krishnarjuna, T. Suguki, R. A. V. Morales, J. Seow, T. Fujiwara, K.  
725 L. Wilde, R. S. Norton, C. A. MacRaidl, *Submitted* **2018**. 726
- [81] J. Seow, R. A. Morales, C. A. MacRaidl, B. Krishnarjuna, S.  
727 McGowan, T. Dingjan, G. Jaipuria, R. Rouet, K. L. Wilde, H. S.  
728 Atreya, J. S. Richards, R. F. Anders, D. Christ, N. Drinkwater, R. S.  
729 Norton, *J. Mol. Biol.* **2017**, 429, 836. 730
- [82] A. Irimia, A. Sarkar, R. L. Stanfield, I. A. Wilson, *Immunity* **2016**, 44,  
731 21. 732
- [83] S. Olugbile, C. Kulangara, G. Bang, S. Bertholet, E. Suzarte, V. Villard,  
733 G. Frank, R. Audran, A. Razaname, I. Nebie, O. Awobusuyi, F. Spertini,  
734 A. V. Kajava, I. Felger, P. Druilhe, G. Corradin, *Infect. Immun.* **2009**, 77,  
735 5701. 736
- [84] A. Sette, R. Rappuoli, *Immunity* **2010**, 33, 530. 737

## AUTHOR BIOGRAPHIES

738



From left to right: Chris MacRaidl, Sreedam Das, Jeff Seow, Ray Norton 739  
740

**DR CHRIS MACRAILD** received his Ph.D. from the University of Melbourne 741  
in 2004. His recent research has addressed diverse aspects of the struc- 742  
tural biology of the malaria parasite, with the ultimate goal of establish- 743  
ing new therapeutic opportunities against the devastating disease it 744  
causes. A particular ongoing interest is the structural and antigenic 745

746 properties of the disordered proteins which constitute an important  
747 class of malaria antigen.

published over 350 articles, received numerous national awards, and is an  
inventor on several patents. 754  
755

748 **PROFESSOR RAY NORTON** holds a personal chair at the Monash Institute of  
749 Pharmaceutical Sciences. He graduated with BSc(Honours) from the Uni-  
750 versity of Melbourne and obtained his Ph.D. in chemistry from the Aus-  
751 tralian National University. His group at Monash employs a range of  
752 biophysical approaches (NMR, SPR, ITC and X-ray crystallography) in  
753 studies of peptide and protein toxins and infectious diseases. He has  
761

**How to cite this article:** MacRaidl CA, Seow J, Das SC, Norton  
RS. Disordered epitopes as peptide vaccines. *Peptide Science*.  
2018;e24067. <https://doi.org/10.1002/pep2.24067>

757  
758  
759  
760

WILEY  
Author Proof

# Appendix II

Lipid interactions modulate the structural and antigenic properties of the C-terminal domain of the malaria antigen merozoite surface protein 2

# Lipid interactions modulate the structural and antigenic properties of the C-terminal domain of the malaria antigen merozoite surface protein 2

Sreedam C. Das<sup>1</sup>, Rodrigo A.V. Morales<sup>1</sup>, Jeffrey Seow<sup>1</sup>, Bankala Krishnarjuna<sup>1</sup>, Ravindu Dissanayake<sup>2</sup>, Robin F. Anders<sup>2</sup>, Christopher A. MacRaild<sup>1</sup> and Raymond S. Norton<sup>1</sup>

<sup>1</sup> Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Melbourne, Australia

<sup>2</sup> Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Australia

## Keywords

antigenicity; lipid interaction; liposome; malaria; merozoite surface protein 2

## Correspondence

R. S. Norton and C. A. MacRaild, Monash Institute of Pharmaceutical Sciences, 381 Royal Parade, Parkville, Vic. 3052, Australia  
Tel: +61 399039167 and +61 399039101  
Fax: +61 399039582  
E-mails: ray.norton@monash.edu, chris.macraild@monash.edu

(Received 14 February 2017, revised 18 May 2017, accepted 12 June 2017)

doi:10.1111/febs.14135

Merozoite surface protein 2 (MSP2) is a highly abundant, GPI-anchored antigen on the malaria parasite *Plasmodium falciparum*. MSP2 induces an immune response in the context of natural infections and vaccine trials, and these responses are associated with protection from parasite infection. Recombinant MSP2 is highly disordered in solution but antigenic analyses suggest that it is more ordered on the merozoite surface. We have shown previously that the interaction of recombinant full-length MSP2 with lipid surfaces induces a conformational change in the conserved N-terminal region of MSP2, which contributes to epitope masking in this region. To explore the impacts of lipid interactions on the conformation and antigenicity of the conserved C-terminal region of MSP2, a construct corresponding to this domain, MSP2<sub>172–221</sub>, was designed. NMR studies indicate that many residues in MSP2<sub>172–221</sub> interact with DPC micelles, including some in epitopes recognised by C-terminal-specific monoclonal antibodies, but, in contrast to the MSP2 N-terminus, there is no indication of stable helical conformation. The binding affinities of a panel of monoclonal antibodies indicate that MSP2<sub>172–221</sub> is antigenically similar to full-length MSP2 and show that liposome conjugation alters the antigenicity in a manner that may mimic native MSP2 on the merozoite surface. These findings highlight the impact of lipid interactions on the conformation and antigenicity of MSP2<sub>172–221</sub> and will assist in the design of recombinant MSP2 immunogens for use as malaria vaccine candidates.

## Databases

Resonance assignments are available in the BioMagResBank (BMRB) database under the accession number 27134.

## Introduction

Malaria is a leading cause of global mortality and morbidity, resulting in more than 200 million clinical episodes and approximately 438,000 deaths annually

worldwide [1]. Among the five species of *Plasmodium* that infect humans, *Plasmodium falciparum* is responsible for the majority of mortality, predominantly in

## Abbreviations

ABTS, 2,2-azino-bis (3-ethylbenzthiazolinesulfonic acid); DOGS, 1,2-di-(9Z-octadecenoyl)-sn-glycero-3-[(N-(5-amino-1-carboxypentyl)iminodiacetic acid) succinyl]; DPC, dodecylphosphocholine; DTT, dithiothreitol; ELISA, enzyme-linked immunosorbent assay; GPI, glycosylphosphatidylinositol; HRP, horseradish peroxidase; HSQC, heteronuclear single quantum coherence; IPTG, isopropyl β-D-1-thiogalactoside; mAb, monoclonal antibody; MSP2, merozoite surface protein 2; NMR, nuclear magnetic resonance; NTA, nitrilotriacetic acid; OD, optical density; PBS, phosphate-buffered saline; POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; RBC, red blood cell; RT, room temperature; SPR, surface plasmon resonance; TEV, tobacco etch virus; Trx, thioredoxin.

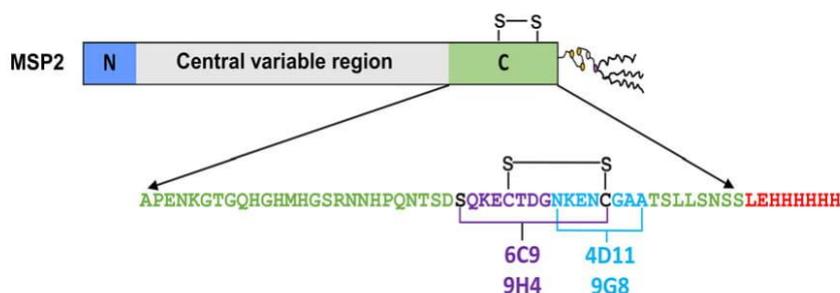
children below five years of age [2]. Vector control and combination drug therapies have seen success in reducing the malaria burden in some regions [3,4], but global eradication of malaria is unlikely to be achievable by these strategies alone, especially in areas of intense transmission. The development of safe and effective vaccines remains an urgent unmet need.

Antigens on the surface of the *P. falciparum* merozoite are attractive candidates for inclusion in an asexual blood-stage malaria vaccine [5]. Merozoite surface protein 2 (MSP2) of *P. falciparum* is one such merozoite antigen, and is a potential component of a malaria vaccine [6,7]. MSP2 is a ~23 kDa protein that is highly abundant and C-terminally glycosylphosphatidylinositol (GPI)-anchored to the surface of the invasive blood stage of the malaria parasite [8,9]. The exact role of MSP2 is not clear but it has been shown to be essential for growth of asexual parasites *in vitro* [6–8]. MSP2 is highly polymorphic, although these polymorphisms are limited to the central variable region of the protein, which is flanked by conserved N- and C-terminal regions (Fig. 1). The patterns of diversity in the central variable region, which evolve under immune pressure from the human host [10,11], have led to all MSP2 alleles being classified into two families, FC27 and 3D7 [12].

Merozoite surface protein 2 is known to be a target of antibodies induced by vaccination or infection [9,13]. In a Phase I-IIb trial in Papua New Guinea, the Combination B vaccine containing the 3D7 allele of MSP2 reduced *P. falciparum* parasite densities by 62% in the vaccinated children [6]. However, the vaccine trial showed protection in a strain-specific manner, with the reduction in parasitaemia being observed only for parasites bearing the 3D7 MSP2 allele while the occurrence of morbidity associated with parasites bearing the FC27 MSP2 allele increased. This strain-specific protection suggested that variable-region

epitopes are immunodominant and these epitopes contribute to vaccine-induced protection [6,14,15]. Antibodies induced by vaccination with MSP2-C1, consisting of both 3D7 and FC27 MSP2, promoted antibody-dependent cellular inhibition (ADCI) of parasite growth [7] and opsonic phagocytosis of merozoites [16]. Anti-MSP2 antibodies were also able to inhibit red blood cell (RBC) invasion via antibody-mediated complement-dependent inhibition through C1q fixation, pointing to another possible mechanism mediating the protective efficacy of an MSP2-based vaccine [17]. Protective antibodies induced in the context of natural infection by *P. falciparum* [18,19] appear to be largely directed towards the central variable region of MSP2, consistent with the strain-specific efficacy seen in the Combination B trial [6].

In contrast, epitopes in the conserved N- and C-terminal regions appear to be the most immunogenic regions of recombinant MSP2 [20], although these regions are less accessible in the parasite antigen [21], implying that the conformation of parasite MSP2 may differ from that of the recombinant protein. However, very little is known about the conformation of MSP2 on the merozoite surface, in particular how it interacts with the merozoite plasma membrane, to which it is GPI anchored. Although MSP2 is intrinsically disordered in solution [22], nuclear magnetic resonance (NMR) studies revealed that both the N- and C-terminal conserved regions of full-length recombinant MSP2 interact with lipid when the protein is C-terminally tethered to mimic GPI anchoring of native MSP2, but the central variable region is minimally affected [23]. In the presence of lipid, the conserved N-terminal region of full-length MSP2 or a construct containing only that region (MSP2<sub>1–25</sub>) adopts a helical conformation [22,24]. This conformational change in recombinant MSP2, induced by lipid interactions, masks the conserved N-terminal epitope of the mouse mAb 6D8



**Fig. 1.** Schematic representation of MSP2 and MSP2<sub>172–221</sub> indicating linear epitopes of different C-terminal region specific mAbs. The N- and C-termini are conserved, while the central variable region differs between two allelic families. The sequence of MSP2<sub>172–221</sub> shows the location of the single, conserved disulfide bond and overlapping linear epitopes of mAbs 4D11, 9G8, 6C9 and 9H4.

[23]. As mAb 6D8 also fails to bind MSP2 on the parasite surface [21], the lipid interactions of recombinant MSP2 may mimic interactions between native MSP2 and the merozoite membrane. Together, these results suggest that lipid interactions may contribute to the masking of conserved epitopes in the N-terminal region, and possibly others in the conserved C-terminal region [21]. These results suggest that lipid-conjugated MSP2 may represent a useful structural and antigenic mimic of native MSP2, and may therefore be worthy of consideration as a suitable formulation for an MSP2-based malaria vaccine. However, analyses of the conformational changes induced by lipid in full-length recombinant MSP2 by NMR were impeded by substantial line broadening [23,24]. Consequently, a shorter MSP2<sub>172–221</sub> construct, consisting of only the conserved C-terminal region of MSP2, was generated to permit NMR analyses of the lipid interactions of this C-terminal region. We have also used this construct and full-length MSP2 to investigate how lipid interactions influence antibody recognition of conserved C-terminal epitopes.

## Results

### Expression and purification of MSP2<sub>172–221</sub>

MSP2<sub>172–221</sub> was cloned into the pET32a vector with a TEV cleavage site engineered between the thioredoxin (Trx)-tag and the N-terminus of the protein. This construct has two extra residues (Leu, Glu) from the vector between the protein and His<sub>6</sub>-tag, giving a total of 58 residues (Fig. 1). The numbering of MSP2<sub>172–221</sub> is based on mature full-length FC27 MSP2. MSP2<sub>172–221</sub> was expressed and purified as a Trx-fusion protein in *Escherichia coli*, with a final yield of approximately 7 mg·L<sup>-1</sup> of bacterial culture. The molecular mass of MSP2<sub>172–221</sub> was 6300 Da as determined by LC-MS, which was close to the theoretical mass of 6298 Da and the purity was greater than 98%. MSP2<sub>172–221</sub> has two cysteines, and the oxidised state of the two cysteines of the lyophilised protein was confirmed by reduction using 100 mM dithiothreitol (DTT), which increased the mass by 2 Da; NMR chemical shifts also confirmed the oxidation state of the cysteines [25].

### Backbone resonance assignments for MSP2<sub>172–221</sub>

A 1D <sup>1</sup>H NMR spectrum (Fig. 2A) and a <sup>1</sup>H-<sup>15</sup>N heteronuclear single quantum coherence (HSQC) spectrum (Fig. 2B) of <sup>15</sup>N-labelled MSP2<sub>172–221</sub> were recorded at 25 °C in 20 mM sodium acetate, pH 4.5, 5% <sup>2</sup>H<sub>2</sub>O. The amide <sup>1</sup>H chemical shifts were

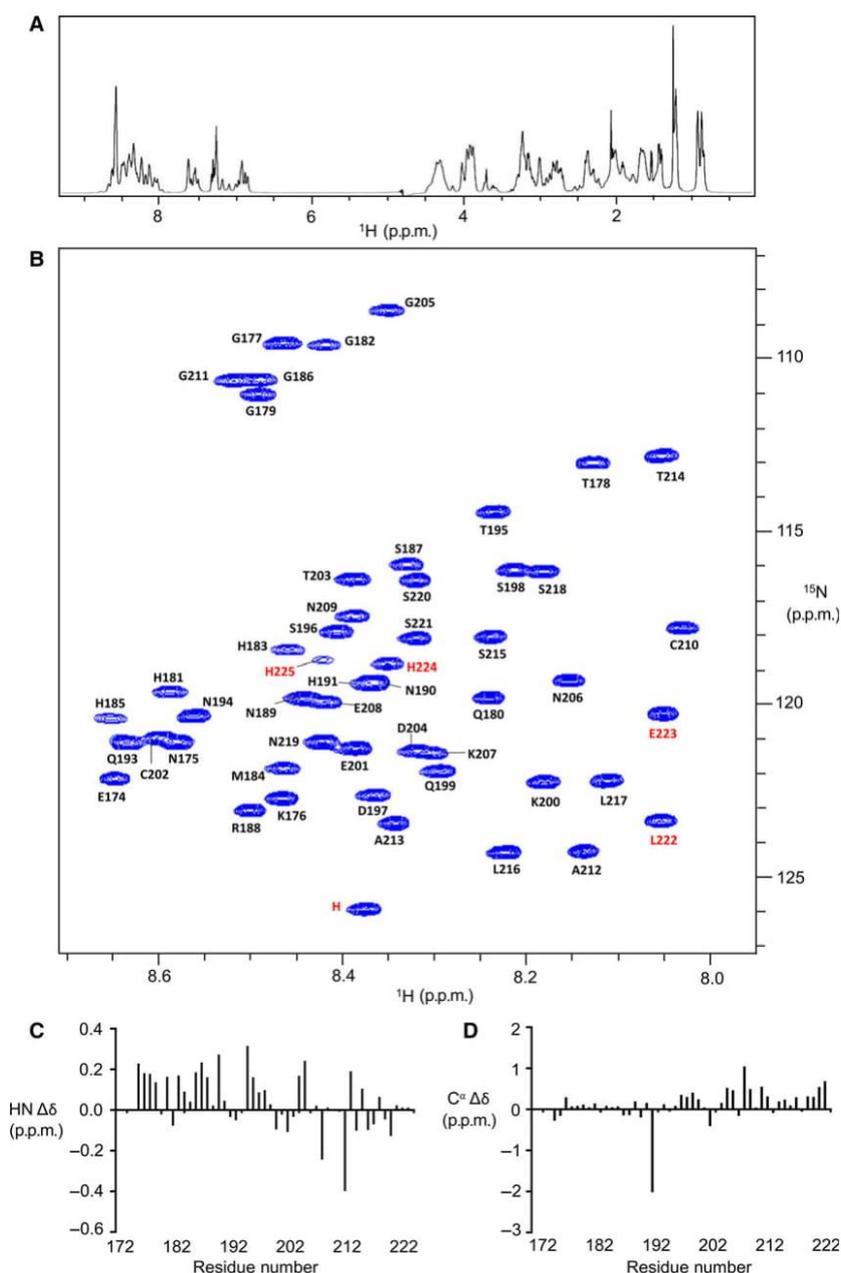
dispersed over a narrow range, 8.0–8.7 ppm, consistent with a disordered and flexible conformation [22].

Except for the N-terminal Ala and four His residues from the His<sub>6</sub>-tag, all non-proline residues were assigned (Fig. 2B) using HNCACB, HNCA, HN(CO)CACB, and HNCO spectra. Deviations of HN and C<sup>α</sup> chemical shifts from random-coil values are plotted in Fig. 2C,D. The chemical shifts were mostly close to random-coil values, as expected. Rather than uniform deviations close to zero, positive deviations were found for most of the HN and C<sup>α</sup> resonances, suggesting a slight propensity for a helical conformation [26]. The large chemical shift deviation of C<sup>α</sup> for His191 is due to the subsequent Pro192 [27]. Around residues Asp197-Ser221 there were some relatively large deviations distributed irregularly, presumably due to conformational restriction caused by the single disulfide bond [22]. Additionally, the pH difference between the random coil shifts (acquired at pH 5.0) [25] and our data (pH 4.5) could have had some effect.

The backbone amide regions of <sup>1</sup>H-<sup>15</sup>N HSQC spectra of full-length FC27 MSP2 and MSP2<sub>172–221</sub> are superimposed in Fig. 3A. Most peaks of MSP2<sub>172–221</sub> have chemical shifts similar to those of the corresponding residues in the full-length protein, with significant changes observed only for the N- and C-termini (Fig. 3B). These changes reflect the fact that the two constructs differ at both termini, with the full-length construct lacking the His<sub>6</sub>-tag present in MSP2<sub>172–221</sub>. Negligible deviations were observed for other residues. These results imply that the C-terminal region is not involved in any stable interactions with other regions of full-length MSP2, consistent with our previous inference from a comparison of chemical shifts and relaxation rates of FC27 and 3D7 MSP2 [20].

### Interaction with lipid micelles

Since native MSP2 is GPI-anchored to the merozoite membrane through its C-terminus, this region could interact with the merozoite surface membrane. <sup>2</sup>H<sub>38</sub>-dodecylphosphocholine (DPC) was used as a membrane mimetic because it is a useful model membrane system for solution NMR, and both the RBC and parasite membrane are rich in phosphatidylcholine-containing lipid [28,29]. A comparison of <sup>1</sup>H-<sup>15</sup>N HSQC spectra of MSP2<sub>172–221</sub> in the presence and absence of DPC micelles containing nickel-chelating lipid 1,2-di-(9Z-octadecenoyl)-*sn*-glycero-3-[(N-(5-amino-1-carboxypentyl) iminodiacetic acid)succinyl] (DOGS-NTA) and Ni<sup>+2</sup> revealed substantial changes in both line-widths and chemical shifts, indicative of interactions

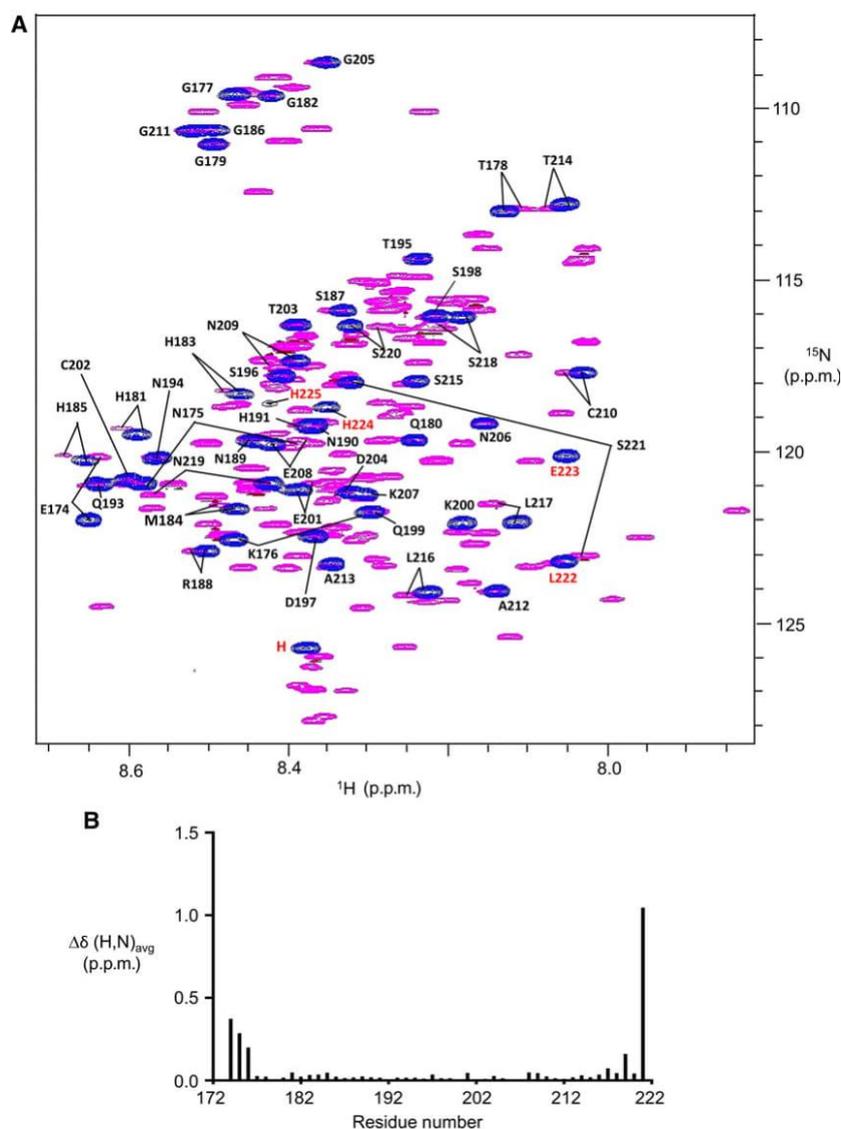


**Fig. 2.**  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum and secondary chemical shifts of MSP2<sub>172–221</sub>. (A)  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum (600 MHz) of MSP2<sub>172–221</sub> in 20 mM sodium acetate, pH 4.5 at 25 °C. Peaks are labelled with residue numbers according to mature full-length FC27 MSP2. Extra residues coming from vector (L222, E223) or His<sub>6</sub>-tag (H224, H225) are labelled red and unassigned His residues from the His<sub>6</sub>-tag are labelled as H. Secondary chemical shifts HN (B) and C $^\alpha$  (C) along the sequence of MSP2<sub>172–221</sub> in 20 mM sodium acetate, pH 4.5. Random coil chemical shifts are from Wishart *et al.* [25].

between MSP2<sub>172–221</sub> and the DPC micelle.  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of MSP2<sub>172–221</sub> in the presence and absence of DPC at pH 4.5 are superimposed in Fig. 4A. Using HNCACB and HNCOCA spectra of MSP2<sub>172–221</sub> in the presence of DPC, assignments were made for all non-proline residues except the N-terminal Ala and four His residues from the His<sub>6</sub>-tag. In the presence of DPC, many peaks corresponding to residues 176–221, as well as several unassigned side chain peaks were shifted (Fig. 4A,B), suggesting that lipid interactions result in significant perturbations to

the conformation or environment of residues throughout the C-terminal region of MSP2.

The differences in backbone amide and C $^\alpha$  chemical shifts between MSP2<sub>172–221</sub> in the absence and presence of DPC are plotted in Fig. 4C,D. The largest changes were observed near the MSP2 C-terminus (Ala212 to Ser221), but significant amide and C $^\alpha$  chemical shift perturbations were observed throughout MSP2<sub>172–221</sub>. C $^\alpha$  chemical shift changes associated with lipid interaction were predominantly positive, but were smaller than observed previously for the N-terminus [24], and



**Fig. 3.** Superposition of backbone amide regions of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra (600 MHz) of full-length FC27 MSP2 (pink) and MSP2<sub>172-221</sub> (blue) at 25 °C, pH 4.5. MSP2<sub>172-221</sub> was prepared using 20 mM sodium acetate, pH: 4.5 and FC27 MSP2 was in water with the pH was adjusted to 4.5 using 1 M acetic acid. (A) Noticeable shifted peaks are labelled and shown with lines; additional peaks of FC27 MSP2 (pink) are from the conserved N-terminus and central variable region. Extra residues of MSP2<sub>172-221</sub> (blue) coming from vector (L222, E223) or His<sub>6</sub>-tag (H224, H225) are labelled red and unassigned residues from the His<sub>6</sub>-tag are labelled as H. (B) Backbone amide chemical shift differences ( $\Delta\delta = [(\Delta\delta_{\text{N}}/5)^2 + (\Delta\delta_{\text{HN}})^2]^{1/2}$ ) between MSP2<sub>172-221</sub> and corresponding residues of full-length FC27 MSP2.

not consistent with the formation of a stable helical structure. Nonetheless, the perturbed residues of MSP2<sub>172-221</sub> include epitopes recognised by C-terminal-specific mAbs, and are consistent with the possibility of lipid-induced changes to the conformation of these epitopes. We then investigated whether these interactions affected the antigenicity of both MSP2<sub>172-221</sub> and full-length MSP2.

### MSP2<sub>172-221</sub> is antigenically similar to full-length MSP2

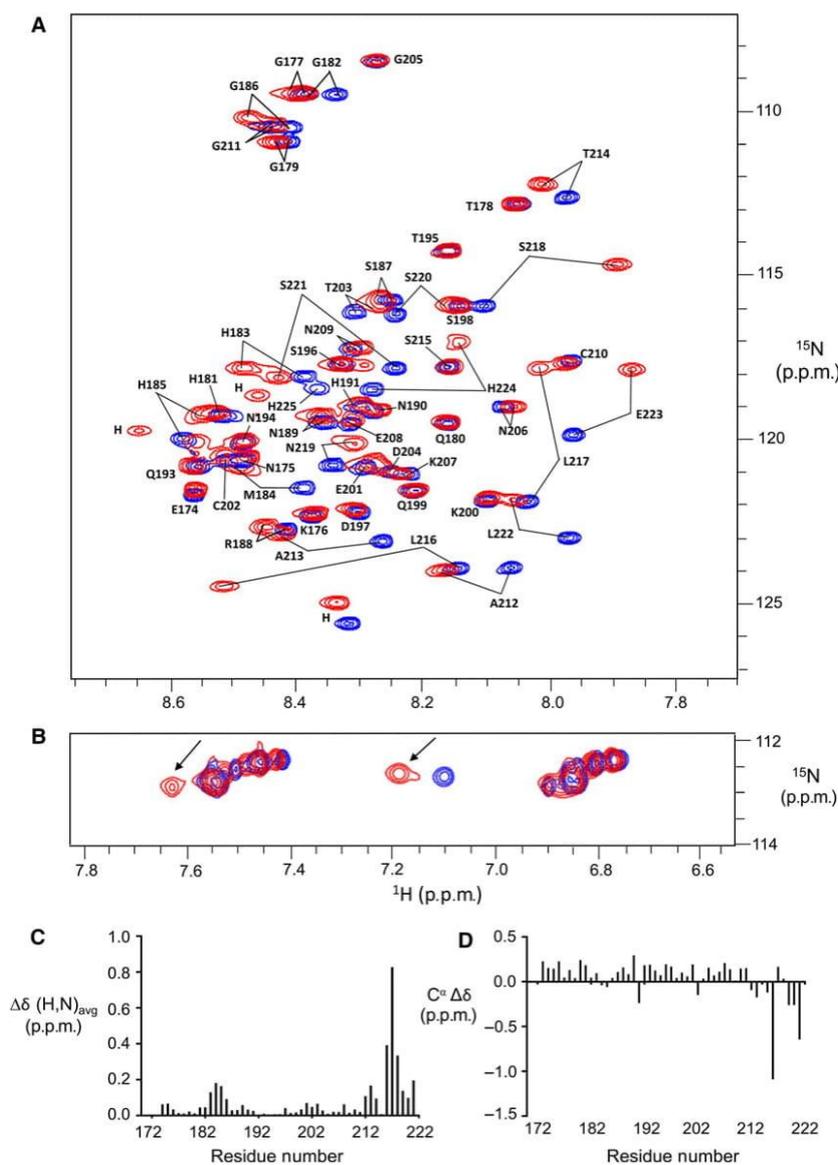
Competitive enzyme-linked immunosorbent assays (ELISA) experiments were used to assess the relative capacity of the three different forms of recombinant

MSP2 antigens (3D7 MSP2, FC27 MSP2 and MSP2<sub>172-221</sub>) to bind to four mAbs recognising conserved C-terminal region epitopes. All three antigens effectively blocked the binding of all four mAbs to immobilised 3D7 MSP2 in a concentration-dependent manner (Fig. 5). MSP2<sub>172-221</sub> and 3D7 MSP2 showed identical concentration dependence in their ability to block mAb binding, consistent with the NMR data indicating that the C-terminal region is not involved in any stable interactions with other regions of full-length MSP2. In contrast, FC27 MSP2 was a more effective inhibitor of the binding of all four mAbs, an effect that was slight for 4D11 and 9G8 (Fig. 5A,B) but more marked for 6C9 and 9H4 (Fig. 5C,D). Similar results were obtained using an MSP2<sub>172-221</sub>-coated plate (data not shown).

### Liposome conjugation alters antigenicity of MSP2<sub>172-221</sub> and full-length MSP2

Previously we have shown that the conserved N-terminal domain of recombinant MSP2 adopts a helical conformation in the presence of lipid that precludes the binding of N-terminal specific mAb 6D8 [23]. To investigate the relationship between lipid interaction and antibody recognition for the conserved C-terminal domain, we developed an ELISA approach in which biotinylated 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) liposomes were immobilised on a streptavidin-coated plate, and C-terminally His<sub>6</sub>-tagged antigens were conjugated to liposomes via lipid molecules containing the nitrilotriacetic acid (NTA) head

group moiety (DOGS-NTA) and Ni<sup>2+</sup>. In this way, we mimicked the C-terminal GPI anchoring of MSP2 on the merozoite surface. Antibody binding to these lipid-anchored antigens was then assessed by direct ELISA. In the case of mAbs 4D11 and 9G8, there were similar trends in relative apparent antibody affinities as were seen in the absence of liposomes (Fig. 6A, B). Specifically, these antibodies bind FC27 MSP2 with slightly higher apparent affinity than MSP2<sub>172-221</sub> and 3D7 MSP2. In contrast, mAbs 6C9 and 9H4 showed strikingly different binding for the three lipid-bound antigens, with antibody binding to 3D7 MSP2 being very much weaker than that to FC27 MSP2, and binding of MSP2<sub>172-221</sub> being almost undetectable (Fig. 6C,D). This clearly demonstrates that the



**Fig. 4.** Superposition of <sup>1</sup>H-<sup>15</sup>N HSQC spectra (600 MHz) of MSP2<sub>172-221</sub> at pH 4.5 in the absence (blue) and presence (red) of 100 mM DPC + 1 mM DOGS-NTA + 0.2 mM Ni<sup>2+</sup> at 25 °C. (A) Largely shifted peaks are labelled and shown with lines. Extra residues coming from the vector (L222, E223) and two from His<sub>6</sub>-tag (H224, H225) are also assigned and labelled, and unassigned His residues from His<sub>6</sub>-tag are labelled as H. (B) The spectral region containing resonances from side chains of Asn, Gln and Arg; the arrows highlight side chain peaks that are shifted in the presence of DPC. Backbone amide chemical shift differences (C) ( $\Delta\delta = [(\Delta\delta_{N/5})^2 + (\Delta\delta_{HN})^2]^{1/2}$ ) and C<sup>α</sup> differences (D) of MSP2<sub>172-221</sub> in the presence of DPC.

recognition of these antigens is strongly affected when the antigens are C-terminally tethered to lipid.

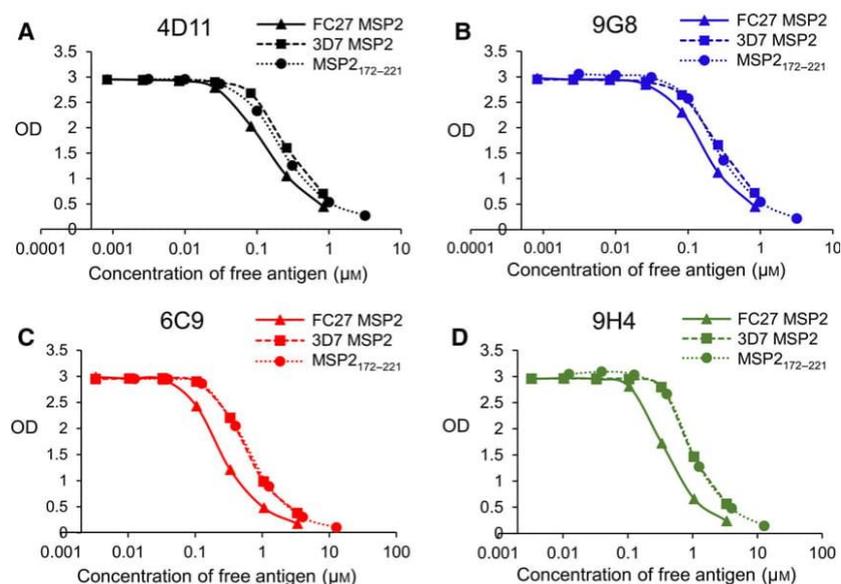
Among the C-terminal conserved-region specific mAbs investigated here, both 4D11 and 9G8 show strong reactivity to parasite MSP2, whereas 6C9 and 9H4, which recognise overlapping epitopes, react only weakly with native MSP2, suggesting that their epitopes are less accessible on the merozoite surface [21]. Our ELISA results show that 4D11 and 9G8 have equivalently strong binding to all three lipid-bound antigens whereas the 6C9 and 9H4 epitopes are markedly less accessible in lipid-bound MSP2<sub>172–221</sub> and 3D7 MSP2 than in lipid-bound FC27 MSP2. These results indicate that lipid interactions can mask epitopes in both the C-terminal and N-terminal conserved regions of recombinant MSP2, albeit in an antigen-dependent manner.

### Alteration of antigenicity of liposome conjugated MSP2<sub>172–221</sub> and full-length MSP2

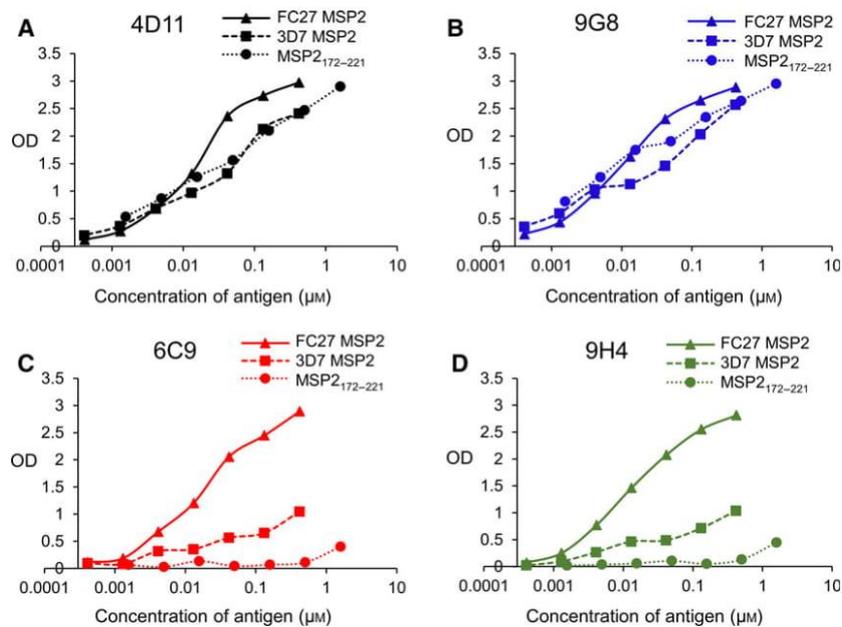
To further examine the role of lipid interactions in masking conserved C-terminal epitopes, we investigated the binding of C-terminal specific mAbs with MSP2<sub>172–221</sub> and both alleles of full-length MSP2 in the absence (Fig. 7A–C) and presence (Fig. 7D–F) of POPC liposomes by surface plasmon resonance (SPR) [23]. For binding studies in the presence of lipid, POPC liposomes containing nickel-chelating lipid, DOGS-NTA, were immobilised on an L1 chip and loaded with Ni<sup>2+</sup>, which bound the C-terminal His<sub>6</sub>-tag on the antigen, whereas antigen in the absence of lipid was immobilised on a CM5 chip. Since it was

difficult to interpret the sensorgram in terms of  $K_D$  values because of slow antibody dissociation, we considered the response of 4D11 as 100% at the highest concentration of antibody and determined responses of other mAbs relative to 4D11 (Fig. 8). In the absence of liposomes, the relative binding responses of 9G8 were 42%, 50% and 44% for FC27 MSP2, 3D7 MSP2 and MSP2<sub>172–221</sub>, respectively, showing limited variation across the three antigens (Fig. 8). Similarly, the relative binding responses of 6C9 for three different antigens were within a very narrow range (14–15%), whereas 9H4 showed relative binding responses of 30%, 17% and 14% for FC27 MSP2, 3D7 MSP2 and MSP2<sub>172–221</sub> respectively. Therefore, in the absence of liposomes, most of the mAbs had similar binding responses for the three antigens, indicating very similar antigenicities, except for 9H4, which showed nearly twice the response for FC27 MSP2 as for 3D7 MSP2 or MSP2<sub>172–221</sub>. This pattern of relative response was consistent with the relative antibody affinities inferred from the competition ELISA experiments. By way of comparison, we have recently determined the affinity of 4D11 IgG for lipid-free MSP2 to be 0.9  $\mu\text{M}$ , using an SPR competition assay [30]. Following the same protocol, the  $K_D$  values of 9G8, 6C9 and 9H4 for lipid-free MSP2 are 0.019, 2.0 and 2.0  $\mu\text{M}$  respectively.

We then investigated how lipid interactions affected the relative binding of these mAbs. Intriguingly, the most strongly affected mAb in terms of response relative to 4D11 was 9G8, where the relative response to all three lipid-bound antigens was approximately twice as great compared to the liposome-free experiment (Fig. 8). The relative responses of 6C9 and 9H4 were



**Fig. 5.** Competition ELISA test to characterise antigenicity using C-terminal specific mAbs. The plate is coated with 3D7 MSP2 and the binding pattern of 4D11 (A), 9G8 (B), 6C9 (C) and 9H4 (D) are shown for MSP2<sub>172–221</sub> (···) or 3D7 MSP2 (---) or FC27 MSP2 (—) in solution as competing antigen. For 6C9 and 9H4, competing antigen concentrations were four times more than with 4D11 and 9G8. The optimal concentrations of 4D11 and 9G8 were 0.20  $\mu\text{M}$  and for 6C9 and 9H4 0.33  $\mu\text{M}$ . The optical density (OD) was measured at 414 nm.



**Fig. 6.** Reactivity of C-terminal region specific mAbs with three different antigens attached to POPC liposomes as measured by ELISA. The reactivity of 4D11 (A), 9G8 (B), 6C9 (C) and 9H4 (D) for MSP2<sub>172-221</sub> (---●---), 3D7 MSP2 (---■---) or FC27 MSP2 (---▲---) are shown. The optimal constant concentration of mAbs was 2  $\mu\text{g}\cdot\text{mL}^{-1}$  (13.2 nM). The OD was measured at 414 nm.

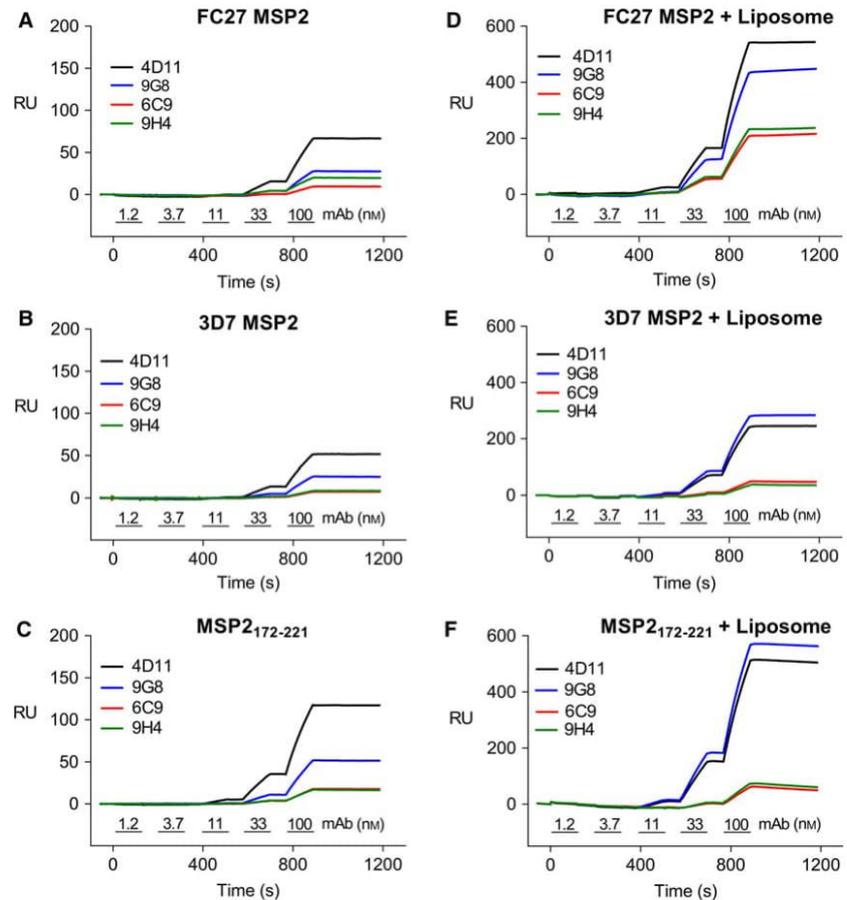
enhanced for FC27 MSP2, but only weakly or not at all for 3D7 MSP2 and MSP2<sub>172-221</sub>. Thus, in the presence of liposomes, 9G8 shows a strong relative response while 6C9 and 9H4 show weak relative responses with recombinant antigens, consistent with both our ELISA results and the reactivity pattern of these mAbs with the parasite antigen [21]. These results provide additional evidence that lipid interactions alter the antigenicity of recombinant MSP2 in a way that mimics native MSP2 on the merozoite surface.

## Discussion

Merozoite surface protein 2 is a vaccine candidate targeting the asexual blood-stages of the malaria parasite *P. falciparum*. As a component of the Combination B vaccine, recombinant MSP2 induced an immune response that protected against *P. falciparum* in a strain-specific manner. To better understand the basis of this strain specificity, and ultimately to develop strain-transcendent immunogens based on MSP2, it is necessary to understand the conformational and antigenic properties of the native protein on the parasite surface.

Tethering the conserved C-terminal domain of MSP2 to micelles using a Ni-chelating lipid caused substantial changes in its NMR spectra, although these changes were not consistent with the formation of ordered secondary structure. This is in contrast to previous observations for the conserved N-terminal

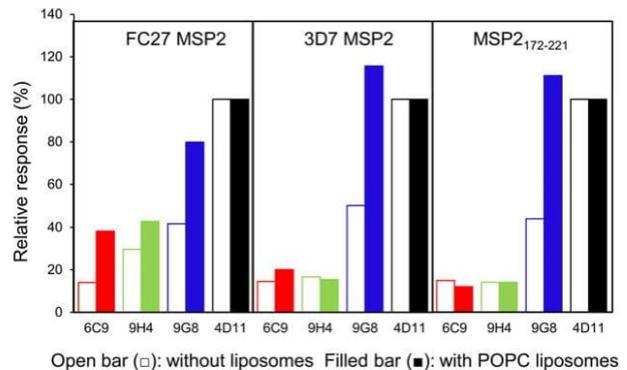
domain, which adopts a helical structure upon lipid binding [23]. Nonetheless, the altered antigenicity of liposome-conjugated MSP2<sub>172-221</sub> and full-length MSP2, as detected by ELISA (Fig. 6C,D), clearly indicates that antibody recognition of these antigens is strongly affected when they are C-terminally tethered to lipid. The epitopes recognised by 6C9 and 9H4 were partially masked for 3D7 MSP2 and almost completely masked for MSP2<sub>172-221</sub> (Fig. 6C,D). It also appears that in the presence of liposomes the conformation of the 9G8 epitope necessary for mAb binding is favoured, as indicated by the increase in relative response (Fig. 8). These results in the presence of liposomes are consistent with the pattern of binding of these mAbs to the parasite antigen; 4D11 and 9G8 bind strongly to the parasite antigen while 6C9 and 9H4 bind weakly, if at all [21]. The selective masking of the 6C9 and 9H4 epitopes is all the more remarkable because the 4D11 and 9G8 epitopes are closer to the lipid-tethered C-terminus and for this reason may have been expected to be more strongly influenced by lipid binding. The experimental model used here for lipid tethering is clearly different from the GPI anchoring of native MSP2, but appears to be a very effective model for studying antigenically relevant conformations of recombinant and parasite MSP2. The GPI moiety has been reported to influence the conformation and antigenicity of other GPI-anchored proteins, for example, the trypanosome variant surface glycoprotein [31] and Thy-1 protein [32]. It remains to be established to what extent the antigenic characteristics



**Fig. 7.** Surface plasmon resonance sensorgrams demonstrating the impact of lipid interactions on the antigenicity of FC27 MSP2 (A,D), 3D7 MSP2 (B,E) and MSP2<sub>172-221</sub> (C,F). For binding studies in the absence of lipid (A–C), the three different antigens were immobilised to a CM5 sensor chip, while for binding studies in the presence of lipid (D–F), POPC vesicles containing DOGS-NTA were immobilised to an L1 chip and loaded with Ni<sup>2+</sup> and C-terminally His<sub>6</sub>-tagged antigens. Each of four mAbs, 4D11 (black), 9G8 (blue), 6C9 (red) and 9H4 (green), was analysed in single-cycle mode, with five successive 120 s injections, spanning concentrations from 1.2 nM to 100 nM as indicated below the curves.

of parasite MSP2 reflect effects of the C-terminal GPI anchor and/or interactions with the merozoite plasma membrane or other merozoite coat components.

The different mAb binding properties of liposome-conjugated MSP2<sub>172-221</sub> and 3D7 MSP2 compared with FC27 MSP2 suggest that residues in the central variable region preceding the conserved C-terminal region may influence antigenicity. Notwithstanding this, our NMR data suggest that, in the absence of lipid, the conformation of the C-terminal region is essentially identical in MSP2<sub>172-221</sub>, 3D7 and FC27 MSP2. It is possible that the ways in which the three constructs interact with lipid differ, and that this contributes to the antigenic differences; this is not readily assessable due to the extensive line-broadening observed for full-length MSP2 in the presence of lipid. However, the three antigens also differ, albeit more subtly, in the absence of lipid, suggesting that antigenic differences here may be driven by other effects. Previously we have shown that mAb 6D8 shows strain-specific affinity for MSP2 even though it recognises a conserved epitope [23]. Just a few residues of the central variable region adjacent to the conserved N-terminus make transient



**Fig. 8.** Binding response of different mAbs for three different antigens calculated from SPR experiment (Fig. 7). The relative binding responses of 6C9, 9H4 and 9G8 are calculated in proportion to 4D11 for each antigen, where the open bars (□) represent absence and filled bars (■) the presence of POPC liposomes.

interactions with regions of 6D8 outside the structurally defined paratope, but these interactions are sufficient to induce this strain specificity [23]. In a similar manner the central variable region residues immediately preceding the conserved C-terminal region may

influence the binding of C-terminal-specific antibodies without altering the conformation of the epitope. Whatever the mechanism, this is most pronounced with FC27 MSP2, which shows higher affinity towards 6C9 and 9H4 in comparison to 3D7 MSP2 and MSP2<sub>172–221</sub>, both in the presence and absence of liposomes (Figs 5C,D and 6C,D). The similar antigenicity of 3D7 MSP2 and MSP2<sub>172–221</sub> suggests that the variable region of 3D7 MSP2 may not influence antibody binding in the absence of liposomes to the same extent as the variable region of FC27 MSP2 (Fig. 5).

In summary, we have shown that lipid interactions influence the antigenicity of C-terminally lipid-tethered MSP2<sub>172–221</sub> or full-length MSP2. Importantly, it appears that the pattern of reactivity of mAbs with recombinant antigens in the presence of liposomes is similar to their reactivity with the parasite antigen [21]. The method by which antigen is conjugated to lipid is potentially important, and may exert dramatic effects on the resulting immune response [33,34], but the non-covalent tethering utilised here appears to be satisfactory mimic of native MSP2 on the parasite surface. Our results support the further exploration of lipid-conjugated MSP2 as a potential malaria vaccine candidate. Because lipid-tethering induces an antigenic state that is more similar to the parasite antigen, we expect that a lipid-based vaccine formulation may induce a more effective protective immune response than that of soluble MSP2. Clinical trials of liposome-formulated vaccines have been reported [33,35–39] and there is increased interest in liposome-conjugated vaccine technology. Liposome tethering not only serves as an adjuvant but also offers the prospect of enhancing the immune response to conserved regions of MSP2 that are exposed on the parasite surface, thus favouring a strain-transcending protective response.

## Materials and methods

### Cloning of His-tagged conserved C-terminal domain

Full-length MSP2 was used as a template for amplification of the C-terminal domain of MSP2. Two sets of primers were designed where first round primers, excluding the restriction sites, were used to amplify only the C-terminal residues, and *KpnI* and *XhoI* restriction sites were incorporated at the 5' terminus of the second round forward and reverse primers, respectively, required for downstream ligation into the bacterial expression vector pET32a (Novagen, San Diego, CA, USA). The second round forward primer also has additional tobacco etch virus (TEV) cleavage site (ENLYFQ). The primers were as follow:

1st Round Primers:

Forward: 5'-GCACCAGAGAATAAAGGTACAG-3'

Reverse: 5'-ACTAGAGTTACTTAAGAGGGGATGTTGC-3'

2nd Round Primers:

Forward: 5'-TCTAGTGGTACCGAAAATTTATATTT  
TCAAGCACCAGAGAATAAAGGTACAG-3'

Reverse: 5'-TCTAGTCTCGAGACTAGAGTTACTTA  
AGAGGGGATG-3'

Amplification was carried out using Platinum<sup>®</sup> PCR SuperMix (Invitrogen, Carlsbad, CA, USA), where full-length FC27 MSP2 was used as template with first round primers. The amplified DNA fragment was purified and used as a template for the second round of amplification. Briefly, the amplification cycle for MSP2<sub>172–221</sub> genes consisted of a hot start of 94 °C for 5 min, followed by 35 cycles consisting of 94 °C for 30 s, annealing at 55 °C for 30 s, and final elongation at 72 °C for 30 s. The amplification cycles were performed using a thermal cycler (Techne, TC-3000, Bibby Scientific Ltd, Stone, Staffordshire, UK). The purified fragment was digested and ligated into *KpnI* and *XhoI* sites of pET32a vector for expression as Trx-fusion proteins, which provided six His residues at the C-terminus of the expressed protein. Following ligation by T4 DNA ligase, the ligated vector was transformed into *E. coli* XL10-gold cells and selected against ampicillin antibiotic. The construct was subsequently verified by DNA sequencing.

### Expression and purification of MSP2<sub>172–221</sub>

The plasmid was transformed into *E. coli* BL21 (DE3) cells and protein was expressed and purified as described previously [20,22]. Expression was induced by adding 1 mM isopropyl β-D-1-thiogalactoside (IPTG) to the cultures when OD<sub>600</sub> ≈ 0.8 and was carried out at 37 °C for 6 h with 150 rpm. The cells were then harvested by centrifugation at 6000 g for 10 min at 4 °C. The bacterial cell pellet was resuspended in resuspension buffer (20 mM Tris, pH 8.0, 200 mM NaCl) and cells were lysed by boiling at 100 °C for 15 min with gentle mixing in between. The cells were sonicated three times for 30 s and the cell lysate was centrifuged at 48 500 g for 25 min at 4 °C. The supernatant was filtered, isolated on His-trap affinity chromatography column (HisTrap 5 mL, GE Healthcare, Piscataway, NJ, USA) and eluted with imidazole. The Trx-fusion protein was cleaved at room temperature (RT) for 6 h with 1% (w/w) TEV protease in the presence of 0.3 mM oxidised and 3 mM reduced glutathione respectively. After cleavage reaction, the desired protein was separated from Trx-tag using size exclusion chromatography. Final purification of His<sub>6</sub>-tagged MSP2<sub>172–221</sub> was performed using a C<sub>18</sub> column (Phenomenex, Macclesfield, Cheshire, UK) with a 5–60% linear gradient of 80% acetonitrile in 0.1% TFA and finally the purified product was lyophilised. <sup>15</sup>N-labelled and <sup>13</sup>C-<sup>15</sup>N-labelled MSP2<sub>172–221</sub> were prepared in M9 minimal medium with 1 g·L<sup>-1</sup> <sup>15</sup>NH<sub>4</sub>Cl or/and 4 g·L<sup>-1</sup> <sup>13</sup>C glucose

(Sigma-Aldrich, St. Louis, MO, USA) for sole nitrogen and carbon sources respectively. Expression and purification conditions were as described for the unlabelled protein. The final recombinant MSP2<sub>172–221</sub> has two extra residues at C-terminal from the vector immediately preceding the His<sub>6</sub>-tag.

### NMR spectroscopy

The sample used for recording <sup>1</sup>H-<sup>15</sup>N HSQC spectra was prepared by dissolving lyophilised uniformly <sup>15</sup>N-labelled recombinant MSP2<sub>172–221</sub> in 20 mM sodium acetate, pH 4.5, containing 5% <sup>2</sup>H<sub>2</sub>O at a concentration of 100 μM. <sup>13</sup>C-<sup>15</sup>N-labelled MSP2<sub>172–221</sub> at a concentration of 1.25 mM was used for 3D heteronuclear triple-resonance experiments (HNCACB, HN(CO)CACB and HNCA) to make resonance assignments. Two-dimensional <sup>1</sup>H-<sup>15</sup>N HSQC spectra were acquired with 2048 (<sup>1</sup>H) and 256 (<sup>15</sup>N) points and spectral widths of 10 ppm (<sup>1</sup>H) and 28 ppm (<sup>15</sup>N), respectively. In all 3D spectra, the spectral widths along the <sup>1</sup>H and <sup>15</sup>N dimensions were the same as in 2D <sup>1</sup>H-<sup>15</sup>N HSQC spectra, and the number of points in *F3* for most experiments was 2048 acquired using nonuniform sampling. In the absence of lipid, HNCACB and HN(CO)CACB spectra were recorded with 32 scans, <sup>13</sup>C spectral width of 75 ppm, and 40 (<sup>15</sup>N) and 128 (<sup>13</sup>C) data points. An HNCA spectrum was recorded with 32 scans, <sup>13</sup>C spectral width of 32 ppm, and 40 (<sup>15</sup>N) and 128 (<sup>13</sup>C) data points. An HNCO spectrum was recorded with <sup>13</sup>C spectral width of 12 ppm, and 40 (<sup>15</sup>N) and 128 (<sup>13</sup>C) data points.

For interaction studies with DPC, <sup>15</sup>N-labelled MSP2<sub>172–221</sub> was dissolved in 20 mM sodium acetate at a concentration of 100 μM, pH 4.5, containing 100 mM DPC, 1 mM DOGS-NTA and 0.2 mM NiSO<sub>4</sub>. To prepare the sample, 19.5 mg of <sup>2</sup>H<sub>38</sub>-DPC and 0.50 mg of the nickel-chelating lipid DOGS-NTA were dissolved in a methanol and chloroform mixture (1 : 1) with gentle shaking. The solvent was evaporated under N<sub>2</sub> and then dried overnight under vacuum. 500 μL of 100 μM of MSP2<sub>172–221</sub> in 20 mM sodium acetate, pH 4.5, was used to dissolve DPC-DOGS-NTA and the foamy sample was centrifuged for 30 s at 20 000 *g* before transferring to an NMR tube. NiSO<sub>4</sub> was added to a final concentration of 0.2 mM. For acquisition of HNCACB and HNCOCA spectra to make resonance assignments, the sample was prepared in the same way and the concentrations of <sup>13</sup>C-<sup>15</sup>N-labelled MSP2<sub>172–221</sub>, DPC, DOGS-NTA and Ni<sup>+2</sup> were 1.25 mM, 200 mM, 2 mM and 1.5 mM, respectively, in 20 mM sodium acetate, pH 4.5. In the presence of lipid, an HN(CO)CA spectrum was recorded with 64 (<sup>15</sup>N) and 128 (<sup>13</sup>C) data points and a <sup>13</sup>C spectral width of 32 ppm. The HNCACB spectrum was recorded with 40 (<sup>15</sup>N) and 128 (<sup>13</sup>C) data points and a <sup>13</sup>C spectral width of 75 ppm. HN(CO)CA and HNCACB spectra were acquired with 16 and 64 scans, respectively. All spectra were recorded at 25 °C on a 600 MHz Bruker

Avance III spectrometer, processed using TOPSPIN (version 2.3, Bruker Biospin), and analysed using CCPNMR analysis (version 2.4.2) [40].

### Enzyme-linked immunosorbent assays

#### Competition ELISA

Four different solutions (coating, blocking, dilution and washing) were prepared and used for ELISA experiments. The coating solution was phosphate-buffered saline (PBS) that contained 2.7 M NaCl, 54 mM KCl and 220 mM phosphate buffer, pH 7.4 (AMRESCO LLC, USA). The blocking solution was 1% bovine serum albumin (BSA) (w/v) and 0.05% (v/v) Tween 20 in PBS. The dilution solution containing 0.1% BSA (w/v) and 0.05% Tween 20 (v/v) in PBS was maintained at RT and the same composition was used as washing solution. Titration curves were prepared for each of the mAbs by plotting the absorbance values as a function of the concentration of serially diluted mAb (data not shown). The concentration of each of the mAbs for the competition ELISA was chosen to give absorbance values in the linear portion of the curve. For the titration curve, wells of three different microtiter plates (Nunc-Immuno C96 Maxisorp, Thermo Fisher Scientific, Kamstrupvej, Roskilde, Denmark) were coated with 2 μg·mL<sup>-1</sup> of MSP2<sub>172–221</sub>, 3D7 MSP2 and FC27 MSP2, respectively, overnight in a moist box at 4 °C. To initiate the assay, plates were washed with washing solution and Milli-Q water three times and nonspecific binding sites were blocked using blocking solution for 1 h with gentle shaking at RT. Each of the C-terminal region specific mAbs, as primary antibodies, was added to antigen-coated wells (100 μL per well) at 6.6 nM in duplicate and serially diluted at half log<sub>10</sub> dilution with dilution buffer except for the blank wells. After 1 h incubation with gentle shaking at RT and washing, the plates were coated with 100 μL per well of goat anti-mouse IgG antibody (1 : 2000 dilution) and incubated for 1 h. After similar washing, the plates were coated with 150 μL per well of freshly prepared 2,2-azinobis(3-ethylbenzthiazolinesulfonic acid (ABTS) substrate (1 mM) in citric acid buffer (pH 4.2) containing 0.03% H<sub>2</sub>O<sub>2</sub> for horseradish peroxidase (HRP) and incubated 45 m with gentle shaking. The absorbance was read at 414 nm using a microplate reader.

For competition ELISA, the polystyrene wells of two different microtiter plates were coated with 2 μg·mL<sup>-1</sup> (100 μL per well) of MSP2<sub>172–221</sub> and full-length 3D7 MSP2, respectively, overnight in a moist box at 4 °C. The plates were washed with washing solution and Milli-Q water (Millipore SAS, Molsheim, France) three times to remove unbound protein. Subsequently, blocking solution (100 μL per well) was added to the wells to prevent non-specific binding, which was followed by 1 h incubation with

gentle shaking at RT. Antigen (MSP2<sub>172–221</sub> or 3D7 MSP2) of varying concentrations diluted in dilution buffer was added and immediately 100  $\mu\text{L}$  per well of mAbs (0.5  $\mu\text{g}\cdot\text{mL}^{-1}$ ) were loaded onto the plate in duplicate at a concentration chosen from titration curve (4D11/9G8: 0.20 nM and 6C9/9H4: 0.33 nM) and incubated for 1 h at RT with gentle shaking. After washing, plates were coated with 100  $\mu\text{L}$  per well of HRP-conjugated goat anti-mouse IgG antibody (1 : 2000 dilution) and subsequently freshly prepared 150  $\mu\text{L}$  per well of ABTS substrate (1 mM) in citric acid buffer (pH 4.2) containing 0.03%  $\text{H}_2\text{O}_2$  for HRP with 1 h incubation and washing in between. The optical density (OD) at 414 nm was measured 60 min after addition of substrate. The mean of duplicate wells was calculated and background was subtracted for each sample.

### ELISA in presence of liposome conjugated to antigen

For ELISA studies of liposome-conjugated antigens, three microtiter plates were coated with 100  $\mu\text{L}$  per well of streptavidin (5  $\mu\text{g}\cdot\text{mL}^{-1}$ ) (New England Biolabs, Ipswich, MA, USA) in PBS and incubated overnight at 4 °C. After washing three times with PBS and Milli-Q water, respectively, the plates were coated with 100  $\mu\text{L}$  per well of 0.6 mM biotinylated-POPC liposome containing DOGS-NTA (1 mol %) diluted using 0.1% BSA in PBS. The plates were incubated for 1 h with gentle shaking at RT followed by similar washing. 1 mM  $\text{NiSO}_4$  was added to each well of microtiter plate and incubated for 15 min with gentle shaking, followed by washing four times thoroughly only with Milli-Q water to prevent precipitation with PBS. Antigens were added to the wells at a concentration of 10  $\mu\text{g}\cdot\text{mL}^{-1}$  and serially diluted using a half  $\log_{10}$  titration with 0.1% BSA in PBS as diluent. MAbs were added to the duplicate wells (100  $\mu\text{L}$  per well) at a concentration of 2  $\mu\text{g}\cdot\text{mL}^{-1}$  (13.2 nM) in PBS containing 0.1% BSA. After washing with PBS and Milli-Q water, the plates were incubated 1 h with 100  $\mu\text{L}$  per well of the goat anti-Mouse IgG antibody (1 : 2000 dilution). After similar washing, 150  $\mu\text{L}$  per well of freshly prepared ABTS substrate (1 mM) in citric acid buffer (pH 4.2) containing 0.03%  $\text{H}_2\text{O}_2$  was added and plates were incubated with gentle shaking at RT, and OD was measured at 414 nm after 45 min.

### SPR study in absence and presence of liposomes

The binding affinities of different C-terminal region specific mAbs (4D11, 9G8, 6C9 and 9H4) for MSP2<sub>172–221</sub> in the presence and absence of POPC liposomes were determined by SPR (Biacore T200, GE Healthcare). POPC vesicles containing DOGS-NTA were immobilised on an L1 chip (GE Healthcare) and loaded with  $\text{Ni}^{2+}$ , which bound the His<sub>6</sub>-tagged MSP2<sub>172–221</sub>. In the absence of lipid, MSP2<sub>172–221</sub> was immobilised on a CM5 chip (GE Healthcare) by conventional amide coupling following the manufacturer's

instructions and the binding affinity to C-terminal specific mAbs was measured.

For SPR assays of antibody binding, lipid vesicles were prepared from POPC and DOGS-NTA at a 100 : 1 molar ratio. The lipids were mixed in chloroform:methanol (1 : 1), dried under  $\text{N}_2$  and vacuum, and resuspended to 26.3 mM lipid in buffer (20 mM HEPES, 150 mM NaCl, pH 7.4). The lipid mixture was sonicated until a clear solution was obtained. Freshly prepared POPC lipid vesicles were diluted and injected from 0.65 mM stock at 2  $\mu\text{L}\cdot\text{min}^{-1}$  for 5 min onto the reference and active cells of a Biacore L1 chip followed by washing with 50 mM NaOH (30  $\mu\text{L}\cdot\text{mL}^{-1}$ ). 50 mM  $\text{NiSO}_4$  was injected for 3 min at 2  $\mu\text{L}\cdot\text{min}^{-1}$  and then running buffer (20 mM HEPES, 150 mM NaCl, pH 7.4) was used to rinse the prepared surface at 30  $\mu\text{L}\cdot\text{min}^{-1}$  for 2 min. Recombinant C-terminally His<sub>6</sub>-tagged protein (3D7 MSP2 or FC27 MSP2 or MSP2<sub>172–221</sub>) was loaded onto the lipid surface from a 150  $\mu\text{g}\cdot\text{mL}^{-1}$  stock at 2  $\mu\text{L}\cdot\text{min}^{-1}$  for 5 min. Antibody binding was assessed with five injections (2 min each) of increasing antibody concentration ranging from 1.2–100 nM at 30  $\mu\text{L}\cdot\text{min}^{-1}$ . The chip surface was regenerated using a 30 s injection of 30 mM NaOH in 40% (v/v) isopropanol between cycles. Similarly, for binding affinity studies in the absence of liposomes, antigens from a 6  $\mu\text{g}\cdot\text{mL}^{-1}$  stock were immobilised to a comparable level on CM5 chip and identical injections of mAb were performed.

Since the intramolecular disulfide bonding within the C-terminal region of MSP2 does not influence the binding of 9G8 or 4D11 but is crucial for 6C9 and 9H4 binding [21], the presence of free thiol in three different antigens was investigated using Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB) with cysteine as positive control and methionine as negative control. No MSP2 samples investigated here showed any colour, indicating that all antigens were in their oxidised form.

### Acknowledgements

RSN acknowledges fellowship support from the National Health and Medical Research Council of Australia. This work was supported in part by the National Health and Medical Research Council (project grant 1042520). SCD is supported by a Faculty of Pharmacy and Pharmaceutical Sciences and Monash Graduate Scholarship.

### Author contributions

SCD, CAM, RFA and RSN designed the experiments. SCD, BK and CAM performed the experiments with contributions of JS, RD and RAVM. SCD, JS, BK, RAVM, CAM, RFA and RSN analysed data. SCD,

CAM, RFA and RSN wrote the paper with input from all authors.

## Conflicts of interest

The authors declare no conflicts of interest.

## References

- WHO (2015) World Malaria Report.
- Elliott SR & Beeson JG (2008) Estimating the burden of global mortality in children aged <5 years by pathogen-specific causes. *Clin Infect Dis* **46**, 1794–1795.
- Thera MA & Plowe CV (2012) Vaccines for malaria: how close are we? *Annu Rev Med* **63**, 345–357.
- Von Seidlein L & Bejon P (2013) Malaria vaccines: past, present and future. *Arch Dis Child* **98**, 981–985.
- Anders RF, Adda CG, Foley M & Norton RS (2010) Recombinant protein vaccines against the asexual blood stages of *Plasmodium falciparum*. *Hum Vaccin* **6**, 39–53.
- Genton B, Betuela I, Felger I, Al-Yaman F, Anders R & Saul A (2002) A recombinant blood-stage malaria vaccine reduces *Plasmodium falciparum* density and exerts selective pressure on parasite populations in a phase 1-2b trial in Papua New Guinea. *J Infect Dis* **185**, 820–827.
- McCarthy JS, Marjason J, Elliott S, Fahey P, Bang G, Malkin E, Tierney E, Aked-Hurditch H, Adda C, Cross N *et al.* (2011) A phase 1 trial of MSP2-C1, a blood-stage malaria vaccine containing 2 isoforms of MSP2 formulated with Montanide(R) ISA 720. *PLoS One* **6**, e24413.
- Gilson PR, Nebl T, Vukcevic D, Moritz RL, Sargeant T, Speed TP, Schofield L & Crabb BS (2006) Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics* **5**, 1286–1299.
- Smythe JA, Peterson MG, Coppel RL, Saul AJ, Kemp DJ & Anders RF (1990) Structural diversity in the 45-kilodalton merozoite surface antigen of *Plasmodium falciparum*. *Mol Biochem Parasitol* **39**, 227–234.
- Conway DJ & Polley SD (2002) Measuring immune selection. *Parasitology* **125** (Suppl), S3–S16.
- Weedall GD & Conway DJ (2010) Detecting signatures of balancing selection to identify targets of anti-parasite immunity. *Trends Parasitol* **26**, 363–369.
- Smythe JA, Coppel RL, Day KP, Martin RK, Oduola AM, Kemp DJ & Anders RF (1991) Structural diversity in the *Plasmodium falciparum* merozoite surface antigen 2. *Proc Natl Acad Sci USA* **88**, 1751–1755.
- Fenton B, Clark JT, Khan CM, Robinson JV, Walliker D, Ridley R, Scaife JG & McBride JS (1991) Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite surface antigen (MSA-2) of the malaria parasite *Plasmodium falciparum*. *Mol Cell Biol* **11**, 963–971.
- Flück C, Smith T, Beck HP, Irion A, Betuela I, Alpers MP, Anders R, Saul A, Genton B & Felger I (2004) Strain-specific humoral response to a polymorphic malaria vaccine. *Infect Immun* **72**, 6300–6305.
- Flück C, Schopflin S, Smith T, Genton B, Alpers MP, Beck HP & Felger I (2007) Effect of the malaria vaccine Combination B on merozoite surface antigen 2 diversity. *Infect Genet Evol* **7**, 44–51.
- Osier FH, Feng G, Boyle MJ, Langer C, Zhou J, Richards JS, McCallum FJ, Reiling L, Jaworowski A, Anders RF *et al.* (2014) Opsonic phagocytosis of *Plasmodium falciparum* merozoites: mechanism in human immunity and a correlate of protection against malaria. *BMC Med* **12**, 108.
- Boyle MJ, Reiling L, Feng G, Langer C, Osier FH, Aspelting-Jones H, Cheng YS, Stubbs J, Tetteh KK, Conway DJ *et al.* (2015) Human antibodies fix complement to inhibit *Plasmodium falciparum* invasion of erythrocytes and are associated with protection against malaria. *Immunity* **42**, 580–590.
- Osier FH, Murungi LM, Fegan G, Tuju J, Tetteh KK, Bull PC, Conway DJ & Marsh K (2010) Allele-specific antibodies to *Plasmodium falciparum* merozoite surface protein-2 and protection against clinical malaria. *Parasite Immunol* **32**, 193–201.
- Taylor RR, Smith DB, Robinson VJ, McBride JS & Riley EM (1995) Human antibody response to *Plasmodium falciparum* merozoite surface protein 2 is serogroup specific and predominantly of the immunoglobulin G3 subclass. *Infect Immun* **63**, 4382–4388.
- MacRaild CA, Zachrdla M, Andrew D, Krishnarjuna B, Novacek J, Zidek L, Sklenar V, Richards JS, Beeson JG, Anders RF *et al.* (2015) Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2. *PLoS One* **10**, e0119899.
- Adda CG, MacRaild CA, Reiling L, Wycherley K, Boyle MJ, Kienzle V, Masendycz P, Foley M, Beeson JG, Norton RS *et al.* (2012) Antigenic characterization of an intrinsically unstructured protein, *Plasmodium falciparum* merozoite surface protein 2. *Infect Immun* **80**, 4177–4185.
- Zhang X, Perugini MA, Yao S, Adda CG, Murphy VJ, Low A, Anders RF & Norton RS (2008) Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2. *J Mol Biol* **379**, 105–121.
- Morales RA, MacRaild CA, Seow J, Krishnarjuna B, Drinkwater N, Rouet R, Anders RF, Christ D, McGowan S & Norton RS (2015) Structural basis for epitope masking and strain specificity of a conserved

- epitope in an intrinsically disordered malaria vaccine candidate. *Sci Rep* **5**, 10103.
- 24 MacRaild CA, Pedersen MØ, Anders RF & Norton RS (2012) Lipid interactions of the malaria antigen merozoite surface protein 2. *Biochim Biophys Acta* **1818**, 2572–2578.
- 25 Wishart DS, Bigam CG, Holm A, Hodges RS & Sykes BD (1995)  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomol NMR* **5**, 67–81.
- 26 Wishart DS & Sykes BD (1994) The  $^{13}\text{C}$  Chemical-Shift Index: a simple method for the identification of protein secondary structure using  $^{13}\text{C}$  chemical-shift data. *J Biomol NMR* **4**, 171–180.
- 27 Schwarzingler S, Kroon GJ, Foss TR, Chung J, Wright PE & Dyson HJ (2001) Sequence-dependent correction of random coil NMR chemical shifts. *J Am Chem Soc* **123**, 2970–2978.
- 28 Op den Kamp JAF, Roelofsen B & van Deenen LLM (1985) Structural and dynamic aspects of phosphatidylcholine in the human erythrocyte membrane. *Trends Biochem Sci* **10**, 320–323.
- 29 Sherman IW (1979) Biochemistry of *Plasmodium* (malaria parasites). *Microbiol Rev* **43**, 453–495.
- 30 Seow J, Morales RA, MacRaild CA, Krishnarjuna B, McGowan S, Dingjan T, Jaipuria G, Rouet R, Wilde KL, Atreya HS *et al.* (2017) Structure and characterisation of a key epitope in the conserved C-terminal domain of the malaria vaccine candidate MSP2. *J Mol Biol* **429**, 836–846.
- 31 Bütikofer P, Malherbe T, Boschung M & Roditi I (2001) GPI-anchored proteins: now you see 'em, now you don't. *FASEB J* **15**, 545–548.
- 32 Barboni E, Rivero BP, George AJ, Martin SR, Renoup DV, Hounsell EF, Barber PC & Morris RJ (1995) The glycophosphatidylinositol anchor affects the conformation of Thy-1 protein. *J Cell Sci* **108** (Pt 2), 487–497.
- 33 Watson DS, Endsley AN & Huang L (2012) Design considerations for liposomal vaccines: influence of formulation parameters on antibody and cell-mediated immune responses to liposome associated antigens. *Vaccine* **30**, 2256–2272.
- 34 Watson DS, Platt VM, Cao L, Venditto VJ & Szoka FC (2011) Antibody response to polyhistidine-tagged peptide and protein antigens attached to liposomes via lipid-linked nitrilotriacetic acid in mice. *Clin Vaccine Immunol* **18**, 289–297.
- 35 Butts C, Murray N, Maksymiuk A, Goss G, Marshall E, Soulieres D, Cormier Y, Ellis P, Price A, Sawhney R *et al.* (2005) Randomized phase IIB trial of BLP25 liposome vaccine in stage IIIB and IV non-small-cell lung cancer. *J Clin Oncol* **23**, 6674–6681.
- 36 Bovier PA (2008) Epaxal: a virosomal vaccine to prevent hepatitis A infection. *Expert Rev Vaccines* **7**, 1141–1150.
- 37 Fries LF, Gordon DM, Richards RL, Egan JE, Hollingdale MR, Gross M, Silverman C & Alving CR (1992) Liposomal malaria vaccine in humans: a safe and potent adjuvant strategy. *Proc Natl Acad Sci USA* **89**, 358–362.
- 38 Agnandji S, Lell B, Soulanoudjingar S, Fernandes J, Abossolo B & Conzelmann C (2011) First results of phase 3 trial of RTS, S/AS01 malaria vaccine in African children. *N Engl J Med* **365**, 1863–1875.
- 39 van Dissel JT, Joosten SA, Hoff ST, Soonawala D, Prins C, Hokey DA, O'Dee DM, Graves A, Thierry-Carstensen B, Andreassen LV *et al.* (2014) A novel liposomal adjuvant system, CAF01, promotes long-lived *Mycobacterium tuberculosis*-specific T-cell responses in human. *Vaccine* **32**, 7098–7107.
- 40 Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J & Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696.