

High-resolution NMR studies of malaria proteins as a basis for designing better drugs and vaccines

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

By

Krishnarjuna Bankala

Bachelor of Science (Microbiology) Sri Krishnadevaraya University, India

Master of Science (Biochemistry) Sri Venkateswara University, India



Monash University

Medicinal Chemistry
Monash Institute of Pharmaceutical Sciences
Melbourne, Victoria, Australia

2016

Copyright notice

© The author (2016). Except as provided in the Copyright Act 1968, this thesis may not be reproduced in any form without the written permission of the author.

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.

Monash University

Declaration for thesis based or partially based on conjointly published or unpublished work

General Declaration

In accordance with Monash University Doctorate Regulation 17.2 Doctor of Philosophy and Research Master's regulations the following declarations are made:

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 five original papers published in peer reviewed journals and one traditional chapter. The core theme of the thesis is: Characterization of small molecule interactions with a malaria drug candidate *Plasmodium falciparum* apical membrane antigen 1 and development of merozoite surface protein 2 based vaccine constructs for malaria caused by *P. falciparum*. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the Department of Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences under the supervision of Professor Raymond S. Norton.

[The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.]

In the case of Chapters 2-7 my contribution to the work involved the following:

Thesis chapter	Publication title	Publication status	Nature and extent of candidate's contribution
2	Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials	Published (<i>J. Mol. Recognit.</i> 2016)	Designed the experiments, prepared ^2H , ^{15}N , ^{13}C labelled AMA1 samples, performed NMR experiments, determined NMR assignments, performed AMA1 binding experiments with small molecules, analysed data and prepared manuscript Contribution: 50%
3	Structure and dynamics of apical membrane antigen 1 from <i>Plasmodium falciparum</i> FVO	<i>Biochemistry</i> 2014, 53:7310-7320	Prepared ^2H , ^{15}N , ^{13}C labelled AMA1 samples, recorded [^1H - ^{15}N]-TROSY and TROSY-HNCO NMR experiments, analysed NMR data and provided intellectual input in writing the manuscript Contribution: 20%
4	Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2	<i>PLoS ONE</i> 2015, 10: e0119899, 1-18	Cloned, expressed and purified ^{15}N , and ^{15}N , ^{13}C labelled 3D7 MSP2 samples and collaborated in preparation of manuscript Contribution: 20%
5	Strain-transcending immune response generated by chimeras of the malaria vaccine candidate merozoite surface protein 2	<i>Sci. Rep.</i> 2016, 6:20613, 1-12	Designed, cloned, expressed, purified and characterized the proteins used in this study. Analysed the results from mice immunization experiments and prepared manuscript Contribution: 50%
6	Structural basis for epitope masking and strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate	<i>Sci. Rep.</i> 2015, 5:10103, 1-10	Prepared 3D7 MSP2, performed and analysed LC-MS of the proteins used in this study and provided intellectual input in writing the manuscript

			Contribution: 15%
7	NMR characterization of transient interactions modulating the strain-specificity of 6D8 monoclonal antibody binding	Traditional chapter	Purified ^2H , ^{15}N , ^{13}C labelled 6D8 scFv, recorded triple-resonance NMR experiments and determined NMR assignments for both epitope free and epitope bound 6D8 scFv, performed 6D8 scFv binding experiments with 3D7MSP ₂₁₄₋₃₄ and FC27MSP ₂₁₄₋₃₄ peptides, analysed NMR data and drafted the results Contribution: 80%

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

Signed:



Date: 30/03/2016

Publications

Lim SS, Yang W, Krishnarjuna B, Kannan Sivaraman K, Chandrashekar IR, Kass I, MacRaild CA, Devine SM, Debono CO, Anders RF, Scanlon MJ, Scammells PJ, Scanlon MJ, Norton RS and McGowan S (2014) Structure and dynamics of apical membrane antigen 1 from *Plasmodium falciparum* FVO. *Biochemistry* **53**:7310-7320 (11 pages).

MacRaild CA, Zachrdla M, Andrew D, Krishnarjuna B, Nováček J, Žídek L, Sklenář V, Richards JS, Beeson JG, Anders RF and Norton RS (2015) Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2. *PLoS ONE* **10**, e0119899 (18 pages).

Morales RAV, MacRaild CA, Seow J, Krishnarjuna B, Drinkwater N, Rouet R, Anders RF, Christ D, McGowan S and 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 (10 pages).

Krishnarjuna B, Andrew D, MacRaild CA, Morales RAV, Beeson JG, Anders RF, Richards JS and Norton RS (2016) Strain-transcending immune response generated by chimeras of the malaria vaccine candidate merozoite surface protein 2. *Sci Rep* **6**, 20613 (12 pages).

Krishnarjuna B, Lim SS, Devine SM, Debono CO, Lam R, Chandrashekar IR, Jaipuria G, Yagi H, Atreya H, Scanlon MJ, MacRaild CA, Scammells PJ and Norton RS (2016) Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials. *J Mol Recognit* **29**:281-291 (11 pages).

Acknowledgments

Over the course of these projects I have had the privilege of working with a number of brilliant individuals. Without their help these projects would simply not have happened.

First and foremost, I would like to express my sincere gratitude to Professor Raymond S. Norton for his leadership in directing the projects and for his unprecedented ability to bring together the collaborations of outstanding people required to ensure their success. He not only provided me with a lot of intellectual support and constructive criticism throughout my Ph.D., but was also a good role model as a highly enthusiastic, professional and successful scientist. His guidance and undivided support have allowed me to build myself into a more independent researcher and I am very grateful to him for giving me a lot of opportunities during my PhD years to expand my skill sets and to improve myself.

I would like to especially thank Dr Christopher A. MacRaid, Mr Dean Andrew, Dr Indu R. Chandrashekar, Dr Shane M. Devine, Dr Garima Jaipuria and Dr Rodrigo A. V. Morales for their recommendations and suggestions, which have been invaluable for my research progress. Without their help I could not have got this far. Special thanks also to Karyn Wilde from ANSTO for proving the bacterial pellets expressing deuterated 6D8 scFv.

I also wish to extend my gratitude to Emeritus Professor Robin F. Anders, Dr. Jack S. Richards, Associate Professor Martin J. Scanlon, Professor James G. Beeson, Associate Professor Hanudatta S. Atreya, Dr Nyssa Drinkwater and Dr Sheena McGowan, whose support has been invaluable. To the members of the Scanlon group, especially Dr Martin L. Williams and Dr Stephen J. Headey, thank you for your constant help and advice.

I also thank my fellow colleagues Gaurav Sharma, Keith Khoo, Samuel Robinson, Steven Yap, Tony Wang, Eleanor Leung, San Lim, Stephen Drane, Jeff Seow and Cael Debono, for their support and help.

Last but not least I wish to avail myself of this opportunity to express a sense of gratitude and love to my friends and my beloved family members for their support, strength, help and for everything.

Table of Contents

Copyright notice	ii
General Declaration	iii
Publications.....	vi
Acknowledgments	vii
Abbreviations.....	xi
Abstract.....	xiii
CHAPTER 1: Introduction.....	1
1.1 Introduction.....	2
1.1.1 The malaria life cycle.....	2
1.1.2 Antimalarial agents and drug resistance	3
1.1.3 Malaria vaccine	4
1.2 Apical membrane antigen 1	6
1.2.1 Erythrocyte invasion mechanism	6
1.2.2 Structure of apical membrane antigen 1.....	8
1.2.2.1 Hydrophobic cleft.....	8
1.2.3 Invasion inhibitory molecules	10
1.2.3.1 Anti-AMA1 antibodies	10
1.2.3.2 Invasion inhibitory peptides	11
1.2.4 AMA1 as a therapeutic target	12
1.2.5 Small molecule inhibitors of AMA1	13
1.2.6 Fragment-based drug design	13
1.2.7 AMA1 as a vaccine candidate.....	14
1.3 Merozoite surface protein 2	16
1.3.1 Structure of MSP2.....	17
1.3.2 Immunogenicity of MSP2.....	17
1.3.3 MSP2 epitopes to mouse mAbs	18
1.3.4 MSP2 as a vaccine candidate	19
1.3.5 Chimeric malaria antigens as vaccines	20
1.4 Project Scope and Aims.....	20
CHAPTER 2: Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials	24
2.1 Declaration of Thesis Chapter 2	25
2.2 Introduction.....	26

CHAPTER 3: Structure and dynamics of apical membrane antigen 1 from <i>Plasmodium falciparum</i> FVO	27
3.1 Declaration of Thesis Chapter 3	28
3.2 Introduction	30
CHAPTER 4: Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2	31
4.1 Declaration of Thesis Chapter 4	32
4.2 Introduction	33
CHAPTER 5: Strain-transcending immune response generated by chimeras of the malaria vaccine candidate merozoite surface protein 2	34
5.1 Declaration of Thesis Chapter 5	35
5.2 Introduction	36
CHAPTER 6: Structural basis for epitope masking and strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate	37
6.1 Declaration of Thesis Chapter 6	38
6.2 Introduction	39
CHAPTER 7: NMR characterization of transient interactions modulating the strain-specificity of 6D8 monoclonal antibody binding	40
7.1 Introduction	41
7.2 Materials and Methods	41
7.2.1 ² H, ¹³ C and ¹⁵ N-labelled 6D8 scFv expression	41
7.2.2 6D8 scFv purification	42
7.2.2.1 Periplasmic extraction	42
7.2.2.2 Preparation of MSP2 affinity column	42
7.2.2.3 6D8 scFv affinity purification	43
7.2.3 Size-exclusion chromatography (SEC)	43
7.2.4 Solution NMR spectroscopy	43
7.2.4.1 Sequential NMR assignments	43
7.2.4.2 Peptide binding experiments by NMR	44
7.3 Results	45
7.3.1 Sequential backbone resonance assignments of MSP ₁₄₋₂₂ free 6D8 scFv	45
7.3.2 Sequential backbone resonance assignments for MSP ₂₁₄₋₂₂ -bound 6D8 scFv	48
7.3.3 H ^N -water exchange in presence and absence of MSP ₂₁₄₋₂₂	52
7.3.4 Identification of 6D8 scFv residues that interact with MSP2 variable region	54

7.4 Discussion	56
7.5 Conclusion and future work.....	58
CHAPTER 8: Conclusions and Future Work.....	60
CHAPTER 9: References	67
Appendix I	82
Appendix II.....	84
Appendix III	85
Appendix IV	87
Appendix V.....	92

Abbreviations

ACT	Artemisinin-based combination therapy
ADCI	Antibody dependent cellular inhibition
AMA1	Apical membrane antigen 1
CDR	Controlled human malaria infection
CHIMI	Complementarity determining region
CPMG	Carr-Purcell-Meiboom-Gill
CSI	Chemical shift index
CSP	Chemical shift perturbation
CTR	C-terminal region
ELISA	Enzyme linked immune assay
<i>E. coli</i>	<i>Escherichia coli</i>
DMSO	Dimethyl sulfoxide
DPC	Dodecylphosphocholine
FBLD	Fragment based ligand design
FPLC	Fast protein liquid chromatography
Gdn-HCl	Guanidine hydrochloride
GIA	Growth inhibition assay
GSH	Reduced glutathione
GSSG	Oxidised glutathione
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HSQC	Heteronuclear single-quantum coherence
HTS	High-throughput screening
IC50	Half maximal inhibitory concentration
IEX	Ion exchange
IFA	Immunofluorescence assay
Ig	Immunoglobulin
ITC	Isothermal calorimetry
LB	Luria broth
LE	Ligand efficiency
mAb	Monoclonal antibody
MD	Molecular dynamics
MJ	Moving junction
MSP1	Merozoite surface protein 1
MSP2	Merozoite surface protein 2
NMR	Nuclear magnetic resonance
NTR	N-terminal region
NUS	Non-uniform sampling
PDB	Protein data bank
PAN	Plasminogen-apple-nematode
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
<i>Pf</i>	<i>Plasmodium falciparum</i>
PPI	Protein-protein interaction
<i>Pv</i>	<i>Plasmodium vivax</i>
RESA	Ring-infected erythrocyte surface antigen
RON	Rhoptry neck protein
RU	Response unit
SAR	Structure-activity relationship

SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SPR	Surface plasmon resonance
STD	Saturation transfer difference
SUB2	Subtilisin-like protease 2
TEV	Tobacco etch virus
TFA	Tri-fluoro acetic acid
TFE	Thioredoxin
Trx	Trifluoroethanol
<i>Tg</i>	<i>Toxoplasma gondii</i>
TROSY	Transverse relaxation optimised spectroscopy
TM	Transmembrane domain
VR	Variable region
WHO	World Health Organisation

Abstract

Malaria is a life-threatening disease caused by *Plasmodium* spp., particularly *Plasmodium falciparum* (Pf). Although current therapies have been very successful in controlling malaria, resistant parasites have already emerged. No vaccine is currently available for complete prevention of the disease. Therefore, the design of drugs with new modes of action to combat this threat and developing a vaccine for disease prevention remain high priorities. Antigenic polymorphism is a major limiting factor in the development of drugs and vaccines for treating malaria. This thesis focuses on the development of small molecule inhibitors for *P. falciparum* apical membrane antigen 1 (PfAMA1) as new antimalarials and an optimized merozoite surface protein 2 (MSP2)-based vaccine construct to achieve a strain-transcending immune response.

PfAMA1 is an essential component of the moving junction (MJ) complex that forms between erythrocytes and parasites during the invasion by *P. falciparum*. PfAMA1 has a conserved hydrophobic cleft that is the binding site for rhoptry neck protein 2 (RON2). As anti-AMA1 antibodies and other inhibitory molecules that target this hydrophobic cleft are able to block the invasion process, PfAMA1 is an attractive target for the development of new antimalarials. The first part of the thesis describes the determination of the nuclear magnetic resonance (NMR) assignments for two variants (3D7 and FVO) of PfAMA1 in order to facilitate the screening and identification of small molecules that bind to the hydrophobic cleft. Additionally, the first crystal structure of FVO PfAMA1 was determined to understand the role of polymorphism on AMA1 structure and to facilitate the design of small molecule inhibitors. The binding sites for small molecules that bind to the hydrophobic cleft were identified using [¹⁵N-¹H]-TROSY perturbation experiments. Among the small molecules tested, the binding sites for benzimidazoles on both FVO and 3D7 PfAMA1 were similar, suggesting that these compounds are promising starting scaffolds for the development of potential PfAMA1 inhibitors.

The second part of my PhD involves the design and development of an optimized MSP2-based vaccine construct for malaria. MSP2 is an intrinsically disordered glycosylphosphatidylinositol-anchored antigen expressed on the merozoite surface. Vaccine trials with recombinant MSP2 showed protection against *P. falciparum*. However, the implications of MSP2 disorder in producing a protective immune response are unknown. Therefore, the conformational and dynamic properties of two allelic forms of MSP2 (3D7 and FC27) were studied using NMR spectroscopy and it was found that these properties were identical in conserved regions but significantly different between the variable regions.

It was observed that the conformationally restricted regions are more antigenic than the more flexible regions, suggesting that the immunogenic efficacy of disordered antigens can be improved by increasing the conformational order of their flexible regions.

The immune response to MSP2 provides protection, but the protection is strain-specific owing to the polymorphic nature of MSP2. In order to achieve a strain-transcending immune response to MSP2, a series of chimeric MSP2 constructs was designed, characterized and used in immunization experiments in mice. Overall, a robust immune response to both conserved and variable regions of all constructs was generated in mice. Among the constructs tested, V_{3D7}V_{FC27}C chimera (V_{3D7}: 3D7 variable region, V_{FC27}: FC27 variable region and C: conserved C-terminal region) was able to induce a uniform immune response to both conserved and variable region epitopes of 3D7 and FC27 MSP2 epitopes. Total IgG levels to this construct were equivalent to that obtained from immunizing with co-administered 3D7+FC27, and generated stronger IgG2b and IgG2c responses. Constructs lacking the conserved regions showed a reduced immune response, suggesting the importance of these regions in generating a robust immune response.

Structural information on antibody-bound MSP2 epitopes would help to understand the native MSP2 conformation as well as the strain-specific protection of MSP2. Additionally, such conformations could be incorporated into either full-length or chimeric MSP2 to increase their conformational order, which would lead to improved immunogenicity. 6D8 is a monoclonal antibody that recognizes a conserved epitope in recombinant MSP2 but cannot recognize native MSP2. The crystal structure of 6D8 scFv-bound 6D8 epitope was determined and compared with the structure of the lipid-bound epitope. When bound to lipid the conserved residue Arg22 is part of a helix and unable to make interactions with 6D8 that are required for recognition. This suggests a mechanism by which native MSP2 escapes antibody recognition. Moreover, NMR binding experiments identified transient and strain-specific interactions between 6D8 and the variable regions of 3D7 and FC27 MSP2 that lead to different affinities for 6D8.

In summary, small molecules that bind to the *Pf*AMA1 hydrophobic cleft were identified and chemical elaboration of these molecules could generate high affinity strain-transcending *Pf*AMA1 inhibitors. On the other hand, chimeric constructs of MSP2 can be used to induce a strain-transcending immune response to *P. falciparum* and represent viable choices to include in a future malaria vaccine. Additionally, the role of conformational order in determining antigenicity of MSP2 was investigated and, together with the information from

the antibody-bound MSP2 epitopes, could have important implications for the design of a better MSP2 vaccine.

CHAPTER 1

Introduction

1.1 Introduction

Malaria is a life-threatening disease caused by parasites of the genus *Plasmodium*. The disease in humans is caused by five different *Plasmodium* species, *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae* and *P. knowlesi*. Malaria due to *P. falciparum*, which predominates in Africa, is the most deadly form. On the other hand, malaria caused by *P. vivax* is less dangerous but more widespread, while malaria cases involving the other three species are less common. Malaria parasites are transmitted to humans *via* the bite of infected female mosquitoes of more than 30 anopheline species. According to the World Health Organization (WHO) malaria report in 2015, around 214 million malaria cases with approximately 438,000 deaths annually were reported worldwide. Populations living in sub-Saharan Africa were found to have the highest risk of acquiring malaria; approximately 80% of malaria cases and 90% of deaths were reported in the African region, with children under five years of age and pregnant women being most severely affected [1]. This consequently hampered both social and economic development in the affected regions [2].

1.1.1 The malaria life cycle

Plasmodium species have a complex life cycle (**Figure 1.1**) in humans, where parasites infect hepatocytes and erythrocytes. As the female *Plasmodium*-infected mosquito sucks blood from a human host, the haploid (n) form of the parasite known as sporozoites, are injected into the host along with anticoagulant-containing saliva. Sporozoites then travel to the human liver, where they rapidly invade hepatocytes and undergo asexual multiplication over a period of up to two weeks, generating several thousand haploid merozoites.

Merozoites released from the infected hepatocytes into the bloodstream subsequently invade erythrocytes, where they replicate. After 48-72 h of erythrocyte invasion, the mature parasites, known as schizonts, rupture and release 8-32 new merozoites and toxins. These new merozoites then infect new erythrocytes and the released toxins cause fever, sweats, headache and chills in the human host [3, 4]. The newly-infected erythrocyte contains a ring stage, which is named after the signet ring appearance of the chromatin dot and blue circle of parasite cytoplasm due to haemozoin [5]. The ring stage proceeds to the trophozoite stage, during which the microscopically visible haemozoin or malarial pigment appears in the cytoplasm. The schizont stage commences when the nucleus replicates. The schizonts release merozoites, which infect new erythrocytes.

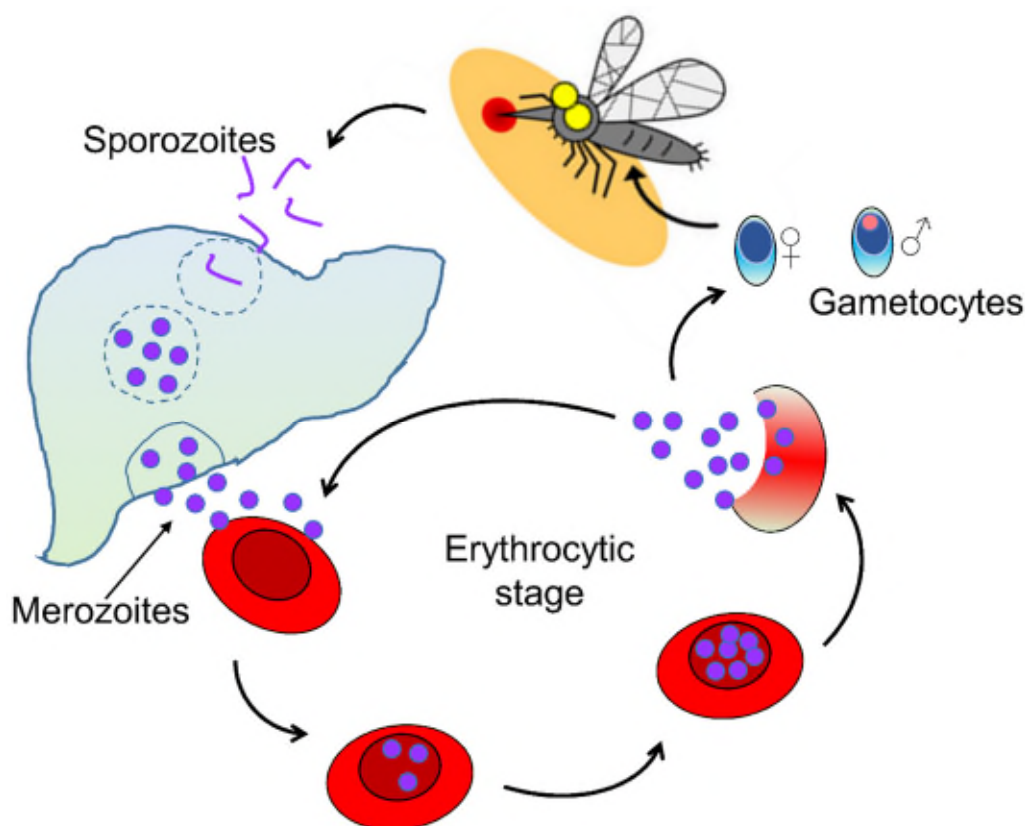


Figure 1.1: Life cycle of *Plasmodium* parasites in the human host.

The human host experiences periodic cycles of fever and chills during the asexual life cycle. Fewer than 1% of the asexual blood-stage parasites will differentiate into male and female gametocytes. When a mosquito bites an infected human, mature gametocytes enter the insect's midgut, where the male gametocytes ex-flagellate and form gametes, which fuse with the female gametes to form diploid ookinetes. The ookinetes then penetrate through the midgut wall and form oocysts. Once oocysts mature, they rupture and release haploid sporozoites, which then migrate to the salivary glands of the mosquito. When the mosquito takes a blood meal, sporozoites enter the human host and the asexual life cycle starts again [4].

1.1.2 Antimalarial agents and drug resistance

Chloroquine and mefloquine, two of the earliest antimalarial agents, function by interfering with haem sequestration. Despite the widespread resistance to both drugs, haem sequestration remains a viable target for development of new antimalarials [6, 7]. Recent reports revealed that 4-amino-7-chloroquinoline-based molecules possess growth inhibitory activity on chloroquine-resistant parasites [8].

Pyrimethamine and sulfadoxine combination therapy has long been used as an important antimalarial combination [9, 10]. Pyrimethamine exerts its pharmacological action by inhibiting parasite dihydrofolate reductase, thereby blocking the biosynthesis of purines and pyrimidines. Sulfadoxine inhibits the dihydropteroate synthetase, preventing the utilisation of *para*-aminobenzoic acid in the synthesis of dihydropteroic acid. This anti-folate combination is highly effective, cheap and well-tolerated. Nevertheless, its clinical role in malaria therapy has declined owing to the occurrence of resistance to the inhibitors [7]. It is now known that mutations in *pfdhfr* and *pfdhps* genes are responsible for conferring malaria resistance to pyrimethamine and sulfadoxine [11].

The WHO has recommended the use of artemisinin-based combination therapies (ACT) as the first-line treatment to clear uncomplicated malaria and it continues to be very successful in most regions where the disease was endemic [12-15]. Members of the artemisinin group of antimalarials contain an endoperoxide moiety that is required for its antimalarial activity [16, 9, 10]. Activation of the endoperoxide group occurs in the presence of free Fe^{3+} , which is released during digestion of haemoglobin. This is followed by production of a carbon-centered radical that alkylates and damages macromolecules in the parasites. Implementation of ACTs together with the use of mosquito nets has dramatically reduced both malaria cases and the mortality rate in the last 10 years. However, artemisinin-resistant parasites have been observed in some parts of the world including Africa and Southeast Asia and appear to be spreading to other parts of the world [17-24]. The spread of drug-resistant parasites challenges the efficacy of current treatments and impedes efforts to control malaria [25].

1.1.3 Malaria vaccine

While antimalarials are designed to treat the disease, a vaccine could prevent the development of malaria altogether. Development of an effective malaria vaccine, however, is challenging due to variation in antigens expressed during different stages of the complex life cycle of *Plasmodium* parasites. Liver-stage (pre-erythrocytic), blood-stage and sexual-stage are important targets for the human immune system [26-28]. The advantage of pre-erythrocytic vaccines over blood-stage and sexual-stage vaccines is that only a small population (tens to a few hundreds) of parasites (sporozoites) need to be targeted during the liver infection. Cytotoxic T-cells are able to kill infected hepatocytes (liver cells) and thus block the development of parasites and their transmission to initiate the blood-stage life cycle [29-32]. One candidate malaria vaccine targeting the liver-stage of the parasite is the

circumsporozoite protein (CSP)-based vaccine RTS,S/AS01E [33], which is the first malaria vaccine to have completed a Phase III clinical trial [34]. It was found that the administration of vaccine in four doses reduced clinical malaria by 39% and severe malaria by 31.5% in 5-17 month old children [34, 1, 35]. Although it has high potential in reducing cases of severe disease, it has limitations in eliciting long-term protection [34, 36, 33, 37]. This vaccine is considered as a complementary intervention based on the results from the Phase III trial, and WHO has recommended that research into developing new drugs and vaccines for malaria be continued.

In addition to the liver-stage, the blood-stage of the parasite is also a major target for the human immune system and is currently an important focus for developing new vaccines. Clinical trials have been carried out using the most promising blood-stage vaccine candidate antigens, apical membrane antigen 1 (AMA1) and merozoite surface protein 2 (MSP2) [38]. Both are polymorphic antigens; AMA1 is expressed during liver-stage as well as blood-stage and MSP2 is expressed only during blood-stage. Both these surface antigens are potential targets for antibodies, induced by either vaccination or the malaria parasite itself. The efficacies of vaccines containing strain-specific AMA1 or MSP2 antigens, however, are limited due to the extensive polymorphic nature of these antigens [39-41], as discussed further below.

As antigens like CSP are expressed only during the liver stage [42, 43], while other antigens such as MSP2 are expressed only during blood-stage [44-47] of the parasite life cycle, the immune response against the hepatic stage of the parasite does not affect the blood-stage of the parasite, and *vice versa*, as both sporozoites and merozoites express different sets of antigens. As such, a vaccine that targets multiple stages of parasite life cycle (for example, AMA1) [48-51] would be a more sensible approach. However, again, AMA1 polymorphic nature limits its efficacy as a vaccine component. In order to overcome the limitation of strain-specific immune responses, a vaccine with many variants of a particular antigen (AMA1 or MSP2) or a combination of different antigens could be used. However, antigens like AMA1 require refolding after expression in *Escherichia coli* (*E. coli*) to obtain their functional form and the refolding of these proteins is a challenging and time-consuming process. Alternatively, a chimeric vaccine comprised of different alleles of a polymorphic antigen (preferably disordered proteins like MSP2 which does not require refolding) may be feasible to overcome allelic polymorphism, and this is explored in this thesis.

Although a vaccine can potentially prevent disease, no effective vaccine for malaria has been developed to date. In addition to the existence of drug-resistant parasites (section 1.1.2),

antigenic polymorphism in *Plasmodium* antigens was found to be one of the limiting steps in developing both drugs [52] as well as vaccines for malaria [53-57]. These suggest that new drugs with novel mode of action for treatment and effective vaccines to eradicate malaria are required.

The focus of this thesis will be on AMA1 and MSP2, describing the properties of each that make them suitable as potential drug and vaccine candidates.

1.2 Apical membrane antigen 1

Apicomplexan parasites including *P. falciparum* and *Toxoplasma gondii* are obligate intracellular parasites that share a unique erythrocyte invasion mechanism which involves the formation of a ring-like structure between the erythrocyte and the parasite known as the moving junction (MJ) [58, 59]. *PfAMA1* is essential for the erythrocyte invasion by merozoites [60, 61]. The conditional knock-down of *PfAMA1* severely impaired the parasite's ability to invade erythrocytes [62]. A complete gene knock-out is not viable in *P. falciparum*, providing direct genetic evidence that *PfAMA1* plays a key role in invasion [63, 61]. In support of this, anti-AMA1 antibodies are able to block the merozoite invasion *in vitro* [64, 65]. On the other hand, the role of AMA1 in invasion has been questioned recently, noting that AMA1-depleted parasites (*T. gondii* and *P. berghei*) are able to form a functional MJ [66, 67]. However, further studies found that AMA1 homologues in *T. gondii* that are absent in *P. falciparum* enable the erythrocyte invasion when AMA1 was inactivated [68], possibly accounting for these divergent results [69]. Moreover, a subsequent study demonstrated that AMA1 knockdown *P. falciparum* was unable to invade erythrocytes, thus strongly supporting the importance of AMA1 for erythrocyte invasion [62].

1.2.1 Erythrocyte invasion mechanism

The invasion process starts with attachment of the merozoite to the host cell membrane through merozoite surface proteins, EBA175, Ebl1, EBA140, *PfRh4* and *PfRh5* [70]. Once attached to the host cell, the merozoite reorients itself in such a manner that its anterior end (i.e. with rhoptries and micronemes) faces towards the erythrocyte surface [71]. During the schizont stage, the parasites synthesize and transport secretory vesicles called rhoptries and micronemes to the apical region of merozoite (**Figure 1.2**) [72]. AMA1 is stored in the micronemes and is released onto the merozoite surface prior to the invasion process. On the other hand, RON proteins, which are derived from rhoptries, are localised on the host

erythrocytes during invasion. It has been shown recently that AMA1 uses rhoptry neck protein 2 (RON2) as a receptor to promote invasion and that this AMA1-RON2 interaction is equally important for the invasive process by *P. falciparum* and *T. gondii* [73-76].

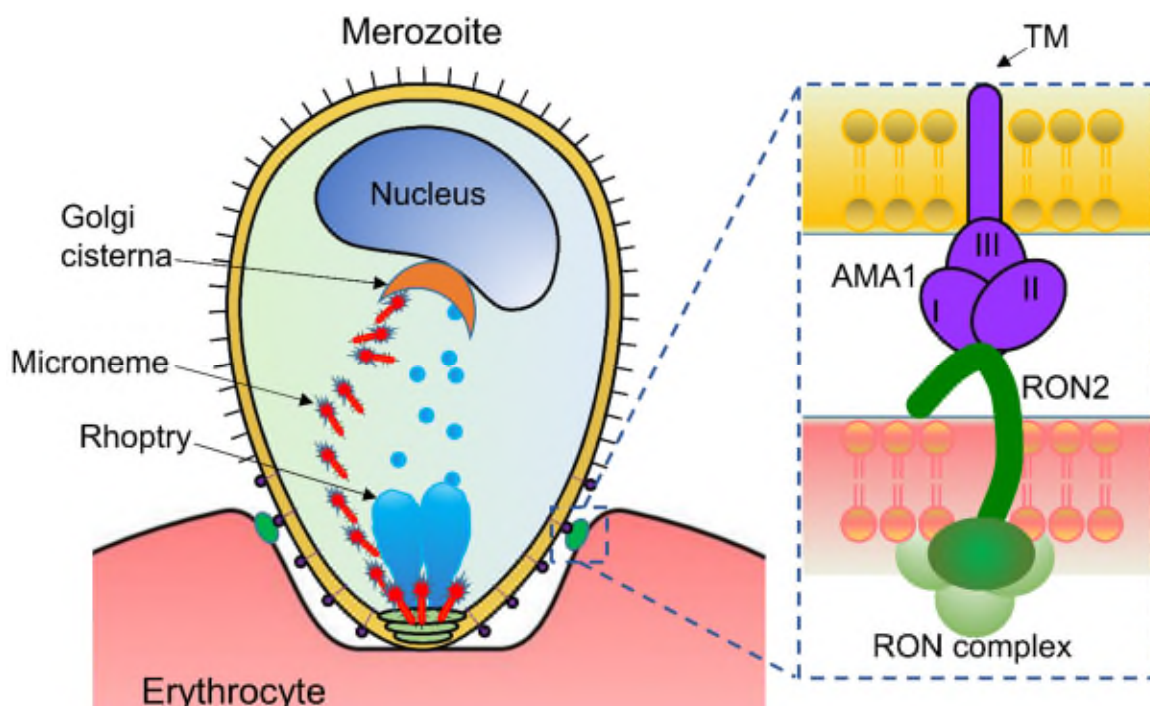


Figure 1.2: Graphical representation of AMA1-RON2 interaction during the erythrocyte invasion by merozoite. Secretory vesicles, micronemes and rhoptries carry the AMA1 and RON proteins to the merozoite apical region. During invasion, RON2 is released and integrates into the erythrocyte membrane and acts as a receptor for AMA1. AMA1 is then released from micronemes and forms a complex with RON2. The AMA1-RON complex is an important component of the MJ. AMA1 consists of three extracellular domains (DI, DII and DIII), a transmembrane domain (TM) and a short cytoplasmic region. DI and DII together form a binding site for RON2.

During this invasion process, RON2 integrates into the host cell plasma membrane, while other RON proteins 4, 5 and 8 localize to its cytoplasmic face (**Figure 1.2**) [77, 58]. AMA1 is then released from micronemes and interacts with RON2, leading to the formation of MJ [74, 76], which is a crucial step in establishing a bridge between the parasite and the erythrocyte. Once the link is established, the MJ moves along the parasite surface towards its posterior end, powered by an actomyosin motor present below the parasite membrane, which then leads to the formation of a parasitophorous vacuole around the parasite [71]. Subsequently, the parasite internalizes into the host cell with the help of the MJ.

1.2.2 Structure of apical membrane antigen 1

AMA1 is a structurally conserved, 66 kDa type I integral membrane protein consisting of an *N*-terminal ectoplasmic region, a single transmembrane region and a small *C*-terminal cytoplasmic region [78, 61]. It has eight intramolecular disulphide bonds, which define the three domains in the ectoplasmic region: an *N*-terminal domain I (DI); a central domain II (DII); and a *C*-terminal domain III (DIII) (**Figure 1.3**) [79, 80]. In *P. falciparum*, AMA1 is synthesized as an 83 kDa precursor protein and converted to a 66 kDa protein prior to the invasion process. Around the point of invasion, the 66 kDa AMA1 is further processed to a 48 kDa protein by a subtilisin-like ‘shedase’ PfSUB2 [81].



Figure 1.3. Schematic representation of *P. falciparum* AMA1.

Nuclear magnetic resonance (NMR) solution structures of individual DII and DIII of PfAMA1 [82, 83], and X-ray crystal structures of the PvAMA1 ectodomain and PfAMA1 containing DI and DII (**Figure 1.4**) have been solved previously [84, 80]. These structures revealed that DI (residues 1-308) and DII (residues 309-436) in AMA1 possess a plasminogen-apple-nematode (PAN) motif consisting of five-stranded β -sheets that pack against a 12-residue α -helix [84, 85]. There are many flexible loops (**Figure 1.4**), located predominantly in DI (Ia-If), and a large non-polymorphic loop (351-387) in DII [86]. The middle region of the DII loop is structured and stabilized by interactions with residues from DI. The electron densities for three flexible loop regions (172-GNQYL, 265-KDESKRNSM and 383-GAFK) were not observed in the crystal structure of 3D7 PfAMA1 [84]. The DII loop is displaced upon binding by RON2 as well as the invasion-inhibitory peptide R1 (section 1.2.3.2) [87, 88, 59].

1.2.2.1 Hydrophobic cleft

RON2 binds to a large conserved region on PfAMA1 called the hydrophobic cleft (**Figure 1.4**). In its unbound state the DII loop in PfAMA1 is intimately associated with the hydrophobic cleft (**Figure 1.5A**), but undergoes a large conformational change and is displaced upon PfRON2 binding (**Figure 1.5B**). The hydrophobic cleft consists of 12 hydrophobic amino acid residues: Val169, Leu176, Phe183, Met190, Tyr202, Val208, Met224, Tyr251, Ile252, Met273, Leu357, and Phe367 surrounded by loop regions [65].

Tyr251 is an essential residue for the function of AMA1 [84, 86, 89] and its mutation to alanine abrogates AMA1 binding to RON2 [64, 75]. Met190 is a dimorphic residue while all other cleft residues are highly conserved across different strains of *Plasmodium* (Appendix I). There are five highly polymorphic residues in *Pf*AMA1, at positions 187, 197, 200, 230 and 243. Residues at positions 187, 197, 200 and 230 from DI surround one end of the conserved hydrophobic cleft and have been shown to be important in mediating immune evasion from naturally-acquired and vaccine-induced anti-AMA1 antibodies [90-92, 56]. The other end of the cleft is highly conserved and is formed by interactions between DI and DII loop.

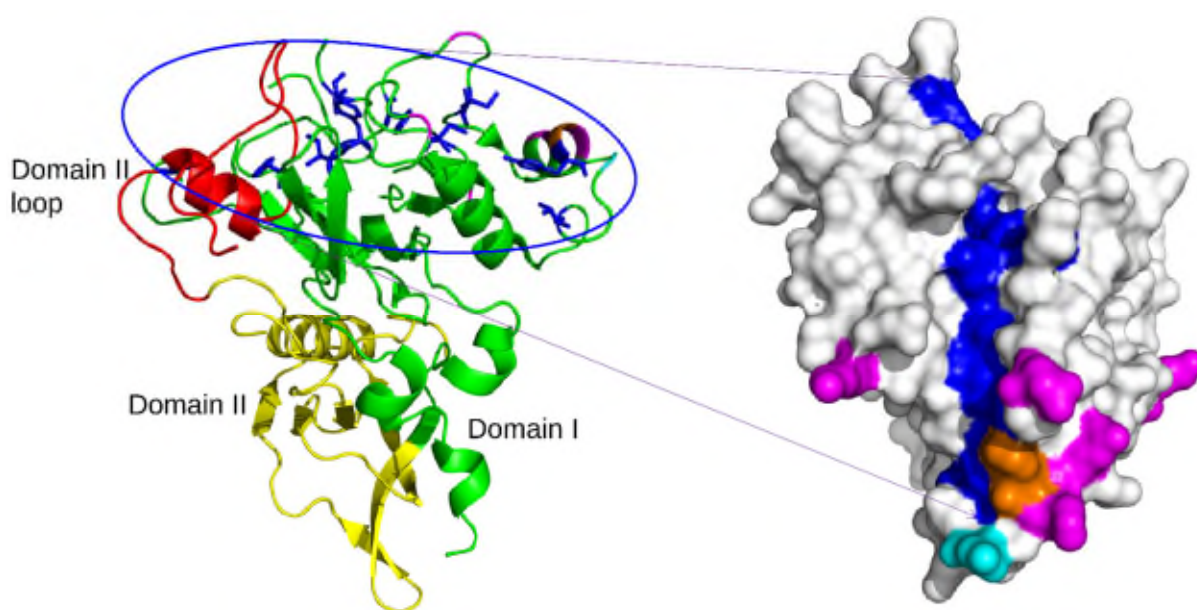


Figure 1.4: Crystal structure of 3D7 *Pf*AMA1 (DI+DII) (PDB ID: 1Z40) [84]. DI (green), DII (yellow), DII loop (red), highly polymorphic residues (magenta), less polymorphic residues (orange), dimorphic residue (cyan) and the hydrophobic cleft is highlighted in blue.

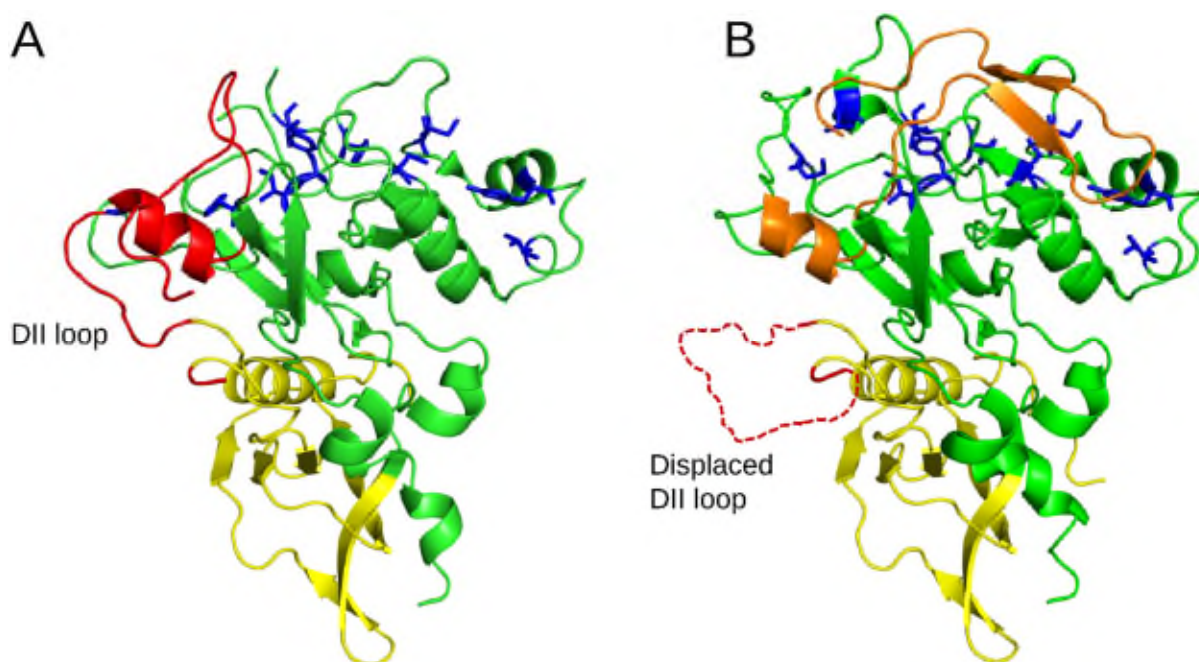


Figure 1.5. Displacement of *Pf*AMA1 DII loop upon *Pf*RON2 binding. A) *Pf*AMA1 with ordered DII loop (PDB ID 1Z40) [84], B) DII loop is displaced by binding to *Pf*RON2 (PDB ID 3ZWZ) [59]. DI (green), DII (yellow), hydrophobic cleft (blue), DII loop (red) and RON2 (orange).

1.2.3 Invasion inhibitory molecules

1.2.3.1 Anti-AMA1 antibodies

Antibodies targeting the hydrophobic cleft on AMA1 are able to block the invasion process [93, 64, 65, 52, 94]. These antibodies, however, exhibit considerable strain-specificity because of extensive polymorphism that exists in AMA1 [95, 65, 52]. For example, 1F9 interacts specifically with an epitope on 3D7 *Pf*AMA1 that includes residues in the hydrophobic cleft as well as highly polymorphic Glu197 and His200, less polymorphic Phe201, and dimorphic Asp204 on loops surrounding the cleft [96, 95] (**Figure 1.6A**). Mutation of these residues, particularly Glu197, abolish 1F9 binding [96]. Similarly, a shark monoclonal antibody (mAb) IgNAR (**Figure 1.6B**) specifically inhibits invasion of erythrocytes by the 3D7 strain of *P. falciparum* but not the W2mef or HB3 strains [65, 94]. The limited cross-reactivity of these antibodies is the result of their binding to the polymorphic end of the cleft. In contrast, 4G2 antibody, which binds to the other end of the hydrophobic cleft that is devoid of polymorphic residues, inhibits merozoite invasion in a strain-independent manner [93, 64, 52]. These antibodies can inhibit the erythrocyte invasion by binding to the hydrophobic cleft on *Pf*AMA1 and have helped establish *Pf*AMA1 as a drug and vaccine target.

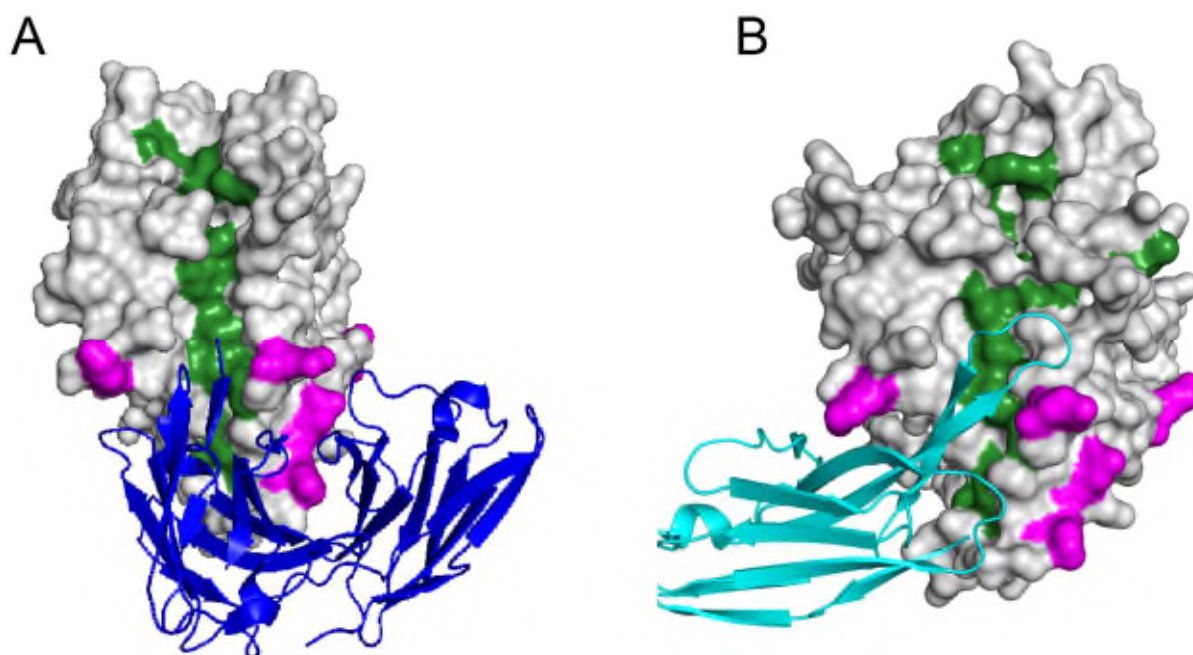


Figure 1.6: Co-crystal structures of (A) *Pf*AMA1-1F9 (PDB ID: 2Q8A) [96] and (B) *Pf*AMA1-IgNAR (PDB ID: 1VES) [65] complexes. Hydrophobic cleft is coloured green, polymorphic residues are coloured magenta, 1F9 coloured blue and IgNAR coloured cyan.

1.2.3.2 Invasion inhibitory peptides

Invasion inhibitory peptides F1 and R1 were identified by screening a random peptide library displayed on the surface of phage [88, 97, 98]. Both F1 and R1 bind strain-specifically to a conformational epitope on the 3D7 *Pf*AMA1 hydrophobic cleft, with affinities of 5 μ M and 100 nM affinity, respectively [88]. It was also reported that F1 was able to compete with 4G2 for binding with the 3D7 *Pf*AMA1 while R1 competed with 1F9 and 4G2 mAbs, in which the binding epitopes lie at the opposite ends of the hydrophobic cleft of AMA1 [98, 52, 99]. These observations for R1 binding are further supported by solution NMR studies of *Pf*AMA1 with methionine methyl groups as probes, which showed that R1 occupies the full length of the hydrophobic cleft [99]. Surprisingly, the crystal structure of R1 in complex with 3D7 *Pf*AMA1 shows that two molecules of R1 occupy the cleft; R1-major, which binds deeply in the hydrophobic groove, and R1-minor, which lies above the R1-major (**Figure 1.7**) [59]. Overall, the binding poses of R1 and *Pf*RON2 peptides were similar (**Figure 1.7**). Further investigation into the crystal structure of the complex reveals that Arg15 of R1-major makes most contacts with 3D7 *Pf*AMA1 and is similar to the contacts made by Arg2041 from *Pf*RON2 [59]. Polymorphic residues at positions 175 and 225 are critical for R1 binding and, thus, limit its reactivity with other alleles of *Pf*AMA1 [59]. NMR studies, however, revealed that only one R1 molecule bound to 3D7 *Pf*AMA1, suggesting that the observed

R1-minor in the crystal structure of the R1 and 3D7 *Pf*AMA1 complex is probably just a crystallization artifact [100].

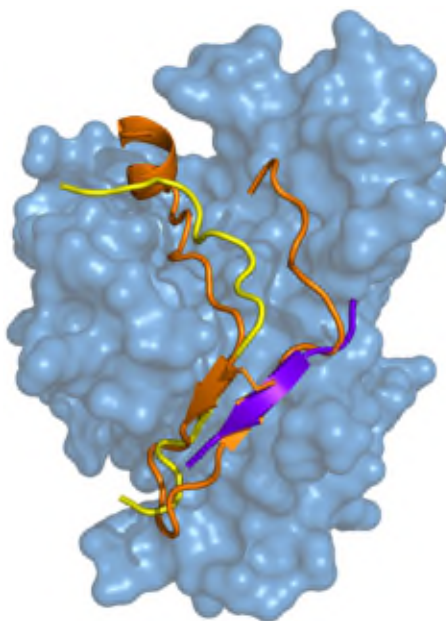


Figure 1.7: Crystal structures of 3D7 *Pf*AMA1 (sky-blue) complexed with *Pf*RON2 (orange) (PDB ID: 3ZWZ) [59] and R1-major (yellow) and R1-minor (purple) (PDB ID: 3SRJ) [59] peptides.

1.2.4 AMA1 as a therapeutic target

AMA1 is an important target for antimalarials as erythrocyte invasion by parasites can be inhibited by blocking the AMA1-RON2 interaction in the presence of inhibitors that target the AMA1 hydrophobic cleft [96, 95, 93, 64, 101, 88]. As AMA1 is expressed and acts on the parasite surface before and during the invasion process, respectively, delivering AMA1 inhibitors through a parasite membrane, which can be very challenging, is not required. Furthermore, the issue of development of drug-resistance due to parasite-mediated drug-efflux mechanisms can be avoided [102]. Moreover, no homologues of AMA1 exist in the human host, making it possible to design inhibitors that can selectively target the *Plasmodium* parasites. In addition, AMA1 has also been reported to be involved in the invasion of hepatocytes [51]. Therefore, a therapeutic drug targeting AMA1 can not only suppress the development of symptomatic malaria, but may also prevent liver-stage infections.

1.2.5 Small molecule inhibitors of AMA1

Small molecule inhibitors of AMA1 offer a basis for developing a new class of antimalarials to treat malaria [103, 104]. Similar to peptide-based *Pf*AMA1 inhibitors, it has been reported recently that small molecules targeting and disrupting the *Pf*AMA1-*Pf*RON2 interaction were able to inhibit erythrocyte invasion [105, 104]. Screening of 21,733 compounds using an AlphaScreen assay led to the identification of three compounds that were able to block invasion *in vitro*, with IC₅₀ values in the range of 21-29 μ M [104]. Subsequent investigation of analogues of one of these three compounds showed strain-transcending inhibitory activity with improved IC₅₀ values of 6 to 9.8 μ M [104]. However, a recent study from our group found that these compounds have low affinity for *Pf*AMA1 with lower invasion inhibitory activity compared to the reported results and the compounds also suffered from unattractive qualities, such as aggregation. The unfavorable physicochemical properties of these compounds impede them as lead *Pf*AMA1 inhibitors [105].

Previous work by Lim *et al.* [103] from our group has also identified small molecules that bind to the hydrophobic cleft. A fragment library containing 1140 small molecules were screened using surface plasmon resonance (SPR), saturation transfer difference (STD) and Carr-Purcell-Meiboom-Gill (CPMG) NMR experiments with R1 competition to identify small molecules that target the *Pf*AMA1 cleft. It was suggested that small molecules that are able to compete with R1 for binding to *Pf*AMA1 may be useful as building blocks for further development into potent AMA1 inhibitors [103].

1.2.6 Fragment-based drug design

Target-based drug design is an efficient approach to identify novel inhibitory molecules for therapeutic protein targets [106]. This approach is based on the screening of low complexity fragments to identify low μ M binders. The activity of the hit compounds identified from initial screening is further improved by chemical modifications guided by structure-activity relationship (SAR) with their target protein to produce potential drug-like molecules for clinical evaluation. Pharmaceutical companies use high-throughput screening (HTS) methods to screen thousands to millions of compounds to identify low μ M (K_D) compounds for targeting a particular disease or symptom [107, 108]. However, designing an automated system for screening such a large library of compounds by HTS method remains a challenge [108]. Additionally, an extra validation procedure in tandem to HTS is required as the chances of getting false positive hits with HTS are significantly high [109].

An alternative method for hit identification is fragment-based lead design (FBLD), which has recently gained momentum in both pharmaceutical companies as well as academic institutions in discovering novel drug candidates [110, 111, 107, 112, 113]. The number of compounds in a library used for screening using this approach is typically in the range of hundreds to a few thousand [114, 115]. Despite their low affinities, fragments have higher ligand efficiency (LE) compared to HTS hits as a high proportion of their atoms are involved in interaction with the target protein. The affinity of these small molecules for their target protein can be subsequently improved by either fragment linking or the addition of extra functional groups.

Solution NMR experiments like CPMG and STD are routinely used to detect binding hits [114, 116]. However, specific interactions between ligands and protein target cannot be determined by ligand-detected experiments such as STD and CPMG. This leads to misinterpretation of the binding data. Therefore, it is necessary to map the binding sites of the fragments on the target protein. X-ray crystallography or proton-detected two-dimensional (2D) NMR experiments such as ^{15}N - and ^{13}C -HSQC are the preferred techniques to determine the binding pose of fragments on a target protein. In NMR, chemical shift perturbations in the spectrum of the target protein are generally used to identify the fragment binding sites.

1.2.7 AMA1 as a vaccine candidate

In addition to being a drug target, *Pf*AMA1 is also an important candidate antigen for inclusion in a malaria vaccine. Anti-AMA1 antibodies are able to inhibit erythrocyte invasion *in vitro* (Section 1.2.3.1) and are associated with protection in humans [117]. Correctly-folded AMA1 ectodomain with intact disulphide bonds is essential to induce a functional immune response [118-120] and the inhibitory antibodies are mainly specific for DI and DII [121-123].

Phase I clinical trial of the FMP2.1 vaccine containing only 3D7 *Pf*AMA1 showed that it is safe and immunogenic [124]. In a Phase IIa trial, vaccine formulated in AS01B and AS02A adjuvants produced a functional immune response as suggested by growth inhibition assay *in vitro* [125]. Nonetheless, Phase II studies with this vaccine revealed problems of restricted strain specificity, as the vaccine contained only 3D7 *Pf*AMA1 [38, 125], and similar results were observed in a field trial [41]. Likewise, only partial protection was observed in rodent models immunized with a single form of AMA1 from *P. chabaudi* [119]. The restricted strain-specificity was due to the polymorphic nature of AMA1 [84, 126-129].

Furthermore, FMP2.1 vaccine formulated in AS01 was tested in controlled human malaria infection (CHMI) models and demonstrated no efficacy [130]. However, the purified antibodies showed functional growth inhibitory activity against 3D7 *P. falciparum* [130].

To overcome the problem of polymorphism, Malkin *et al.* [131] designed the AMA1 vaccine, AMA1-combination 1 (AMA1-C1), which contains equal amounts of 3D7 and FVO *PfAMA1* formulated in Alhydrogel. A phase I study of this formulation in semi-immune adults produced higher antibody titres that showed activity in growth inhibition assays *in vitro* [131]. Additionally, a Phase I study of this vaccine in 2-3 year old children suggested that the vaccine is safe and immunogenic [132]. Nevertheless, Phase II trials of this vaccine in Malian children of the same age showed no impact on primary (rate of *P. falciparum* >3000/μL/day at risk) and secondary (frequency of clinical episodes per time at risk) biological endpoints [133]. Therefore an immune enhancer, CpG 7909, was added to the AMA1/Alhydrogel formulation to improve the vaccine immunogenicity and functionality of the antibodies. [134-136]. Its efficacy was tested in malaria-naïve adults challenged with infected blood [137] and it was found that the vaccine is safe and immunogenic [138]. However, no significant relationship was observed between vaccine-induced growth inhibition activity (GIA) and parasite multiplication rate in vaccinated subjects [138]. Another Phase II trial with AMA1-C1 (formulated with aluminium hydroxide) in Malian children showed no evidence of vaccine selection or strain-specific efficacy [139].

Immunization of rabbits with three AMA1 alleles produced a broad functional immune response to a small panel of laboratory stains [140]. Inclusion of several strains resulted in dilution of relative strain-specific responses and increased the response of cross-reactive Abs to both conserved as well as polymorphic epitopes termed strain-specific epitope dilution [141, 142]. Another study in which four alleles of AMA1 were included was reported in which the immune response to these four alleles was able to inhibit 22 non-laboratory strains of *P. falciparum* [142].

Recently, Miura *et al.* [143] reported the advantage of immunization with a mixture of allelic forms of AMA1 over immunization with a single allelic form. In these experiments, a group of rabbits was immunized with a single form of AMA1, whereas another group of rabbits was immunized with a mixture of four (3D7, CAMP, HP47 and M24), five (3D7, CAMP, HP47, M24 and L32) and six (3D7, CAMP, HP47, M24, L32 and HP22) different allelic forms of AMA1. In functional assays, the antibodies raised against multiple AMA1 alleles inhibited parasite growth significantly compared to antibodies raised against a single AMA1 allele [143].

A novel approach in which the efficacy of anti-AMA1 antibodies raised in mice immunized with an AMA1-RON2 functional complex was compared with the immune response raised against the individual antigens. The anti-AMA1 antibodies generated showed improved inhibitory activity [144]. Detailed analysis suggests that a switch in the proportion of antibodies against highly polymorphic DId loop to less polymorphic loop regions (DIe and DI f), indicating that the functional complex can generate more potent inhibitory antibodies to AMA1. Thus, the antibodies against less polymorphic regions can be more inhibitory than those against highly polymorphic epitopes [142, 144].

1.3 Merozoite surface protein 2

Merozoite surface protein 2 (MSP2), is the second most abundant merozoite surface coat protein antigen and is a potential component of a malaria vaccine [40, 45, 145-147]. MSP2 is an essential protein for survival of the asexual blood-stages of *P. falciparum* [148]. It is one of the most polymorphic surface antigens, with a central variable region (VR) that comprises 60% of the protein, flanked by conserved *N*-terminal (NTR, 25 residues) and *C*-terminal (CTR, 50 residues) regions. All forms of MSP2 can be classified into two allelic families, 3D7 and FC27, based on the dimorphic sequences within the VR (**Figure 1.8**) [44, 149, 147, 46, 47, 150]. The dimorphic sequences are interspersed by polymorphic tandem repeats, with 3D7 being characterized by highly variable GSA-rich repeats, and the FC27 allele by 32- and 12-mer repeats [151]. The NTR is fully conserved across all strains of malaria, while the CTR region is conserved except for a single point mutation (Appendix II) [44, 47]. MSP2 may help the parasite to interact with the erythrocyte surface during the invasion process [147, 152]. Nevertheless, the exact function of MSP2 is unclear.

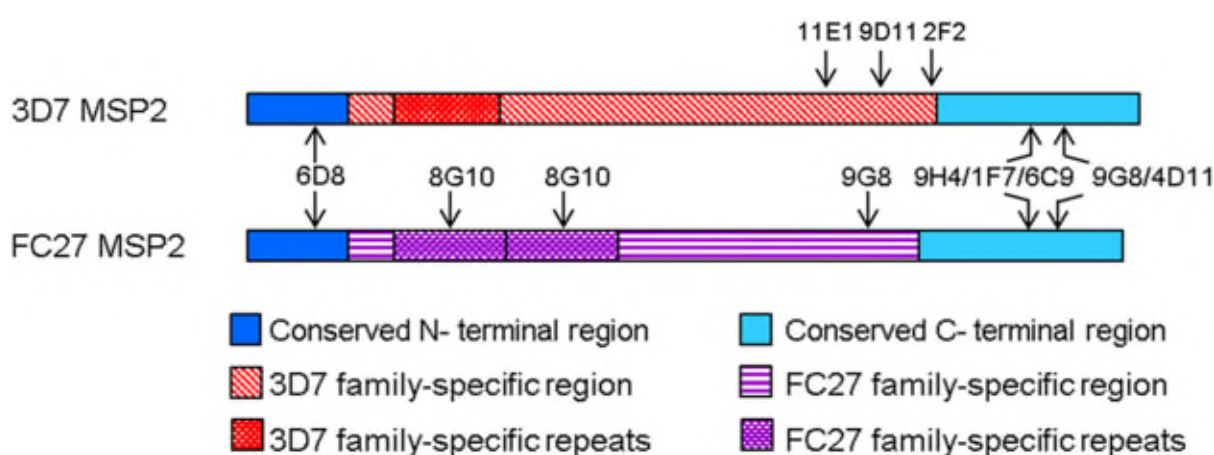


Figure 1.8: Schematic showing the location of mouse mAb epitopes identified by peptide mapping in the two forms of MSP2 alleles [153].

1.3.1 Structure of MSP2

MSP2 of *P. falciparum* is a ~23 kDa antigen that is C-terminally GPI-anchored to the surface of the merozoite [45, 47]. It is an intrinsically disordered protein (IDP) which is also highly hydrophilic and lacks defined secondary structure [151]. MSP2 forms amyloid-like fibrils under physiological conditions due to the aggregation propensity of the NTR [154-158]. As MSP2 is an intrinsically disordered antigen, its conformational and antigenic properties are defined by local sequence features [159, 160, 151]. This property is very suitable for making chimeric antigens as ligating two disordered proteins does not affect the conformation and immunogenicity of individual epitope sequences. Additionally, as mentioned above (section 1.1.3), complicated refolding protocols are not required for preparation of MSP2 and it does not require any special storage conditions as it is highly stable even at elevated temperatures (80 °C). NMR studies in the presence of dodecylphosphocholine (DPC) micelles showed that the conserved N-terminal region interacts with lipids and adopts a helical conformation [145]. This suggests that the variable region may be exposed on the merozoite surface while the conserved NTR make interactions directly and CTR makes interactions through a GPI-anchor with the merozoite membrane. (**Figure 1.9**).

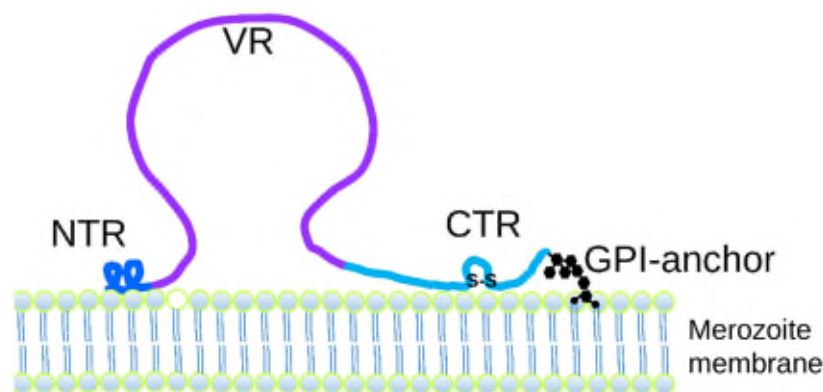


Figure 1.9: A possible model structure of MSP2 on merozoite surface [145].

1.3.2 Immunogenicity of MSP2

Anti-MSP2 antibodies in natural infections are associated with protection in adults and children over age of 5 years [161, 162]. Sero-epidemiological studies also suggest that there is a correlation between the anti-MSP2 immune response and resistance to infections [163-165, 162]. Despite indirect evidence that the anti-MSP2 antibodies to natural infections inhibit the erythrocyte invasion to some extent [166], they were mainly found to function through antibody-dependent cellular inhibition (ADCI), opsonic phagocytosis and complement-mediated mechanisms [167, 168, 146, 169]. The immune response to MSP2 is

mostly directed against polymorphic and dimorphic regions within the variable region [146, 150]. Anti-MSP2 antibodies alone or in combination with antibodies targeting other antigens of *Plasmodium* provide protection in clinical infections [162]. Nevertheless, there is no detailed analysis of the implications of MSP2 disorder for generating an effective immune response.

1.3.3 MSP2 epitopes to mouse mAbs

mAbs to MSP2 were generated in mice to characterize its immunogenic regions. Epitopes for ten different mAbs on both FC27 and 3D7 MSP2 alleles were mapped (**Figure 1.8** and **Table 1.1**) using a peptide array composed of 84 biotinylated 13-residue linear peptides, with an 8-residue overlap, covering the entire 3D7 and FC27 MSP2 sequences [153].

Table 1.1 MSP2 mAbs with corresponding epitope sequences are shown [153].

	mAb	Epitope sequence
NTR specific	6D8	FINNAYNMSIRRS
3D7 specific	11E1	NPKGKGGEVQEPNQ
	9D11	KSNVPPTQDADTK
	2F2	QTESPELQSAPEN
FC27 specific	8G10	IASGSQRSTNSASTS
CTR specific	1F7, 6C9 and 9H4	SQKECTDGNKENC
	4D11	NKENCGAA
	9G8	NKENCGAA

Among ten mAbs, 6D8 recognized an epitope within the conserved *N*-terminal region (preferentially monomeric form of MSP2), while five other mAbs (i.e. 1F7, 4D11, 6C9, 9G8 and 9H4) recognized one of two overlapping epitopes within the conserved *C*-terminal region of recombinant MSP2. 6D8, however, does not react with native MSP2 on the parasite, as suggested by western blot analysis and immunofluorescence assay (IFA). 4D11 and 9G8 react strongly with native MSP2, while 6C9, 1F7 and 9H4 reacted weakly. On the other hand, mAbs 11E1, 9D11 and 2F2 (which bind to the variable region of 3D7 MSP2) and 8G10 (which binds to an epitope within the 32-residue repeat of the FC27 variable region) (**Figure 1.8**), strongly recognize native MSP2. Western blot analysis suggested that all mAbs reacted strongly with recombinant MSP2 despite having differences in reactivity with the native MSP2 [153]. This clearly indicates that MSP2 adopts a different

conformation on the parasite surface and/or that the epitopes from conserved regions may have less accessibility to the immune system on the parasite surface [153, 170]. These observations also imply that the disordered VRs of both MSP2 alleles are exposed and are crucial for a protective immune response, while the conserved regions interact with the parasite membrane (**Figure 1.9**) [153, 171, 40, 145, 172]. As the conformation of native MSP2 is not well characterized, the structures of MSP2 epitopes in complex with their mAbs should be informative, and are expected to contribute to the design of a more robust MSP2-based vaccine.

1.3.4 MSP2 as a vaccine candidate

There is evidence in the literature that supports MSP2 as a potential target for the human immune system and its role in providing protection against *P. falciparum* infections [173, 163, 174, 161, 175, 40, 169]. To illustrate, a Phase I/IIb clinical trial of Combination B vaccine consisting of MSP1, the 3D7 allele of MSP2, and the ring-infected erythrocyte surface antigen (RESA) significantly reduced the parasite densities [40]. However, a higher prevalence of morbidity was noticed in the vaccine group associated with *P. falciparum* infections of the FC27-type MSP2 alleles. These observations suggest that the vaccine effectively suppressed the growth of parasites expressing 3D7 MSP2 but not the parasites expressing FC27 MSP2. This was due to selective pressure on the parasite as the immune system produced antibodies against the polymorphic epitopes of 3D7 MSP2. Therefore, a clinical trial of MSP2-C1, a multivalent vaccine containing recombinant forms of both MSP2 alleles, 3D7 and FC27, in equal amounts formulated with Montanide ISA720 as a water-in-oil emulsion was conducted [146]. The antibody levels were similar for both the MSP2 alleles and exhibited functional activity in *in vitro* experiments including ADCI and complement-mediated assays [167, 146]. Vaccination was stopped after two doses even though the vaccine was sufficiently immunogenic. This was a consequence of adjuvant-related adverse effects in a few subjects at the injection site. Thus, the vaccine formulation needs to be optimized to reduce its side effects before proceeding to further vaccine trials [146]. Sequence polymorphism is a problem for using MSP2 as a vaccine candidate. A chimeric vaccine approach is an alternative to a multivalent vaccine to overcome the restricted strain-specificity of the immune response.

1.3.5 Chimeric malaria antigens as vaccines

Chimeric proteins encompassing different isotypes of a polymorphic antigen would be useful in overcoming the strain-specific immune response and also have other advantages as explained here. Higher immune responses and product yields were observed for the chimeric vaccine, *Pf*CP2.9 (which comprises DIII of AMA1 and a 19 kDa C-terminal fragment of the MSP1) compared to vaccines containing the individual antigens [176]. Sera from animals immunized with this chimeric vaccine completely inhibited parasite growth under *in vitro* conditions [176]. A phase I clinical trial of *Pf*CP-2.9 showed higher antibody responses, which remained elevated over the course of the study [177]. However, the authors found that these antibodies reacted poorly in parasite GIA. Hence, they concluded that *Pf*CP-2.9 was safe and immunogenic but that antibodies raised against AMA1 DIII might not be as functional as antibodies raised against the intact protein [177]. Similarly, in a separate study, the antibody responses were improved with a chimeric vaccine containing MSP1 and MSP8 antigens compared to vaccine containing either one of these antigens [178]. Thus, there is evidence that immunogenic properties can be improved irrespective of antigenic polymorphism among different strains by using chimeric antigens as a vaccine components [179, 177, 178]. Compared with other chimeric antigens and existing MSP2 vaccine formulations, an MSP2 chimeric antigen has several advantages, such as cheaper and more rapid production since there are no refolding requirements with MSP2 being an intrinsically disordered protein. Additionally, as there are only two major alleles of MSP2, it is easy to accommodate them in a single construct.

1.4 Project Scope and Aims

*Pf*AMA1 plays an important role in the invasion of erythrocytes. Molecules blocking the AMA1-RON2 interaction by targeting the AMA1 hydrophobic cleft inhibit invasion, suggesting the potential of *Pf*AMA1 as a novel drug target. Compared to the inhibitory peptides and mAbs, small molecule inhibitors have advantages that include lower production costs and suitability for oral administration [180, 181]. Previously, our group has identified a range of scaffolds that bind to the *Pf*AMA1 hydrophobic cleft as defined by competition with the R1 peptide which is known to bind this cleft [103]. It is, however, not clear whether these compounds interact with polymorphic residues or conserved residues surrounding the cleft. This is an important question as compounds that interact with polymorphic residues are not good starting points for developing strain-transcending inhibitors of *Pf*AMA1.

Therefore, it is critical to determine the specific interactions of these scaffolds on *Pf*AMA1 before proceeding to their chemical elaboration in the lead discovery pipeline. This is a basis for developing potential *Pf*AMA1 inhibitors as new antimalarials.

To determine the critical interactions between *Pf*AMA1 and small molecules, the present work focuses on mapping the binding sites of small molecules on *Pf*AMA1 using two-dimensional [^1H - ^{15}N]-TROSY perturbation experiments. To map the small molecule binding sites on *Pf*AMA1 using NMR spectroscopy, sequence-specific resonance assignments of *Pf*AMA1 were first determined. A series of TROSY based triple-resonance NMR experiments acquired on triple-labelled (^2H , ^{13}C and ^{15}N) AMA1 were used to obtain the sequential assignments. Two-dimensional [^1H - ^{15}N]-TROSY spectra recorded on 3D7 *Pf*AMA1 and FVO *Pf*AMA1 in the presence and absence of small molecules were used to map the binding sites. These results are presented in Chapter 2.

The first crystal structure of *Pf*AMA1 from the FVO strain was solved and compared with the crystal structures of the 3D7 *Pf*AMA1 and *Pv*AMA1. The flexibility of the loop regions in DI and DII was investigated using a combination of molecular dynamics simulations and normalized B-factor analysis in an effort to understand their role in determining the strain-specificity of human immune responses and inhibitory peptides. This structural information together with other existing crystal structures provided valuable insights in designing a strain-transcending inhibitors targeting *Pf*AMA1. The results are presented in Chapter 3.

As described in section 1.1.3, drugs are limited to treating the disease, hence, an efficient vaccine is required for complete eradication of malaria. This work is particularly focused on an intrinsically disordered merozoite surface antigen MSP2, as a vaccine candidate. The immune response against MSP2 has been associated with protection, albeit in a strain-specific manner. Nonetheless, the effect of MSP2 disorder in generating a protective immune response is unknown. Therefore, the aim of this study is to understand the correlation between MSP2 disorder and its immunogenicity, and subsequently design and develop an MSP2-based construct that can generate a strain-transcending immune response to all strains of *P. falciparum*. Furthermore, this work is extended to determine the structure of antibody-bound MSP2 epitopes in an effort to understand the native MSP2 conformation.

Understanding the conformational properties features of MSP2 is important in designing an optimized MSP2 vaccine. Therefore, detailed conformational dynamics of both alleles of MSP2 were characterized using NMR spectroscopy and related to the immunogenicity of different regions of MSP2. It was observed that the highly disordered regions of MSP2

(GGSA repeats) were less immunogenic compared to other regions that were more strongly immunogenic. The results of these studies are presented in Chapter 4.

To achieve a strain-transcending immune response to MSP2, a series of MSP2 chimeric constructs was generated and characterized by various biophysical techniques such as size-exclusion chromatography (SEC), liquid chromatography-mass spectrometry (LC-MS), sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) etc. Mice immunization experiments were carried out for these constructs formulated in Montanide ISA 720, an adjuvant commonly used in human trials [40, 182, 183, 177, 184-187]. All sera were characterized by mapping the epitopes onto the MSP2 constructs and the total IgG and IgG-subtype responses to all the MSP2 constructs used in this study were measured using enzyme linked immune assay (ELISA). Both an improved immune response and a uniform reactivity of the sera against these new chimeras were observed in ELISA. Additionally, total IgG and IgG subtypes in the sera against chimeric constructs were higher compared to IgG levels in the sera raised against 3D7 and FC27 mixture. Removal of repeat sequences had an impact on the immunogenicity to other epitopes nearby. Additionally, a reduced immune response was observed for constructs lacking either NTR or CTR, suggesting the importance of the conserved regions. These results are described in Chapter 5.

The MSP2 *N*-terminal specific antibody 6D8 does not recognise MSP2 on the parasite surface, suggesting a mechanism by which native MSP2 escapes host immune recognition. This result was further supported by SPR binding experiments where 6D8 showed no binding with MSP2 formulated in lipids. Isothermal calorimetry (ITC) experiments suggested that the recombinant 3D7 MSP2 and FC27 MSP2 proteins have different affinities for 6D8, but the crystal structure of 6D8 in complex with its epitope did not explain the different affinities for 3D7 MSP2 and FC27 MSP2. Solution NMR experiments suggest that these affinity differences are due to transient interactions between 6D8 and residues beyond the conserved 6D8 epitope. The crystal structure of 6D8 in complex with its epitope and the preliminary results on affinity differences are discussed in Chapter 6.

Subsequent work on the backbone NMR assignments of 6D8 scFv in the presence and absence of 6D8 minimal epitope, as well as extended epitopes, were carried out. A set of triple-resonance NMR experiments was recorded on triple-labelled 6D8 scFv. About 80% of backbone assignments including C β were completed for both epitope-free and epitope-bound 6D8 scFv. CLEANEX NMR experiments were measured to map the solvent-exposed residues. These results are described in Chapter 7.

Finally, all major findings from the work described in this thesis (including some preliminary findings outlined in Appendices III and IV) as well as potential future directions of these projects are summarized in Chapter 8.

CHAPTER 2

Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials

2.1 Declaration of Thesis Chapter 2

Declaration by candidate

In the case of **Chapter 2**, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Designed the experiments, prepared ^2H , ^{15}N , ^{13}C labelled AMA1 samples, performed NMR experiments, determined NMR assignments, performed AMA1 binding experiments with small molecules, analysed data and prepared manuscript	50

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
San Sui Lim	Performed experiments, manuscript preparation	
Shane M. Devine	Fragment synthesis, manuscript preparation	
Cael O. Debono	Fragment synthesis	2.5
Raymond Lam	Fragment synthesis	2.5
Indu R. Chandrashekar	Intellectual input	
Garima Jaipuria	Intellectual input	
Hiromasa Yagi	Intellectual input	
Hanudatta S. Atreya	Intellectual input	
Martin J. Scanlon	Intellectual input	
Christopher A. MacRaid	Intellectual input	
Peter J. Scammells	Intellectual input	
Raymond S. Norton	Intellectual input, manuscript preparation	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

**Candidate's
Signature**

	Date 30/03/2016
---	---------------------------

**Main
Supervisor's
Signature**

	Date 30/03/2016
---	---------------------------

2.2 Introduction

The development and spread of drug-resistant malaria parasites are limiting the efficacy of existing drugs. For this reason, it is imperative to discover and develop next generation drugs with new modes of action. *PfAMA1-PfRON2* interaction is essential for the formation of the moving junction during erythrocyte invasion. *PfRON2* binds to a conserved hydrophobic cleft on *PfAMA1* and inhibitory molecules targeting the cleft are able to block invasion. Thus *PfAMA1* is a drug candidate for the development of strain-transcending antimalarials. The overall goal of this work was mapping the binding sites for small molecules as a basis for further chemical elaboration and development of potential inhibitors for *PfAMA1*. Solution NMR spectroscopy was used as a tool for studying AMA1-small molecule interactions. Sequential ^1H , ^{13}C and ^{15}N NMR assignments were determined for *PfAMA1* from 3D7 and FVO strains and then used to map the binding sites for small molecules including benzimidazoles, pyrazoles and 2-aminothiozoles. It was found that the benzimidazoles are promising scaffolds for the development of more potent *PfAMA1* inhibitors. As the results of this chapter have been published in the *Journal of Molecular Recognition*, they are presented in the format of a published article in the immediate section below.

Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials

Bankala Krishnarjuna^{a†}, San Sui Lim^{a†}, Shane M. Devine^a, Cael O. Debono^a, Raymond Lam^a, Indu R. Chandrashekar^a, Garima Jaipuria^b, Hiromasa Yagi^a, Hanudatta S. Atreya^b, Martin J. Scanlon^a, Christopher A. MacRaid^a, Peter J. Scammells^a and Raymond S. Norton^{a*}



Plasmodium falciparum apical membrane antigen 1 (PfAMA1) plays an important role in the invasion by merozoites of human red blood cells during a malaria infection. A key region of PfAMA1 is a conserved hydrophobic cleft formed by 12 hydrophobic residues. As anti-apical membrane antigen 1 antibodies and other inhibitory molecules that target this hydrophobic cleft are able to block the invasion process, PfAMA1 is an attractive target for the development of strain-transcending antimalarial agents. As solution nuclear magnetic resonance spectroscopy is a valuable technique for the rapid characterization of protein–ligand interactions, we have determined the sequence-specific backbone assignments for PfAMA1 from two *P. falciparum* strains, FVO and 3D7. Both selective labelling and unlabelling strategies were used to complement triple-resonance experiments in order to facilitate the assignment process. We have then used these assignments for mapping the binding sites for small molecules, including benzimidazoles, pyrazoles and 2-aminothiazoles, which were selected on the basis of their affinities measured from surface plasmon resonance binding experiments. Among the compounds tested, benzimidazoles showed binding to a similar region on both FVO and 3D7 PfAMA1, suggesting that these compounds are promising scaffolds for the development of novel PfAMA1 inhibitors. Copyright © 2016 John Wiley & Sons, Ltd.

Additional supporting information may be found in the online version of this article at the publisher's web site.

Keywords: AMA1; NMR; resonance assignments; isotopic labelling; SPR; fragments

INTRODUCTION

Malaria is a significant global health burden, causing nearly 600 000 deaths per annum (WHO, 2014). The causative agents are *Plasmodium* parasites, with the majority of the disease burden arising from infection by *Plasmodium falciparum*. Importantly, this parasite has started to develop resistance to available antimalarial drugs (White, 2004), including first-line therapy artemisinin combination therapies (Bremar, 2012). Hence, there is a pressing need to develop new antimalarial agents targeting resistant strains of *P. falciparum* parasites. A key step in the invasion by *Plasmodium* parasites of host red blood cells during the asexual blood stage of malaria is the formation of a highly conserved functional complex called the moving junction (Aikawa *et al.*, 1978). Two parasite proteins, *P. falciparum* apical membrane antigen 1 (PfAMA1) and *P. falciparum* rhoptry neck protein 2 (PfRON2), play a crucial role in the formation of the moving junction (Lamarque *et al.*, 2011; Srinivasan *et al.*, 2011; Tonkin *et al.*, 2011; Vulliez-Le Normand *et al.*, 2012). During host cell invasion, PfRON2 is secreted onto the red blood cell membrane and forms strong interactions with PfAMA1 expressed on the parasite surface membrane (Vulliez-Le Normand *et al.*, 2012). Both PfAMA1 and PfRON2 are essential for the invasion process, and hence, small molecule inhibitors

targeting this complex have considerable potential to be used as new antimalarials to combat resistant *P. falciparum* parasites (MacRaid *et al.*, 2011).

Plasmodium falciparum apical membrane antigen 1 is a type I integral membrane protein with three extracellular domains (DI + II + III). PfAMA1 has a hydrophobic cleft, which is the site of interactions with PfRON2 (Lamarque *et al.*, 2011; Srinivasan *et al.*, 2011; Tonkin *et al.*, 2011; Vulliez-Le Normand *et al.*, 2012). The opposing ends of the cleft possess different properties that play

* Correspondence to: R. S. Norton, Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, Victoria 3052, Australia.
E-mail: ray.norton@monash.edu

† These authors contributed equally.

a B. Krishnarjuna, S. S. Lim, S. M. Devine, C. O. Debono, R. Lam, I. R. Chandrashekar, H. Yagi, M. J. Scanlon, C. A. MacRaid, P. J. Scammells, R. S. Norton
Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, Victoria, 3052, Australia

b G. Jaipuria, H. S. Atreya
NMR Research Centre, Indian Institute of Science, Bangalore 560012, India

distinct functional roles in the complex formation (Bai *et al.*, 2005). At one end, there is a conserved domain II loop (DII loop) that undergoes conformational changes upon PfPRON2 binding. The monoclonal antibody 4G2, which binds to the base of this DII loop, exhibits strain-transcending inhibitory effects on red cell invasion by *P. falciparum* parasites (Collins *et al.*, 2007). At the opposite end, there are numerous polymorphic residues that surround the PfAMA1 hydrophobic cleft (Bai *et al.*, 2005; Lim *et al.*, 2014). These polymorphic residues are important in mediating immune escape from natural antibodies in the human host, and they play important roles in limiting the cross-strain inhibitory activities of ligands such as R1 (Harris *et al.*, 2005; Harris *et al.*, 2009) and F1 (Li *et al.*, 2002; Keizer *et al.*, 2003) peptides, as well as the monoclonal antibody 1 F9 (Coley *et al.*, 2006; Coley *et al.*, 2007).

A fragment screening campaign has identified a range of scaffolds that bind to the PfAMA1 hydrophobic cleft as defined by competition with the peptide R1, which is known to bind to this cleft (Harris *et al.*, 2005; Harris *et al.*, 2009; Richard *et al.*, 2010; Lee *et al.*, 2011; Vulliez-Le Normand *et al.*, 2012; Lim *et al.*, 2013; Ge *et al.*, 2014; Wang *et al.*, 2014). Whether these compounds interact with some of the polymorphic residues surrounding the cleft is an important question as compounds that interact with such residues would be considered poor starting points for chemical elaboration given that our goal is to develop strain-transcending inhibitors of PfAMA1. Therefore, it is critical to determine the specific binding sites of the R1-competing fragments early in the lead discovery pipeline. The aim of this work is to use nuclear magnetic resonance (NMR) chemical shift perturbations (CSPs) to identify the residues that are perturbed upon fragment binding and to map them onto relevant PfAMA1 crystal structures to define the binding sites.

MATERIALS AND METHODS

Protein expression and purification

Uniformly ^{15}N -labelled, ^{13}C -labelled, ^2H -labelled FVO PfAMA1_[104–438] and 3D7 PfAMA1_[104–438] proteins were expressed and purified as described previously (Lim *et al.*, 2013). Briefly, the protein was expressed in *Escherichia coli* BL21 (DE3) strain cultured in $^2\text{H}_2\text{O}$ -based M9 medium containing ^{13}C -glucose and $^{15}\text{NH}_4\text{Cl}$ as the sole carbon and nitrogen sources, respectively, using the high-cell-density method (Sivashanmugam *et al.*, 2009). The cells were lysed in 20 mM Tris pH 8.0 containing lysozyme, and the protein was purified as inclusion bodies, followed by solubilization in 6 M Gdn-HCl. The protein was purified under denaturing conditions on a Ni^{2+} -nitrilotriacetic acid (NTA) affinity column and subjected to refolding for 24 h in a buffer containing 20 mM Tris (pH 8.0), 0.5 M urea, 50 mM NaCl, 100 mM reduced glutathione and 25 mM oxidized glutathione. The folded material was purified by anion-exchange chromatography, subjected to tobacco etch virus (TEV) cleavage to remove the 6His tag, and finally purified by anion-exchange chromatography. The fractions containing pure protein according to sodium dodecyl sulfate polyacrylamide gel electrophoresis analysis were lyophilized in 20 mM ammonium bicarbonate buffer and stored at -80°C . Uniformly ^2H -labelled and selectively ^{15}N -Lys-labelled FVO PfAMA1 was prepared by adding 200 mg/l of ^{15}N -labelled Lys (Sigma-Aldrich, St. Louis, MO, USA) into the high-cell-density $^2\text{H}_2\text{O}$ -based M9 medium 1 h prior to isopropyl β -D-1-thiogalactopyranoside (IPTG) induction. Similarly Lys, Arg, Asn and Ile, Leu, Val selectively unlabelled 3D7 PfAMA1 samples were prepared by expressing the protein in $^2\text{H}_2\text{O}$ -based M9 medium containing 1 g/l of the

respective unlabelled amino acids and 1 g/l of $^{15}\text{NH}_4\text{Cl}$ (Krishnarjuna *et al.*, 2011).

Surface plasmon resonance binding experiments

The binding affinities of small molecules including benzimidazoles (MIPS-0001404, MIPS-0000865 and MIPS-0008405), pyrazoles (MIPS-0001160 and MIPS-0001176) and 2-aminothiazoles (MIPS-0000620 and MIPS-0008939) (Figure 1) were evaluated using a BiAcCore T200 biosensor. 3D7 PfAMA1 was immobilized on a CM5 sensor chip using amine coupling to achieve an immobilization level of 10 000 RU. A reference flow cell was prepared in a similar fashion, but in the absence of protein. R1 peptide (GL Biochem, Shanghai, China) was injected over the immobilized surface at a flow rate of 60 $\mu\text{l}/\text{min}$ in an HBS-EP buffer (50 mM HEPES, 150 mM NaCl, 3.4 mM EDTA, 0.05% Tween 20, pH 7.4) containing 5% dimethyl sulfoxide at concentrations ranging from 31.3 to 1000 nM with a contact time of 2 min. Compounds were screened in the same buffer at concentrations ranging from 7 to 250 μM . All sensorgrams were double referenced and solvent corrected prior to analysis. Binding affinities were estimated based on steady-state analysis of the dose–response using a fixed small molecule R_{max} . The small molecule R_{max} was calculated with Eqn 1 using the R_{max} of R1 peptide derived from fitting the binding responses of R1 at steady state to a 1:1 binding model

$$R_{\text{max}}[\text{small molecule}] = R_{\text{max}}[\text{R1 peptide}] \times \left(\frac{\text{MW}_{[\text{small molecule}]}}{\text{MW}_{[\text{R1 peptide}]}} \right)$$

MW = Molecular weight

(1)

Experimental procedures for the synthesis of small molecules used in this work are described in the Supporting Information.

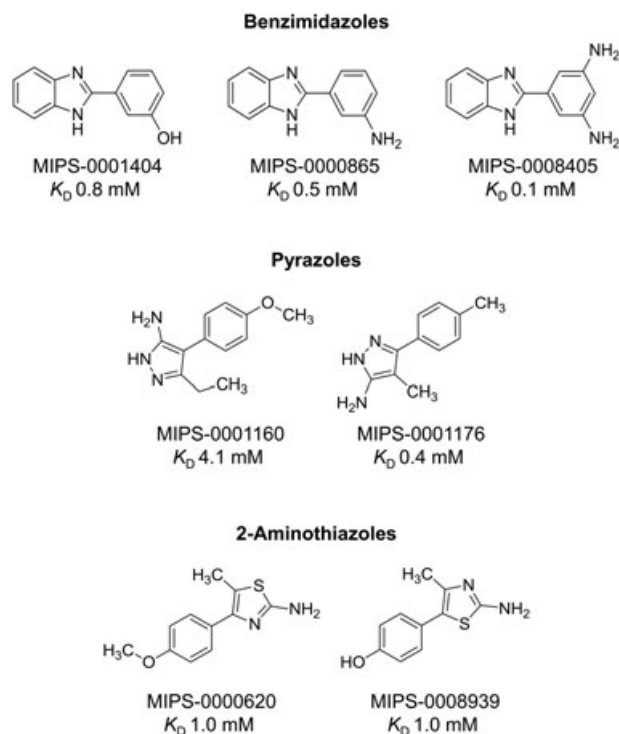


Figure 1. Structures of compounds, with their affinities (K_D) for apical membrane antigen 1 as measured by surface plasmon resonance.

Nuclear magnetic resonance spectroscopy and resonance assignments

Nuclear magnetic resonance spectra for FVO *PfAMA1* was acquired at 35 °C on a Bruker Avance III 600-MHz spectrometer (Billerica, MA, USA) and for 3D7 *PfAMA1* at 30 °C on a Bruker Avance III 800-MHz spectrometer, in both cases with transverse relaxation optimized spectroscopy (TROSY). CSP experiments for FVO and 3D7 *PfAMA1* were carried out on a Bruker Avance III 600-MHz spectrometer at 35 and 30 °C, respectively. The buffer used for acquisition of NMR spectra was 20 mM phosphate buffer, pH 6.8 (FVO) and pH 7.0 (3D7), containing 50 mM L-arginine, 50 mM L-glutamic acid, 0.2% (w/v) protease inhibitors cocktail (Roche, Mannheim, Germany), 0.01% (w/v) sodium azide and 10% (v/v) ²H₂O. Buffer at pH 7.0 was used for 3D7 *PfAMA1* instead of pH 6.8 as this protein precipitates at pH 6.8. Two-dimensional [¹H-¹⁵N]-TROSY spectra were acquired with 32 scans at 2048 and 256 datapoints for the ¹H and ¹⁵N dimensions, respectively. For FVO *PfAMA1*, triple-resonance experiments including HNCA and HN(CO)CA spectra were acquired with a sample concentration of 150 μM and 32 scans, and HN(CA)CB and HN(COCA)CB spectra were acquired with a sample concentration of 250 μM and 64 scans. All 3D NMR experiments were recorded with 2048, 46 and 80 datapoints for ¹H, ¹⁵N and ¹³C dimensions, respectively. For 3D7 *PfAMA1*, HNCA and HN(CO)CA spectra were acquired with 16 and 32 scans, respectively, and HN(CA)CB and HN(COCA)CB spectra were recorded with 24 scans. All 3D NMR experiments for 3D7 *PfAMA1* were recorded with a sample concentration of 250 μM and using non-uniform sampling conditions with 2048, 64, 120 datapoints for the ¹H, ¹⁵N and ¹³C dimensions, respectively. Data for FVO *PfAMA1* were processed using Bruker Topspin 3.2, and non-uniform sampling data for 3D7 *PfAMA1* were reconstructed with compressed sensing using qMDD and analysed using Sparky (v3.114; University of California, San Francisco).

The stability of FVO *PfAMA1* at 35 °C was monitored using 2D [¹H-¹⁵N]-TROSY experiments at days 0, 3 and 6. Similarly, the stability of 3D7 *PfAMA1* at 30 °C was monitored at days 0, 1 and 2. Two-dimensional [¹H-¹⁵N]-TROSY spectra were acquired with eight scans and 1024 and 64 datapoints for the ¹H and ¹⁵N dimensions, respectively, and the spectra were processed as described earlier.

Phase-modulated CLEAN chemical exchange (CLEANEX) TROSY (Hwang *et al.*, 1998) experiments were recorded for 3D7 and FVO *PfAMA1* at 35 °C on a Bruker Avance 600-MHz spectrometer to identify amide protons in rapid chemical exchange with water. The mixing time for the CLEANEX spinlock was set to 45 ms. Spectra were acquired with 128 scans and 2048 and 256 datapoints for the ¹H and ¹⁵N dimensions, respectively. The spectra were recorded on ²H-labelled and ¹⁵N-labelled 3D7 *PfAMA1* and FVO *PfAMA1* samples in the absence and presence of *PfRON2* peptide in 20 mM sodium phosphate buffer, pH 7.4 at sample concentrations of 100 and 70 μM, respectively.

Chemical shift perturbations

Two-dimensional [¹H-¹⁵N]-TROSY spectra of ¹⁵N-labelled and ²H-labelled *PfAMA1* in the presence and absence of small molecules were recorded with 128 scans and 256 increments. The concentrations of *PfAMA1* and small molecule were 25 μM and 2 mM, respectively. The buffer used in these experiments contained 20 mM phosphate buffer, pH 6.8 (FVO) and pH 7.0 (3D7), with 50 mM L-arginine, 50 mM L-glutamic acid

and 10% (v/v) ²H₂O. Weighting of chemical shifts from ¹H and ¹⁵N resonances was performed using the following equation (Williamson, 2013):

$$\text{Weighted CSP } (\Delta\delta) = \sqrt{\frac{1}{2} [\delta_H^2 + (0.14 \cdot \delta_N)^2]} \quad (2)$$

where δ_H and δ_N are the CSP (ppm) at ¹H and ¹⁵N dimensions, respectively. Small molecules were considered to be binding to a specific site on *PfAMA1* based on two criteria: (1) weighted CSP >0.01 ppm and (2) more than one peak perturbed in the spectrum.

RESULTS

Backbone resonance assignments

Both the FVO and 3D7 *PfAMA1* (DI+II) constructs consisted of 335 residues, including 19 proline residues, with molecular mass of around 38 kDa. Biomolecules of this size have fast T₂ relaxation rate with resulting signal broadening. To overcome this issue, a combination of protein deuteration (Gardner & Kay, 1998; Sattler & Fesik, 1996) and the TROSY scheme (Riek *et al.*, 2000; Salzmann *et al.*, 1998) in pulse sequences was used to improve the sensitivity of all heteronuclear NMR experiments (Figure S1). In our previous studies (Lim *et al.*, 2014), we observed 261 (~83%) and 250 (~79%) of the 316 amide backbone peaks expected for FVO (Figure 2) and 3D7 (Figure 3) *PfAMA1*, respectively. Overall, the 2D [¹H-¹⁵N]-TROSY spectra of both proteins showed well-dispersed amide backbone peaks that are consistent with a single folded conformer.

To determine if the protein samples (FVO and 3D7 *PfAMA1*) were suitable for use in 3D NMR experiments with long acquisition times (HNCA, HN(CO)CA, HN(CA)CB and HN(COCA)CB), their stability was assessed using 2D [¹H-¹⁵N]-TROSY experiments. No significant changes of the resonances were observed in the FVO *PfAMA1* 2D [¹H-¹⁵N]-TROSY spectrum over a period of 6 days at 35 °C, indicating that the FVO *PfAMA1* sample possesses good long-term stability (Figure S2). In contrast, it was found that a few new resonances appeared over time in the 3D7 *PfAMA1* spectrum (Figure S2), which could be a consequence of partial unfolding of the protein with time or limited proteolysis (despite the presence of protease inhibitors).

In the FVO *PfAMA1* 3D NMR spectra, totals of 246 Cα_i and 244 Cα_{i-1} resonances were observed in the HNCA and HN(CO)CA, respectively. Owing to the lower sensitivity, fewer signals were detected in the Cβ-detected experiments, with 205 Cβ_i and 202 Cβ_{i-1} signals observed in the HN(CA)CB and HN(COCA)CB, respectively. An example of an assignment strip plot from the 3D HNCA, HN(CO)CA, HN(CA)CB and HN(COCA)CB spectra recorded on FVO *PfAMA1* is shown in Figure S3. Two-dimensional [¹H-¹⁵N]-TROSY spectra recorded on the ¹⁵N-Lys selectively labelled *PfAMA1* were also used to minimize ambiguity in the amide backbone resonance assignments (Figure 4). Lysine was selected because it is one of the most abundant residues in *PfAMA1* that can be selectively labelled without the problem of scrambling to other amino acids during expression in *E. coli* (Muchmore *et al.*, 1989; Goto & Kay, 2000; Ohki & Kainosho, 2008; Krishnarjuna *et al.*, 2011; Jaipuria *et al.*, 2012).

In addition to manual peak assignments, the assignments were also confirmed by the Resonance Assignment by chemical Shift Prediction algorithm (MacRaild & Norton, 2014). A total of

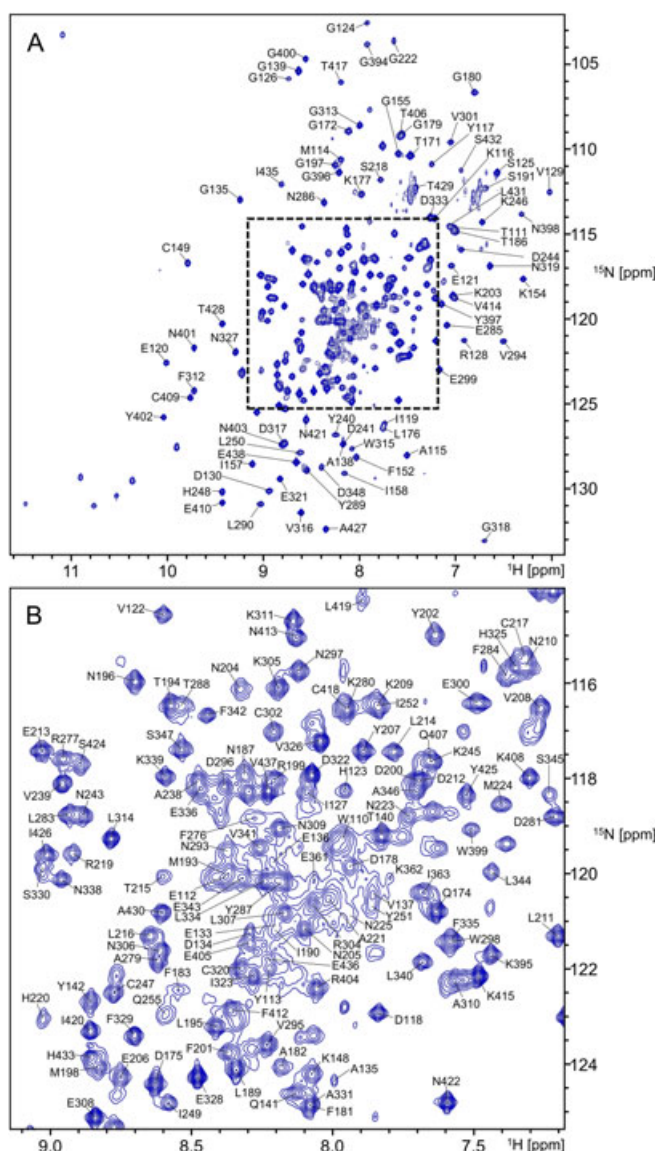


Figure 2. Two-dimensional [^1H - ^{15}N]-transverse relaxation optimized spectroscopy spectrum of FVO *Plasmodium falciparum* apical membrane antigen 1 (DI + DII) with resonance assignments shown. (A) Full spectrum. (B) Enlarged view of the middle region (dotted box in A). The experiments were performed at 35 °C, pH 6.8, on a Bruker Avance III 600-MHz spectrometer.

220 peaks in 2D [^1H - ^{15}N]-TROSY could be assigned for FVO PfAMA1 (BMRB ID 26648), which corresponds to 84% of the observed peaks (Figures 2, 5A and 5B). Notably, eight of the 12 residues in the hydrophobic cleft and 31 of the 40 polymorphic residues, including one of the most important residues that mediate immune escape against antibodies, Gly197, were assigned (Takala *et al.*, 2009; Ouattara *et al.*, 2013). Most of the unassigned resonances (except Arg143 and Val151 in structured regions) correspond to residues in the loop regions (104–108, 131–132, 144, 146–147, 159–167, 169, 227–236, 253–254, 257–259, 261–275, 291–292, 332, 349, 351–360, 365–380 and 382–393) of FVO PfAMA1 (Lim *et al.*, 2014). These unassigned regions include the entire If loop (residues 264–273), most of the le loop (residues 224–238) and the DII loop (residues 351–387) that surround the PfAMA1 hydrophobic cleft (Figure 5B). Assignments could not be obtained in these regions because of weak or missing resonances in the NMR spectra that were probably a

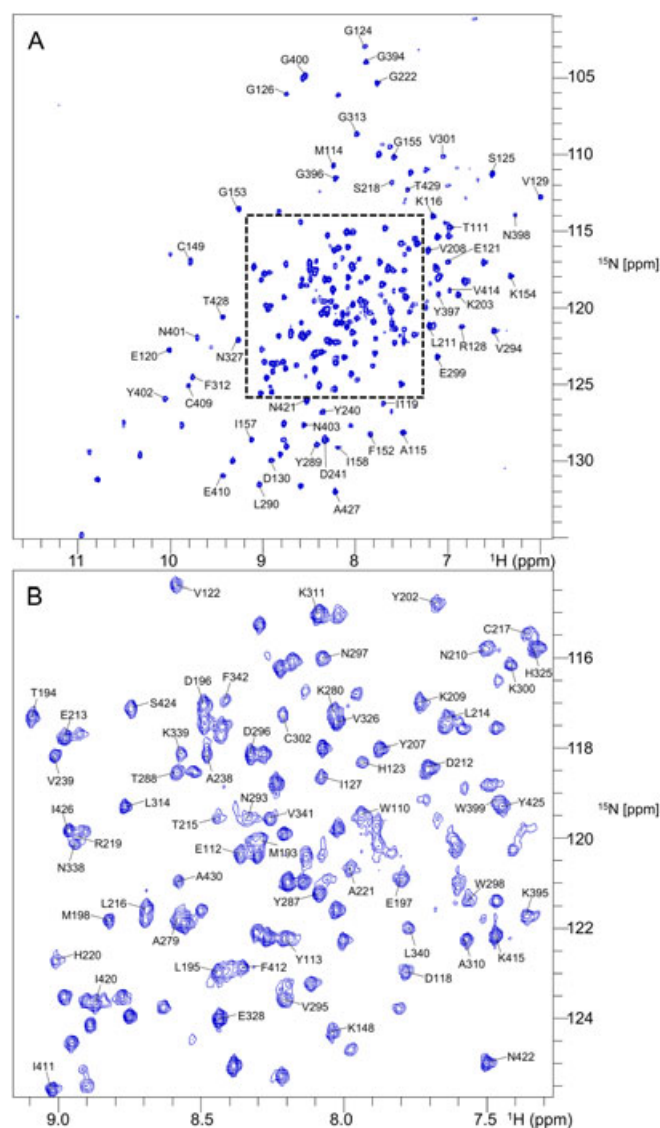


Figure 3. Two-dimensional [^1H - ^{15}N]-transverse relaxation optimized spectroscopy spectrum of 3D7 *Plasmodium falciparum* apical membrane antigen 1 (DI + DII) with resonance assignments shown. (A) Full spectrum. (B) Enlarged view of the middle region (dotted box in A). The experiments were performed at 30 °C, pH 7.0, on a Bruker Avance III 800-MHz spectrometer.

result of peak broadening caused by intermediate conformational exchange or the relatively high amide proton exchange in solution at pH 6.8. Previous studies (Lim *et al.*, 2014) found that the 3D7 PfAMA1 spectrum had a greater number of missing resonances than did the FVO PfAMA1 spectrum owing to broadening by intermediate timescale conformational exchange and amide proton exchange at pH 7 (Figures 2 and 3).

To further probe the number of amide protons in rapid exchange with solvent, CLEANEX experiments (Hwang *et al.*, 1998) were performed on both 3D7 and FVO PfAMA1. More peaks were observed in the CLEANEX spectrum of 3D7 PfAMA1 than in that of FVO PfAMA1 (Figure S4A). In the presence of a peptide corresponding to the apical membrane antigen 1 binding site on PfRON2 (Tonkin *et al.*, 2011), we observed an increase in the number of resonances in the CLEANEX spectra of both 3D7 and FVO PfAMA1 (Figure S4B and S4C). Most of these newly observed resonances are highly overlapped in the centre of the spectrum, suggesting that they may be from amides in the DII loop,

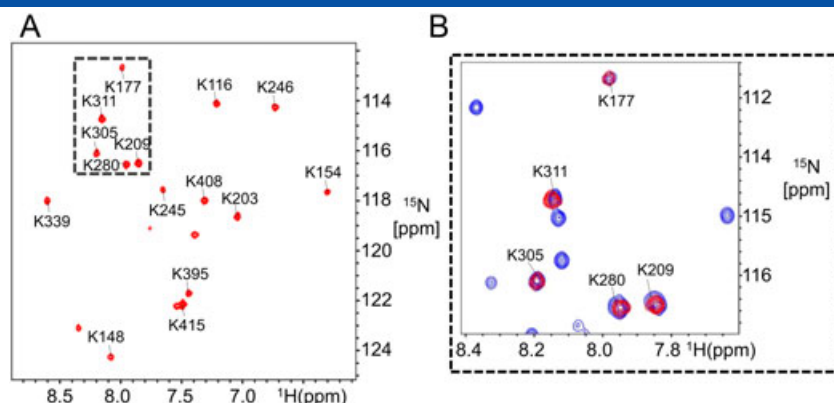


Figure 4. Lysine selective labelling of FVO *Plasmodium falciparum* apical membrane antigen 1 (PfAMA1). (A) 2D $[^1\text{H}-^{15}\text{N}]$ -transverse relaxation optimized spectroscopy spectrum of ^{15}N -Lys selectively labelled FVO PfAMA1. (B) Overlay of the 2D $[^1\text{H}-^{15}\text{N}]$ -transverse relaxation optimized spectroscopy spectra of ^{15}N -Lys selectively labelled (red) and uniformly ^{15}N -labelled (blue) FVO PfAMA1 samples.

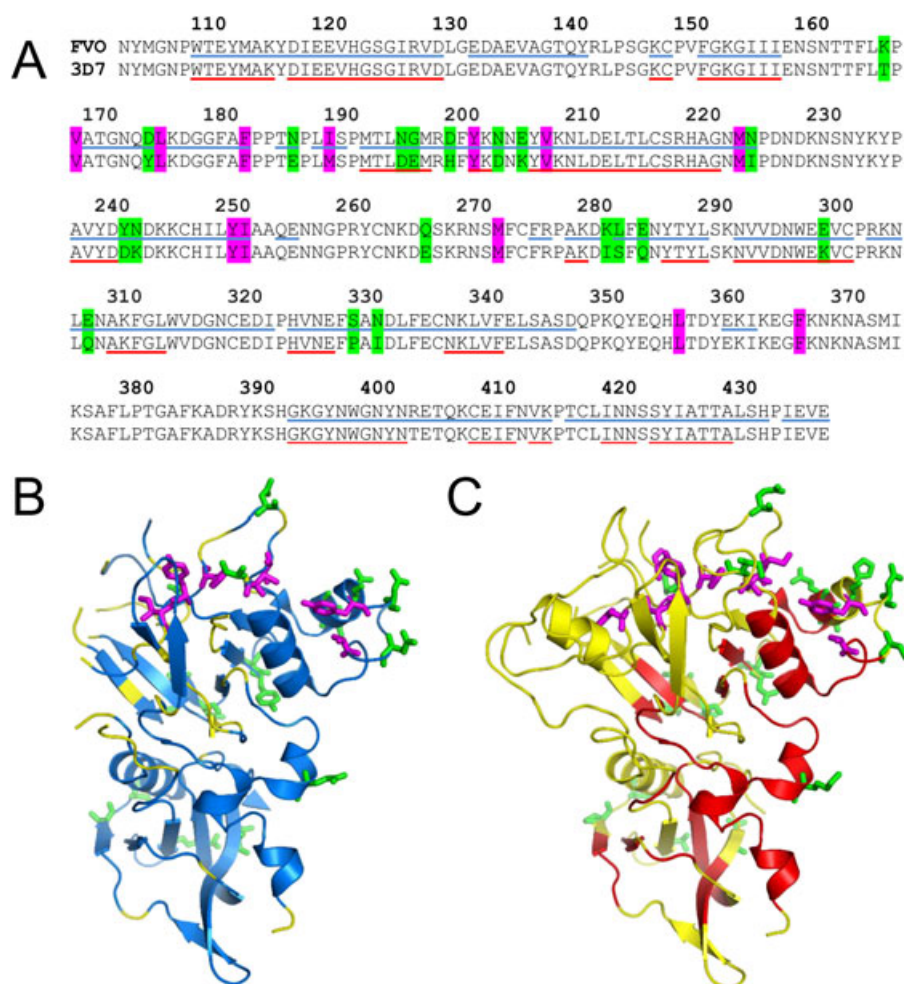


Figure 5. Comparison of nuclear magnetic resonance assignments of FVO *Plasmodium falciparum* apical membrane antigen 1 (PfAMA1) (DI + DII) and 3D7 PfAMA1 (DI + DII). (A) Amino acid sequences of FVO and 3D7 PfAMA1. Residues that line the hydrophobic cleft are coloured purple. Residues that are different between the two forms of PfAMA1 are highlighted in green. Assigned residues on FVO and 3D7 PfAMA1 are underlined in blue and red, respectively. The crystal structures of (B) FVO PfAMA1 (Lim *et al.*, 2014) (PDB ID 4R1A) and (C) 3D7 PfAMA1 (Bai *et al.*, 2005) (PDB ID 1Z40). The residues that line the cleft are shown as purple sticks. Residues that are different between the two forms of PfAMA1 are shown as green sticks. Regions that are not assigned are coloured yellow in both structures. Assigned regions are coloured blue and red for FVO and 3D7 PfAMA1, respectively.

which is displaced by *PfRON2* binding to the hydrophobic cleft to become more flexible and solvent exposed (Ge *et al.*, 2014). Some

peaks observed in CLEANEX are difficult to observe in 2D $[^1\text{H}-^{15}\text{N}]$ -TROSY and other H^{N} -detected triple-resonance NMR experiments

because of their rapid exchange with solvent. These undesirable properties in 3D7 PfAMA1, together with signal overlap and poor stability in solution, resulted in poor-quality triple-resonance NMR spectra, which presented a major hurdle in assigning the amide backbone resonances. To partially overcome this problem, multiple selectively unlabelled 3D7 PfAMA1 (Lys, Arg, Asn and Ile, Leu, Val) samples were prepared, and 2D [^1H - ^{15}N]-TROSY spectra were recorded for each sample (Figure S5) (Krishnarjuna *et al.*, 2011). The unlabelling approach was employed in this case because of the cost of preparing multiple specifically labelled samples.

Residue-specific assignments were obtained by overlaying the spectra of selectively unlabelled 3D7 PfAMA1 samples on that of uniformly ^{15}N -labelled 3D7 PfAMA1. The missing resonances in the spectra of selectively unlabelled samples were assigned to the corresponding amino acid (Figure S5) that was added to M9 medium in an unlabelled (^{14}N) form. These residue-specific assignments were used as starting points to obtain sequential assignments and to confirm resonance assignments from triple-resonance experiments (Figure S6 and Table S1). For 3D7 PfAMA1, totals of 212 $\text{C}\alpha_i$ and 207 $\text{C}\alpha_{i-1}$ signals were observed in HNCA and HN(CO)CA spectra, respectively, and 190 $\text{C}\beta_i$ and 150 $\text{C}\beta_{i-1}$ signals, respectively, in HN(CA)CB and HN(COCA)CB spectra. Although many resonances were not observed in the 3D7 PfAMA1 spectra, 114 (~45% of the observed peaks in 2D [^1H - ^{15}N]-TROSY), including two from residues in the cleft, were assigned (Figures 3, 5A and 5C). The assigned chemical shifts (HN, N, C_α and C_β) for 3D7 PfAMA1 are summarized in Table S1.

Chemical shift perturbations

Selected small molecule classes that displayed some structure–activity relationship in the initial fragment screening campaign (Lim *et al.*, 2013) were chosen for further evaluation in the CSP study. These scaffolds include benzimidazoles, pyrazoles and 2-aminothiazoles. These series bind weakly to 3D7 PfAMA1 with K_D values ranging from 0.1 to 4.1 mM, as determined by surface plasmon resonance (SPR) (Figure 1 and Figure S7). The SPR sensorgrams indicate that all of the compounds bind to PfAMA1 with very fast association and dissociation rates (Figure S7). Amongst the different small molecule classes tested, the benzimidazole series showed higher binding affinities than the pyrazoles and 2-aminothiazoles (Figure 1). Of the compounds tested in 2D [^1H - ^{15}N]-TROSY experiments, all but MIPS-0000620 induced CSPs of more than one backbone amide resonance of FVO PfAMA1 (Table S2). The number of peaks perturbed by MIPS-0000865, MIPS-0008405, MIPS-0001404, MIPS-0001160, MIPS-0001176 and MIPS-0008939 ranges from 5 to 19, as summarized in Table S2.

Resonances perturbed by the benzimidazoles MIPS-0000865, MIPS-0001404 and MIPS-0008405 were mapped onto both ends (DII loop and polymorphic end) of the PfAMA1 hydrophobic cleft (Table S2). Larger numbers and magnitudes of CSPs were found at the DII loop end of the hydrophobic cleft for these compounds, suggesting that this region is the primary binding site (Figures 5B and 6). All resonances (except D175 for MIPS-0008405 and MIPS-0000865) perturbed by these compounds

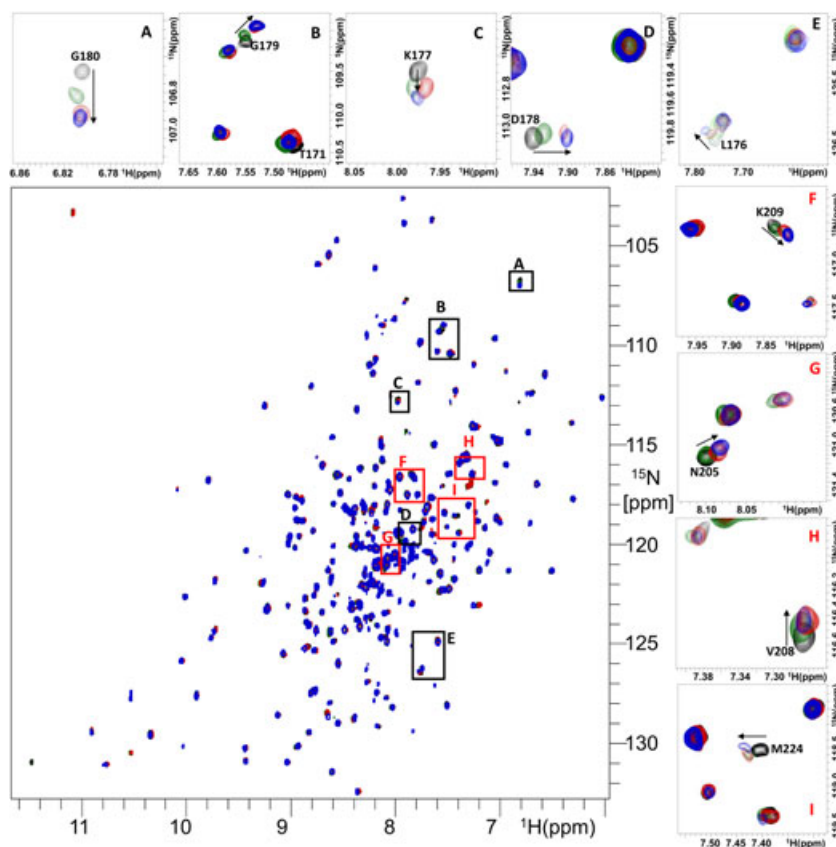


Figure 6. Two-dimensional [^1H - ^{15}N]-transverse relaxation optimized spectroscopy perturbations of FVO *Plasmodium falciparum* apical membrane antigen 1 (black), by MIPS-0001404 (green), MIPS-0000865 (red) and MIPS-0008405 (blue). Perturbed residues in A–E and F–I were mapped onto the conserved and polymorphic ends of the hydrophobic cleft, respectively. The black arrows indicate changes in the chemical shifts.

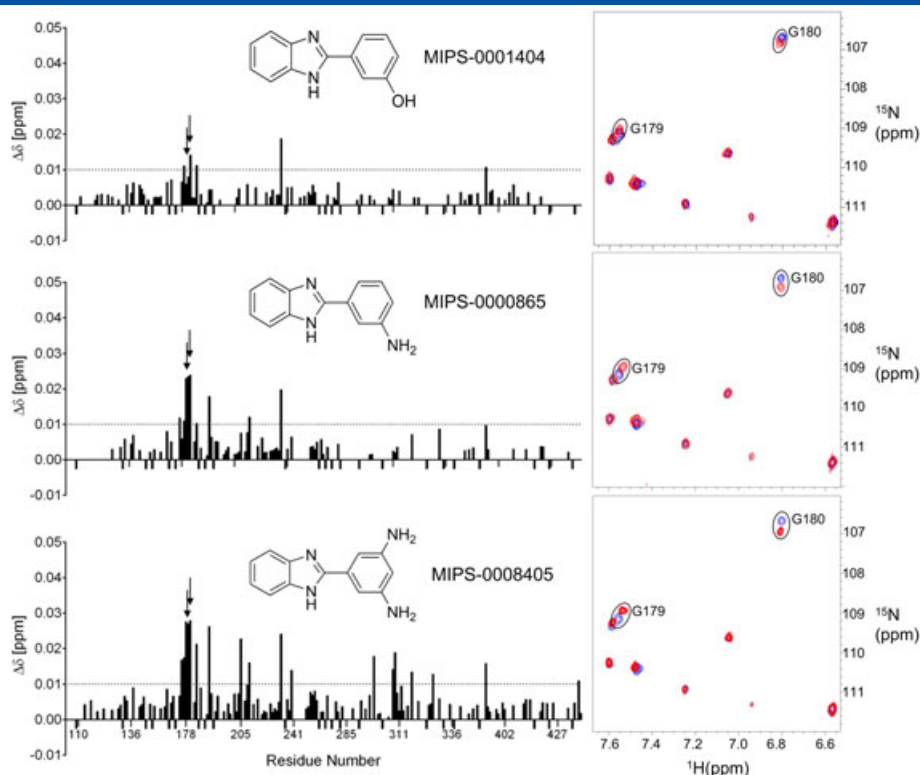


Figure 7. Chemical shift perturbations ($\Delta\delta$) of FVO *Plasmodium falciparum* apical membrane antigen 1 upon binding to benzimidazoles. Left: chemical shift perturbations per residue induced by the compounds shown. Gly179 and Gly180 are highlighted with arrows. Right: expansion of 2D [^1H - ^{15}N]-transverse relaxation optimized spectroscopy spectra highlighting perturbed resonances of Gly179 and Gly180. Blue and red spectra were recorded in the absence and presence of the benzimidazole shown in the left panel. The thick negative bars indicate the regions of unassigned residues (BMRB ID 26648).

correspond to conserved residues near the end of DII loop. Overall, the magnitude of the CSPs induced by benzimidazole compounds, particularly MIPS-0008405, was larger (Figures 6 and 7) compared with that of the CSPs induced by pyrazoles and 2-aminothiazoles (Table S2). MIPS-0008405 perturbed more resonances than other benzimidazoles (Figure 7). The perturbed residues are mapped onto the crystal structure of FVO PfAMA1 (PDB ID 4R1A) in (Figure 8A and 8B). These results were also compared with the R1 and PfRON2 peptide binding sites on 3D7 PfAMA1 (Figure 8C and 8D). Small-molecule binding sites overlap partially with the binding sites of either R1 or PfRON2. Only MIPS-0008405, which showed higher affinity for 3D7 PfAMA1 in SPR binding experiments (Figure 1), was used in 2D [^1H - ^{15}N]-TROSY experiment with 3D7 PfAMA1; CSPs are compared with those for FVO PfAMA1 in Figure 9. The similar perturbations observed in the spectra of FVO PfAMA1 and 3D7 PfAMA1 imply that FVO and 3D7 PfAMA1 have very similar binding sites for this compound.

The pyrazoles MIPS-0001160 and MIPS-0001176 also showed binding to the DII loop end of the cleft. Apart from their interactions with polymorphic residue 172, these small molecules contact mostly conserved residues in this region on PfAMA1 (Table S2). MIPS-0001160 also interacts with residues 419 and 420, which are located outside the hydrophobic cleft.

The 2-aminothiazole MIPS-0000620 induced CSPs at only one resonance, corresponding to Thr171 at the DII end of the cleft, which is also part of the binding site of R1 peptide. We have found that MIPS-0000620 competes with R1 (Lim *et al.*, 2013), and this observation, together with the 2D [^1H - ^{15}N]-TROSY perturbation results, suggests a significant binding activity of MIPS-0000620 at

the DII end of the hydrophobic cleft. Unlike MIPS-0000620, the other 2-aminothiazole, MIPS-0008939, showed extensive CSPs from multiple sites on PfAMA1 that include both the DII loop and polymorphic ends of the cleft (Table S2 of the Supporting Information). This observation is in agreement with our previous study, which showed that 2-aminothiazoles can be problematic promiscuous binders (Devine *et al.*, 2015).

DISCUSSION

Structural information on ligand binding may be obtained using either X-ray crystallography or NMR spectroscopy. The former is usually the method of choice as it provides detailed structural information about ligand-protein interactions that is essential for designing potent inhibitors (Arkin & Wells, 2004; Congreve *et al.*, 2008). However, the study of small molecule-protein crystals in a static environment may not always be fully informative given that the *in vivo* interactions occur in a dynamic solution environment (Ishima & Torchia, 2000; Davis *et al.*, 2003; Davis *et al.*, 2008). Moreover, in the case of apical membrane antigen 1, there is evidence of conformational flexibility (Ge *et al.*, 2014; Lim *et al.*, 2014) and ligand-induced conformational changes (Tonkin *et al.*, 2011; Vulliez-Le Normand *et al.*, 2012; Ge *et al.*, 2014). Mapping CSPs in 2D [^1H - ^{15}N]-TROSY spectra of the target protein represents a rapid and informative means of defining ligand binding sites in solution (Shuker *et al.*, 1996; Williamson, 2013) and can serve as a valuable complement to available crystallographic data. In situations where an accurate small molecule

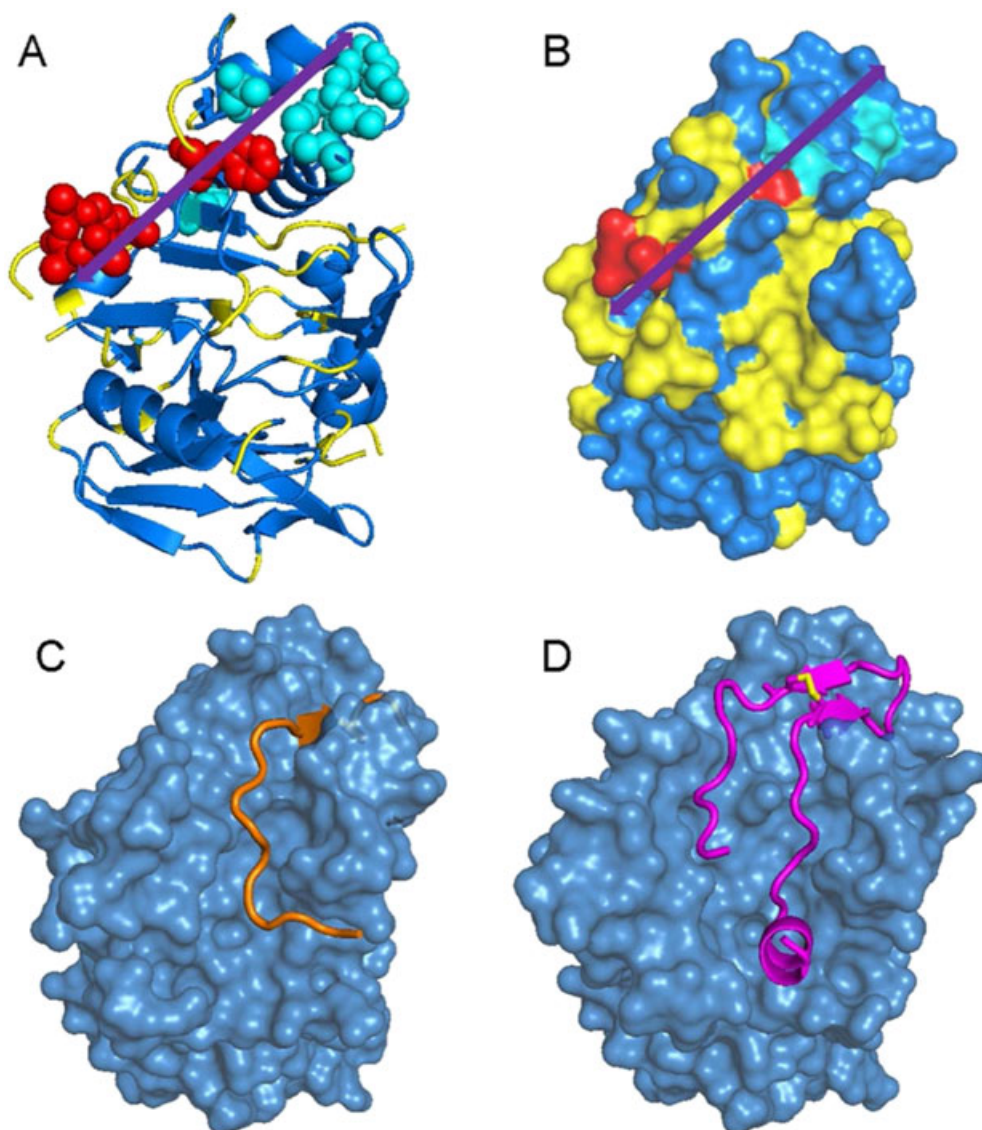


Figure 8. Mapping the chemical shift perturbations of benzimidazoles (MIPS-0001404, MIPS-0000865 and MIPS-0008405) onto the FVO *PfAMA1* crystal structure (Lim *et al.*, 2014) (PDB ID 4R1A). (A) Cartoon view and (B) surface view. The hydrophobic cleft is indicated by purple lines. The blue and yellow regions correspond to assigned and unassigned resonances, respectively. Resonances from residues at the conserved and polymorphic ends of the cleft that were perturbed by these compounds in 2D [^1H - ^{15}N]-transverse relaxation optimized spectroscopy spectra are coloured red and cyan, respectively. Three compounds perturb resonances from residues at two sites in the cleft, with primary and secondary binding sites towards the conserved and polymorphic ends of the cleft, respectively. (C) Crystal structure of 3D7 *PfAMA1* (sky blue) in the presence of R1 peptide (orange) (PDB ID 3SRJ) (Vulliez-Le Normand *et al.*, 2012) and (D) crystal structure of 3D7 *PfAMA1* (sky blue) in the presence of *P. falciparum* rhostry neck protein 2 (magenta) with a disulfide bond highlighted in yellow (PDB ID 3ZWZ). All structures were aligned using the PyMOL (version 1.2r3pre Schrodinger, LLC. Coordinates) alignment function.

binding pose is difficult to obtain, such as low-resolution crystal structures or missing fragment electron densities, mapping the binding sites of fragment hits using CSPs will provide valuable guidance to subsequent medicinal chemistry elaboration.

Two-dimensional [^1H - ^{15}N]-TROSY perturbation experiments can be conducted with relatively small amounts of *PfAMA1* (~0.3 mg). The FVO and 3D7 *PfAMA1* backbone amide resonances in 2D [^1H - ^{15}N]-TROSY spectra were assigned using triple-resonance NMR experiments, complemented by both selective labelling and selective unlabelling methods (Jaipuria *et al.*, 2012; Krishnarjuna *et al.*, 2011). Using all these approaches, we could assign ~84% and ~45%, respectively, of the peaks in 2D [^1H - ^{15}N]-TROSY spectra of FVO and 3D7

PfAMA1. The assignments for FVO *PfAMA1* include most of the conserved hydrophobic cleft residues, as well as the polymorphic residues surrounding the cleft.

In both 3D7 and FVO *PfAMA1*, the number of resonances observable in CLEANEX spectra increased upon addition of *PfRON2* peptide (Vulliez-Le Normand *et al.*, 2012). We believe that these resonances arise from amides in the flexible DII loop, which is displaced by *PfRON2* binding to the hydrophobic cleft (Ge *et al.*, 2014; Lim *et al.*, 2014). Several factors restricted the number of backbone assignments for 3D7 *PfAMA1*, including the instability of the protein (Figure S2), line broadening due to intermediate exchange (Lim *et al.*, 2014) and the greater extent of solvent exchange identified

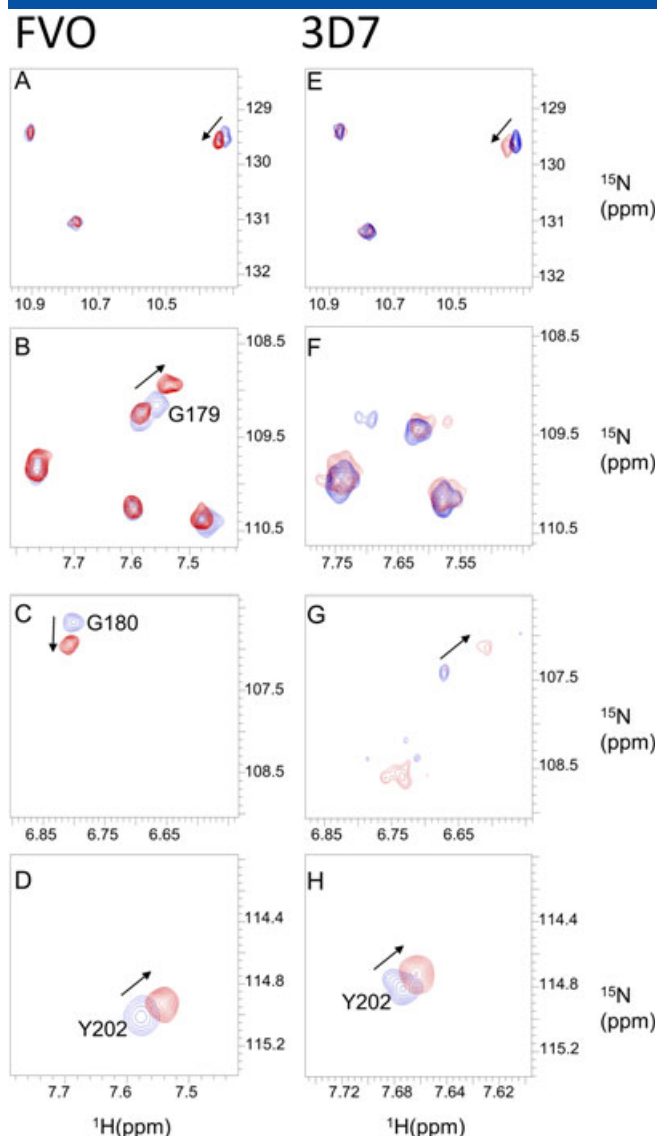


Figure 9. Comparison of the chemical shift perturbations in the spectra of FVO *Plasmodium falciparum* apical membrane antigen 1 (PfAMA1) and 3D7 PfAMA1 in the presence of MIPS-0008405. Left panel (A–D) represents the different regions of the FVO PfAMA1 spectra, and right panel (E–H) represents the same regions for 3D7 PfAMA1 spectra. Blue and red spectra are in the absence and presence, respectively, of MIPS-0008405. The black arrows indicate changes in the chemical shifts.

in CLEANEX spectra (Figure S4). For this reason, the binding sites were mapped on FVO PfAMA1 rather than 3D7 PfAMA1 for most of the small molecules tested here.

The DII end of the hydrophobic cleft has been proposed as a promising site for small-molecule intervention owing to its largely conserved nature (Bai *et al.*, 2005; MacRaild *et al.*, 2011; Lim *et al.*, 2014). The benzimidazole and pyrazole series showed binding to this region and therefore represent promising starting scaffolds for the design of PfAMA1 lead compounds. Although the 2-aminothiazole series also showed binding to the DII end of the cleft, their promiscuous binding, as shown in previous (Devine *et al.*, 2015) and current studies, makes them poor starting points for subsequent fragment elaboration. MIPS-0008405 and MIPS-0000865 small molecules showed interaction

with residue 175, which is one of the very few polymorphic residues found at this end of the cleft. This site can vary between Asp and Tyr in PfAMA1 and is known to limit the cross reactivity of R1 (3D7 PfAMA1-specific peptide inhibitor) against different strains of PfAMA1, with the Y175D substitution reducing R1 peptide binding (Chesne-Seck *et al.*, 2005; Vulliez-Le Normand *et al.*, 2012). Asp is the predominant residue in the PfAMA1 alleles, and as a result of this trait, potent inhibitory activity of R1 is limited to a small subset of *Plasmodium* strains. Two small molecules, MIPS-0008405 and MIPS-0000865, were identified initially from screening against 3D7 PfAMA1, which has Tyr residue at position 175 (Lim *et al.*, 2013). We have shown that these compounds also bound to FVO PfAMA1, indicating that they are able to bind to PfAMA1 with either Asp or Tyr residues at position 175, perhaps through their interaction with the backbone. The benzimidazole series also perturbed resonances near the polymorphic end of the cleft. As opposed to the conserved end of the cleft, this binding site represents a less attractive site for fragment development (MacRaild *et al.*, 2011) as the polymorphic residues would limit the possible chemical space that can be exploited for the design of potent inhibitors with broad strain specificity. It is also possible that the observed CSPs in this region are a result of allosteric effects or non-specific interactions upon small molecule binding to the conserved region at the end of the cleft. Future analogue synthesis would aim to introduce functional groups that favour binding to the DII end of the cleft and disfavour binding to the polymorphic end of the cleft.

Both of the pyrazole compounds contact a dimorphic residue at position 172 in the DII region that can vary between Gly and Glu, with these two residues occurring at approximately the same frequency (Gly = 46% and Glu = 54%). Both 3D7 and FVO PfAMA1 have Gly at this position. It is unclear at this stage if MIPS-0001160 and MIPS-0001176 could still bind to PfAMA1 when Gly is substituted by Glu; this would require further investigation with either PfAMA1 mutation experiments or binding studies of PfAMA1 alleles carrying Glu172.

CONCLUSION

In conclusion, we have successfully assigned 84% of the expected backbone resonances for FVO PfAMA1 and used these for mapping the binding sites for several small molecules. This information will aid in elaborating these small molecules into high-affinity ligands that bind to the conserved regions of PfAMA1. Among the compounds tested, MIPS-0008405 showed stronger affinity by SPR and the largest perturbations in the PfAMA1 spectrum. This suggests that benzimidazoles would be suitable as a basis for designing new PfAMA1 inhibitors targeting the conserved regions of the hydrophobic cleft.

Acknowledgements

We thank Robin Anders for valuable discussions. This work was supported in part by an Australian National Health and Medical Research Council (NHMRC) project grant (1025150). R.S.N. acknowledges fellowship support from the NHMRC. The facilities provided by NMR Research Centre at IISc, supported by the Department of Science and Technology (DST), India, are gratefully acknowledged.

REFERENCES

- Aikawa M, Miller L, Johnson J, Rabbege J. 1978. Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. *J. Cell Biol.* **77**: 72–82.
- Arkin MR, Wells JA. 2004. Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **3**: 301–317.
- Bai T, Becker M, Gupta A, Strike P, Murphy VJ, Anders RF, Batchelor AH. 2005. Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 12736–12741.
- Breman JG. 2012. Resistance to artemisinin-based combination therapy. *Lancet Infect. Dis.* **12**: 820–822.
- Chesne-Seck M-L, Pizarro JC, Normand BV-L, Collins CR, Blackman MJ, Faber BW, Remarque EJ, Kocken CHM, Thomas AW, Bentley GA. 2005. Structural comparison of apical membrane antigen 1 orthologues and paralogues in apicomplexan parasites. *Mol. Biochem. Parasitol.* **144**: 55–67.
- Coley AM, Gupta A, Murphy VJ, Bai T, Kim H, Anders RF, Foley M, Batchelor AH. 2007. Structure of the malaria antigen AMA1 in complex with a growth-inhibitory antibody. *PLoS Pathog.* **3**: e138.
- Coley AM, Parisi K, Masciantonio R, Hoeck J, Casey JL, Murphy VJ, Harris KS, Batchelor AH, Anders RF, Foley M. 2006. The most polymorphic residue on *Plasmodium falciparum* apical membrane antigen 1 determines binding of an invasion-inhibitory antibody. *Infect. Immun.* **74**: 2628–2636.
- Collins CR, Withers-Martinez C, Bentley GA, Batchelor AH, Thomas AW, Blackman MJ. 2007. Fine mapping of an epitope recognized by an invasion-inhibitory monoclonal antibody on the malaria vaccine candidate apical membrane antigen 1. *J. Biol. Chem.* **282**: 7431–7441.
- Congreve M, Chessari G, Tisi D, Woodhead AJ. 2008. Recent developments in fragment-based drug discovery. *J. Med. Chem.* **51**: 3661–3680.
- Davis AM, St-Galley SA, Kleywegt GJ. 2008. Limitations and lessons in the use of X-ray structural information in drug design. *Drug Discov. Today* **13**: 831–841.
- Davis AM, Teague SJ, Kleywegt GJ. 2003. Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angew. Chem. Int. Ed.* **42**: 2718–2736.
- Devine SM, Mulcair MD, Debono CO, Leung EWW, Nissink JWM, Lim SS, Chandrashekar IR, Vazirani M, Mohanty B, Simpson JS, Baell JB, Scammells PJ, Norton RS, Scanlon MJ. 2015. Promiscuous 2-aminothiazoles (PrATs): a frequent hitting scaffold. *J. Med. Chem.* **58**: 1205–1214.
- Gardner KH, Kay LE. 1998. The use of ^2H , ^{13}C , ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **27**: 357–406.
- Ge X, MacRaid CA, Devine SM, Debono CO, Wang G, Scammells PJ, Scanlon MJ, Anders RF, Foley M, Norton RS. 2014. Ligand-induced conformational change of *Plasmodium falciparum* AMA1 detected using ^{19}F NMR. *J. Med. Chem.* **57**: 6419–6427.
- Goto NK, Kay LE. 2000. New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Curr. Opin. Struct. Biol.* **10**: 585–592.
- Harris KS, Casey JL, Coley AM, Karas JA, Sabo JK, Tan YY, Dolezal O, Norton RS, Hughes AB, Scanlon D, Foley M. 2009. Rapid optimization of a peptide inhibitor of malaria parasite invasion by comprehensive N-methyl scanning. *J. Biol. Chem.* **284**: 9361–9371.
- Harris KS, Casey JL, Coley AM, Masciantonio R, Sabo JK, Keizer DW, Lee EF, McMahon A, Norton RS, Anders RF, Foley M. 2005. Binding hot spot for invasion inhibitory molecules on *Plasmodium falciparum* apical membrane antigen 1. *Infect. Immun.* **73**: 6981–6989.
- Hwang T-L, van Zijl PM, Mori S. 1998. Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical exchange (CLEANEX-PM) approach with a fast-HSQC (FHSQC) detection scheme. *J. Biomol. NMR* **11**: 221–226.
- Ishima R, Torchia DA. 2000. Protein dynamics from NMR. *Nat. Struct. Biol.* **7**: 740–743.
- Jaipuria G, Krishnarjuna B, Mondal S, Dubey A, Atreya HS. 2012. Amino acid selective labeling and unlabeled for protein resonance assignments, Isotope labeling in biomolecular NMR, Atreya HS (ed). Springer Netherlands. pp 95–118.
- Keizer DW, Miles LA, Li F, Nair M, Anders RF, Coley AM, Foley M, Norton RS. 2003. Structures of phage-display peptides that bind to the malarial surface protein, apical membrane antigen 1, and block erythrocyte invasion. *Biochemistry* **42**: 9915–9923.
- Krishnarjuna B, Jaipuria G, Thakur A, D'Silva P, Atreya HS. 2011. Amino acid selective unlabeled for sequence specific resonance assignments in proteins. *J. Biomol. NMR* **49**: 39–51.
- Lamarque M, Besteiro S, Papoin J, Roques M, Normand BV-L, Morlon-Guyot J, Dubremetz J-F, Fauquenoy S, Tomavo S, Faber BW, Kocken CH, Thomas AW, Boulanger MJ, Bentley GA, Lebrun M. 2011. The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites. *PLoS Pathog.* **7**: e1001276.
- Lee EF, Yao S, Sabo JK, Fairlie WD, Stevenson RA, Harris KS, Anders RF, Foley M, Norton RS. 2011. Peptide inhibitors of the malaria surface protein, apical membrane antigen 1: identification of key binding residues. *Biopolymers* **95**: 354–364.
- Li F, Dluzewski A, Coley AM, Thomas A, Tilley L, Anders RF, Foley M. 2002. Phage-displayed peptides bind to the malarial protein apical membrane antigen-1 and inhibit the merozoite invasion of host erythrocytes. *J. Biol. Chem.* **277**: 50303–50310.
- Lim SS, Debono CO, MacRaid CA, Chandrashekar IR, Dolezal O, Anders RF, Simpson JS, Scanlon MJ, Devine SM, Scammells PJ, Norton RS. 2013. Development of inhibitors of *Plasmodium falciparum* apical membrane antigen 1 based on fragment screening. *Aust. J. Chem.* **66**: 1530–1536.
- Lim SS, Yang W, Krishnarjuna B, Kannan Sivaraman K, Chandrashekar IR, Kass I, MacRaid CA, Devine SM, Debono CO, Anders RF, Scanlon MJ, Scammells PJ, Norton RS, McGowan S. 2014. Structure and dynamics of apical membrane antigen 1 from *Plasmodium falciparum* FVO. *Biochemistry* **53**: 7310–7320.
- MacRaid CA, Norton RS. 2014. RASP: rapid and robust backbone chemical shift assignments from protein structure. *J. Biomol. NMR* **58**: 155–163.
- MacRaid CA, Anders RF, Foley M, Norton RS. 2011. Apical membrane antigen 1 as an anti-malarial drug target. *Curr. Top. Med. Chem.* **11**: 2039–2047.
- Muchmore DC, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW. 1989. Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance, Methods enzymol, Norman JO, Thomas LJ (eds). Academic Press: London. pp 44–73.
- Ohki S-y, Kainosho M. 2008. Stable isotope labeling methods for protein NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **53**: 208–226.
- Ouattara A, Takala-Harrison S, Thera MA, Coulibaly D, Niangaly A, Saye R, Tolo Y, Dutta S, Heppner DG, Soisson L, Diggs CL, Vekemans J, Cohen J, Blackwelder WC, Dube T, Laurens MB, Doumbo OK, Plowe CV. 2013. Molecular basis of allele-specific efficacy of a blood-stage malaria vaccine: vaccine development implications. *J. Infect. Dis.* **207**: 511–519.
- Richard D, MacRaid CA, Riglar DT, Chan J-A, Foley M, Baum J, Ralph SA, Norton RS, Cowman AF. 2010. Interaction between *Plasmodium falciparum* apical membrane antigen 1 and the roptry neck protein complex defines a key step in the erythrocyte invasion process of malaria parasites. *J. Biol. Chem.* **285**: 14815–14822.
- Riek R, Pervushin K, Wüthrich K. 2000. TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. *Trends Biochem. Sci.* **25**: 462–468.
- Salzmann M, Pervushin K, Wider G, Senn H, Wüthrich K. 1998. TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc. Natl. Acad. Sci. U. S. A.* **95**: 13585–13590.
- Sattler M, Fesik SW. 1996. Use of deuterium labeling in NMR: overcoming a sizeable problem. *Structure* **4**: 1245–1249.
- Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. 1996. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**: 1531–1534.
- Sivashanmugam A, Murray V, Cui CX, Zhang YH, Wang JJ, Li QQ. 2009. Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci.* **18**: 936–948.
- Srinivasan P, Beatty WL, Diouf A, Herrera R, Ambroggio X, Moch JK, Tyler JS, Narum DL, Pierce SK, Boothroyd JC, Haynes JD, Miller LH. 2011. Binding of *Plasmodium* merozoite proteins RON2 and AMA1 triggers commitment to invasion. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 13275–13280.
- Takala SL, Coulibaly D, Thera MA, Batchelor AH, Cummings MP, Escalante AA, Ouattara A, Traoré K, Niangaly A, Djimdé AA, Doumbo OK, Plowe

- CV. 2009. Extreme polymorphism in a vaccine antigen and risk of clinical malaria: implications for vaccine development. *Sci. Transl. Med.* **1**: 2ra5–2ra5.
- Tonkin ML, Roques M, Lamarque MH, Pugnière M, Douguet D, Crawford J, Lebrun M, Boulanger MJ. 2011. Host cell invasion by apicomplexan parasites: insights from the co-structure of AMA1 with a RON2 peptide. *Science* **333**: 463–467.
- Vulliez-Le Normand B, Tonkin ML, Lamarque MH, Langer S, Hoos S, Roques M, Saul FA, Faber BW, Bentley GA, Boulanger MJ, Lebrun M. 2012. Structural and functional insights into the malaria parasite moving junction complex. *PLoS Pathog.* **8**: e1002755.
- Wang G, MacRaild CA, Mohanty B, Mobli M, Cowieson NP, Anders RF, Simpson JS, McGowan S, Norton RS, Scanlon MJ. 2014. Molecular insights into the interaction between *Plasmodium falciparum* apical membrane antigen 1 and an invasion-inhibitory peptide. *PLoS One* **9**: e109674.
- White NJ. 2004. Antimalarial drug resistance. *J. Clin. Invest.* **113**: 1084–1092.
- WHO. 2014. World malaria report. *WHO Library Cataloguing*. 1–242.
- Williamson MP. 2013. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**: 1–16.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web site.

Supplementary Information

Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials


Bankala Krishnarjuna^{a†}, San Sui Lim^{a†}, Shane M. Devine^a, Cael O. Debono^a, Raymond Lam^a, Indu R. Chandrashekarana^a, Garima Jaipuria^b, Hiromasa Yagi^a, Hanudatta S. Atreya^b, Martin J. Scanlon^a, Christopher A. MacRaild^a, Peter J. Scammells^a, Raymond S. Norton^{a*}

a Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, Victoria 3052, Australia.

b NMR Research Centre, Indian Institute of Science, Bangalore 560012, India.

[†]These authors contributed equally.

** Correspondence to: Raymond S. Norton, Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, Victoria 3052, Australia.*



CONTENTS

Table S1. 3D7 *Pf*AMA1 chemical shifts

Table S2. Chemical shift perturbations

Figure S1. Comparison of non-deuterated and deuterated spectra

Figure S2. *Pf*AMA1 stability studies by 2D TROSY

Figure S3. Strip plots for sequential assignment of FVO *Pf*AMA1

Figure S4. CLEANEX experiments

Figure S5. Amino acid selective unlabelling

Figure S6. Sequential assignments for 3D7 *Pf*AMA1 using unlabelling strategy

Figure S7. SPR binding experiments

S18 Chemistry – General experimental

Synthetic chemistry procedures

Table S1. 3D7 *Pf*AMA1 chemical shifts.

Sample conditions: 250 μ M 3D7 *Pf*AMA1, 20 mM phosphate buffer, pH 7.0 and temperature 303 K.

Residue number	HN	N	C $_{\alpha}$	C $_{\beta}$
P109	-	-	63.39	30.61
W110	7.97	119.23	57.46	28.95
T111	7.00	114.52	67.21	68.34
E112	8.39	120.14	59.33	27.96
Y113	8.24	122.02	60.88	38.60
M114	8.26	110.44	53.10	27.96
A115	7.51	127.91	55.88	17.03
K116	7.18	113.79	56.71	30.63
D118	7.81	122.72	51.80	37.66
I119	7.72	126.03	64.53	37.70
E120	10.03	122.52	60.50	29.47
E121	7.02	116.80	57.60	29.57
V122	8.61	114.12	64.20	31.06
H123	7.97	118.07	57.90	33.94
G124	7.92	102.73	46.11	-
S125	6.54	110.98	53.84	64.15
G126	8.76	105.80	44.35	-
I127	8.11	118.42	63.50	38.11
R128	6.87	121.00	59.15	27.84
V129	6.02	112.57	61.61	32.30
D130	8.93	129.68	51.45	40.71
K148	8.07	124.08	58.52	31.52
C149	9.81	116.74	51.27	39.27
F152	7.86	127.96	60.73	37.34
G153	9.28	113.30	44.34	-
K154	6.35	117.69	54.71	33.02

G155	7.60	109.96	42.94	-
I156	3.39	116.28	57.29	37.85
I157	9.14	128.36	60.64	38.59
I158	8.22	128.86	60.70	36.53
P192	-	-	63.07	32.54
M193	8.32	119.87	55.40	36.58
T194	9.12	117.08	60.83	70.26
L195	8.47	122.73	58.41	40.37
D196	8.52	116.79	57.22	39.61
E197	7.82	120.68	58.89	28.54
M198	8.85	121.62	60.12	34.03
Y202	7.70	114.58	58.11	36.69
K203	6.91	118.91	59.17	30.56
Y207	7.90	117.81	59.21	37.23
V208	7.24	116.06	63.02	31.75
K209	7.76	116.76	58.86	30.22
N210	7.52	115.57	52.55	38.83
L211	7.22	120.94	54.76	42.55
D212	7.74	118.25	53.45	41.40
E213	9.02	117.57	59.22	28.71
L214	7.66	117.12	57.16	38.88
T215	8.47	119.37	66.98	65.79
L216	8.72	121.35	58.19	40.22
C217	7.38	115.22	56.39	33.51
S218	7.62	111.56	60.84	62.45
R219	8.94	119.71	56.48	27.05
H220	9.04	122.45	59.11	30.48
A221	8.02	120.5	54.35	18.81
G222	7.79	105.10	45.74	-
P237	-	-	61.52	31.76
A238	8.50	117.88	50.30	24.30

V239	9.03	117.93	59.53	35.02
Y240	8.37	126.56	56.15	40.66
D241	8.35	128.42	51.22	41.65
P278	-	-	60.91	32.20
A279	8.61	121.55	52.00	22.75
K280	8.06	116.97	55.46	31.42
Y287	8.12	121.01	53.10	37.59
T288	8.61	118.34	61.85	69.86
Y289	8.44	128.70	58.77	39.88
L290	9.06	131.33	52.98	42.66
N293	8.36	119.36	50.71	38.84
V294	6.53	121.26	64.12	30.35
V295	8.24	123.39	60.95	32.64
D296	8.36	117.94	54.93	39.61
N297	8.10	115.80	50.92	35.56
W298	7.59	121.10	59.73	27.17
E299	7.14	122.96	59.09	27.45
K300	7.44	115.94	56.49	32.79
V301	7.07	109.87	61.65	32.39
C302	8.24	117.00	53.97	42.72
A310	7.61	122.01	50.56	21.98
K311	8.11	114.77	52.02	35.56
F312	9.78	124.25	57.95	39.88
G313	8.01	108.44	43.67	-
L314	8.79	119.04	52.33	44.21
P324	-	-	64.36	30.99
H325	7.34	115.72	54.88	29.50
V326	8.04	117.17	59.16	3.79
N327	9.29	121.85	51.83	39.36
E328	8.47	123.80	55.22	31.52
N338	8.96	119.85	59.12	39.13

K339	8.59	117.94	60.13	31.31
L340	7.80	121.76	57.66	41.31
V341	8.29	119.34	65.97	30.20
F342	8.44	116.68	62.16	38.11
G394	7.13	118.86	44.69	-
K395	7.40	121.5	56.47	31.68
G396	8.24	111.28	44.66	-
Y397	7.13	118.86	57.03	34.55
N398	6.29	113.69	52.56	39.21
W399	7.50	118.97	57.68	33.78
G400	8.56	104.58	44.68	-
N401	9.73	121.68	51.14	41.08
Y402	10.07	125.69	55.68	38.49
N403	8.58	127.39	50.94	37.59
C409	9.82	124.83	53.58	37.19
E410	9.45	130.69	55.28	28.44
I411	9.05	125.37	62.72	39.12
F412	8.39	122.66	56.33	41.18
V414	7.00	118.62	58.64	33.03
K415	7.50	121.93	52.44	30.12
I420	8.89	123.42	59.84	39.85
N421	8.56	125.83	53.30	37.75
N422	7.53	124.72	53.70	39.63
S424	8.77	116.87	58.86	62.78
Y425	7.47	119.05	57.65	43.86
I426	8.99	119.62	60.29	41.81
A427	8.23	131.76	51.32	16.99
T428	9.45	120.33	60.00	71.90
T429	7.46	112.08	59.50	72.92
A430	8.60	120.72	54.59	20.06

Table S2. Specific binding sites of fragment hits determined using 2D [^1H - ^{15}N]- TROSY perturbations. Site 1 and 2 correspond to the conserved and polymorphic ends, respectively, of the FVO *Pf*AMA1 hydrophobic cleft. Others are outside the cleft. Only one perturbation was observed for MIPS-0000620. The CSP values (ppm) are shown in brackets. **Polymorphic residues are highlighted red.**

Binding regions	MIPS-0008405	MIPS-0000865	MIPS-0001404	MIPS-0001160	MIPS-0001176	MIPS-0008939
Site 1	L176(0.01) K177(0.01) D178(0.03) G179(0.03) G180(0.03) F183(0.02)	D175 (0.01) K177(0.01) D178(0.02) G179(0.02) G180(0.02) F183(0.01)	K177(0.01) G180(0.01) F183(0.01)	G172 (0.01) K177(0.02) G180(0.02)	G172 (0.02) G180(0.01) I363 (0.01)	T171(0.01) L176 (0.02) K177(0.02) D178(0.01) G179(0.01) G180(0.02) F183(0.01)
Site 2	I190(0.02) N205(0.02) V208(0.01) K209(0.01) M224 (0.02) Y240(0.01)	I190(0.02) K209(0.01) M224 (0.02)	M224 (0.02)			S191(0.01) Q255(0.01)
others	E299 (0.02) E308 (0.01) N309(0.02) D317(0.01) N327(0.01) I363(0.01) V437(0.01)	I363(0.01)	I363(0.01)	I363 (0.01) L419(0.01)		L307(0.01) E308 (0.01) N309(0.01) A310(0.01) I363 (0.01) N421(0.01)

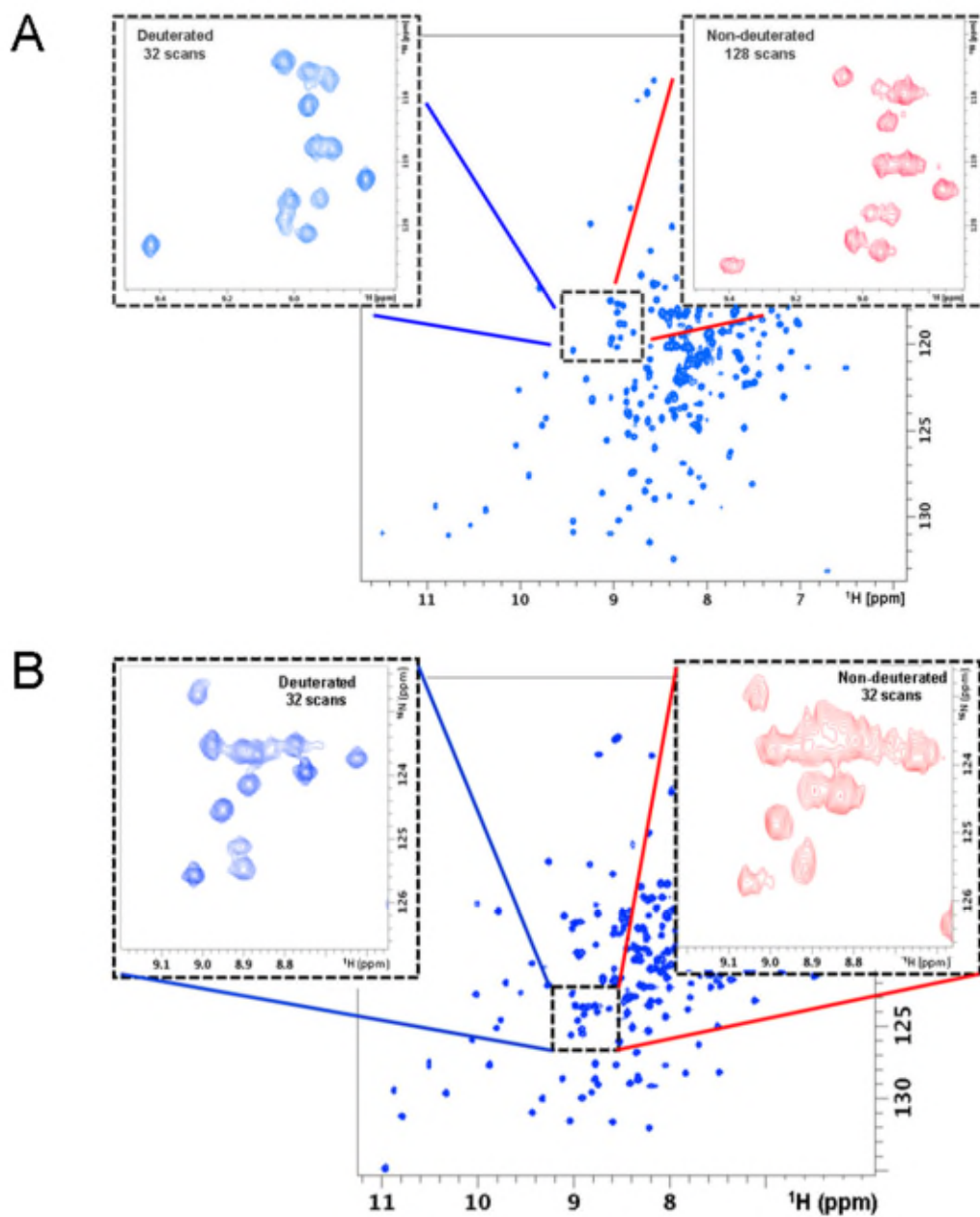


Figure S1. Comparison between 2D [^1H - ^{15}N]-TROSY spectra of deuterated (blue) and non-deuterated (red) A) FVO *Pf*AMA1 and B) 3D7 *Pf*AMA1.

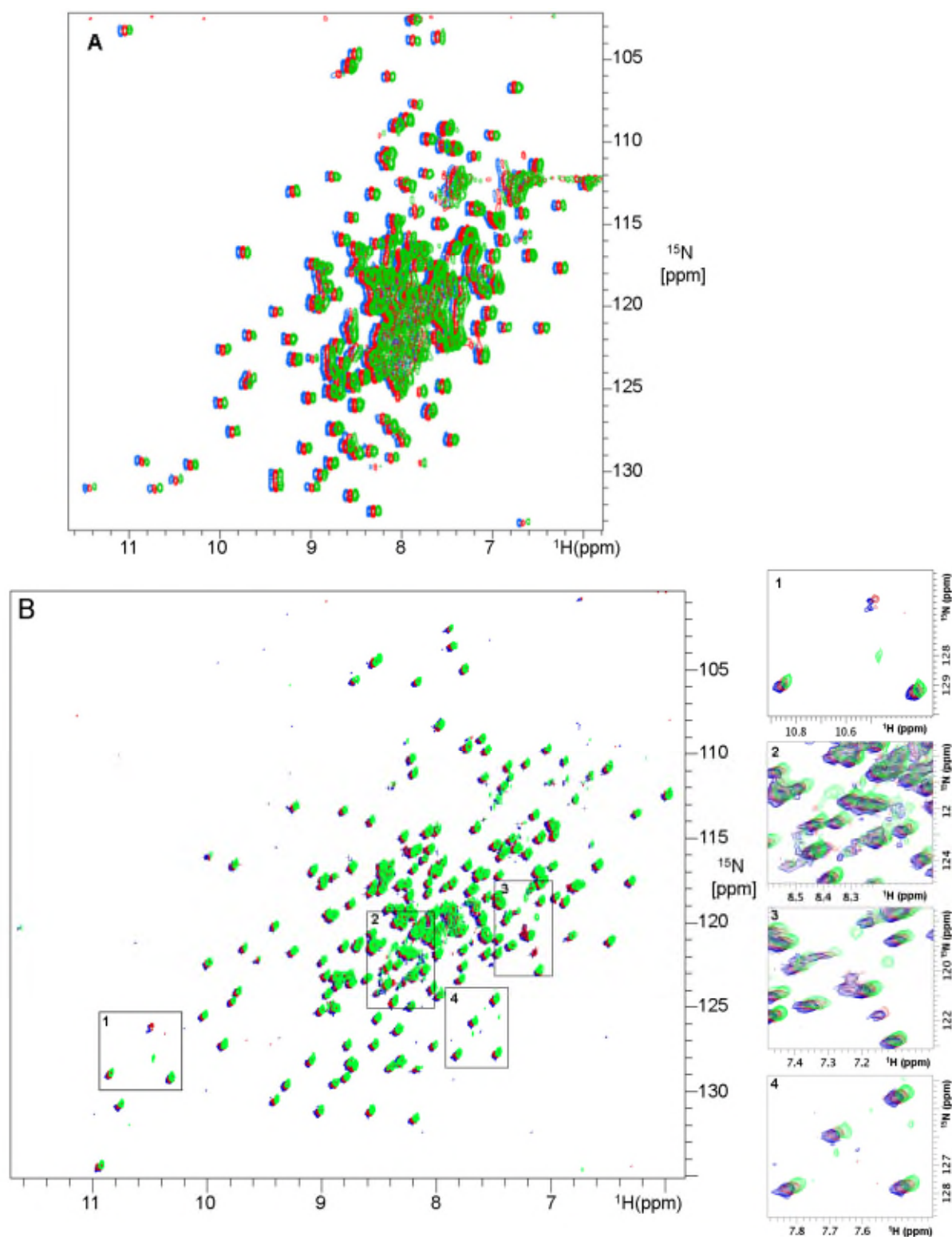


Figure S2. Stability studies of A) ^1H - ^{15}N -labelled FVO *PfAMA1* and B) ^2H - ^{15}N -labelled 3D7 *PfAMA1* using 2D [^1H - ^{15}N]-TROSY experiments at 35 °C and 30 °C, respectively. Spectra were acquired at Day 0 (blue), Day 3 (red) and Day 6 (green) for FVO *PfAMA1* and Day 0 (blue), Day 1 (red) and Day 2 (green) for 3D7 *PfAMA1*. For clarity, the TROSY spectra were offset by -0.05 and -0.1 ppm in the ^1H and ^{15}N dimensions, respectively. The spectra were recorded in buffer containing protease inhibitor cocktail (Roche) 50 mM of each L-arginine

and L-glutamate. The appearance or disappearance of resonances is highlighted in the boxes and the same expanded regions are shown on right side of the spectrum. This suggests that a part of the 3D7 *Pf*AMA1 protein becomes misfolded with time. No changes were observed over 6 days in the FVO *Pf*AMA1 spectra.

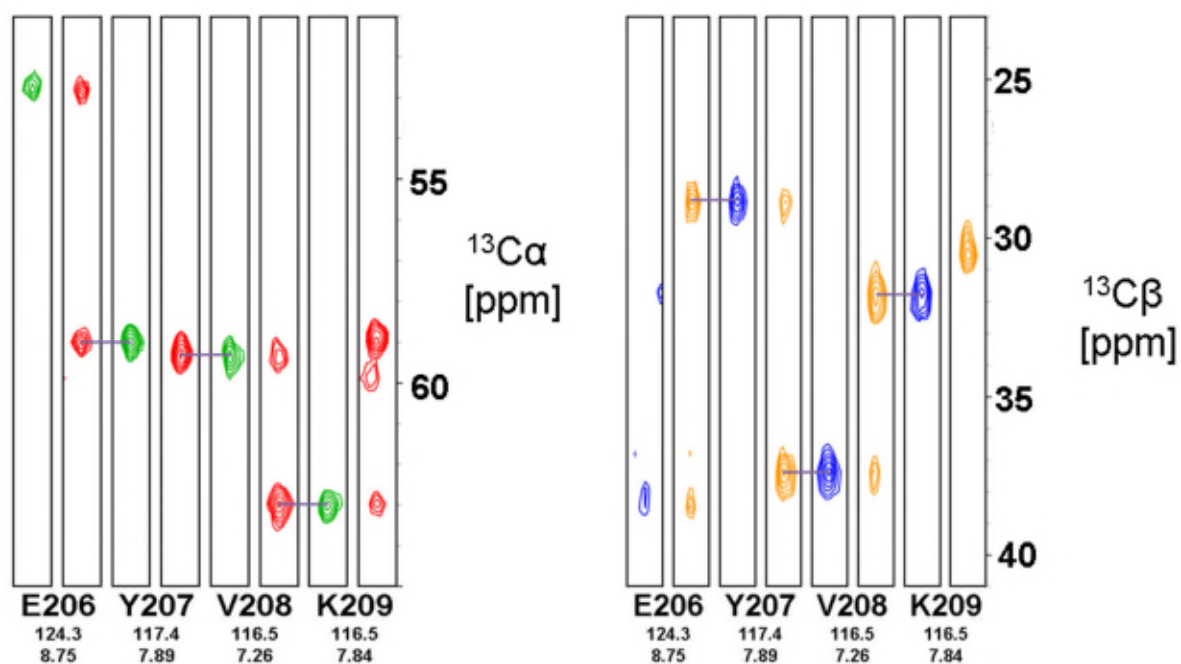


Figure S3. Strip plots for sequential assignment of FVO *PfAMA1*. The sequential connectivities of C_α and C_β chemical shift from residues E206 to K209 are shown in the left and right panels, respectively. HNCA, HN(CO)CA, HN(CA)CB and HN(COCA)CB spectra are indicated in red, green, orange and blue, respectively.

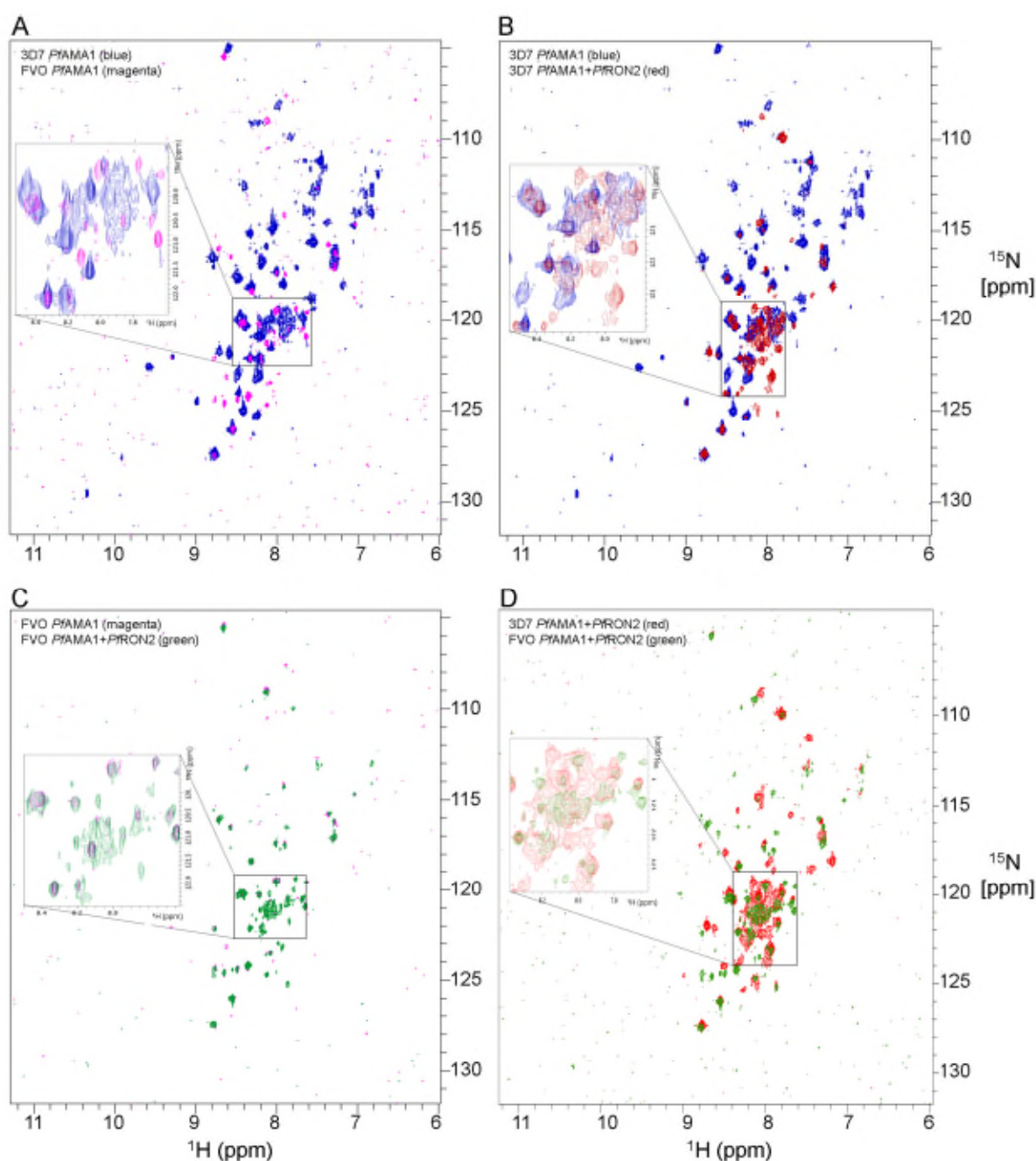
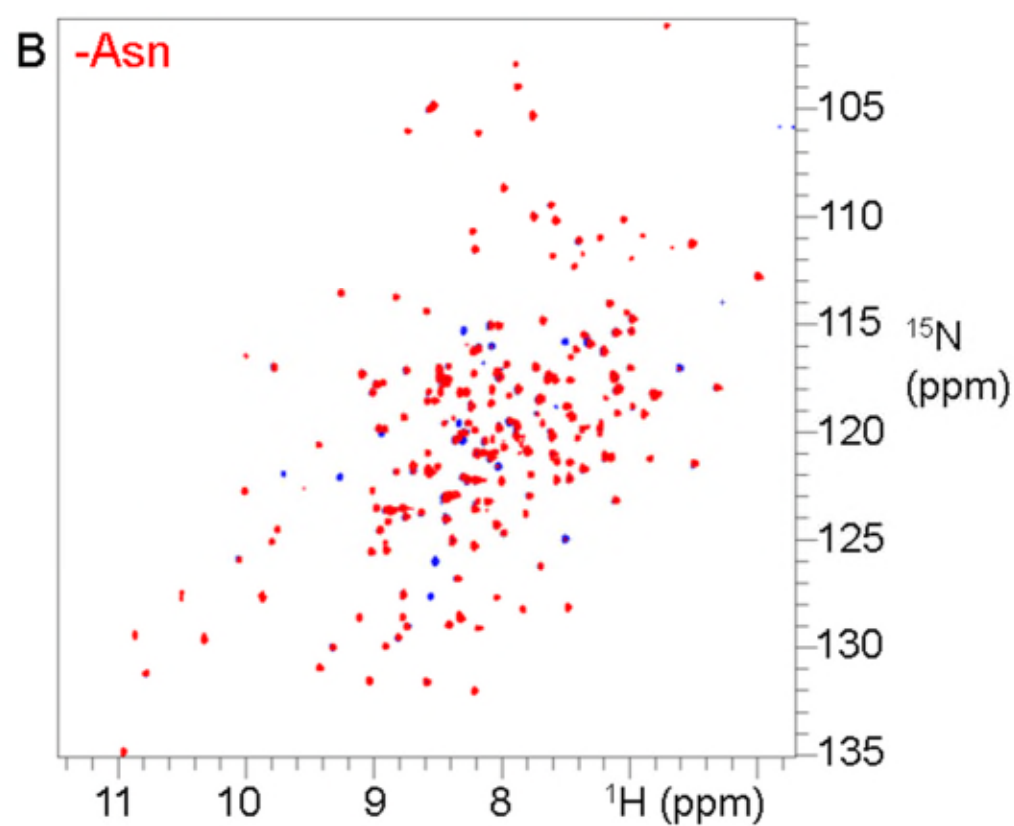
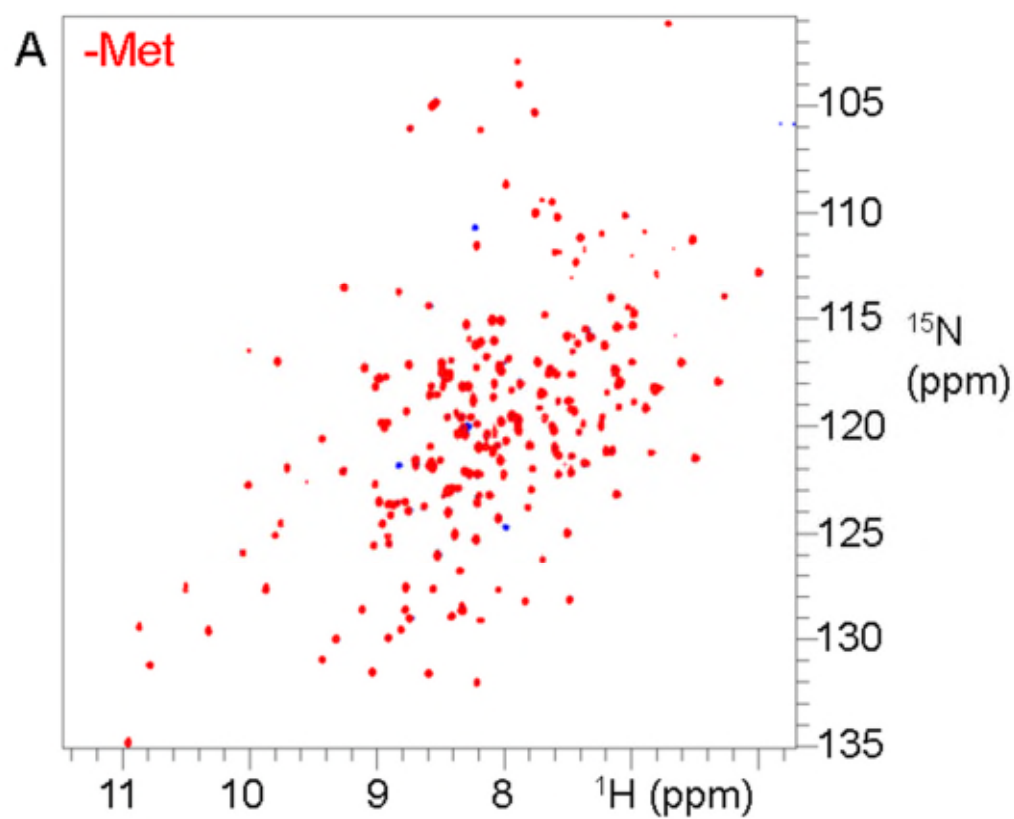


Figure S4. Comparison of solvent exchange in 3D7 and FVO *PfAMA1* in the absence and presence of RON2 peptide, by CLEANEX TROSY. Overlay of A) 3D7 *PfAMA1* and FVO *PfAMA1*, B) 3D7 *PfAMA1* and 3D7 *PfAMA1*+RON2, C) FVO *PfAMA1* and FVO *PfAMA1*+RON2 and D) 3D7 *PfAMA1*+RON2 and FVO *PfAMA1*+RON2. The blue and red coloured spectra are from 3D7 *PfAMA1* and 3D7 *PfAMA1*+RON2, respectively. The magenta and green coloured spectra are from FVO *PfAMA1* and FVO *PfAMA1*+RON2, respectively. The expanded regions show the highly overlapped resonances of the spectra. The mixing time was 45 ms. The number of scans was 128 for each increment, with a total of 256 t_1 increments.



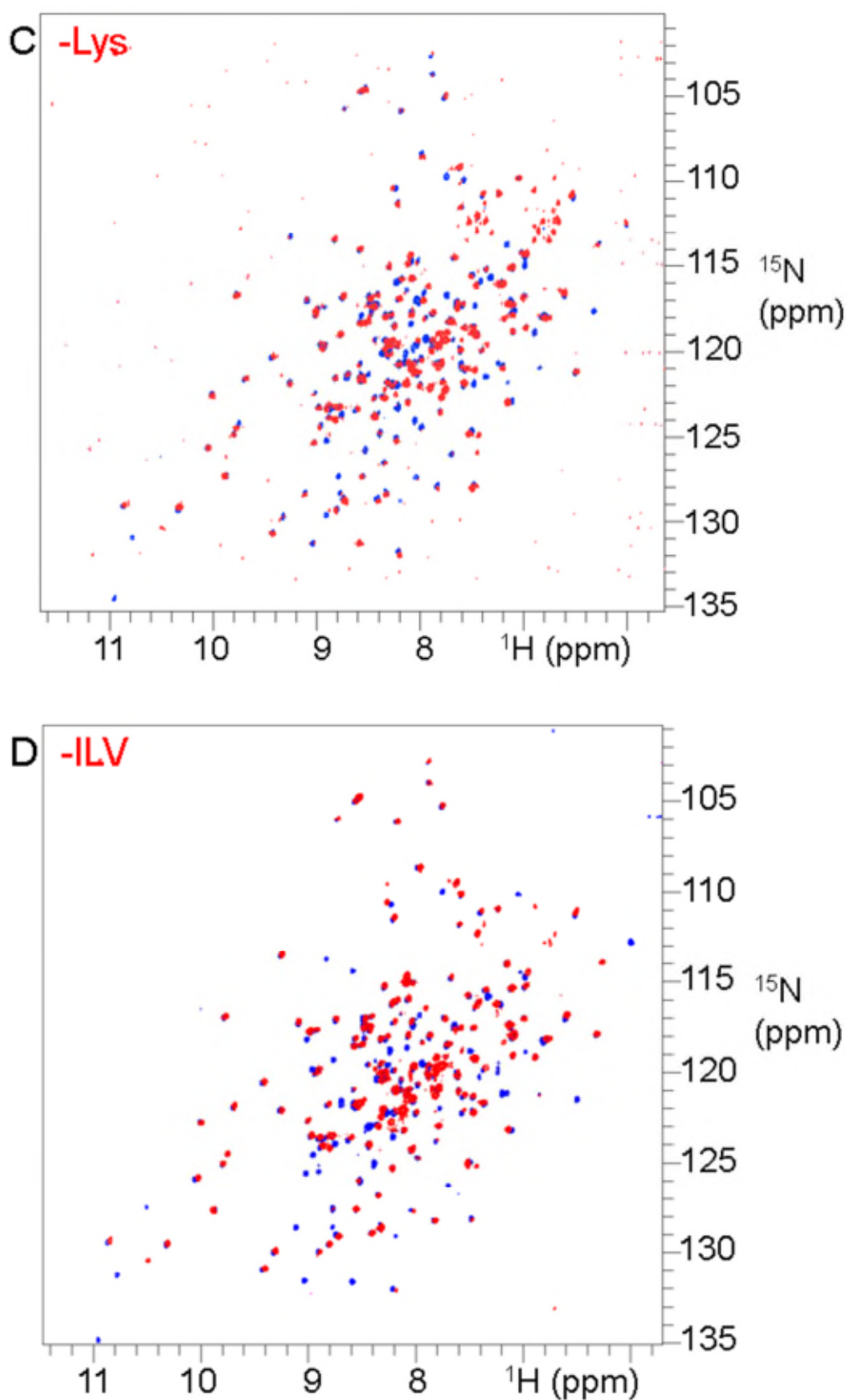


Figure S5. Amino acid selective unlabelling for residue-specific assignments in 3D7 PfAMA1. 2D [^1H - ^{15}N]-TROSY spectrum (blue) recorded on uniformly ^2H , ^{15}N -labelled sample is overlaid with the spectra (red) recorded on different selectively-unlabelled samples. A, B, C

and D represent the Met, Asn, Lys and ILV unlabelled spectra, respectively. Resonances from uniformly-labelled protein (blue) that are missing in the red spectra arise from unlabelled residues.

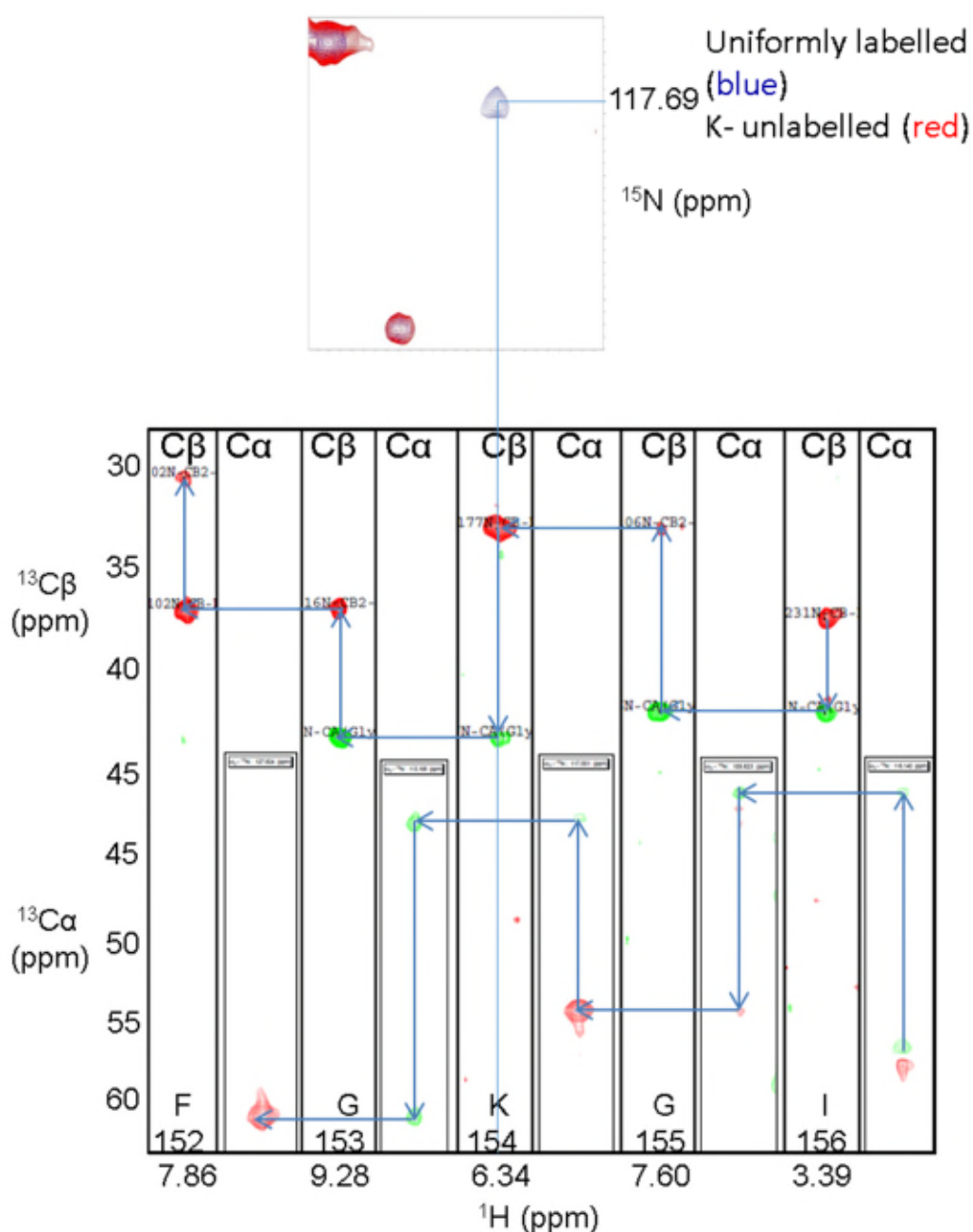


Figure S6. Amino acid selective unlabelling for residue-specific and sequential assignments of 3D7 *PfAMA1*. A) Overlay of the 2D ^1H - ^{15}N -TROSY spectrum of Lys selectively unlabelled sample on the spectrum of uniformly ^{15}N -labelled sample. B) Sequential backbone assignments using the residue-specific assignments as starting points in the spectrum of the Lys selectively unlabelled sample (*cf.* Figure S5C).

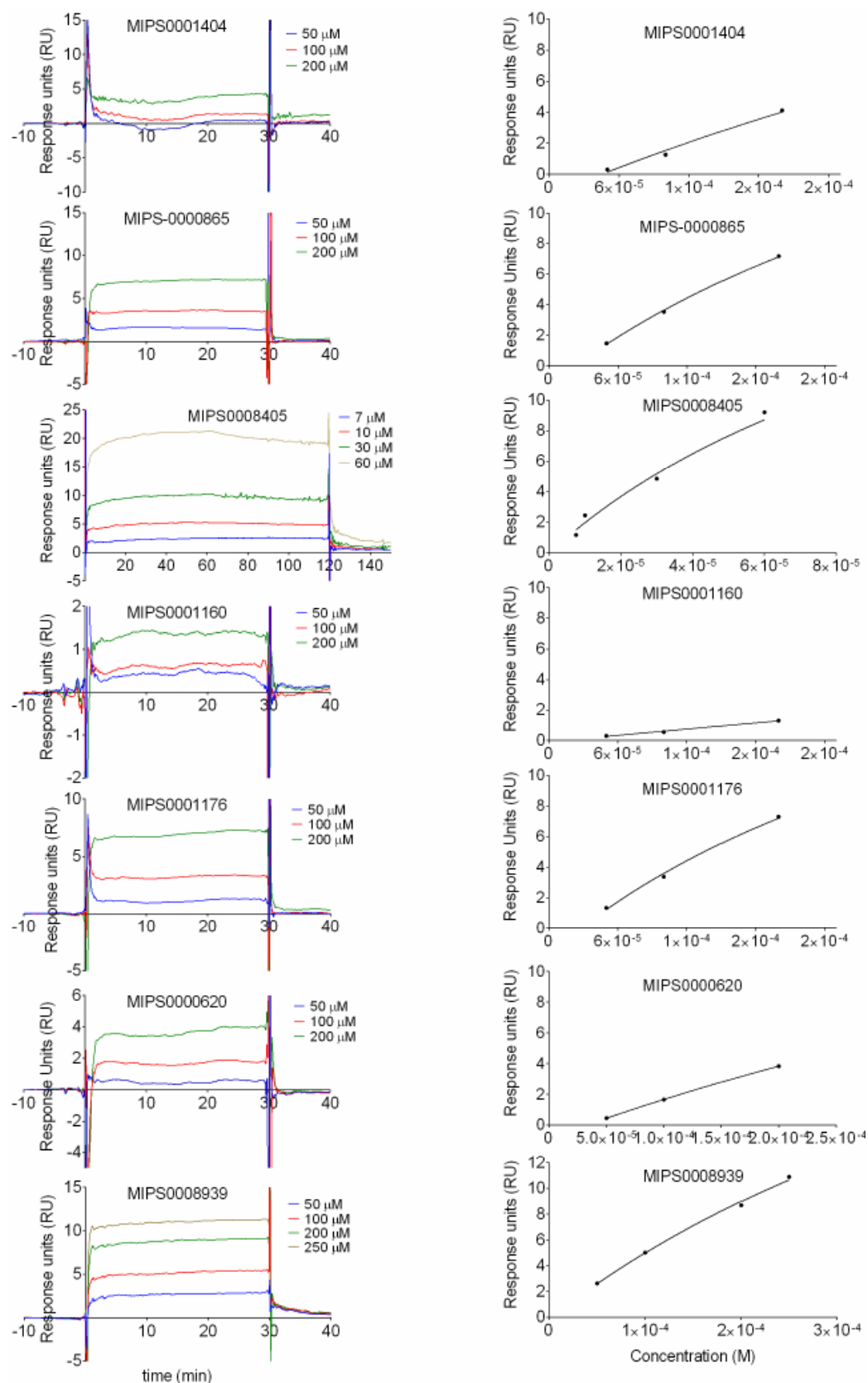


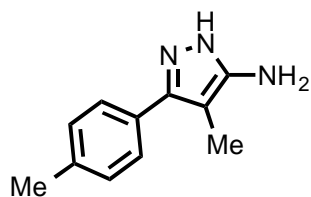
Figure S7: Surface plasmon resonance (SPR) results for the benzimidazoles (MIPS-0001404, MIPS-0000865 and MIPS-0008405), pyrazoles (MIPS-0001160 and MIPS-0001176) and 2-aminothiazoles (MIPS-0000620 and MIPS-0008939). Compounds were tested at concentrations shown on the right side of sensorgrams.

Chemistry – General Experimental

NMR spectra (^1H , ^{13}C) were recorded on a Bruker Avance Nanobay III 400 MHz Ultrashield Plus spectrometer coupled to a BACS 60 automatic sample changer at 400.13 and 100.62 MHz, respectively, at 25 °C. Chemical shifts (δ) are recorded in parts per million (ppm) by correction with reference to the chemical shift of the solvent, according to the procedure described by Gottlieb.⁴⁵ Coupling constants (J) are recorded in Hz, and the significant multiplicities described by singlet (s), doublet (d), triplet (t), quadruplet (q), broad (br), multiplet (m), doublet of doublets (dd), triplet of doublets (td) and doublet of triplets (dt). LC-MS were run to verify reaction outcome and purity using an Agilent 6100 series Single Quad coupled to an Agilent 1200 series HPLC. The following buffers were used: buffer A, 0.1% formic acid in H_2O ; buffer B, 0.1% formic acid in MeCN. The following gradient was used with a Phenomenex Luna 3 μM C8(2) 15 mm \times 4.6 mm column, and a flow rate of 0.5 mL/min and total run time of 12 min; 0–4 min 95% buffer A and 5% buffer B, 4–7 min 0% buffer A and 100% buffer B, 7–12 min 95% buffer A and 5% buffer B. Mass spectra were acquired in positive and negative ion mode with a scan range of 0–1000 m/z at 5 V. UV detection was carried out at 254 nm. All compounds were of >95% purity. Thin layer chromatography was conducted on 0.2 mm plates using Merck silica gel 60 F₂₅₄. Column chromatography was achieved using Merck silica gel 60 (particle size 0.063–0.200 μm , 70–230 mesh).

Synthetic procedures

4-Methyl-3-(*p*-tolyl)-1*H*-pyrazol-5-amine (MIPS-0001176)

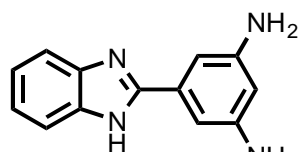


NaH (60% in mineral oil, 365 mg, 9.14 mmol) was suspended in dry toluene (10 mL) under N_2 , followed by the addition of ethyl 4-methylbenzoate (488 μL , 3.05 mmol). The reaction vessel was sealed under an N_2 atmosphere and heated to 70 °C, at which point dry propionitrile (1.09 mL, 15.2 mmol) was added dropwise. The reaction mixture was then heated to 80 °C for 11 h. The reaction mixture was then cooled to room temperature and diluted with petroleum spirits (10 mL). The precipitate was filtered and washed with petroleum spirits, then allowed to dry. The precipitate was dissolved in water (20 mL) and acidified to pH 1 with 1M HCl. The resultant solution was extracted with EtOAc (3 \times 20 mL). The combined organic fractions were re-extracted with sat. NaHCO_3 (20 mL) and sat. NaCl (20 mL), dried over MgSO_4 , filtered, evaporated at reduced pressure and gave 3-(4-methylphenyl)-2-methyl-3-oxopropanenitrile as a yellow oil (401 mg, 76%). This was used without further purification. ^1H NMR (400 MHz, CDCl_3): δ 7.91 – 7.86 (m, 2H), 7.32 (d, J = 8.0 Hz, 2H), 4.34 (q, J = 7.2 Hz, 1H), 2.44 (s, 3H), 1.64 (d, J = 7.2 Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3): δ 190.4 (C), 145.9 (C), 131.4 (C), 129.9 (2 \times CH), 129.1 (2 \times CH), 118.5 (C), 33.7 (CH), 21.9 (CH_3), 15.1 (CH_3).

3-(4-Methylphenyl)-2-methyl-3-oxopropanenitrile (350 mg, 2.02 mmol) was dissolved in EtOH (15 mL) followed by the addition of $\text{N}_2\text{H}_4\cdot\text{H}_2\text{O}$ (196 μL , 4.04 mmol) and AcOH (185 μL , 3.23 mmol). The reaction mixture was heated under reflux for 2 h, cooled to room temperature and evaporated at reduced pressure. Water (5 mL) was added to the residue and the resultant solution basified to pH 11 with 1M NaOH. The resulting mixture was extracted with EtOAc (3 \times 15 mL). The combined organic fractions were re-extracted with sat. NaCl (20 mL), dried over MgSO_4 , filtered, and evaporated at reduced pressure to give the crude product as a yellow oil. The residue was purified by column chromatography (DCM/MeOH/ NH_4OH 94:5:1) and gave 4-methyl-3-(*p*-tolyl)-1*H*-pyrazol-5-amine (MIPS-0001176) as a yellow oil (234 mg, 62%). ^1H NMR (400 MHz, CDCl_3): δ 7.36 (d, J = 8.1 Hz, 2H), 7.25 (d, J = 7.9 Hz,

2H), 2.39 (s, 3H), 2.04 (s, 3H). ^{13}C NMR (101 MHz, CDCl_3): δ 153.8 (C), 141.8 (C), 138.3 (C), 129.7 (2 \times CH), 127.9 (C), 127.1 (2 \times CH), 99.3 (C), 21.4 (CH_3), 7.9 (CH_3). ESI-TOF HRMS m/z ($\text{M}+\text{H}$) $^+$ $\text{C}_{11}\text{H}_{14}\text{N}_3^+$ calc. 188.1182, found 188.1191. The oil was converted to its corresponding hydrochloride salt. The amine was dissolved in EtOAc (20 mL) and HCl (1.0 M in Et_2O , 2.0 eq) was added dropwise whilst stirring. The mixture was evaporated at reduced pressure and the resultant solid suspended in Et_2O , filtered and washed with Et_2O . ^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ 7.50 (d, J = 8.0 Hz, 2H), 7.34 (d, J = 7.7 Hz, 2H), 2.36 (s, 3H), 2.07 (s, 3H).

5-(1*H*-Benzo[d]imidazol-2-yl)benzene-1,3-diamine (MIPS-0008405)

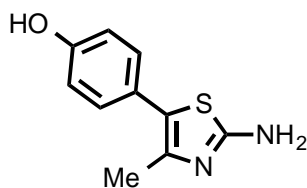


o-Phenylenediamine (1.00 g, 9.24 mmol), EDCI.HCl (2.13 g, 11.1 mmol), HOBT (1.50 g, 11.1 mmol), DIPEA (4.00 mL, 23.1 mmol) were dissolved in DMF (20 mL). 3,5-Dinitrobenzoic acid (1.96 g, 9.24 mmol) was added portionwise and the solution stirred for 3 h. The solution was poured onto ice/water (250 mL), stirred and the resultant precipitate was filtered and washed with water to give *N*-(2-aminophenyl)-3,5-dinitrobenzamide (872 mg, 31%) as a brown solid. ^1H NMR (400 MHz, DMSO): δ 10.31 (s, 1H), 9.19 (d, J = 2.1 Hz, 2H), 9.00 (t, J = 2.1 Hz, 1H), 7.14 (dd, J = 7.8, 1.4 Hz, 1H), 7.02 (td, J = 8.1, 1.5 Hz, 1H), 6.78 (dd, J = 8.1, 1.3 Hz, 1H), 6.60 (td, J = 7.7, 1.4 Hz, 1H), 5.11 (br s, 2H).

N-(2-Aminophenyl)-3,5-dinitrobenzamide (870 mg, mmol) was taken up in AcOH (10 mL) and heated at 100 $^\circ\text{C}$ for 16 h. The yellow solution was cooled to room temperature, poured onto ice/water (200 mL) and the mixture was carefully neutralised with 5M NaOH. The mixture was filtered and washed with water to give 2-(3,5-dinitrophenyl)-1*H*-benzo[d]imidazole (687 mg, 84%) as a light brown solid. ^1H NMR (400 MHz, DMSO): δ 13.60 (s, 1H), 9.33 (d, J = 2.1 Hz, 2H), 8.85 (t, J = 2.1 Hz, 1H), 7.76 (d, J = 7.8 Hz, 1H), 7.62 (d, J = 7.8 Hz, 1H), 7.30 (dt, J = 15.2, 6.4 Hz, 2H).

2-(3,5-Dinitrophenyl)-1*H*-benzo[d]imidazole (169 mg, 0.59 mmol) was taken up in AcOH (15 mL) and NH_4HCO_2 (187 mg, 2.97 mmol) added and sealed with a Suba-Seal. After repeated cycles of evacuation by vacuum and charging with N_2 , 10 % Pd/C (17 mg) was added and the mixture was heated at 50 $^\circ\text{C}$ for 16 h. The reaction was cooled to room temperature and after careful degassing, EtOH (100 mL) was added, filtered through CeliteTM and evaporated at reduced pressure to give a mixture of product, mono- and diacetamido products (129 mg). This mixture was taken up in a EtOH/1M NaOH (1:1, 20 mL) and heated at 100 $^\circ\text{C}$ for 16 h. The EtOH was evaporated at reduced pressure and the aqueous extracted with EtOAc (3 \times 10 mL), washed with sat. NaCl, dried over MgSO_4 , filtered and the filtrate evaporated at reduced pressure to give the crude product. The residue was purified by flash chromatography (DCM/MeOH/ NH_4OH , 89:10:1) and gave 5-(1*H*-benzo[d]imidazol-2-yl)benzene-1,3-diamine (MIPS-0008405) as a light brown solid (77 mg, 58%). ^1H NMR (400 MHz, MeOD): δ 7.59 – 7.53 (m, 2H), 7.25 – 7.18 (m, 2H), 6.79 (d, J = 2.0 Hz, 2H), 6.27 (t, J = 2.0 Hz, 1H). ^{13}C NMR (101 MHz, MeOD): δ 154.4 (C), 150.3 (4 \times C), 132.2 (C), 123.6 (2 \times CH), 115.6 (2 \times CH), 105.4 (2 \times CH), 103.5 (CH). ESI-TOF HRMS m/z ($\text{M}+\text{H}$) $^+$ $\text{C}_{13}\text{H}_{13}\text{N}_4^+$ calc. 225.1135, found 225.1130.

4-(2-Amino-4-methylthiazol-5-yl)phenol (MIPS-0008939)



2-Acetamido-4-methylthiazole (1.00 g, 6.40 mmol) was dissolved in MeCN (20 mL), NIS (2.16 g, 9.60 mmol) added and the mixture stirred at 25 °C for 30 min. The succinimide that formed was filtered, washed with MeCN and the filtrate evaporated at reduced pressure. The residue was dissolved in EtOAc (40 mL), washed with 1M Na₂S₂O₃ (2 × 20 mL), dried over MgSO₄, filtered and the filtrate removed at reduced to give *N*-(5-iodo-4-methylthiazol-2-yl)acetamide (1.15 g, 64%) as a yellow solid. ¹H NMR (400 MHz, CDCl₃): δ 9.99 (s, 1H), 2.35 (s, 3H), 2.23 (s, 3H).

To a degassed biphasic solution of THF (1.5 mL) and 1M Na₂CO₃ (0.5 mL), was added *N*-(5-iodo-4-methylthiazol-2-yl)acetamide (200 mg, 0.71 mmol), 4-hydroxyphenylboronic acid (103 mg, 0.74 mmol) and PdCl₂(PPh₃)₂ (50 mg, 0.07 mmol) and the sealed mixture heated at 100 °C for 2 h. the reaction mixture was cooled to room temperature, diluted with EtOAc (3 mL), the organic layer filtered through cotton wool and evaporated at reduced pressure. The residue was purified by column chromatography (hexane:EtOAc, 1:1) to give *N*-(5-(4-hydroxyphenyl)-4-methylthiazol-2-yl)acetamide (68 mg, 39%) as a light yellow solid. ¹H NMR (400 MHz, DMSO): δ 11.99 (s, 1H), 9.61 (s, 1H), 7.24 (d, *J* = 8.0 Hz, 2H), 6.82 (d, *J* = 8.0 Hz, 2H), 2.28 (s, 3H), 2.11 (s, 3H).

N-(5-(4-Hydroxyphenyl)-4-methylthiazol-2-yl)acetamide (68 mg, 0.27 mmol) was dissolved in a 1:1 mixture of EtOH/1M NaOH (5 mL) and heated at 100 °C for 16 h. The EtOH was evaporated at reduced pressure and the aqueous extracted with EtOAc (3 × 10 mL), washed with sat. NaCl, dried over MgSO₄, filtered, and the filtrate evaporated at reduced pressure to give the crude product. The residue was purified by column chromatography (hexane:EtOAc, 1:1) to give 4-(2-amino-4-methylthiazol-5-yl)phenol (**MIPS-0008939**) (45 mg, 80%) as a light yellow solid. ¹H NMR (400 MHz, MeOD): δ 7.16 (d, *J* = 8.0 Hz, 2H), 6.79 (d, *J* = 8.0 Hz, 2H), 2.17 (s, 3H). ¹³C NMR (101 MHz, DMSO): δ 164.9 (C), 155.9 (C), 141.5 (C), 129.4 (2 × CH), 123.6 (C), 117.8 (C), 115.4 (2 × CH), 15.9 (CH₃). ESI-TOF HRMS *m/z* (M+H)⁺ C₁₀H₁₁N₂OS⁺ calc. 207.0587, found 207.0582.

CHAPTER 3

Structure and dynamics of apical membrane antigen 1 from *Plasmodium falciparum* FVO

3.1 Declaration of Thesis Chapter 3

Declaration by candidate

In the case of **Chapter 3**, the nature and extent of my contribution to the work was the following:


Nature of contribution	Extent of contribution (%)
Prepared ^2H , ^{15}N , ^{13}C labelled AMA1 samples, recorded [^1H - ^{15}N]-TROSY and TROSY-HNCO NMR experiments, analysed NMR data and provided intellectual input in writing the manuscript	20

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
San S. Lim	Performed experiments, data analysis, manuscript preparation	
Wei Yang	Performed experiments, data analysis	
Komagal K. Sivaraman	Performed experiments	
Indu R. Chandrashekar	Manuscript preparation	
Itamar Kass	Performed experiments, data analysis, manuscript preparation	
Christopher A. MacRaid	Manuscript preparation	
Shane M. Devine	Manuscript preparation	
Cael O. Debono	Performed experiments	2.5
Robin F. Anders	Intellectual input, manuscript preparation	
Martin J. Scanlon	Intellectual input	
Peter J. Scammells	Intellectual input	
Raymond S. Norton	Intellectual input, manuscript preparation	
Sheena McGowan	Performed experiments, intellectual input, manuscript preparation	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

**Candidate's
Signature**

	Date 30/03/2016
---	---------------------------

**Main
Supervisor's
Signature**

	Date 30/03/2016
---	---------------------------

3.2 Introduction

Chapter 2 revealed interactions of small molecules with *Pf*AMA1 using NMR spectroscopy. In conjunction with the work presented in Chapter 2, the first X-ray crystal structure of FVO *Pf*AMA1 was determined and used for mapping the binding sites for small molecules as described in Chapter 2.

The new structure of FVO *Pf*AMA1 was compared with the structures of 3D7 *Pf*AMA1 and *P. vivax* AMA1. Overall, the FVO *Pf*AMA1 structure was found to be very similar to those from 3D7 *P. falciparum* and *P. vivax*. The DII loop end of the hydrophobic cleft appears to be conserved across different allelic forms of AMA1, representing an attractive site for strain-transcending therapeutic interventions. Polymorphic residues surrounding one end of the cleft (residues 175, 197, 200, 201, 204 and 225) were shown to restrict the cross-reactivity of inhibitory antibodies and peptides such as 1F9, IgNAR and R1. It has been postulated that the sequence polymorphism may result in local secondary structure change. Hence, the dynamics of flexible loops were analyzed to understand how their sequence polymorphism play a role in limiting the cross-reactivity of AMA1 antibodies and inhibitory peptides. However, it was observed that polymorphism did not result in any structural changes. As the results of this chapter have been published in *Biochemistry*, they are presented in the format of a published article in the immediate section below.

Structure and Dynamics of Apical Membrane Antigen 1 from *Plasmodium falciparum* FVO

San Sui Lim,[†] Wei Yang,[‡] Bankala Krishnarjuna,[†] Komagal Kannan Sivaraman,[‡] Indu R. Chandrashekar,[†] Itamar Kass,^{‡,§} Christopher A. MacRaild,[†] Shane M. Devine,[†] Cael O. Debono,[†] Robin F. Anders,^{†,||} Martin J. Scanlon,^{†,⊥} Peter J. Scammells,[†] Raymond S. Norton,^{*,†} and Sheena McGowan^{*,‡}

[†]Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

[‡]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia

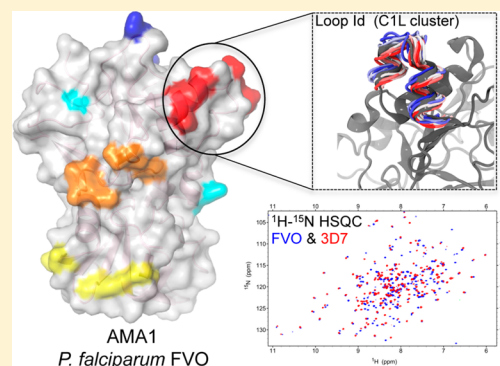
[§]Victorian Life Sciences Computation Initiative Life Sciences Computation Centre, Monash University, Clayton, Victoria 3800, Australia

^{||}Department of Biochemistry, La Trobe Institute for Molecular Science, La Trobe University, Bundoora, Victoria 3086, Australia

[⊥]Centre of Excellence for Coherent X-ray Science, Monash University, Parkville, Victoria 3052, Australia

Supporting Information

ABSTRACT: Apical membrane antigen 1 (AMA1) interacts with RON2 to form a protein complex that plays a key role in the invasion of host cells by malaria parasites. Blocking this protein–protein interaction represents a potential route to controlling malaria and related parasitic diseases, but the polymorphic nature of AMA1 has proven to be a major challenge to vaccine-induced antibodies and peptide inhibitors exerting strain-transcending inhibitory effects. Here we present the X-ray crystal structure of AMA1 domains I and II from *Plasmodium falciparum* strain FVO. We compare our new structure to those of AMA1 from *P. falciparum* 3D7 and *Plasmodium vivax*. A combination of normalized B factor analysis and computational methods has been used to investigate the flexibility of the domain I loops and how this correlates with their roles in determining the strain specificity of human antibody responses and inhibitory peptides. We also investigated the domain II loop, a key region involved in inhibitor binding, by comparison of multiple AMA1 crystal structures. Collectively, these results provide valuable insights that should contribute to the design of strain-transcending agents targeting *P. falciparum* AMA1.



Malaria is one of the most widespread infections, with more than 40% of the global population at risk of contracting the disease.^{1,2} Each year, there are approximately 250 million clinical cases of malaria that result in more than 600 000 deaths worldwide.² The majority of these deaths are due to *Plasmodium falciparum* infections occurring in young children in sub-Saharan Africa.² Although much less likely to cause death, *Plasmodium vivax* infections also contribute to a substantial malaria burden across the globe, with 70–80 million cases occurring annually.³ Although current artemisinin combination therapies have been highly effective against *Plasmodium* parasites, signs of resistance have already emerged.⁴ There is an urgent need to combat this threat using therapeutic agents that act against a broad range of parasite strains, especially those that have become resistant to available therapies.

Apical membrane antigen 1 (AMA1) forms part of the moving junction complex essential for erythrocyte invasion by *Plasmodium* merozoites, and ligands that disrupt AMA1 function inhibit the growth *in vitro* of *P. falciparum* asexual blood stages.^{5–9} Further, a conditional knockdown of PfAMA1

severely impaired the parasite's ability to invade red cells,¹⁰ and a complete gene knockout is not viable in *P. falciparum*.¹¹ AMA1 is a type I integral membrane protein,¹² the extracellular region of which consists of three domains based on the connectivities of its eight intramolecular disulfide bonds: an N-terminal domain I, a central domain II, and a C-terminal domain III.^{13,14} Antibodies to AMA1 can block red cell invasion by *P. falciparum* *in vitro* and contribute to the adaptive immune response that partially protects exposed individuals against malaria. AMA1 has been a leading candidate for inclusion in a vaccine against *P. falciparum*,¹⁵ but in a Phase IIb clinical trial in Mali, a 3D7 PfAMA1 vaccine provided protection against only a subset of *P. falciparum* AMA1 genotypes, reflecting the extensive sequence polymorphisms in this antigen.¹⁶ A bivalent 3D7 and FVO PfAMA1 vaccine, also tested in Phase IIb trials, failed to protect because of the poor immunogenicity of the

Received: September 25, 2014

Revised: October 30, 2014

Published: October 31, 2014

alum formulation.¹⁷ Currently, preclinical studies with multivalent vaccines (four to six AMA1 alleles) show promise, inducing a more broadly cross-reactive antibody response.^{18,19}

The two most extensively studied forms of PfAMA1, 3D7 and FVO, have 24 amino acid differences, and these polymorphic residues have been grouped into domain I–III clusters based on their spatial proximity on the X-ray crystal structure of 3D7 PfAMA1.²⁰ The domain I cluster is the most important of the three in mediating escape from inhibitory antibodies and was further classified into subclusters C1–C3. Within C1, the region termed C1L is particularly important for immune escape.^{18,21} In the preclinical studies of PfAMA1 vaccines, it was noted that antisera to 3D7 PfAMA1 were more strain-specific than antisera to FVO PfAMA1.

AMA1 has a hydrophobic cleft that is the site of interactions with its protein-binding partner RON2.^{9,22} The cleft is surrounded by six loops from domain I (loops Ia–If) and an extended loop from domain II (DII loop).²³ The DII loop appears to contain a strain-transcending epitope as the monoclonal antibody 4G2, which binds to the base of the DII loop, exhibits strain-independent inhibition of *P. falciparum*.⁵ We and others have proposed that small molecules targeted to the hydrophobic cleft may interrupt the AMA1–RON2 protein–protein interaction and provide a route to novel therapeutics in the form of protein–protein interaction inhibitors.^{24–26} However, polymorphic regions C1 and C3 surrounding one end of the cleft were shown to restrict the cross-reactivity of inhibitory antibodies and peptides such as IgNAR,⁷ 1F9,^{27,28} and R1.^{6,9,29} It has been postulated that sequence variation in the highly polymorphic C1L region (of the C1 subcluster) may result in local secondary structure changes.^{15,21} In particular, the presence of a two-turn helix in the Id loop has been questioned for the FVO allele because of the presence of a glycine residue at position 197.¹⁵

To answer fundamental questions regarding the impact of sequence diversity on AMA1, we have determined the first X-ray crystal structure of PfAMA1 from FVO. We show through X-ray crystallography, all-atom molecular dynamics, and nuclear magnetic resonance (NMR) spectroscopy that the sequence divergence does not result in structural changes that account for the strain-specific effects documented for inhibitory antibodies and peptides.

■ EXPERIMENTAL PROCEDURES

Protein Expression and Purification. Domains I and II of the ectodomains of FVO and 3D7 PfAMA1 (residues 104–438) were produced according to the protocol described by Lim et al.²⁴ except that the hexa-His tag was removed using TEV protease. The pure His-tagged proteins were dialyzed using membrane tubing (Spectra/Por 3, 3.5 kDa MWCO) against a 100-fold volume of 50 mM Tris (pH 8.0) under constant stirring at 4 °C overnight, and then TEV protease was added to the sample at a 1:30 ratio (w/w) and cleavage allowed to proceed at 4 °C for 48 h with gentle agitation. The mixture was filtered (0.2 μm) and loaded onto a 5 mL CHT I ceramic hydroxyapatite column (Bio-Rad). The cleaved protein was eluted using a linear gradient of 10 to 150 mM phosphate buffers (Na₂HPO₄ and NaH₂PO₄·H₂O) (pH 7.4) over 15 column volumes. The pooled fractions were concentrated and buffer exchanged into 20 mM Tris (pH 8.0) using an Amicon Ultra-4 centrifugal unit with an Ultracel-10 membrane (Millipore).

¹H NMR Spectroscopy and Size-Exclusion Chromatography. FVO PfAMA1 purified from hydroxyapatite chromatography was buffer exchanged into 20 mM phosphate buffer (Na₂HPO₄ and NaH₂PO₄·H₂O) and 50 mM NaCl (pH 7.4) containing 10% ²H₂O using a PD-10 desalting column (GE Healthcare). The sample was subsequently concentrated as described above to a final protein concentration of 50 μM. Part of this final product was used for ¹H NMR. A ¹H-detected pulse program incorporating the excitation sculpting scheme for water suppression was employed to characterize FVO PfAMA1. A total of 128 scans and 16K data points were acquired at 600 MHz on a Bruker Avance III spectrometer at 35 °C. The data were processed in Topspin 3.2 using an exponential multiplication function with 2 Hz line broadening. The water signal was used to reference the ¹H NMR spectrum. The final product was loaded onto a Superdex 200 10/30 GL column and eluted isocratically with 20 mM phosphate buffer (Na₂HPO₄ and NaH₂PO₄·H₂O) and 50 mM NaCl (pH 7.4) at a flow rate of 0.5 mL/min.

²H₂O Escherichia coli Adaptation. Fifty microliters of a competent *E. coli* BL21(DE3) glycerol stock, previously frozen at –80 °C, was thawed on ice for 10 min prior to adding 1 μL of plasmid carrying expression vector pPROEX HTb (Novagen) with FVO or 3D7 PfAMA1_[104–438] sequences. The mixture was left on ice for a further 30 min and then in water at 42 °C for 45 s. One milliliter of Luria broth (LB) was added to each sample, and the culture was incubated at 37 °C while being constantly shaken at 225 rpm. After 45 min, 50 μL of culture was spread over a LB plate containing 50% (v/v) ²H₂O and 100 μg/mL ampicillin and then incubated at 37 °C overnight. A single colony of the freshly transformed cells was inoculated into 10 mL of LB medium with 50% (v/v) ²H₂O and 100 μg/mL ampicillin. The culture was grown for 24 h at 37 °C while being constantly shaken at 225 rpm. The cell mixture was spread on a culture plate, and subsequently, a single colony of cells was incubated in growth medium as described above, except that the LB plate and medium prepared with 75% (v/v) ²H₂O were used instead. The final cell culture containing 75% (v/v) ²H₂O was stored at –80 °C with 20% (v/v) glycerol.

Isotopically Labeled AMA1. A scrape of the glycerol stock of ²H₂O-adapted *E. coli* was inoculated into LB medium with 75% (v/v) ²H₂O and 100 μg/mL ampicillin. The culture was incubated overnight at 37 °C while being constantly shaken at 225 rpm. The overnight culture was then centrifuged at 1500g for 15 min. The supernatant was decanted, and the cell pellets were resuspended in 2 volumes of optimized minimal medium³⁰ prepared with 100% (v/v) ²H₂O, 1 g/L ¹⁵NH₄Cl, and 8 g/L [¹³C]glucose. The cells were allowed to grow for 3 h while being shaken at 37 °C before being induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 24 h. The protein was then purified as described above.

¹H–¹⁵N HSQC and Three-Dimensional (3D) HNCO Experiments. ²H-, ¹⁵N-, and ¹³C-labeled 3D7 and FVO PfAMA1 were dissolved at concentrations of 300 and 75 μM, respectively, in 20 mM phosphate buffer (Na₂HPO₄ and NaH₂PO₄·H₂O) (pH 7.0) containing 50 mM L-arginine, 50 mM L-glutamic acid, 0.2% (w/v) protease inhibitor cocktail (Roche), 0.01% (v/v) sodium azide, and 10% (v/v) ²H₂O. The FVO PfAMA1 spectrum was acquired on a Bruker Avance III 600 MHz spectrometer at 35 °C. A spectrum of 3D7 PfAMA1 was acquired on a Bruker Avance III 800 MHz spectrometer at 30 °C. Both the ¹H–¹⁵N HSQC and 3D HNCO experiments

were conducted using pulse sequences with transverse relaxation-optimized spectroscopy (TROSY) effects.^{31,32} ¹H–¹⁵N HSQC spectra were acquired with 64 scans at 2048 and 256 data points for the ¹H and ¹⁵N dimensions, respectively. A total of 32 scans was recorded for the 3D HNCO experiments, with 2048, 128, and 128 data points for the ¹H, ¹⁵N, and ¹³C dimensions, respectively. Both the direct and indirect dimensions of ¹H–¹⁵N HSQC and 3D HNCO data were processed using a QSINE window with a phase shift of 2. Linear prediction using 32 coefficients was applied to all indirect dimensions.

R1–FVO PfAMA1 Interactions. A Biacore T200 biosensor was employed to measure the interaction between recombinant FVO PfAMA1 DI and DII and the R1 peptide (GL Biochem). Surface plasmon resonance (SPR) experiments were conducted essentially as described previously²⁴ except that dimethyl sulfoxide was not included in the running buffer. Approximately 8000 RU of protein was coupled in a single flow cell (1000 RU = 1 ng of protein/mm²). The binding of R1 peptide^{6,29} to FVO PfAMA1 was evaluated using a 2-fold serial dilution ranging in concentration from 15.6 to 500 μM.

Sequence Alignment and Analysis. All sequence alignments and analyses were performed using the UniProt online tool (<http://www.uniprot.org/>). Accession numbers for AMA1 sequences used in this study are as follows: UniProt entry Q9TY48 for FVO PfAMA1 and UniProt entry Q7KQK5 for 3D7 PfAMA1. The residues defining the AMA1 hydrophobic cleft and polymorphic sites were obtained from published literature.^{7,23,33–35}

Crystallization, X-ray Data Collection, Structure Determination, and Refinement. Crystallization conditions for FVO PfAMA1 were identified following a robotic broad screen using the IndexHT (Hampton Research) and JCSGPlus (Molecular Dimensions) crystal screens. Optimization of a single initial hit from the Index screen used the hanging drop vapor diffusion method, with a 1:1 (v/v) ratio of protein to mother liquor (well volume of 0.5 mL). Small, stacked crystals appeared after three months in 25% (v/v) polyethylene glycol 3350, 0.1 M HEPES (pH 7.5), and 0.2 M MgCl₂. A single crystal was separated from the stacked cluster and cryo-protected by the addition of 10% glycerol prior to data collection.

3D7 PfAMA1 crystals were grown in 12–15% (v/v) polyethylene glycol 3350, 0.02 M MES (pH 6.0), and 10 mM MnCl₂ as detailed in ref 23. 3D7 PfAMA1 crystals were dehydrated overnight in a reservoir solution with an increased level [35% (v/v)] of polyethylene glycol 3350 before cryo-stabilization in 38% (v/v) polyethylene glycol 3350, 0.088 M MES (pH 6.0), and 44 mM MnCl₂ for 6–8 h prior to data collection. For crystals used to test soaking solvents, 5% (v/v) methanol or Milli-Q water was added to the stabilization solution.

Data were collected at 100 K for all crystals using the Australian Synchrotron micro crystallography MX2 beamline 3ID1. Diffraction images were processed using XDS³⁶ and AIMLESS³⁷ from the CCP4 suite.³⁸ Five percent of each data set was flagged for calculation of R_{Free}^{39} with neither a sigma nor a low-resolution cutoff applied to the data. A summary of data collection statistics is provided in Table S1 of the Supporting Information.

Structure determination proceeded using the Molecular Replacement method and the program PHASER.⁴⁰ A search model for FVO PfAMA1 was constructed by removing the

solvent and flexible loops from a crystal structure of 3D7 PfAMA1 [Protein Data Bank (PDB) entry 1Z40]. A single clear peak in both the rotation and translation functions was evident and packed well within the asymmetric unit. Together with the unbiased features in the initial electron density maps, the correctness of the molecular replacement solution was confirmed. All subsequent model building and structural validation for FVO and 3D7 PfAMA1 structures was conducted using Phenix^{41,42} and COOT.⁴³ Solvent molecules were added only if they had acceptable hydrogen bonding geometry contacts of 2.5–3.5 Å with protein atoms or with existing solvent and were in good $2F_o - F_c$ and $F_o - F_c$ electron density. Hydrogen bonds (excluding water-mediated bonds) and salt bridges were calculated using PDBEPIA.⁴⁴ The coordinates and structure factors are available from the Protein Data Bank (entries 4R1A, 4R19, 4R1B, and 4R1C). Raw data and images are available from TARDIS⁴⁵ (www.tardis.edu.au).

B Factor Analysis. The B factors obtained from PDB files cannot be used directly, because the values may be on different scales because of the application of different refinement procedures.⁴⁶ To compare the B factors from different structures, the values were normalized as described by Parthasarathy et al.⁴⁷ The α B factor values were extracted from FVO PfAMA1 (PDB entry 4R1A) and *P. vivax* AMA1 (PDB entry 1W81) as well as the four 3D7 PfAMA1 crystal structures (PDB entries 1Z40, 4R19, 4R1B, and 4R1C), and normalized using the equation $B_{\text{normalized}} = (B_{\text{factor}} - B_{\text{mean}})/\sigma(B)$, where B_{mean} and $\sigma(B)$ are the mean value and standard deviation, respectively, of the distribution of observed thermal factors. The average values and standard deviations of normalized B factors were calculated for α -helical and β -sheet regions, loops Ia–If, and the DII loop. Average values were not calculated for regions in which more than half of the amino acid sequence was missing in the crystal structures.

Molecular Dynamics Simulations. The dynamics of three different AMA1 structures were studied. Atomic coordinates of *P. vivax* AMA1 (PDB entry 1W81, residues 49–383) and 3D7 PfAMA1 (PDB entry 1Z40, residues 104–438) were obtained from the Protein Data Bank. Coordinates for FVO PfAMA1 (residues 104–438) were obtained in this study. Missing atoms and residues were modeled using MOE 2012.10.⁴⁸ Each protein was solvated in a water cubic box consisting of TIP3P water molecules⁴⁹ with Na⁺ ions added to neutralize any charge. The minimum distance from the surface of the protein to any face of the water box was set to 12 Å for each simulation.

All-atom molecular dynamics (MD) simulations were performed using the NAMD 2.9 MD package⁵⁰ on the IBM Blue Gene/Q supercomputer of the Victorian Life Sciences Computation Initiative (VLSCI). Proteins were defined by the newly published and tested AMBER force field, FF12SB.^{51–53} Equilibration was performed in three stages. First, potential steric clashes in the initial configurations were relieved with 50000 steps of energy minimization. Initial velocities for each system were assigned randomly according to a Maxwell–Boltzmann distribution at 100 K. Each system was then heated to 300 K over 0.1 ns, with the protein harmonically restrained (10 kcal mol^{−1} Å^{−2}) under the canonical ensemble (NVT) conditions. Following this, each system was simulated for an additional 0.1 ns under the isothermal–isobaric ensemble (NPT) conditions with all heavy protein atoms harmonically restrained (10 kcal mol^{−1} Å^{−2}). Thereafter, each system was subjected to 250 ns of free simulation.

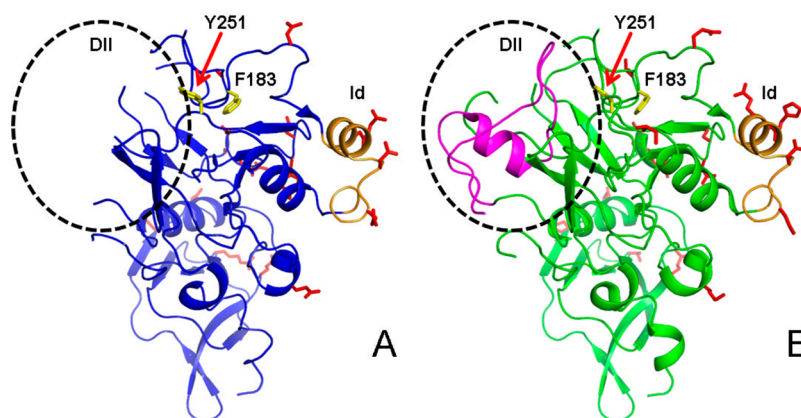


Figure 1. X-ray crystal structures of (A) FVO (PDB entry 4R1A) and (B) 3D7 (PDB entry 1Z40) *PfAMA1*. Residues that vary between FVO and 3D7 *PfAMA1* proteins are shown as red sticks. The Id loops are colored orange in both structures. The DII loop is indicated by the dotted circle in both panels, and the structure of the DII loop in 1Z40 is colored magenta (B). Two residues critical for the AMA1–RON2 interaction (Phe183 and Tyr251) are shown as yellow sticks in both structures. The hydrophobic cleft runs across the top of the molecule in this view, from the Id loop at one end to the DII loop at the other.

For all simulations, an integration time of 2 fs was used and the nonbonded cutoff length was set at 1 nm. All simulations were conducted at a constant temperature (300 K) and pressure (1 atm), using a Langevin damping coefficient of 0.5 fs^{−1}. For each simulation system, periodic boundary conditions (PBC) were used together with the particle mesh Ewald (PME) method for electrostatics interactions.⁵⁴ Electrostatics and VDW nonbonded forces were cut off at 1 nm. For each protein, three trajectories were run in parallel, differing only in the distribution of their initial velocities. System conformations were saved every 10 ps for subsequent analysis.

All the analyses were performed using the GROMACS 4.0.7 simulation software package.⁵⁵ Prior to MD analyses, translational and rotational motions were eliminated by superposition of each frame onto the initial conformation. The root-mean-square deviations (rmsds) of the backbone heavy atoms in each system were calculated relative to their corresponding initial minimized structures. Backbone root-mean-square fluctuations (rmsfs) were calculated for the productive phase (50–250 ns) of each simulation. All images were created by VMD version 1.9.1⁵⁶ or PyMOL version 1.3r2 (Schrodinger, LLC, 2010, The PyMOL Molecular Graphics System).

Electrostatic Surface Potential Calculations. Protein electrostatic potentials were calculated using APBS version 1.3.⁵⁷ Atom electrostatic charges were taken from the FF12SB force field. The electrostatic potential was visualized using PyMOL 1.3r2 (Schrodinger, LLC) with positive potential in blue and negative potential in red over the range of $-3k_bT/e_c$ to $3k_bT/e_c$, where k_b is the Boltzmann constant, T is the temperature (set to 300 K), and e_c is the electron charge.

Effect of Mutations on the Stability of the R1–3D7 *PfAMA1* Complex. The contributions of specific mutations to the overall thermodynamic stability of the R1–3D7 *PfAMA1* complex structure⁹ were estimated *in silico* with FoldX using default settings.⁵⁸ The reported Gibbs free energies are the differences between those of wild-type and mutated 3D7 *PfAMA1* in the context of the complex.

RESULTS

X-ray Crystal Structure of FVO *PfAMA1*. To generate protein crystals of FVO *PfAMA1*, a construct of FVO domains I and II (DI + II) equivalent to that of 3D7 *PfAMA1* was

produced.²³ The quality of our recombinant FVO *PfAMA1* protein was assessed using ¹H NMR spectroscopy and size-exclusion chromatography (Figure S1 of the Supporting Information). Overall, good signal dispersion was observed in the ¹H NMR spectrum, with methyl proton signals at −0.48 and −0.52 ppm as well as amide proton signals beyond 9 ppm (Figure S1A of the Supporting Information); the spectrum was consistent with a single folded product. During size-exclusion chromatography, FVO *PfAMA1* eluted as a single peak consistent with a monomeric form of the protein (38 kDa) (Figure S1B of the Supporting Information).

The X-ray crystal structure of FVO *PfAMA1* was determined to 2.0 Å with final R and R_{free} values of 19.5 and 25.5%, respectively (Figure 1A and Table S1 of the Supporting Information). FVO *PfAMA1* crystallized with one molecule in the asymmetric unit in space group C2₁. Seven α -helical and 16 β -sheet regions were identified in the final FVO *PfAMA1* structure. Similar to the 3D7 *PfAMA1* (Figure 1B, PDB entry 1Z40) and *P. vivax* AMA1 structures (PDB entry 1W81), both DI and DII of FVO *PfAMA1* formed PAN folds that consist of a two-turn α -helix packed against a five-stranded β -sheet. The two PAN folds pack against each other to form the protein core, as seen in the 3D7 *PfAMA1*²³ (PDB entry 1Z40; 0.27 Å rmsd over 219 C α atoms) and *P. vivax*⁵⁹ AMA1 (PDB entry 1Z40; 0.29 Å rmsd over 197 C α atoms) structures. The DI + II sequences of 3D7 and FVO *PfAMA1* have 21 amino acid residue differences [sharing 94% sequence identity (Figure S2 of the Supporting Information)].

All residue differences found in the C1 (residues 187, 190, 196, 197, 200, 204, 206, and 225) and C2 (residues 242, 243, 282, 283, and 285) clusters can be observed in our FVO *PfAMA1* structure (Figure S2 of the Supporting Information). Structural analysis of regions in the vicinity of the C1L (residues 196, 197, 200, 204, and 206), C2, and DII clusters, as well as residues 167 and 300, did not reveal any significant structural differences between FVO and 3D7 *PfAMA1* DI + II (Figure S3 of the Supporting Information). Structural differences were also not found in regions around residues 187 and 190 within the C1 cluster. Of the regions in the proximity of residue 225, residues 226–232 are disordered in FVO *PfAMA1* but ordered in 3D7 *PfAMA1*. Structural comparisons could not be performed for C3 because residues 175 and 267, which

constitute this cluster, are disordered in both structures. Nine of the 12 residues that define the hydrophobic cleft could be observed in our FVO *Pf*AMA1 structure (V169, L176, F183, I190, Y202, V208, M224, Y251, and I252);⁷ the three remaining hydrophobic residues (M273 from DI and L357 and F367 from DII) are disordered in the FVO *Pf*AMA1 structure.

The structure of FVO *Pf*AMA1 showed numerous disordered loops in both domains. In DI, the disordered residues not observed in the density (160–163, 173–176, 226–232, and 258–273) correspond to loops Ib and If and part of the Ie loop (Figure 1A). There are 38 residues missing in DII [351–388 (Figure 1A)], which correspond to most of the loop DII structure found in the 3D7 *Pf*AMA1 structure (Figure 1B).

Flexibility of the Loops in Different Forms of AMA1.

To ascertain whether the sequence polymorphisms might provide a structure-based “escape” route from the host immune response and inhibitor binding, we analyzed the loop flexibility of our FVO *Pf*AMA1 structure in comparison to those of the published 3D7 *Pf*AMA1 (PDB entry 1Z40) and *P. vivax* AMA1 (PDB entry 1W81) structures. We compared normalized *B* factors from all *Cα* atoms from each AMA1. The normalized *B* values were expressed in units of standard deviations about the mean *Cα* *B* factor for the corresponding structure; therefore, regions that are more rigid in a protein would have low normalized *B* factors, whereas flexible regions would have high normalized *B* values. This analysis showed that loops Ib and If (polymorphic cluster C3) are highly flexible regions in all three structures (Figure 2 and Table S2 of the Supporting Information). Loops Ia, Ic (residues 187 and 190 of the C1 cluster), and Ie (residue 225 of the C1 cluster) are mobile in

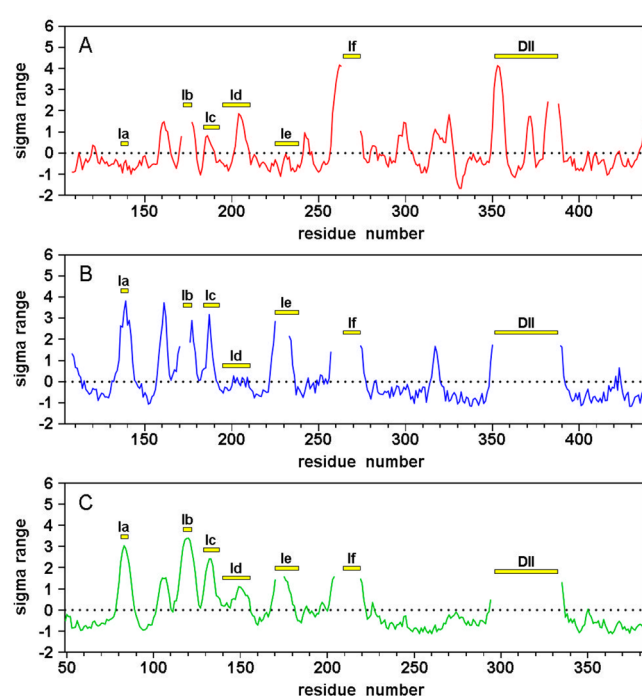


Figure 2. Flexibility of Ia–If and DII loops (identified with yellow bars). Normalized *Cα* *B* factors of (A) 3D7 *Pf*AMA1 (PDB entry 1Z40), (B) FVO *Pf*AMA1 (PDB entry 4R1A), and (C) *P. vivax* AMA1 (PDB entry 1W81). Residues 49–383 for the *P. vivax* AMA1 sequence are equivalent to residues 104–438 in the FVO and 3D7 *Pf*AMA1 sequences.

FVO *Pf*AMA1 and *P. vivax* AMA1 but exhibit limited mobility in 3D7 *Pf*AMA1. The Id loop (C1L cluster) appears to be more rigid in FVO *Pf*AMA1 than in 3D7 and *P. vivax* AMA1, but inferring the biological relevance of this result is difficult because of the presence of extensive crystal contacts made by residues within the Id loop of FVO *Pf*AMA1 (Figure S4 of the Supporting Information; see below).

In an effort to further assess the flexibility of the three AMA1 proteins, we undertook all-atom molecular dynamics simulations. Throughout our MD simulations, all systems were found to be stable following an initial structural rearrangement that took place early in the simulations (30–50 ns). Therefore, all subsequent analyses were conducted for the last 200 ns of each simulation (productive stage). Calculated rmsd values for the productive stage of MD simulations [3D7 *Pf*AMA1, 0.19 ± 0.04 nm; FVO *Pf*AMA1, 0.20 ± 0.03 nm; and *P. vivax* AMA1, 0.30 ± 0.05 nm (Figure S5 of the Supporting Information)] indicated that all systems were stable and that the *P. vivax* protein was slightly more flexible than the other proteins [particularly in the region of loop Ia (Figure S6 of the Supporting Information)]. The rmsf results (Figure S6 of the Supporting Information) show that the fluctuation patterns of the three proteins were similar, with peak fluctuations occurring in the same loops of each AMA1 structure (loops Ib, Ic, Ie, and If), which is consistent with the *B* factor analysis. The Ie loop appears to be more flexible in FVO *Pf*AMA1 than in 3D7 *Pf*AMA1. This observation may explain why the Ie loop is disordered in the FVO *Pf*AMA1 crystal structure but not in 3D7 *Pf*AMA1. On the basis of the rmsf results, the Id loop is flexible in all three AMA1 structures, which supports our earlier conclusions that the rigidity of the Id loop (Figure 2B) seen in FVO *Pf*AMA1 was a consequence of crystal contacts.

Crystallization artifacts, such as crystal contacts, can often complicate structural comparisons and implications drawn from static structures. The presence of stabilizing crystal contacts was identified in the original 3D7 *Pf*AMA1 structure by Bai et al.²³ In the search for suitable crystallization conditions of 3D7 *Pf*AMA1 for our fragment screening campaign,²⁴ we noticed that such crystal contacts could be modulated by soaking crystals in different solvents. This allowed us to examine the effect of these contacts on the conformation and/or flexibility of the surface loops in 3D7 *Pf*AMA1. We reproduced the original crystal conditions of Bai et al.,²³ determined the 1.8 Å X-ray crystal structure (PDB entry 4R19), and showed that it was identical to the published structure (PDB entry 1Z40; 0.118 Å rmsd over 275 *Cα* atoms). Under these conditions, 3D7 *Pf*AMA1 crystallized in a *P*3₁ space group with two molecules per asymmetric unit and showed crystal contacts that potentially stabilize the Ia–If and DII loops surrounding the hydrophobic cleft (Table S3 of the Supporting Information). Subjecting the crystals to 5% methanol (PDB entry 4R1C) or water (PDB entry 4R1B) during their stabilization immediately prior to data collection produced changes in the space group and unit cell dimensions (Table S1 of the Supporting Information), and these two new 3D7 *Pf*AMA1 structures showed different degrees of flexibility within the Ia, Ic, and Ie loops relative to the published 3D7 *Pf*AMA1 structure (Table S2 of the Supporting Information). The 3D7 *Pf*AMA1 structures from the Milli-Q water- and methanol-treated crystals have higher average normalized *B* factors for the Ia, Ic, and Ie loops compared to those of the 1Z40 structure (Figures 3 and 4). The Ie loop of FVO *Pf*AMA1, which is disordered, is ordered in all the 3D7 *Pf*AMA1 structures.

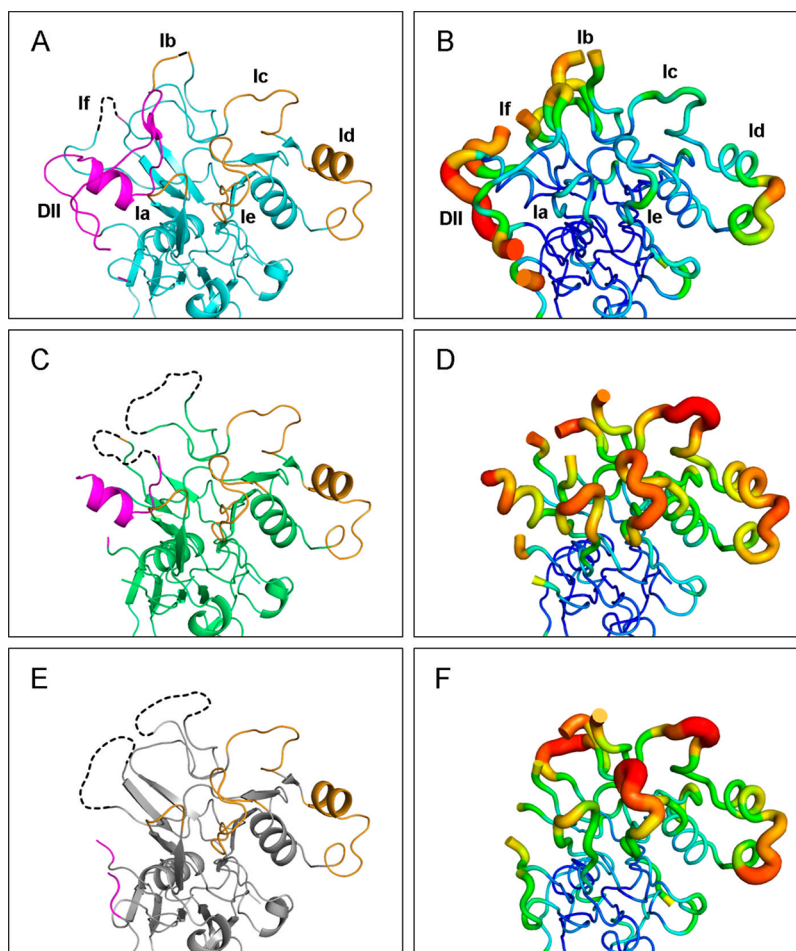


Figure 3. Cartoon model (left) and B factor putty (right) of the X-ray crystal structures of 3D7 *PfAMA1*: (A and B) 3D7 crystal structure from ref 23, (C and D) 3D7 crystal treated with Milli-Q water (PDB entry 4R1B), and (E and F) 3D7 crystal treated with methanol (PDB entry 4R1C).

However, it is difficult to determine if there is a real difference in the flexibility of the Ie loop as crystal packing is found in all the 3D7 *PfAMA1* structures (Table S3 of the Supporting Information). The Ib and If loops are disordered in both the water- and methanol-treated 3D7 *PfAMA1* crystals, indicating that these regions are highly mobile in the protein (Figures 3 and 4 and Table S2 of the Supporting Information).

Conformational Flexibility of the DII Loop. The DII loop in both *P. vivax* AMA1 and our FVO *PfAMA1* (Figure 1A) structure is disordered. This is in contrast to the published 3D7 *PfAMA1* structure in which only five residues are missing from the DII loop (residues 383–387) (Figure 1B). However, there was missing density for DII loop residues 370–387 in our Milli-Q water-treated 3D7 *PfAMA1* structure and residues 351–387 in the methanol-treated 3D7 *PfAMA1* structure, similar to our FVO *PfAMA1* structure (Figures 1B and 3). MD simulations showed that, although the N- and C-termini of the DII loop are highly mobile for FVO *PfAMA1*, 3D7 *PfAMA1*, and *P. vivax* AMA1 (Figure S6 of the Supporting Information), large conformational changes or movements of the α -helix at the center of the DII loop were not observed in any of these proteins. This implies that, despite its flexibility, the DII loop undergoes slow conformational exchange, beyond the time scale sampled in our MD simulations.

The original 3D7 *PfAMA1* structure (PDB entry 1Z40) provides support for the DII loop being ordered as a consequence of crystal contacts.²³ Our investigation of these

contacts found that there were 9 and 13 residues from neighboring molecules close to (<4.0 Å) the DII loop of reference 3D7 *PfAMA1* chains A and E, respectively (Figure S7A of the Supporting Information). In particular, we found that both Glu354 (O ϵ 1) and His356 (NH) of chain E formed hydrogen bonds with the main and side chains of Ser423 of a symmetry-related molecule (Figure S7A of the Supporting Information). The DII loop of chain A was stabilized by polar interactions between Lys363 (N ζ) and Asp317 of a symmetry-related molecule. In our FVO structure, there are no crystal contacts close to the DII loop, possibly allowing the loop to populate different positions within the crystal lattice (Figure S7B of the Supporting Information).

NMR studies of FVO and 3D7 *PfAMA1* were undertaken to further assess their flexibility in solution. Both 2D ^1H – ^{15}N HSQC and 3D HNCOC spectra were acquired. Of the 316 amide backbone NMR resonances expected in both forms of *PfAMA1*, only 261 and 250 peaks were identified in the HSQC spectra of FVO and 3D7 *PfAMA1*, respectively (Figure S8 of the Supporting Information). The discrepancies between the expected and observed resonances are due to regions of the protein that undergo conformational exchange on the intermediate time scale (microseconds to milliseconds).⁶⁰ NMR signals from regions with such motion are often broadened beyond detection in multidimensional NMR experiments. In the HNCOC spectra, there were fewer resonances in 3D7 *PfAMA1* than in FVO *PfAMA1* [248 and

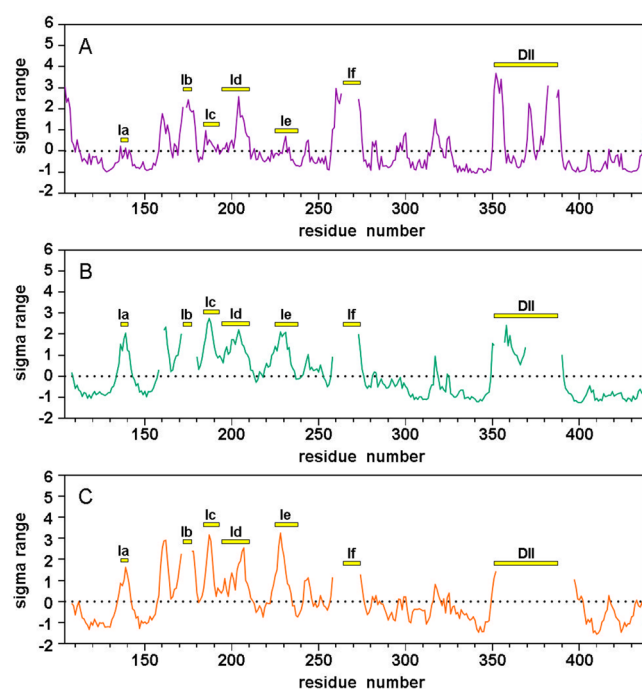


Figure 4. Flexibility of Ia–If and DII loops (identified with yellow bars) in 3D7 *PfAMA1*. Normalized Ca B factors of (A) the 3D7 original condition,²³ (B) the 3D7 crystal treated with Milli-Q water, and (C) the 3D7 crystal treated with methanol.

216 peaks in FVO and 3D7 *PfAMA1*, respectively (Figure S9 of the Supporting Information)]. This difference suggests that 3D7 *PfAMA1* has a slightly greater number of backbone resonances that are broadened by intermediate conformational exchange than FVO *PfAMA1*. This is in contrast to what is predicted from comparison of the published 3D7 *PfAMA1*

structure (PDB entry 1Z40), which is highly ordered throughout the molecule, with our crystal structures of both FVO *PfAMA1* and 3D7 *PfAMA1* in different solvents. In summary, these data indicate that in solution both FVO *PfAMA1* and 3D7 *PfAMA1* contain significant regions of disordered structure, some of which are undergoing conformational exchange on an intermediate time scale that produces substantial broadening of NMR resonances.⁶⁰

Mapping the Strain Variation of the Inhibitory Peptide R1. The R1 peptide,^{6,29} identified by phage display, inhibits red cell invasion by merozoites of 3D7 *P. falciparum* and related strains with a 50% inhibitory concentration (IC_{50}) of $\sim 0.1 \mu\text{M}$. X-ray crystal structures^{23,25} show that R1 contacts three polymorphic residues (Tyr175, Met224, and Ile225) in 3D7 *PfAMA1*, with residues at positions 175 and 225 being important determinants of R1 strain specificity. Substitution of these residues as in W2mef (I225E) or HB3 (Y175D and I225N) significantly reduced the peptide's inhibitory effect.⁹ FVO *PfAMA1* also has Y175D and I225N substitutions (Figure S2 of the Supporting Information). Accordingly, we found that R1 binds weakly to FVO *PfAMA1*, with an estimated K_d of $\geq 500 \mu\text{M}$ (Figure S10 of the Supporting Information).

To investigate why R1 binds so weakly to FVO *PfAMA1*, we estimated the effect of mutations on binding of R1 by FVO *PfAMA1*. The crystal structure of the 3D7 *PfAMA1*–R1 complex (PDB entry 3SRJ) was used as a template, and four FVO *PfAMA1* sequence variations found in the R1-binding cleft (Y175D, M190I, D204N, and I225N) were generated by FoldX 3.0^{61,62} to mimic an apparent FVO *PfAMA1*–R1 complex-binding structure (Figure 5). Single mutations were also generated to compare the individual effects on the binding energy. Analysis of point mutations shows that the changes at positions 175 and 225 (Y to D and I to N, respectively) had the largest effect on stability, with increases in free energy of 6 and

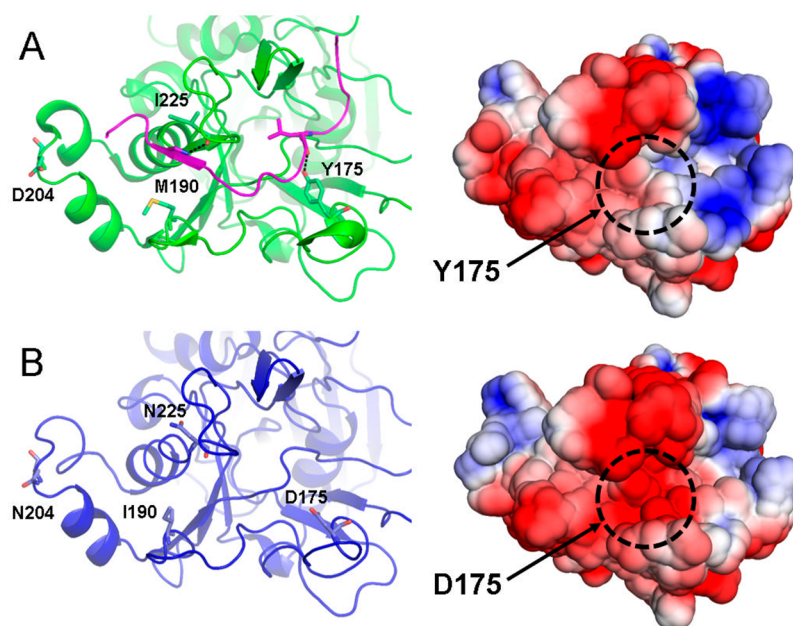


Figure 5. Comparison between the R1–3D7 *PfAMA1* complex and free FVO *PfAMA1*. (A) R1–3D7 *PfAMA1* crystal structure (PDB entry 3SRJ). (B) FVO *PfAMA1* model. Electrostatic potentials mapped on the solvent accessible surface for the R1-bound *PfAMA1* structures are shown on the right. Sequence variations between 3D7 and FVO *PfAMA1* are shown as sticks. R1 is shown as a magenta ribbon. Protein surfaces are color-coded according to electrostatic potential gradient, where positively and negatively charged areas are colored blue and red (iso-values from $3k_bT/e_c$ to $-3k_bT/e_c$), respectively.

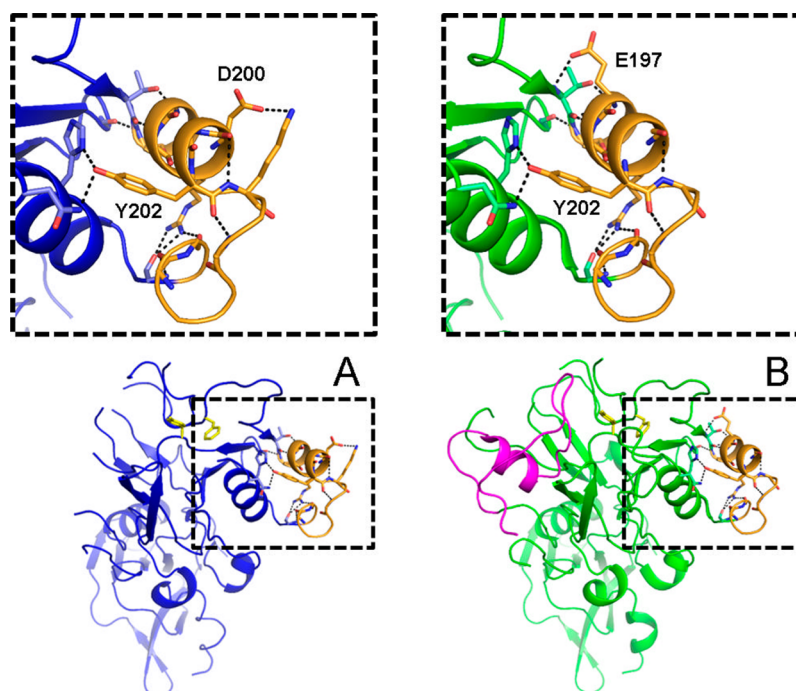


Figure 6. Cartoon showing polar interactions that stabilize the Id loop (orange) in (A) FVO (blue) and (B) 3D7 (green) *PfAMA1*. The DII loop is colored magenta. Residues Phe183 and Tyr251 are shown as yellow sticks.

3 kcal/mol, respectively (Table S4 of the Supporting Information). These energy changes are presumably due to the loss of two hydrogen bonds between R1 and AMA1 at these two positions (Figure 5). The M190I change resulted in a small increase in free energy (1 kcal/mol), suggesting that the increase in cleft volume and hydrophobicity resulting from the M190I change has an effect on R1 binding. Distal to the R1 interaction sites, the D204N change has no effect on the complex. Taken together, the four changes were estimated to contribute to a total increase in ΔG of 10 kcal/mol and significant changes in the hydrophobic cleft, creating an unfavorable environment for the R1 peptide (Table S4 of the Supporting Information).

DISCUSSION

AMA1 is implicated in the invasion of host cells by malaria parasites as well as other apicomplexan parasites.^{22,25,63,64} Sequence comparison of FVO and 3D7 *PfAMA1* domains I and II identifies 21 amino acid differences that occur exclusively at the polymorphic face of AMA1.^{23,33,35} However, despite earlier hypotheses,^{15,21} our FVO structure shows that these changes do not influence the overall fold of AMA1. The Ia–If loops that surround most of the hydrophobic cleft in AMA1 all display some level of disorder in the electron density from the FVO *PfAMA1*, 3D7 *PfAMA1*, and *P. vivax* structures. This apparent loop flexibility coupled with the polymorphic residues found in most of these regions provides AMA1 with an effective means of restricting cross-strain inhibitory activities of various ligands such as the R1 peptide and poses a challenge to efforts to design a vaccine or therapeutic agents effective against a broad range of strains and species of *Plasmodium* parasites.

Polymorphic residues within loop Id (C1L cluster) are important in mediating escape against AMA1 antibodies induced by *P. falciparum* infections or in vaccine trials.^{23,25,65} Residue 197 appears to be one of the most important residues in AMA1 responsible for immune escape.^{65,66} In this study, we

have shown that residue changes, including a Gly at position 197, do not result in structural changes in this region. This includes the two-turn helix within C1L. Residues 196, 204, and 206 are not engaged in any polar interactions and are unlikely to be important in stabilizing the structure of the Id loop (Table S5 of the Supporting Information). The main chain atoms of residues 197 and 200 in both strains form polar interactions with main chain atoms of Thr194 and Lys203, respectively. Substitution of Glu with Gly at position 197 in FVO *PfAMA1* prevents the polar interaction between Thr194 [O] and Glu197 [O γ 1] observed in 3D7. The H200D change at position 200 results in an additional side chain electrostatic interaction between Asp200 [O δ 1] and Lys203 [N ζ] in the FVO form of AMA1. The extensive pattern of hydrogen bonds stabilizing the Id loop is largely conserved in both strains (Figure 6 and Table S5 of the Supporting Information). These interactions appear to stabilize the structure and permit radical changes at polymorphic sites in the Id loop without causing significant conformational changes. This implies that the polymorphic nature of this region does not affect the structure of the protein and that immune escape arises largely from changes in properties of individual side chains.

Unlike the published structures of 3D7 *PfAMA1* and one 1F9–3D7 *PfAMA1* complex (PDB entries 1Z40 and 2Q8A, respectively), in which the conserved DII loop is partially ordered, the DII loop of FVO *PfAMA1* is completely disordered. This is similar to the case in another crystal form of the 1F9–3D7 *PfAMA1* complex (PDB entry 2Q8B)²⁷ and the *P. vivax* AMA1 structure (PDB entry 1W81).⁵⁹ The flexibility of the DII loop is linked to a conformational change that allows the protein to interact with its protein-binding partner, RON2 (PDB entry 3ZWZ).²⁷ By sequence and structural superposition, it appears that the intraprotein contacts involved in stabilizing the local secondary structure of the DII loop are maintained in FVO *PfAMA1*. The differences observed between the crystal structures of the two

forms of AMA1 reflect different individual conformational states captured under the different crystallization conditions rather than an inherently greater level of disorder in the FVO PfAMA1 DII loop. This view is further supported by our 3D7 PfAMA1 structures, in which different extents of crystal contacts gave rise to different degrees of order in the DII loop.

In conclusion, our results show that the overall structure, including the flexible nature of the DII loop, is conserved between the FVO and 3D7 forms of AMA1. The interacting interfaces between DI and the DII loop consist of invariant residues across all *P. falciparum* strains. The structural conservation of DI and DII appears likely to be conserved across all allelic forms of AMA1 and hence represents an attractive site for strain-transcending therapeutic interventions. Given that DII loop displacement is associated with formation of the AMA1–RON2 complex, it is conceivable that stabilizing the DII loop in its ordered state would be inhibitory to AMA1 function.⁹ Recently, we have undertaken a fragment-based screening campaign against AMA1 to identify chemical scaffolds capable of inhibiting the protein–protein interactions.²⁴ The structure described here will help guide the design of small molecule inhibitors of AMA1 with broad strain specificity.

■ ASSOCIATED CONTENT

■ Supporting Information

¹H NMR and size-exclusion chromatography of FVO PfAMA1 (Figure S1), sequence alignment of 3D7 and FVO PfAMA1 (Figure S2), 3D7 and FVO PfAMA1 sequence variations in the D1 and D2 polymorphic clusters (Figure S3), crystal packing of the FVO PfAMA1 Id loop (Figure S4), rmsds of FVO PfAMA1, 3D7 PfAMA1, and *P. vivax* AMA1 (Figure S5), rmsfs of FVO PfAMA1, 3D7 PfAMA1, and *P. vivax* AMA1 (Figure S6), crystal packing against the PfAMA1 DII loop (Figure S7), ¹H–¹⁵N HSQC spectra of FVO and 3D7 PfAMA1 (Figure S8), ¹H–¹³C projection of HNCO spectra of FVO and 3D7 PfAMA1 (Figure S9), SPR of R1–FVO PfAMA1 interaction (Figure S10), X-ray data collection and refinement statistics (Table S1), average values of normalized B factors for Ia–If and DII loops (Table S2), crystal packing in AMA1 crystal structures (Table S3), FoldX-calculated energies (Table S4), and polar interactions involved in stabilizing the AMA1 Id loops (Table S5). This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: ray.norton@monash.edu. Telephone: +613 9903 9167. Fax: +613 9903 9582.

*E-mail: sheena.mcgowan@monash.edu. Telephone: +613 9902 9309. Fax: +613 9902 9500.

Funding

S.M. is an Australian Research Council Future Fellow (FT100100690). R.S.N. acknowledges fellowship support from the National Health and Medical Research Council of Australia. This work was supported by the National Health and Medical Research Council (Project Grant 1025150). This research was supported by the Victorian Life Sciences Computation Initiative Life Sciences Computation Centre (RAS Grant VR0316), a collaboration among Melbourne, Monash, and La Trobe Universities and an initiative of the Victorian Government of Australia.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Nyssa Drinkwater for technical assistance. We thank the Australian Synchrotron for beam time and technical assistance.

■ ABBREVIATIONS

AMA1, apical membrane antigen 1; NMR, nuclear magnetic resonance; HSQC, heteronuclear single-quantum coherence; Pf, *P. falciparum*; TROSY, transverse relaxation-optimized spectroscopy; MD, molecular dynamics; SPR, surface plasmon resonance; MWCO, molecular weight cutoff; TEV, tobacco etch virus; LB, Luria broth; IPTG, isopropyl β -D-1-thiogalactopyranoside; rmsd, root-mean-square deviation; rmsf, root-mean-square fluctuation; PEG, polyethylene glycol; MES, 2-(N-morpholino)ethanesulfonic acid; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; PME, particle mesh Ewald.

■ REFERENCES

- (1) White, N. J., Dondorp, A. M., Faiz, A., Mishra, S., and Hien, T. T. (2012) New global estimates of malaria deaths. *Lancet* 380, 559–560.
- (2) *World Malaria Report* (2013) World Health Organization, Geneva.
- (3) Mendis, K., Sina, B. J., Marchesini, P., and Carter, R. (2001) The neglected burden of *Plasmodium vivax* malaria. *Am. J. Trop. Med. Hyg.* 64, 97–106.
- (4) Duffy, P. E., and Sibley, C. H. (2005) Are we losing artemisinin combination therapy already? *Lancet* 366, 1908–1909.
- (5) Collins, C. R., Withers-Martinez, C., Bentley, G. A., Batchelor, A. H., Thomas, A. W., and Blackman, M. J. (2007) Fine mapping of an epitope recognized by an invasion-inhibitory monoclonal antibody on the malaria vaccine candidate apical membrane antigen 1. *J. Biol. Chem.* 282, 7431–7441.
- (6) Harris, K. S., Casey, J. L., Coley, A. M., Karas, J. A., Sabo, J. K., Tan, Y. Y., Dolezal, O., Norton, R. S., Hughes, A. B., Scanlon, D., and Foley, M. (2009) Rapid optimization of a peptide inhibitor of malaria parasite invasion by comprehensive N-methyl scanning. *J. Biol. Chem.* 284, 9361–9371.
- (7) Henderson, K. A., Streltsov, V. A., Coley, A. M., Dolezal, O., Hudson, P. J., Batchelor, A. H., Gupta, A., Bai, T., Murphy, V. J., Anders, R. F., Foley, M., and Nuttall, S. D. (2007) Structure of an IgNAR-AMA1 complex: Targeting a conserved hydrophobic cleft broadens malarial strain recognition. *Structure* 15, 1452–1466.
- (8) Li, F., Dlugowski, A., Coley, A. M., Thomas, A., Tilley, L., Anders, R. F., and Foley, M. (2002) Phage-displayed peptides bind to the malarial protein apical membrane antigen-1 and inhibit the merozoite invasion of host erythrocytes. *J. Biol. Chem.* 277, 50303–50310.
- (9) Vulliez-Le Normand, B., Tonkin, M. L., Lamarque, M. H., Langer, S., Hoos, S., Roques, M., Saul, F. A., Faber, B. W., Bentley, G. A., Boulanger, M. J., and Lebrun, M. (2012) Structural and functional insights into the malaria parasite moving junction complex. *PLoS Pathog.* 8, e1002755.
- (10) Yap, A., Azevedo, M. F., Gilson, P. R., Weiss, G. E., O'Neill, M. T., Wilson, D. W., Crabb, B. S., and Cowman, A. F. (2014) Conditional expression of apical membrane antigen 1 in *Plasmodium falciparum* shows it is required for erythrocyte invasion by merozoites. *Cell. Microbiol.* 16, 642–656.
- (11) Triglia, T., Healer, J., Caruana, S. R., Hodder, A. N., Anders, R. F., Crabb, B. S., and Cowman, A. F. (2000) Apical membrane antigen 1 plays a central role in erythrocyte invasion by *Plasmodium* species. *Mol. Microbiol.* 38, 706–718.
- (12) Waters, A. P., Thomas, A. W., Deans, J. A., Mitchell, G. H., Hudson, D. E., Miller, L. H., McCutchan, T. F., and Cohen, S. (1990) A merozoite receptor protein from *Plasmodium knowlesi* is highly

conserved and distributed throughout *Plasmodium*. *J. Biol. Chem.* 265, 17974–17979.

(13) Hodder, A. N., Crewther, P. E., Matthew, M. L., Reid, G. E., Moritz, R. L., Simpson, R. J., and Anders, R. F. (1996) The disulfide bond structure of *Plasmodium* apical membrane antigen-1. *J. Biol. Chem.* 271, 29446–29452.

(14) Nair, M., Hinds, M. G., Coley, A. M., Hodder, A. N., Foley, M., Anders, R. F., and Norton, R. S. (2002) Structure of domain III of the blood-stage malaria vaccine candidate, *Plasmodium falciparum* apical membrane antigen 1 (AMA1). *J. Mol. Biol.* 322, 741–753.

(15) Anders, R. F., Adda, C. G., Foley, M., and Norton, R. S. (2010) Recombinant protein vaccines against the asexual blood stages of *Plasmodium falciparum*. *Hum. Vaccines* 6, 39–53.

(16) Thera, M. A., Doumbo, O. K., Coulibaly, D., Laurens, M. B., Ouattara, A., Kone, A. K., Guindo, A. B., Traore, K., Traore, I., Kouriba, B., Diallo, D. A., Diarra, I., Daou, M., Dolo, A., Tolo, Y., Sissoko, M. S., Niangaly, A., Sissoko, M., Takala-Harrison, S., Lyke, K. E., Wu, Y., Blackwelder, W. C., Godeaux, O., Vekemans, J., Dubois, M. C., Ballou, W. R., Cohen, J., Thompson, D., Dube, T., Soisson, L., Diggs, C. L., House, B., Lanar, D. E., Dutta, S., Heppner, D. G., Jr., and Plowe, C. V. (2011) A field trial to assess a blood-stage malaria vaccine. *N. Engl. J. Med.* 365, 1004–1013.

(17) Ouattara, A., Mu, J. B., Takala-Harrison, S., Saye, R., Sagara, I., Dicko, A., Niangaly, A., Duan, J. H., Ellis, R. D., Miller, L. H., Su, X. Z., Plowe, C. V., and Doumbo, O. K. (2010) Lack of allele-specific efficacy of a bivalent AMA1 malaria vaccine. *Malar. J.* 9, 175.

(18) Dutta, S., Dlugosz, L. S., Drew, D. R., Ge, X., Ababacar, D., Rovira, Y. I., Moch, J. K., Shi, M., Long, C. A., Foley, M., Beeson, J. G., Anders, R. F., Miura, K., Haynes, J. D., and Batchelor, A. H. (2013) Overcoming antigenic diversity by enhancing the immunogenicity of conserved epitopes on the malaria vaccine candidate apical membrane antigen-1. *PLoS Pathog.* 9, e1003840.

(19) Miura, K., Herrera, R., Diouf, A., Zhou, H., Mu, J., Hu, Z., MacDonald, N. J., Reiter, K., Nguyen, V., Shimp, R. L., Jr., Singh, K., Narum, D. L., Long, C. A., and Miller, L. H. (2013) Overcoming allelic specificity by immunization with five allelic forms of *Plasmodium falciparum* apical membrane antigen 1. *Infect. Immun.* 81, 1491–1501.

(20) Dutta, S., Lee, S. Y., Batchelor, A. H., and Lanar, D. E. (2007) Structural basis of antigenic escape of a malaria vaccine candidate. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12488–12493.

(21) Harris, K. S., Adda, C. G., Khore, M., Drew, D. R., Valentini-Gatt, A., Fowkes, F. J., Beeson, J. G., Dutta, S., Anders, R. F., and Foley, M. (2014) Immunodampening to overcome diversity in the malaria vaccine candidate apical membrane antigen 1. *Infect. Immun.* DOI: 10.1128/IAI.02061-14.

(22) Tonkin, M. L., Crawford, J., Lebrun, M. L., and Boulanger, M. J. (2013) *Babesia divergens* and *Neospora caninum* apical membrane antigen 1 structures reveal selectivity and plasticity in apicomplexan parasite host cell invasion. *Protein Sci.* 22, 114–127.

(23) Bai, T., Becker, M., Gupta, A., Strike, P., Murphy, V. J., Anders, R. F., and Batchelor, A. H. (2005) Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12736–12741.

(24) Lim, S. S., Debono, C. O., MacRaid, C. A., Chandrashekar, I. R., Dolezal, O., Anders, R. F., Simpson, J. S., Scanlon, M. J., Devine, S. M., Scammells, P. J., and Norton, R. S. (2013) Development of inhibitors of *Plasmodium falciparum* apical membrane antigen 1 based on fragment screening. *Aust. J. Chem.* 66, 1530–1536.

(25) MacRaid, C. A., Anders, R. F., Foley, M., and Norton, R. S. (2011) Apical membrane antigen 1 as an anti-malarial drug target. *Curr. Top. Med. Chem.* 11, 2039–2047.

(26) Srinivasan, P., Yasgar, A., Luci, D. K., Beatty, W. L., Hu, X., Andersen, J., Narum, D. L., Moch, J. K., Sun, H., Haynes, J. D., Maloney, D. J., Jadhav, A., Simeonov, A., and Miller, L. H. (2013) Disrupting malaria parasite AMA1–RON2 interaction with a small molecule prevents erythrocyte invasion. *Nat. Commun.* 4, 2261.

(27) Coley, A. M., Gupta, A., Murphy, V. J., Bai, T., Kim, H., Foley, M., Anders, R. F., and Batchelor, A. H. (2007) Structure of the malaria

antigen AMA1 in complex with a growth-inhibitory antibody. *PLoS Pathog.* 3, 1308–1319.

(28) Coley, A. M., Parisi, K., Masciantonio, R., Hoeck, J., Casey, J. L., Murphy, V. J., Harris, K. S., Batchelor, A. H., Anders, R. F., and Foley, M. (2006) The most polymorphic residue on *Plasmodium falciparum* apical membrane antigen 1 determines binding of an invasion-inhibitory antibody. *Infect. Immun.* 74, 2628–2636.

(29) Harris, K. S., Casey, J. L., Coley, A. M., Masciantonio, R., Sabo, J. K., Keizer, D. W., Lee, E. F., McMahon, A., Norton, R. S., Anders, R. F., and Foley, M. (2005) Binding hot spot for invasion inhibitory molecules on *Plasmodium falciparum* apical membrane antigen 1. *Infect. Immun.* 73, 6981–6989.

(30) Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., and Li, Q. (2009) Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci.* 18, 936–948.

(31) Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (1998) TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins. *Proc. Natl. Acad. Sci. U.S.A.* 95, 13585–13590.

(32) Riek, R., Pervushin, K., and Wuthrich, K. (2000) TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. *Trends Biochem. Sci.* 25, 462–468.

(33) Chesne-Seck, M. L., Pizarro, J. C., Vulliez-Le Normand, B., Collins, C. R., Blackman, M. J., Faber, B. W., Remarque, E. J., Kocken, C. H., Thomas, A. W., and Bentley, G. A. (2005) Structural comparison of apical membrane antigen 1 orthologues and paralogues in apicomplexan parasites. *Mol. Biochem. Parasitol.* 144, 55–67.

(34) Gunasekera, A. M., Wickramarachchi, T., Neasey, D. E., Ganguli, I., Perera, L., Premaratne, P. H., Hartl, D., Handunnetti, S. M., Udagama-Randeniya, P. V., and Wirth, D. F. (2007) Genetic diversity and selection at the *Plasmodium vivax* apical membrane antigen-1 (PvAMA-1) locus in a Sri Lankan population. *Mol. Biol. Evol.* 24, 939–947.

(35) Remarque, E. J., Faber, B. W., Kocken, C. H., and Thomas, A. W. (2008) Apical membrane antigen 1: A malaria vaccine candidate in review. *Trends Parasitol.* 24, 74–84.

(36) Kabsch, W. (2010) Xds. *Acta Crystallogr. D* 66, 125–132.

(37) Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. D* 69, 1204–1214.

(38) Collaborative Computational Project, No. 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* 50, 760–763.

(39) Brunger, A. T. (1993) Assessment of phase accuracy by cross validation: The free R value. Methods and applications. *Acta Crystallogr. D* 49, 24–36.

(40) McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* 61, 458–464.

(41) Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* 66, 213–221.

(42) Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* 68, 352–367.

(43) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D* 60, 2126–2132.

(44) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797.

(45) Androulakis, S., Schmidberger, J., Bate, M. A., DeGori, R., Beitz, A., Keong, C., Cameron, B., McGowan, S., Porter, C. J., Harrison, A., Hunter, J., Martin, J. L., Kobe, B., Dobson, R. C., Parker, M. W., Whisstock, J. C., Gray, J., Treloar, A., Groenewegen, D., Dickson, N.,

and Buckle, A. M. (2008) Federated repositories of X-ray diffraction images. *Acta Crystallogr. D* 64, 810–814.

(46) Tronrud, D. E. (1996) Knowledge-based B-factor restraints for the refinement of proteins. *J. Appl. Crystallogr.* 29, 100–104.

(47) Parthasarathy, S., and Murthy, M. R. N. (1997) Analysis of temperature factor distribution in high-resolution protein structures. *Protein Sci.* 6, 2561–2567.

(48) *Molecular Operating Environment (MOE)* (2012) CCGI, Montreal, QC.

(49) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935.

(50) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781–1802.

(51) Case, D., Darden, T., Cheatham, T., III, Simmerling, C., Wang, J., Duke, R., Luo, R., Walker, R., Zhang, W., and Merz, K. (2012) *AMBER 12*, University of California, San Francisco.

(52) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* 65, 712–725.

(53) Wickstrom, L., Okur, A., and Simmerling, C. (2009) Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. *Biophys. J.* 97, 853–856.

(54) Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* 103, 8577–8593.

(55) Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4, 435–447.

(56) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics* 14, 33–38.

(57) Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10037–10041.

(58) Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005) The FoldX web server: An online force field. *Nucleic Acids Res.* 33, W382–W388.

(59) Pizarro, J. C., Vulliez-Le Normand, B., Chesne-Seck, M. L., Collins, C. R., Withers-Martinez, C., Hackett, F., Blackman, M. J., Faber, B. W., Remarque, E. J., Kocken, C. H., Thomas, A. W., and Bentley, G. A. (2005) Crystal structure of the malaria vaccine candidate apical membrane antigen 1. *Science* 308, 408–411.

(60) Ge, X., MacRaild, C. A., Devine, S. M., Debono, C. O., Wang, G., Scammells, P. J., Scanlon, M. J., Anders, R. F., Foley, M., and Norton, R. S. (2014) Ligand-induced conformational change of *Plasmodium falciparum* AMA1 detected using ¹⁹F NMR. *J. Med. Chem.* 57, 6419–6427.

(61) Guerois, R., Nielsen, J. E., and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *J. Mol. Biol.* 320, 369–387.

(62) Schymkowitz, J. W. H., Rousseau, F., Martins, I. C., Ferkinghoff-Borg, J., Stricher, F., and Serrano, L. (2005) Prediction of water and metal binding sites and their affinities by using the Fold-X force field. *Proc. Natl. Acad. Sci. U.S.A.* 102, 10147–10152.

(63) Hehl, A. B., Lekutis, C., Grigg, M. E., Bradley, P. J., Dubremetz, J. F., Ortega-Barria, E., and Boothroyd, J. C. (2000) *Toxoplasma gondii* homologue of *Plasmodium* apical membrane antigen 1 is involved in invasion of host cells. *Infect. Immun.* 68, 7078–7086.

(64) Lamarque, M., Besteiro, S., Papoin, J., Roques, M., Vulliez-Le Normand, B., Morlon-Guyot, J., Dubremetz, J. F., Fauquenoy, S., Tomavo, S., Faber, B. W., Kocken, C. H., Thomas, A. W., Boulanger, M. J., Bentley, G. A., and Lebrun, M. (2011) The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites. *PLoS Pathog.* 7, e1001276.

(65) Takala, S. L., Coulibaly, D., Thera, M. A., Batchelor, A. H., Cummings, M. P., Escalante, A. A., Ouattara, A., Traore, K., Niangaly, A., Djimde, A. A., Doumbo, O. K., and Plowe, C. V. (2009) Extreme polymorphism in a vaccine antigen and risk of clinical malaria: Implications for vaccine development. *Sci. Transl. Med.* 1, 2ra5.

(66) Ouattara, A., Takala-Harrison, S., Thera, M. A., Coulibaly, D., Niangaly, A., Saye, R., Tolo, Y., Dutta, S., Heppner, D. G., Soisson, L., Diggs, C. L., Vekemans, J., Cohen, J., Blackwelder, W. C., Dube, T., Laurens, M. B., Doumbo, O. K., and Plowe, C. V. (2013) Molecular basis of allele-specific efficacy of a blood-stage malaria vaccine: Vaccine development implications. *J. Infect. Dis.* 207, 511–519.

Supplementary Information

Structure and Dynamics of Apical Membrane Antigen 1 from *Plasmodium falciparum* FVO

San Sui Lim¹, Wei Yang², Bankala Krishnarjuna¹, Komagal Kannan Sivaraman², Indu R. Chandrashekar¹, Itamar Kass^{2,3}, Christopher A. MacRaild¹, Shane M. Devine¹, Cael O. Debono¹, Robin F. Anders^{1,4}, Martin J. Scanlon^{1,5}, Peter J. Scammells¹, Raymond S. Norton¹, and Sheena McGowan²

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

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

³ Victorian Life Sciences Computation Initiative Life Sciences Computation Centre, Monash University, Clayton, VIC 3800, Australia.

⁴ Department of Biochemistry, La Trobe University, Bundoora, Victoria 3086, Australia.

⁵ Centre of Excellence for Coherent X-Ray Science, Monash University, Parkville, Victoria 3052, Australia.

CONTENTS

Figure S1. ¹H NMR and size-exclusion chromatography of recombinant FVO *Pf*AMA1 DI+II

Figure S2. Sequence alignment of FVO and 3D7 strains of *Pf*AMA1

Figure S3. Sequence variations between 3D7 and FVO *Pf*AMA1 in the D1 and D2 polymorphic clusters

Figure S4. Crystal packing of FVO *Pf*AMA1 Id loop.

Figure S5. Backbone Root-mean-square deviation (RMSD) of different systems studied.

Figure S6. Root-mean-square fluctuation (RMSF) of 3D7 *Pf*AMA1, FVO *Pf*AMA1 and *P. vivax* AMA1.

Figure S7. Crystal packing against the *Pf*AMA1 DII loop.

Figure S8. ¹H-¹⁵N HSQC spectra of FVO and 3D7 *Pf*AMA1 domain I + II.

Figure S9. ^1H - ^{13}C projection of HNCO spectra of FVO and 3D7 *Pf*AMA1 domain I + II.

Figure S10. R1-FVO *Pf*AMA1 binding affinity determined by surface plasmon resonance.

Table S1. X-ray data collection and refinement statistics

Table S2. Average values of normalized B-factors for Ia – f and DII loops

Table S3. Crystal packing in AMA1 crystal structures

Table S4. The FoldX calculated energy differences between the 3D7 *Pf*AMA1-R1 structure and the 5 mutate models

Table S5. Polar interactions involved in stabilizing the Id loops of 3D7 and FVO *Pf*AMA1

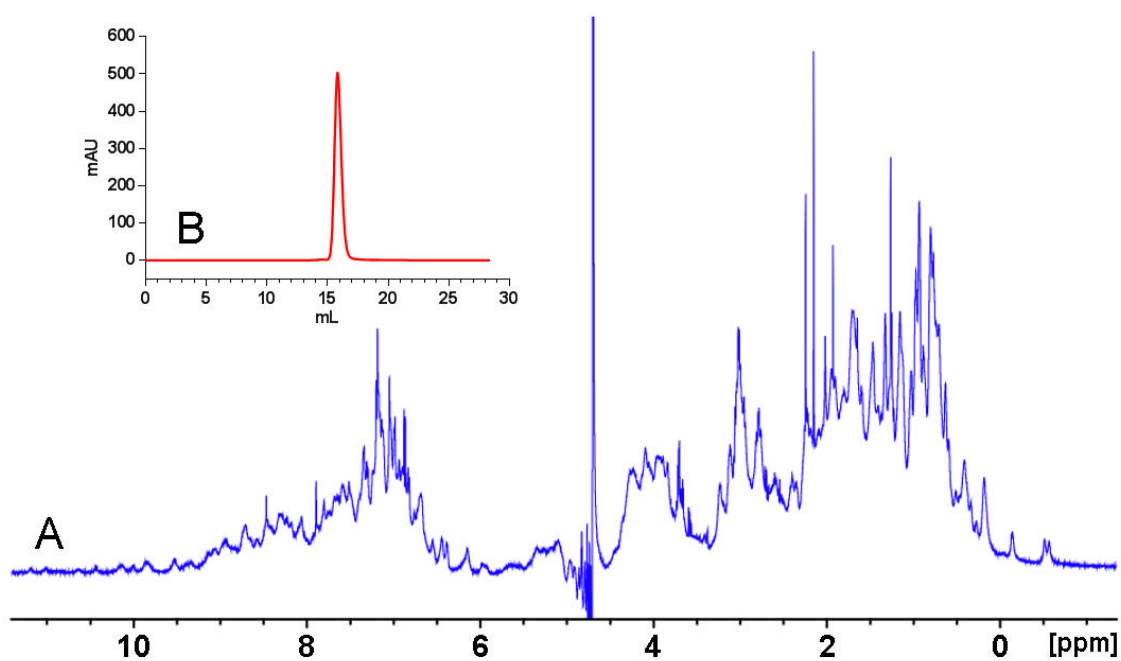


Figure S1. Characterisation of recombinant FVO *PfAMA1* DI +II. (A) ^1H NMR spectrum acquired at 35 °C, pH 7.4 to evaluate the protein folding. (B) Size-exclusion chromatography confirmed that the protein is monomeric.

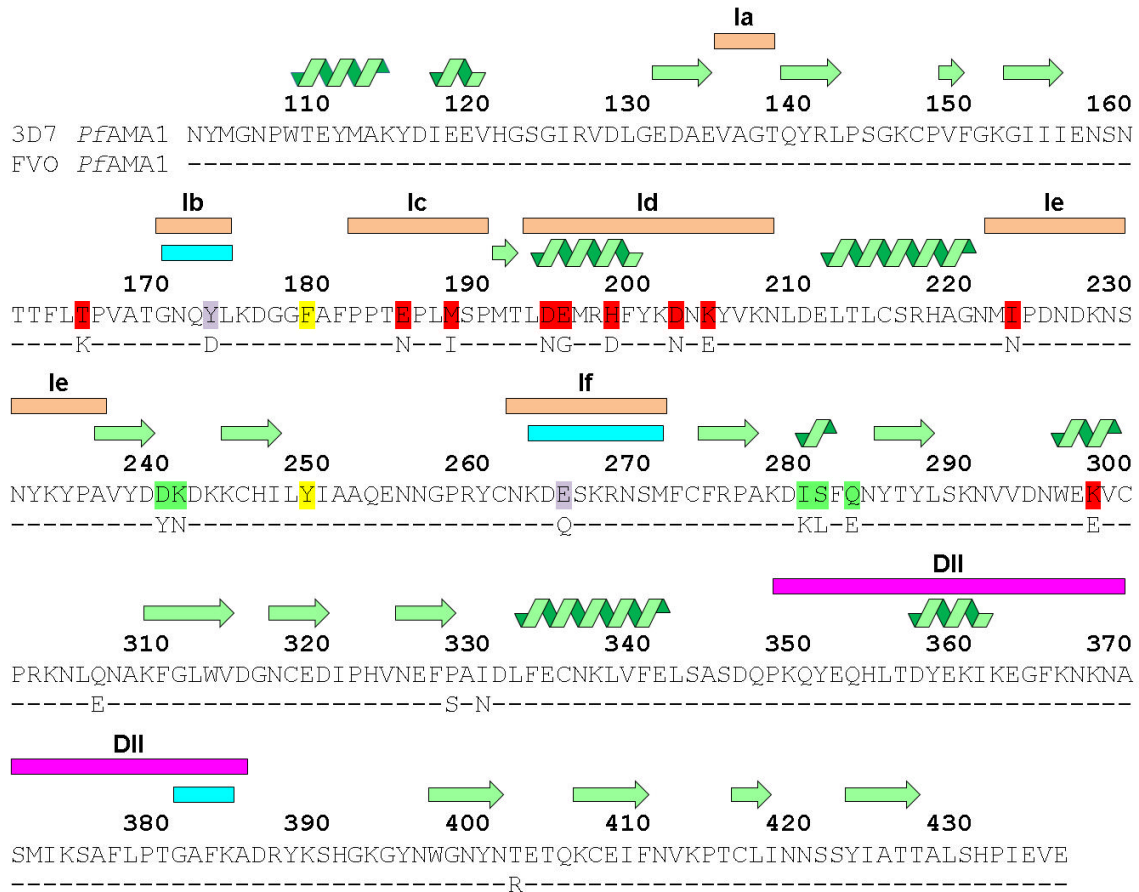


Figure S2. Amino acid sequence of 3D7 *PfAMA1* (top) with sequence polymorphisms in FVO shown. Secondary structure elements are shown above, with arrows depicting β -sheets, coils depicting α -helices and orange rectangles depicting loop regions. The disordered regions are coloured cyan. Phe183 and Tyr251 are highlighted in yellow. Polymorphic residues in domain I are divided into three clusters: C1 (residues 187, 190, 196, 197, 200, 204, 206, 225, boxed in red), C2 (residues 242, 243, 282, 283 and 285, boxed in green) and C3 (residues 175 and 267, boxed in lilac).

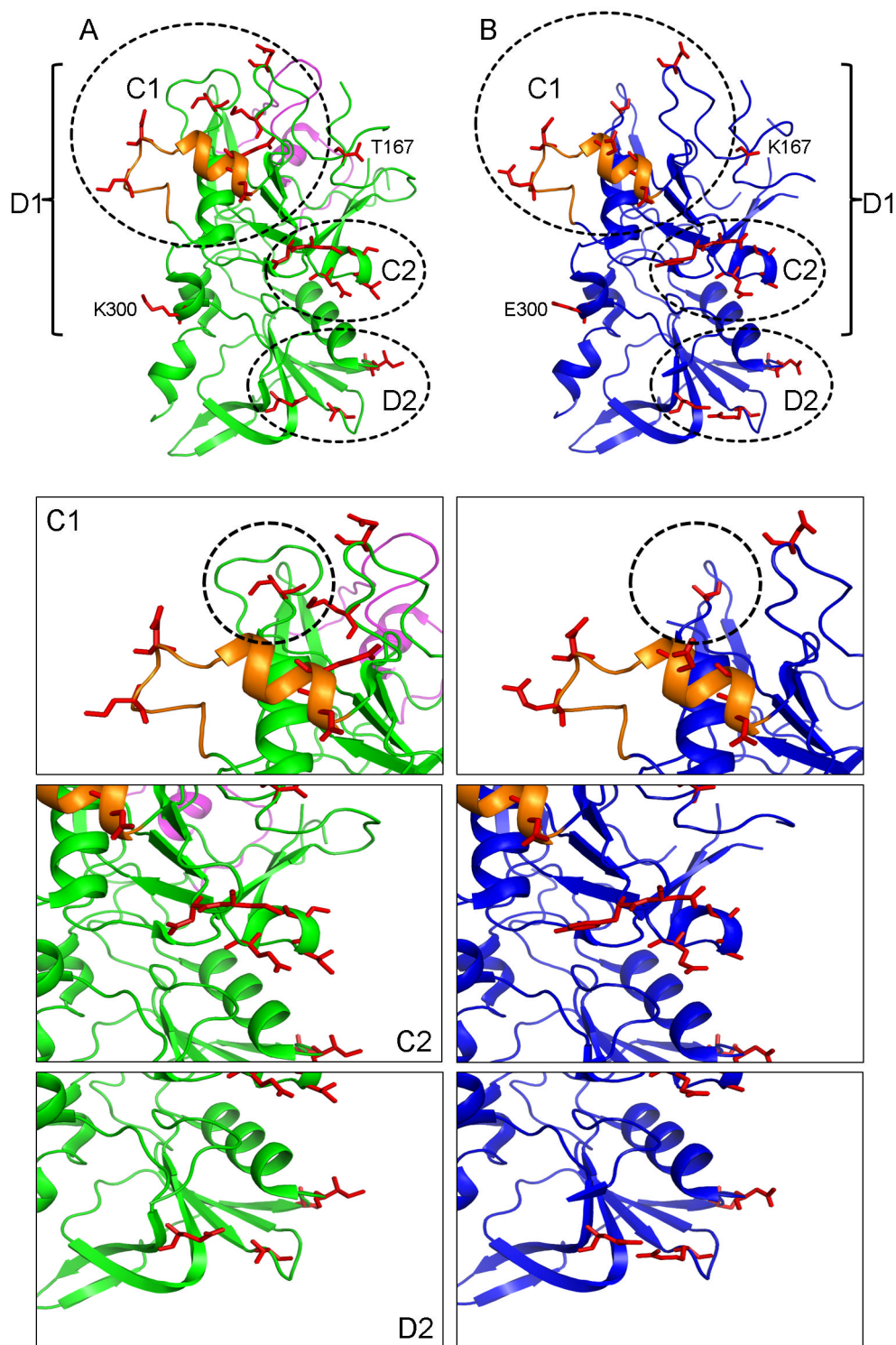


Figure S3. Sequence variations between 3D7 (green) and FVO (blue) *PfAMA1* in the D1 and D2 polymorphic clusters. The C1 and C2 clusters as well as residues 167 and 300 of D1 are shown but the C3 cluster was not seen in either structure. The Id loop (C1L) is coloured orange in both structures. Sequence polymorphisms are shown as red sticks. The DII loop is coloured magenta. Residues 226 – 232 are shown in dotted circles.

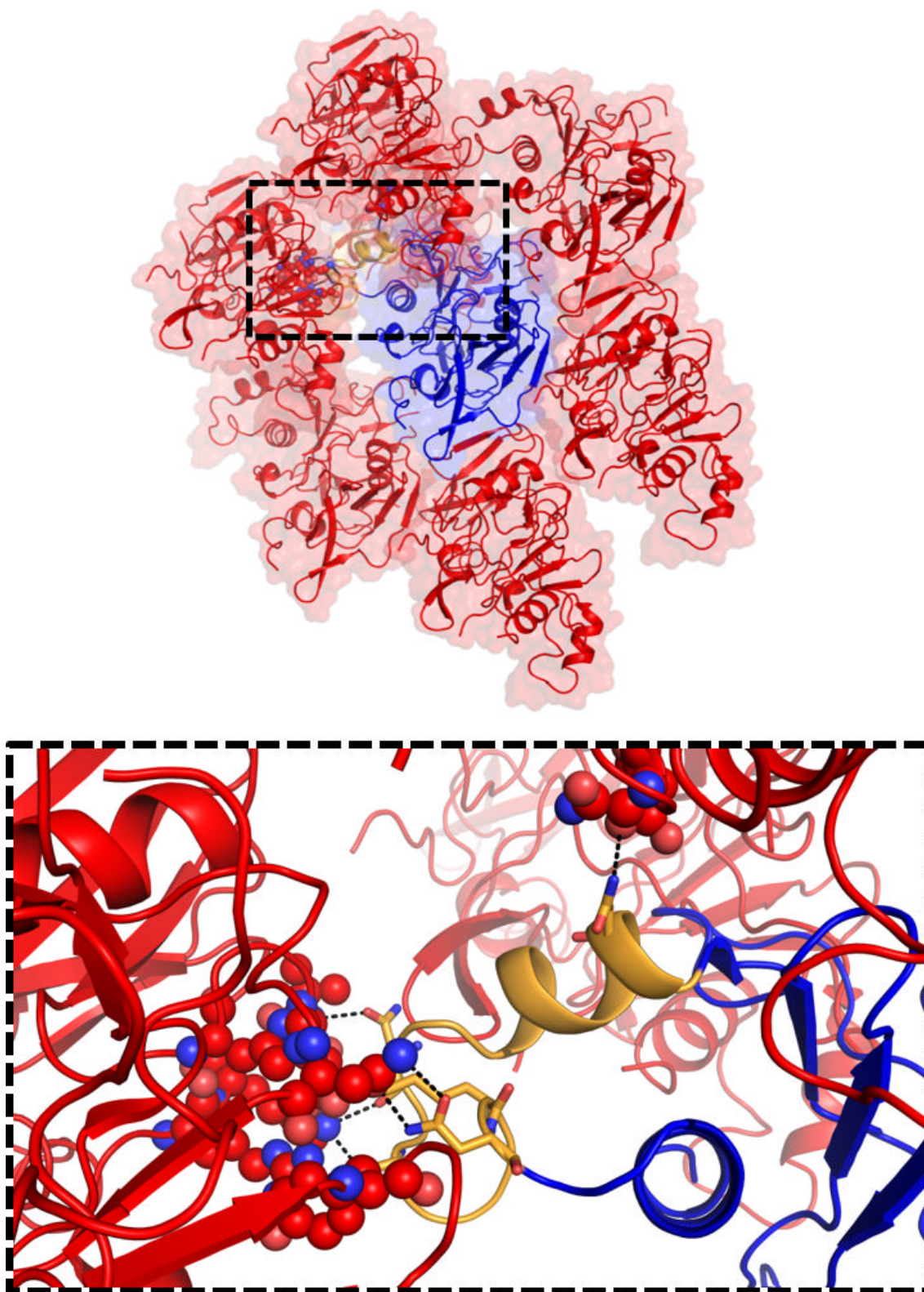


Figure S4. Crystal packing of FVO *PfAMA1* Id loop (orange). Residues that are within 4 Å of the Id loop are shown as spheres.

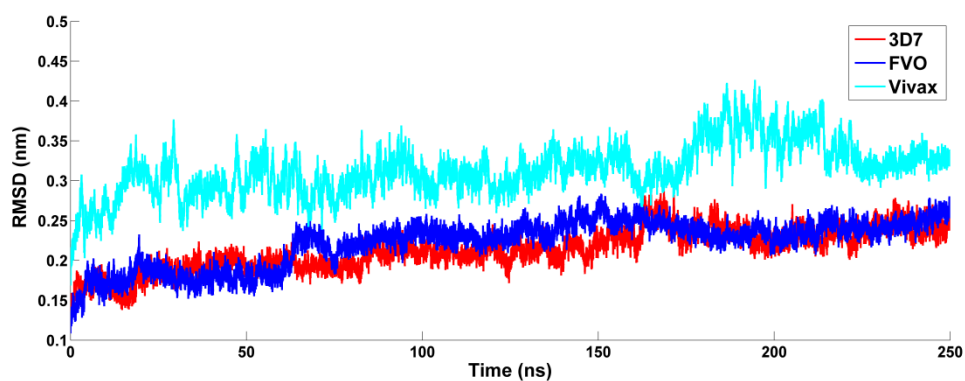


Figure S5. Backbone root-mean-square deviation (RMSD) of different systems studied. RMSD plot as a function of time of backbone atoms for representative trajectories of 3D7 *Pf*AMA1 (red), FVO *Pf*AMA1 (blue) and *P. vivax* AMA1 (cyan). This plot was produced by Matlab 2012b.⁽²⁾

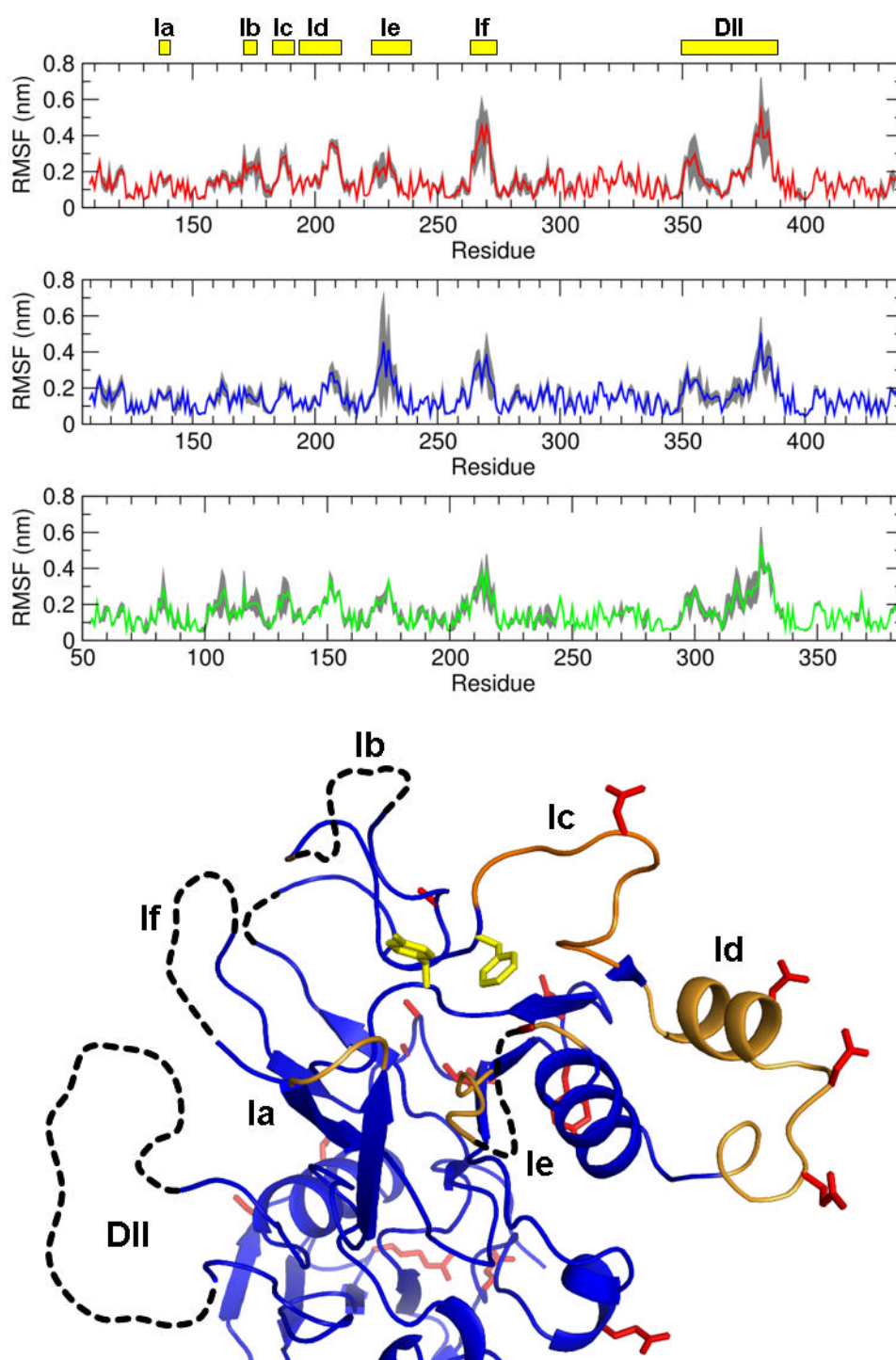


Figure S6. Root-mean-square fluctuation (RMSF) of 3D7 *PfAMA1* (red), FVO *PfAMA1* (blue) and *P. vivax* AMA1 (green). All molecular dynamics (MD) simulations were performed in triplicate. The average values of RMSF – over three repeats of MD simulations – and its corresponding standard deviations are shown as colored lines and grey shades, respectively. Crystal structure of FVO *PfAMA1* with Ia – f and DII loops shown on the bottom. Red sticks represent sequence variations between FVO and 3D7 *PfAMA1*. Yellow sticks correspond to Phe183 and Tyr 251. Dotted lines are disordered regions.

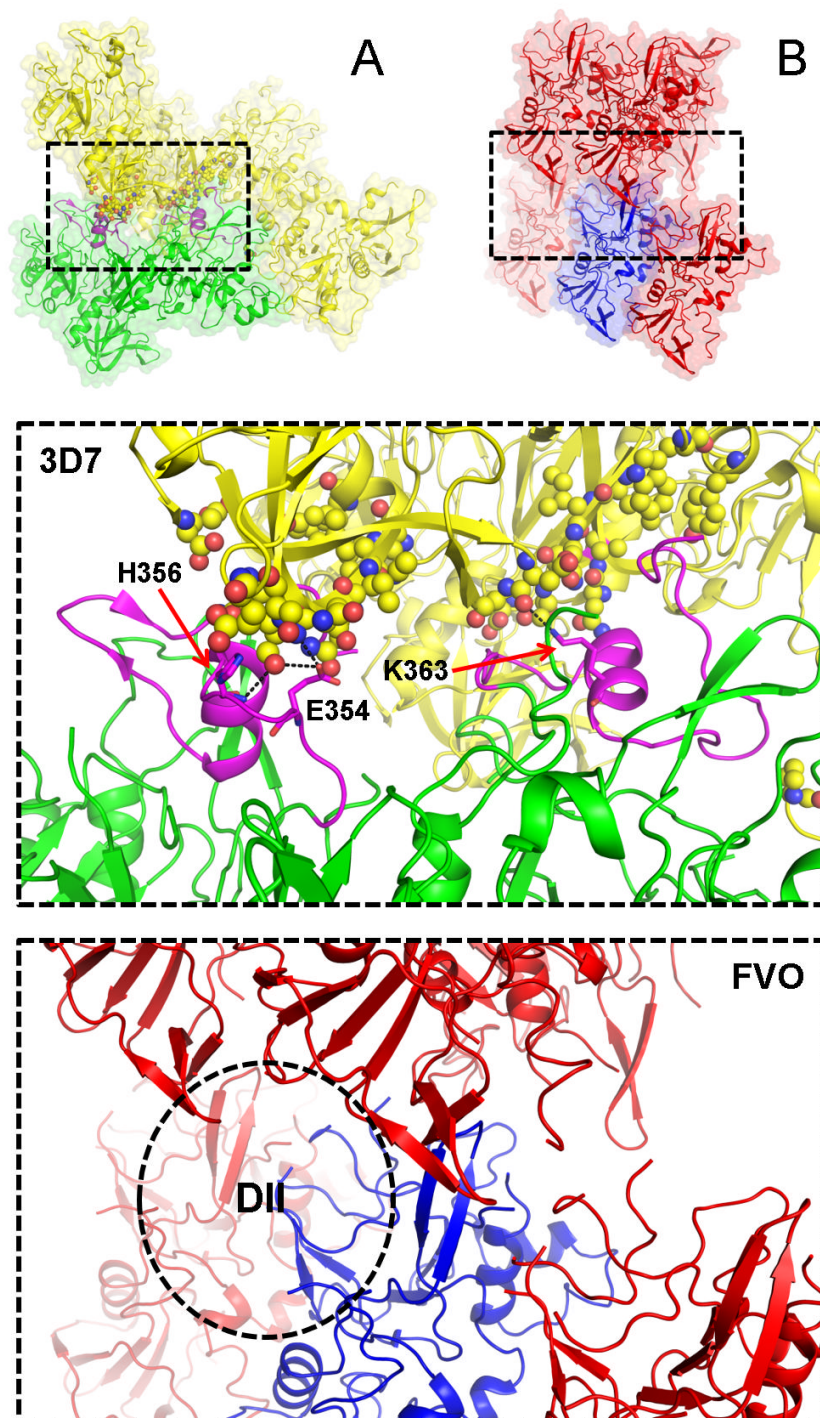


Figure S7. Crystal packing against the *PfAMA1* DII loop. (A) The presence of crystal contacts in the 3D7 *PfAMA1* DII loop (PDB ID: 1Z40). The reference and neighboring proteins are colored green and yellow, respectively. The DII loops are highlighted in magenta and residues that are within 4 Å of the DII loops are shown as spheres. (B) The FVO *PfAMA1* DII loop is disordered in the absence of crystal packing (PDB ID 4R1A); no residues were found within 4 Å of the DII loops. There appears to be sufficient space in the lattice for the DII loop to adopt a closed conformation as in 3D7 *PfAMA1* (PDB ID: 1Z40) if this were its preferred conformation in FVO *PfAMA1*. Blue and red cartoon structures represent the reference and neighbouring AMA1 molecules, respectively. Black dotted lines indicate where the DII loop of FVO *PfAMA1* should be.

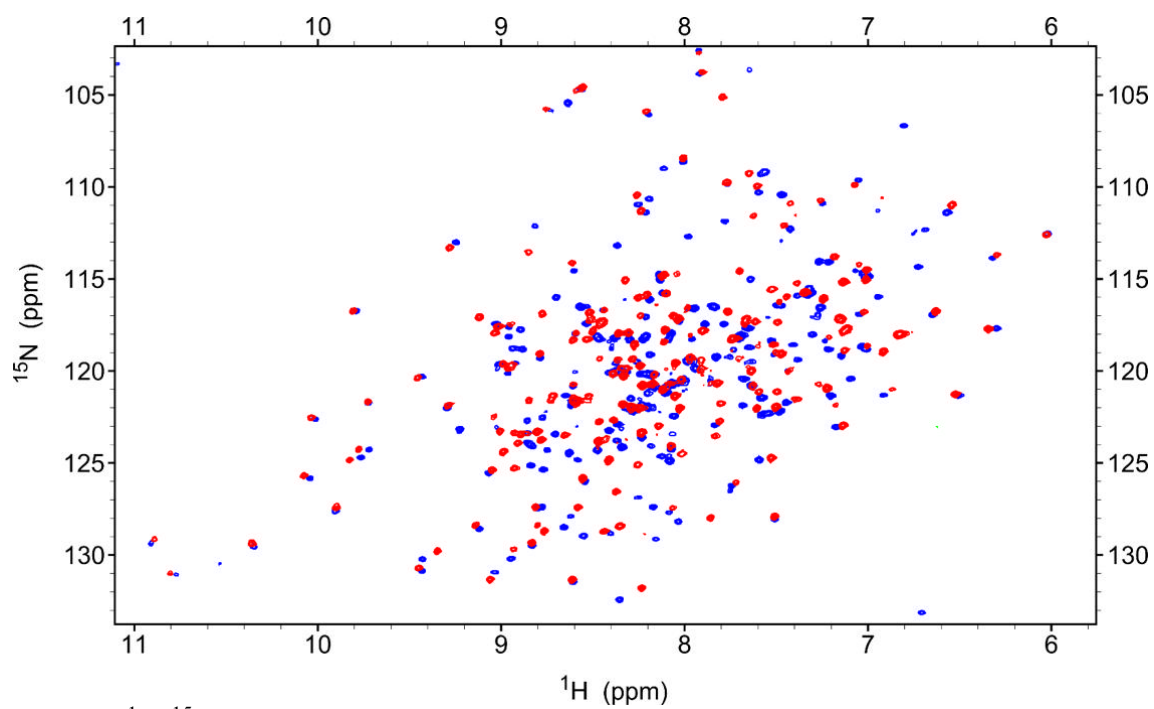


Figure S8. ^1H - ^{15}N HSQC spectra of FVO (blue) and 3D7 (red) *Pf*AMA1 domain I + II.

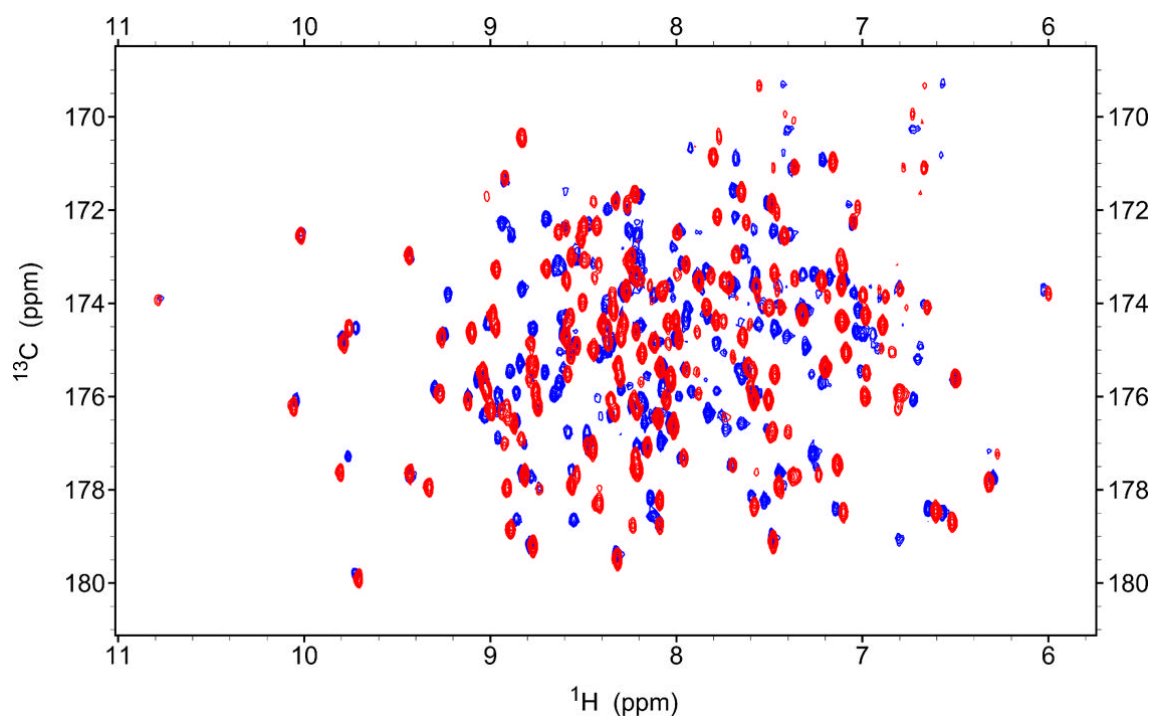


Figure S9. ^1H - ^{13}C projection of HNCO spectra of FVO (blue) and 3D7 (red) *Pf*AMA1 domain I + II.

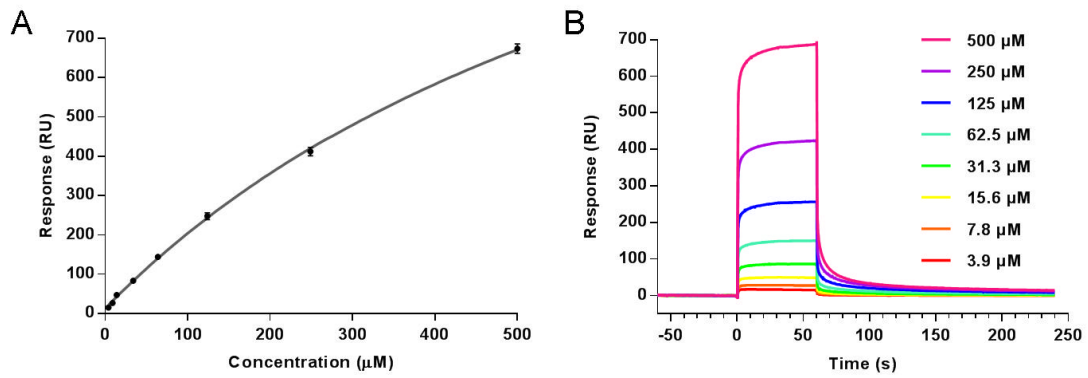


Figure S10. R1-FVO *PfAMA1* binding affinity determined by surface plasmon resonance. (A) Affinity fit for experiments performed in triplicate (B) SPR sensorgram. Although concentration-dependent responses were observed, super-stoichiometric interaction, with the ligand R_{max} (650 RU) exceeding the theoretical R_{max} (440 RU at immobilization level of ~ 8000 RU) was found. The observed over-binding may compromise the accuracy of the K_d obtained from the fit.

Table S1. Data Collection and refinement statistics

Data collection	FVO	3D7 (original)	3D7 (5% MeOH)	3D7 (MQ)
Space Group	$C2_1$	$P3_1$	$C2_1$	$C2_1$
Cell dimensions (Å)	a=109.8 b=37.8 c=71.7, $\alpha = \gamma = 90^\circ$; $\beta = 94^\circ$	a=54.4 b=54.4 c=214.3, $\alpha = \beta = 90^\circ$; $\gamma = 120^\circ$	a=156.5 b=54.5 c=67.9, $\alpha = \gamma = 90^\circ$; $\beta = 91.4^\circ$	a=74.9 b=51.4 c=87.6, $\alpha = \gamma = 90^\circ$; $\beta = 114.3^\circ$
Resolution (Å)	35.72 – 2.00 (2.05 – 2.00)	47.07 – 1.80 (1.84 – 1.80)	41.2 – 2.0 (2.05 – 2.0)	41.07 – 1.60 (1.63 – 1.60)
Total observations	75353	562849	182343	680213
Unique observations	19576	65700	38986	40241
Multiplicity	3.8 (4.0)	8.6 (8.5)	4.7 (4.7)	16.9 (9.1)
Data	97.5 (96.4)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Completeness (%)				
$\langle I/\sigma_I \rangle$	5.1 (1.1)	6.7 (1.4)	9.7 (2.2)	11.7 (1.1)
CC(1/2)	0.987 (0.551)	0.994 (0.487)	0.996 (0.305)	0.997 (0.526)
R_{pim} (%) ^b	10.7 (67.9)	6.6 (80.2)	8.4 (119.0)	5.4 (80.7)
PDB	4R1A	4R19	4R1B	4R1C
Structure refinement				
Molecules in a.u.	1	2	2	1
Non hydrogen atoms				
Protein	2038	4959	4569	2188
Solvent (HOH)	145	452	317	300
R_{free} (%)	25.9	20.3	24.8	17.9
R_{cryst} (%)	20.9	17.2	20.3	20.7

Bond lengths (Å)	0.003	0.010	0.006	0.012
Bond angles (°)	0.65	1.17	0.98	1.30
<i>Ramachandran</i>				
<i>plot</i>	95.6	96.8	96.3	96.7
Favoured (%)	0.0	0.0	0.0	0.0
Outliers (%)				
<i>B factors (Å²)</i>				
Mean protein	30.9	35.0	45.2	29.1
Mean water molecule	36.7	41.8	49.6	39.4
Molprobity Score ^c	1.10 (100 th percentile (N=12522, 2.0 Å ± 0.25 Å)	1.28 (98 th percentile (N=11444, 1.8 Å ± 0.25 Å)	1.53 (96 th percentile (N=12522, 2.0 Å ± 0.25 Å)	1.51 (89 th percentile (N=7200, 1.6 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 .

Table S2. Average values of normalized C α B-factors for α -helical and β -sheet regions, and Ia – f and DII loops. D = disordered region with more than half the residues missing.

AMA1	α -helix	B-sheet	Ia (137 – 140)	Ib (172 – 176)	Ic (184 – 192)	Id (195 – 210)	Ie (224 – 238)	If (264 – 273)	DII (351 – 387)
3D7 (1Z40)	-0.30 \pm 0.54	-0.37 \pm 0.48	-0.64 \pm 0.27	D	0.23 \pm 0.47	0.34 \pm 0.94	-0.58 \pm 0.30	D	0.75 \pm 1.69
FVO (4R1A)	-0.36 \pm 0.48	-0.16 \pm 0.85	3.11 \pm 0.58	D	1.12 \pm 1.06	-0.12 \pm 0.25	1.13 \pm 1.26	D	D
<i>P. vivax</i> (1W81)	-0.40 \pm 0.42	-0.33 \pm 0.76	2.82 \pm 0.21	3.27 \pm 0.17	1.62 \pm 0.71	0.61 \pm 0.32	0.72 \pm 0.72	D	D
3D7 (4R19)	-0.27 \pm 0.48	-0.51 \pm 0.42	-0.19 \pm 0.23	2.13 \pm 0.20	0.34 \pm 0.28	0.85 \pm 0.74	-0.18 \pm 0.38	D	0.78 \pm 1.46
3D7 (4R1B)	-0.34 \pm 0.68	-0.37 \pm 0.58	1.16 \pm 0.37	D	1.56 \pm 1.05	1.18 \pm 0.08	1.05 \pm 1.22	D	D
3D7 (4R1C)	-0.06 \pm 0.96	-0.43 \pm 0.66	1.57 \pm 0.41	D	1.81 \pm 0.68	1.39 \pm 0.43	1.14 \pm 0.78	D	D

Table S3. Crystal packing in AMA1 crystal structures. The number of residues from the neighbouring proteins in close vicinity ($< 4.0 \text{ \AA}$) of the Ia – f and DII loops of the respective AMA1 crystal structures. The numbers of polar contacts with symmetry-related molecules are indicated as bold numbers inside the brackets. Where no crystal packing was present, this is indicated as a long dash.

AMA1	Ia (137 – 140)	Ib (172 – 176)	Ic (184 – 192)	Id (195 – 210)	Ie (224 – 238)	If (264 – 273)	DII (351 – 387)
FVO (4R1A)	–	–	3 (1)	9 (7)	–	–	–
3D7 (1Z40)	–	–	5 (1)	7 (1)	5 (2)	–	9 (1)
<i>P. vivax</i> (1W81)	–	1 (0)	1 (0)	3 (0)	3 (2)	–	–
3D7 (4R19)	–	–	6 (0)	4 (1)	5 (2)	–	12 (2)
3D7 (4R1B)	1 (0)	–	2 (0)	5 (3)	2 (1)	–	–
3D7 (4R1C)	–	–	–	3 (0)	1 (0)	–	–

Table S4. FoldX calculated free energy differences between the 3D7-R1 structure and the 5 mutant models (Positive numbers indicate lower stability).

Model	Total energy change (kcal/mol)
Y175D	6
M190I	1
D204N	0
I225N	3
FVO model	10

Table S5. Polar interactions involved in stabilizing the α -helical and loop components of FVO and 3D7 *Pf*AMA1 1d loop (residues 195 to 210). **Red text** indicates residues that are different between the two strains.

FVO α-helix	Interacting residues	Distance (Å)	3D7 α-helix	Interacting residues	Distance (Å)
L195 [N]	K245 [O]	2.9	L195 [N]	K245 [O]	2.9
G 197 [N]	T194 [O γ 1]	3.0	E 197 [N]	T194 [O]	3.1
			E 197 [N]	T194 [O γ 1]	3.2
			E 197 [O ϵ 1]	T194 [O]	3.0
M198 [N]	T194 [O]	3.0	M198 [N]	T194 [O]	2.9
R199 [NH1]	L211 [O]	2.8	R199 [NH1]	L211 [O]	2.7
R199 [NH1]	K209 [O]	2.7	R199 [NH1]	K209 [O]	2.8
R199 [NH2]	L211 [O]	2.9	R199 [NH2]	L211 [O]	3.0
D 200 [O]	K203 [N]	3.1	H 200 [O]	K203 [N]	3.2
D 200 [O δ 1]	K203 [N ζ]	2.9			
Y202 [O]	N205 [N]	2.8	Y202 [O]	N205 [N]	2.9
Y202 [OH]	H220 [N δ 1]	2.7	Y202 [OH]	H220 [N δ 1]	2.6
Y202 [OH]	N223 [O δ 1]	3.3	Y202 [OH]	N223 [O δ 1]	3.0
FVO loop	Interacting residues	Distance (Å)	3D7 loop	Interacting residues	Distance (Å)
K203[N]	D200 [O]	3.1	K203[N]	H200 [O]	3.2
K 203 [N ζ]	D200 [O δ 1]	2.9			
N205[N]	Y202 [O]	2.8	N205[N]	Y202 [O]	2.9
V208 [O]	L211 [N]	3.2	V208 [O]	L211 [N]	2.8
K209 [O]	R199 [N η 1]	2.7	K209 [O]	R199 [N η 1]	3.1

REFERENCE

1. Dutta, S., Lee, S. Y., Batchelor, A. H., and Lanar, D. E. (2007) Structural basis of antigenic escape of a malaria vaccine candidate, *Proc. Natl. Acad. Sci. USA* 104, 12488-12493.

CHAPTER 4

Conformational dynamics and antigenicity in the
disordered malaria antigen merozoite surface
protein 2

4.1 Declaration of Thesis Chapter 4

Declaration by candidate

In the case of **Chapter 4**, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Cloned, expressed and purified ^{15}N , and ^{15}N , ^{13}C labelled 3D7 MSP2 samples and collaborated in preparation of manuscript	20

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Christopher A. MacRaid	Manuscript preparation	
Milan Zachrdla	Performed experiments	
Dean Andrew	Performed experiments	
Jiří Nováček,	Intellectual input	
Lukáš Židek	Performed experiments	
Vladimír Sklenář	Intellectual input	
Jack S. Richards	Intellectual input	
James G. Beeson	Intellectual input	
Robin F. Anders	Intellectual input	
Raymond S. Norton	Intellectual input, manuscript preparation	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

**Candidate's
Signature**

	Date 30/03/2016
---	---------------------------

**Main
Supervisor's
Signature**

	Date 30/03/2016
---	---------------------------

4.2 Introduction

Previous chapters (Chapter 2 and Chapter 3) revealed the structure of *Pf*AMA1 and interactions with small molecules using NMR spectroscopy. This was to identify small molecules that bind to the hydrophobic cleft on *Pf*AMA1 with the intent of developing more potent *Pf*AMA1 inhibitors as novel antimalarials. Although drugs are important for the treatment of malaria, a vaccine is required for the prevention of the disease. In view of this, the following chapters are focused on a malaria vaccine candidate, merozoite surface protein 2 (MSP2). MSP2 is an intrinsically disordered, polymorphic merozoite surface antigen, the function of which is unknown. However, antibodies to MSP2 have been shown to provide protection in both natural infections as well as in previous clinical trials of MSP2, albeit in a strain-specific manner. The native conformation of MSP2 is not known, although NMR suggests that MSP2 make interactions with the parasite membrane thorough its *N*- and *C*-terminal regions, while the central variable region is exposed.

Therefore, in this chapter, a detailed analysis of the conformational dynamics of 3D7 and FC27 MSP2 alleles was carried out using NMR spectroscopy. Together with experiments with antisera to recombinant MSP2, it was found that a strong correlation exists between the conformational dynamics and antigenicity of MSP2. A new expression protocol for 3D7 MSP2 was established using an *E. coli* expression system. Conformationally restricted regions were shown to be more immunogenic than the more flexible regions, suggesting that MSP2 immunogenicity can be enhanced by increasing its conformational order. As the results of this chapter have been published in the journal *PLOS ONE*, they are presented in the format of a published article in the immediate section below.

RESEARCH ARTICLE

Conformational Dynamics and Antigenicity in the Disordered Malaria Antigen Merozoite Surface Protein 2

Christopher A. MacRaid^{1*}, Milan Zachrdla^{2,3}, Dean Andrew⁴, Bankala Krishnarjuna¹, Jiří Nováček^{2,3}, Lukáš Židek^{2,3}, Vladimír Sklenář^{2,3}, Jack S. Richards⁴, James G. Beeson⁴, Robin F. Anders⁵, Raymond S. Norton¹

1 Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, 3052, Australia, **2** NCBR, Faculty of Science, Masaryk University, Kamenice 5, 62500, Brno, Czech Republic, **3** CEITEC, Masaryk University, Kamenice 5, 62500, Brno, Czech Republic, **4** Centre for Biomedical Research, Burnet Institute, Melbourne, Victoria, 3004, Australia, **5** Department of Biochemistry, La Trobe University, Victoria, 3086, Australia

* These authors contributed equally to this work.



OPEN ACCESS

Citation: MacRaid CA, Zachrdla M, Andrew D, Krishnarjuna B, Nováček J, Židek L, et al. (2015) Conformational Dynamics and Antigenicity in the Disordered Malaria Antigen Merozoite Surface Protein 2. PLoS ONE 10(3): e0119899. doi:10.1371/journal.pone.0119899

Academic Editor: Takafumi Tsuboi, Ehime University, JAPAN

Received: September 18, 2014

Accepted: January 16, 2015

Published: March 5, 2015

Copyright: © 2015 MacRaid et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Assigned backbone chemical shifts for 3D7 MSP2 have been submitted to the BMRB (accession number 25431). All other relevant data are within the paper and its Supporting Information files.

Funding: This work was supported in part by an Australian National Health and Medical Research Council (NHMRC) project grant (1042520) and by the project "CEITEC - Central European Institute of Technology" (grant CZ.1.05/1.1.00/02.0068) from European Regional Development Fund. JSR, JGB, and RSN acknowledge fellowship support from the

Abstract

Merozoite surface protein 2 (MSP2) of *Plasmodium falciparum* is an abundant, intrinsically disordered protein that is GPI-anchored to the surface of the invasive blood stage of the malaria parasite. Recombinant MSP2 has been trialled as a component of a malaria vaccine, and is one of several disordered proteins that are candidates for inclusion in vaccines for malaria and other diseases. Nonetheless, little is known about the implications of protein disorder for the development of an effective antibody response. We have therefore undertaken a detailed analysis of the conformational dynamics of the two allelic forms of MSP2 (3D7 and FC27) using NMR spectroscopy. Chemical shifts and NMR relaxation data indicate that conformational and dynamic properties of the N- and C-terminal conserved regions in the two forms of MSP2 are essentially identical, but significant variation exists between and within the central variable regions. We observe a strong relationship between the conformational dynamics and the antigenicity of MSP2, as assessed with antisera to recombinant MSP2. Regions of increased conformational order in MSP2, including those in the conserved regions, are more strongly antigenic, while the most flexible regions are minimally antigenic. This suggests that modifications that increase conformational order may offer a means to tune the antigenicity of MSP2 and other disordered antigens, with implications for vaccine design.

Introduction

Recent decades have seen an increasing recognition that many proteins naturally lack a defined folded state, and that their function depends instead on conformational disorder [1,2]. Such proteins are termed intrinsically unstructured or disordered proteins, and are found across all

NHMRC. The Burnet Institute is supported by funding from the NHMRC Infrastructure for Independent Research Institutes Support Scheme, and Victoria State Government Operational Infrastructure Support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: James Beeson is a PLOS ONE Editorial Board member. This does not alter the authors' adherence to PLOS ONE Editorial policies and criteria.

of biology. In particular, intrinsically disordered proteins are abundant in a range of pathogenic organisms. The proteomes of some viruses are predicted to be almost entirely disordered [3], and several parasite species also have an unusually high proportion of disordered proteins [4]. Nonetheless, the implications of protein disorder for immune recognition by B cells and antibodies have received remarkably little attention [5]. On the one hand, it has been suggested that intrinsically disordered proteins generally elicit weak immune responses or are even completely non-immunogenic [6]. It has been observed that functionally important sites on protein antigens are highly flexible, or are surrounded by flexible loops. This flexibility is proposed in some instances to serve as a means of immune evasion [7]. In sharp contrast to this view, however, it has been suggested that disordered antigens are in some contexts immunodominant [8], but that they fail to contribute to an effective immune response. Thus, they are believed to function for some pathogens as a smoke screen, diverting the immune system from targets with greater protective potential [9]. Nonetheless, numerous B-cell epitopes have been characterised in disordered proteins, and many of these appear to contribute to functional immune responses and therefore represent potential vaccine candidates [5,10–17]. For example, the protective effects of RTS,S, the most advanced malaria vaccine in clinical development, appear to be mediated by antibodies to the disordered repeats of the circumsporozoite protein [15,18].

In order to better understand the effects of conformational disorder on the immune response, and to contribute to the development of a malaria vaccine, we have investigated merozoite surface protein 2 (MSP2). MSP2 is an abundant component of the surface coat of the *Plasmodium falciparum* merozoite, the form of the parasite that invades red blood cells during the blood-stage of infection, which is responsible for symptomatic and severe malaria. Although the specific function of MSP2 has not been defined, it appears to play an essential role in blood-stage replication; it is retained on the merozoite surface during invasion and then degraded soon after invasion is complete [19]. An extensive body of evidence implicates MSP2 as a potential target of protective immunity against *P. falciparum* infection [20–26]. Antibodies to MSP2 have been associated with protection from malaria in prospective longitudinal studies [27–29] and MSP2 antibodies promote opsonic phagocytosis of merozoites and antibody-dependent cellular inhibition of blood-stage replication [26,30,31].

MSP2 is highly polymorphic, with conserved N- and C-terminal domains flanking a central variable region, which contains tandemly arrayed repetitive sequences [32,33]. All MSP2 alleles have been categorized into two families typified by the 3D7 and FC27 alleles, respectively, because of differences in the repeats and flanking variable sequences (Fig. 1) [32,34,35]. Indeed, the sequence variability within each allelic family is limited to the repeat regions and to a few localised regions of heterogeneity within the regions flanking the repeats (green and pink in Fig. 1).

MSP2 is a candidate for inclusion in a malaria vaccine [36], and the 3D7 allele of MSP2 was a component of a subunit vaccine that significantly reduced parasite densities in a clinical trial in Papua New Guinea [25]. This vaccine showed protective efficacy against infections with parasites expressing the vaccine-like 3D7-type MSP2 sequence, indicating that vaccine efficacy was mediated by strain-specific responses to MSP2 [23]. Efforts to elicit protective antibodies against the conserved regions of MSP2 are complicated by the observation that anti-MSP2 antibodies induced by infection with *P. falciparum* are largely directed against epitopes in the central variable region of the molecule [37,38], and that many conserved-region epitopes are cryptic on the parasite surface [10]. As such, the generation of a broadly effective MSP2-based vaccine is likely to require fine control of the specificity of the induced immune response. In this context, we have undertaken a detailed study of the conformational dynamics of MSP2, with the goal of establishing the extent to which these properties might contribute to the

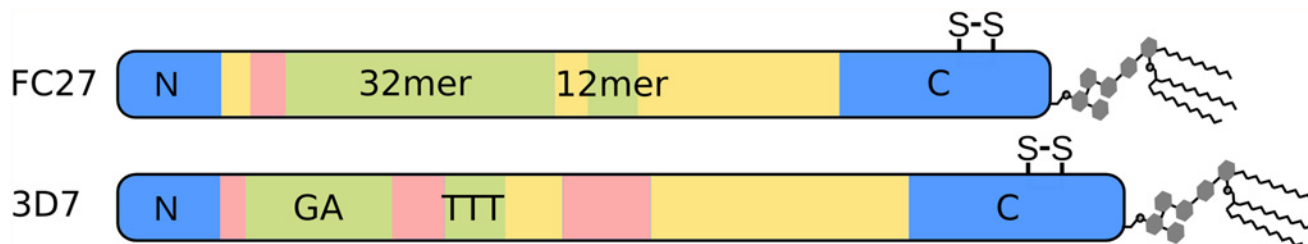


Fig 1. Schematic depiction of the primary structure of the two allelic families of MSP2. The conserved N- and C-terminal regions of MSP2 are in blue, while the allele-specific central region is composed of polymorphic repeats (green) and non-repetitive sequences (pink) as well as dimorphic regions (yellow) that differ between the allelic families but are conserved within them. The position of the conserved disulfide bond in the C-terminal regions is indicated.

doi:10.1371/journal.pone.0119899.g001

observed patterns of antigenicity and immunogenicity against MSP2, and the extent to which they might be exploited to fine-tune the specificity of the antibody response against MSP2.

Methods

Materials

Untagged full-length FC27 MSP2 was expressed and purified using a strategy specific for recombinantly expressed disordered proteins, as described previously [39]. A synthetic gene encoding 3D7 MSP2, codon optimised for expression in *Escherichia coli* (Genescript), was cloned into pET32a (Novagen) using KpnI and NcoI. The resulting construct contains an N-terminal thioredoxin (Trx) and His₆-tag for affinity purification. Bacterial cell pellets were lysed by heating, as for FC27 MSP2 [39]. The expressed fusion protein was isolated on a His-TrapFF affinity column (GE Healthcare), eluted with imidazole and cleaved with 1% (w/w) TEV protease. The released Trx-tag and any uncleaved fusion protein were subsequently removed by a second passage through the His-trap column. Final purification of 3D7 MSP2 was by HPLC, using a C18 column (0.9 x 25 cm, Zorbax) and a linear acetonitrile gradient in 0.1% TFA. Isotopically enriched 3D7 and FC27 MSP2 for NMR studies was prepared by growing expression cultures in M9 minimal medium, with 1 g/L ¹⁵N ammonium chloride and/or 4 g/L ¹³C glucose as the sole nitrogen and carbon sources, respectively. The final recombinant 3D7 MSP2 has an N-terminal Gly derived from the TEV cleavage site whereas recombinant FC27 MSP2 has an N-terminal Met derived from the start codon.

NMR spectroscopy

NMR samples contained 0.4 mM 3D7 or FC27 MSP2 in 50 mM sodium acetate, pH 4.5, with 7% ²H₂O included for the spectrometer lock. All of the data used for resonance assignments were acquired on a 700 MHz Bruker Avance III spectrometer equipped with the ¹H/¹³C/¹⁵N TXO cryogenic probehead with z-axis gradients at 25°C. The HNCQ spectrum was acquired with spectral widths set to 9800 (aq) x 2500 (¹⁵N) x 2000 (¹³C') Hz, and with maximal evolution times of 80 ms (¹³C') and 80 ms (¹⁵N) in the indirectly detected dimensions. The inter-scan delay was set to 1.1 s, and 4 transients per free induction decay (FID) were cumulated. The overall number of 2048 complex points was acquired in the acquisition dimension, whereas 600 hypercomplex points were randomly distributed over the indirectly-detected dimensions. The experiment was acquired in 3.5 h, which represents 1.9% of the time needed for a conventional experiment with similar settings. The 5D HN(CA)CONH experiment was acquired with the spectral widths set to 9800 (aq) x 2500 (¹⁵N) x 2000 (¹³C') x 2800 (¹⁵N) x 8000 (¹H) Hz [40]. The maximal acquisition times were adjusted to 15 ms for the ¹H indirectly-detected dimension, to 27 ms and 40 ms for ¹⁵N dimensions, and to 30 ms for the ¹³C' dimension.

The experiment was acquired with 2048 complex points in the acquisition dimension and 1750 hypercomplex points were randomly distributed over the indirectly detected dimensions. The inter-scan delay was set to 1.25 s and 4 transients per FID were collected. The experimental time of 46 h represents 0.0036% of the time needed for a similar experiment using conventional settings. The 5D HabCabCONH experiment was acquired with spectral widths set to 9800 (aq) x 2500 (^{15}N) x 2000 (^{13}C) x 10000 ($^{13}\text{C}^{\text{aliph}}$) x 5000 ($^1\text{H}^{\text{aliph}}$) [40]. The maximal evolution times were set to 12 ms for $^1\text{H}^{\text{aliph}}$, 6.5 ms for $^{13}\text{C}^{\text{aliph}}$, 30 ms for ^{13}C , and 22 ms for ^{15}N indirect dimensions. The total number of 1536 complex points was measured in the directly-detected dimension, and 1750 hypercomplex points were randomly distributed in the indirectly-detected dimensions. The experiment was acquired with 4 transients per collected FID and an interscan delay of 1.25 s. The overall experimental time of 46 h represents 0.008% of the time needed for acquisition of the conventional experiment providing similar resolution.

NMR relaxation experiments were performed on a 600 MHz Bruker Avance III NMR spectrometer equipped with a QCI-P cryogenic probehead with z-axis gradients at 25°C. Temperature was calibrated according to the chemical shift differences of pure methanol peaks. Spectral widths were set to 8370 (aq) x 1428 (^{15}N) Hz. The overall number of 2048 complex points was acquired in the acquisition dimension and 400 complex points were acquired in the indirect dimension for auto-relaxation rates R_1 , R_2 , cross-correlated relaxation rates Γ_x , Γ_z and steady state ^{15}N - ^1H nuclear Overhauser effect (NOE) [41]. Standard experiments were used for the measurement of R_1 (relaxation delays 11.2, 56, 134.4, 235.2, 380.8, 560, 896*, 1344, 1848, and 2352 ms) and R_2 (relaxation delays 0, 14.4, 28.8*, 43.2, 57.6, 72*, 86.4, 115.2, and 144 ms) [42]. Asterisks denote spectra recorded twice in order to estimate experimental error. Experiments based on symmetrical reconversion were performed for determination of transverse cross-correlated relaxation rates Γ_x (relaxation delays 30, 50, and 70 ms) and longitudinal cross-correlated relaxation rates Γ_z (relaxation delays 100, 150, 200, and 250 ms) [43,44].

Antigenicity

Antigenicity across the MSP2 sequence was determined using sera from mice and rabbits immunised with full-length recombinant 3D7 or FC27 MSP2 (Genebank JN248383 and JN248384). Both proteins were expressed in *E. coli* with C-terminal His₆ tags and purified by metal-chelating, anion-exchange and reverse-phase chromatography [31]. Animals were immunised with the recombinant MSP2 formulated in Montanide ISA720. Mice (C57Bl/6) were immunised with 10 µg subcutaneously and rabbits were immunised with 100 µg intramuscularly on two occasions with a four-week interval between immunisations. Serum samples used in antigenic analyses were obtained from blood samples collected two weeks after the second immunization. Immunisations were approved by the La Trobe University Animal Ethics Committee and were conducted in accord with the policies of the National Health and Medical Research Council, Australia. Reactivity to a panel of 13-residue biotinylated peptides covering the sequence of both antigens with an 8-residue overlap, was measured by ELISA, as described previously [10,19]. The panel contains one copy of the first three peptides common to both 3D7 and FC27 MSP2, but because the central variable regions of 3D7 and FC27 MSP2 are different lengths, the two peptide sets (3D7 and FC27) extended through the conserved C-terminal region to give two sets of peptides covering the same sequence but out of frame with respect to each other, as described previously [10]. Sixteen sera (four per condition) were analysed in triplicate, with all sera tested at 1:1000 dilution and secondary antibody diluted 1:2000. Responses from unimmunised animals were also measured, and a background signal three standard deviations greater than the mean of these responses was subtracted from all results. Agreement between animals was assessed using Pearson's correlation coefficients for

pairwise comparisons of mean ELISA results for each animal. Permutation tests were used to estimate two-tailed p-values. Within each condition, each residue in MSP2 was accorded the average response of all peptides in which that residue is represented and the resulting antigenicity profiles were normalised.

Results

Backbone resonance assignments for 3D7 MSP2

The assignment of observed spectral frequencies (chemical shifts) in an NMR spectrum to specific atoms in the protein is a prerequisite for detailed structural analysis by NMR, allowing measured spectral parameters to be ascribed to specific structural features. We have previously determined near-complete backbone assignments for FC27 MSP2, but expression yields for 3D7 MSP2 were insufficient to permit the detailed analysis of that allelic form [39]. Here, we employ a new expression system, based on a thioredoxin fusion strategy, that yields ~ 10 mg 3D7 MSP2 per litre of culture medium. An assignment strategy tailored to repetitive disordered proteins and exploiting two 5D experiments, HN(CA)CONH and HabCabCONH, was employed to assign the resonance frequencies of 3D7 MSP2 [45]. All the non-proline residues were successfully assigned, although residues 37–58, within the GGSA repeats, show degenerate backbone chemical shifts, as do residues 77, 78, and 84–87 within the TTT repeats.

The amide chemical shifts of 3D7 MSP2 show minimal dispersion (Fig. 2), and backbone shifts are close to those expected for a disordered protein (Fig. 3). This demonstrates that 3D7 MSP2, like FC27, is extensively disordered, consistent with our previous analyses [39,46]. Comparison of the backbone chemical shifts of the FC27 and 3D7 forms of MSP2 reveals almost perfect correspondence between the shifts of the conserved N and C-terminal regions, indicating that the conformational propensities of these regions are identical in the two allelic forms (Figs. 2 and 3). In particular, slightly elevated α secondary chemical shifts are seen in the N-terminal region of both MSP2 forms, indicating a weak preference for helical conformation in this region (Fig. 3) [39,47,48]. $\text{C}\beta$ chemical shifts of the two Cys residues confirm the presence of the single disulfide in MSP2 [49].

Conformational dynamics probed by ^{15}N relaxation

NMR relaxation rates are sensitive to fast conformational dynamics [51], and as such are valuable probes of the extent of disorder in unstructured proteins [52]. We have measured relaxation rates of the backbone amides of MSP2 to determine conformational dynamics at ps-ns timescales and at single-residue resolution (Fig. 4). Values of the spectral density function $J(\omega)$ at zero frequency and at the ^{15}N and ^1H Larmor frequencies were calculated from ^{15}N relaxation rates (Fig. 5) [53]. These values represent the direct link between the experimental data and the conformational dynamics of the protein, with a larger value of $J(\omega)$ indicating a larger contribution to relaxation from dynamic processes with frequency ω . In order to identify residues exhibiting dynamics at μs -ms timescales, the $J(0)$ and $J(\omega_{\text{N}})$ values were calculated from both auto-correlated (R_1 , R_2 , steady-state $[^1\text{H}]-^{15}\text{N}$ NOE) and cross-correlated (Γ_z , Γ_x) relaxation data [54].

The results of the relaxation measurements are consistent with those expected for a disordered protein; relaxation rates are uniformly low, while the steady-state $[^1\text{H}]-^{15}\text{N}$ NOE is generally negative. These results reflect conformational dynamics dominated by local processes on timescales shorter than 1 ns. In contrast, the relaxation of a conventionally structured protein is dominated by overall rotational diffusion (on a timescale >10 ns, for a protein the size of MSP2), with relatively small contributions from faster local processes. For FC27 MSP2 we observe relaxation properties that are in excellent agreement with those we reported previously

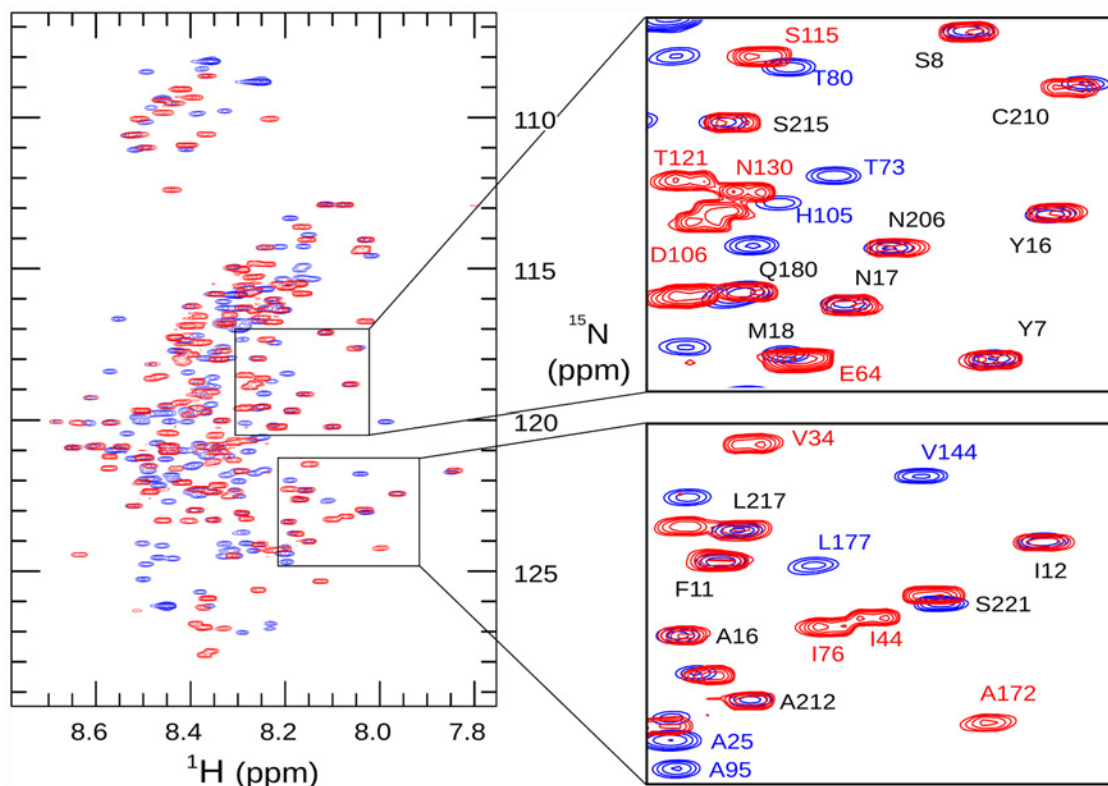


Fig 2. ^1H , ^{15}N heteronuclear single-quantum correlation spectra of 3D7 (blue) and FC27 (red) MSP2. Assigned peaks are labelled in the expanded regions, highlighting the similarity of chemical shift for these residues in the conserved N- and C-terminal regions (labelled in black type, FC27 numbering) of the two allelic forms of MSP2.

doi:10.1371/journal.pone.0119899.g002

under more acidic conditions [39]. Under both conditions, we observe more rapid relaxation and smaller magnitude (and in some cases small positive) ^1H - ^{15}N NOEs, consistent with a degree of conformational constraint, in the following three distinct regions: throughout the conserved N-terminus, in part of the C-terminal region coincident with the single disulfide bond in MSP2, and in part of the FC27-specific dimorphic region, between residues 140 and 150 (Figs. 4 and 5). It should be stressed that, although these regions are more ordered than the rest of MSP2, they do not represent regions of folded regular structure as both their relaxation properties and chemical shifts are indicative of significant residual disorder, well beyond that observed in conventional structured proteins. Rather, the flexibility of these regions is weakly constrained by transient helical structure in the N-terminal region, and by the disulfide in the C-terminal region. For a few residues in the C-terminal conserved region, values of $J(0)$ calculated from auto-correlated relaxation data are larger than those obtained from the cross-correlated relaxation (Fig. 5). This is suggestive of exchange contributions to the measured R_2 relaxation rates for these residues, and may imply the existence of a weakly populated metastable conformational state with a lifetime in the μs -ms range [55]. The repeat regions of FC27 MSP2 show somewhat variable dynamic properties, with elevated values of $J(\omega_{\text{H}})$, indicating more extensive sub-ns dynamics than observed in the rest of the dimorphic and C-terminal regions, but with significant variation in the lower-frequency spectral densities across the 32-residue repeat (Fig. 5).

The variable region of 3D7 MSP2 shows greater diversity in its dynamic properties, as reported by relaxation measurements. The largest region of polymorphism in 3D7 MSP2, the

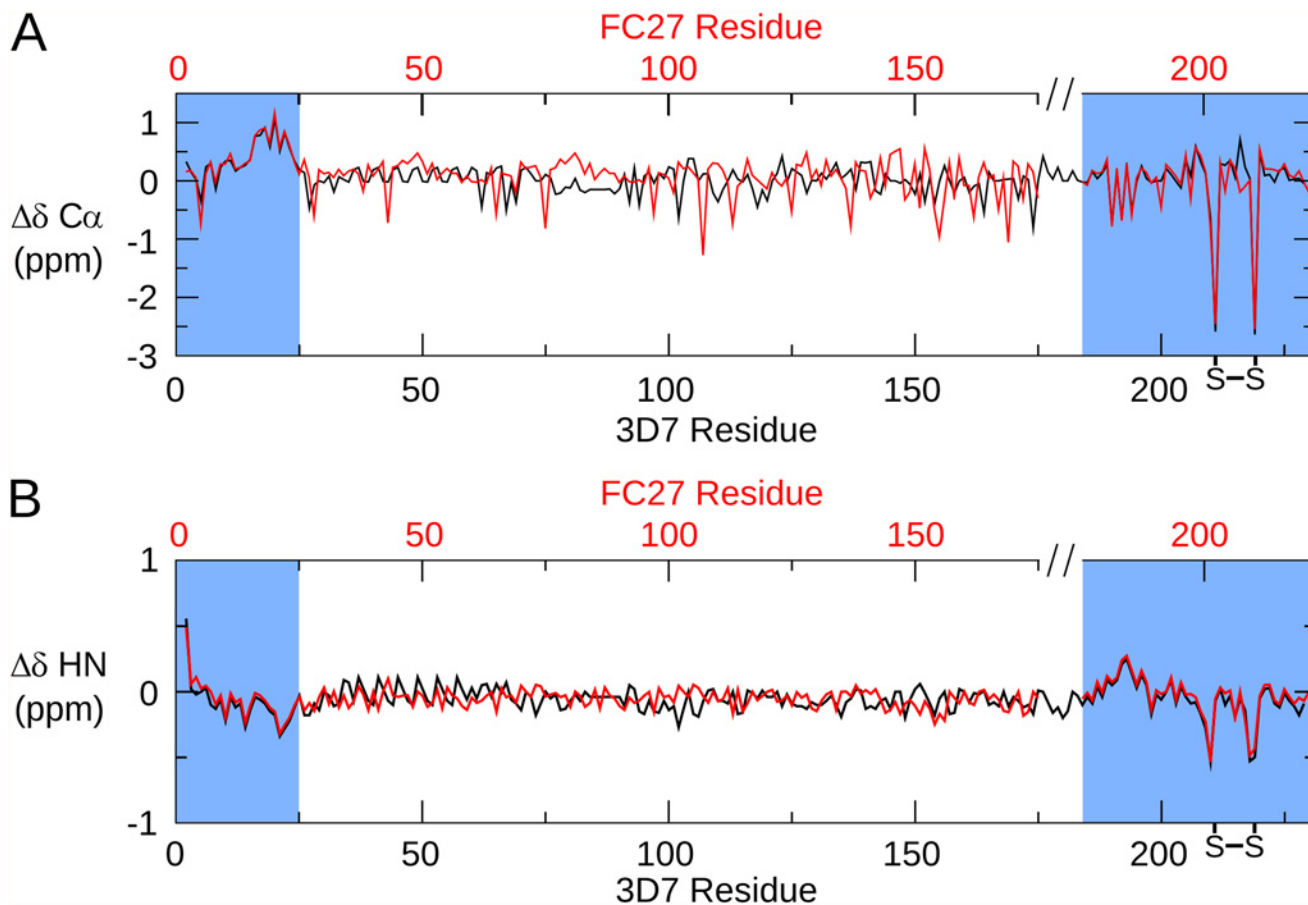


Fig 3. Secondary chemical shifts of MSP2. The difference between the observed C α (A) and HN (B) chemical shifts and those predicted for a disordered protein by the method of Tamiola et al. [50] is plotted for 3D7 (black) and FC27 (red) MSP2. The data for FC27 are plotted on a broken axis (top) in order to correctly align the conserved regions.

doi:10.1371/journal.pone.0119899.g003

GGSA repeats, is exceptionally flexible, with relaxation properties indistinguishable from the extreme termini of the protein (Fig 5; residues 32–63). Presumably this flexibility is a consequence of the uniformly small side chains in this region, and the correspondingly small steric barriers to backbone reorganisation. It is noteworthy that, despite the high levels of polymorphism, this region is consistently rich in small residues, with Gly, Ser and Ala representing over 90% of residues seen in this region across all 3D7 MSP2 alleles characterised. In contrast, the next largest region of polymorphism, residues 103–122 appears relatively ordered, to essentially the same degree as the more ordered region in the FC27 dimorphic domain, residues 140–150. Likewise, the degree of order in the 3D7 dimorphic region (residues 123–180; yellow in Figs 4 and 5) is comparable to the remainder of the FC27 dimorphic region, and the 3D7 TTT repeats are comparable to the FC27 32- and 12-residue repeats.

In contrast to the variable regions, the dynamic properties of the conserved regions of MSP2 are indistinguishable in the two allelic forms at ps-ns timescales, as indicated by identical relaxation rates, with the regions of reduced flexibility within the N- and C-terminal regions being the most ordered regions in both alleles (Fig 3). Together with the perfect correspondence of backbone chemical shifts, this agreement indicates that the ensembles of rapidly interconverting conformational states sampled by these regions are identical in FC27 and 3D7 MSP2. The intervening variable regions exert no perceptible influence on these properties.

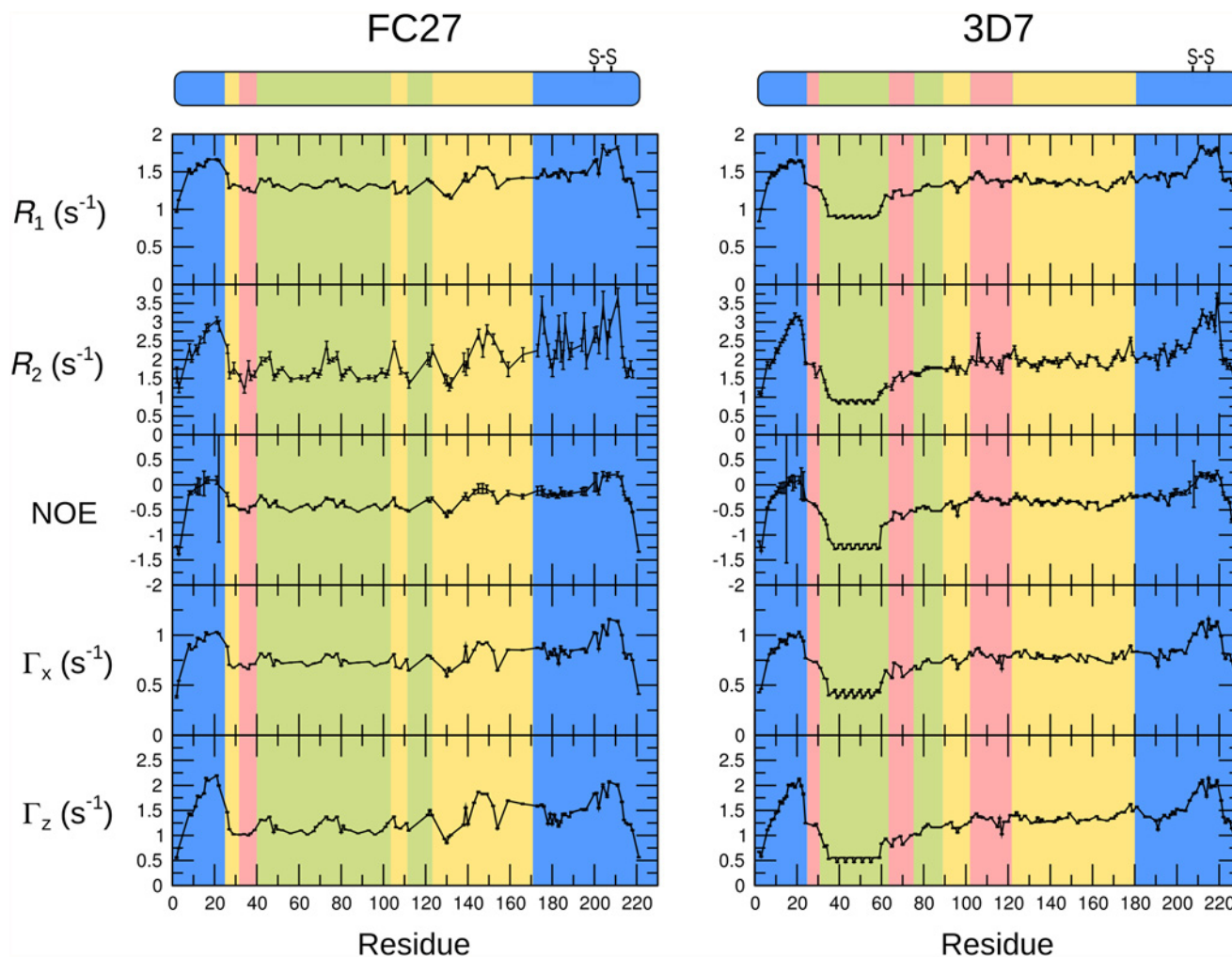


Fig 4. The ^{15}N auto-relaxation rates R_1 , R_2 , cross-correlated relaxation rates Γ_x , Γ_z , and steady-state ^{15}N - ^1H NOE for FC27 (left) and 3D7 (right) MSP2. The sequence regions of MSP2 are colour-coded as in Fig. 1.

doi:10.1371/journal.pone.0119899.g004

More generally, this suggests that the conformational properties of MSP2 are entirely locally determined. This inference is consistent with our observation that each member of a large panel of monoclonal antibodies recognises a simple linear epitope [10].

For 3D7 MSP2, the $J(0)$ values calculated from auto-correlated relaxation data are never significantly larger than those obtained from cross-correlated relaxation (Fig. 5), suggesting that the μs -ms dynamics that was inferred for FC27 MSP2 C-terminal conserved region are absent in this region of 3D7 MSP2. This suggests that the variable region may influence the population or lifetime of meta-stable conformations in the conserved C-terminal domain. Remarkably, this appears to occur without affecting the ps-ns dynamics, or the overall conformational preferences (as reported by chemical shift) of either region.

Antigenicity is correlated with local dynamics

Potential correlations between the conformational dynamics characterised above and the antigenicity of MSP2 have been investigated by examining the patterns of local antigenicity in sera of mice and rabbits immunised with recombinant 3D7 and FC27 MSP2. The reactivity of

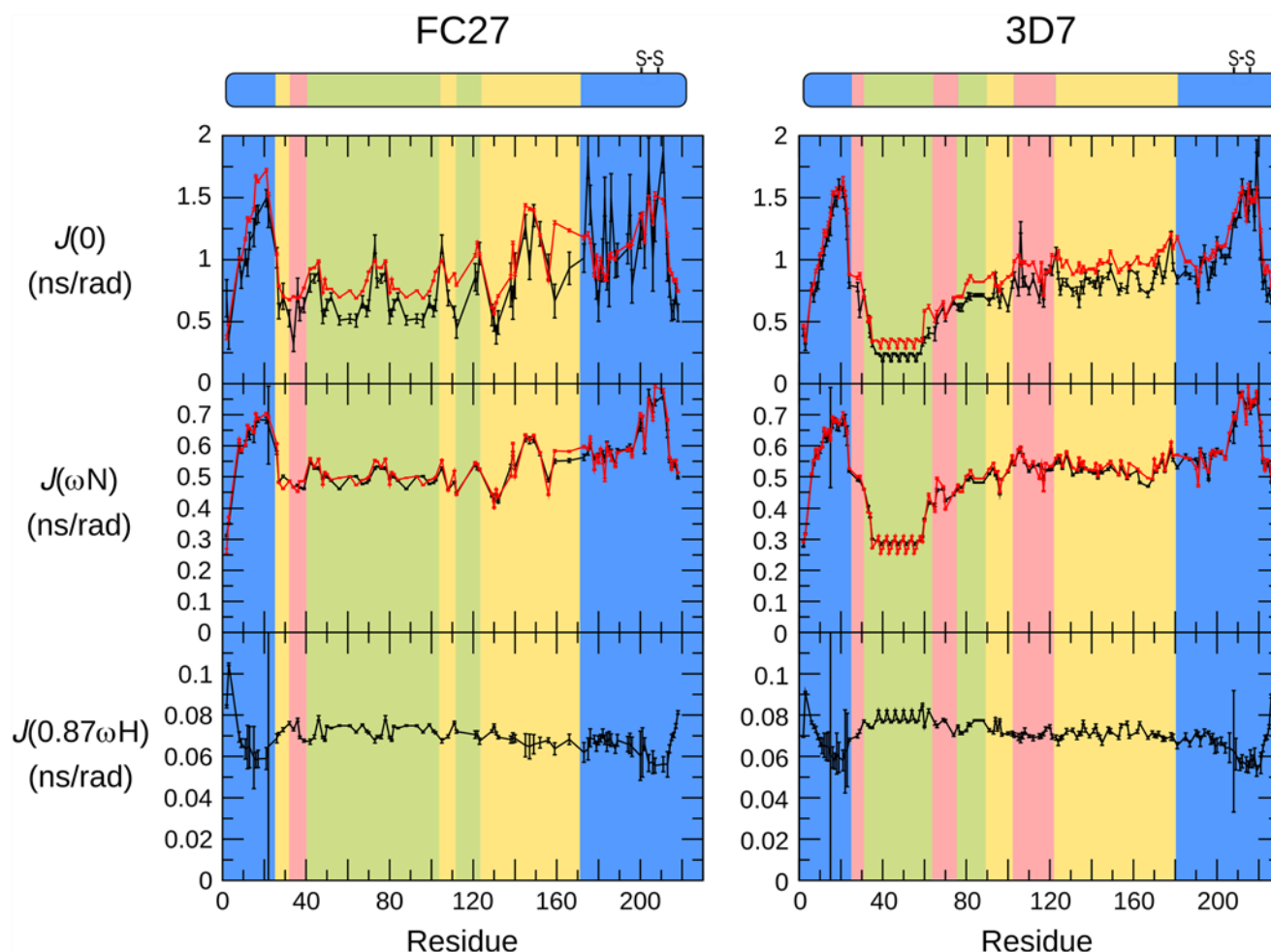


Fig 5. Values of the ^{15}N spectral density functions for FC27 (left) and 3D7 (right) MSP2, as determined from reduced spectral density mapping from the auto-relaxation rates and steady-state ^{15}N - ^1H NOE (black) and cross-correlated relaxation rates (red). The sequence regions of MSP2 are colour-coded as in Fig. 1.

doi:10.1371/journal.pone.0119899.g005

MSP2 antisera to an array of overlapping peptides covering the entire sequences of 3D7 and FC27 MSP2 was measured by ELISA [10,19]. To enable direct comparison between these results and our NMR measurements, which are resolved at the level of individual residues, we adopted a scoring scheme in which each residue was scored according to the average reactivity of each of the peptides in which that residue was represented. There was good agreement across the individual mice immunised against each antigen. Pairwise comparisons of ELISA results from individual mice immunised with 3D7 MSP2 yielded average correlation coefficients of 0.7 ± 0.1 and for mice immunised with FC27 MSP2 the average correlation was 0.6 ± 0.1 (Pearson's r ; $p < 10^{-5}$ for all comparisons). For rabbits, there is substantially greater variation between individuals, with correlation coefficients of 0.3 ± 0.2 for each antigen ($p < 0.05$ for seven of 12 comparisons). Nonetheless there was reasonable qualitative agreement across all antigenicity profiles for each antigen, with most regions identified to be antigenic in mice also antigenic in at least one rabbit, and vice versa (Fig. 6). The following analysis therefore considers a single average profile for each antigen in each species (Fig. 7A).

Several lines of evidence give rise to confidence that these profiles of antigenicity are robust estimations of the intrinsic immunogenicity of MSP2. First, there is excellent agreement

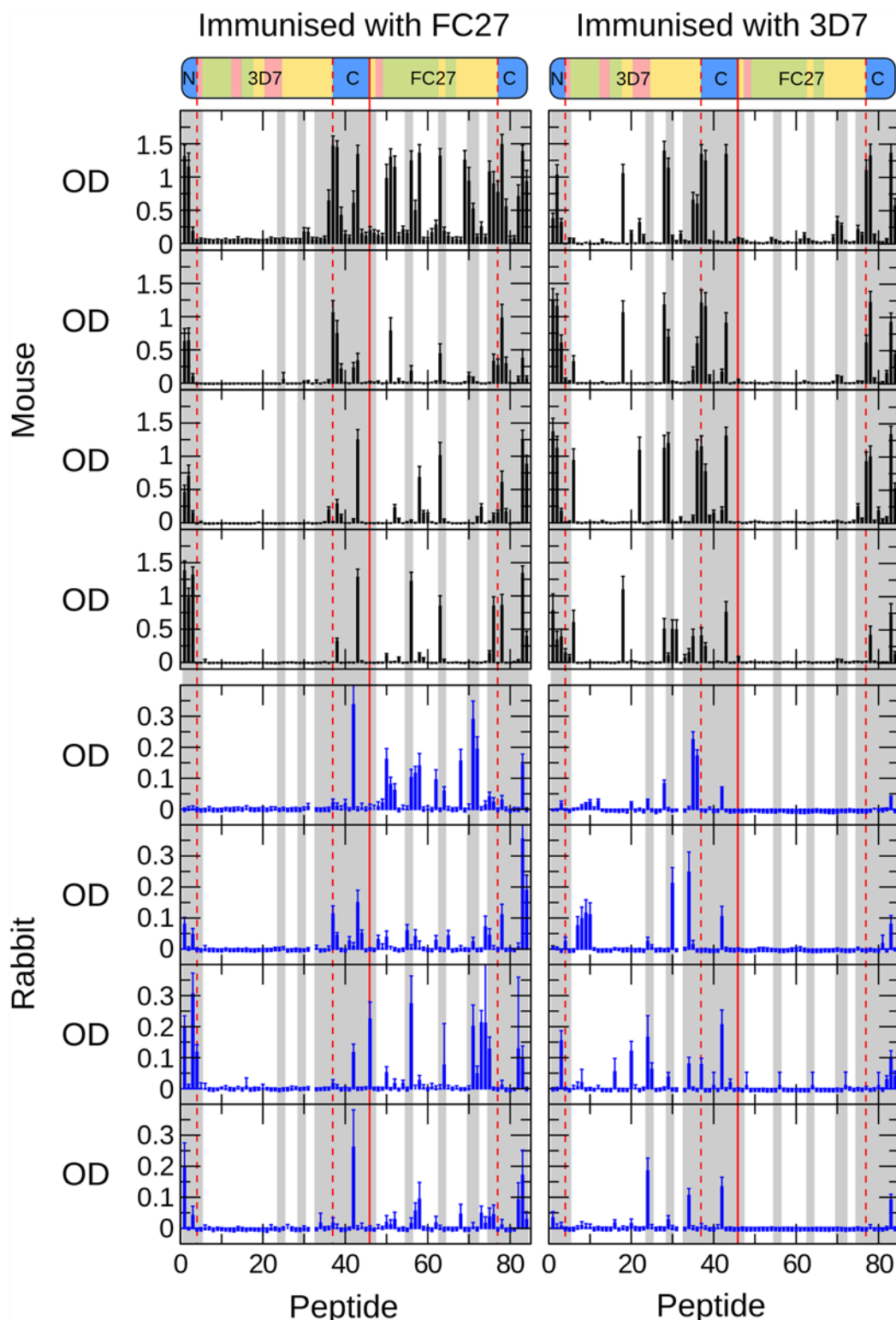


Fig 6. Antigenic profile of MSP2 mapped by ELISA. Four mice (top panels, black bars) and four rabbits (bottom panels, blue bars) were immunised with either FC27 (left panels) or 3D7 (right panels) MSP2. Individual immune sera were tested against a single panel of overlapping peptides covering the sequences of both 3D7 MSP2 (peptides 1–45) and FC27 MSP2 (peptides 46–84), as shown schematically above. The conserved N terminal (peptides 1–3) and C terminal (peptides 37–45 and 77–84) regions are common to both 3D7 and FC27 MSP2 and are delineated with dashed red lines. Peptides showing greater than the median level of conformational restriction are shaded grey (Table 1). Mean optical density from triplicate assays is plotted for each serum, corrected for the response from unimmunised control sera. Error bars are one standard deviation.

doi:10.1371/journal.pone.0119899.g006

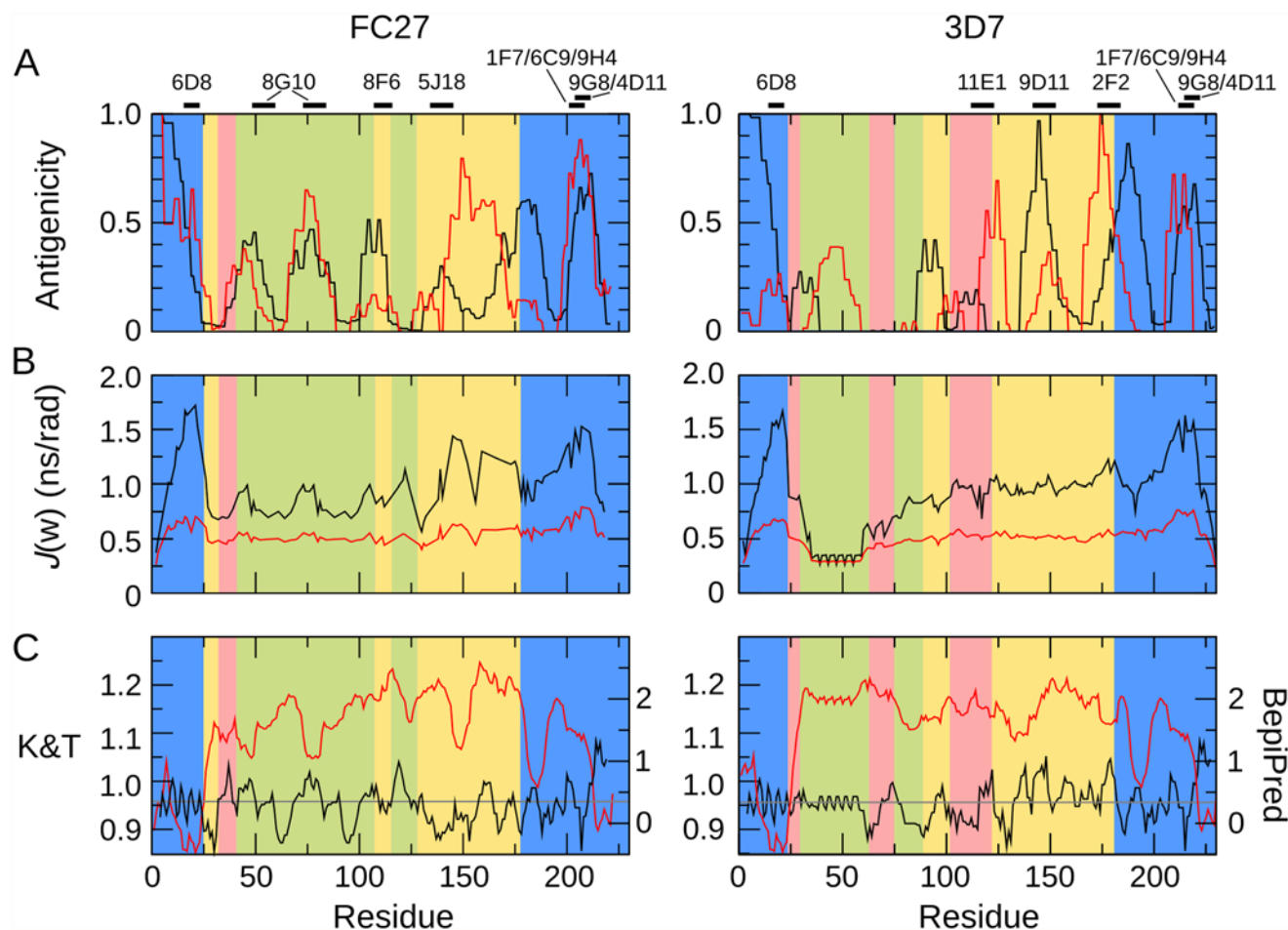


Fig 7. Comparison of experimental patterns of antigenicity, predicted antigenicity, and conformational dynamics, for FC27 (left) and 3D7 (right) MSP2. A. Antigenicity profiles of MSP2 inferred from experimental immunisation of mice (black) and rabbits (red) are plotted against the sequence. Black bars (top) denote the location of epitopes of a panel of monoclonal antibodies to MSP2. B. Conformational flexibility of MSP2 as measured by the spectral density functions derived from the ^{15}N relaxation data. Spectral density functions are plotted at zero frequency (black line) and at the ^{15}N Larmor frequency (red line). C. Antigenicity of MSP2 as predicted using BepiPred [60] (red, right axis) and the method of Kolaskar and Tongaonkar [61] (black, left axis). The threshold for epitope prediction for both methods is denoted by the grey line.

doi:10.1371/journal.pone.0119899.g007

Table 1. Conformationally constrained peptides are more antigenic.

		No. animals responding ^b				No. peptides with > 2 animals responding ^c		
	$J(0)^a$	mice	rabbits	total		mice	rabbits	total
Constrained	42 1.06–1.88	2.2 ± 0.4	1.1 ± 0.2	1.7 ± 0.3		14	4	18
Flexible	42 0.36–1.06	0.50 ± 0.15	0.45 ± 0.10	0.48 ± 0.09		3	0	3

a Range of maximum $J(0)$ values defining each peptide class.

b The number of sera generating a background-corrected response greater than 0.3 OD (mice) or 0.05 OD (rabbits) to individual peptides, averaged (\pm SEM) over each class.

c Number of peptides in each class to which more than two sera respond.

doi:10.1371/journal.pone.0119899.t001

between the current 3D7 profile and those derived previously from experimental immunisations of mice and humans with recombinant 3D7 MSP2 [56,57]. Second, the conserved regions of MSP2 show similar patterns of antigenicity in both the FC27 and 3D7 profiles. Finally, the epitopes of an extensive panel of monoclonal antibodies to MSP2 [10,30,58,59] all coincide with peaks in the antigenicity profile (Fig. 7A).

Strikingly, both of the regions of marked conformational restriction in MSP2, the conserved N-terminal region and the region around the disulfide in the conserved C-terminal region, coincide with peaks in the experimental antigenicity profiles of both 3D7 and FC27 MSP2, and with the epitopes of several monoclonal antibodies (Fig. 7). Likewise, antigenic regions within the repeats and dimorphic regions of FC27 correspond to those that show slightly elevated low-frequency spectral densities, indicative of conformational restriction. Although the GGSA repeat, which is the most flexible region in 3D7 MSP2, shows some antigenicity, this arises from a significant response in only a single rabbit (Fig. 6), suggesting that this very flexible region is only rarely antigenic (Fig. 7). Indeed, there is a significant correlation between the antigenicity profile and relaxation-based measures of conformational flexibility: Spearman's ρ for the comparison of the average antigenicity profiles over all mice with $J(0)$ are 0.35 and 0.54 for FC27 and 3D7 MSP2, respectively, and for rabbits 0.30 and 0.21 (two-tailed $p < 0.005$ for all comparisons, by permutation). Thus, it appears that restricted conformational disorder within MSP2 may be a robust predictor of local antigenicity. To explore this further, we divided the peptides into two equal groups according to the maximum value of $J(0)$ measured for the residues in each peptide, representing the conformationally constrained and flexible regions of MSP2 (Table 1). The peptides from constrained regions are almost four times as likely as the peptides from flexible regions to be significantly antigenic, while 85% of peptides that show significant responses in more than two animals (of either species) are from conformationally constrained regions of MSP2.

In contrast, sequence-based predictors of B-cell epitopes [60,61] perform poorly when applied to MSP2, showing weak and in some cases negative correlation with the experimental antigenicity, and failing to predict monoclonal antibody epitopes (Fig. 7). The Bepipred predictor [60] predicts 80 of 84 peptides in our array to contain B-cell epitopes, when in fact only 21 peptides reacted significantly with more than two antisera (Fig. 6), and these 21 peptides included two of the four peptides not predicted to be epitopes by this method. The approach of Kolaskar and Tongaonkar [61] performs only slightly better, predicting 35 peptides to contain epitopes, including 10 that reacted with more than two antisera.

Discussion

Intrinsically disordered proteins are increasingly attracting interest as potential vaccine candidates against malaria [11–15,36] and other pathogens [16,17]. In spite of this, little is known about the implications of conformational disorder for the development of an effective immune response. In the case of MSP2, the recombinant protein used for both experimental immunisation and clinical trials is highly disordered, as demonstrated previously [39] and characterised further here. On the other hand, the conformation of the native GPI-anchored protein is likely to be constrained, to a greater or lesser extent, by interactions with the merozoite membrane [48] or by self-association [36,46,47,62]. These interactions also modulate the accessibility of certain epitopes on the parasite surface [10]. In light of these observations, it is evident that the efficacy of an MSP2-based vaccine is likely to depend on the appropriate targeting of epitopes that exist in an accessible form on the parasite surface. Achieving this will require an improved understanding of the way antigen conformation and flexibility modulates the specificity of the immune response. As a first step to addressing this problem, we have compared the local

conformational dynamics of MSP2, as reported by ^{15}N relaxation measurements, with local antigenicity as inferred from experimental animal immunisations. We find that regions of MSP2 that are most antigenic correspond to those regions in which conformational flexibility is somewhat constrained, whereas those regions that are most flexible appear to be the least antigenic.

In contrast, we find no evidence that the polymorphic regions of MSP2 are particularly antigenic. Indeed, the most polymorphic region of MSP2, the GGSA repeats of 3D7, is also the most flexible and amongst the least antigenic regions. Other polymorphic regions (green and pink in Fig. 7) are no more antigenic than are the dimorphic and conserved regions. There is evidence that the polymorphisms within these regions are selectively favoured, and although details of these selective processes are unclear, they are expected to involve host immune pressure [63,64]. As such, the lack of obvious antigenic bias towards these regions is surprising, and may highlight important immunogenic differences between recombinant MSP2 and the native parasite antigen [10].

Previous studies of structured antigens have established that increased epitope flexibility tends to increase antigenicity [65,66], in contrast to the current findings. An important distinction is that these studies have addressed epitopes that are variably flexible loops in largely structured proteins. The most flexible of these loops are unlikely to be as flexible as even the least flexible regions of MSP2. In the model structured antigen lysozyme, all residues show positive steady-state ^1H - ^{15}N NOE values greater than 0.6 [67], reflecting markedly more constrained sub-ns dynamics than is seen for any region of MSP2 (Fig 4). One possible explanation for the apparent discrepancy, therefore, may be that a moderate degree of flexibility is optimal for antigenicity, with epitopes that are either too rigid, or too flexible, being less effective. Alternatively, the determinants of antigenicity in structured and disordered proteins may differ in a more fundamental way. For example, it has been suggested that the correlation between flexibility and antigenicity observed in structured proteins reflects accessibility, rather than flexibility *per se* [68], whereas the accessibility of potential epitopes in a disordered antigen is likely to be uniformly high. Perhaps consistent with this interpretation is our observation that epitope predictors, parameterised primarily on the basis of structured antigens, perform poorly for MSP2.

The consistency of the antigenic profiles we have measured here between animals and with other previous studies in mice and in humans, strongly suggests that these profiles are determined by the intrinsic immunogenicity of the recombinant MSP2 antigen. As such, the correlation we observe between conformational restriction and antigenicity probably reflects a tendency for more flexible regions of MSP2 to be less immunogenic. The mechanistic basis underlying this tendency is currently unclear, though several possible explanations are worthy of consideration. It has been suggested that the unusual residue composition of disordered and repetitive antigens may give rise to extensively cross-reactive responses, which fail to mature into high-affinity and specific antibodies [69]. Alternatively, it may be that conformational disorder itself frustrates the process by which a mature antibody response develops. Any disordered antigen exists in a vast ensemble of distinct conformations, but a developing antibody is likely to be limited in the range of conformations it is capable of recognising. The conformational diversity of disordered antigens may therefore impose a significant barrier to antibody maturation, as proposed for the *Staphylococcus aureus* fibronectin binding protein [70]. This effect may be viewed as a conformational analogue of the epitope dilution effect recently described in the context of a polyvalent vaccine of the polymorphic antigen apical membrane antigen 1 (AMA1) [71,72]. In that context, polymorphic epitopes are ‘diluted’ relative to conserved epitopes by the combination of diverse allelic forms of AMA1, resulting in an antibody response that favours conserved epitopes. In the current context, we envisage that epitope conformations

are diluted to an extent determined by the degree of disorder present in the epitope, with the result that the antibody response is biased towards more ordered epitopes.

Little is known about which MSP2 epitopes contribute to a protective immune response. Vaccine-derived protection mediated by MSP2 appears to be strain specific [23,73], suggesting that variable epitopes dominate. However this does may not be the case for the natural immune response to MSP2, where strain-specific protection has not been detected [37,74]. Nonetheless, a protective, strain-independent response is clearly desirable in the context of vaccine development. As such, our observation that conserved N- and C-terminal epitopes are amongst the most immunogenic regions of MSP2 is encouraging, although it is likely that not all of these epitopes will be accessible on the parasite surface [10].

The correlation established here begs the question of causation: is it possible to modulate the immunogenicity or antigenicity of a disordered antigen by altering its flexibility? Antigen flexibility could be modulated by directly modifying the antigen by addition of bulky residues or disulfide bonds at sites flanking a target epitope. Alternatively, simply changing the formulation of the antigen may have the desired effect. For example, the N-terminal region of MSP2 can be conformationally stabilised by interactions with lipid membranes, in a way that may better reflect the conformation of MSP2 on the merozoite surface [48]. These possibilities have important implications for the development of vaccines based on MSP2, where it is desirable to tune antigenicity towards epitopes that are conserved and exposed on the parasite surface [10,36].

Author Contributions

Conceived and designed the experiments: CAM JSR JN VS RFA RSN. Performed the experiments: MZ DA LZ. Analyzed the data: CAM MZ DA JN LZ JSR JGB. Contributed reagents/materials/analysis tools: CAM BK RFA. Wrote the paper: CAM MZ DA BK JN LZ VS JSR JGB RFA RSN.

References

1. Uversky VN, Oldfield CJ, Dunker AK. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. 2005; J Mol Recognit 18: 343–384. PMID: [16094605](#)
2. Oldfield CJ, Dunker AK. Intrinsically disordered proteins and intrinsically disordered protein regions. 2014; Annu Rev Biochem 83: 553–584. doi: [10.1146/annurev-biochem-072711-164947](#) PMID: [24606139](#)
3. Xue B, Blocquel D, Habchi J, Uversky AV, Kurgan L, Uversky VN, et al. Structural disorder in viral proteins. 2014; Chem Rev 114: 6880–6911. doi: [10.1021/cr4005692](#) PMID: [24823319](#)
4. Feng ZP, Zhang X, Han P, Arora N, Anders RF, Norton RS. Abundance of intrinsically unstructured proteins in *P. falciparum* and other apicomplexan parasite proteomes. 2006; Mol Biochem Parasitol 150: 256–267. PMID: [17010454](#)
5. Pavlovic MD, Jandric DR, Mitic NS. Epitope distribution in ordered and disordered protein regions. Part B—Ordered regions and disordered binding sites are targets of T- and B-cell immunity. 2014; J Immunol Methods 407: 90–107. doi: [10.1016/j.jim.2014.03.027](#) PMID: [24726865](#)
6. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z. Intrinsic disorder and protein function. 2002; Biochemistry 41: 6573–6582. PMID: [12022860](#)
7. Kwong PD, Doyle ML, Caspar DJ, Cicala C, Leavitt SA, Majeed S, et al. HIV-1 evades antibody-mediated neutralisation through conformational masking of receptor binding sites. 2002; Nature 420: 678–682. PMID: [12478295](#)
8. Stahl HD, Coppel RL, Brown GV, Saint R, Lingelbach K, Cowman AF, et al. Differential antibody screening of cloned *Plasmodium falciparum* sequences expressed in *Escherichia coli*: procedure for isolation of defined antigens and analysis of human antisera. 1984; Proc Natl Acad Sci U S A 81: 2456–2460. PMID: [6371814](#)
9. Kemp DJ, Coppel RL, Anders RF. Repetitive proteins and genes of malaria. 1987; Annu Rev Microbiol 41: 181–208. PMID: [3318667](#)

10. Adda CG, MacRaild CA, Reiling L, Wycherley K, Boyle MJ, Kienzie V, et al. Antigenic characterization of an intrinsically unstructured protein, *Plasmodium falciparum* merozoite surface protein 2. 2012; Infect Immun 80: 4177–4185. doi: [10.1128/IAI.00665-12](https://doi.org/10.1128/IAI.00665-12) PMID: [22966050](https://pubmed.ncbi.nlm.nih.gov/22966050/)
11. Raj DK, Nixon CP, Nixon CE, Dvorin JD, DiPetrillo CG, Pond-Tor S, et al. Antibodies to PfSEA-1 block parasite egress from RBCs and protect against malaria infection. 2014; Science 344: 871–877. doi: [10.1126/science.1254417](https://doi.org/10.1126/science.1254417) PMID: [24855263](https://pubmed.ncbi.nlm.nih.gov/24855263/)
12. Richards JS, Stanisic DI, Fowkes FJ, Tavul L, Dabod E, Thompson JK, et al. Association between naturally acquired antibodies to erythrocyte-binding antigens of *Plasmodium falciparum* and protection from malaria and high-density parasitemia. 2010; Clin Infect Dis 51: e50–60. doi: [10.1086/656413](https://doi.org/10.1086/656413) PMID: [20843207](https://pubmed.ncbi.nlm.nih.gov/20843207/)
13. Olugbile S, Kulangara C, Bang G, Bertholet S, Suzarte E, Villard V, et al. Vaccine potentials of an intrinsically unstructured fragment derived from the blood stage-associated *Plasmodium falciparum* protein PFF0165c. 2009; Infect Immun 77: 5701–5709. doi: [10.1128/IAI.00652-09](https://doi.org/10.1128/IAI.00652-09) PMID: [19786562](https://pubmed.ncbi.nlm.nih.gov/19786562/)
14. Yagi M, Bang G, Tougan T, Palacpac NM, Arisue N, Aoshi T, et al. Protective epitopes of the *Plasmodium falciparum* SERA5 malaria vaccine reside in intrinsically unstructured N-terminal repetitive sequences. 2014; PLoS One 9: e98460. doi: [10.1371/journal.pone.0098460](https://doi.org/10.1371/journal.pone.0098460) PMID: [24886718](https://pubmed.ncbi.nlm.nih.gov/24886718/)
15. Foquet L, Hermsen CC, van Gemert GJ, Van Braeckel E, Weening KE, Sauerwein R, et al. Vaccine-induced monoclonal antibodies targeting circumsporozoite protein prevent *Plasmodium falciparum* infection. 2014; J Clin Invest 124: 140–144. PMID: [24292709](https://pubmed.ncbi.nlm.nih.gov/24292709/)
16. Muster T, Steindl F, Purtscher M, Trkola A, Klima A, Himmeler G, et al. A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. 1993; J Virol 67: 6642–6647. PMID: [7692082](https://pubmed.ncbi.nlm.nih.gov/7692082/)
17. Foucault M, Mayol K, Receveur-Brechot V, Bussat MC, Klinguer-Hamour C, Verrier B, 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. 2010; Proteins 78: 1441–1456. doi: [10.1002/prot.22661](https://doi.org/10.1002/prot.22661) PMID: [20034112](https://pubmed.ncbi.nlm.nih.gov/20034112/)
18. Dyson HJ, Satterthwait AC, Lerner RA, Wright PE. Conformational preferences of synthetic peptides derived from the immunodominant site of the circumsporozoite protein of *Plasmodium falciparum* by ¹H NMR. 1990; Biochemistry 29: 7828–7837. PMID: [2261440](https://pubmed.ncbi.nlm.nih.gov/2261440/)
19. Boyle MJ, Langer C, Chan JA, Hodder AN, Coppel RL, Anders RF, et al. Sequential processing of merozoite surface proteins during and after erythrocyte invasion by *Plasmodium falciparum*. 2014; Infect Immun 82: 924–936. doi: [10.1128/IAI.00866-13](https://doi.org/10.1128/IAI.00866-13) PMID: [24218484](https://pubmed.ncbi.nlm.nih.gov/24218484/)
20. al-Yaman F, Genton B, Anders R, Taraika J, Ginny M, Mellor S, et al. Assessment of the role of the humoral response to *Plasmodium falciparum* MSP2 compared to RESA and SPf66 in protecting Papua New Guinean children from clinical malaria. 1995; Parasite Immunol 17: 493–501. PMID: [8552419](https://pubmed.ncbi.nlm.nih.gov/8552419/)
21. Aucan C, Traore Y, Tall F, Nacro B, Traore-Leroux T, Fumoux F, et al. High immunoglobulin G2 (IgG2) and low IgG4 levels are associated with human resistance to *Plasmodium falciparum* malaria. 2000; Infect Immun 68: 1252–1258. PMID: [10678934](https://pubmed.ncbi.nlm.nih.gov/10678934/)
22. Ekala MT, Jouin H, Lekoulou F, Mercereau-Puijalon O, Ntoumi F. Allelic family-specific humoral responses to merozoite surface protein 2 (MSP2) in Gabonese residents with *Plasmodium falciparum* infections. 2002; Clin Exp Immunol 129: 326–331. PMID: [12165090](https://pubmed.ncbi.nlm.nih.gov/12165090/)
23. Flück C, Smith T, Beck HP, Irion A, Betuela I, Alpers MP, et al. Strain-specific humoral response to a polymorphic malaria vaccine. 2004; Infect Immun 72: 6300–6305. PMID: [15501757](https://pubmed.ncbi.nlm.nih.gov/15501757/)
24. al-Yaman F, Genton B, Anders RF, Falk M, Triglia T, Lewis D, et al. Relationship between humoral response to *Plasmodium falciparum* merozoite surface antigen-2 and malaria morbidity in a highly endemic area of Papua New Guinea. 1994; Am J Trop Med Hyg 51: 593–602. PMID: [7985752](https://pubmed.ncbi.nlm.nih.gov/7985752/)
25. Genton B, Betuela I, Felger I, Al-Yaman F, Anders RF, Saul A, 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. 2002; J Infect Dis 185: 820–827. PMID: [11920300](https://pubmed.ncbi.nlm.nih.gov/11920300/)
26. Osier FH, Feng G, Boyle MJ, Langer C, Zhou J, Richards JS, et al. Opsonic phagocytosis of *Plasmodium falciparum* merozoites: mechanism in human immunity and a correlate of protection against malaria. 2014; BMC Med 12: 108. doi: [10.1186/1741-7015-12-108](https://doi.org/10.1186/1741-7015-12-108) PMID: [24980799](https://pubmed.ncbi.nlm.nih.gov/24980799/)
27. Stanisic DI, Richards JS, McCallum FJ, Michon P, King CL, Schoepflin S, et al. Immunoglobulin G subclass-specific responses against *Plasmodium falciparum* merozoite antigens are associated with control of parasitemia and protection from symptomatic illness. 2009; Infect Immun 77: 1165–1174. doi: [10.1128/IAI.01129-08](https://doi.org/10.1128/IAI.01129-08) PMID: [19139189](https://pubmed.ncbi.nlm.nih.gov/19139189/)
28. Taylor RR, Allen SJ, Greenwood BM, Riley EM. IgG3 antibodies to *Plasmodium falciparum* merozoite surface protein 2 (MSP2): increasing prevalence with age and association with clinical immunity to malaria. 1998; Am J Trop Med Hyg 58: 406–413. PMID: [9574783](https://pubmed.ncbi.nlm.nih.gov/9574783/)

29. Osier FH, Fegan G, Polley SD, Murungi L, Verra F, Tetteh KK, et al. Breadth and magnitude of antibody responses to multiple *Plasmodium falciparum* merozoite antigens are associated with protection from clinical malaria. 2008; *Infect Immun* 76: 2240–2248. doi: [10.1128/IAI.01585-07](https://doi.org/10.1128/IAI.01585-07) PMID: [18316390](https://pubmed.ncbi.nlm.nih.gov/18316390/)
30. Stubbs J, Olugbile S, Saidou B, Simpore J, Corradin G, Lanzavecchia A. Strain-transcending Fc-dependent killing of *Plasmodium falciparum* by merozoite surface protein 2 allele-specific human antibodies. 2011; *Infect Immun* 79: 1143–1152. doi: [10.1128/IAI.01034-10](https://doi.org/10.1128/IAI.01034-10) PMID: [21189324](https://pubmed.ncbi.nlm.nih.gov/21189324/)
31. McCarthy JS, Marjason J, Elliott S, Fahey P, Bang G, Malkin E, et al. A phase 1 trial of MSP2-C1, a blood-stage malaria vaccine containing 2 isoforms of MSP2 formulated with Montanide ISA 720. 2011; *PLoS One* 6: e24413. doi: [10.1371/journal.pone.0024413](https://doi.org/10.1371/journal.pone.0024413) PMID: [21949716](https://pubmed.ncbi.nlm.nih.gov/21949716/)
32. Smythe JA, Coppel RL, Day KP, Martin RK, Oduola AM, Kemp DJ, et al. Structural diversity in the *Plasmodium falciparum* merozoite surface antigen 2. 1991; *Proc Natl Acad Sci U S A* 88: 1751–1755. PMID: [2000383](https://pubmed.ncbi.nlm.nih.gov/2000383/)
33. Fenton B, Clark JT, Khan CM, Robinson JV, Walliker D, Ridley R, et al. Structural and antigenic polymorphism of the 35- to 48-kilodalton merozoite surface antigen (MSA-2) of the malaria parasite *Plasmodium falciparum*. 1991; *Mol Cell Biol* 11: 963–971. PMID: [1990294](https://pubmed.ncbi.nlm.nih.gov/1990294/)
34. Smythe JA, Peterson MG, Coppel RL, Saul AJ, Kemp DJ, Anders RF. Structural diversity in the 45-kilodalton merozoite surface antigen of *Plasmodium falciparum*. 1990; *Mol Biochem Parasitol* 39: 227–234. PMID: [2181307](https://pubmed.ncbi.nlm.nih.gov/2181307/)
35. Thomas AW, Carr DA, Carter JM, Lyon JA. Sequence comparison of allelic forms of the *Plasmodium falciparum* merozoite surface antigen MSA2. 1990; *Mol Biochem Parasitol* 43: 211–220. PMID: [2090943](https://pubmed.ncbi.nlm.nih.gov/2090943/)
36. Anders RF, Adda CG, Foley M, Norton RS. Recombinant protein vaccines against the asexual blood-stages of *Plasmodium falciparum*. 2010; *Hum Vaccin* 6: 1–15.
37. Osier FH, Murungi LM, Fegan G, Tuju J, Tetteh KK, Bull PC, et al. Allele-specific antibodies to *Plasmodium falciparum* merozoite surface protein-2 and protection against clinical malaria. 2010; *Parasite Immunol* 32: 193–201. doi: [10.1111/j.1365-3024.2009.01178.x](https://doi.org/10.1111/j.1365-3024.2009.01178.x) PMID: [20398182](https://pubmed.ncbi.nlm.nih.gov/20398182/)
38. Taylor RR, Smith DB, Robinson VJ, McBride JS, Riley EM. Human antibody response to *Plasmodium falciparum* merozoite surface protein 2 is serogroup specific and predominantly of the immunoglobulin G3 subclass. 1995; *Infect Immun* 63: 4382–4388. PMID: [7591074](https://pubmed.ncbi.nlm.nih.gov/7591074/)
39. Zhang X, Perugini MA, Yao S, Adda CG, Murphy VJ, Low A, et al. Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2. 2008; *J Mol Biol* 379: 105–121. doi: [10.1016/j.jmb.2008.03.039](https://doi.org/10.1016/j.jmb.2008.03.039) PMID: [18440022](https://pubmed.ncbi.nlm.nih.gov/18440022/)
40. Kazimierczuk K, Zawadzka-Kazimierczuk A, Kozminski W. Non-uniform frequency domain for optimal exploitation of non-uniform sampling. 2010; *J Magn Reson* 205: 286–292. doi: [10.1016/j.jmr.2010.05.012](https://doi.org/10.1016/j.jmr.2010.05.012) PMID: [20547466](https://pubmed.ncbi.nlm.nih.gov/20547466/)
41. Ferrage F, Cowburn D, Ghose R. Accurate sampling of high-frequency motions in proteins by steady-state $^{15}\text{N}\{-^1\text{H}\}$ nuclear Overhauser effect measurements in the presence of cross-correlated relaxation. 2009; *J Am Chem Soc* 131: 6048–6049. doi: [10.1021/ja809526q](https://doi.org/10.1021/ja809526q) PMID: [19358609](https://pubmed.ncbi.nlm.nih.gov/19358609/)
42. Korzhnev DM, Billeter M, Arseniev AS, Orekhov VY. NMR studies of Brownian tumbling and internal motions in proteins. 2001; *Prog Nucl Magn Reson Spectrosc* 38: 197–266.
43. Pelupessy P, Espallargas GM, Bodenhausen G. Symmetrical reconversion: measuring cross-correlation rates with enhanced accuracy. 2003; *J Magn Reson* 161: 258–264. PMID: [12713978](https://pubmed.ncbi.nlm.nih.gov/12713978/)
44. Pelupessy P, Ferrage F, Bodenhausen G. Accurate measurement of longitudinal cross-relaxation rates in nuclear magnetic resonance. 2007; *J Chem Phys* 126: 134508. PMID: [17430048](https://pubmed.ncbi.nlm.nih.gov/17430048/)
45. Motackova V, Novacek J, Zawadzka-Kazimierczuk A, Kazimierczuk K, Zidek L, Sanderova H, et al. Strategy for complete NMR assignment of disordered proteins with highly repetitive sequences based on resolution-enhanced 5D experiments. 2010; *J Biomol NMR* 48: 169–177. doi: [10.1007/s10858-010-9447-3](https://doi.org/10.1007/s10858-010-9447-3) PMID: [20890634](https://pubmed.ncbi.nlm.nih.gov/20890634/)
46. Adda CG, Murphy VJ, Sunde M, Waddington LJ, Schloegel J, Talbo GH, et al. *Plasmodium falciparum* merozoite surface protein 2 is unstructured and forms amyloid-like fibrils. 2009; *Mol Biochem Parasitol* 166: 159–171. doi: [10.1016/j.molbiopara.2009.03.012](https://doi.org/10.1016/j.molbiopara.2009.03.012) PMID: [19450733](https://pubmed.ncbi.nlm.nih.gov/19450733/)
47. Low A, Chandrashekar IR, Adda CG, Yao S, Sabo JK, Zhang X, et al. Merozoite surface protein 2 of *Plasmodium falciparum*: expression, structure, dynamics, and fibril formation of the conserved N-terminal domain. 2007; *Biopolymers* 87: 12–22. PMID: [17516503](https://pubmed.ncbi.nlm.nih.gov/17516503/)
48. MacRaild CA, Pedersen MØ, Anders RF, Norton RS. Lipid interactions of the malaria antigen merozoite surface protein 2. 2012; *Biochim Biophys Acta* 1818: 2572–2578. doi: [10.1016/j.bbame.2012.06.015](https://doi.org/10.1016/j.bbame.2012.06.015) PMID: [22749949](https://pubmed.ncbi.nlm.nih.gov/22749949/)
49. Sharma D, Rajarathnam K. ^{13}C NMR chemical shifts can predict disulfide bond formation. 2000; *J Biomol NMR* 18: 165–171. PMID: [11101221](https://pubmed.ncbi.nlm.nih.gov/11101221/)

50. Tamiola K, Acar B, Mulder FAA. Sequence-specific random coil chemical shifts of intrinsically disordered proteins. 2010; J Am Chem Soc 132: 18000–18003. doi: [10.1021/ja105656t](https://doi.org/10.1021/ja105656t) PMID: [21128621](https://pubmed.ncbi.nlm.nih.gov/21128621/)
51. Palmer AG. NMR characterization of the dynamics of biomacromolecules. 2004; Chemical Reviews 104: 3623–3640. PMID: [15303831](https://pubmed.ncbi.nlm.nih.gov/15303831/)
52. Jensen MR, Zweckstetter M, Huang JR, Blackledge M. Exploring Free-Energy Landscapes of Intrinsically Disordered Proteins at Atomic Resolution Using NMR Spectroscopy. 2014; Chem Rev 114: 6632–6660. doi: [10.1021/cr400688u](https://doi.org/10.1021/cr400688u) PMID: [24725176](https://pubmed.ncbi.nlm.nih.gov/24725176/)
53. Farrow NA, Zhang O, Szabo A, Torchia DA, Kay LE. Spectral density function mapping using ^{15}N relaxation data exclusively. 1995; J Biomol NMR 6: 153–162. PMID: [8589604](https://pubmed.ncbi.nlm.nih.gov/8589604/)
54. Kaderavek P, Zapletal V, Rabatinova A, Krasny L, Sklenar V, Zidek L. Spectral density mapping protocols for analysis of molecular motions in disordered proteins. 2014; J Biomol NMR 58: 193–207. doi: [10.1007/s10858-014-9816-4](https://doi.org/10.1007/s10858-014-9816-4) PMID: [24515886](https://pubmed.ncbi.nlm.nih.gov/24515886/)
55. Mittermaier AK, Kay LE. Observing biological dynamics at atomic resolution using NMR. 2009; Trends Biochem Sci 34: 601–611. doi: [10.1016/j.tibs.2009.07.004](https://doi.org/10.1016/j.tibs.2009.07.004) PMID: [19846313](https://pubmed.ncbi.nlm.nih.gov/19846313/)
56. 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. 2000; Parasite Immunol 22: 211–221. PMID: [10792760](https://pubmed.ncbi.nlm.nih.gov/10792760/)
57. Saul A, Lawrence G, Smillie A, Rzepczyk CM, Reed C, Taylor D, et al. Human phase I vaccine trials of 3 recombinant asexual stage malaria antigens with Montanide ISA720 adjuvant. 1999; Vaccine 17: 3145–3159. PMID: [10462251](https://pubmed.ncbi.nlm.nih.gov/10462251/)
58. Epping RJ, Goldstone SD, Ingram LT. An epitope recognised by inhibitory monoclonal antibodies that react with a 51 kilodalton merozoite surface antigen in *Plasmodium falciparum*. 1988; Mol Biochem Parasitol 28: 1–10. PMID: [2453800](https://pubmed.ncbi.nlm.nih.gov/2453800/)
59. Saul A, Lord R, Jones G, Geysen HM, Gale J, Mollard R. Cross-reactivity of antibody against an epitope of the *Plasmodium falciparum* second merozoite surface antigen. 1989; Parasite Immunol 11: 593–601. PMID: [2482473](https://pubmed.ncbi.nlm.nih.gov/2482473/)
60. Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. 2006; Immunome Res 2: 2. PMID: [16635264](https://pubmed.ncbi.nlm.nih.gov/16635264/)
61. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. 1990; FEBS Lett 276: 172–174. PMID: [1702393](https://pubmed.ncbi.nlm.nih.gov/1702393/)
62. Yang X, Adda CG, MacRaild CA, Low A, Zhang X, Zeng W, et al. Identification of key residues involved in fibril formation by the conserved N-terminal region of *Plasmodium falciparum* merozoite surface protein 2 (MSP2). 2010; Biochimie 92: 1287–1295. doi: [10.1016/j.biochi.2010.06.001](https://doi.org/10.1016/j.biochi.2010.06.001) PMID: [20542076](https://pubmed.ncbi.nlm.nih.gov/20542076/)
63. Putaporntip C, Jongwutiwes S, Hughes AL. Differential selective pressures on the merozoite surface protein 2 locus of *Plasmodium falciparum* in a low endemic area. 2008; Gene 427: 51–57. doi: [10.1016/j.gene.2008.09.009](https://doi.org/10.1016/j.gene.2008.09.009) PMID: [18840512](https://pubmed.ncbi.nlm.nih.gov/18840512/)
64. Zilversmit MM, Volkman SK, DePristo MA, Wirth DF, Awadalla P, Hartl DL. Low-complexity regions in *Plasmodium falciparum*: missing links in the evolution of an extreme genome. 2010; Mol Biol Evol 27: 2198–2209. doi: [10.1093/molbev/msq108](https://doi.org/10.1093/molbev/msq108) PMID: [20427419](https://pubmed.ncbi.nlm.nih.gov/20427419/)
65. Ofek G, Guenaga FJ, Schief WR, Skinner J, Baker D, Wyatt R, et al. Elicitation of structure-specific antibodies by epitope scaffolds. 2010; Proc Natl Acad Sci U S A 107: 17880–17887. doi: [10.1073/pnas.1004728107](https://doi.org/10.1073/pnas.1004728107) PMID: [20876137](https://pubmed.ncbi.nlm.nih.gov/20876137/)
66. Westhof E, Altschuh D, Moras D, Bloomer AC, Mondragon A, Klug A, et al. Correlation between segmental mobility and the location of antigenic determinants in proteins. 1984; Nature 311: 123–126. PMID: [6206398](https://pubmed.ncbi.nlm.nih.gov/6206398/)
67. Buck M, Boyd J, Redfield C, MacKenzie DA, Jeenes DJ, Archer DB, et al. Structural determinants of protein dynamics: analysis of ^{15}N NMR relaxation measurements for main-chain and side-chain nuclei of hen egg white lysozyme. 1995; Biochemistry 34: 4041–4055. PMID: [7696270](https://pubmed.ncbi.nlm.nih.gov/7696270/)
68. Novotny J, Handschumacher M, Haber E, Brucoleri RE, Carlson WB, Fanning DW, et al. Antigenic determinants in proteins coincide with surface regions accessible to large probes (antibody domains). 1986; Proc Natl Acad Sci U S A 83: 226–230. PMID: [2417241](https://pubmed.ncbi.nlm.nih.gov/2417241/)
69. Anders RF. Multiple cross-reactivities amongst antigens of *Plasmodium falciparum* impair the development of protective immunity against malaria. 1986; Parasite Immunol 8: 529–539. PMID: [3543808](https://pubmed.ncbi.nlm.nih.gov/3543808/)
70. Penkett CJ, Redfield C, Jones JA, Dodd I, Hubbard J, Smith RA, et al. Structural and dynamical characterization of a biologically active unfolded fibronectin-binding protein from *Staphylococcus aureus*. 1998; Biochemistry 37: 17054–17067. PMID: [9836601](https://pubmed.ncbi.nlm.nih.gov/9836601/)
71. Dutta S, Dlugosz LS, Drew DR, Ge X, Ababacar D, Rovira YI, et al. Overcoming antigenic diversity by enhancing the immunogenicity of conserved epitopes on the malaria vaccine candidate apical

membrane antigen-1. 2013; PLoS Pathog 9: e1003840. doi: [10.1371/journal.ppat.1003840](https://doi.org/10.1371/journal.ppat.1003840) PMID: [24385910](https://pubmed.ncbi.nlm.nih.gov/24385910/)

72. Chaudhury S, Reifman J, Wallqvist A. Simulation of B Cell Affinity Maturation Explains Enhanced Antibody Cross-Reactivity Induced by the Polyvalent Malaria Vaccine AMA1. 2014; J Immunol 193: 2073–2086. doi: [10.4049/jimmunol.1401054](https://doi.org/10.4049/jimmunol.1401054) PMID: [25080483](https://pubmed.ncbi.nlm.nih.gov/25080483/)
73. Flück C, Schopflin S, Smith T, Genton B, Alpers MP, Beck HP, et al. Effect of the malaria vaccine Combination B on merozoite surface antigen 2 diversity. 2007; Infect Genet Evol 7: 44–51. PMID: [16647307](https://pubmed.ncbi.nlm.nih.gov/16647307/)
74. Scopel KK, da Silva-Nunes M, Malafronte RS, Braga EM, Ferreira MU. Variant-specific antibodies to merozoite surface protein 2 and clinical expression of *Plasmodium falciparum* malaria in rural Amazonians. 2007; Am J Trop Med Hyg 76: 1084–1091. PMID: [17556615](https://pubmed.ncbi.nlm.nih.gov/17556615/)

CHAPTER 5

Strain-transcending immune response generated
by chimeras of the malaria vaccine candidate
merozoite surface protein 2

5.1 Declaration of Thesis Chapter 5

Declaration by candidate

In the case of **Chapter 5**, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Designed, cloned, expressed, purified and characterized the proteins used in this study. Analysed the results from mice immunization experiments and prepared manuscript	50

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Dean Andrew	Performed experiments	
Christopher A. MacRaid	Intellectual input and manuscript preparation	
Rodrigo A. V. Morales	Manuscript preparation	
James G. Beeson	Intellectual input	
Robin F. Anders	Intellectual input and manuscript preparation	
Jack S. Richards	Intellectual input and manuscript preparation	
Raymond S. Norton	Intellectual input and manuscript preparation	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

**Candidate's
Signature**

	Date 30/03/2016
---	---------------------------

**Main
Supervisor's
Signature**

	Date 30/03/2016
---	---------------------------

5.2 Introduction

Chapter 4 revealed that the immunogenicity of MSP2 can be enhanced by increasing its conformational order and that this may have implications in designing better MSP2 based malaria vaccines. The immune response to MSP2 is protective but strain-specific owing to the polymorphic nature of MSP2. Moreover, the role of conserved regions in generating the protective immune response is unclear and it would seem logical that targeting these conserved epitopes could mediate strain-transcending immunity.

Therefore, in this chapter, immunization experiments were carried out with a series of MSP2 constructs including chimeras covering 3D7 and FC27 MSP2 alleles as well as truncated MSP2 devoid of either *N*-terminal, *C*-terminal or both conserved regions. All the constructs were expressed and purified, followed by characterization using LC-MS, analytical HPLC and endotoxin assay etc. Antigens were formulated in Montanide ISA720 and the immunization experiments were carried out in C57BL/6 mice. The specificity of antibodies to all the constructs was characterized using a peptide array and total IgG and IgG sub-types were quantified using ELISA. Furthermore, the antisera for each construct were characterized by Western blot to characterize the ability of antibodies in recognizing native MSP2. Among the chimeras tested, V_{3D7}V_{FC27}C showed a uniform distribution of epitopes across 3D7 and FC27 MSP2 alleles. This construct was able to induce IgG levels that were equivalent to that obtained by immunizing with separate 3D7+FC27 antigens. The highest IgG2b and IgG2c responses that mediate the Fc receptor-dependent processes like complement-mediated growth inhibition and opsonic phagocytosis, were also observed for this construct. Removal of conserved regions reduced the immune response to the variable region epitopes, suggesting that conserved regions are important for a robust immune response. As these results have been published in *Scientific Reports* journal, they are presented in the format of a published article in the immediate section below.

SCIENTIFIC REPORTS

OPEN

Strain-transcending immune response generated by chimeras of the malaria vaccine candidate merozoite surface protein 2

Bankala Krishnarjuna^{1,*}, Dean Andrew^{2,*}, Christopher A. MacRaild¹, Rodrigo A. V. Morales¹, James G. Beeson^{2,3,4}, Robin F. Anders⁵, Jack S. Richards^{2,3,4} & Raymond S. Norton¹

MSP2 is an intrinsically disordered protein that is abundant on the merozoite surface and essential to the parasite *Plasmodium falciparum*. Naturally-acquired antibody responses to MSP2 are biased towards dimorphic sequences within the central variable region of MSP2 and have been linked to naturally-acquired protection from malaria. In a phase IIb study, an MSP2-containing vaccine induced an immune response that reduced parasitemias in a strain-specific manner. A subsequent phase I study of a vaccine that contained both dimorphic forms of MSP2 induced antibodies that exhibited functional activity *in vitro*. We have assessed the contribution of the conserved and variable regions of MSP2 to the generation of a strain-transcending antibody response by generating MSP2 chimeras that included conserved and variable regions of the 3D7 and FC27 alleles. Robust anti-MSP2 antibody responses targeting both conserved and variable regions were generated in mice, although the fine specificity and the balance of responses to these regions differed amongst the constructs tested. We observed significant differences in antibody subclass distribution in the responses to these chimeras. Our results suggest that chimeric MSP2 antigens can elicit a broad immune response suitable for protection against different strains of *P. falciparum*.

Malaria causes over 200 million cases and nearly 650,000 deaths per year worldwide¹, with the most severe disease caused by *Plasmodium falciparum*. Recent gains in malaria control and progress towards elimination and eventual eradication are under threat because of the rise of drug-resistant parasites² and insecticide-resistant mosquitoes³. Eventual global eradication is likely to depend on the development of multi-stage vaccines that prevent or limit blood-stage infection and block ongoing transmission. Previous clinical trials of blood-stage antigens such as merozoite surface protein 2 (MSP2) and apical membrane antigen 1 (AMA1) suggested that vaccine-induced immune responses can reduce the parasite burden, although their efficacy is potentially limited by the antigenic polymorphisms of *Plasmodium* blood-stage antigens^{4,5}.

MSP2 is a ~23-kDa surface coat protein essential for survival of the asexual blood-stages of *P. falciparum*^{6–8}. MSP2 consists of a central variable region (VR) that constitutes 60% of the protein, flanked by conserved N-terminal and C-terminal regions (NTR and CTR, respectively)^{9,10}. All forms of MSP2 can be grouped into two allelic families, 3D7 and FC27, based on dimorphic sequences within the VR^{9–13}. These sequences are interspersed by more polymorphic repeat sequences, with the 3D7 allele characterized by highly variable GSA-rich repeats, and FC27-MSP2 harbouring tandem repeats of 32- and 12-residue sequences (Fig. 1A)^{9,14–16}. MSP2 is highly disordered, lacking any well-defined conformation in solution¹⁷, and the conformational and antigenic properties of various MSP2 regions are defined entirely by local sequence features^{17,18}. In particular, all MSP2 epitopes mapped to date consist of short linear sequences¹⁹.

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

²Centre for Biomedical Research, Burnet Institute, Melbourne, Victoria 3004, Australia. ³Department of Microbiology, Monash University, Clayton 3800, Australia. ⁴Department of Medicine, University of Melbourne, Parkville 3052, Australia. ⁵Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086, Australia. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to R.S.N. (email: ray.norton@monash.edu) or J.S.R. (email: richards@burnet.edu.au)

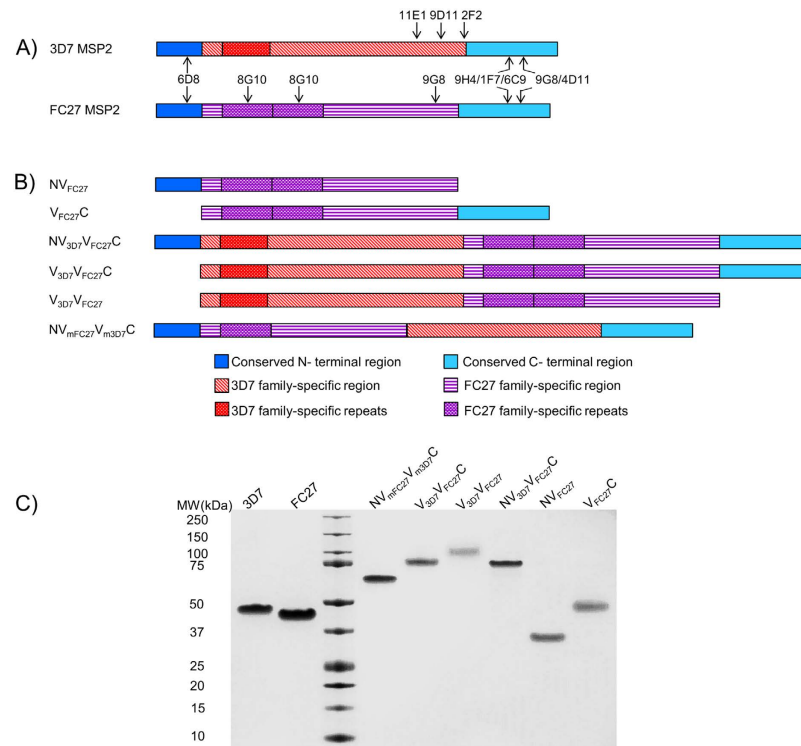


Figure 1. (A) Schematic representation of recombinant 3D7 and FC27 MSP2 constructs and the binding regions of previously identified mouse monoclonal antibodies¹⁹. (B) Representation of truncated and chimeric MSP2 constructs used in this work. Truncated constructs were derived from the excision of either conserved N-terminal or C-terminal regions. Chimeric constructs were generated by the fusion of 3D7 and FC27 variable regions in the presence or absence of the conserved N- or C-terminus. (C) SDS-PAGE analysis of purified full-length 3D7 MSP2, FC27 MSP2 and other MSP2 constructs: NV_{FC27} - FC27 MSP2 without the NTR (FC27 MSP2₁₋₁₇₁); V_{FC27}C - FC27 MSP2 without NTR (FC27 MSP2₂₆₋₂₂₁); NV_{3D7}V_{FC27}C - NTR, 3D7 MSP2₁₋₁₇₉ followed by FC27 MSP2₂₆₋₁₇₁ (with two 12-residue repeats) and CTR; V_{3D7}V_{FC27}C - 3D7 MSP2₂₆₋₁₇₉ followed by the FC27 MSP2₂₆₋₁₇₁ and CTR; V_{3D7}V_{FC27} - VR of 3D7 MSP2₂₆₋₁₇₉ followed by the VR of FC27 MSP2₂₆₋₁₇₁; NV_{FC27}V_{m3D7}C - NTR, followed by the modified variable regions of FC27 MSP2₆₄₋₁₇₁, 3D7 MSP2₅₈₋₁₇₉ and CTR. Gel was loaded with 1.25 µg of protein per well and stained with Coomassie Blue.

Anti-MSP2 antibodies are generated following natural infections and correlate with the age-dependent protection observed in adults and children over the age of 5 years^{20,21}. Anti-MSP2 antibodies are mostly of the IgG1 and IgG3 subtypes and function through antibody-dependent cellular inhibition (ADCI), opsonic phagocytosis and complement-mediated mechanisms²²⁻²⁴. The acquisition of anti-MSP2 antibodies alone or in combination with antibodies targeting other *P. falciparum* antigens is an important indicator of a reduced risk of clinical infection^{20,23,25}.

The Combination B vaccine, composed of full-length 3D7 MSP2 (Ag1624), MSP1 (190LCS.T3, N-terminal end, K1 allele) and the C-terminal 70% of RESA (Ag1505H), adjuvanted in Montanide ISA720, reduced parasite densities by 62% when tested in a phase I-IIb trial in Papua New Guinean children²⁶. However, the reduction in parasitemia was biased towards parasites containing the 3D7 MSP2 allele^{26,27}, and consequently a higher incidence of morbidity associated with FC27 MSP2-type parasites was reported. A subsequent phase I trial of a combination of 3D7 and FC27 alleles of full-length MSP2, adjuvanted in Montanide ISA720²⁴, exhibited functional activity *in vitro* that included ADCI²⁴ and complement-mediated inhibition of parasite growth²². These trials suggested that MSP2 vaccines can induce functional responses that may mediate *in vivo* protection, but highlighted the need to evaluate the benefits of including the VR from both alleles, as well as the conserved NTR and CTR.

The functional role of NTR and CTR epitopes is unclear but it would seem logical that targeting these conserved epitopes could mediate strain-transcending immunity. Some of the antibodies to these conserved epitopes do not recognize native MSP2, despite recognizing recombinant MSP2^{19,28}. Moreover, it has been observed that the NTR can contribute to fibril formation of recombinant MSP2^{29,30}. It is therefore important to determine which aspects of the NTR and CTR are advantageous in construct design. In order to overcome the allele-specific immune response, it has been proposed that the VR of the two allelic types of MSP2 (3D7 and FC27) should be included in an MSP2 vaccine^{24,31}. In this study, we sought to determine the effects of immunizing with six truncated and chimeric MSP2 constructs compared to using a combination of both alleles of full-length MSP2³². Chimeric antigens also offer the potential advantages of simplified manufacture, simplified quality assurance and reduced costs, which may be particularly important in the context of future multi-component vaccines targeting multiple stages of the *Plasmodium* life cycle³³.

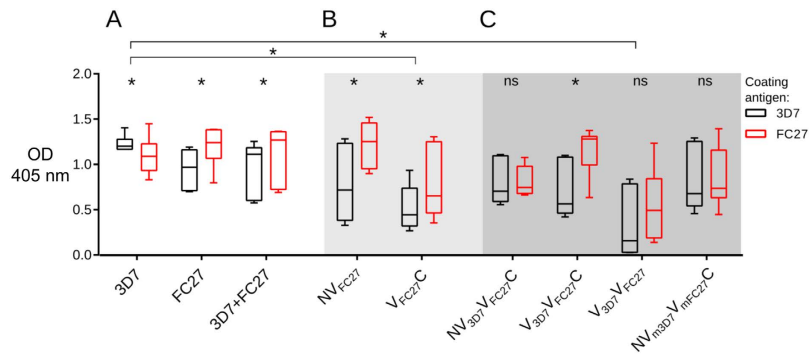


Figure 2. Anti-MSP2 responses of mice immunized with recombinant MSP2 constructs. Total IgG levels tested against 3D7 (black) and FC27 (red) MSP2 measured by ELISA. Responses to 3D7 and FC27 MSP2 were compared with those to the mixture of the two alleles (A), to N- and C-terminally truncated FC27 MSP2 (B) and to MSP2 chimeras (C). The difference in responses to 3D7 and FC27 allelic variants was compared using the Wilcoxon matched pairs test, while between-group comparisons were made by the Kruskal-Wallis test. * denotes $p < 0.05$; OD, optical density.

In this work we have exploited the highly disordered structure of MSP2, in which all epitopes mapped to date consist of short linear sequences¹⁹, to investigate the roles of the conserved and variable regions of 3D7 and FC27 MSP2 in shaping the antibody repertoire against these proteins. A series of MSP2 constructs was designed to accommodate both allelic forms of the protein as well as permutations in their conserved, polymorphic and repeat regions. These chimeras allowed us to test whether, by manipulating the composition of these chimeras, we could enhance the immune response, target it toward potentially protective epitopes, remove the propensity of MSP2 for aggregation, and modulate the IgG subclass distribution elicited by MSP2.

Results

Rationale, design and production of engineered MSP2 constructs. We designed a set of chimeric antigens (Fig. 1B) consisting of different regions of 3D7 and FC27 MSP2 in order to address two issues: 1) could chimeric MSP2 constructs induce effective antibody responses to both 3D7 and FC27 forms of MSP2, and 2) what impact did the conserved regions have on the immunogenicity and aggregation propensity of MSP2. We name these constructs using a scheme in which the NTR and CTR are denoted as N and C, respectively, and the VR of 3D7 and FC27 are denoted as V_{3D7} and V_{FC27}, respectively. Accordingly constructs, NV_{FC27} and V_{FC27}C, represent FC27 MSP2 lacking the conserved CTR or NTR, respectively, while NV_{3D7}V_{FC27}C is a simple chimera containing both VRs flanked by the NTR and CTR, and V_{3D7}V_{FC27}C and V_{3D7}V_{FC27}C lack one or both conserved regions, respectively. In NV_{mFC27}V_{m3D7}C, the order of the VRs is inverted, and the number of repeat regions reduced in order to assess the role of the VR in aggregation propensity as well as the significance of tandem repeat sequences for immunogenicity of other epitopes of MSP2 (Fig. 1B). The complete amino acid sequences of all constructs are given in Table S1.

All constructs were produced in an *E. coli* expression system optimised for high-yield expression of MSP2¹⁸. The purified proteins migrated as single bands in SDS-PAGE (Fig. 1C) and were pure by analytical HPLC (Fig. S1). All MSP2 constructs used had anomalous relative molecular masses on SDS-PAGE (Fig. 1C), as expected for highly hydrophilic intrinsically disordered proteins³⁴. However, LC-MS data confirmed that all MSP2 constructs had the expected molecular mass (Table S2) and purity >99% (Fig. S1). Estimated endotoxin levels were less than 0.1 EU/μg protein (Table S2).

Protein aggregation is reduced in chimeric constructs. Both allelic forms of full-length recombinant MSP2 are intrinsically disordered and are prone to aggregation and fibril formation in solution^{17,35}. This aggregation is driven by the conserved NTR of MSP2, but FC27 MSP2 is much more prone to aggregation than 3D7 MSP2, indicating that variable-region sequences also play a role in modulating aggregation propensity^{29,30,35,36}. Size-exclusion chromatographic analysis of the six designed MSP2 constructs used in this study detected aggregation only in NV_{FC27} (Fig. S2), which consists of the NTR and the VR of FC27, but lacks the CTR. NV_{3D7}V_{FC27}C, which contains the conserved NTR followed by 3D7 VR, was not prone to aggregation (Fig. S2), confirming that residues following the NTR modulate aggregation.

Immunogenicity of the MSP2 constructs. Montanide ISA720 was selected as the adjuvant for the mouse immunogenicity studies because of its use in previous phase I and II studies of MSP2^{24,37}. Endpoint titres (EPT) were calculated for total IgG responses and compared with a single sera dilution of 1:20,000 to ensure that this sera dilution was within the linear range of the assay. EPT and the single sera dilution at 1:20,000 were highly correlated (Fig. S3), so single sera dilutions were used for the subsequent analyses. IgG levels in sera of mice immunized with the mixture of full-length 3D7 and FC27 MSP2 were not significantly different from those in sera of mice immunized with either 3D7 or FC27 MSP2 alone (Fig. 2A). Unexpectedly, the response generated by the mixture of the two alleles displayed a degree of strain-specificity, with a significantly higher ELISA signal seen on plates coated with FC27 MSP2 than on those coated with 3D7. IgG responses elicited by the C-terminally

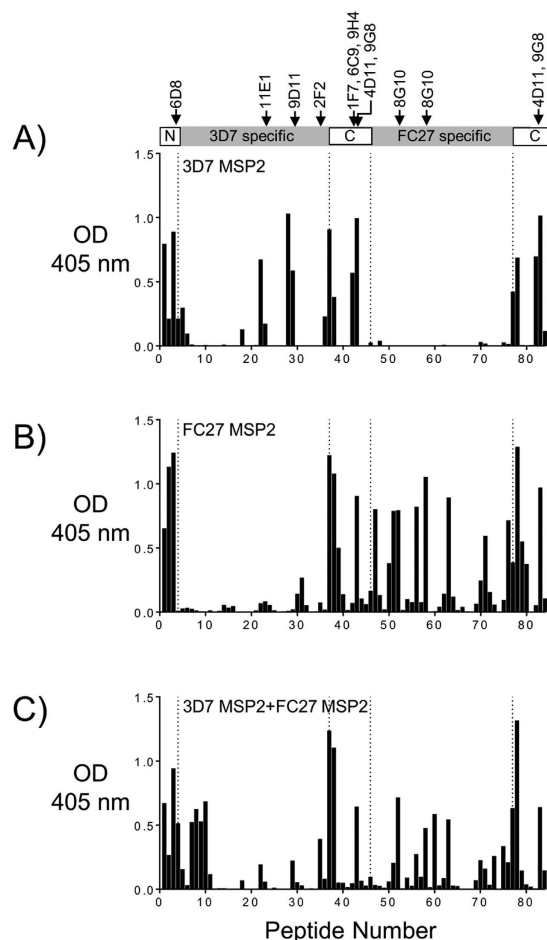


Figure 3. Epitope mapping for 3D7 and FC27 MSP2, and for the mixed immunization, by peptide array using ELISA¹⁹. Linear epitopes along 3D7 MSP2 and FC27 MSP2 were mapped using pooled sera from mice immunized with recombinant full length 3D7 or FC27 MSP2 and a mixture of 3D7 and FC27 MSP2 (A–C, respectively). The corresponding epitopes of known mouse mAbs are indicated. Antibody reactivity towards 3D7 and FC27 is uniform towards their conserved and variable regions (A,B) in contrast to immunizations using a mixture of both alleles which shows broad epitope coverage albeit lower overall reactivity (C). In all peptide arrays (including Figs 4 and 5), the small differences in the reactivity to the conserved C-terminal epitopes are due to a small difference between the linear peptide sequences extending from the central 3D7 and FC27 dimorphic regions. OD, optical density.

truncated FC27 construct NV_{FC27} were similar to those elicited by full-length FC27, whereas the N-terminal truncation, V_{FC27}C, generated a somewhat weaker response (Fig. 2B).

The majority of the chimeric MSP2 constructs induced total anti-MSP2 IgG levels similar to those induced by immunization with full-length 3D7 or FC27 MSP2 (Fig. 2C). Moreover, in contrast to the response to the mixture of FC27 and 3D7 MSP2, most chimeras generated responses that were not strain-specific. Lower total IgG levels were induced by V_{FC27}C and V_{3D7}V_{FC27}, both of which lacked the conserved NTR. However, V_{3D7}V_{FC27}C, which also lacks the NTR, induced high IgG levels. V_{3D7}V_{FC27} differs from V_{3D7}V_{FC27}C only in the absence of the conserved CTR, but generates lower IgG responses than V_{3D7}V_{FC27}C and all other constructs (Fig. 2). These results suggest that both conserved regions are important for the IgG response, albeit in a somewhat context-dependent manner.

Specificity of the response to a mixture of full-length 3D7 and FC27 MSP2. The epitope specificity of the antibody responses was determined using an overlapping peptide array that included the NTR, VR of both allelic variants, and the CTR. Epitope-specific responses for animals immunized with full-length 3D7 and FC27 (Fig. 3A,B) were in agreement with the pattern of antigenicity described previously in mice, rabbits and in humans^{18,38,39}. Robust responses were observed to conserved epitopes spanning the entire NTR, to multiple epitopes within the VRs of both MSP2 alleles, and to two distinct sequences within the CTR (Fig. 3A–C). The response to the mixture of 3D7 and FC27 MSP2 showed broadly similar specificity to those induced by the two individual immunizations (Fig. 3C). However, although responses to conserved-region epitopes were identical across the three immunization groups, the responses to VR epitopes were somewhat weaker in the 3D7 + FC27 MSP2 immunization group. This was particularly true of the 3D7 MSP2 variable region (except

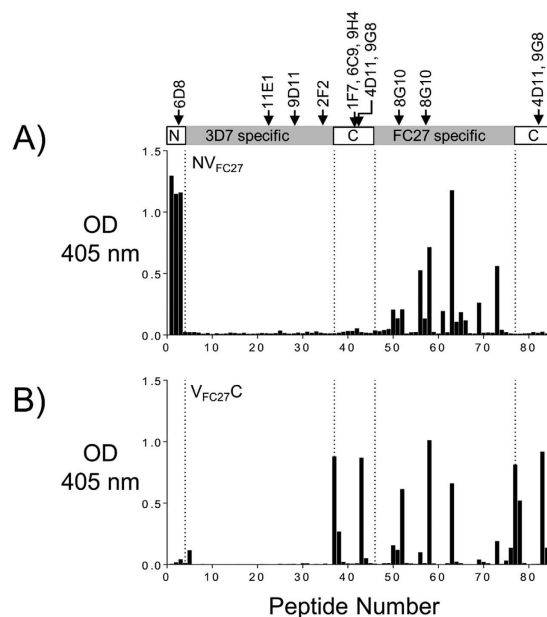


Figure 4. Epitope mapping for NV_{FC27} and $V_{FC27}C$ constructs by peptide array using ELISA¹⁹. Linear epitopes along 3D7 MSP2 and FC27 MSP2 were mapped by peptide array for sera from mice immunized with NV_{FC27} and $V_{FC27}C$ (A,B respectively) constructs. The respective regions of known mouse mAbs are indicated in the schematic. Epitope mapping assays were conducted using pooled mice sera from each of the immunization groups. The overall reactivity to the VRs is reduced when compared to the reactivity to the full-length FC27 MSP2 (Fig. 3B). OD, optical density.

in responses against 3D7-specific repeats, peptides 7–11), consistent with the strain specificity observed in this response (Fig. 3C). Unexpectedly, the response induced by the mixture retained significant strain-specificity for FC27 MSP2.

The impact of MSP2 conserved regions on the specificity of the antibody response to variable region epitopes. Removal of the conserved regions of MSP2 resulted in reduced immunogenicity of the FC27 MSP2 constructs, as measured by total IgG levels, with this reduction being more marked for the construct lacking the NTR (Fig. 2B). To examine this effect in more detail, epitope mapping of responses from mice immunized with each of the chimeras was also undertaken using the peptide array. The antibody responses evoked by FC27 MSP2 constructs lacking either the N- or C-terminal region (i.e. $V_{FC27}C$ and NV_{FC27}) were broadly similar (Fig. 4), but these responses were weak compared to the strong response to numerous VR epitopes induced by full-length FC27 MSP2 (compare peptides 47–76 in Figs 3B and 4). This result suggests that both the N- and C-terminal conserved regions are necessary to induce robust antibody responses to the VR, at least in the context of FC27 MSP2.

Chimeric MSP2 antigens elicit antibody responses to both conserved and variable region epitopes. The chimeric 3D7-FC27 MSP2, $NV_{3D7}V_{FC27}C$, elicited antibody responses against epitopes in both the NTR and CTR as well as epitopes in the VR of both alleles (Fig. 5A). In contrast to the antibodies elicited by immunization with the mixture of 3D7 and FC27 MSP2, which reacted more strongly with conserved epitopes (Fig. 3C), the strongest antibody responses to $NV_{3D7}V_{FC27}C$ were to variable region epitopes. Thus, the response induced by the chimeric antigen $NV_{3D7}V_{FC27}C$ appeared to offer an improved balance between conserved and variable region epitopes.

Removal of the N-terminal conserved region was anticipated to decrease the aggregation and fibril-formation tendencies of MSP2 (Fig. S2), although we found that it also affected the response to the FC27-specific epitopes in $V_{FC27}C$ (Fig. 4B). Accordingly, we examined the effects of removal of that region from the chimeric antigen (Fig. 5B). In striking contrast to the results with the N-terminally deleted FC27 MSP2 ($V_{FC27}C$), the antibody response to the chimeric antigen lacking the N-terminus ($V_{3D7}V_{FC27}C$) showed improved reactivity with epitopes in the VR of both 3D7 and FC27 MSP2, as well as in the CTR (compare Figs 4B and 5B). These results indicate that removal of the conserved NTR can not only minimize the tendency for protein aggregation but, in this construct, also enhance antibody responses to the VR and CTR.

To further examine the role of conserved regions in the development of a robust antibody response, we immunized animals with $V_{3D7}V_{FC27}$, which consists of only the VR regions of both 3D7 and FC27. The resulting sera showed significantly reduced reactivity to the epitopes in the VR (Figs 2C and 5C), confirming the importance of the CTR for a robust antibody response against MSP2. Finally, we examined the response to $NV_{mFC27}V_{m3D7}C$, in which the two VRs were presented in the opposite order, relative to the previous chimeras, and in which the tandem repeat regions were truncated (Fig. 1, Table S1). The observed response was similar in magnitude and

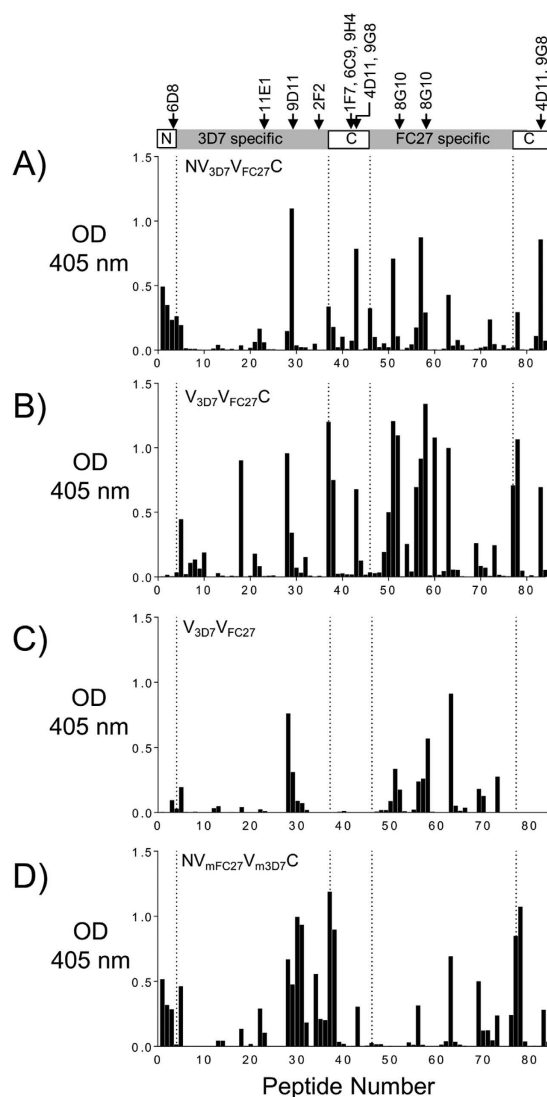


Figure 5. Epitope mapping for MSP2 chimeras by peptide array using ELISA¹⁹. Linear epitopes along 3D7 MSP2 and FC27 MSP2 were mapped for sera from mice immunized with MSP2 chimeric constructs. The corresponding epitopes of known mouse mAbs are indicated. The overall reactivity to the 3D7 and FC27 epitopes is comparable for the NV_{3D7}V_{FC27}C (A) and improved for the V_{3D7}V_{FC27}C construct (B). The reactivity is reduced for V_{3D7}V_{FC27} (C), which is lacking conserved NTR and CTR, and improved and reduced, respectively, for the 3D7 and FC27 VR regions when the GSA repeats and 32-residue repeats were removed in the chimera NV_{mFC27}V_{m3D7}C (D). OD, optical density.

fine specificity to that seen for NV_{3D7}V_{FC27}C (Figs 2C and 5D). However, truncation of one 32-residue repeat sequence in FC27 VR suppressed the immunogenicity of the remaining 32-residue repeat sequence (peptides 50–62, Fig. 5D), while removal of the complete GGSA repeat sequence of 3D7 VR improved the immunogenicity of the remaining 3D7 VR epitopes. This suggests that the number of 32-residue repeats in FC27 MSP2 predominantly affects reactivity to epitopes encoded by the 32-mer, whereas the GGSA repeat sequence in 3D7 MSP2 may compromise the immunogenicity of other 3D7 VR epitopes.

IgG subclass responses to different MSP2 constructs. Human anti-MSP2 antibodies may mediate protection by Fc receptor-mediated processes including ADCC²⁴ and opsonic phagocytosis²³, and by complement-mediated growth inhibition²². Therefore it is likely that a human MSP2-based malaria vaccine will need to induce and sustain high titres of IgG1 and/or IgG3 antibodies, which mediate these effector mechanisms, as is seen with naturally-acquired anti-MSP2 responses^{20,40,41}. IgG1 was the dominant subclass response to all MSP2 constructs, including the four chimeras studied here, in mice (Fig. 6A), but, in contrast to human IgG1, murine IgG1 effector functions are not mediated by interactions with Fc receptors or complement. Lower levels of IgG2b and IgG2c antibodies were induced by the MSP2 constructs although the highest IgG2b and IgG2c responses were observed in mice immunized with V_{3D7}V_{FC27}C (Fig. 6B,C). These murine subclasses, which share

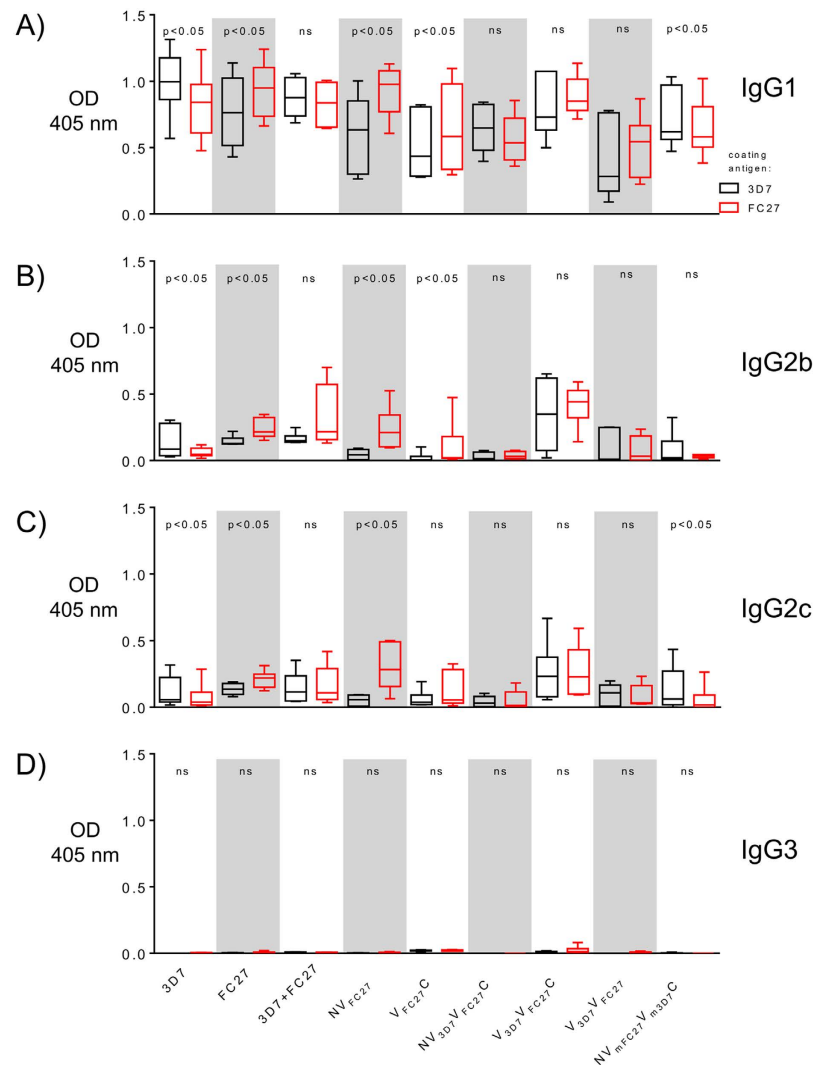


Figure 6. Anti-MSP2 IgG subtype (A–D corresponding IgG1, IgG2b, IgG2c and IgG3, respectively) responses of mice sera following immunization with recombinant 3D7, FC27 and other MSP2 constructs. Total IgG subclass levels measured by ELISA against 3D7 MSP2 (black), FC27 MSP2 (red) coated antigens using sera from mice immunized with full-length recombinant 3D7, FC27 and other MSP2 constructs. Data shown here are for sera diluted at 1:20,000. Five mice for NV_{3D7}V_{FC27}C construct and six mice for all other constructs were included in each of the immunization groups, with the mean and standard deviation indicated. IgG1 and IgG2b response was biased towards FC27 MSP2 for FC27 MSP2, NV_{FC27} and NV_{3D7}V_{FC27}C constructs, while the IgG1 response was biased towards 3D7 MSP2 and for the 3D7 MSP2 and NV_{mFC27}V_{m3D7}C. IgG1 and IgG2c response was biased towards the 3D7 MSP2 for the NV_{mFC27}V_{m3D7}C construct. There was no difference in the IgG3 subclass observed. On the other hand, more balanced IgG subclasses were observed for the mixture of 3D7 and FC27 MSP2, NV_{3D7}V_{FC27}C, V_{3D7}V_{FC27}C and V_{3D7}V_{FC27}C constructs despite the presence or absence of the conserved NTR and CTR. The difference in responses to 3D7 and FC27 allelic variants was compared using the Wilcoxon Matched Pairs Test. OD, optical density.

characteristics with human IgG3 and IgG1, respectively, are cytophilic and complement fixing⁴². IgG3 responses were negligible (Fig. 6D).

Chimeric MSP2 antigens elicit antibodies capable of strain-independent recognition of parasite extracts. An important requirement of a protective immune response against MSP2 is the ability to recognize the protein on the parasite surface. We have characterized several epitopes on recombinant MSP2 that appear to be significantly less accessible on the parasite surface¹⁹, and have shown that the masking of these epitopes is recapitulated in parasite-derived material examined by western blot¹⁹. Accordingly, we have employed this approach to test the antisera raised against each of our MSP2 constructs, using parasite strains expressing 3D7 and FC27 MSP2 (Fig. 7). The antisera raised against individual MSP2 alleles showed cross-reactivity in western blot, indicating that, despite the extensive masking of conserved region epitopes, some antibodies to the conserved regions are able to recognize MSP2 on the parasite surface. In contrast, no such cross-reactivity is observed

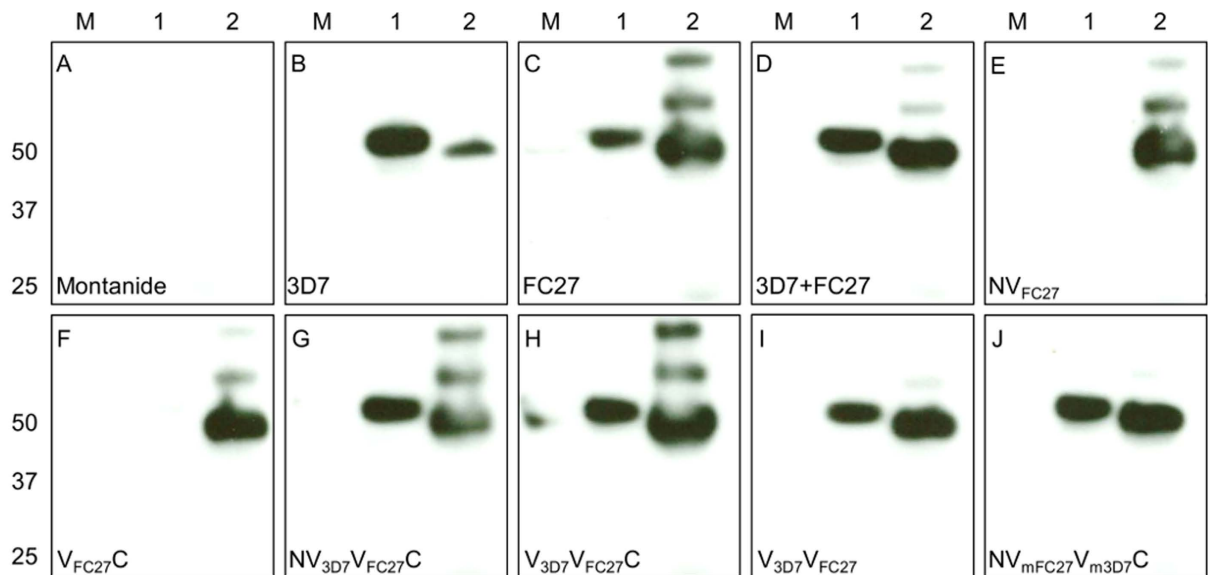


Figure 7. Western blot analysis of the Reactivity of mice sera against *P. falciparum* lysates following immunization with MSP2 constructs. Immunization with different MSP2 constructs promoted a strong antibody response that recognized the native MSP2 in 3D7 and FC27 (D10 clone) parasite extracts. Removal of either the conserved N-terminal or C-terminal regions of FC27 MSP2 abolished the cross-reactivity to the 3D7 parasite (NV_{FC27} and V_{FC27C}). The labels at the left bottom side of the gel indicate the corresponding constructs used in mice immunization experiments. M = protein marker and 1 and 2 indicated the 3D7 and FC27 (D10 clone) parasite extracts, respectively. Panels (C–H) show that antibodies cross-react with aggregated forms of FC27, albeit more weakly than with the monomeric species. (Panel I) shows that this cross-reactivity is even weaker in the absence of the conserved N- and C-terminal conserved regions. It is also suppressed in the NV_{mFC27}V_{m3D7}C chimera (panel J).

for the truncated versions of FC27 MSP2 lacking either the conserved N- or C-terminal regions, despite the presence of antibodies reactive to the remaining conserved region in each case, as judged from the peptide array data (Fig. 4) and the cross-reactivity of total IgG and IgG subtypes (Figs S4 and S5). This suggests that generating antibodies to the conserved epitopes of MSP2 on the parasite requires the presence of both conserved regions, at least in the context of FC27 MSP2 (Fig. 7C, E and F). Antibodies elicited by the mixture of 3D7 and FC27 MSP2 allelic forms, as well as by all chimeric constructs, robustly recognize parasites of both strains, confirming that either approach represents an effective strategy for the generation of strain-transcending antibody responses to native MSP2.

Discussion

The phase IIb Combination B vaccine trial strongly suggested that immune responses against MSP2 are protective^{26,27,43}. The strain-specific protection observed in that study indicated that epitopes within the central variable regions are likely to mediate such responses and that vaccine formulations will need to include multiple MSP2 alleles to afford protection against different parasite strains^{26,44}. Fortunately, epitopes in the variable region are known to be naturally immunogenic²¹ and much of the variability is dimorphic¹³, suggesting that a bivalent vaccine may be sufficient to cover the natural antigenic diversity of MSP2. In a subsequent phase I study, a mixture of two representative allelic forms of MSP2²⁴ induced antibody responses to the central variable region as well as to conserved C-terminal epitopes. Antibodies from this study were shown to inhibit *in vitro* parasite growth when acting in conjunction with complement or monocytes^{22,24}, suggesting that responses to C-terminal epitopes may also be functionally important. Our study sought to determine if vaccine delivery would be more effective using two recombinant antigens or a chimeric vaccine.

Studies of chimeric antigens based on other vaccine candidates have demonstrated that antigenic diversity can effectively be addressed using this approach^{45,46}. Both strategies of immunization (chimeric antigens and a mixture of antigens) have yielded evidence for distinct antibody specificity and, in some cases, distinct functional responses, compared with those elicited by monovalent antigens^{45–48}. Nonetheless, the relative merits of these two strategies are not well understood. The series of immunizations described in our study has been designed to assess these issues in the context of MSP2, and to evaluate the significance of the conserved NTR and CTR for the immunogenicity and aggregation propensity of MSP2.

One key advantage proposed for chimeric antigens over multivalent vaccines arises from the simpler production, formulation, quality assurance process and reduced costs likely to be associated with the development of a single chimeric construct. This is expected to be particularly significant in the context of a future multi-component malaria vaccine. On the other hand, the yields of some chimeric constructs are dramatically lower than their constituent antigens⁴⁵, presumably due to challenges associated with the correct folding of

structured antigens in the context of larger chimeras. In this respect, the highly disordered nature of MSP2, which obviates the need for folding, represents a key advantage. Indeed, the yields of all constructs examined here are comparable to those of recombinantly expressed 3D7 and FC27 MSP2. Moreover, the specificity of the antibody response across all of these constructs is consistent with our previous observations¹⁸. We observed that antibody responses to all constructs were focused on a relatively small number of epitopes corresponding to regions of MSP2 that we and others have shown to be particularly immunogenic^{18,32,38,49}. This result suggests that the conformational and antigenic properties of MSP2 are for the most part preserved in the context of these chimeras. However, we did observe some unexpected and context-dependent effects with the inclusion of the NTR and CTR on the overall immunogenicity of these chimeras and the fine specificity of the antibody response they elicit. The basis of these effects is unclear, but they may be mediated by T-cell epitopes present within the conserved regions of MSP2^{50,51}.

For the structured antigen AMA1, immunization with mixtures of diverse alleles elicited a bias towards conserved region epitopes in rabbits^{48,52}. That was rationalised in terms of the relative dilution of polymorphic epitopes, and its impact on B-cell affinity maturation against AMA1^{52,53}. The desirability of this effect will ultimately reflect a trade-off between the extent of antigenic diversity on the one hand, and the relative functional capacities of antibodies to conserved and variable epitopes on the other. In the case of MSP2, most variable region epitopes are essentially dimorphic and so are effectively covered by the two major allelic forms. Epitopes in the MSP2 variable region appear to be important targets for protective immunity, while many conserved region epitopes are masked on the parasite surface and therefore unlikely to play a protective role^{19,28}. However, the importance of conserved MSP2 epitopes to functional antibody responses is yet to be established.

MSP2 has a propensity to form fibrillar aggregates^{29,35}. While aggregated forms of some antigens are more immunogenic than their monomeric forms^{54–56}, aggregated antigens may also generate low-affinity or autoreactive antibodies^{57,58}. Our MSP2 constructs showed reduced aggregation propensity compared to the native protein, presumably a consequence of the low fibril propensity imposed by the arrangement of the MSP2 sequences in the constructs used in this work. For example, aggregation is abolished in NV_{3D7}V_{FC27}C and NV_{mFC27}V_{m3D7}C where the NTR is followed by 3D7 VR and a stretch of residues from the 3D7 VR, respectively. This supports the conclusion that residues immediately C-terminal to the conserved N-terminal region in native FC27 MSP2 contribute to its aggregation. Furthermore, it demonstrates that aggregation can be ameliorated in chimeric MSP2 antigens by appropriate construct design, and represents another advantage of the chimeric antigens over a mixture of individual antigens.

Variation in the number of tandem repeat sequences could affect their immunogenicity¹⁴. This is further confirmed in this work, where the removal of one 32-residue repeat diminished the immune response to the other 32-residue repeat in NV_{mFC27}V_{m3D7}C. Surprisingly, in the same construct the removal of 3D7 allele-specific GSA-rich sequences improved the response to the epitopes in 3D7 MSP2 dimorphic region. This suggests that the GSA repeats may modulate the immunogenicity of the epitopes in the 3D7 MSP2 dimorphic region and may explain why 3D7 MSP2 is less immunogenic than FC27 MSP2 in the mixture of the two antigens.

Anti-MSP2 antibodies associated with protection in natural infections are mainly IgG1 and IgG3 subtypes^{11,20,40,41,43}, with IgG1 predominant in children and IgG3 predominant in older malaria-exposed individuals⁴¹. These cytophilic subclass responses are likely to be important in eliciting complement- and monocyte-mediated effector responses. The IgG subtypes elicited by the MSP2 constructs tested in this work are biased mainly towards murine IgG1 and IgG2b subclasses. Murine IgG2b shares many features with human IgG3⁴². In our experiments, the chimeric construct V_{3D7}V_{FC27}C was able to induce strong antibody responses reactive to the conserved C-terminal and VR region epitopes of both 3D7 and FC27 MSP2 alleles. This construct also provided a more balanced IgG subtype response, shifting the dominance of IgG1 towards IgG2b without significantly affecting the total IgG titre. A better understanding of the mechanisms responsible for protection by antibodies targeting MSP2^{22,23} and of the IgG subclasses responsible for those effector responses will enable the design of optimal MSP2-based constructs for inclusion in a future malaria vaccine.

These results demonstrate that our chimeric approach enables both the fine epitope specificity and the distribution of IgG subclasses to be optimised. The V_{3D7}V_{FC27}C chimera was able to induce IgG levels that were the equivalent of that obtained from immunizing with co-administered 3D7 + FC27, but was able to induce stronger IgG2b and IgG2c responses. This chimera induced strong responses against the 3D7- and FC27-specific epitopes, as well as strong response to the C-terminal conserved epitopes. The V_{3D7}V_{FC27}C chimera demonstrates that such approaches can achieve responses that are equivalent to or better than those of admixed recombinant antigens, while providing potential advantages in manufacture and cost.

Methods

Expression of MSP2 chimeras and truncated FC27 MSP2 proteins. Full-length 3D7 MSP2 (JN248383) and FC27 MSP2 (JN248384) alleles (Fig. 1A) expressed in *E. coli* to GMP standards were available from a recent phase I clinical trial²⁴. Codon-optimized gene sequences for six MSP2 constructs (truncated FC27 MSP2 and chimeras) were obtained from GenScript (Fig. 1B), as follows: NV_{FC27} - FC27 MSP2 without the CTR (FC27 MSP2_{1–171}); V_{FC27}C - FC27 MSP2 without NTR (FC27 MSP2_{26–221}); NV_{3D7}V_{FC27}C - NTR, 3D7 MSP2_{1–179} followed by FC27 MSP2_{26–171} (with two 12-residue repeats) and CTR; V_{3D7}V_{FC27}C - 3D7 MSP2_{26–179} followed by the FC27 MSP2_{26–171} and CTR; V_{3D7}V_{FC27}C - VR of 3D7 MSP2_{26–179} followed by the VR of FC27 MSP2_{26–171}; NV_{mFC27}V_{m3D7}C - NTR, followed by the modified variable regions of FC27 MSP2_{64–171}, 3D7 MSP2_{58–179} and CTR. The amino acid sequences of all MSP2 constructs are shown in Table S1.

Synthetic genes were sub-cloned into a pET32a vector at KpnI and NcoI restriction sites for expression as thioredoxin-fusion proteins. The sequence of each fusion protein was confirmed by DNA sequencing. MSP2 plasmids were transformed into *E. coli* BL21 (DE3) cells and protein expression was carried out at 37 °C for 3 h in the

presence of 1 mM isopropyl β -D-1-thiogalactoside. The cells were then harvested by centrifugation at 6000 rpm for 15 min and the cell pellets were stored at -80°C until protein purification.

Protein purification. Bacterial cell pellets were resuspended in 20 mM Tris (pH 8.0) buffer containing 100 mM NaCl and 20 mM imidazole in the presence of protease inhibitors (Roche Life Sciences) and lysed by boiling for 10 min⁵⁹. The fusion proteins were purified by affinity chromatography as described earlier^{17,18}. Following TEV cleavage of the fusion proteins at 34°C overnight, the His-tagged thioredoxin was removed by affinity chromatography. The flow-through containing MSP2 was further purified by anion-exchange chromatography and reverse-phase high-pressure liquid chromatography (RP-HPLC). HPLC purification was carried out using a C18 column (250 mm \times 10 mm, ZORBAX) equilibrated with solvent A (0.1% TFA in water) and the proteins were loaded at a flow rate of 2 mL/min. The bound proteins were eluted using a linear gradient of acetonitrile (solvent B: 80% acetonitrile/0.1% TFA) over 45 min. The HPLC fractions containing protein were analysed by SDS-PAGE and LC-MS. Final yields of all purified proteins ranged from 2–10 mg/L of bacterial culture. The purified proteins were lyophilized and stored at -80°C until further use.

Aggregation propensity of MSP2 constructs. The aggregation propensities of all MSP2 constructs were characterized by size-exclusion chromatography prior to immunization experiments. Each construct was incubated at 1 mg/mL in 25 mM potassium phosphate buffer (pH 6.5) for four weeks at 4°C prior to testing for aggregation. Samples were also tested after heating at 90°C for 10 min just before loading onto the size-exclusion column. For size-exclusion chromatography a Superdex-75 column (GE) equilibrated with 25 mM of potassium phosphate buffer was used with a flow rate of 0.5 mL/min.

Protein preparation and mice immunization experiments. Full-length 3D7 and FC27 MSP2 and the six designed MSP2 constructs were reconstituted as stock solutions of 1 mg/mL in 10 mM phosphate buffered saline (PBS) and tested for the presence of endotoxin using a Limulus Amebocyte Lysate (LAL) assay kit (GenScript). The kit has a minimum detection limit of 0.005 EU/mL. In this work, the maximum acceptable endotoxin value was set to 0.1 EU/ μg protein⁶⁰. Protein stock solutions were stored at -80°C until further use.

Prior to immunizations, MSP2 stock solutions were diluted to 0.35 mg/mL in PBS, heated at 100°C for 5 min, cooled, then filtered through a $0.02\text{ }\mu\text{m}$ filter to remove MSP2 aggregates. All forms of MSP2 were formulated with Montanide ISA720 at 3:7 ratio (antigen:adjuvant) to a final concentration of 0.1 mg/mL. Female C57BL/6 mice ($n = 5$ per group for NV_{3D7}V_{FC27}C and $n = 6$ per group for other constructs) were inoculated subcutaneously with $100\text{ }\mu\text{L}$ containing $10\text{ }\mu\text{g}$ of formulated antigen at weeks 0, 4 and 8, then euthanized at week 10. The 3D7 MSP2 + FC27 MSP2 group received $10\text{ }\mu\text{g}$ of each antigen (i.e. $20\text{ }\mu\text{g}$ total protein). Sera were collected and stored at -20°C . All procedures were approved by AMREP AEC (Approval No. E/1417/2013/F) and all methods were carried out in accordance with the approved guidelines.

Antibody characterization with a peptide array. The specificity profile of antibodies induced by immunization with each MSP2 construct was characterized using a peptide array composed of 84 biotinylated 13-mer peptides, with an 8-residue overlap, covering the entire 3D7 and FC27 MSP2 sequences, as described earlier^{18,19}. Nunc MaxiSorp flat-bottom 96-well plates were coated with $1\text{ }\mu\text{g/mL}$ streptavidin in PBS and incubated overnight at 4°C . Plates were washed with PBS containing 0.05% Tween-20, blocked with 0.5% BSA at 37°C for 2 h and coated with 1:500 biotinylated peptides. Sera samples were pooled for each immunization group, diluted 1:1000 in blocking solution and added to the peptide array. Antibody detection was performed by the addition of goat anti-mouse IgG-HRP and developed using ABTS (2, 2'-azino-bis[3-ethyl benzenethiazoline-6-sulfonic acid]). The colour reaction was quenched by the addition of 1% SDS solution and the optical density was determined at 405 nm. Sera from mice inoculated with adjuvant alone were used as the background control.

Quantification of total IgG and IgG subclasses. Anti-MSP2 IgG was quantified by ELISA using Nunc MaxiSorp flat-bottom 96-well plates coated with full-length recombinant MSP2 (3D7, FC27) and the other MSP2 constructs at $1.25\text{ }\mu\text{g/mL}$ overnight at 4°C . Plates were washed with PBS containing 0.05% Tween-20 and blocked with 0.1% casein at 37°C for 2 h. Sera collected from individual mice were serially diluted to calculate total IgG endpoint titres (EPT) for the two full-length MSP2 antigens and the six designed MSP2 constructs. These EPT were compared with a single sera dilution of 1:20,000 to ensure that this sera dilution was within the linear range of the assay. EPT and the single sera dilution at 1:20,000 were highly correlated (Fig. S3), so single sera dilutions tested in duplicate were used for the subsequent analyses. Total IgG and IgG subclasses were quantified using goat anti-mouse IgG-HRP, IgG1-HRP, IgG2b-HRP, IgG2c-HRP and IgG3-HRP diluted 1:2,000 in blocking buffer. HRP activity was visualised with ABTS read at 405 nm. Sera from mice inoculated with adjuvant alone were used as the background control. A titration of pooled sera from MSP2 immunized mice was added to each plate to allow normalisation of values across different plates.

Antibody reactivity with parasite lysate. *P. falciparum* 3D7 and FC27 (D10 clone) strains were cultured in human erythrocytes at a hematocrit of 3% in RPMI media supplemented with 10% AlbuMAX II. Infected erythrocytes were harvested at the schizont stage of parasite development at 10% parasitaemia, then lysed with 0.1% saponin. Parasite lysates were fractionated by SDS-PAGE, then transferred to nitrocellulose membranes. Blots were blocked with 5% skim milk, incubated with a 1:500 dilution in 5% skim milk of pooled sera from each group of immunized mice, washed with PBS containing 0.05% Tween-20 (PBS-T), incubated with 1:5,000 goat anti-mouse IgG HRP, washed with PBS-T and developed with Amersham ECL Western Blot detection kit (GE Healthcare).

Statistical analysis. Total IgG and IgG subclasses (IgG1, IgG2b, IgG2c, IgG3) from each group were compared statistically between recombinant 3D7 and FC27 MSP2 coated ELISA plates using Wilcoxon matched pairs signed rank test. Tests were performed one-sided when there was an expectation of specificity, and two-sided otherwise. Additionally, the Kruskal-Wallis test with Dunn's post test was performed for each immunization group against recombinant 3D7, FC27 MSP2 and designed MSP2 constructs across total IgG and IgG subclasses.

References

- WHO. World malaria report. 1–242 (2014).
- Tun, K. M. *et al.* Spread of artemisinin-resistant *Plasmodium falciparum* in Myanmar: a cross-sectional survey of the K13 molecular marker. *Lancet Infect. Dis.* **15**, 415–421 (2015).
- Nkya, T. *et al.* Insecticide resistance mechanisms associated with different environments in the malaria vector *Anopheles gambiae*: a case study in Tanzania. *Malar. J.* **13**, 28 (2014).
- Barry, A. E. & Arnott, A. Strategies for designing and monitoring malaria vaccines targeting diverse antigens. *Front. Immunol.* **5**, 1–16 (2014).
- Ferreira, M. U., da Silva Nunes, M. & Wunderlich, G. Antigenic diversity and immune evasion by malaria parasites. *Clin. Diagn. Lab. Immunol.* **11**, 987–995 (2004).
- Smythe, J. A. *et al.* Identification of two integral membrane proteins of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA* **85**, 5195–5199 (1988).
- Sanders, P. R. *et al.* A set of glycosylphosphatidyl inositol-anchored membrane proteins of *Plasmodium falciparum* is refractory to genetic deletion. *Infect. Immun.* **74**, 4330–4338 (2006).
- Gilson, P. R. *et al.* Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite *Plasmodium falciparum*. *Mol. Cell. Proteomics* **5**, 1286–1299 (2006).
- 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**, 963–971 (1991).
- Smythe, J. A. *et al.* Structural diversity in the 45-kilodalton merozoite surface antigen of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **39**, 227–234 (1990).
- 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**, 4382–4388 (1995).
- Gerold, P., Schofield, L., Blackman, M. J., Holder, A. A. & Schwarz, R. T. Structural analysis of the glycosyl-phosphatidylinositol membrane anchor of the merozoite surface proteins-1 and -2 of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **75**, 131–143 (1996).
- Smythe, J. A. *et al.* Structural diversity in the *Plasmodium falciparum* merozoite surface antigen 2. *Proc. Natl. Acad. Sci. USA* **88**, 1751–1755 (1991).
- Ranford-Cartwright, L. C. *et al.* Differential antibody recognition of FC27-like *Plasmodium falciparum* merozoite surface protein MSP2 antigens which lack 12 amino acid repeats. *Parasite Immunol.* **18**, 411–420 (1996).
- Scopel, K. K. G., Silva-Nunes, M. D., Malafronte, R. S., Braga, É. M. & Ferreira, M. U. Variant-specific antibodies to merozoite surface protein 2 and clinical expression of *Plasmodium falciparum* malaria in rural amazonians. *Am. J. Trop. Med. Hyg.* **76**, 1084–1091 (2007).
- Eisen, D., Billman-Jacobe, H., Marshall, V. F., Fryauff, D. & Coppel, R. L. Temporal variation of the merozoite surface protein-2 gene of *Plasmodium falciparum*. *Infect. Immun.* **66**, 239–246 (1998).
- 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).
- MacRaild, C. A. *et al.* Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2. *PLoS ONE* **10**, e0119899 (2015).
- Adda, C. G. *et al.* Antigenic characterization of an intrinsically unstructured protein, *Plasmodium falciparum* merozoite surface protein 2. *Infect. Immun.* **80**, 4177–4185 (2012).
- Taylor, R. R., Allen, S. J., Greenwood, B. M. & Riley, E. M. IgG3 antibodies to *Plasmodium falciparum* merozoite surface protein 2 (MSP2): increasing prevalence with age and association with clinical immunity to malaria. *Am. J. Trop. Med. Hyg.* **58**, 406–413 (1998).
- Ekala, M. T., Jouin, H., Lekoulou, F., Mercereau-Puijalon, O. & Ntoumi, F. Allelic family-specific humoral responses to merozoite surface protein 2 (MSP2) in Gabonese residents with *Plasmodium falciparum* infections. *Clin. Exp. Immunol.* **129**, 326–331 (2002).
- Boyle, M. J. *et al.* Human antibodies fix complement to inhibit *Plasmodium falciparum* invasion of erythrocytes and are associated with protection against malaria. *Immunity* **42**, 580–590 (2015).
- Osier, F. *et al.* Opsonic phagocytosis of *Plasmodium falciparum* merozoites: mechanism in human immunity and a correlate of protection against malaria. *BMC Med.* **12**, 108 (2014).
- McCarthy, J. S. *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**, e24413 (2011).
- Besteiro, S., Michelin, A., Poncet, J., Dubremetz, J.-F. & Lebrun, M. Export of a *Toxoplasma gondii* rhoptry neck protein complex at the host cell membrane to form the moving junction during invasion. *PLoS Pathog.* **5**, e1000309 (2009).
- 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**, 820–827 (2002).
- Genton, B. *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**, 30–41 (2003).
- 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 (2015).
- Yang, X. *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**, 1287–1295 (2010).
- 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**, 1380–1387 (2012).
- Flück, C. *et al.* Strain-specific humoral response to a polymorphic malaria vaccine. *Infect. Immun.* **72**, 6300–6305 (2004).
- 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).
- Fowkes, F., Simpson, J. & Beeson, J. Implications of the licensure of a partially efficacious malaria vaccine on evaluating second-generation vaccines. *BMC Med.* **11**, 232 (2013).
- Tomba, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27**, 527–533 (2002).
- 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).

36. Chandrashekar, I. R., Adda, C. G., MacRaild, C. A., Anders, R. F. & Norton, R. S. Inhibition by flavonoids of amyloid-like fibril formation by *Plasmodium falciparum* merozoite surface protein 2. *Biochemistry* **49**, 5899–5908 (2010).
37. Genton, B. *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**, 2504–2511 (2000).
38. Balam, S. *et al.* *Plasmodium falciparum* merozoite surface protein 2: epitope mapping and fine specificity of human antibody response against non-polymorphic domains. *Malar. J.* **13**, 510 (2014).
39. Boyle, M. J. *et al.* Sequential processing of merozoite surface proteins during and after erythrocyte invasion by *Plasmodium falciparum*. *Infect. Immun.* **82**, 924–936 (2014).
40. Metzger, W. G. *et al.* Serum IgG3 to the *Plasmodium falciparum* merozoite surface protein 2 is strongly associated with a reduced prospective risk of malaria. *Parasite Immunol.* **25**, 307–312 (2003).
41. Duah, N. O., Miles, D. J. C., Whittle, H. C. & Conway, D. J. Acquisition of antibody isotypes against *Plasmodium falciparum* blood stage antigens in a birth cohort. *Parasite Immunol.* **32**, 125–134 (2010).
42. Tongren, J. E., Corran, P. H., Jarra, W., Langhorne, J. & Riley, E. M. Epitope-specific regulation of immunoglobulin class switching in mice immunized with malarial merozoite surface proteins. *Infect. Immun.* **73**, 8119–8129 (2005).
43. Polley, S. D. *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**, 4233–4246 (2006).
44. Weisman, S. *et al.* Antibody responses to infections with strains of *Plasmodium falciparum* expressing diverse forms of merozoite surface protein 2. *Infect. Immun.* **69**, 959–967 (2001).
45. Faber, B. W. *et al.* Diversity covering AMA1-MSP119 fusion proteins as malaria vaccines. *Infect. Immun.* **81**, 1479–1490 (2013).
46. Etemad, B. *et al.* An envelope domain III-based chimeric antigen produced in *Pichia pastoris* elicits neutralizing antibodies against all four dengue virus serotypes. *Am. J. Trop. Med. Hyg.* **79**, 353–363 (2008).
47. Faber, B. W. *et al.* Malaria vaccine-related benefits of a single protein comprising *Plasmodium falciparum* apical membrane antigen 1 domains I and II fused to a modified form of the 19-kilodalton C-terminal fragment of merozoite surface protein 1. *Infect. Immun.* **75**, 5947–5955 (2007).
48. Dutta, S. *et al.* Overcoming antigenic diversity by enhancing the immunogenicity of conserved epitopes on the malaria vaccine candidate apical membrane antigen-1. *PLoS Pathog.* **9**, e1003840 (2013).
49. Saul, A. *et al.* Human phase I vaccine trials of 3 recombinant asexual stage malaria antigens with Montanide ISA720 adjuvant. *Vaccine* **17**, 3145–3159 (1999).
50. Rzepczyk, C. M. *et al.* Amino acid sequences recognized by T cells: studies on a merozoite surface antigen from the FCQ-27/PNG isolate of *Plasmodium falciparum*. *Immunol. Lett.* **25**, 155–163 (1990).
51. Rzepczyk, C. M. *et al.* Comparative study of the T cell response to two allelic forms of a malarial vaccine candidate protein. *J. Immunol.* **148**, 1197–1204 (1992).
52. Kusi, K. A., Faber, B. W., Thomas, A. W. & Remarque, E. J. Humoral immune response to mixed PfAMA1 alleles; multivalent PfAMA1 vaccines induce broad specificity. *PLoS ONE* **4**, e8110 (2009).
53. Chaudhury, S., Reifman, J. & Wallqvist, A. Simulation of B cell affinity maturation explains enhanced antibody cross-reactivity induced by the polyvalent malaria vaccine AMA1. *J. Immunol.* **193**, 2073–2086 (2014).
54. Denis, J. *et al.* Immunogenicity of papaya mosaic virus-like particles fused to a hepatitis C virus epitope: Evidence for the critical function of multimerization. *Virology* **363**, 59–68 (2007).
55. Qian, F. *et al.* Immunogenicity of self-associated aggregates and chemically cross-linked conjugates of the 42 kDa *Plasmodium falciparum* merozoite surface protein-1. *PLoS ONE* **7**, e36996 (2012).
56. Rudra, J. S., Tripathi, P. K., Hildeman, D. A., Jung, J. P. & Collier, J. H. Immune responses to coiled coil supramolecular biomaterials. *Biomaterials* **31**, 8475–8483 (2010).
57. Kessler, M., Goldsmith, D. & Schellekens, H. Immunogenicity of biopharmaceuticals. *Nephrol. Dial. Transplant.* **21**, v9–v12 (2006).
58. Bachmann, M. F. & Zinkernagel, R. M. Neutralizing antiviral B cell responses. *Annu. Rev. Immunol.* **15**, 235–270 (1997).
59. Kalthoff, C. A novel strategy for the purification of recombinantly expressed unstructured protein domains. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* **786**, 247–254 (2003).
60. Brito, L. A. & Singh, M. Acceptable levels of endotoxin in vaccine formulations during preclinical research. *J. Pharm. Sci.* **100**, 34–37 (2011).

Acknowledgements

This work was partially supported by the Indo-Australian Biotechnology fund (BF050053) and the National Health and Medical Research Council of Australia (APP1042520). RSN, JSR and JGB acknowledge fellowship support from the National Health and Medical Research Council of Australia.

Author Contributions

B.K., C.A.M., R.F.A., J.S.R. and R.S.N. designed the experiments; B.K. and D.A. performed the experiments; B.K., D.A., C.A.M., R.A.V.M., R.F.A., J.S.R. and R.S.N. analysed the data; B.K., C.A.M., R.A.V.M., R.F.A., J.S.R. and R.S.N. wrote the manuscript with additional input from J.G.B.

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: Krishnarajuna, B. *et al.* Strain-transcending immune response generated by chimeras of the malaria vaccine candidate merozoite surface protein 2. *Sci. Rep.* **6**, 20613; doi: 10.1038/srep20613 (2016).



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

Strain-transcending immune response generated by chimeras of the malaria vaccine candidate merozoite surface protein 2

Bankala Krishnarjuna¹⁺, Dean Andrew²⁺, Christopher A. MacRaid¹, Rodrigo A. V. Morales¹,
James G. Beeson^{2,3,4}, Robin F. Anders⁵, Jack S. Richards^{2,3,4*} and Raymond S. Norton^{1*}

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

²Centre for Biomedical Research, Burnet Institute, Melbourne, Victoria 3004, Australia

³Department of Microbiology, Monash University, Clayton 3800, Australia

⁴Department of Medicine, University of Melbourne, Parkville, 3052, Australia

⁵Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe
University, Melbourne, Victoria 3086, Australia

⁺These authors contributed equally.

*Corresponding authors Raymond S. Norton [REDACTED] Jack S. Richards
[REDACTED]

Table S1. Amino acid sequences for the truncated FC27 MSP2 and different MSP2 chimeras. The sequences are colour-coded to match those in Fig. 1B and the repeat sequences are underlined. The N-terminal conserved region is fully conserved across all strains of malaria, while the C-terminal conserved region is conserved except for a single point mutation (Asn to Ser, bold and underlined in the sequence of V_{FC27C}), as documented by Fenton et al. and Smythe et al. (9, 10).

MSP2 constructs	Amino acid sequences
NV _{FC27}	<u>GIKNE</u> <u>SKYSNTFINNAYNMSIRRS</u> <u>MANEGSNTNSV</u> <u>GANAPNADT</u> <u>IASGSQRSTNSASTSTTNNGESQTTTPTAADT</u> <u>IASGSQRSTNSAST</u> <u>STTNNGESQTTTPTAADTPTATESNSPSP</u> <u>PITTT</u> <u>TESSSSGNAPNKT</u> <u>DGKGEESEKQNELNE</u> <u>STEEGPRAPQEPQTAENENPA</u>
V _{FC27C}	<u>GNEGSNTNSV</u> <u>GANAPNADT</u> <u>IASGSQRSTNSASTSTTNNGESQTT</u> <u>TPTAADT</u> <u>IASGSQRSTNSASTSTTNNGESQTTTPTAADTPTATES</u> <u>NSPSP</u> <u>PITTT</u> <u>TESSSSGNAPNKT</u> <u>DGKGEESEKQNELNE</u> <u>STEEGPRAP</u> <u>QEPQTAENENPA</u> <u>APENKGTGQHGHMHGSRNNHPQNTSDSQKE</u> <u>CTDGNKENC</u> <u>GAATSLLS</u> <u>SNSS</u>
NV _{3D7} V _{FC27C}	<u>GIKNE</u> <u>SKYSNTFINNAYNMSIRRS</u> <u>MAESKPSTGAGGSAGGSAGG</u> <u>SAGGSAGGSAGGSAGSGDGNGADAEGSSSTPATTTTTTKTTTTTT</u> <u>TTND</u> <u>AEASTST</u> <u>SSENPNHKNAETNPKGKGEVQEPNQANKETQN</u> <u>NSNVQQDSQTKSNVPPTQDADTKSPTAQPEQAENSAPTAEQTES</u> <u>PELQS</u> <u>NEGSNTNSV</u> <u>GANAPNADT</u> <u>IASGSQRSTNSASTSTTNNGES</u> <u>QTTTPTAADT</u> <u>IASGSQRSTNSASTSTTNNGESQTTTPTAADTPTA</u> <u>TESNSPSP</u> <u>PITTT</u> <u>TESNSPSP</u> <u>PITTT</u> <u>TESSSSGNAPNKT</u> <u>DGKGEESEKQ</u> <u>NELNE</u> <u>STEEGPRAPQEPQTAENENPA</u> <u>APENKGTGQHGHMHGSR</u> <u>NNHPQNTSDSQKE</u> <u>CTDGNKENC</u> <u>GAATSLLS</u> <u>SNSS</u>
V _{3D7} V _{FC27C}	<u>GESKPSTGAGGSAGGSAGGSAGGSAGGSAGGSAGSGDGNGAD</u> <u>AEGSSSTPATTTTTTKTTTTTTTTND</u> <u>AEASTST</u> <u>SSENPNHKNAETN</u> <u>PKGKGEVQEPNQANKETQNNSNVQQDSQTKSNVPPTQDADTKS</u> <u>PTAQPEQAENSAPTAEQTESPELQS</u> <u>NEGSNTNSV</u> <u>GANAPNADT</u> <u>I</u> <u>ASGSQRSTNSASTSTTNNGESQTTTPTAADT</u> <u>IASGSQRSTNSAST</u> <u>STTNNGESQTTTPTAADTPTATESNSPSP</u> <u>PITTT</u> <u>TESSSSGNAPNKT</u> <u>DGKGEESEKQNELNE</u> <u>STEEGPRAPQEPQTAENENPA</u> <u>APENKGTG</u> <u>QHGHMHGSRNNHPQNTSDSQKE</u> <u>CTDGNKENC</u> <u>GAATSLLS</u> <u>SNSS</u>
V _{3D7} V _{FC27}	<u>GESKPSTGAGGSAGGSAGGSAGGSAGGSAGGSAGSGDGNGAD</u> <u>AEGSSSTPATTTTTTKTTTTTTTTND</u> <u>AEASTST</u> <u>SSENPNHKNAETN</u> <u>PKGKGEVQEPNQANKETQNNSNVQQDSQTKSNVPPTQDADTKS</u> <u>PTAQPEQAENSAPTAEQTESPELQS</u> <u>NEGSNTNSV</u> <u>GANAPNADT</u> <u>I</u> <u>ASGSQRSTNSASTSTTNNGESQTTTPTAADT</u> <u>IASGSQRSTNSAST</u> <u>STTNNGESQTTTPTAADTPTATESNSPSP</u> <u>PITTT</u> <u>TESSSSGNAPNKT</u> <u>DGKGEESEKQNELNE</u> <u>STEEGPRAPQEPQTAENENPA</u>
NV _{mFC27} V _{m3D7C}	<u>GIKNE</u> <u>SKYSNTFINNAYNMSIRRS</u> <u>MAESKPSTGAESQTTTPTAAD</u> <u>TIASGSQRSTNSASTSTTNNGESQTTTPTAADTPTATESNSPSP</u> <u>PIT</u> <u>TTESSSSGNAPNKT</u> <u>DGKGEESEKQNELNE</u> <u>STEEGPRAPQEPQTAE</u> <u>NENPA</u> <u>GSGDGNGADAEGSSSTPATTTTTTKTTTTTTTTND</u> <u>AEAST</u> <u>STSSENPNHKNAETNPKGKGEVQEPNQANKETQNNSNVQQDSQ</u> <u>TKSNVPPTQDADTKSPTAQPEQAENSAPTAEQTESPELQS</u> <u>APEN</u>

	KGTGQHGHMHGSRNNHPQNTSDSQKECTDGNKENC GAATSL LNSS
--	--

Table S2. Molecular masses of different MSP2 constructs and estimated endotoxin levels. The molecular masses were determined using LC-MS and endotoxin levels were estimated using an endotoxin assay kit from GenScript.

MSP2 constructs	Expected mass (kDa)	Observed mass (kDa)	Endotoxin levels (EU/ μ g)
3D7 MSP2	24189.0	24195.0	0.02
FC27 MSP2	23842.0	23849.0	0.02
NV _{FC27}	17644.2	17654.7	0.08
V _{FC27} C	19926.3	19933.6	0.04
NV _{3D7} V _{FC27} C	38999.0	39048.2	0.01
V _{3D7} V _{FC27} C	34851.4	34908.0	0.02
V _{3D7} V _{FC27}	29334.0	29342.0	0.02
NV _{mFC27} V _{m3D7} C	32515.6	32530.2	0.03

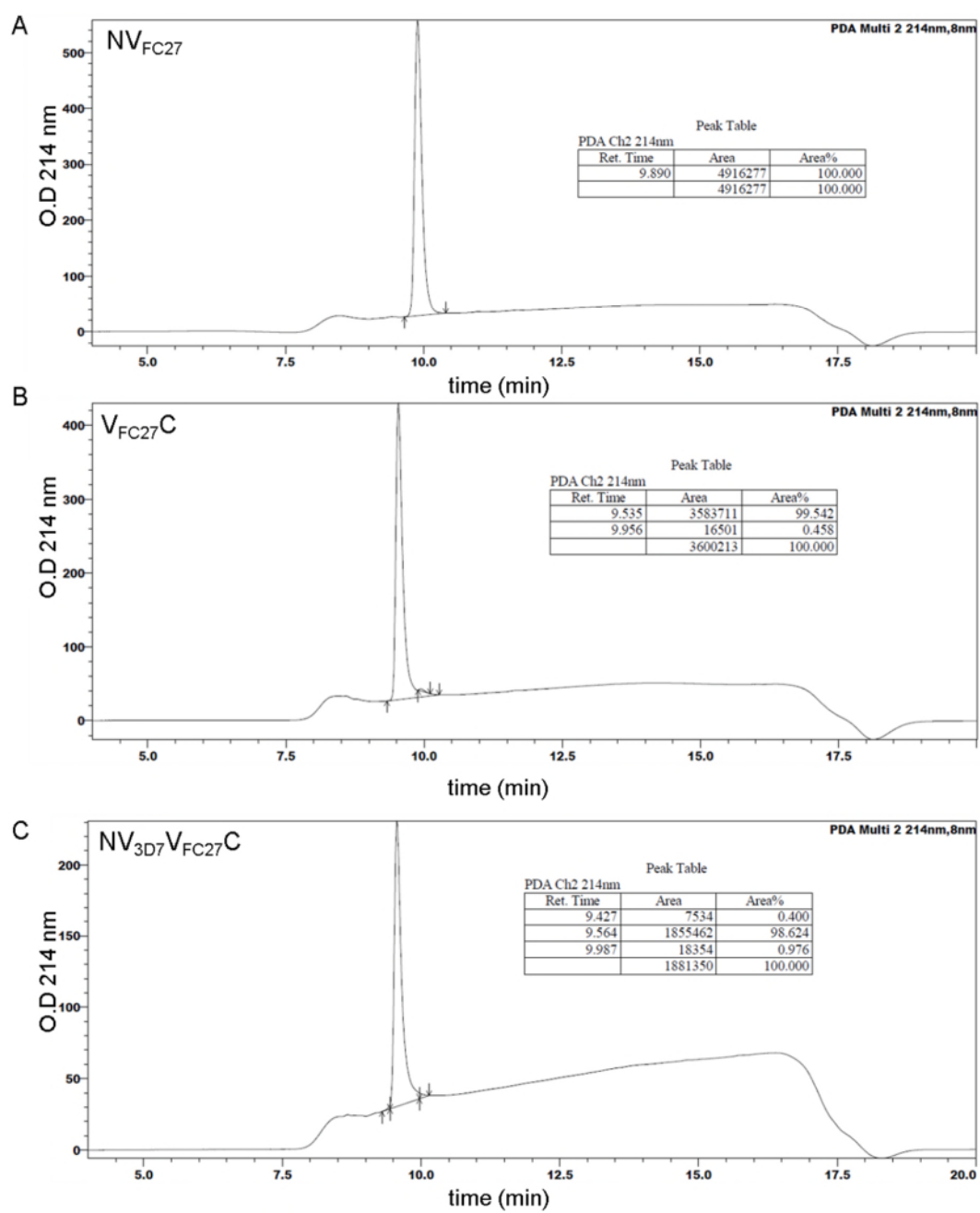


Figure S1 (continued)

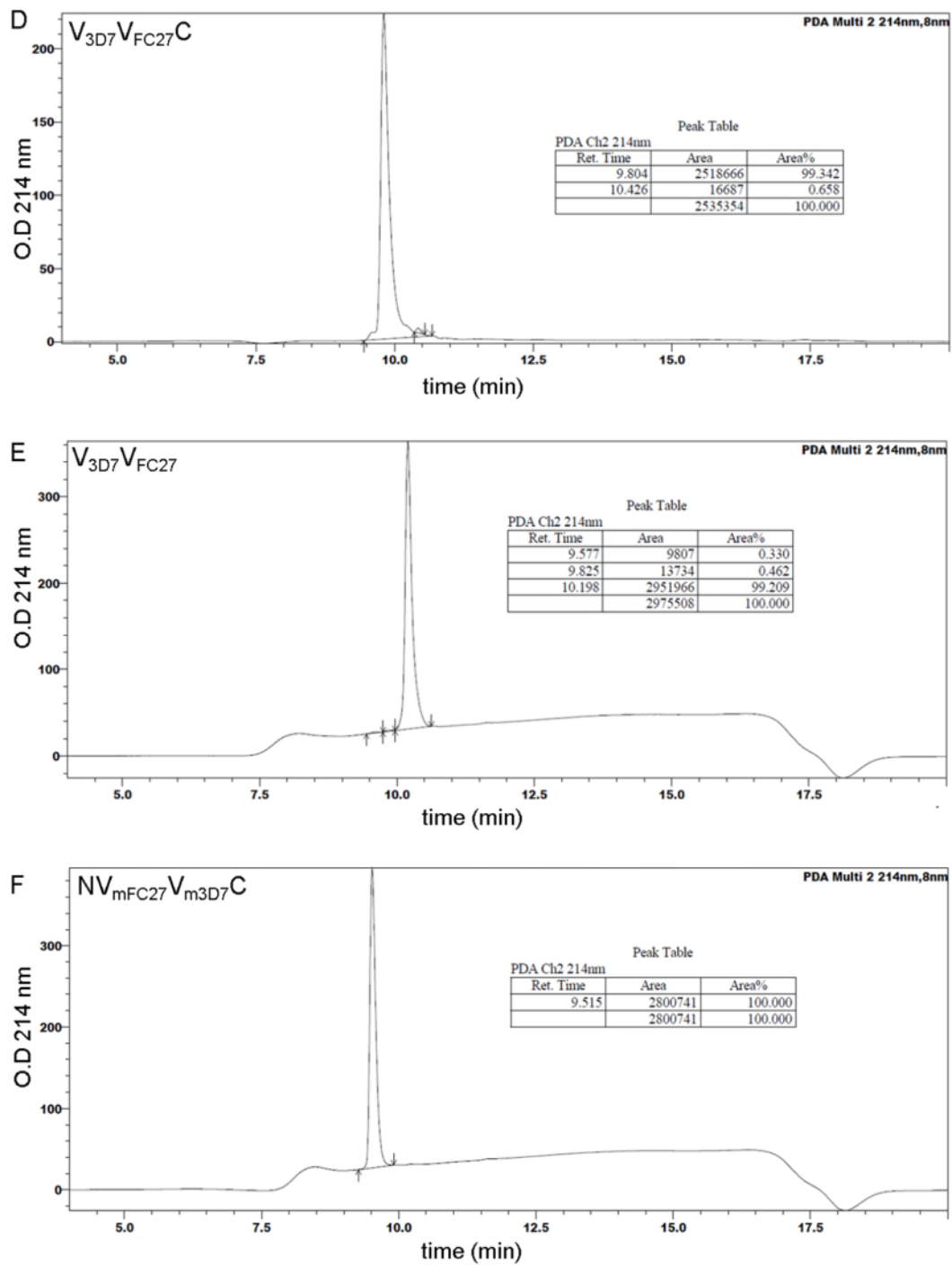


Figure S1. Purity of the MSP2 variants tested by analytical HPLC. The table in each panel shows the purity level of the corresponding MSP2 construct. Samples were assayed on a C4 analytical HPLC column using a gradient of 0-80% acetonitrile containing 0.1% formic acid.

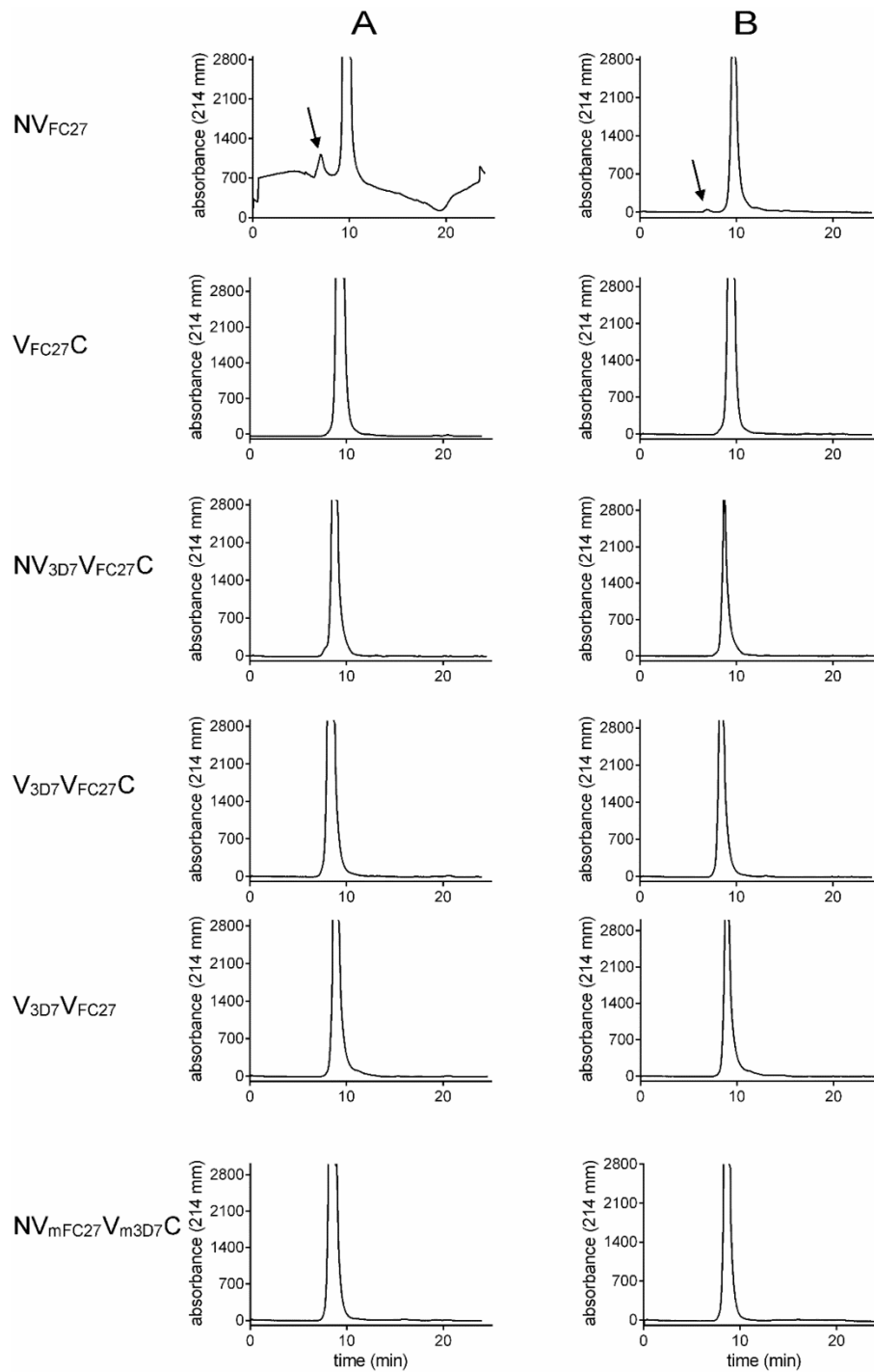


Figure S2. Aggregation propensity of MSP2 constructs (Fig. 1B) as analysed by size-exclusion chromatography. The left panel (A) represents chromatograms of the samples loaded directly onto the size-exclusion column, and the right panel (B) the samples that were preheated before loading onto the size-exclusion column. The arrows indicate aggregated protein in the NV_{FC27} construct.

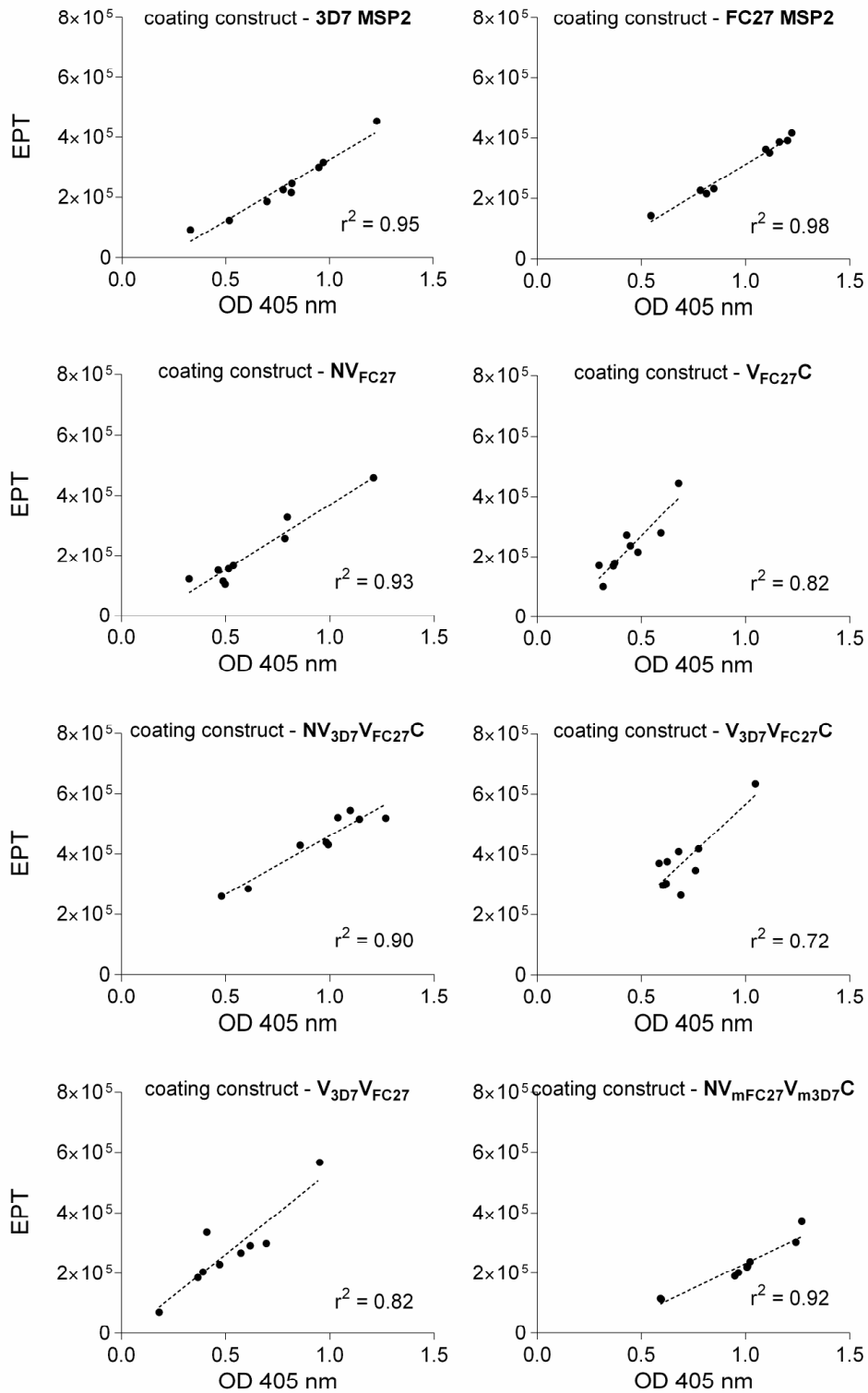


Figure S3. Correlations between the endpoint titres (EPT) and single sera dilutions for total IgG responses. The EPT and optical density (OD) were determined using sera from individual mice in each of the immunisation groups against each of the coating antigens (as indicated). The median EPT and median OD for each of these immunisation groups were plotted and the coefficient of determination calculated.

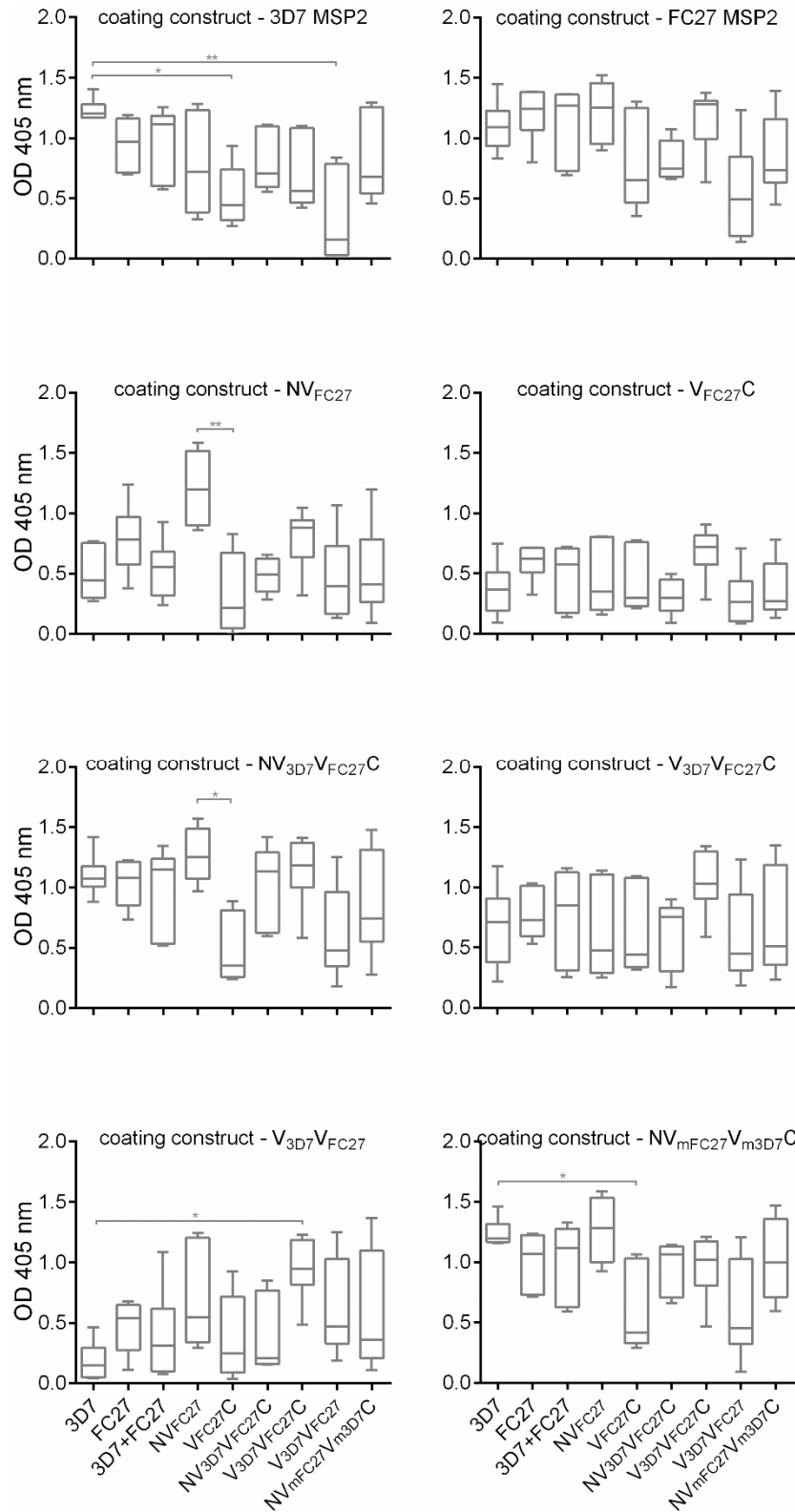


Figure S4. Cross-reactivity of the total IgG raised against different MSP2 constructs (coating construct) with recombinant 3D7 MSP2, FC27 MSP2 and other MSP2 constructs (Fig. 1B). The differences between the groups that are statistically significant (p value ≤ 0.05) are represented by asterisks (*). OD, optical density.

A

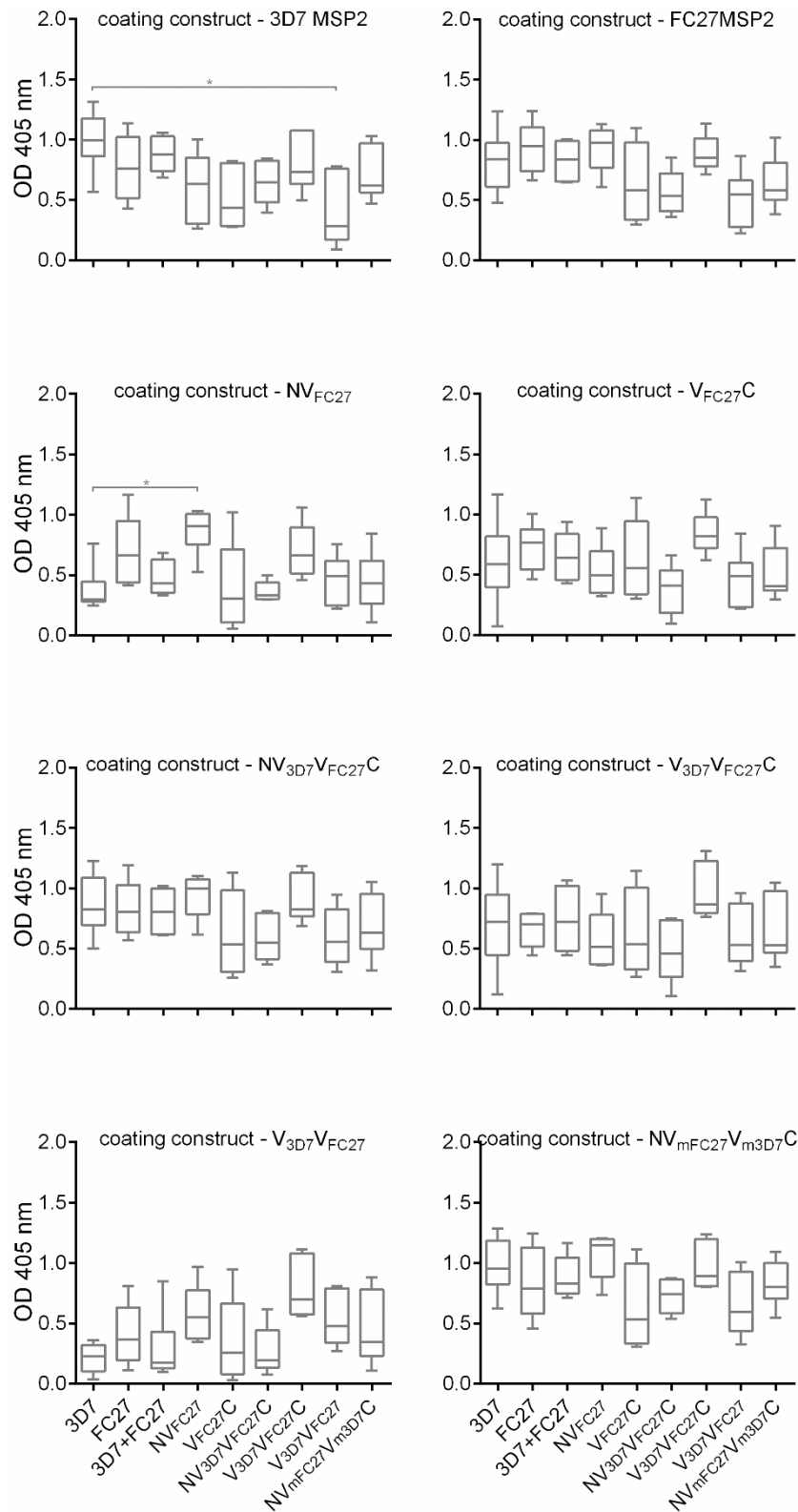


Figure S5 (continued)

B

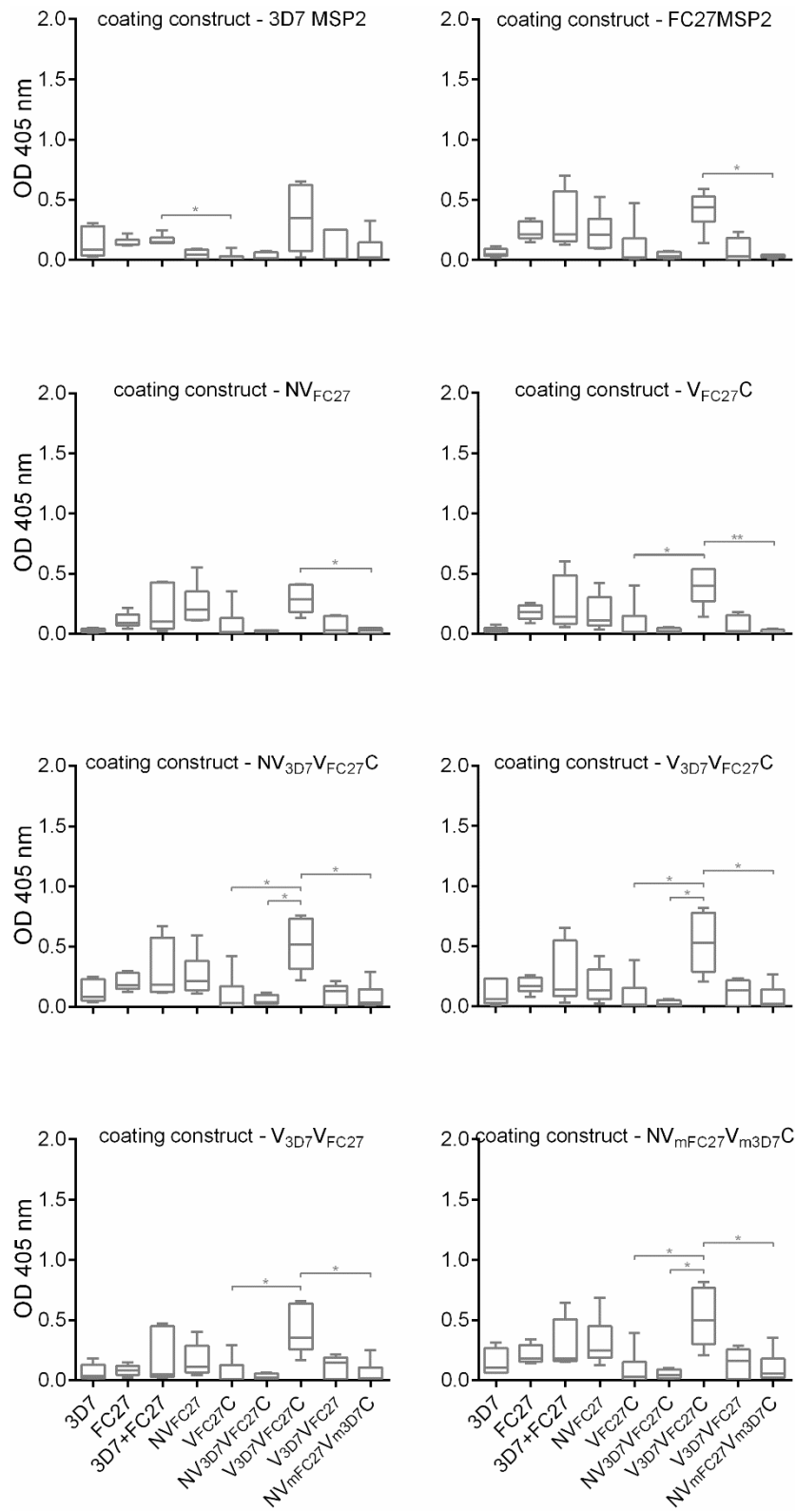


Figure S5 (continued)

C

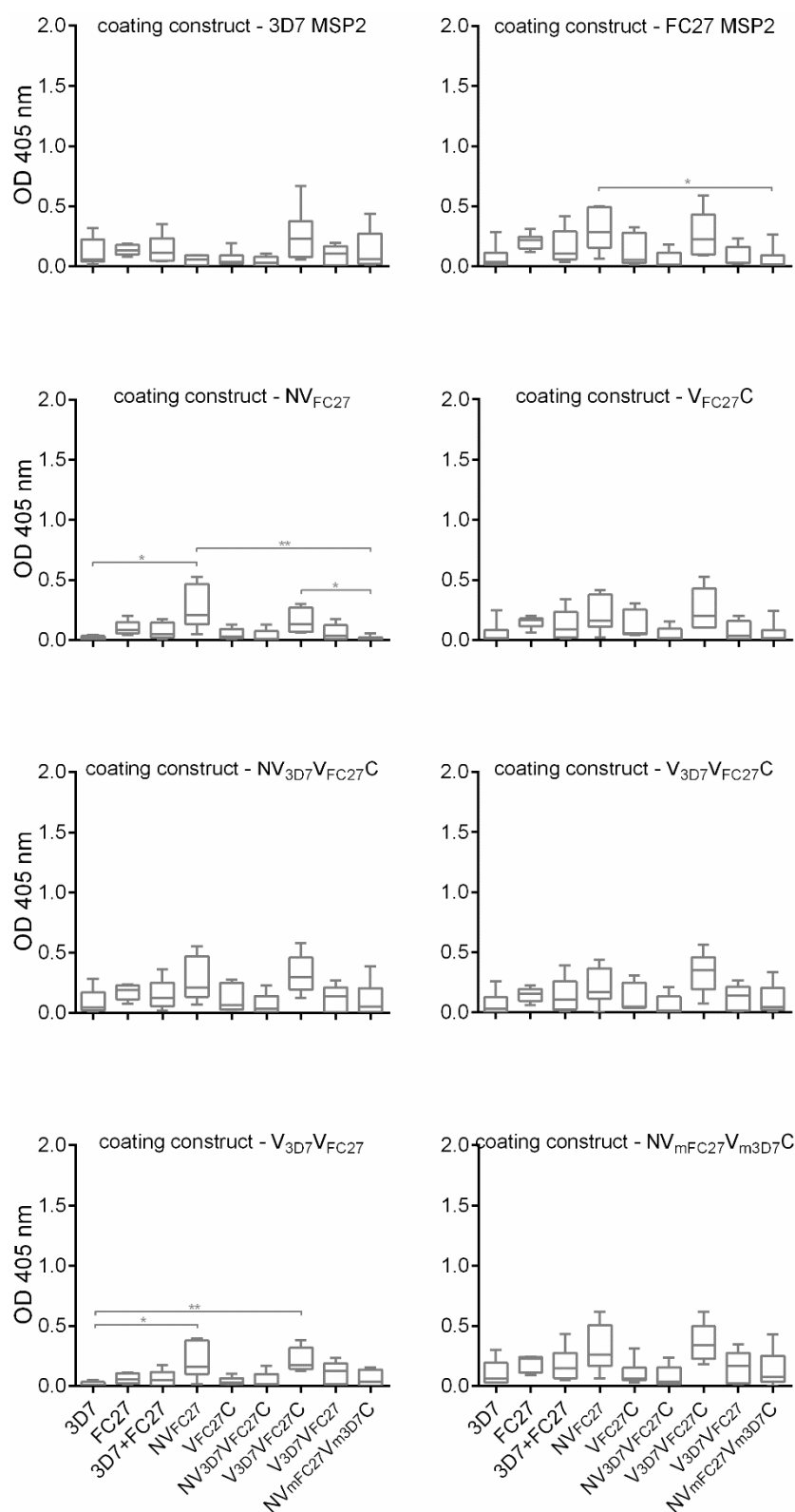


Figure S5 (continued)

D

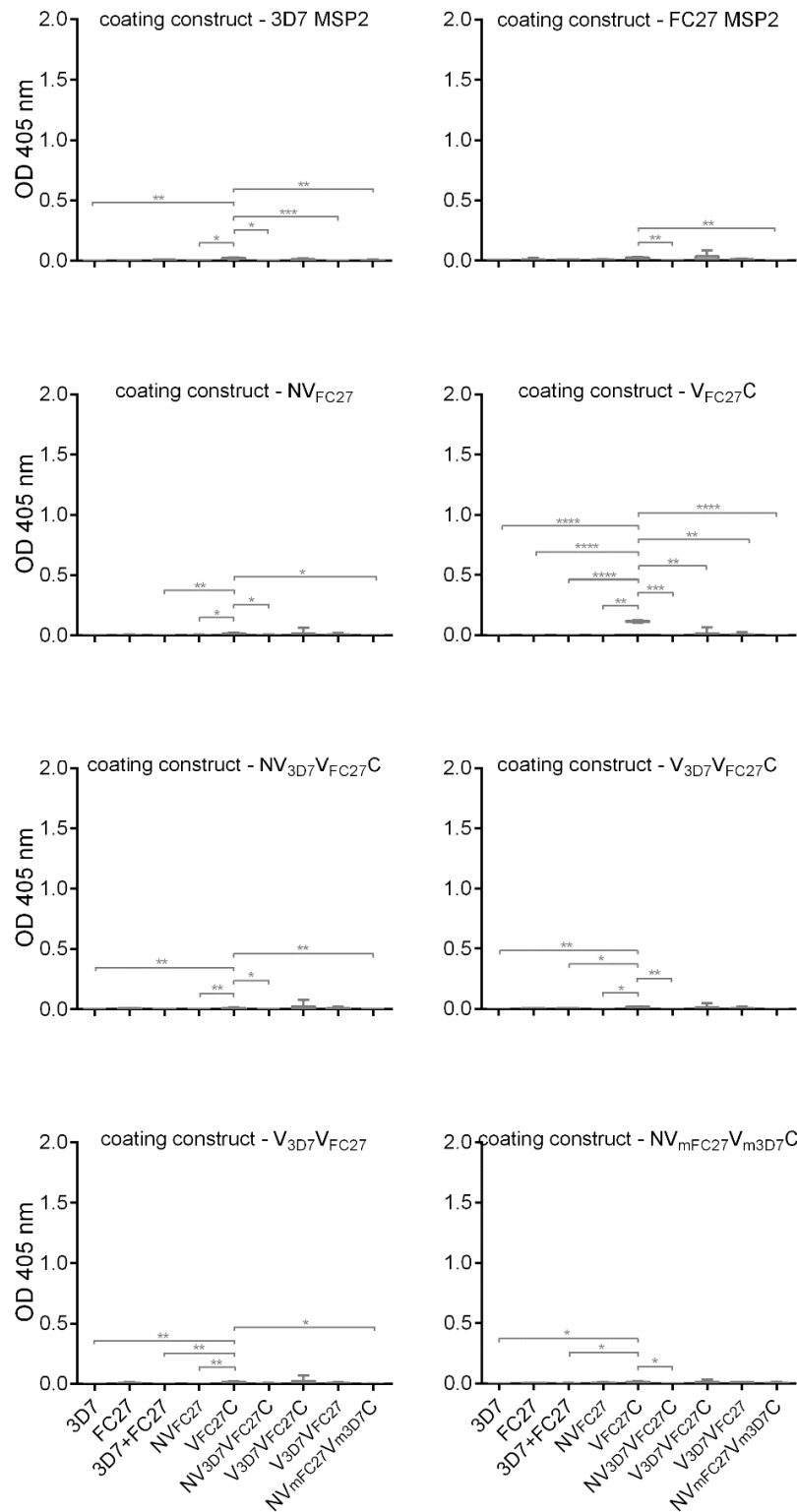


Figure S5. Cross-reactivity of the (A) IgG1, (B) IgG2b, (C) IgGc and (D) IgG3 raised against different MSP2 antigens (coating construct) with recombinant 3D7 MSP2, FC27 MSP2 and other MSP2 constructs. The differences between the groups that are statistically significant are represented as * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$ and **** $p \leq 0.0001$. OD, optical density.

CHAPTER 6

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

6.1 Declaration of Thesis Chapter 6

Declaration by candidate

In the case of **Chapter 6**, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Prepared 3D7 MSP2, performed and analysed LC-MS of the proteins used in this study and provided intellectual input in writing the manuscript	15

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Rodrigo A. V. Morales	Designed and performed the experiments, analysed the data and wrote the manuscript	
Christopher A. MacRaild	Designed and performed the experiments, analysed the data and wrote the manuscript	
Jeffrey Seow	Performed the experiments	15
Nyssa Drinkwater	Performed the experiments	
Romain Rouet	Performed the experiments	
Robin F. Anders	Designed the experiments and Intellectual input	
Daniel Christ	Designed the experiments and Intellectual input	
Sheena McGowan	Designed the experiments and Intellectual input	
Raymond S. Norton	Designed the experiments, analysed the data and wrote the manuscript	

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

**Candidate's
Signature**

	Date 30/03/2016
---	---------------------------

**Main
Supervisor's
Signature**

	Date 30/03/2016
---	---------------------------

6.2 Introduction

Chapter 5 revealed that a strain-transcending immune response to the MSP2 can be generated by using a chimera containing epitopes covering both alleles of MSP2 (3D7 and FC27). Furthermore, the removal of conserved regions was found to affect the immune response to variable region epitopes, presumably owing to a lack of T-cell epitopes present in the conserved regions and possibly also to conformational changes in the variable regions upon removal of conserved regions. Chapter 4 described the effect of conformational flexibility on the antigenicity of MSP2 and suggested that the antigenicity can be improved by increasing the conformational order of MSP2 epitopes. It is possible to improve the conformational order of MSP2 by introducing modified epitopes that mimic the structures of antibody-bound MSP2 epitopes. The modified epitopes could be incorporated into the chimeras without affecting their antigenicity, to achieve an improved and strain-transcending immune response to MSP2.

In this chapter, the crystal structure of 6D8 was solved in complex with its MSP2 epitope. 6D8 is a monoclonal antibody recognizing an epitope within the conserved *N*-terminus of MSP2. The structure of the 6D8-bound epitope is incompatible with the structure of the lipid-bound conformation of that region. This suggests that when the *N*-terminal region of MSP2 is interacting with the membrane it is not accessible to antibodies raised against recombinant MSP2. On the other hand, although 6D8 binds to a highly conserved region (MSP2₁₄₋₂₂), it has different affinities for recombinant 3D7 and FC27 MSP2 as confirmed by ITC binding experiments. Binding experiments carried out with the 6D8 epitopes containing just the first five residues of the variable region of each allele (3D7 and FC27 MSP2₁₄₋₃₀) also showed similar results. This suggests that the affinity differences are due to extra interactions between 6D8 and variable region residues. While crystal structures for 6D8 scFv in complex with the extended epitopes did not explain the basis for these affinity differences, two-dimensional [¹⁵N-¹H]-TROSY experiments suggested that the affinity differences were due to the existence of transient interactions between 6D8 and the residues from the variable region that is close to the 6D8 epitope. As these results have been published in *Scientific Reports*, they are presented in the format of a published article in the immediate section below. Although my work in the 3D7 MSP2 preparation and LC-MS analysis of all the proteins formed a small part of this study, this publication was included as a chapter in my thesis to provide essential background for the subsequent chapter.

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. Morales¹, Christopher A. MacRaild¹, Jeffrey Seow¹, Bankala Krishnarjuna¹, Nyssa Drinkwater², Romain Rouet³, Robin F. Anders⁴, Daniel Christ³, Sheena McGowan² & Raymond S. Norton¹

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 genera¹, and several have been identified as potential vaccine candidates for malaria^{2–7} and other diseases^{8,9}. 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^{6,10}. The interactions of intrinsically disordered proteins with their binding partners differ from those of structured proteins in terms of specificity and kinetics¹¹, as well as the physicochemical properties of the interacting residues¹², and are functionally important for a variety of cellular processes¹³. 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^{14,15}. 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^{14,16}. Human vaccine trial subjects immunized with recombinant 3D7 MSP2

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

³Garvan Institute of Medical Research, Darlinghurst, Sydney, NSW 2010, Australia. ⁴Department 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)

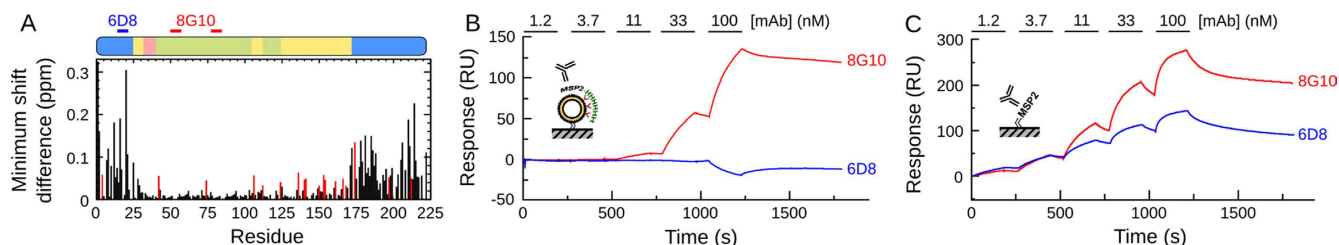


Figure 1. Membrane anchoring recapitulates the solvent exposure of the central variable region and the masking of an N-terminal epitope in recombinant MSP2. **(A)** NMR chemical shift perturbations of His-tagged MSP2 in DOGS-NiNTA/DPC micelles indicate marked differences at the conserved N and C-terminal 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, allele-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 parasitemia¹⁷. However, this vaccine preferentially targeted parasites expressing a 3D7-type MSP2 sequence, indicating that vaccine efficacy was mediated by strain-specific responses to MSP2^{18,19}. Consistent with this result, the polymorphic region appears to be immunodominant in the natural immune response to MSP2^{20,21} and some conserved region epitopes are cryptic on the parasite surface^{22,23}. 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 MSP2²⁴ but does not recognize the parasite surface,²² 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^{25,26}. These interactions, although weak, were sufficient to stabilize the 25-residue N-terminal peptide as an α -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 ¹H-¹⁵N 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-NTA) 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 MSP2²⁵ 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_a of these residues, and because the directions of the chemical shift changes

Peptide	Sequence ^Δ	K_d (nM) ^o
MSP2 ₁₁₋₂₃	FINNAYNMSIRRS	6±3
MSP2 ₁₁₋₁₈	FINNAYNM	>1000
MSP2 ₁₂₋₁₉	INNAYNMS	>1000
MSP2 ₁₃₋₂₀	NNAYNMSI	>1000
MSP2 ₁₄₋₂₁	NAYNMSIR	>1000
MSP2 ₁₅₋₂₂	AYNMSIRR	87±21
MSP2 ₁₆₋₂₃	YNMSIRRS	>1000
MSP2 ₁₅₋₂₃	AYNMSIRRS	34±12
MSP2 ₁₄₋₂₂	NAYNMSIRR	6±2
MSP2 ₁₃₋₂₂	NNAYNMSIRR	14±5
MSP2 ₁₄₋₂₃	NAYNMSIRRS	16±7

Table 1. Mapping of the optimal binding region of the mAb 6D8 on the MSP2₁₁₋₂₃ sequence using a panel of overlapping synthetic peptides. ^o K_d determined by SPR ± standard deviation of three replicate experiments. ^ΔSynthetic peptides were N-terminally acetylated and C-terminally amidated; ^{*}Broad epitope mapped previously²².

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 Ni²⁺ 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^{22,27}, 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 MSP2²⁴ but fails to bind native MSP2 on the parasite surface²², 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 MSP2²⁴, with the epitope localized to residues 11 to 23 (MSP2₁₁₋₂₃)²². This region of full-length MSP2 undergoes structural changes in the presence of lipids as demonstrated previously^{25,26} 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 MSP2₁₁₋₂₃ (Table 1) and tested against immobilized mAb 6D8 by SPR. MSP2₁₅₋₂₂ was the only 8-residue peptide of this initial panel to show detectable binding.

The affinity of MSP2₁₅₋₂₂ for mAb 6D8 was found to be 87 nM, which is more than ten-fold weaker than for MSP2₁₁₋₂₃ (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 MSP2₁₅₋₂₂ (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 MSP2₁₅₋₂₃ with MSP2₁₆₋₂₃ and MSP2₁₄₋₂₁ with MSP2₁₄₋₂₂).

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^{25,28}. This region of MSP2 encompasses the entire 6D8 epitope, suggesting that secondary structure formation could play a role in antibody binding²². The α-helical propensities of MSP2₁₁₋₂₃ and the equipotent minimal epitope, MSP2₁₄₋₂₂, were investigated by circular dichroism. The minimal peptide epitope was substantially less helical than the 13-residue peptide (MSP2₁₁₋₂₃) 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.

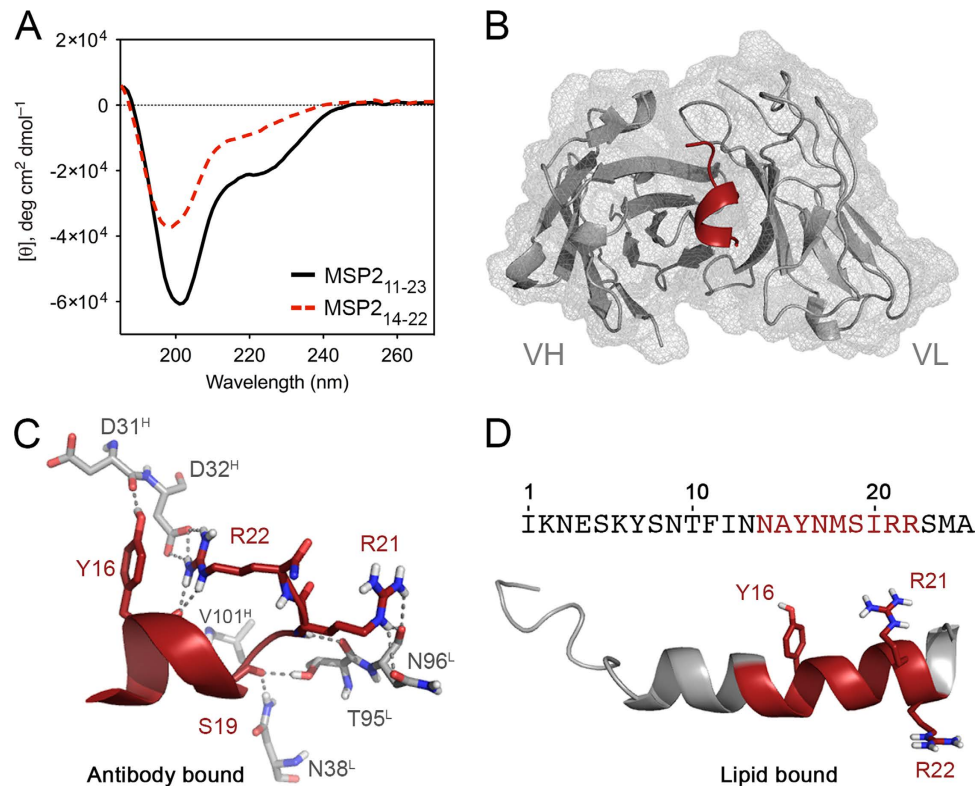


Figure 2. Crystal structure to 1.2 Å resolution of 6D8 Fv bound to the 9-mer peptide MSP2_{14–22}. (A) Helical propensity of synthetic epitope-bearing peptides MSP2_{11–23} and MSP2_{14–22}. Both peptides showed ellipticity at 190, 208 and 220 nm, consistent with some helical content in water, with the longer peptide having significantly greater tendency for helical conformation. (B) Crystal structure of MSP2_{14–22} at the paratope cleft formed by the antibody CDRs. (C) MSP2_{14–22} adopts a single turn of helix stabilized by hydrogen bonds with residues Asp31^H, Asp32^H, Val101^H and Asn38^L. In solution, Arg22 contacts residues Thr91^L and Asn92^L 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_{1–25} indicating the relative orientation of residues Tyr16, Arg21 and Arg22. The α-helical configuration of the lipid-bound MSP2_{1–25} 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_{14–22} 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 α-helical propensity of the 6D8 epitope in MSP2 and the stabilizing effect of lipid interactions²⁵, 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^L and Thr95^L (6D8 is numbered according to Kabat²⁹, 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^ε 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-π interactions with the Tyr16 side chain and an inter-molecular salt bridge with Asp32^H (Fig. 2C). The complex is further stabilized by hydrogen bonds between peptide residues Tyr16, Ser19, Arg22 and 6D8 residues Asp31^H, Val101^H and Asn96^L, 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 solution²⁶, 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)^{25,28}. 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

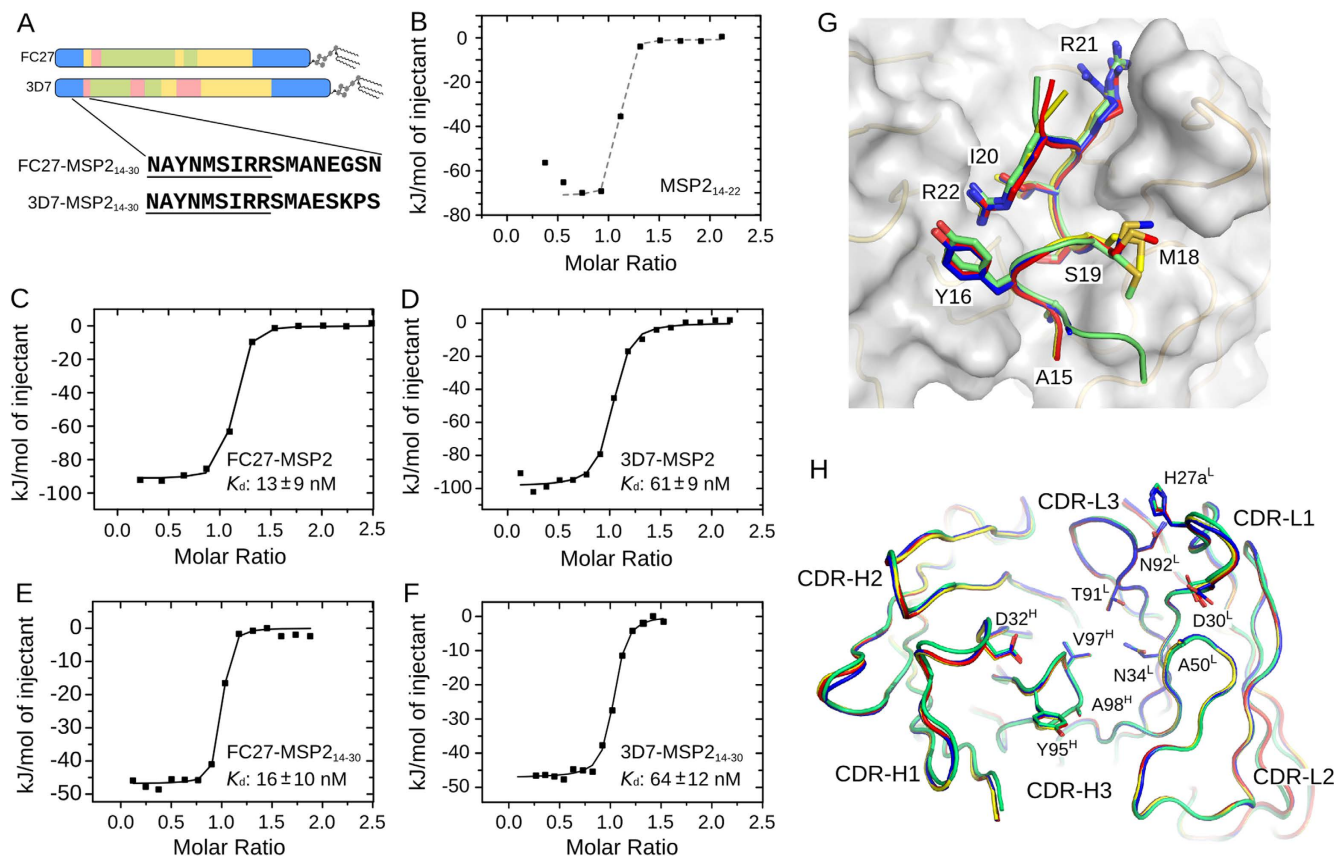


Figure 3. 6D8 displays strain-specific binding, even though it recognizes a conserved epitope sequence. (A) Schematic of the sequence features of MSP2, showing conserved regions (blue), repeats (green), allelic-specific dimorphic regions (yellow) and polymorphic regions (pink), together with the sequences of the two MSP2₁₄₋₃₀ peptides **b-e**: ITC titrations of MSP2₁₄₋₂₂ (B), full-length FC27 (C) and 3D7 (D) MSP2, FC27-MSP2₁₄₋₃₀ (E) and 3D7-MSP2₁₄₋₃₀ (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 MSP2₁₄₋₂₂ (blue), MSP2₁₁₋₂₃ (green), 3D7-MSP2₁₄₋₃₀ (yellow) and FC27-MSP2₁₄₋₃₀ (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 MSP2₁₄₋₂₂ (blue), MSP2₁₁₋₂₃ (green), 3D7-MSP2₁₄₋₃₀ (yellow) and FC27-MSP2₁₄₋₃₀ (red). CDRs and side chains making significant contacts with the antigen are shown.

6D8, including Tyr16, Met18 and Ile20. The structure of MSP2₁₄₋₂₂ 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 MSP2₁₁₋₂₃ 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, MSP2₁₄₋₂₂, 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 MSP2₁₄₋₃₀; 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).

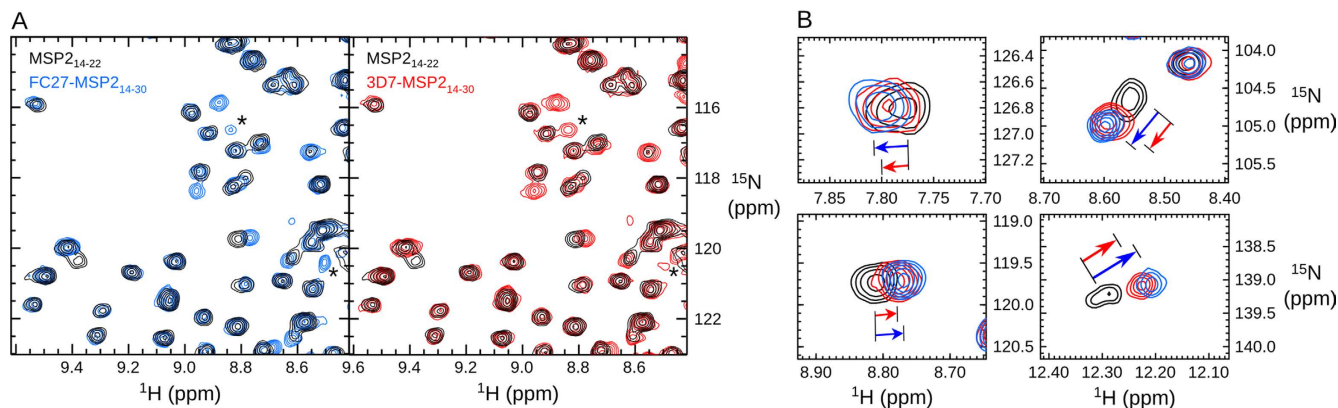


Figure 4. ^1H , ^{15}N HSQC spectra of 6D8 scFv reveal strain-specific transient interactions. (A) Comparison of the spectra of 6D8 in the MSP2₁₄₋₂₂ complex (black) with the FC27-MSP2₁₄₋₃₀ complex (blue) and with the 3D7-MSP2₁₄₋₃₀ complex (red) shows extensive spectral changes. A small number of peaks (*) 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 (MSP2₁₄₋₃₀) at resolutions of 1.4 Å and 1.6 Å, respectively. Surprisingly, the basic mode of binding observed in the MSP2₁₄₋₂₂-6D8 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-MSP2₁₄₋₃₀, 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 MSP2₁₄₋₂₂ as the minimal 6D8 epitope. The unresolved region includes all of the residues that differ between the FC27 and 3D7 MSP2₁₄₋₃₀ peptides (Fig. 3A). In contrast, the peptide is resolved until Ser28 in the 3D7-MSP2₁₄₋₃₀-6D8 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-MSP2₁₄₋₃₀ and the side chains of Asp52^H and Asn55^H of 6D8 Fv. Plainly, this interaction depends on one of the carboxylate moieties of Glu27 or Asp52^H 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_a 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^L 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 NH₂ groups of 6D8 scFv bound to the peptides 3D7-MSP2₁₄₋₃₀ and FC7-MSP2₁₄₋₃₀ (Fig. S6C), suggesting the absence of strain-specific interactions involving Asn55^H or other NH₂ groups under these conditions. Thus, we conclude that the interaction between Glu27 of 3D7-MSP2₁₄₋₃₀ and Asp52^H 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 ^1H , ^{15}N HSQC spectra, despite the tendency of scFv to dimerise at high concentrations³⁰. Comparisons of the spectrum of 6D8 scFv bound to MSP2₁₄₋₂₂ with those of 6D8 scFv bound to FC27-MSP2₁₄₋₃₀ (Fig. 4A & S7) and 3D7-MSP2₁₄₋₃₀ (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-MSP2₁₄₋₃₀ and 3D7-MSP2₁₄₋₃₀ 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_{14–22} 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, and thus the chemical shift reports on the relative populations 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^{1,31}. 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 structure⁵. 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^{11,13,32}. 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 degraded³³, unlike many merozoite surface antigens, which are released into the blood-stream³⁴. 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 interactions²⁵ or oligomerization²⁴ 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_{1–25}) binds lipid, adopting an extended helical conformation spanning at least residues 10–22²⁵. 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^{24,35} play an important role in determining the antigenic characteristics of MSP2 on the merozoite surface²².

Coupled folding and binding is often important in the functional interactions of disordered proteins³⁶ and has also been demonstrated for disordered antigens^{37–39}. Indeed, functional properties of antibodies targeting disordered epitopes have been shown to depend on the specific conformation recognized by these antibodies³⁸. 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^{40,41}. Nonetheless, antibodies able to recognize the N-terminal region of MSP2 on the merozoite surface have been characterized^{23,33}. 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^{13,42}. Transient or ‘fuzzy’ interactions allow disordered epitopes to interact with multiple partners in biological roles otherwise not accessible to structurally rigid proteins^{13,32,43,44}. 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 interactions¹³, 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^{45–47}. This hypothesis has been challenged because the majority of disordered apicomplexan proteins are not antigenic⁴⁸. 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 MSP2⁴⁹. 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 regions¹. 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^{26,50}. 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 Ni²⁺ 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_{11–23} and MSP2_{14–22} and the allele-specific N-terminal peptides 3D7 MSP2_{14–30} and FC27 MSP2_{14–30} 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. ¹³C, ¹⁵N-labelled 6D8 scFv was dissolved in 20 mM sodium citrate, pH 6.5, 7% ²H₂O, in the absence of ligand to a concentration of approximately 100 μM. Peptides were added to 20% molar excess. ¹⁵N-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 NiCl₂. ¹H-¹⁵N 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 ¹H 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).
- Tomba, 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-1 β 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. Morales¹, Christopher A. MacRaid¹, Jeffrey Seow¹, Bankala Krishnarjuna¹,
Nyssa Drinkwater², Romain Rouet³, Robin F. Anders⁴, Daniel Christ³, Sheena McGowan² and
Raymond S. Norton^{*1}

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 (Genbank 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,^{1,2} 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^{3,4} using sequence-specific primers.⁵ Primer-derived errors found at the first conserved N-terminal residues of the VH sequence were corrected based on the closely-related V_H mouse germline, IGHV1-78 (<http://www.imgt.org/>) (Fig. S8). Variable heavy (V_H) and light (V_L) 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.⁶ 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 ¹⁵N-ammonium chloride and ¹³C-glucose at 1 g/L and 4 g/L respectively for isotopic labeling.⁷ 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 MgCl₂ 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 hydroxylsuccinidyl-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 MSP2₁₁₋₂₃, MSP2₁₁₋₁₈, MSP2₁₂₋₁₉, MSP2₁₃₋₂₀, MSP2₁₄₋₂₁, MSP2₁₅₋₂₂, MSP2₁₆₋₂₃, MSP2₁₄₋₂₃, MSP2₁₄₋₂₂, MSP2₁₃₋₂₃, MSP2₁₄₋₂₃ and the allele-specific regions 3D7 MSP2₁₄₋₃₀ and FC27 MSP2₁₄₋₃₀ were prepared in-house using standard 9-fluorenylmethoxycarbonyl (Fmoc) solid-phase chemistry.⁸ 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 MSP2₁₂₋₁₉ and MSP2₁₃₋₂₀ were resuspended in solvent A (0.1% TFA in water) to a final concentration of 2 mg/mL for HPLC purification. Peptides MSP2₁₂₋₁₉ and MSP2₁₃₋₂₀ 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 facility³ 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). NiCl₂ was added to 1 mM and 7 % ²H₂O added, and this solution was used to dissolve ¹⁵N-FC27-MSP2-6His to a final concentration of 0.2 mM. ¹H-¹⁵N SOFAST HMQC spectra⁹ 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 Ni^{2+} with a 3 min injection of 50 mM 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 μM 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 MSP2₁₄₋₂₂ and MSP2₁₁₋₂₃ and the allele-specific N-terminal regions 3D7 MSP2₁₄₋₃₀ and FC27 MSP2₁₄₋₃₀. 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 μm spin filter at 4°C. Antibody concentrations were determined based on the absorbance at 280 nm ($A_{280\text{nm}}^{1\%} = 1.6$) 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 XDS¹⁰ and Aimless from the CCP4 suite.¹¹ 5% of each dataset was flagged for calculation of R_{free} ¹² 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.¹³ 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 + MSP2₁₄₋₂₂ 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 MSP2₁₄₋₂₂ peptide ligand between the two protein chains. Automated model building was performed using the program ARP/warp.¹⁴ All subsequent model building and structural validation was undertaken using Phenix^{15,16} and COOT.¹⁷ 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 + MSP2₁₁₋₂₃, 6D8 Fv + 3D7 MSP2₁₄₋₃₀ and 6D8 Fv + FC27 MSP2₁₄₋₃₀) 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.¹⁸

Circular dichroism. The secondary structure of the MSP2 peptides MSP2₁₁₋₂₃ and MSP2₁₄₋₂₂ 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 θ 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 ₁₄₋₂₂	22 % (w/v) PEG8000 0.1 M sodium acetate (pH 4.7) 0.2 M sodium chloride	1.2
6D8 Fv + 3D7 MSP2 ₁₄₋₃₀	30 % (w/v) PEG8000 0.1 M sodium acetate (pH 4.2) 0.2 M sodium chloride	1.4
6D8 Fv + FC27 MSP2 ₁₄₋₃₀	30 % (w/v) PEG8000 0.1 M sodium acetate (pH 4.4) 0.2 M sodium chloride	1.6
6D8 Fv + MSP2 ₁₁₋₂₃	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 ₁₄₋₂₂	6D8 Fv + 3D7 MSP2 ₁₄₋₃₀	6D8 Fv + FC27 MSP2 ₁₄₋₃₀	6D8 Fv + MSP2 ₁₁₋₂₃
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 _{pim} (%) ^b	6.4 (51.9)	22 (366)	21 (96)	17.0 (148.0)
PDB	4QYO	4QY8	4QXT	4R3S
Structure refinement				
Non hydrogen atoms				
Protein	1821	1829	1808	1847
Solvent (HOH)	377	369	326	370
R _{free} (%)	17.0	16.5	17.8	20.0
R _{cryst} (%)	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 ²)				
Mean protein	13.5	15.7	17.0	9.10
Mean water molecule	30.6	29.3	32.0	23.8
Molprobtity Score ^c	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 .

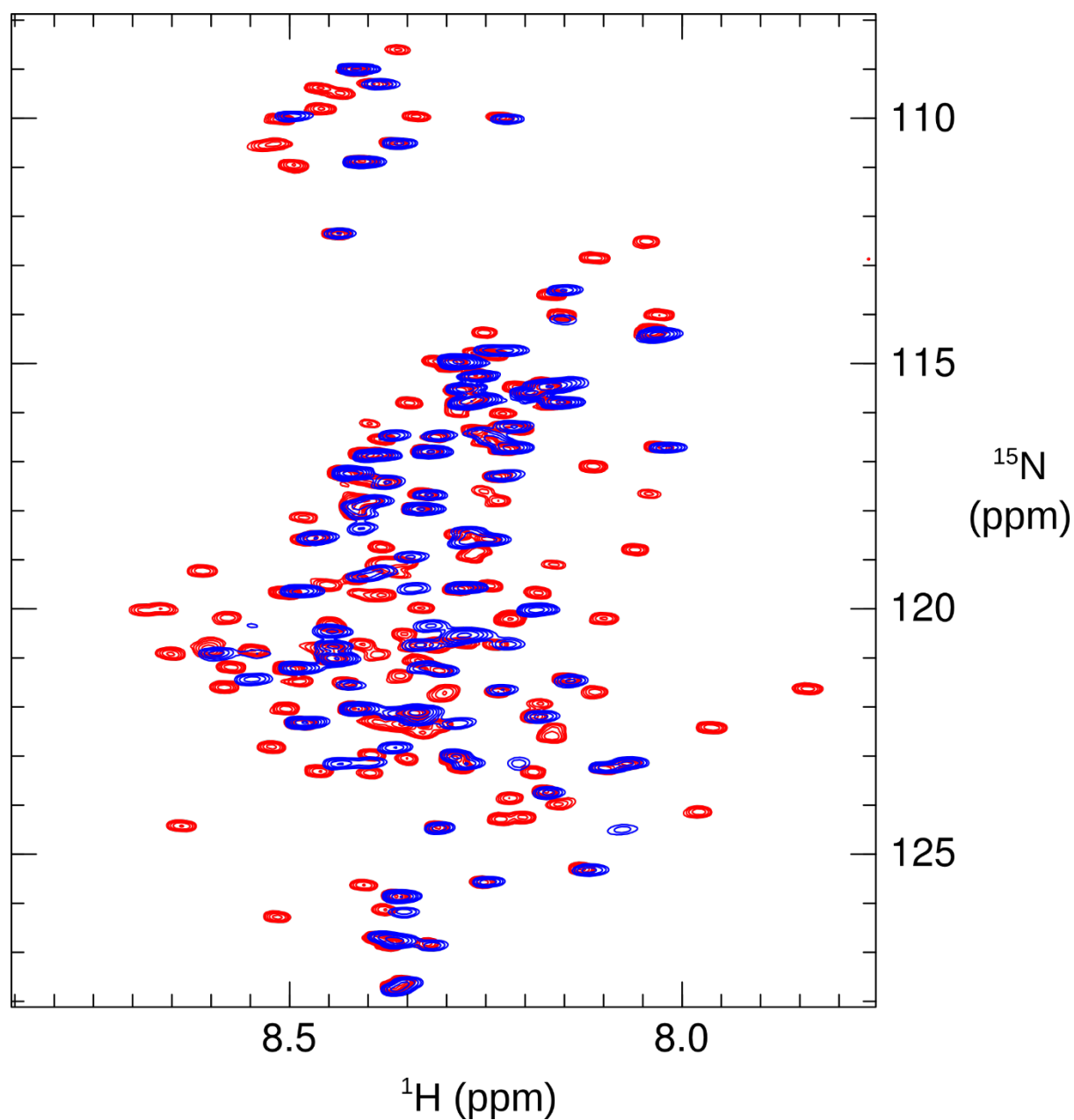


Figure S1. ^1H - ^{15}N SOFAST-HMQC spectra of FC27-MSP2-6His in the presence (blue) and absence (red) of dodecylphosphocholine (DPC) micelles containing 1 mol % of the Ni^{2+} -charged chelating lipid 1,2-di-(9Z-octadecenoyl)-*sn*-glycero-3-[(N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl (DOGS-NTA).

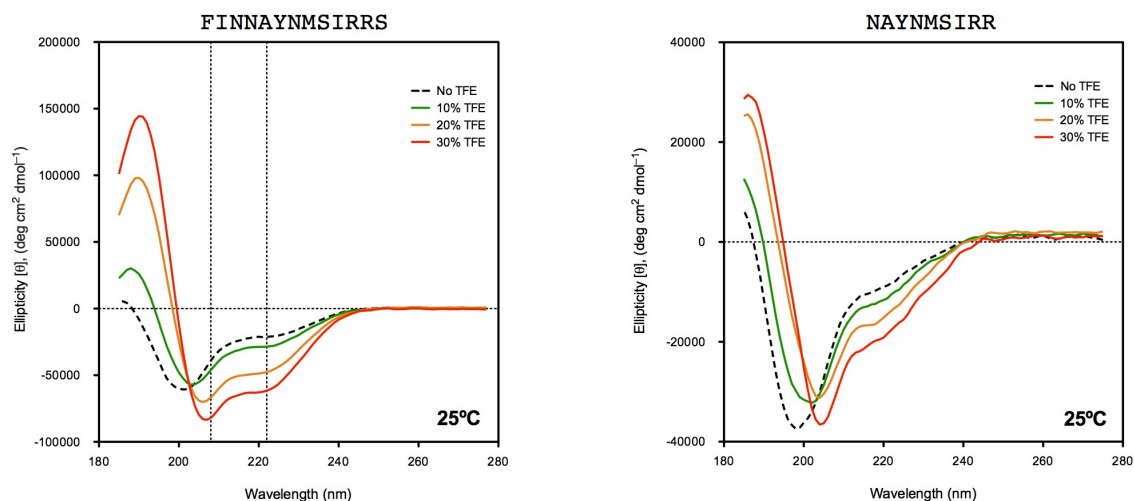


Figure S2. Helical propensity of synthetic epitope-bearing peptides MSP2₁₁₋₂₃ and MSP2₁₄₋₂₂. Both peptides showed decreases in ellipticity at 208 and 220 nm and an increase in peak height at 190 nm, characteristic of α -helical motifs adopted upon trifluoroethanol addition (0 to 30%). The minimal peptide epitope MSP2₁₄₋₂₂ was less helical than the 13-mer peptide (MSP2₁₁₋₂₃) 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.

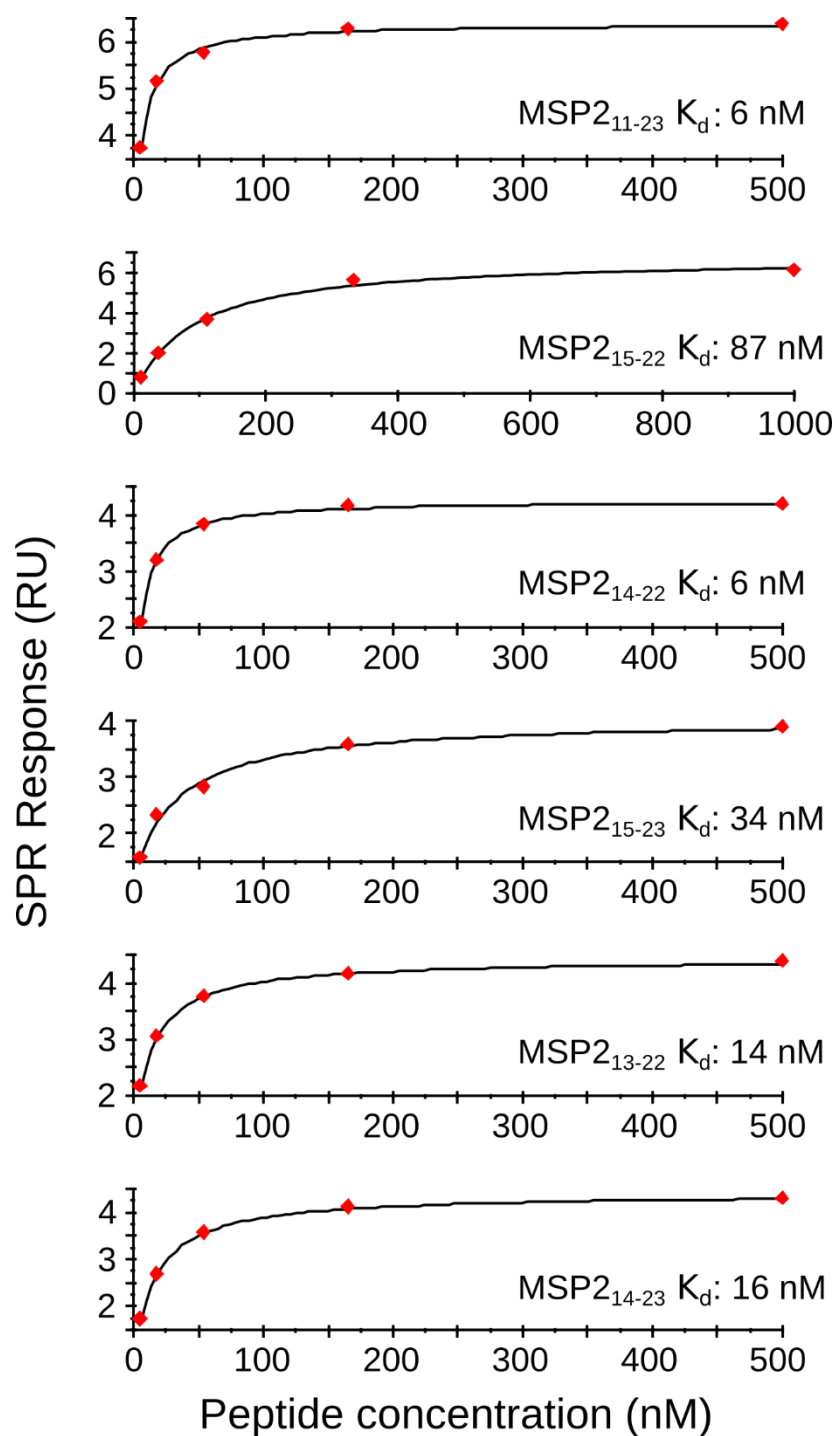


Figure S3. Representative SPR data for the determination of peptide binding affinities for immobilised 6D8.

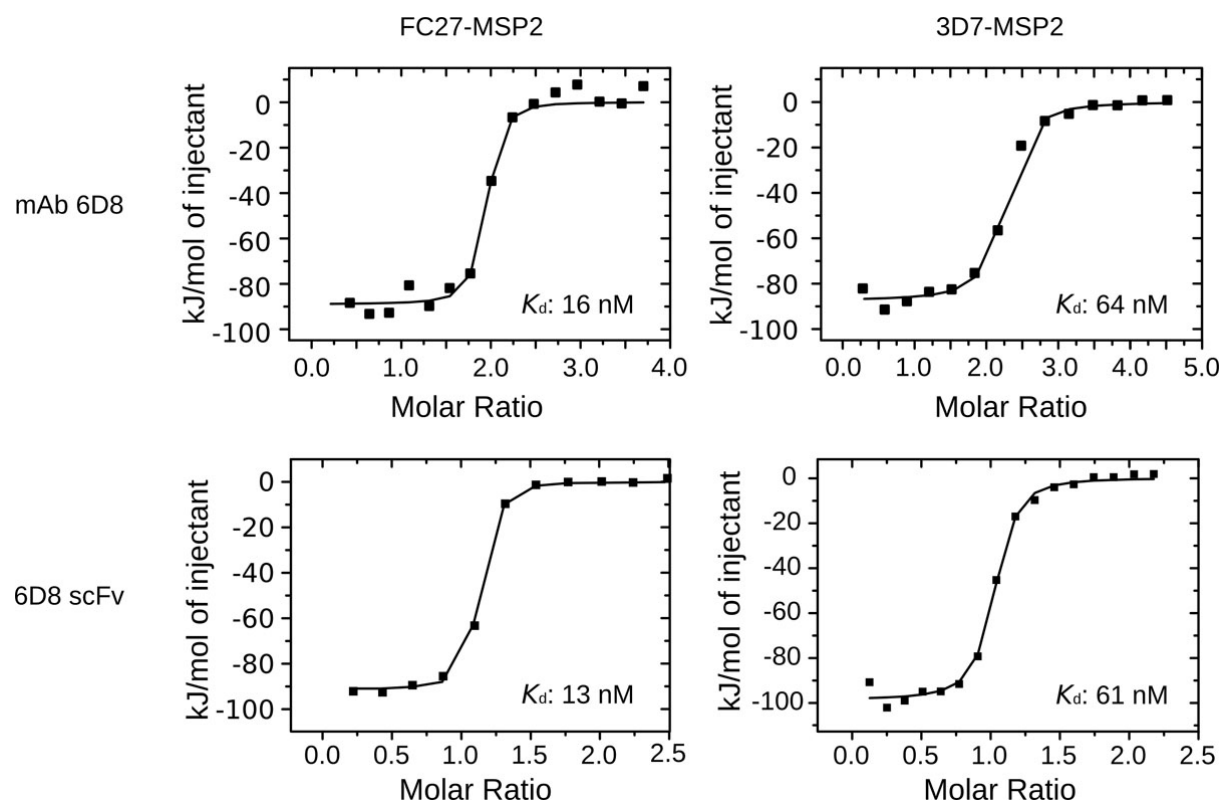


Figure S4. ITC titration of MSP2 to the mAb 6D8 (IgG) and 6D8 scFv. Binding data indicate that recognition of full-length 3D7 and FC27 MSP2 alleles is unaffected by the conversion of the full-length IgG antibody into a single-chain fragment. The strain-selectivity observed in 6D8 IgG is also fully reproduced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

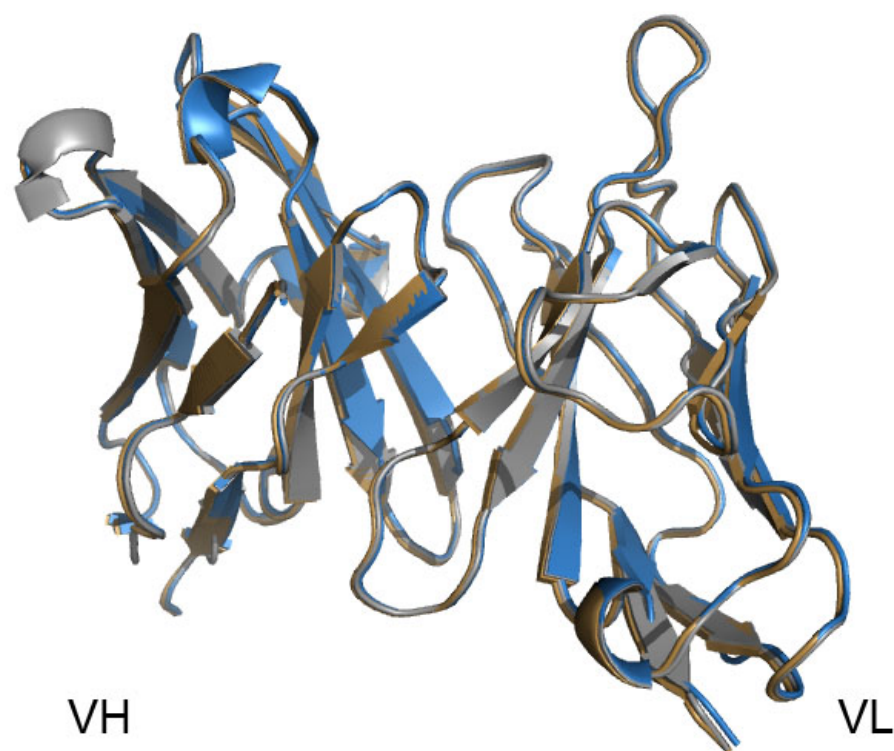


Figure S5. Overlay of the bound antibody structure of crystal complexes 6D8 Fv-MSP2₁₄₋₂₂, 6D8 Fv-3D7 MSP2₁₄₋₃₀ and 6D8 Fv-FC27 MSP2₁₄₋₃₀ shown in grey, blue and brown respectively. Despite marked differences in affinity, side-chain deviations across the three crystal structures were minimal and unlikely to explain affinity differences (C α RMSD of 1.55 Å between 6D8 Fv-MSP2₁₄₋₂₂ and 6D8 Fv-3D7 MSP2₁₄₋₃₀ complexes and C α RMSD of 1.53 Å between 6D8 Fv-MSP2₁₄₋₂₂ and 6D8 Fv-FC27 MSP2₁₄₋₃₀).

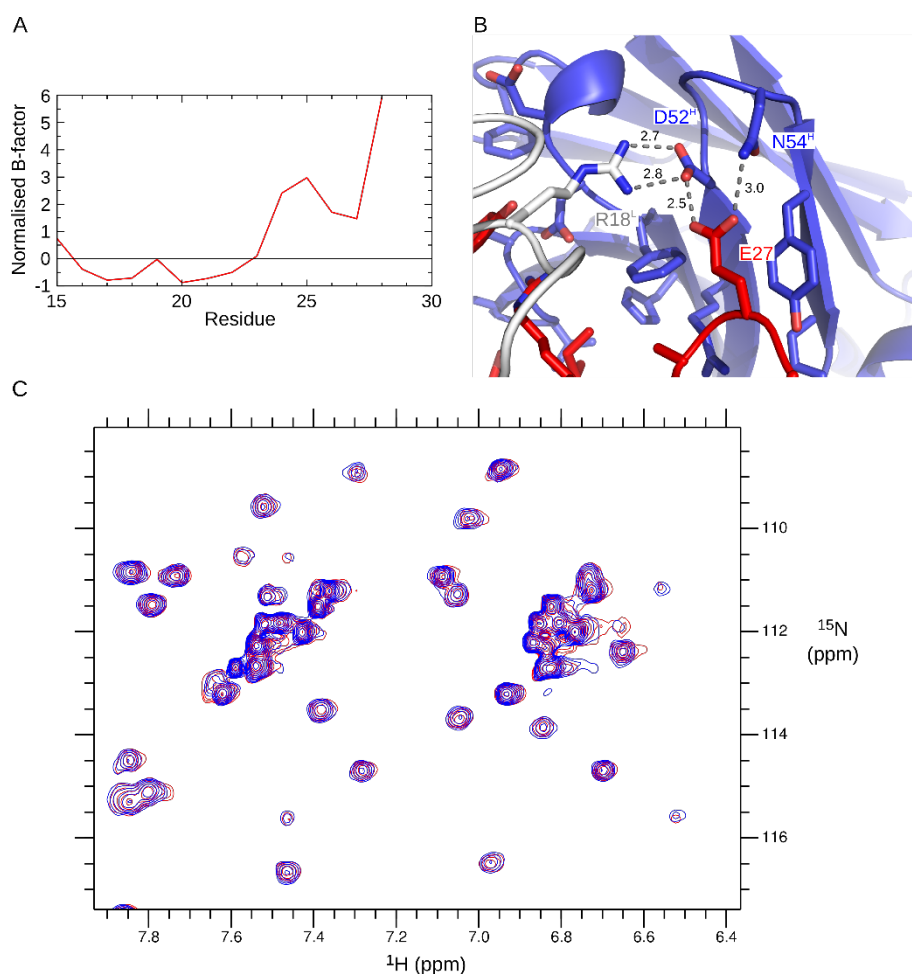


Figure S6. Crystal-contact-mediated interactions of 3D7-MSP2₁₄₋₃₀ with 6D8. **(A).** Normalised B-factors of C α atoms of 3D7-MSP2₁₄₋₃₀ in complex with 6D8 Fv. Isotropic B-factors are normalized as $(B - \langle B \rangle) / \sigma$ where $\langle B \rangle$ and σ are the average and standard deviation, respectively, of all C α B-factors in the complex. **(B).** H-bonds between 3D7-MSP2 residue E27 (red) and 6D8 V_H (D52^H and N54^H; blue) are stabilized by crystal contacts involving R18^L of a crystallographically adjacent Fv complex (grey). Donor-acceptor distances (Å) for H-bonds and salt bridges are labelled. **(C)** Sidechain NH₂ region of ¹H-¹⁵N HSQC spectra of 6D8 scFv in the presence of 3D7-MSP2₁₄₋₃₀ (blue) and FC27-MSP2₁₄₋₃₀ (red) showing the absence of chemical shift perturbations that are expected if E27 of 3D7 MSP2 interacts with the sidechain of N54^H in solution.

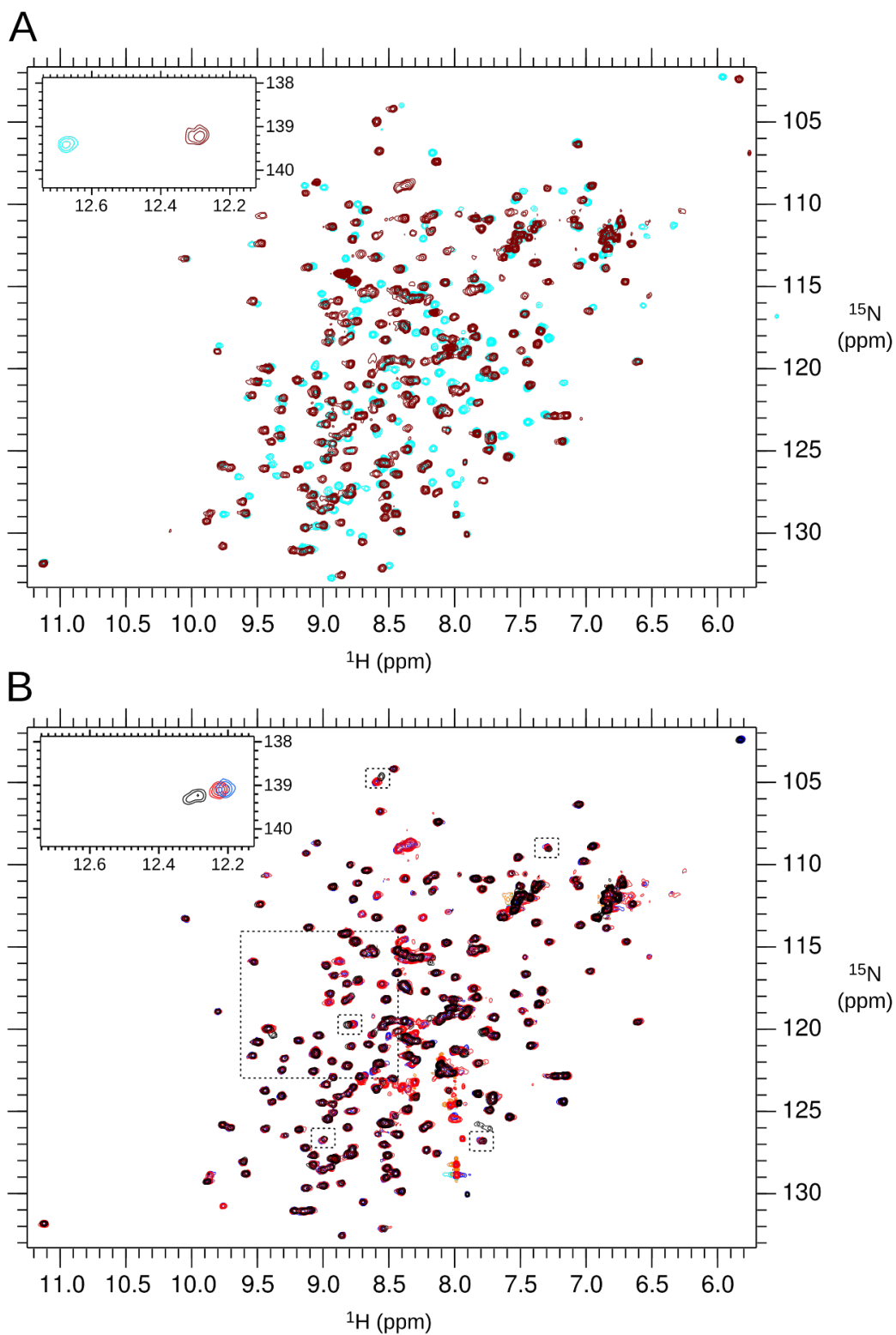


Figure S7. (A) Overlay of ^1H - ^{15}N HSQC spectra of 6D8 scFv in the absence (cyan) and presence (brown) of MSP2₁₄₋₂₂. (B) Overlay of ^1H - ^{15}N HSQC spectra of 6D8 scFv in the presence of MSP2₁₄₋₂₂ (black), FC27-MSP2₁₄₋₃₀ (blue) and 3D7-MSP2₁₄₋₃₀ (red). Spectral regions shown in Fig 4 (main text) are boxed, and an outlying Trp indole NH peak is inset.

		CDR1		CDR2	
IGHV1-78	QVQLQQSGDELVKPGASVKISCKVSGYFTDHTIHWMKQRPEQGLEWIGYIYPRDGSTKY				60
6D8VH	QVQLQQSGDELVKPGASVKLSCTVSGFNKDDFIHWMKQRPEQGLEWIGRIDPANGYTKY				60
	*****:*.**:::.*.***** * * : * **				
	CDR2		CDR3		
IGHV1-78	NEKFKGKATLTADKSSSTAYMQLNSLTSEDSAVYFCAR-----			98	
6D8VH	APKFQDKATMTADTSSNTAYLQLSSLASEDAVYYCATYGVAYWGQGLTVTVSA			114	
	.:*:***.*.***:*.***:***:***:***				

		CDR1		CDR2	
IGKV3-4	DIVLTQSPASLAVSLGQRATISCKASQSVDDGDSYMNWYQQKPGQPPKLLIYAASNL			ES	60
6D8VL	DIVLTQSPASLAVSLGQRATISCKASQSVDDGDSYMNWFQQKPGQSPKLLIYAASNL			ES	60
	*****:*****:*****.*****				
		CDR3			
IGKV3-4	GIPARFSGSGSGTDFTLNIHPVEEEDAATYYCQQSNEPDPTVLQ-----			105	
6D8VL	GIPARFSGSGSGTDFTLNIHPVEEEDAATYYCQQTNEPDPTVFGGGTKLEIK			111	
	*****:*****:***** * *				

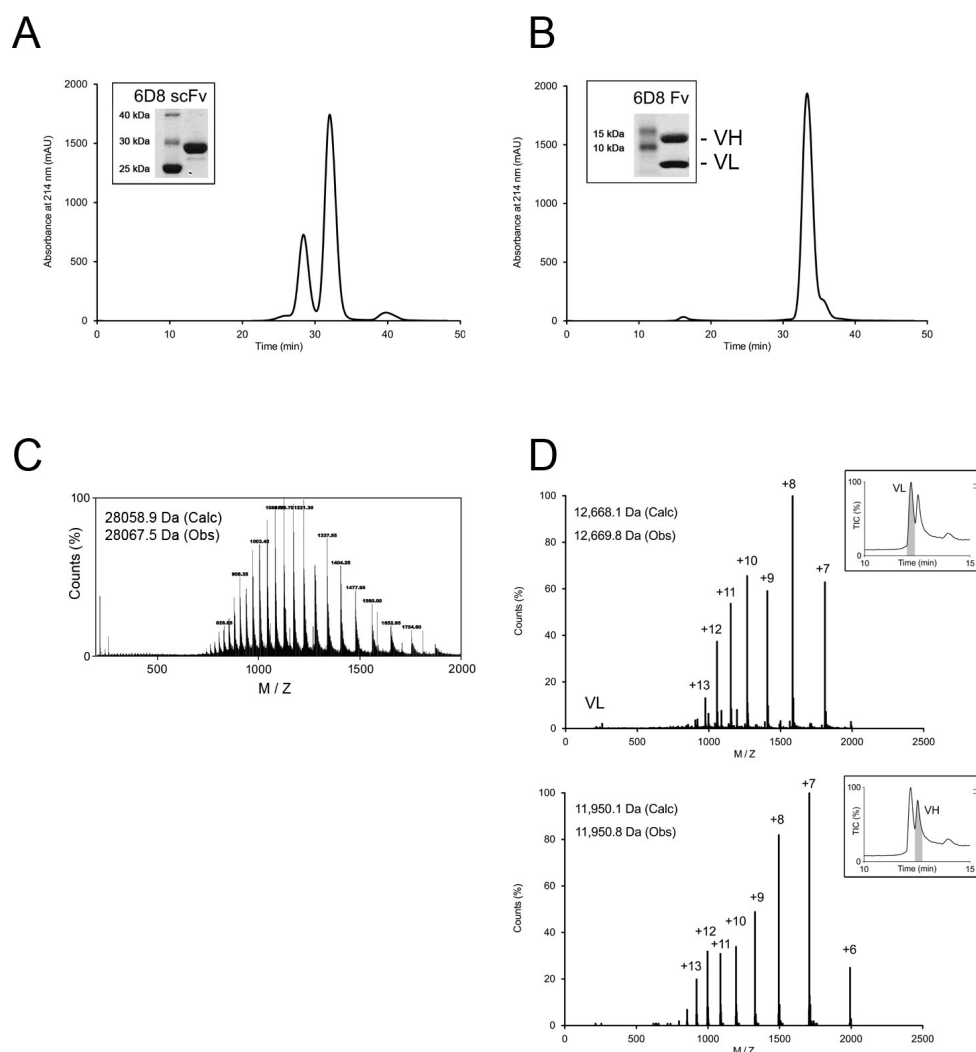


Figure S9. Purification profile of recombinant antibody fragments 6D8 scFv (A) and 6D8 Fv (B) used in this work. Antibody fragments were affinity purified, concentrated and freed from aggregates by gel filtration prior to structural and binding studies. The molecular masses observed for scFv (C) and 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 7

NMR characterization of transient interactions
modulating the strain-specificity of 6D8
monoclonal antibody binding

7.1 Introduction

Chapter 6 revealed that the 6D8 antibody has different affinities for full-length 3D7 MSP2 and FC27 MSP2, even though it recognizes a highly conserved epitope found across all *P. falciparum* strains. Similar affinity results were observed for the 6D8 antibody with 3D7 MSP2₁₄₋₃₀ and FC27 MSP2₁₄₋₃₀. Crystal structures of 6D8 Fv with 3D7 MSP2₁₄₋₃₀ and FC27 MSP2₁₄₋₃₀ epitopes could not explain the basis for this unusual strain-specific affinity of 6D8 because the crystal contacts for the extended regions of 6D8 were not observed. Solution NMR experiments of 6D8 scFv in the presence of 3D7 MSP2₁₄₋₃₀ and FC27 MSP2₁₄₋₃₀ epitopes showed significant differences in the spectra. A few resonances were shown to differ in their chemical shifts and some other resonances differed in their intensities. This suggests that the difference in the affinity of 6D8 for 3D7 MSP2 and FC27 MSP2 could be due to additional interactions between 6D8 and the residues from the MSP2 variable region. Detailed analysis of these interactions is limited due to the unavailability of sequential NMR assignments for the 6D8 scFv.

Hence, in this study, we have determined the resonance assignments for 6D8 scFv in the presence and absence of MSP2₁₄₋₂₂. These assignments were used to analyse transient interactions between 6D8 and the variable regions of 3D7 and FC27 MSP2. Triple-resonance NMR experiments were recorded on triple-labelled 6D8 scFv for determining NMR assignments. ²H, ¹³C, ¹⁵N-labelled triple-labelled 6D8 scFv was used instead of ¹H, ¹³C, ¹⁵N-labelled 6D8 scFv for all the NMR experiments in order to generate high-quality NMR data. We have determined the backbone NMR assignments for 6D8 scFv both in the presence and absence of the 6D8 epitope, MSP2₁₄₋₂₂. 6D8 residues that make transient interactions were identified by comparing the two-dimensional [¹H-¹⁵N]-TROSY spectra of 6D8 scFv recorded in the presence of both 3D7 and FC27 MSP2₁₄₋₃₄. CLEANEX NMR experiments were also recorded to identify residues in exchange with solvent.

7.2 Materials and Methods

7.2.1 ²H, ¹³C and ¹⁵N-labelled 6D8 scFv expression

The 6D8 scFv construct, consisting of the light chain and heavy chain of the mouse 6D8 monoclonal antibody, was chosen for NMR experiments in the presence and absence of its epitope. The 6D8 epitope (MSP2₁₄₋₂₂) extended by up to 9 residues from either the 3D7 or FC27 MSP2 variable regions was used in this study (**Table 7.1**). Bacterial cells expressing

the triple-labelled 6D8 scFv were provided by ANSTO (Australian Nuclear Science and Technology Organization) as cell pellets [188].

Table 7.1. Amino acid sequences for the 6D8 scFv, MSP2₁₄₋₂₂, FC27 MSP2₁₄₋₃₄ and 3D7 MSP2₁₄₋₃₄. Variable heavy chain (VH, 1-114), variable light chain (VL, 131-241). The complementarity determining region (CDR) regions are underlined, linker sequence and the extra residues from the vector along with a 6His-tag in 6D8 scFv are shown in red and the extended regions of the 6D8 epitope sequences are in bold.

Protein/ Peptides	Amino acid sequence
6D8 scFv	QVQLQQSGDELVKPGASVKLSCTVSGFNIKDDFIHWMKQRPEQGLE WIGRIDPANGYTKYAPKFQDKATMTADTSSNTAYLQLSSLASEDAA VYYCATYGVAYWGQGLTVSA GGGGSGGGSGGGG TDIVLTQS PASLAVSLGQRATISCKASQSVDHDGDSYMNWFQQKPGQSPKLLIY <u>AASNLESGIPARFSGSGSGTDFTLNHPVEEEDAATYYCQQTNEDPY</u> <u>TFGGG</u> TKLEIK GSEQKLISEEDLNGSHHHHHH
MSP2 ₁₄₋₂₂	NAYNMSIRR
FC27 MSP2 ₁₄₋₃₄	NAYNMSIRRSMANEGSNTNSV
3D7 MSP2 ₁₄₋₃₄	NAYNMSIRRSMAESKPSTGAG

7.2.2 6D8 scFv purification

7.2.2.1 Periplasmic extraction

The bacterial cell pellet was resuspended gently in sucrose buffer (25 mM Tris pH 8, 2 mM EDTA and 30% sucrose) containing protease inhibitors cocktail (Roche, Mannheim, Germany) and centrifuged at 10,000 g for 20 min. The pellet was resuspended in 30 mL ice-cold MQ water containing 2 mM MgCl₂ and 10 µg/mL DNase and incubated for 15 min to release the periplasmic contents. The supernatant containing periplasmic contents was collected carefully after centrifugation at 20,000 g for 20 min.

7.2.2.2 Preparation of MSP2 affinity column

7 mL of *N*-Hydroxysuccinimidyl-Sepharose beads (Sigma Aldrich) were washed in a step wise manner with 100% isopropanol, 100% ethanol and 50 mM phosphate buffer pH 7. 3D7 MSP2 was dissolved in 50 mM phosphate buffer pH 7 and incubated with the beads for 1 h at room temperature. MSP2 bound beads were washed with 50 mM phosphate buffer pH 7

and 10 mL of 6M Gdn-HCl followed by 50 mL of 0.1M NaOH and finally equilibrated with 50 mM phosphate buffer pH 7.

7.2.2.3 6D8 scFv affinity purification

The supernatant containing the 6D8 scFv was incubated for 1 h on the MSP2 affinity column, which was equilibrated with 50 mM phosphate buffer, pH 7. The column was washed with two column volumes of 50 mM phosphate buffer to remove impurities and unbound misfolded 6D8 scFv. The 6D8 scFv was eluted quickly under denaturing conditions using 10 mL of 0.1 M glycine, pH 2.5, and for the immediate refolding of 6D8 scFv the elution was collected directly into a falcon tube containing 2 mL of 1 M Tris, pH 8. The samples were analyzed by SDS-PAGE and the protein solution was dialyzed for 24 h in 20 mM ammonium bicarbonate buffer and lyophilized. All the purification steps were carried out at 4 °C.

7.2.3 Size-exclusion chromatography (SEC)

All the solutions and buffers used in size exclusion chromatography (SEC) were filtered through a 0.22 µm membrane and degassed for 20 min. The lyophilized protein was dissolved in 20 mM citrate buffer, pH 6.5, containing 50 mM NaCl. A Superdex 75 column (GE) was equilibrated with 20 mM citrate buffer, pH 6.5, containing 50 mM NaCl at a flow rate of 0.5 mL/min. Protein was loaded onto the column at 3 mg/mL to reduce the dimerization apparent at higher concentrations [189-191].

7.2.4 Solution NMR spectroscopy

7.2.4.1 Sequential NMR assignments

^2H , ^{13}C , ^{15}N -triple labelled 6D8 scFv fractions from SEC were concentrated to 250 µM in a buffer containing 20 mM sodium citrate, pH 6.5, 7% $^2\text{H}_2\text{O}$ for NMR studies. NMR spectra were acquired on a Bruker Avance III spectrometer equipped with a cryogenically cooled triple-resonance probe, operating at a frequency of 600 MHz at 308 K. To determine sequential backbone assignments, a set of transverse relaxation optimized spectroscopy (TROSY) [192-194] versions of three-dimensional HNCO [195], HNCA [195], HN(CO)CA [196], HNCACB [196], HN(CA)CB [196] and HN(CO)CACB [196] spectra using echo-anti echo and ^2H -decoupling were acquired using iterative non-uniform sampling (NUS) [197], along with a two dimensional [^{15}N - ^1H]-TROSY. Conventional data were processed using

Bruker Topspin 3.2 and non-uniform sampling data were reconstructed with compressed sensing using qMDD [198, 199]. The backbone assignments were determined manually using CcpNmr Analysis 2.4.2 [200, 201] together with the RASP (Resonance Assignment by chemical Shift Prediction) automatic assignment process [202]. The parameters used for these experiments are presented in **Table 7.2**.

Table 7.2. Number of scans and data points used for acquiring triple-resonance NMR experiments.

Experiment	HNCO	HNCA	HN(CO)CA	HNCACB	HN(CO)CACB	HN(CA)CB
Number of scans	24	64	64	64	56	48
Number of data points						
F3 (^1H)	2048	2048	2048	2048	2048	2048
F1 (^{13}C)	42	64	128	128	128	128
F2 (^{15}N)	128	42	42	42	42	42

7.2.4.2 Peptide binding experiments by NMR

Two-dimensional [^1H - ^{15}N]-TROSY spectra of 6D8 scFv were recorded with 128 scans and 256 increments in the presence and absence of MSP2₁₄₋₂₂, FC27 MSP2₁₄₋₃₄ and 3D7 MSP2₁₄₋₃₄. The concentrations of 6D8 scFv and peptides were 30 and 36 μM , respectively. The buffer used in these experiments contained 20 mM sodium citrate, pH 6.5, 7% (v/v) $^2\text{H}_2\text{O}$. To determine the sequential assignments for MSP2₁₄₋₂₂ bound 6D8 scFv, the same set of triple-resonance experiments (**Table 7.2**) was recorded under identical conditions. The spectra were processed using Bruker Topspin 3.2 and analyzed using CcpNmr Analysis 2.4.2 [200, 201]. The amino acid sequence for 6D8 scFv is shown in **Table 7.1**. Weighting of chemical shifts from ^1H and ^{15}N resonances was performed using the following equation [203, 204]:

$$\text{Weighted CSP } (\Delta\delta) = \sqrt{\frac{1}{2}[\delta_{\text{H}}^2 + (0.14 \cdot \delta_{\text{N}})^2]} \quad (1)$$

Where, δ_{H} and δ_{N} are the chemical shift perturbations (CSPs in ppm) in ^1H and ^{15}N dimensions, respectively.

Phase-modulated CLEANEX (clean chemical exchange) [^1H - ^{15}N]-TROSY [205] experiments were recorded on 6D8 scFv in the presence and absence of MSP2₁₄₋₂₂, 3D7 MSP2₁₄₋₃₄ and 3D7 MSP2₁₄₋₃₄ to identify amide protons that are in chemical exchange with

water. Spectra were acquired with 120 scans and 2048 and 256 data points for ^1H and ^{15}N dimensions, respectively. Samples were prepared at 30 μM concentration in 20 mM sodium citrate buffer, pH 6.5 containing 7% $^2\text{H}_2\text{O}$ and the data were collected with mixing times of 25, 45, and 65 ms.

All the Figures for protein data bank (PDB) files were prepared using PyMOL (version 1.2r3pre Schrodinger, LLC.).

7.3 Results

The 6D8 scFv was successfully purified using a MSP2 affinity column, and SDS-PAGE analysis indicated that the protein was >95% pure (**Figure 7.1A**). Monomeric 6D8 scFv was successfully separated from its dimeric form using SEC (**Figure 7.1B**). The purified samples were concentrated using a Centricon filter (10 kDa MWCO, Millipore) and used directly for NMR experiments.

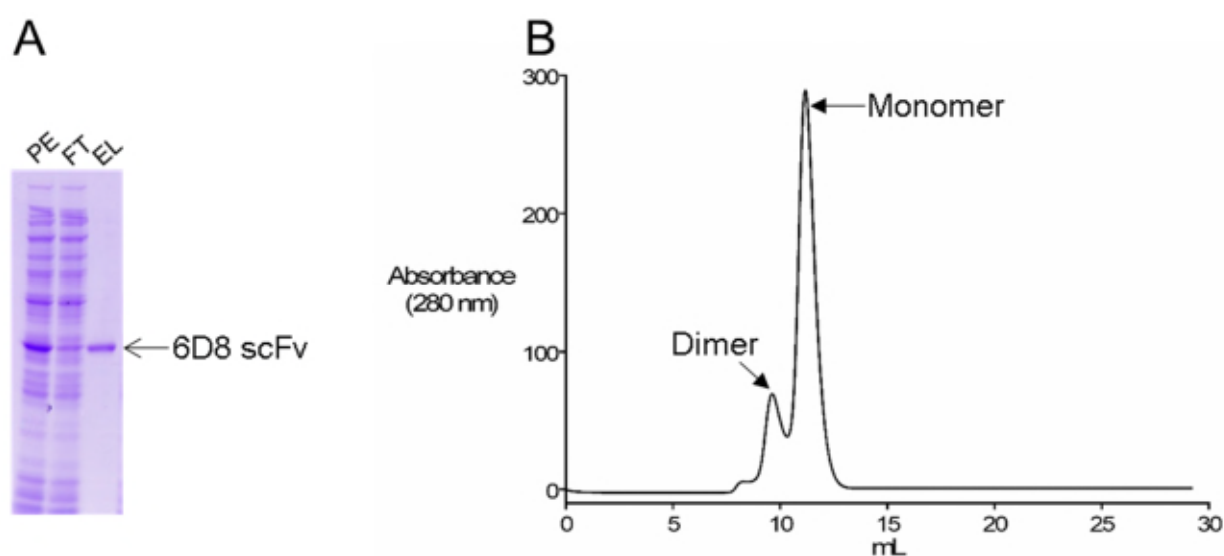


Figure 7.1. A) SDS-PAGE analysis of affinity purified 6D8 scFv. PE=periplasmic extract, FT=flow through and EL=elution containing 6D8 scFv. B) Size-exclusion chromatogram showing the peaks corresponding to the monomeric and dimeric forms of 6D8 scFv. The monomeric form of 6D8 scFv was used for NMR experiments.

7.3.1 Sequential backbone resonance assignments of MSP₁₄₋₂₂ free 6D8 scFv

6D8 scFv consists of 262 residues, including 10 Pro residues, 6His-tag, the 15-residue Gly-rich linker and 15 residues from the vector, with a molecular mass of ~ 28 kDa. Despite having well dispersed resonances in the NMR spectrum, many resonances were overlapping and some were missing in the spectrum recorded on protonated sample (**Figure 7.2A**). In

contrast, an improved spectral resolution and the appearance of more resonances were observed in spectra recorded on deuterated 6D8 scFv compared to the spectrum of protonated 6D8 scFv (**Figure 7.2B**).

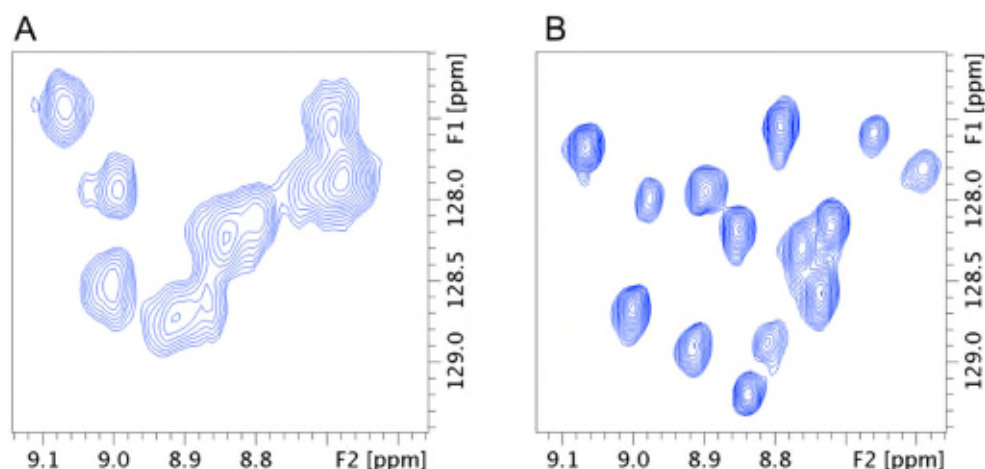


Figure 7.2. Comparison of a region of 2D spectra A) $[^1\text{N}-^1\text{H}]$ -HSQC of non-deuterated 6D8 scFv and B) $[^{15}\text{N}-^1\text{H}]$ -TROSY of deuterated 6D8 scFv. Resolution was improved in the spectrum recorded on deuterated 6D8 scFv compared to the spectrum recorded on the protonated 6D8 scFv. The better resolution and sensitivity in spectrum B as compared to spectrum A is a result of reduced linewidths upon deuteration together with TROSY effect [206, 207].

Well-dispersed resonances in the $[^{15}\text{N}-^1\text{H}]$ -TROSY spectrum and a uniform intensity indicate a single correctly-folded conformation of the 6D8 scFv (**Figure 7.3**). Backbone assignments for 190 residues, corresponding to 85% of observed resonances, were completed for MSP2₁₄₋₂₂ free 6D8 scFv. 50% of backbone assignments were obtained from the RASP automatic assignment process and the other 35% were completed by manual assignment using CcpNmr Analysis 2.4.2 [200, 201]. An example of an assignment strip plot from the 3D HNCACB and HN(CO)CACB spectra is shown in **Figure 7.4**. A full list of chemical shifts is given in Appendix V.

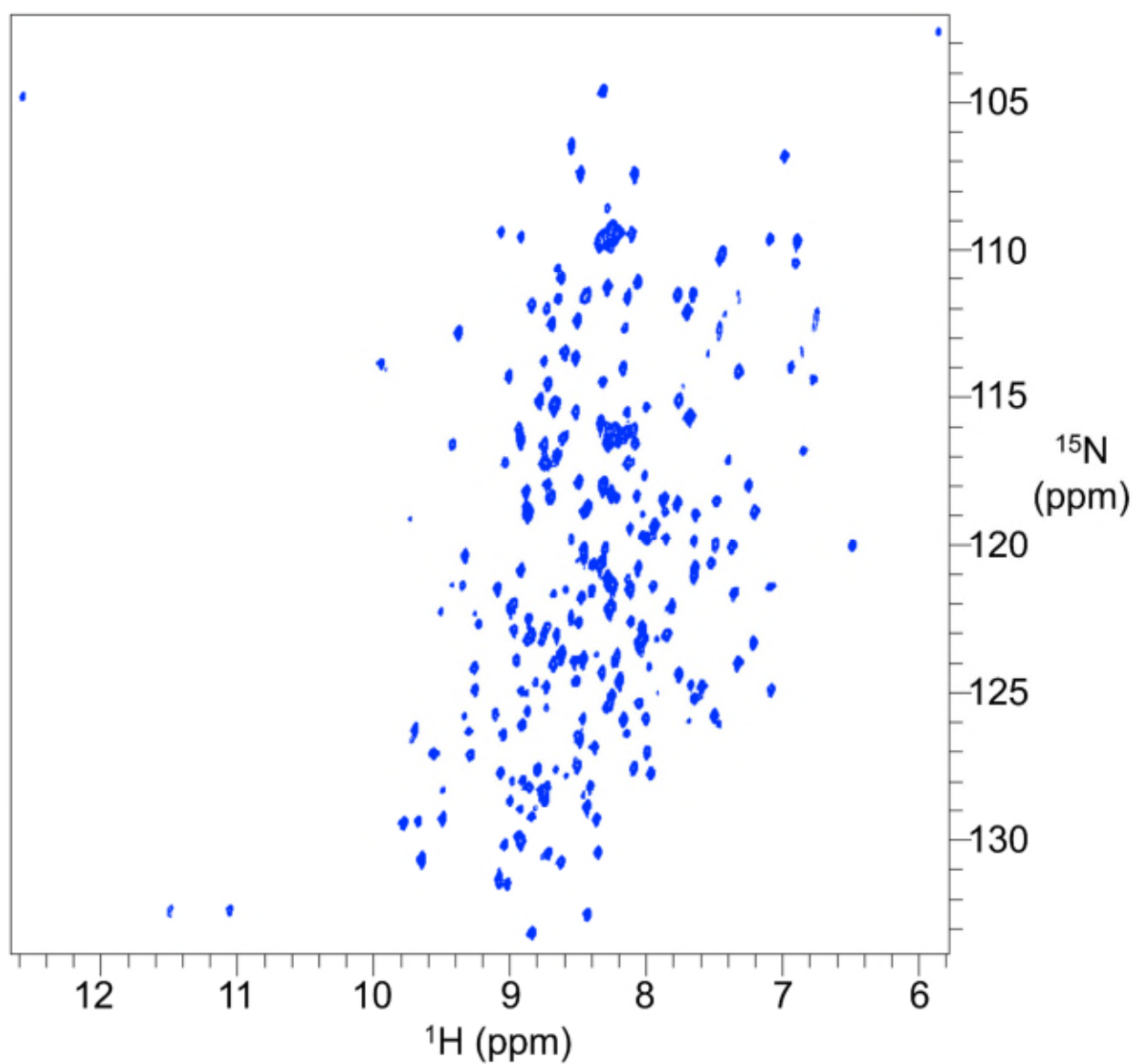


Figure 7.3. Two-dimensional [^1H - ^{15}N]-TROSY spectrum of 6D8 scFv. Well dispersed resonances indicate that the protein is well folded.

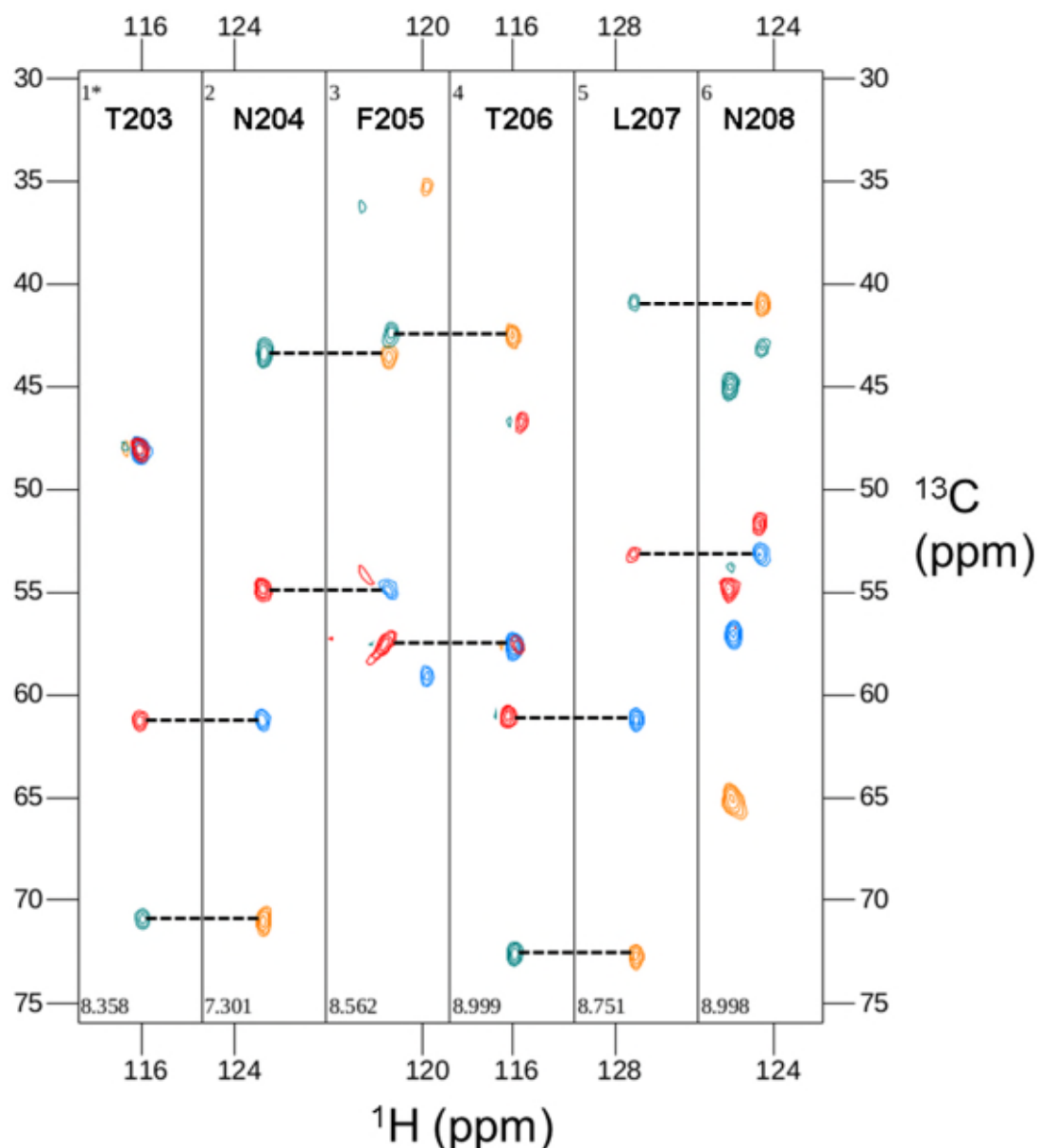


Figure 7.4. A strip plot of HNCACB and HN(CO)CACB showing sequential connectivities for a stretch of six residues from the 6D8 scFv. Green and red peaks (HNCACB) represent the C α and C β chemical shifts from *i* and *i*-1 residues. Blue and orange (HN[CO]CACB) peaks represent the C α and C β resonances from *i*-1 residue.

7.3.2 Sequential backbone resonance assignments for MSP2₁₄₋₂₂-bound 6D8 scFv

NMR assignments for MSP2₁₄₋₂₂-bound 6D8 scFv are required in order to map the residues that make transient interactions with the MSP2 variable region. Initially, most of the assignments for MSP2₁₄₋₂₂-bound 6D8 scFv were obtained directly by comparing two-dimensional [^1H - ^{15}N]-TROSY spectra of 6D8 scFv in the presence and absence of MSP2₁₄₋₂₂. However, there were resonances that showed large chemical shift changes in the presence

of the MSP₁₄₋₂₂ and could not be assigned directly. For this reason, a new set of triple-resonance data collected on MSP₂₁₄₋₂₂-bound 6D8 scFv was recorded for completing the assignments. 80% of assignments were obtained for the MSP₁₄₋₂₂-bound 6D8 scFv. Of the six residues that make H-bonds with the MSP₂₁₄₋₂₂ in the crystal structure, four were assigned (Asp31^H, Asp32^H, Asn168^L, and Thr225^L) and two were unassigned (Val101^H, Asn226^L) (Figures 7.5 and 7.6). The assigned residues of both MSP₂₁₄₋₂₂ free and MSP₁₄₋₂₂ bound 6D8 scFv were mapped onto the crystal structure of MSP₂₁₄₋₂₂ bound 6D8 Fv (Figure 7.7).

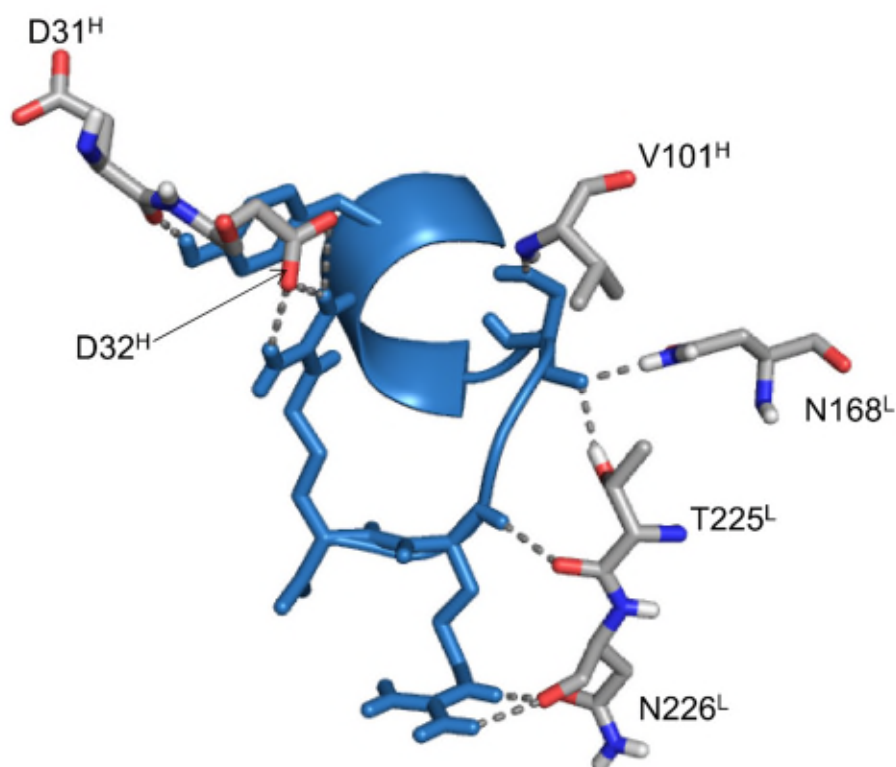


Figure 7.5. Crystal structure showing the H-bonds between the 6D8 epitope and 6D8 Fv (PDB ID: 4QYO) [190]. H and L represents 6D8 scFv variable region heavy and light chains, respectively.

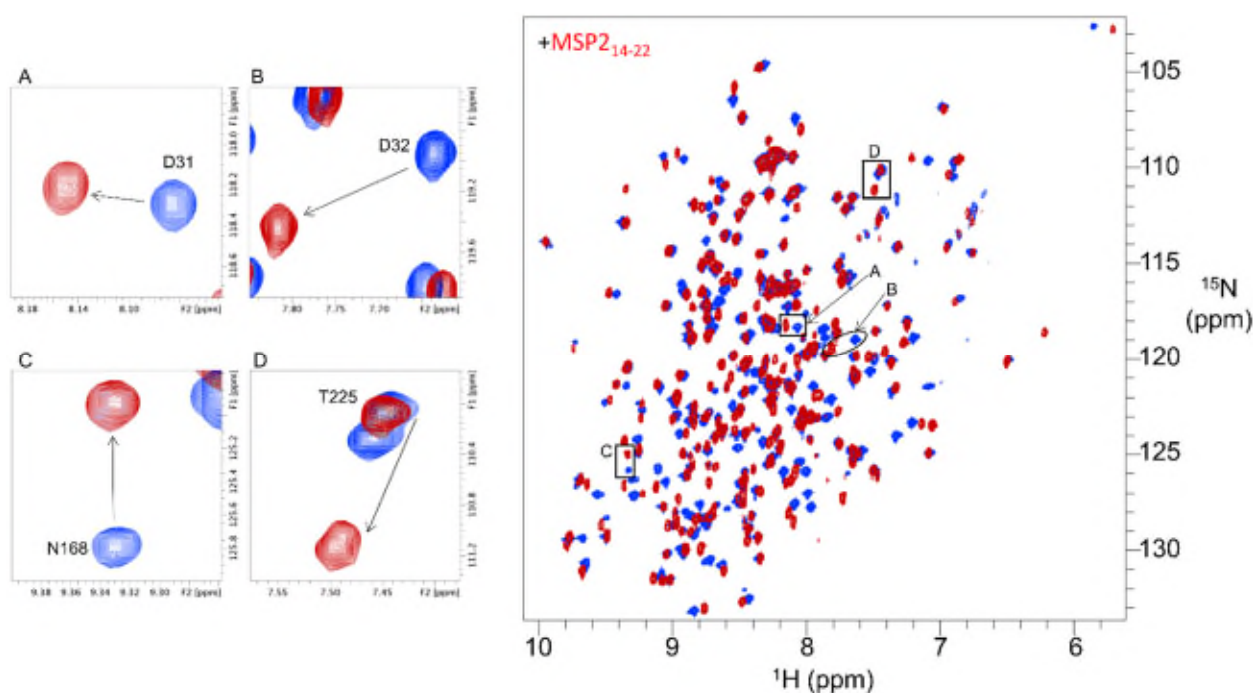


Figure 7.6. Two-dimensional [^1H - ^{15}N]-TROSY spectrum of 6D8 scFv in the presence (red) and absence (blue) of MSP2₁₄₋₂₂. The perturbed resonances shown on the left side are from the residues that made H-bonds and stabilize the helix of MSP2₁₄₋₂₂ in the crystal structure of MSP2₁₄₋₂₂-bound 6D8 Fv (**Figure 7.5**).

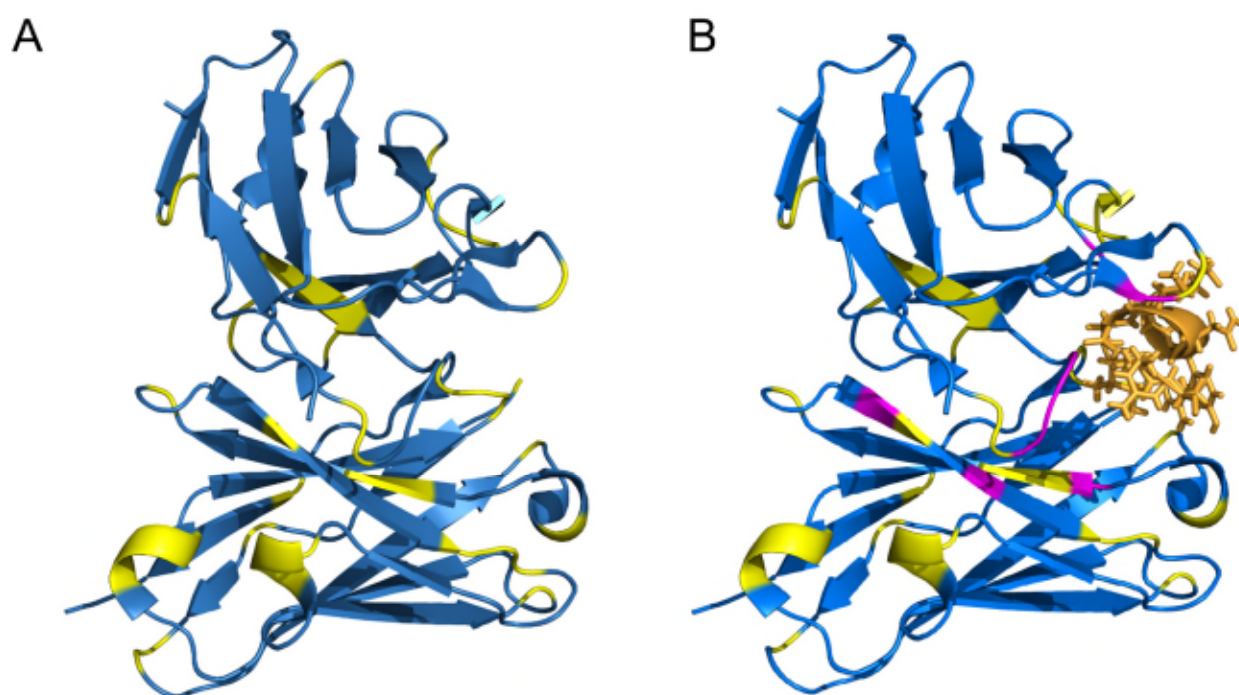


Figure 7.7. Mapping NMR assignments for the 6D8 scFv in the absence (A) and presence (B) of MSP2₁₄₋₂₂ onto the crystal structure of MSP2₁₄₋₂₂ bound 6D8 Fv (PDB ID: 4QYO) [190]. Assigned: blue; unassigned: yellow, unassigned residues in the MSP2₁₄₋₂₂-bound 6D8 scFv that are assigned in the MSP₁₄₋₂₂-free 6D8 scFv: magenta and MSP2₁₄₋₂₂: orange. For clarity MSP₁₄₋₂₂ is not shown in A.

6D8 scFv spectra showed large chemical shift changes in the presence of MSP2₁₄₋₂₂. CSPs were calculated for the perturbed resonances in the [¹H-¹⁵N]-TROSY spectrum of MSP2₁₄₋₂₂ bound 6D8 scFv (**Figures 7.8A**). The perturbed resonances with CSP values >0.05 ppm were mapped onto the crystal structure of MSP2₁₄₋₂₂ bound 6D8 Fv (PDB ID: 4QYO) (**Figure 7.8B**). Chemical shift changes observed for the residues that are far from the MSP2₁₄₋₂₂ binding site indicate a large conformational change in 6D8 scFv due to allosteric effects upon binding to MSP2₁₄₋₂₂.

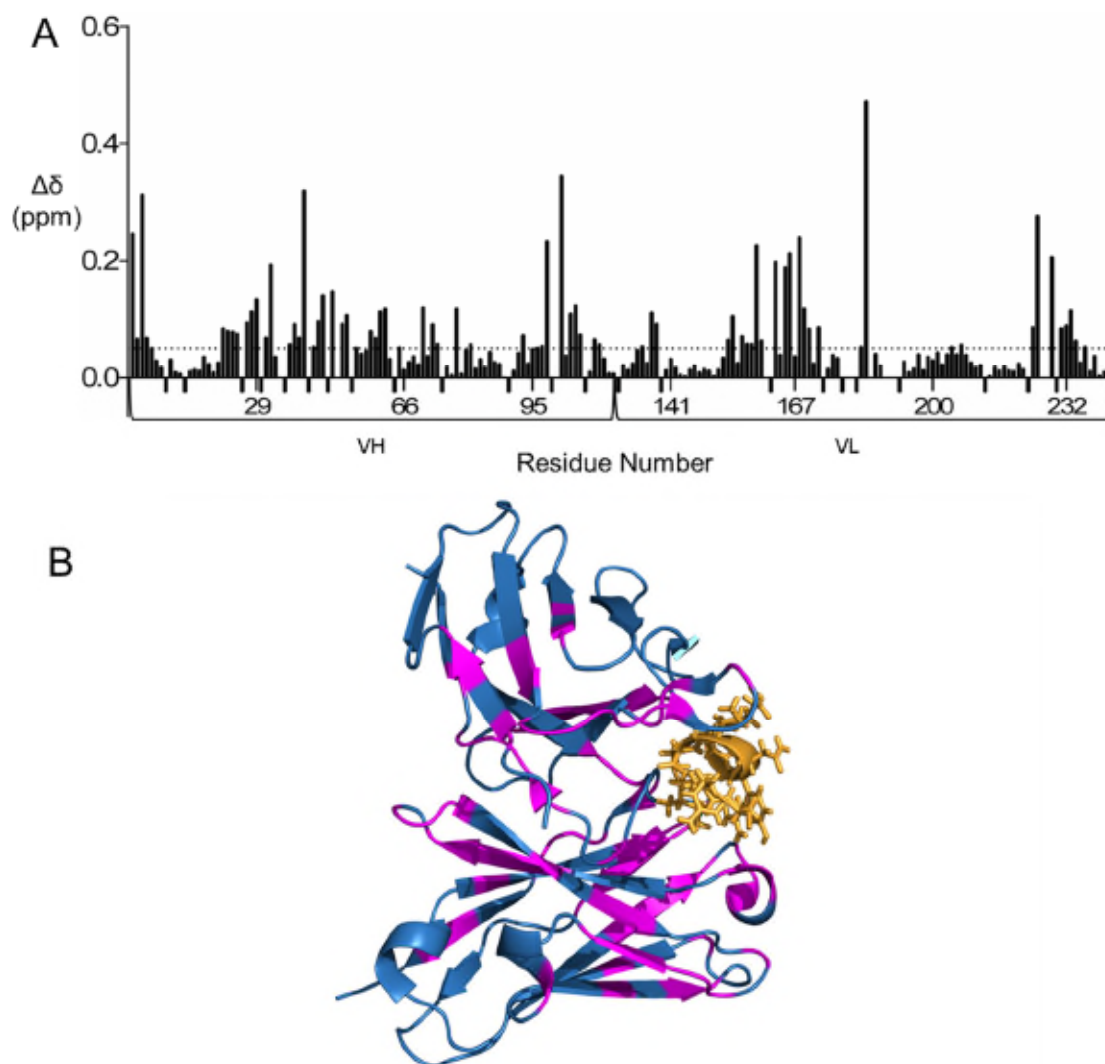


Figure 7.8. A) Weighted chemical shift perturbations ($\Delta\delta$) of 6D8 scFv upon binding to MSP2₁₄₋₂₂. 15 residues from the Gly-rich linker, 15 C-terminal residues from the vector along with a 6His-tag are excluded in the plot. Residues, 26-35, 49-66, 99-103 and 154-169, 184-190, 223-231 are complementarity determining regions in VH and VL domains, respectively. The thick negative bars indicate the regions of unassigned residues. B) The residues with >0.05 CSPs are mapped onto the crystal structure of MSP₁₄₋₂₂ bound 6D8 Fv. 6D8 Fv (blue), perturbed residues (magenta) and MSP2₁₄₋₂₂ (orange).

7.3.3 H^N -water exchange in presence and absence of MSP2₁₄₋₂₂

Most of the unassigned residues were due to signal overlap of resonances coming from flexible loop regions. To probe the number of amide protons that are in chemical exchange with solvent, CLEANEX experiments were performed on 6D8 scFv in the presence and absence of MSP2₁₄₋₂₂. Overall, CLEANEX NMR spectra recorded at different mixing times showed a similar number of resonances but their intensity varied, with stronger resonances at longer mixing times (**Figure 7.9A-C**). The highly overlapped resonances observed in the spectrum correspond to a stretch of Gly and Ser residues from the linker sequence and could also be resonances from the additional cloning-generated residues in the C-terminal region (**Table 7.1**). The assigned residues in the CLEANEX spectrum were mapped onto the crystal structure of MSP2₁₄₋₂₂-bound 6D8 Fv (**Figures 7.9D and 7.10**). The majority of the residues observed in CLEANEX spectrum are from the loop regions and few (eg. Thr23, Ser140 and Gly200) are from the structured regions (**Figure 7.10**). The resonance for Thr23 was stronger in the spectrum recorded with longer mixing time and the Ser140 was observed only at longer mixing time, suggesting that Thr23 was in rapid chemical exchange with solvent compared to Ser140. Some of the resonances that showed perturbations in the two-dimensional [1H - ^{15}N]-TROSY spectrum of MSP2₁₄₋₂₂-bound 6D8 scFv, are also observed in the CLEANEX spectrum of MSP2₁₄₋₂₂-bound 6D8 scFv (**Figure 7.9E**). Perturbations of solvent-exposed residues indicate a large conformational change in the structure of 6D8 scFv in the presence of MSP2₁₄₋₂₂.

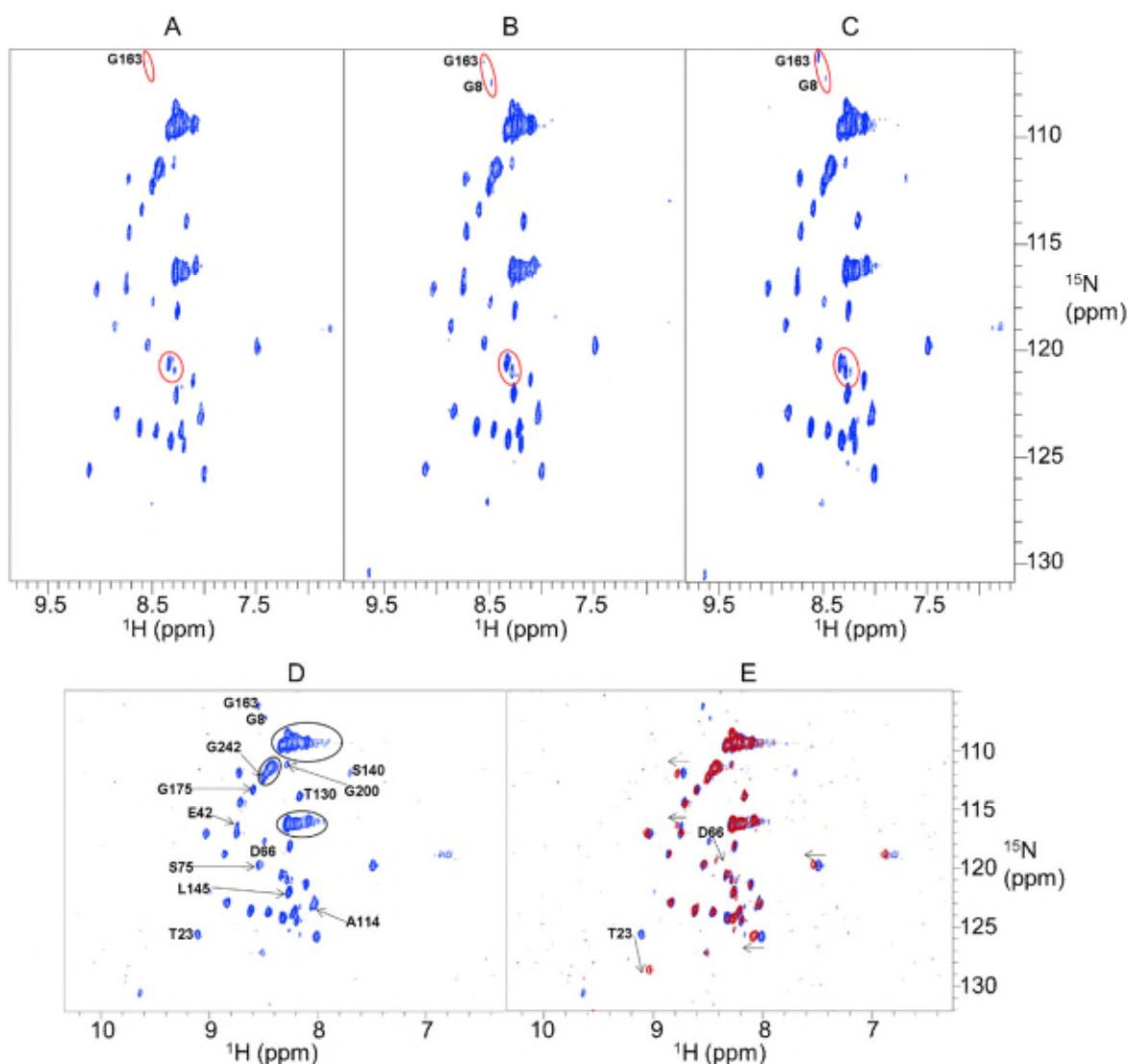


Figure 7.9. H^N -water exchange monitored by CLEANEX experiment. The data were collected at 600 MHz 1H frequency, pH 6.5, 308 K, with a mixing time of 25 ms (A), 45 ms (B) and 65 ms (C). The number of scans was 120 for each increment, with a total of 256 increments. The peak intensity increased with increasing mixing time, for example resonances shown in red circles including Gly8 and Gly163 appeared at higher mixing times. D) CLEANEX-TROSY spectrum of 6D8 scFv. Resonances without labels are unassigned. The overlapped regions are shown in black circles E) Overlay of CLEANEX-TROSY spectra of MSP₂₁₄₋₂₂-free (blue) and MSP₂₁₄₋₂₂-bound (red) 6D8 scFv. The perturbed resonances in the presence of MSP₂₁₄₋₂₂ are shown with arrows. Thr130 and Gly242 are from the linker and cloning-generated sequences, respectively. Resonances not labelled are unassigned.

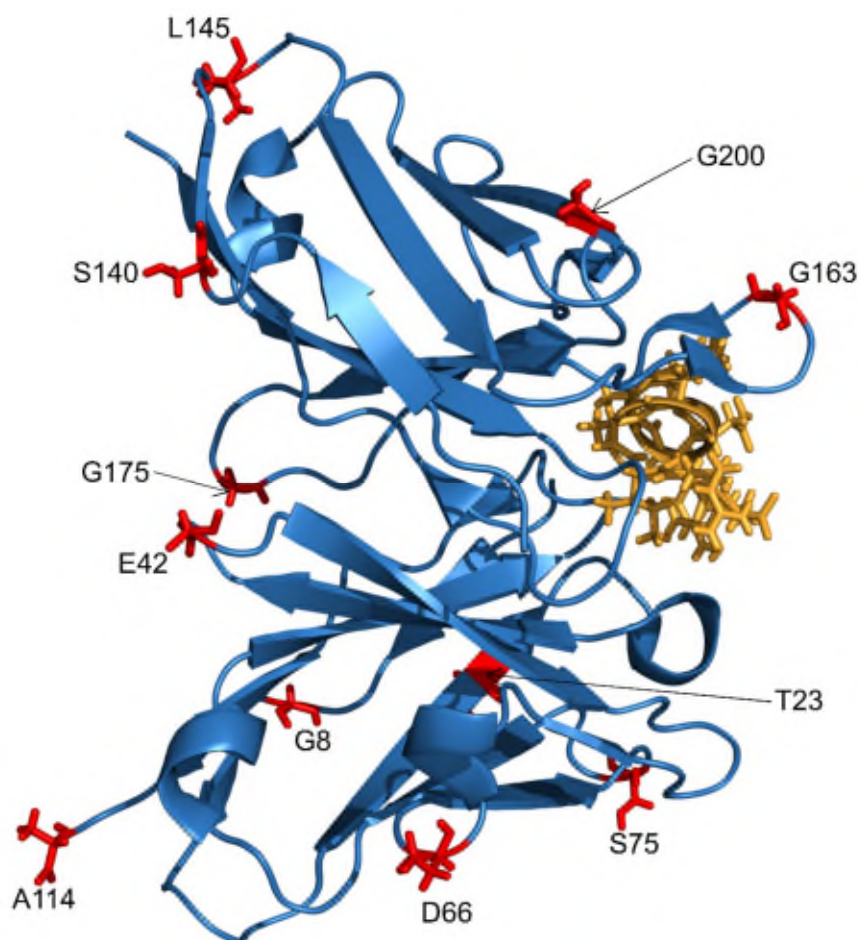


Figure 7.10. Mapping the residues that are in chemical exchange with solvent protons (observed in the CLEANEX spectrum [Figure 7.9]) on to the crystal structure of MSP₂₁₄₋₂₂-bound 6D8 Fv. Only assigned residues in 6D8 scFv that are in exchange with solvent protons are shown (red). 6D8 Fv (blue) and MSP₂₁₄₋₂₂ (orange).

7.3.4 Identification of 6D8 scFv residues that interact with MSP2 variable region

Binding experiments (Chapter 6) of 6D8 antibody showed different affinities for full-length MSP2 (3D7 and FC27), as well as for 3D7 MSP₂₁₄₋₃₄ and FC27 MSP₂₁₄₋₃₄ epitopes. The assignments of MSP₂₁₄₋₂₂-bound 6D8 scFv were used for assigning resonances in the spectra of 6D8 scFv bound to 3D7 MSP₂₁₄₋₃₄ and FC27 MSP₂₁₄₋₃₄. Resonances with different chemical shifts between the spectra of 6D8 scFv bound to 3D7 MSP₂₁₄₋₃₄ and FC27 MSP₂₁₄₋₃₄ were identified by overlaying the corresponding spectra (Figure 7.11). Residues that showed chemical shift changes are mapped onto the MSP₂₁₄₋₂₂-bound 6D8 Fv crystal structure (Figure 7.12). The majority of these residues distributed at one side of the protein and only few of them are close to the MSP₁₄₋₂₂ binding site (Figure 7.12A). Ala218, Thr231

and Gly234 are present on the opposite side of the MSP2₁₄₋₂₂ binding site. All the residues (except Val212 and Gly234) that showed changes between two spectra are from CDRs.

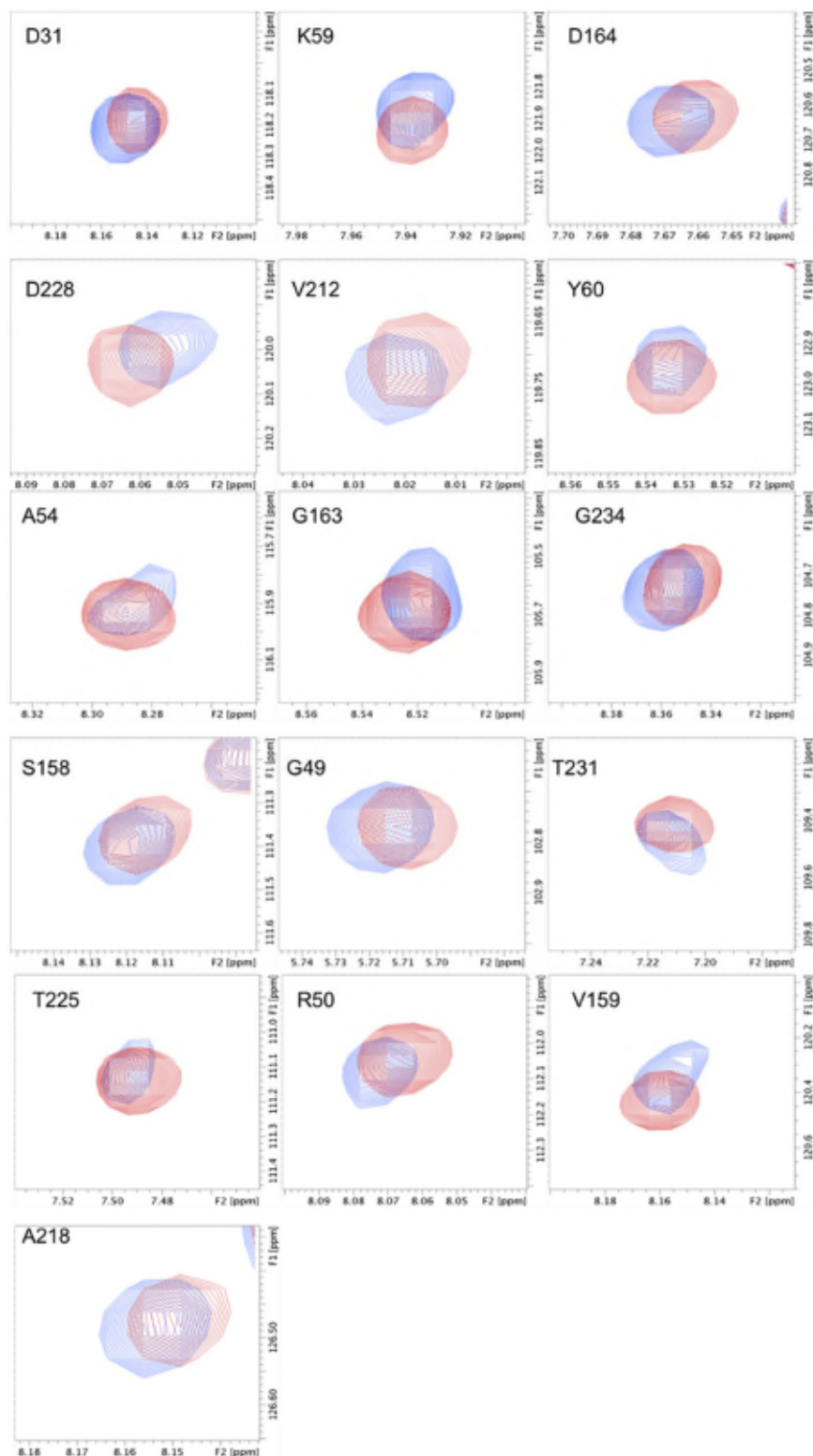


Figure 7.11. Overlay of two-dimensional [^1H - ^{15}N]-TROSY spectra recorded on 6D8 scFv in the presence of 3D7 MSP2₁₄₋₃₄ (blue) and FC27 MSP2₁₄₋₃₄ (red).

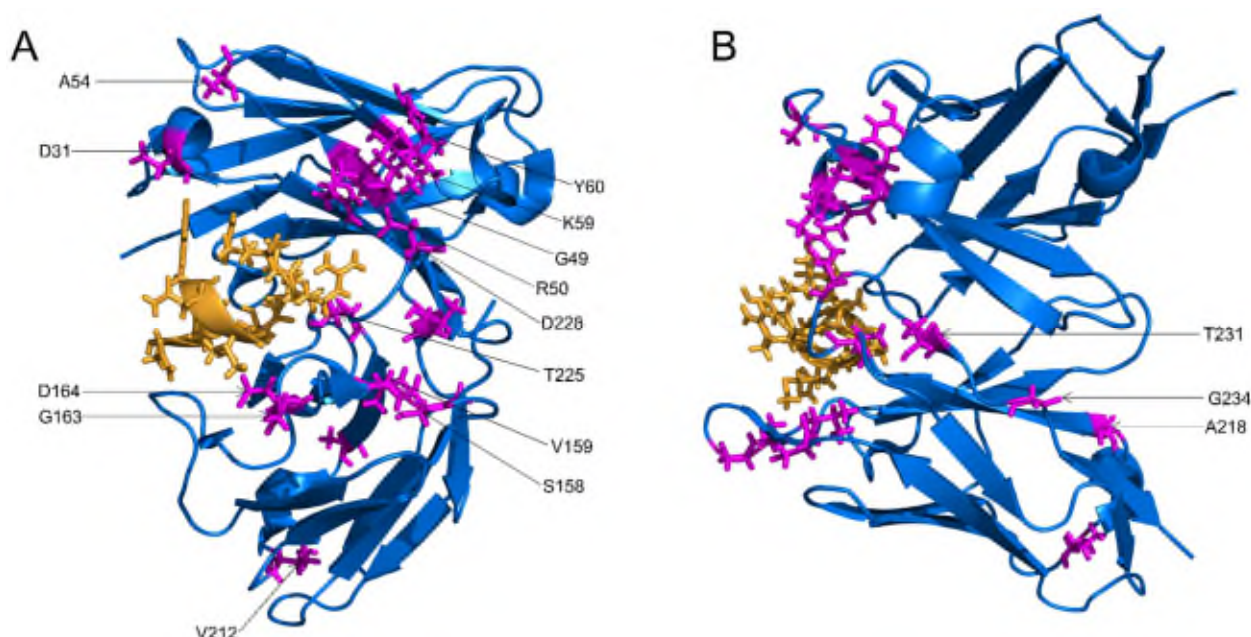


Figure 7.12. Mapping the transient interactions on to the crystal structure of the MSP2₁₄₋₂₂-bound 6D8 Fv. The residues that showed differences in the spectra of 6D8 scFv in the presence of 3D7 MSP2₁₄₋₃₄ and FC27 MSP3₁₄₋₃₄ are coloured magenta. 6D8 Fv (blue) and MSP2₁₄₋₂₂ (orange).

7.4 Discussion

Structural information on protein-protein and protein-ligand interactions can be obtained by either X-ray crystallography or NMR spectroscopy. The former is usually a preferable method as it provides high-resolution structures, despite potentially giving some artefacts [208, 209, 59, 100]. However, many protein-protein/ligand interactions are highly dynamic [210] and are mediated by weak contacts termed ‘transient’ interactions that mostly cannot be observed in crystal structures [211-213, 190]. It is essential to map those transient interactions to understand various biological functions [214]. NMR is a versatile technique, used in understanding a vast number of biological functions [215-217]. It is the method of choice to identify transient interactions in solution [218-221]. NMR chemical shift perturbations are the probes to map the protein-protein or protein-ligand interactions [222, 223]. The information can be easily obtained by analysing the two-dimensional [¹H-¹⁵N]-TROSY spectra of a ligand-free and ligand-bound protein or by comparing the spectra of protein bound to different ligands.

Our previous study (Chapter 6) indicates that 6D8 binds to full-length 3D7 MSP2 and FC27 MSP2 with different affinities, despite recognizing a highly conserved epitope [190]. ITC binding studies confirm that these affinity differences were due to the existence of additional interactions between 6D8 and the variable region residues that are close to the

6D8 epitope. Crystal structures were determined for 3D7 MSP2₁₄₋₃₀ and FC27 MSP2₁₄₋₃₀ epitopes in complex with the 6D8 Fv to identify the residues that mediate the affinity differences. However, these structures were found to be identical with the crystal structure of MSP2₁₄₋₂₂-bound 6D8 Fv. Electron densities for the variable regions of 6D8 epitope were not detected in the crystal structures [190], suggesting that those regions are highly dynamic and have not made any stable crystal contacts with the 6D8 Fv. Preliminary NMR experiments supported the ITC data and confirmed that the affinity differences are due to the existence of transient interactions between 6D8 and the variable regions of MSP2.

In NMR spectroscopy, determining sequential assignments is the first step towards mapping any interactions in proteins. 6D8 scFv is prone to form dimers in solution, and these dimers affect signal intensity in NMR due to their unfavourable relaxation properties (small T_2). To remove these dimers, the protein was purified by SEC. Freshly prepared 6D8 scFv at low concentration (3 mg/mL) was used for SEC to minimize the concentration of dimer and also to improve the resolution in SEC.

Resonance overlap was seen in the HSQC spectrum recorded on protonated 6D8 scFv (**Figure 7.2A**). In general, efficient spin-spin couplings due to the slower tumbling of large proteins affects transverse relaxation rates, which results in line broadening and signal loss [224]. The TROSY technique reduces the signal loss due to increased relaxation rates and thus reduced the line widths by exploiting the relaxation effects arising from chemical shift anisotropy and dipole-dipole interactions [225, 193]. Furthermore, line-broadening due to ^1H - ^1H dipolar relaxation between H^{N} and H^{C} protons can be removed by replacing the H^{C} protons with ^2H by protein deuteration. Hence, TROSY together with deuteration [203] was used in acquiring all the NMR experiments for 6D8 scFv both in the presence and absence of the 6D8 epitope. We observed an increased number of resonances in the spectrum of deuterated 6D8 scFv due to reduced signal loss and improved line widths compared to the spectrum of non-deuterated 6D8 scFv (**Figure 7.2B**). Backbone NMR assignments for 6D8 scFv both in the presence and absence of the 6D8 epitope were determined. Of 220 resonances that were observed in the two-dimensional [^1H - ^{15}N]-TROSY spectrum (**Figure 7.3**), which included the linker sequence, 6His-tag and extra residues that arise from the vector, 190 resonances for MSP2₁₄₋₂₂-free and 180 resonances for MSP2₁₄₋₂₂-bound 6D8 scFv (corresponding to 85% and 80% of observed resonances) were assigned successfully. Assignments are limited to 80-85% due to several factors, including signal overlap, line broadening due to intermediate conformational exchange and the great extent of solvent exchangeable residues identified in CLEANEX spectra (**Figure 7.9**). Larger CSP (>0.05

ppm) values and chemical shift changes for several residues that are away from the MSP2₁₄₋₂₂ binding site indicate a conformational change due to allosteric effects on the other regions of the 6D8 scFv upon binding to the MSP2₁₄₋₂₂ (**Figures 7.6 and 7.8**). This observation is similar to that of many proteins, which, upon binding to their binding partners or ligands, undergo conformational changes as part of the recognition process [222]. This is further supported by the perturbation of solvent-exposed residues observed in the CLEANEX NMR spectrum of MSP2₁₄₋₂₂-bound 6D8 scFv (**Figure 7.9E**).

6D8 scFv residues that form transient interactions with variable regions of MSP2 were identified and mapped onto the crystal structure of MSP2₁₄₋₂₂ bound 6D8 Fv. Some of these residues are far from the MSP2₁₄₋₂₂ binding site, but within the CDRs or close to the CDRs, suggesting that some of these interactions between 6D8 and MSP2 variable regions could be non-specific. Although these are weak interactions, the immunogenicity of the variable region close to the 6D8 epitope might be modulated by these transient interactions [226].

7.5 Conclusion and future work

In conclusion, we have successfully assigned the backbone resonances for 6D8 scFv in the presence and absence of MSP2₁₄₋₂₂ and identified the 6D8 scFv residues that make transient interactions with MSP2 variable regions (**Figure 7.11**). The findings from this work will contribute to a broader understanding of interactions between intrinsically disordered antigens and the human immune system. These results are also relevant to the development of other disordered antigens as vaccine candidates as the *Plasmodium* and many other human pathogens contain numerous disordered antigens. In the context of MSP2 immunogenicity, the efficacy of the variable region epitopes close to the 6D8 epitope might be affected due to the existence of interactions with the high affinity antibody 6D8 (Chapter 6). Increasing the helical propensity of the variable region close to the 6D8 epitope would increase its immunogenicity and have implications in designing an improved MSP2-based vaccine.

This work has identified the residues of 6D8 scFv that interact with the 6D8 extended epitopes on MSP2. Further work will be required for identifying the MSP2 variable region residues that interact with 6D8. Future work will include mapping the extended 6D8 epitope-6D8 scFv intermolecular interactions by NOESY NMR experiments. Moreover, the dynamics of these extended epitopes will be studied by acquiring NMR relaxation experiments on labelled epitopes in the presence and absence of 6D8 scFv. Isotope labelled 6D8 extended epitopes for these experiments will be expressed using an *E. coli* expression

system. Altogether, these findings are expected to have implications in understanding the role of transient interactions in MSP2 immunogenicity, which may be applicable to other intrinsically disordered antigens.

CHAPTER 8

Conclusions and Future Work

The emergence of drug-resistant malaria parasites and lack of an efficacious vaccine highlight the need for the ongoing development of new antimalarials and vaccines. This thesis focused on both of these requirements.

The first part of the thesis aimed at mapping the interactions of small molecules that bind to the *Pf*AMA1 hydrophobic cleft as a basis for developing potential *Pf*AMA1 inhibitors using different chemical approaches. *Pf*AMA1 plays an important role in erythrocyte invasion by *Plasmodium* parasites by forming a moving junction between RBC and parasite in co-ordination with the *Pf*RON2 protein. The *Pf*RON2 binding site on *Pf*AMA1, termed the “hydrophobic cleft”, is central to *Pf*AMA1 function, and therefore ligands that target the *Pf*AMA1-*Pf*RON2 interaction at the parasite-host cell interface are an effective way of inhibiting RBC invasion by the parasite. The initial stage of this project involved preparation of triple-labelled recombinant *Pf*AMA1 (3D7 and FVO) using an *E. coli* BL21 (DE3) expression system. The high-cell-density method of protein expression was used in order to obtain higher protein yields. In this methodology, bacterial cells were adapted to $^2\text{H}_2\text{O}$ by culturing in media prepared using increasing concentrations of $^2\text{H}_2\text{O}$. This effort together with using the high-cell-density method of expression enabled the *E. coli* cells to grow to a high optical density ($\text{OD}_{600} = \sim 5$) in $^2\text{H}_2\text{O}$ -based M9 medium and resulted in high yield expression of isotopically labelled *Pf*AMA1. Furthermore, better refolding results were obtained when ~ 40 mg of unfolded *Pf*AMA1 was included in 1L of refolding buffer. pH 7.0 for 3D7 *Pf*AMA1 and pH 6.8 for FVO *Pf*AMA1 in 20 mM sodium phosphate buffer containing 50 mM each of glutamate and arginine were found to be suitable for NMR experiments. A set of TROSY-based triple-resonance NMR experiments (HNCO, HNCA, HN[CO]CA, HNCACB, HN[CO]CACB and HN[CA]CB) was collected to determine sequential backbone assignments. Different isotope labelling approaches including selective labelling (Lys) and selective unlabelling (Lys, Arg, Val, Leu and Ile) were also applied to support the assignment process. Backbone assignments were confirmed by comparing different spectra of selectively labelled/unlabelled *Pf*AMA1. About 80% of sequential backbone assignments were completed from triple-resonance experiments together with RASP (Resonance Assignment by chemical Shift Prediction).

The first X-ray crystal structure of FVO *Pf*AMA1 was determined to gain insights into the impact of sequence diversity on AMA1 structure and help in the development of strain-transcending inhibitors of *Pf*AMA1. Normalized B-factor analysis together with MD simulations were used for understanding the roles of the Ia-f and DII loops flexibility in limiting the cross-reactivity of human antibody responses and inhibitory peptides against

different AMA1 alleles. Overall, the structure, including the flexible DII loop, is conserved between 3D7 and FVO forms of *Pf*AMA1, suggesting that the sequence divergence does not result in structural changes that account for the strain-specific effects of immune response and inhibitory peptide. The interacting interfaces between DI and DII consist of conserved residues across all *P. falciparum* strains. As the DII loop end of the hydrophobic cleft is conserved across all allelic forms of AMA1, this region represents an attractive site for developing strain-transcending small molecule inhibitors of *Pf*AMA1.

Small molecules that showed binding with *Pf*AMA1 in R1 competition experiments were selected for their affinity measurements by SPR and binding site mapping by NMR. The binding affinities for benzimidazoles, pyrazoles and 2-aminothiazoles were evaluated using SPR and it was found that benzimidazoles bound more strongly compared to pyrazoles and 2-aminothiazoles. 2D [¹H-¹⁵N]-TROSY experiments were acquired in the presence and absence of these small molecules to map their binding sites. Chemical shift perturbations were calculated based on the chemical shift changes observed in NMR spectra. The binding sites for benzimidazoles were mapped onto the FVO *Pf*AMA1 crystal structure and it was found that their binding sites overlap with the binding sites for R1 and RON2 peptides. The similar perturbations observed in the spectra of FVO *Pf*AMA1 and 3D7 *Pf*AMA1 imply that FVO and 3D7 *Pf*AMA1 have very similar binding sites for these compounds.

Future directions of this project will focus on elaborating the benzimidazoles to improve their binding affinities for *Pf*AMA1. This may be approached by either incorporating extra functional groups on this scaffold or by linking with other small molecules that bind close to the benzimidazoles binding site. The specific interactions of the elaborated small molecules with *Pf*AMA1 will be determined rapidly by 2D [¹H-¹⁵N]-TROSY NMR experiments. SPR binding experiments will be used to determine the affinities of small molecules with *Pf*AMA1. Additionally, small molecules with higher affinities will be subjected to X-ray crystallography to define their binding pose on *Pf*AMA1. Finally, small molecules identified in this study will be tested in *in vitro* invasion assays to determine their inhibitory efficacy.

The second part of this thesis focused on developing an MSP2-based vaccine construct that can generate a strain-transcending immune response to different *P. falciparum* strains.

The project started with optimizing the expression yields for 3D7 MSP2 using an *E. coli* expression system. The plasmid for 3D7 MSP2 was designed using a pET32a vector containing the thioredoxin as a fusion domain and achieved final protein yields up to 10-12 mg/L M9 medium. Isotope-labelled MSP2 (3D7 and FC27) was expressed in *E. coli* BL21

(DE3) and purified using various chromatography techniques. The MSP2 spectrum was highly overlapped owing to minimal chemical shift dispersion, indicating its extensive disordered conformation. Triple-resonance experiments particularly designed for IDPs including 5D HN(CA)CONH, 5D HabCabCONH were employed to determine sequential assignments for 3D7 MSP2. Relaxation rates calculated from ^{15}N relaxation measurements were uniformly low along the sequence, as expected for IDPs. Positive ^1H - ^{15}N NOEs were observed for three conformationally restricted regions: *N*-, *C*- and a part of FC27-specific dimorphic sequence. These measurements also suggested that the 3D7-specific GGSA repeats were highly flexible compared to other regions of MSP2. We found a strong correlation between the conformational dynamics of MSP2 and its antigenicity as defined using antisera to recombinant MSP2. The regions of restricted conformational order were significantly more antigenic compared to conformationally non-restricted regions of MSP2. This suggests that the immunogenic properties can be improved by reducing the flexibility of MSP2 epitopes. The flexibility of the MSP2 epitopes could be reduced by introducing a bulky residue or a disulfide bridge connecting the epitope ends [227]. Alternatively, the *N*-terminal epitopes can be stabilized by lipid interactions in such a way as to mimic the native MSP2 conformation [145]. It may also be possible to tune the immune response to MSP2 towards its conserved and exposed epitopes on the merozoite surface.

MSP2 elicited protection in the Combination B vaccine trial, albeit in a strain-specific manner owing to the polymorphic nature of MSP2 and poor immunogenicity of its conserved regions. The limitation of strain-specificity can be overcome by either using a vaccine formulation consisting of a mixture of major alleles (3D7 and FC27) of MSP2 or using a chimeric construct covering both 3D7 and FC27 MSP2 sequence epitopes. Here we used the second approach and generated a group of chimeric MSP2 constructs in an effort to achieve a strain-transcending immune response to MSP2. Additionally, truncated forms of MSP2 lacking either *N*- or *C*- or both conserved regions were generated to study their role in immunogenicity and aggregation propensity of MSP2. Key advantages of MSP2 chimeras are that protein refolding is not required and their production is simpler than for two separate constructs. Furthermore, MSP2 variable regions are essentially dimorphic and the two major alleles, 3D7 and FC27, can be accommodated easily into a single chimeric construct. The new constructs were expressed, purified and characterized by LC-MS, analytical HPLC, size-exclusion chromatography and endotoxin assay. The lower aggregation propensity observed for MSP2 chimeras could be due to the rearrangement of sequences in the constructs.

The regions of all MSP2 constructs that were immunogenic are similar to what has been reported for MSP2 previously, suggesting that the conformational properties of MSP2 epitopes within the chimeras were mostly preserved [228, 229, 160]. However, we observed some context-dependent effects with the inclusion of conserved *N*- and *C*-terminal regions on the overall immunogenicity of the chimeras and the fine specificity of the response they elicit. These effects could be mediated by the T-cell epitopes present within the MSP2 conserved regions [230, 231]. Similarly, removal of one 32-residue repeat diminished the antibody response to the remaining 32-residue repeat in the NV_{mFC27}V_{m3D7}C chimera. Interestingly, in the same construct, the removal of 3D7-specific repeats boosted the response to remaining 3D7-specific epitopes. These observations suggested that the repeat sequences in MSP2 modulate the immunogenicity of nearby epitopes. Another reason for a change in the immune response to truncated FC27 MSP2 and NV_{mFC27}V_{m3D7}C constructs could be a conformational change in variable regions and repeat sequences upon the removal of conserved regions and nearby sequences, respectively. We observed large chemical shift changes for a few resonances of VR epitopes in the NMR spectrum of truncated FC27 MSP2 (Appendix III). However, NMR assignments for the truncated FC27 MSP2 constructs are unavailable. It is, thus, unclear whether these spectral changes indicate any conformational changes in MSP2 variable regions or not. Future work will be focused on determining the NMR assignments for the truncated FC27 MSP2 constructs to address this question. Overall, total IgG levels were similar for all the constructs, while the V_{3D7}V_{FC27}C construct elicited higher IgG2b and IgG2c subtypes responses without affecting the total IgG titre. Understanding the effector mechanisms responsible for protection mediated by the MSP2 IgG subtypes will enable the design of an optimal MSP2-based construct that can be included in a future malaria vaccine.

It appears that the GPI-anchor of MSP2 can be mimicked by adding an extra sequence to the *C*-terminus, which has a propensity of binding with lipids and could help in the formulation of MSP2 in lipids. In another approach, a complex of MSP2 with *Pf*AMA1 can be made by generating a MSP2-*Pf*RON2 chimera. Here the *Pf*RON2 mimics the *C*-terminal GPI-anchor of MSP2 when the *Pf*RON2 binds to *Pf*AMA1. Thus the chimera work was extended to the design and preparation of a complex of MSP2-*Pf*RON2 chimera and *Pf*AMA1. Chimeras of both 3D7 MSP2-*Pf*RON2 and FC27 MSP2-*Pf*RON2 constructs were designed and initial work focused on preparation of the 3D7 MSP2-*Pf*RON2 complex with 3D7 *Pf*AMA1 and FVO *Pf*AMA1. The 3D7 MSP2-*Pf*RON2 chimera was successfully expressed in *E. coli* and purified using different chromatography techniques. Complexes of

3D7 MSP2-*Pf*RON2 with 3D7 *Pf*AMA1 and FVO *Pf*AMA1 were prepared by using a *Pf*AMA1-based affinity column. We found that this complex was unstable as suggested by size-exclusion chromatography, and further optimization will be required to improve the stability of this complex before proceeding to immunization experiments. These initial results are described in Appendix IV.

Our MSP2 work also focused on determining the structure of antibody-bound epitopes. Such structures would help in understanding the conformation of native MSP2, and stabilizing these structures in recombinant MSP2 might improve the immunogenic efficacy of MSP2. 6D8 is a monoclonal antibody recognizing an epitope within the conserved *N*-terminus of MSP2. The structure of the 6D8-bound epitope is incompatible with the structure of the lipid-bound conformation of that region. This suggests that the *N*-terminal region of MSP2 is interacting with the parasite membrane and is not accessible to antibodies raised against recombinant MSP2, thus suggesting a mechanism by which the parasite MSP2 is able to escape recognition from antibodies. On the other hand, although 6D8 binds to a highly conserved region (MSP2₁₄₋₂₂), it has different affinities for recombinant 3D7 and FC27 MSP2, as confirmed by ITC binding experiments. Binding experiments carried out with the 6D8 epitopes containing just the first five residues of the variable region of each allele (3D7 and FC27 MSP2₁₄₋₃₀) also showed similar results. This suggests that the affinity differences are due to extra interactions between 6D8 and variable region residues. While crystal structures for 6D8 scFv in complex with the 6D8 extended epitopes did not explain the basis for these affinity differences, two-dimensional [¹H-¹⁵N]-TROSY experiments suggested that the affinity differences were due to the existence of transient interactions between 6D8 and residues from the variable region that are close to the 6D8 epitope. NMR assignments of 6D8 scFv were determined in the presence and absence of MSP2₁₄₋₂₂ and the 6D8 residues that form transient interactions with the MSP2 variable regions were identified using these assignments. NOESY experiments will be recorded on the extended epitopes in complex with 6D8 scFv in order to identify intermolecular peptide-protein NOEs. Moreover, NMR relaxation experiments on these epitopes in the presence and absence of 6D8 scFv will be acquired to understand the dynamics of the epitopes. These findings will contribute to understanding the role of transient interactions in the immunogenicity of MSP2 and many other intrinsically disordered antigens.

In summary, the results from my studies of *Pf*AMA1-small molecule interactions presented in this thesis provide the basis for the future development of potential *Pf*AMA1

inhibitors as new antimalarials. Similarly, the findings from my studies on MSP2 provide a platform for developing an efficacious malaria vaccine.

CHAPTER 9

References

1. WHO (2015) World malaria report. 1-242.
2. Sachs J., and Malaney P. (2002) The economic and social burden of malaria. *Nature* 415, 680-685.
3. Kwiatkowski D. (1995) Malarial toxins and the regulation of parasite density. *Parasitol. Today* 11, 206-212.
4. Winzeler E. A. (2006) Applied systems biology and malaria. *Nat. Rev. Microbiol.* 4, 145-151.
5. Egan T. J. (2008) Haemozoin formation. *Mol. Biochem. Parasitol.* 157, 127-136.
6. Craft J. C. (2008) Challenges facing drug development for malaria. *Curr. Opin. Microbiol.* 11, 428-433.
7. Petersen I., Eastman R., and Lanzer M. (2011) Drug-resistant malaria: molecular mechanisms and implications for public health. *FEBS Lett.* 585, 1551-1562.
8. Manohar S., Khan S. I., and Rawat D. S. (2013) 4-Aminoquinoline-triazine-based hybrids with improved in vitro antimalarial activity against CQ-sensitive and CQ-resistant strains of *Plasmodium falciparum*. *Chem. Biol. Drug Des.* 81, 625-630.
9. Kappe S. H. I., Vaughan A. M., Boddey J. A., and Cowman A. F. (2010) That was then but this is now: malaria research in the time of an eradication agenda. *Science* 328, 862-866.
10. Wells T. N. C., Alonso P. L., and Gutteridge W. E. (2009) New medicines to improve control and contribute to the eradication of malaria. *Nat. Rev. Drug Discov.* 8, 879-891.
11. Cammack N. (2011) Exploiting malaria drug resistance to our advantage. *Science* 333, 705-706.
12. Dondorp A. M., Nosten F., Yi P., Das D., Phyo A. P., *et al.* (2009) Artemisinin resistance in *Plasmodium falciparum* malaria. *N. Engl. J. Med.* 361, 455-467.
13. Noedl H., Se Y., Schaefer K., Smith B. L., Socheat D., *et al.* (2008) Evidence of artemisinin-resistant malaria in Western Cambodia. *N. Engl. J. Med.* 359, 2619-2620.
14. WHO (2010) Guidelines for the treatment of malaria. *Geneve: WHO.*
15. World Health Organization (WHO). (2011) Malaria, **2011**; All informations available from <http://www.who.int/mediacentre/factsheets/fs094/en/index.html>.
16. Burrows J. N., Chibale K., and Wells T. N. C. (2011) The state of the art in anti-malarial drug discovery and development. *Curr. Top. Med. Chem.* 11, 1226-1254.
17. Ajayi N. A., and Ukwaja K. N. (2013) Possible artemisinin-based combination therapy-resistant malaria in Nigeria: a report of three cases. *Rev. Soc. Bras. Med. Trop.* 46, 525-527.
18. Amaratunga C., Lim P., Suon S., Sreng S., Mao S., *et al.* (2016) Dihydroartemisinin-piperaquine resistance in *Plasmodium falciparum* malaria in Cambodia: a multisite prospective cohort study. *Lancet Infect. Dis.* 16, 357-365.
19. Ashley E. A., Dhorda M., Fairhurst R. M., Amaratunga C., Lim P., *et al.* (2014) Spread of artemisinin resistance in *Plasmodium falciparum* malaria. *N. Engl. J. Med.* 371, 411-423.
20. Breman J. G. (2012) Resistance to artemisinin-based combination therapy. *Lancet Infect. Dis.* 12, 820-822.
21. Duffy P. E., and Sibley C. H. (2005) Are we losing artemisinin combination therapy already? *The Lancet* 366, 1908-1909.
22. Hastings I. M., and Ward S. A. (2005) Coartem (artemether-lumefantrine) in Africa: the beginning of the end? *J. Infect. Dis.* 192, 1303-1304.

23. Noedl H., Se Y., Schaecher K., Smith B. L., Socheat D., *et al.* (2008) Evidence of artemisinin-resistant malaria in Western Cambodia. *N. Engl. J. Med.* 359, 2619-2620.
24. Phyo A. P., Nkhoma S., Stepniewska K., Ashley E. A., Nair S., *et al.* Emergence of artemisinin-resistant malaria on the western border of Thailand: a longitudinal study. *The Lancet* 379, 1960-1966.
25. Enserink M. (2010) Malaria's drug miracle in danger. *Science* 328, 844-846.
26. Offeddu V., Thathy V., Marsh K., and Matuschewski K. (2012) Naturally acquired immune responses against *Plasmodium falciparum* sporozoites and liver infection. *Int. J. Parasitol.* 42, 535-548.
27. Ouédraogo A. L., Roeffen W., Luty A. J. F., de Vlas S. J., Nebie I., *et al.* (2011) Naturally acquired immune responses to *Plasmodium falciparum* sexual stage antigens Pfs48/45 and Pfs230 in an area of seasonal transmission. *Infect. Immun.* 79, 4957-4964.
28. Wipasa J., Elliott S., Xu H., and Good M. F. (2002) Immunity to asexual blood stage malaria and vaccine approaches. *Immunol. Cell Biol.* 80, 401-414.
29. Bertolino P., and Bowen D. G. (2015) Malaria and the liver: immunological hide-and-seek or subversion of immunity from within? *Front. in Microbiol.* 6, 41.
30. Overstreet M. G., Cockburn I. A., and Zavala F. (2008) Protective CD8(+) T cells against *Plasmodium* liver stages: immunobiology of an 'unnatural' immune response. *Immunol. Rev.* 225, 272-283.
31. Van Braeckel-Budimir N., and Harty J. T. (2014) CD8 T-cell-mediated protection against liver-stage malaria: lessons from a mouse model. *Front. in Microbiol.* 5, 272.
32. Weiss W. R., Sedegah M., Beaudoin R. L., Miller L. H., and Good M. F. (1988) CD8+ T cells (cytotoxic/suppressors) are required for protection in mice immunized with malaria sporozoites. *Proc. Natl. Acad. Sci. U.S.A.* 85, 573-576.
33. Agnandji S. T., Lell B., Soulanoudjingar S. S., Fernandes J. F., Abossolo B. P., *et al.* (2011) First results of Phase 3 trial of RTS,S/AS01 malaria vaccine in African children. *N. Engl. J. Med.* 365, 1863-1875.
34. (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. *The Lancet* 386, 31-45.
35. WHO (2016) Malaria vaccine: WHO position paper. 91, 33-52.
36. Agnandji S. T., Lell B., Fernandes J. F., Abossolo B. P., Methogo B. G. N. O., *et al.* (2012) A Phase 3 trial of RTS,S/AS01 malaria vaccine in African infants. *N. Engl. J. Med.* 367, 2284-2295.
37. Schwartz L., Brown G., Genton B., and Moorthy V. (2012) A review of malaria vaccine clinical projects based on the WHO rainbow table. *Malar. J.* 11, 11.
38. Anders R. F., Adda C. G., Foley M., and Norton R. S. (2010) Recombinant protein vaccines against the asexual blood stages of *Plasmodium falciparum*. *Hum. Vaccin.* 6, 39-53.
39. Flück C., Schöpflin S., Smith T., Genton B., Alpers M. P., *et al.* (2007) Effect of the malaria vaccine Combination B on merozoite surface antigen 2 diversity. *Infect. Genet. Evol.* 7, 44-51.
40. Genton B., Betuela I., Felger I., Al-Yaman F., Anders R. F., *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, 820-827.

41. Thera M. A., Doumbo O. K., Coulibaly D., Laurens M. B., Ouattara A., *et al.* (2011) A field trial to assess a blood-stage malaria vaccine. *N. Engl. J. Med.* 365, 1004-1013.
42. Florens L., Washburn M. P., Raine J. D., Anthony R. M., Grainger M., *et al.* (2002) A proteomic view of the *Plasmodium falciparum* life cycle. *Nature* 419, 520-526.
43. Plassmeyer M. L., Reiter K., Shimp R. L., Kotova S., Smith P. D., *et al.* (2009) Structure of the *Plasmodium falciparum* circumsporozoite protein, a leading malaria vaccine candidate. *J. Biol. Chem.* 284, 26951-26963.
44. Fenton B., Clark J. T., Khan C. M., Robinson J. V., Walliker D., *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, 963-971.
45. Gilson P. R., Nebl T., Vukcevic D., Moritz R. L., Sargeant T., *et al.* (2006) Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite *Plasmodium falciparum*. *Mol. Cell. Proteomics* 5, 1286-1299.
46. Smythe J. A., Coppel R. L., Day K. P., Martin R. K., Oduola A. M., *et al.* (1991) Structural diversity in the *Plasmodium falciparum* merozoite surface antigen 2. *Proc. Natl. Acad. Sci. U.S.A.* 88, 1751-1755.
47. Smythe J. A., Peterson M. G., Coppel R. L., Saul A. J., Kemp D. J., *et al.* (1990) Structural diversity in the 45-kilodalton merozoite surface antigen of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 39, 227-234.
48. Mitchell G. H., Thomas A. W., Margos G., Dluzewski A. R., and Bannister L. H. (2004) Apical membrane antigen 1, a major malaria vaccine candidate, mediates the close attachment of invasive merozoites to host red blood cells. *Infect. Immun.* 72, 154-158.
49. Riglar D. T., Richard D., Wilson D. W., Boyle M. J., Dekiwadia C., *et al.* (2011) Super-resolution dissection of coordinated events during malaria parasite invasion of the human erythrocyte. *Cell Host Microbe* 9, 9-20.
50. Schussek S., Trieu A., Apte S. H., Sidney J., Sette A., *et al.* (2013) Immunization with apical membrane antigen 1 confers sterile infection-blocking immunity against *Plasmodium* sporozoite challenge in a rodent model. *Infect. Immun.* 81, 3586-3599.
51. Silvie O., Franetich J.-F., Charrin S., Mueller M. S., Siau A., *et al.* (2004) A role for apical membrane antigen 1 during invasion of hepatocytes by *Plasmodium falciparum* sporozoites. *J. Biol. Chem.* 279, 9490-9496.
52. MacRaild C. A., Anders R. F., Foley M., and Norton R. S. (2011) Apical membrane antigen 1 as an anti-malarial drug target. *Curr. Top. Med. Chem.* 11, 2039-2047.
53. Barry A. E., and Arnott A. (2014) Strategies for designing and monitoring malaria vaccines targeting diverse antigens. *Front. Immunol.* 5, 1-16.
54. Ouattara A., Barry A. E., Dutta S., Remarque E. J., Beeson J. G., *et al.* (2015) Designing malaria vaccines to circumvent antigen variability. *Vaccine* 33, 7506-7512.
55. Sutherland C. (2007) A challenge for the development of malaria vaccines: polymorphic target antigens. *PLoS Med.* 4, e116.
56. Takala S. L., Coulibaly D., Thera M. A., Batchelor A. H., Cummings M. P., *et al.* (2009) Extreme polymorphism in a vaccine antigen and risk of clinical malaria: implications for vaccine development. *Sci. Transl. Med.* 1, 2ra5-2ra5.
57. Takala S. L., and Plowe C. V. (2009) Genetic diversity and malaria vaccine design, testing, and efficacy: Preventing and overcoming “vaccine resistant malaria”. *Parasite Immunol.* 31, 560-573.

58. Tonkin M. L., Roques M., Lamarque M. H., Pugn  re M., Douguet D., *et al.* (2011) Host cell invasion by apicomplexan parasites: insights from the co-structure of AMA1 with a RON2 peptide. *Science* 333, 463-467.
59. Vulliez-Le Normand B., Tonkin M. L., Lamarque M. H., Langer S., Hoos S., *et al.* (2012) Structural and functional insights into the malaria parasite moving junction complex. *PLoS Pathog.* 8, e1002755.
60. Hehl A. B., Lekutis C., Grigg M. E., Bradley P. J., Dubremetz J.-F., *et al.* (2000) *Toxoplasma gondii* homologue of *Plasmodium* apical membrane antigen 1 is involved in invasion of host cells. *Infect. Immun.* 68, 7078-7086.
61. Triglia T., Healer J., Caruana S. R., Hodder A. N., Anders R. F., *et al.* (2000) Apical membrane antigen 1 plays a central role in erythrocyte invasion by *Plasmodium* species. *Mol. Microbiol.* 38, 706-718.
62. Yap A., Azevedo M. F., Gilson P. R., Weiss G. E., O'Neill M. T., *et al.* (2014) Conditional expression of apical membrane antigen 1 in *Plasmodium falciparum* shows it is required for erythrocyte invasion by merozoites. *Cell. Microbiol.* 16, 642-656.
63. Mital J., Meissner M., Soldati D., and Ward G. E. (2005) Conditional expression of toxoplasma gondii apical membrane antigen-1 (TgAMA1) demonstrates that TgAMA1 plays a critical role in host cell invasion. *Mol. Biol. Cell* 16, 4341-4349.
64. Collins C. R., Withers-Martinez C., Hackett F., and Blackman M. J. (2009) An inhibitory antibody blocks interactions between components of the malarial invasion machinery. *PLoS Pathog.* 5, e1000273.
65. Henderson K. A., Streltsov V. A., Coley A. M., Dolezal O., Hudson P. J., *et al.* (2007) Structure of an IgNAR-AMA1 complex: targeting a conserved hydrophobic cleft broadens malarial strain recognition. *Structure* 15, 1452-1466.
66. Bargieri D. Y., Andenmatten N., Lagal V., Thiberge S., Whitelaw J. A., *et al.* (2013) Apical membrane antigen 1 mediates apicomplexan parasite attachment but is dispensable for host cell invasion. *Nat. Commun.* 4, 2552.
67. Giovannini D., Sp  th S., Lacroix C., Perazzi A., Bargieri D., *et al.* (2011) Independent roles of apical membrane antigen 1 and rhoptry neck proteins during host cell invasion by apicomplexa. *Cell Host Microbe* 10, 591-602.
68. Lamarque M. H., Roques M., Kong-Hap M., Tonkin M. L., Rugarabamu G., *et al.* (2014) Plasticity and redundancy among AMA-RON pairs ensure host cell entry of *Toxoplasma* parasites. *Nat. Commun.* 5, 4098.
69. Bargieri D., Lagal V., Andenmatten N., Tardieux I., Meissner M., *et al.* (2014) Host cell invasion by apicomplexan parasites: the junction conundrum. *PLoS Pathog.* 10, e1004273.
70. Cowman A. F., Berry D., and Baum J. (2012) The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *J. Cell Biol.* 198, 961-971.
71. Farrow R. E., Green J., Katsimitsoulia Z., Taylor W. R., Holder A. A., *et al.* (2011) The mechanism of erythrocyte invasion by the malarial parasite, *Plasmodium falciparum*. *Semin. Cell Dev. Biol.* 22, 953-960.
72. Bannister L. H., Hopkins J. M., Dluzewski A. R., Margos G., Williams I. T., *et al.* (2003) *Plasmodium falciparum* apical membrane antigen 1 (PfAMA-1) is translocated within micronemes along subpellicular microtubules during merozoite development. *J. Cell Sci.* 116, 3825-3834.
73. Cao J., Kaneko O., Thongkukiatkul A., Tachibana M., Otsuki H., *et al.* (2009) Rhoptry neck protein RON2 forms a complex with microneme protein AMA1 in *Plasmodium falciparum* merozoites. *Parasitol. Int.* 58, 29-35.

74. Lamarque M., Besteiro S., Papoin J., Roques M., Normand B. V.-L., *et al.* (2011) The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites. *PLoS Pathog.* 7, e1001276.
75. Srinivasan P., Beatty W. L., Diouf A., Herrera R., Ambroggio X., *et al.* (2011) Binding of *Plasmodium* merozoite proteins RON2 and AMA1 triggers commitment to invasion. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13275-13280.
76. Tyler J. S., and Boothroyd J. C. (2011) The C-terminus of *Toxoplasma* RON2 provides the crucial link between AMA1 and the host-associated invasion complex. *PLoS Pathog.* 7, e1001282.
77. Besteiro S., Michelin A., Poncet J., Dubremetz J.-F., and Lebrun M. (2009) Export of a *Toxoplasma gondii* rhoptry neck protein complex at the host cell membrane to form the moving junction during invasion. *PLoS Pathog.* 5, e1000309.
78. Peterson M. G., Marshall V. M., Smythe J. A., Crewther P. E., Lew A., *et al.* (1989) Integral membrane protein located in the apical complex of *Plasmodium falciparum*. *Mol. Cell. Biol.* 9, 3151-3154.
79. Hodder A. N., Crewther P. E., Matthew M. L. S. M., Reid G. E., Moritz R. L., *et al.* (1996) The disulfide bond structure of *Plasmodium* apical membrane antigen-1. *J. Biol. Chem.* 271, 29446-29452.
80. Pizarro J. C., Normand B. V.-L., Chesne-Seck M.-L., Collins C. R., Withers-Martinez C., *et al.* (2005) Crystal structure of the malaria vaccine candidate apical membrane antigen 1. *Science* 308, 408-411.
81. Remarque E. J., Faber B. W., Kocken C. H. M., and Thomas A. W. (2008) Apical membrane antigen 1: a malaria vaccine candidate in review. *Trends Parasitol.* 24, 74-84.
82. Feng Z.-P., Keizer D. W., Stevenson R. A., Yao S., Babon J. J., *et al.* (2005) Structure and inter-domain interactions of domain II from the blood-stage malarial protein, apical membrane antigen 1. *J. Mol. Biol.* 350, 641-656.
83. Nair M., Hinds M. G., Coley A. M., Hodder A. N., Foley M., *et al.* (2002) Structure of domain III of the blood-stage malaria vaccine candidate, *Plasmodium falciparum* apical membrane antigen 1 (AMA1). *J. Mol. Biol.* 322, 741-753.
84. Bai T., Becker M., Gupta A., Strike P., Murphy V. J., *et al.* (2005) Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12736-12741.
85. Tordai H., Bányai L., and Patthy L. (1999) The PAN module: the N-terminal domains of plasminogen and hepatocyte growth factor are homologous with the apple domains of the prekallikrein family and with a novel domain found in numerous nematode proteins. *FEBS Lett.* 461, 63-67.
86. Chesne-Seck M.-L., Pizarro J. C., Normand B. V.-L., Collins C. R., Blackman M. J., *et al.* (2005) Structural comparison of apical membrane antigen 1 orthologues and paralogues in apicomplexan parasites. *Mol. Biochem. Parasitol.* 144, 55-67.
87. Ge X., MacRaid C. A., Devine S. M., Debono C. O., Wang G., *et al.* (2014) Ligand-induced conformational change of *Plasmodium falciparum* AMA1 detected using ¹⁹F NMR. *J. Med. Chem.* 57, 6419-6427.
88. Harris K. S., Casey J. L., Coley A. M., Masciantonio R., Sabo J. K., *et al.* (2005) Binding hot spot for invasion inhibitory molecules on *Plasmodium falciparum* apical membrane antigen 1. *Infect. Immun.* 73, 6981-6989.
89. Crawford J., Tonkin M. L., Grujic O., and Boulanger M. J. (2010) Structural characterization of apical membrane antigen 1 (AMA1) from *Toxoplasma gondii*. *J. Biol. Chem.* 285, 15644-15652.

90. Dutta S., Lee S. Y., Batchelor A. H., and Lanar D. E. (2007) Structural basis of antigenic escape of a malaria vaccine candidate. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12488-12493.
91. Healer J., Murphy V., Hodder A. N., Masciantonio R., Gemmill A. W., *et al.* (2004) Allelic polymorphisms in apical membrane antigen-1 are responsible for evasion of antibody-mediated inhibition in *Plasmodium falciparum*. *Mol. Microbiol.* 52, 159-168.
92. Ouattara A., Takala-Harrison S., Thera M. A., Coulibaly D., Niangaly A., *et al.* (2013) Molecular basis of allele-specific efficacy of a blood-stage malaria vaccine: vaccine development implications. *J. Infect. Dis.* 207, 511-519.
93. Collins C. R., Withers-Martinez C., Bentley G. A., Batchelor A. H., Thomas A. W., *et al.* (2007) Fine mapping of an epitope recognized by an invasion-inhibitory monoclonal antibody on the malaria vaccine candidate apical membrane antigen 1. *J. Biol. Chem.* 282, 7431-7441.
94. Nuttall S. D., Humberstone K. S., Krishnan U. V., Carmichael J. A., Doughty L., *et al.* (2004) Selection and affinity maturation of IgNAR variable domains targeting *Plasmodium falciparum* AMA1. *Proteins: Struct., Funct., Bioinf.* 55, 187-197.
95. Coley A. M., Parisi K., Masciantonio R., Hoeck J., Casey J. L., *et al.* (2006) The most polymorphic residue on *Plasmodium falciparum* apical membrane antigen 1 determines binding of an invasion-inhibitory antibody. *Infect. Immun.* 74, 2628-2636.
96. Coley A. M., Gupta A., Murphy V. J., Bai T., Kim H., *et al.* (2007) Structure of the malaria antigen AMA1 in complex with a growth-inhibitory antibody. *PLoS Pathog.* 3, e138.
97. Lee E. F., Yao S., Sabo J. K., Fairlie W. D., Stevenson R. A., *et al.* (2011) Peptide inhibitors of the malaria surface protein, apical membrane antigen 1: identification of key binding residues. *Biopolymers* 95, 354-364.
98. Li F., Dluzewski A., Coley A. M., Thomas A., Tilley L., *et al.* (2002) Phage-displayed peptides bind to the malarial protein apical membrane antigen-1 and inhibit the merozoite invasion of host erythrocytes. *J. Biol. Chem.* 277, 50303-50310.
99. Richard D., MacRaild C. A., Riglar D. T., Chan J.-A., Foley M., *et al.* (2010) Interaction between *Plasmodium falciparum* apical membrane antigen 1 and the rhoptry neck protein complex defines a key step in the erythrocyte invasion process of malaria parasites. *J. Biol. Chem.* 285, 14815-14822.
100. Wang G., MacRaild C. A., Mohanty B., Mobli M., Cowieson N. P., *et al.* (2014) Molecular insights into the interaction between *Plasmodium falciparum* apical membrane antigen 1 and an invasion-inhibitory peptide. *PLoS ONE* 9, e109674.
101. Harris K. S., Casey J. L., Coley A. M., Karas J. A., Sabo J. K., *et al.* (2009) Rapid optimization of a peptide inhibitor of malaria parasite invasion by comprehensive N-methyl scanning. *J. Biol. Chem.* 284, 9361-9371.
102. Sanchez C. P., Dave A., Stein W. D., and Lanzer M. (2010) Transporters as mediators of drug resistance in *Plasmodium falciparum*. *Int. J. Parasitol.* 40, 1109-1118.
103. Lim S. S., Debono C. O., MacRaild C. A., Chandrashekar I. R., Dolezal O., *et al.* (2013) Development of inhibitors of *Plasmodium falciparum* apical membrane antigen 1 based on fragment screening. *Aust. J. Chem.* 66, 1530-1536.
104. Srinivasan P., Yasgar A., Luci D. K., Beatty W. L., Hu X., *et al.* (2013) Disrupting malaria parasite AMA1-RON2 interaction with a small molecule prevents erythrocyte invasion. *Nat. Commun.* 4, 2261.
105. Devine S. M., Lim S. S., Chandrashekar I. R., MacRaild C. A., Drew D. R., *et al.* (2014) A critical evaluation of pyrrolo[2,3-d]pyrimidine-4-amines as *Plasmodium*

- falciparum* apical membrane antigen 1 (AMA1) inhibitors. *Med. Chem. Comm.* 5, 1500-1506.
106. Singh N., Chevé G., Ferguson D., and McCurdy C. (2006) A combined ligand-based and target-based drug design approach for G-protein coupled receptors: application to salvinorin A, a selective kappa opioid receptor agonist. *J. Comput. Aided Mol. Des.* 20, 471-493.
 107. Hajduk P. J., and Greer J. (2007) A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* 6, 211-219.
 108. Macarron R. (2006) Critical review of the role of HTS in drug discovery. *Drug Discov. Today* 11, 277-279.
 109. Sink R., Gobec S., Pecar S., and Zega A. (2010) False positives in the early stages of drug discovery. *Curr. Med. Chem.* 17, 4231-4255.
 110. Congreve M., Chessari G., Tisi D., and Woodhead A. J. (2008) Recent developments in fragment-based drug discovery. *J. Med. Chem.* 51, 3661-3680.
 111. Erlanson D. A., McDowell R. S., and O'Brien T. (2004) Fragment-based drug discovery. *J. Med. Chem.* 47, 3463-3482.
 112. Rees D. C., Congreve M., Murray C. W., and Carr R. (2004) Fragment-based lead discovery. *Nat. Rev. Drug Discov.* 3, 660-672.
 113. Shuker S. B., Hajduk P. J., Meadows R. P., and Fesik S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 274, 1531-1534.
 114. Reymond J.-L., and Awale M. (2012) Exploring chemical space for drug discovery using the chemical universe database. *ACS Chem. Neurosci.* 3, 649-657.
 115. van Deursen R., and Reymond J.-L. (2007) Chemical space travel. *ChemMedChem* 2, 636-640.
 116. Stockman B. J., and Dalvit C. (2002) NMR screening techniques in drug discovery and drug design. *Prog. Nucl. Magn. Reson. Spectrosc.* 41, 187-231.
 117. Mugenyi C. K., Elliott S. R., McCallum F. J., Anders R. F., Marsh K., *et al.* (2013) Antibodies to polymorphic invasion-inhibitory and non-inhibitory epitopes of *Plasmodium falciparum* apical membrane antigen 1 in human malaria. *PLoS ONE* 8, e68304.
 118. Anders R. F., Crewther P. E., Edwards S., Margetts M., Matthew M. L. S. M., *et al.* (1998) Immunisation with recombinant AMA-1 protects mice against infection with *Plasmodium chabaudi*. *Vaccine* 16, 240-247.
 119. Crewther P. E., Matthew M. L., Flegg R. H., and Anders R. F. (1996) Protective immune responses to apical membrane antigen 1 of *Plasmodium chabaudi* involve recognition of strain-specific epitopes. *Infect. Immun.* 64, 3310-3317.
 120. Salvatore D., Hodder A. N., Zeng W., Brown L. E., Anders R. F., *et al.* (2002) Identification of antigenically active tryptic fragments of apical membrane antigen-1 (AMA1) of *Plasmodium chabaudi* malaria: strategies for assembly of immunologically active peptides. *Vaccine* 20, 3477-3484.
 121. Cortés A., Mellombo M., Masciantonio R., Murphy V. J., Reeder J. C., *et al.* (2005) Allele specificity of naturally acquired antibody responses against *Plasmodium falciparum* apical membrane antigen 1. *Infect. Immun.* 73, 422-430.
 122. Lalitha P. V., Ware L. A., Barbosa A., Dutta S., Moch J. K., *et al.* (2004) Production of the subdomains of the *Plasmodium falciparum* apical membrane antigen 1 ectodomain and analysis of the immune response. *Infect. Immun.* 72, 4464-4470.
 123. Polley S. D., Mwangi T., Kocken C. H. M., Thomas A. W., Dutta S., *et al.* (2004) Human antibodies to recombinant protein constructs of *Plasmodium falciparum* apical membrane antigen 1 (AMA1) and their associations with protection from malaria. *Vaccine* 23, 718-728.

124. Lyke K. E., Daou M., Diarra I., Kone A., Kouriba B., *et al.* (2009) Cell-mediated immunity elicited by the blood stage malaria vaccine apical membrane antigen 1 in Malian adults: results of a Phase I randomized trial. *Vaccine* 27, 2171-2176.
125. Spring M. D., Cummings J. F., Ockenhouse C. F., Dutta S., Reidler R., *et al.* (2009) Phase 1/2a study of the malaria vaccine candidate apical membrane antigen-1 (AMA-1) administered in adjuvant system AS01B or AS02A. *PLoS ONE* 4, e5254.
126. Duan J., Mu J., Thera M. A., Joy D., Kosakovsky Pond S. L., *et al.* (2008) Population structure of the genes encoding the polymorphic *Plasmodium falciparum* apical membrane antigen 1: implications for vaccine design. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7857-7862.
127. Escalante A. A., Grebert H. M., Chaiyaroj S. C., Magris M., Biswas S., *et al.* (2001) Polymorphism in the gene encoding the apical membrane antigen-1 (AMA-1) of *Plasmodium falciparum*. X. Asembo Bay Cohort Project. *Mol. Biochem. Parasitol.* 113, 279-287.
128. Marshall V. M., Zhang L., Anders R. F., and Coppel R. L. (1996) Diversity of the vaccine candidate AMA-1 of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 77, 109-113.
129. Polley S. D., and Conway D. J. (2001) Strong diversifying selection on domains of the *Plasmodium falciparum* apical membrane antigen 1 gene. *Genetics* 158, 1505-1512.
130. Payne R. O., Milne K. H., Elias S. C., Edwards N. J., Douglas A. D., *et al.* (2016) Demonstration of the blood-stage *Plasmodium falciparum* controlled human malaria infection model to assess efficacy of the *P. falciparum* apical membrane antigen 1 vaccine, FMP2.1/AS01. *J. Infect. Dis.*
131. Malkin E. M., Diemert D. J., McArthur J. H., Perreault J. R., Miles A. P., *et al.* (2005) Phase 1 clinical trial of apical membrane antigen 1: an asexual blood-stage vaccine for *Plasmodium falciparum* malaria. *Infect. Immun.* 73, 3677-3685.
132. Dicko A., Sagara I., Ellis R. D., Miura K., Guindo O., *et al.* (2008) Phase 1 study of a Combination AMA1 blood stage malaria vaccine in Malian children. *PLoS ONE* 3, e1563.
133. Sagara I., Dicko A., Ellis R. D., Fay M. P., Diawara S. I., *et al.* (2009) A randomized controlled phase 2 trial of the blood stage AMA1-C1/alhydrogel malaria vaccine in children in Mali. *Vaccine* 27, 3090-3098.
134. Ellis R. D., Mullen G. E. D., Pierce M., Martin L. B., Miura K., *et al.* (2009) A Phase 1 study of the blood-stage malaria vaccine candidate AMA1-C1/Alhydrogel® with CPG 7909, using two different formulations and dosing intervals. *Vaccine* 27, 4104-4109.
135. Mullen G. E. D., Ellis R. D., Miura K., Malkin E., Nolan C., *et al.* (2008) Phase 1 trial of AMA1-C1/Alhydrogel plus CPG 7909: an asexual blood-stage vaccine for *Plasmodium falciparum* Malaria. *PLoS ONE* 3, e2940.
136. Sagara I., Ellis R. D., Dicko A., Niambele M. B., Kamate B., *et al.* (2009) A randomized and controlled Phase 1 study of the safety and immunogenicity of the AMA1-C1/Alhydrogel® + CPG 7909 vaccine for *Plasmodium falciparum* malaria in semi-immune Malian adults. *Vaccine* 27, 7292-7298.
137. www.ClinicalTrials.gov. (2009) A study to assess safety, immunogenicity and parasite growth inhibition of an asexual blood stage vaccine for *P. falciparum* malaria. NCT00984763.
138. Duncan C. J. A., Sheehy S. H., Ewer K. J., Douglas A. D., Collins K. A., *et al.* (2011) Impact on malaria parasite multiplication rates in infected volunteers of the protein-in-adjuvant vaccine AMA1-C1/Alhydrogel+CPG 7909. *PLoS ONE* 6, e22271.

139. Ouattara A., Mu J., Takala-Harrison S., Saye R., Sagara I., *et al.* (2010) Lack of allele-specific efficacy of a bivalent AMA1 malaria vaccine. *Malar. J.* 9, 1-13.
140. Remarque E. J., Faber B. W., Kocken C. H. M., and Thomas A. W. (2008) A diversity-covering approach to immunization with *Plasmodium falciparum* apical membrane antigen 1 induces broader allelic recognition and growth inhibition responses in rabbits. *Infect. Immun.* 76, 2660-2670.
141. Chaudhury S., Reifman J., and Wallqvist A. (2014) Simulation of B cell affinity maturation explains enhanced antibody cross-reactivity induced by the polyvalent malaria vaccine AMA1. *J. Immunol.* 193, 2073-2086.
142. Dutta S., Dlugosz L. S., Drew D. R., Ge X., Ababacar D., *et al.* (2013) Overcoming antigenic diversity by enhancing the immunogenicity of conserved epitopes on the malaria vaccine candidate apical membrane antigen-1. *PLoS Pathog.* 9, e1003840.
143. Miura K., Herrera R., Diouf A., Zhou H., Mu J., *et al.* (2013) Overcoming allelic specificity by immunization with five allelic forms of *Plasmodium falciparum* apical membrane antigen 1. *Infect. Immun.* 81, 1491-1501.
144. Srinivasan P., Ekanem E., Diouf A., Tonkin M. L., Miura K., *et al.* (2014) Immunization with a functional protein complex required for erythrocyte invasion protects against lethal malaria. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10311-10316.
145. MacRaild C. A., Pedersen M. Ø., Anders R. F., and Norton R. S. (2012) Lipid interactions of the malaria antigen merozoite surface protein 2. *BBA-Biomembranes* 1818, 2572-2578.
146. McCarthy J. S., Marjason J., Elliott S., Fahey P., Bang G., *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, e24413.
147. Smythe J. A., Coppel R. L., Brown G. V., Ramasamy R., Kemp D. J., *et al.* (1988) Identification of two integral membrane proteins of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U.S.A.* 85, 5195-5199.
148. Sanders P. R., Kats L. M., Drew D. R., O'Donnell R. A., O'Neill M., *et al.* (2006) A set of glycosylphosphatidyl inositol-anchored membrane proteins of *Plasmodium falciparum* is refractory to genetic deletion. *Infect. Immun.* 74, 4330-4338.
149. Gerold P., Schofield L., Blackman M. J., Holder A. A., and Schwarz R. T. (1996) Structural analysis of the glycosyl-phosphatidylinositol membrane anchor of the merozoite surface proteins-1 and -2 of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 75, 131-143.
150. Taylor R. R., Smith D. B., Robinson V. J., McBride J. S., and Riley E. M. (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.
151. Zhang X., Perugini M. A., Yao S., Adda C. G., Murphy V. J., *et al.* (2008) Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2. *J. Mol. Biol.* 379, 105-121.
152. Wright G. J., and Rayner J. C. (2014) *Plasmodium falciparum* erythrocyte invasion: combining function with immune evasion. *PLoS Pathog.* 10, e1003943.
153. Adda C. G., MacRaild C. A., Reiling L., Wycherley K., Boyle M. J., *et al.* (2012) Antigenic characterization of an intrinsically unstructured protein, *Plasmodium falciparum* merozoite surface protein 2. *Infect. Immun.* 80, 4177-4185.
154. Adda C. G., Murphy V. J., Sunde M., Waddington L. J., Schloegel J., *et al.* (2009) *Plasmodium falciparum* merozoite surface protein 2 is unstructured and forms amyloid-like fibrils. *Mol. Biochem. Parasitol.* 166, 159-171.

155. Chandrashekar I. R., Adda C. G., MacRaild C. A., Anders R. F., and Norton R. S. (2010) Inhibition by flavonoids of amyloid-like fibril formation by *Plasmodium falciparum* merozoite surface protein 2. *Biochemistry* 49, 5899-5908.
156. Chandrashekar I. R., Adda C. G., MacRaild C. A., Anders R. F., and Norton R. S. (2011) EGCG disaggregates amyloid-like fibrils formed by *Plasmodium falciparum* merozoite surface protein 2. *Arch. Biochem. Biophys.* 513, 153-157.
157. Feng Z.-P., Zhang X., Han P., Arora N., Anders R. F., *et al.* (2006) Abundance of intrinsically unstructured proteins in *P. falciparum* and other apicomplexan parasite proteomes. *Mol. Biochem. Parasitol.* 150, 256-267.
158. Zhang X., Adda C. G., Low A., Zhang J., Zhang W., *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, 1380-1387.
159. MacRaild Christopher A., Richards Jack S., Anders Robin F., and Norton Raymond S. (2016) Antibody recognition of disordered antigens. *Structure* 24, 148-157.
160. MacRaild C. A., Zachrdla M., Andrew D., Krishnarjuna B., Nováček J., *et al.* (2015) Conformational dynamics and antigenicity in the disordered malaria antigen merozoite surface protein 2. *PLoS ONE* 10, e0119899.
161. Ekala M. T., Jouin H., Lekoulou F., Mercereau-Puijalon O., and Ntoumi F. (2002) Allelic family-specific humoral responses to merozoite surface protein 2 (MSP2) in Gabonese residents with *Plasmodium falciparum* infections. *Clin. Exp. Immunol.* 129, 326-331.
162. Taylor R. R., Allen S. J., Greenwood B. M., and Riley E. M. (1998) IgG3 antibodies to *Plasmodium falciparum* merozoite surface protein 2 (MSP2): increasing prevalence with age and association with clinical immunity to malaria. *Am. J. Trop. Med. Hyg.* 58, 406-413.
163. Al-Yaman F., Genton B., Anders R. F., Falk M., Triglia T., *et al.* (1994) Relationship between humoral response to *Plasmodium falciparum* merozoite surface antigen-2 and malaria morbidity in a highly endemic area of Papua New Guinea. *Am. J. Trop. Med. Hyg.* 51, 593-602.
164. Metzger W. G., Okenu D. M. N., Cavanagh D. R., Robinson J. V., Bojang K. A., *et al.* (2003) Serum IgG3 to the *Plasmodium falciparum* merozoite surface protein 2 is strongly associated with a reduced prospective risk of malaria. *Parasite Immunol.* 25, 307-312.
165. Polley S. D., Conway D. J., Cavanagh D. R., McBride J. S., Lowe B. S., *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, 4233-4246.
166. Courtin D., Oesterholt M., Huisman H., Kusi K., Milet J., *et al.* (2009) The quantity and quality of African children's IgG responses to merozoite surface antigens reflect protection against *Plasmodium falciparum* malaria. *PLoS ONE* 4, e7590.
167. Boyle M. J., Reiling L., Feng G., Langer C., Osier Faith H., *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.
168. Flueck C., Frank G., Smith T., Jafarshad A., Nebie I., *et al.* (2009) Evaluation of two long synthetic merozoite surface protein 2 peptides as malaria vaccine candidates. *Vaccine* 27, 2653-2661.

169. Osier F., Feng G., Boyle M., Langer C., Zhou J., *et al.* (2014) Opsonic phagocytosis of *Plasmodium falciparum* merozoites: mechanism in human immunity and a correlate of protection against malaria. *BMC Med.* 12, 108.
170. Wright P. E., and Dyson H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* 293, 321-331.
171. Eisen D., Billman-Jacobe H., Marshall V. F., Fryauff D., and Coppel R. L. (1998) Temporal variation of the merozoite surface protein-2 gene of *Plasmodium falciparum*. *Infect. Immun.* 66, 239-246.
172. Weisman S., Wang L., Billman-Jacobe H., Nhan D. H., Richie T. L., *et al.* (2001) Antibody responses to infections with strains of *Plasmodium falciparum* expressing diverse forms of merozoite surface protein 2. *Infect. Immun.* 69, 959-967.
173. Al-Yaman F., Genton B., Anders R., Taraika J., Ginny M., *et al.* (1995) Assessment of the role of the humoral response to *Plasmodium falciparum* MSP2 compared to RESA and SPf66 in protecting Papua New Guinean children from clinical malaria. *Parasite Immunol.* 17, 493-501.
174. Aucan C., Traoré Y., Tall F., Nacro B., Traoré-Leroux T., *et al.* (2000) High immunoglobulin G2 (IgG2) and low IgG4 levels are associated with human resistance to *Plasmodium falciparum* malaria. *Infect. Immun.* 68, 1252-1258.
175. Flück C., Smith T., Beck H.-P., Irion A., Betuela I., *et al.* (2004) Strain-specific humoral response to a polymorphic malaria vaccine. *Infect. Immun.* 72, 6300-6305.
176. Pan W., Huang D., Zhang Q., Qu L., Zhang D., *et al.* (2004) Fusion of two malaria vaccine candidate antigens enhances product yield, immunogenicity, and antibody-mediated inhibition of parasite growth in vitro. *J. Immunol.* 172, 6167-6174.
177. Malkin E., Hu J., Li Z., Chen Z., Bi X., *et al.* (2008) A Phase 1 trial of PfCP2.9: An AMA1/MSP1 chimeric recombinant protein vaccine for *Plasmodium falciparum* malaria. *Vaccine* 26, 6864-6873.
178. Shi Q., Lynch M. M., Romero M., and Burns J. M. (2007) Enhanced protection against malaria by a chimeric merozoite surface protein vaccine. *Infect. Immun.* 75, 1349-1358.
179. Faber B. W., Remarque E. J., Morgan W. D., Kocken C. H. M., Holder A. A., *et al.* (2007) Malaria vaccine-related benefits of a single protein comprising *Plasmodium falciparum* apical membrane antigen 1 domains I and II fused to a modified form of the 19-kilodalton C-terminal fragment of merozoite surface protein 1. *Infect. Immun.* 75, 5947-5955.
180. Imai K., and Takaoka A. (2006) Comparing antibody and small-molecule therapies for cancer. *Nat. Rev. Cancer* 6, 714-727.
181. Kar S., and Kar S. (2010) Control of malaria. *Nat. Rev. Drug Discov.* 9, 511-512.
182. Lawrence G., Cheng Q., Reed C., Taylor D., Stowers A., *et al.* (2000) Effect of vaccination with 3 recombinant asexual-stage malaria antigens on initial growth rates of *Plasmodium falciparum* in non-immune volunteers. *Vaccine* 18, 1925-1931.
183. Lawrence G. W., Saul A., Giddy A. J., Kemp R., and Pye D. (1997) Phase I trial in humans of an oil-based adjuvant SEPPIC Montanide ISA 720. *Vaccine* 15, 176-178.
184. Mbow M. L., De Gregorio E., Valiante N. M., and Rappuoli R. (2010) New adjuvants for human vaccines. *Curr. Opin. Immunol.* 22, 411-416.
185. Saul A., Lawrence G., Allworth A., Elliott S., Anderson K., *et al.* (2005) A human Phase 1 vaccine clinical trial of the *Plasmodium falciparum* malaria vaccine candidate apical membrane antigen 1 in Montanide ISA720 adjuvant. *Vaccine* 23, 3076-3083.

186. Saul A., Lawrence G., Smillie A., Rzepczyk C. M., Reed C., *et al.* (1999) Human Phase I vaccine trials of 3 recombinant asexual stage malaria antigens with Montanide ISA720 adjuvant. *Vaccine* 17, 3145-3159.
187. Toledo H., Baly A., Castro O., Resik S., Laferté J., *et al.* (2001) A Phase I clinical trial of a multi-epitope polypeptide TAB9 combined with Montanide ISA 720 adjuvant in non-HIV-1 infected human volunteers. *Vaccine* 19, 4328-4336.
188. Duff A. P., Wilde K. L., Rekas A., Lake V., and Holden P. J. (2015) Robust high-yield methodologies for ^2H and $^2\text{H}/^{15}\text{N}/^{13}\text{C}$ labeling of proteins for structural investigations using neutron scattering and NMR, *Methods Enzymol.* (Zvi, K., Ed.), 3-25, Academic Press.
189. Ahmad Z. A., Yeap S. K., Ali A. M., Ho W. Y., Alitheen N. B. M., *et al.* (2012) scFv antibody: principles and clinical application. *Clin. Dev. Immunol.* 2012, 15.
190. Morales R. A. V., MacRaild C. A., Seow J., Krishnarjuna B., Drinkwater N., *et al.* (2015) Structural basis for epitope masking and strain specificity of a conserved epitope in an intrinsically disordered malaria vaccine candidate. *Sci. Rep.* 5, 10103.
191. Whitlow M., Bell B. A., Feng S.-L., Filpula D., Hardman K. D., *et al.* (1993) An improved linker for single-chain Fv with reduced aggregation and enhanced proteolytic stability. *Protein Eng.* 6, 989-995.
192. Eletsky A., Kienhöfer A., and Pervushin K. TROSY NMR with partially deuterated proteins. *J. Biomol. NMR* 20, 177-180.
193. Pervushin K., Riek R., Wider G., and Wüthrich K. (1997) Attenuated T(2) relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U.S.A.* 94, 12366-12371.
194. Schulte-Herbrüggen T., and Sørensen O. W. (2000) Clean TROSY: compensation for relaxation-induced artifacts. *J. Magn. Reson.* 144, 123-128.
195. Salzmann M., Pervushin K., Wider G., Senn H., and Wüthrich K. (1998) TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins. *Proc. Natl. Acad. Sci. U.S.A.* 95, 13585-13590.
196. Salzmann M., Wider G., Pervushin K., Senn H., and Wüthrich K. (1999) TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins. *J. Am. Chem. Soc.* 121, 844-848.
197. Orekhov V. Y., and Jaravine V. A. (2011) Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Prog. Nucl. Magn. Reson. Spectrosc.* 59, 271-292.
198. Korzhnev D. M., Ibraghimov I. V., Billeter M., and Orekhov V. Y. MUNIN: Application of three-way decomposition to the analysis of heteronuclear NMR relaxation data. *J. Biomol. NMR* 21, 263-268.
199. Orekhov V. Y., Ibraghimov I., and Billeter M. Optimizing resolution in multidimensional NMR by three-way decomposition. *J. Biomol. NMR* 27, 165-173.
200. Skinner S. P., Goult B. T., Fogh R. H., Boucher W., Stevens T. J., *et al.* (2015) Structure calculation, refinement and validation using CcpNmr Analysis. *Acta Crystallographica Section D* 71, 154-161.
201. Vranken W. F., Boucher W., Stevens T. J., Fogh R. H., Pajon A., *et al.* (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: Struct., Funct., Bioinf.* 59, 687-696.
202. MacRaild C. A., and Norton R. S. (2014) RASP: rapid and robust backbone chemical shift assignments from protein structure. *J. Biomol. NMR* 58, 155-163.
203. Krishnarjuna B., Lim S. S., Devine S. M., Debono C. O., Lam R., *et al.* (2016) Solution NMR characterization of apical membrane antigen 1 and small molecule

- interactions as a basis for designing new antimalarials. *J. Mol. Recognit.*, (Published online).
204. Williamson M. P. (2013) Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* 73, 1-16.
 205. Hwang T.-L., van Zijl P. M., and Mori S. (1998) Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical exchange (CLEANEX-PM) approach with a fast-HSQC (FHSQC) detection scheme. *J. Biomol. NMR* 11, 221-226.
 206. Fernández C., and Wider G. (2003) TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr. Opin. Struct. Biol.* 13, 570-580.
 207. Gardner K. H., and Kay L. E. (1998) The use of ^2H , ^{13}C , ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* 27, 357-406.
 208. Lim S. S., Yang W., Krishnarajuna B., Kannan Sivaraman K., Chandrashekar I. R., *et al.* (2014) Structure and dynamics of apical membrane antigen 1 from *Plasmodium falciparum* FVO. *Biochemistry* 53, 7310-7320.
 209. Søndergaard C. R., Garrett A. E., Carstensen T., Pollastri G., and Nielsen J. E. (2009) Structural artifacts in protein–ligand X-ray structures: implications for the development of docking scoring functions. *J. Med. Chem.* 52, 5673-5684.
 210. Ishima R., and Torchia D. A. (2000) Protein dynamics from NMR. *Nat. Struct. Biol.* 7, 740-743.
 211. Arkin M. R., and Wells J. A. (2004) Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* 3, 301-317.
 212. Davis A. M., St-Gallay S. A., and Kleywegt G. J. (2008) Limitations and lessons in the use of X-ray structural information in drug design. *Drug Discov. Today* 13, 831-841.
 213. Davis A. M., Teague S. J., and Kleywegt G. J. (2003) Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angew. Chem. Int. Ed.* 42, 2718-2736.
 214. Perkins J. R., Diboun I., Dessailly B. H., Lees J. G., and Orengo C. (2010) Transient protein-protein interactions: structural, functional, and network properties. *Structure* 18, 1233-1243.
 215. Mackenzie N. E., and Gooley P. R. (1988) Applications of NMR spectroscopy to biological systems. *Med. Res. Rev.* 8, 57-76.
 216. Oleg. Jardetzky and Roberts G. C. K. (1981) NMR in molecular biology. 1-579.
 217. Wüthrich K. (1995) NMR in structural biology *World scientific series in 20th century chemistry* 5.
 218. Assfalg M., Ragona L., Pagano K., D'Onofrio M., Zanzoni S., *et al.* (2016) The study of transient protein–nanoparticle interactions by solution NMR spectroscopy. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1864, 102-114.
 219. Janowska M. K., Wu K.-P., and Baum J. (2015) Unveiling transient protein-protein interactions that modulate inhibition of alpha-synuclein aggregation by beta-synuclein, a pre-synaptic protein that co-localizes with alpha-synuclein. *Scientific Reports* 5, 15164.
 220. Liu Z., Gong Z., Dong X., and Tang C. (2016) Transient protein–protein interactions visualized by solution NMR. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1864, 115-122.
 221. Worrall J. A. R., Reinle W., Bernhardt R., and Ubbink M. (2003) Transient protein interactions studied by NMR spectroscopy: the case of cytochrome c and adrenodoxin. *Biochemistry* 42, 7068-7076.

222. Foster M. P., Wuttke D. S., Clemens K. R., Jahnke W., Radhakrishnan I., *et al.* Chemical shift as a probe of molecular interfaces: NMR studies of DNA binding by the three amino-terminal zinc finger domains from transcription factor IIIA. *J. Biomol. NMR* 12, 51-71.
223. Zuiderweg E. R. P. (2002) Mapping protein–protein interactions in solution by NMR spectroscopy†. *Biochemistry* 41, 1-7.
224. Keifer P. A. (2000) NMR spectroscopy in drug discovery: tools for combinatorial chemistry, natural products, and metabolism research, *Prog. Drug Res.* (Jucker, E., Ed.), 137-211, Birkhäuser Basel, Basel.
225. David W. (2005) NMR spectroscopy and protein structure determination: Applications to drug discovery and development. *Curr. Pharm. Biotechnol.* 6, 105-120.
226. Dyson H. J. (2016) Making sense of intrinsically disordered proteins. *Biophys. J.* 110, 1013-1016.
227. Yap B. K., Leung E. W. W., Yagi H., Galea C. A., Chhabra S., *et al.* (2014) A potent cyclic peptide targeting SPSB2 protein as a potential anti-infective agent. *J. Med. Chem.* 57, 7006-7015.
228. Balam S., Olugbile S., Servis C., Diakite M., D'Alessandro A., *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.
229. Lawrence N., Stowers A., Mann V., Taylor D., and Saul A. (2000) 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.
230. Rzepczyk C. M., Csurhes P. A., Baxter E. P., Doran T. J., Irving D. O., *et al.* (1990) Amino acid sequences recognized by T cells: studies on a merozoite surface antigen from the FCQ-27/PNG isolate of *Plasmodium falciparum*. *Immunol. Lett.* 25, 155-163.
231. Rzepczyk C. M., Csurhes P. A., Saul A. J., Jones G. L., Dyer S., *et al.* (1992) Comparative study of the T cell response to two allelic forms of a malarial vaccine candidate protein. *J. Immunol.* 148, 1197-1204.
232. Bortnick A., and Allman D. (2013) What is and what should always have been: long-lived plasma cells induced by T cell-independent antigens. *J. Immunol.* 190, 5913-5918.
233. Udhayakumar V., Kariuki S., Kolczack M., Girma M., Roberts J. M., *et al.* (2001) Longitudinal study of natural immune responses to the *Plasmodium falciparum* apical membrane antigen (AMA-1) in a holoendemic region of malaria in western Kenya: Asembo Bay Cohort Project VIII. *Am. J. Trop. Med. Hyg.* 65.

Appendix I

AI.1 Alignment of *Pf*AMA1 protein sequences

FVO_AJ277646_	MRKLYCVLLLSAFEFTYMINFGRGQNYWEHPYQKSDVYHPINEHREHPKEYEYPLHQEHT	60
HB3_AMA1_PROT_	MRKLYCVLLLSAFEFTYMINFGRGQNYWEHPYQNSDVYRPINEHREHPKEYEYPLHQEHT	60
W2MEF_PROT_	MRKLYCVLLLSAFEFTYMINFGRGQNYWEHPYQKSDVYHPINEHREHPKEYQYPLHQEHT	60
3D7_AMA1_PROT_	MRKLYCVLLLSAFEFTYMINFGRGQNYWEHPYQNSDVYRPINEHREHPKEYEYPLHQEHT	60
D10_AMA1_PROT_	MRKLYCVLLLSAFEFTYMINFGRGQNYWEHPYQKSDVYHPINEHREHPKEYQYPLHQEHT	60
	*****:****:*****:*****	
FVO_AJ277646_	YQQEDSGEDENTLQHAYPIDHEGAEPAPQEQLNLFPSIEIVERSSNYMGNPWTEYMAKYDIE	120
HB3_AMA1_PROT_	YQQEDSGEDENTLQHAYPIDHEGAEPAPQEQLNLFSSIEIVERSSNYMGNPWTEYMAKYDIE	120
W2MEF_PROT_	YQQEDSGEDENTLQHAYPIDHEGAEPAPQEQLNLFSSIEIVERSSNYMGNPWTEYMAKYDIE	120
3D7_AMA1_PROT_	YQQEDSGEDENTLQHAYPIDHEGAEPAPQEQLNLFSSIEIVERSSNYMGNPWTEYMAKYDIE	120
D10_AMA1_PROT_	YQQEDSGEDENTLQHAYPIDHEGAEPAPQEQLNLFSSIEIVERSSNYMGNPWTEYMAKYDIE	120
	*****.*****	
FVO_AJ277646_	EVHSGGIRVDLGEDAEVAGTQYRLPSGKCPVFGKGIIENSNTTFLKPVATGNQDLKDGG	180
HB3_AMA1_PROT_	KVHSGGIRVDLGEDAEVAGTQYRLPSGKCPVFGKGIIENSKTTFLTPVATENQDLKDGG	180
W2MEF_PROT_	EVHSGGIRVDLGEDAEVAGTQYRLPSGKCPVFGKGIIENSNTTFLTPVATGNQYLDKGG	180
3D7_AMA1_PROT_	EVHSGGIRVDLGEDAEVAGTQYRLPSGKCPVFGKGIIENSNTTFLTPVATGNQYLDKGG	180
D10_AMA1_PROT_	EVHSGGIRVDLGEDAEVAGTQYRLPSGKCPVFGKGIIENSNTTFLTPVATGNQYLDKGG	180
	:*****:*****.***** ** *	
FVO_AJ277646_	FAPPTNPLISPMTLNGMRDFYKNNEYVKNLDELTLCSRHAGNMNPDNDKNSNYKYPAYV	240
HB3_AMA1_PROT_	FAPPTEPLISPMTLDQMRHLYKDNEYVKNLDELTLCSRHAGNMNPDNDKNSNYKYPAYV	240
W2MEF_PROT_	FAPPTKPLMSPMTLDDMRLLYKDNEVKNLDELTLCSRHAGNMNPDNDKNSNYKYPAYV	240
3D7_AMA1_PROT_	FAPPTEPLMSPMTLDEMRFYKDNKYVKNLDELTLCSRHAGNMIPDNDKNSNYKYPAYV	240
D10_AMA1_PROT_	FAPPTEPLMSPMTLDEMRFYKDNKYVKNLDELTLCSRHAGNMIPDNDKNSNYKYPAYV	240
	*****:****:****: ** :*:*: ***** *****	
FVO_AJ277646_	DYNDKKCHILYIAAQENNGPRYCNKDQSKRNSMFCFRPAKDKLFENYTYLSKNVVDNWEE	300
HB3_AMA1_PROT_	DYEDKKCHILYIAAQENNGPRYCNKDESKRNSMFCFRPAKDKLFENYTYLSKNVVDNWEE	300
W2MEF_PROT_	DYNDKKCHILYIAAQENNGPRYCNKDESKRNSMFCFRPAKDKSFQNYTYLSKNVVDNWEE	300
3D7_AMA1_PROT_	DDKDKKCHILYIAAQENNGPRYCNKDESKRNSMFCFRPAKDISFQNYTYLSKNVVDNWEK	300
D10_AMA1_PROT_	DDKDKKCHILYIAAQENNGPRYCNKDESKRNSMFCFRPAKDISFQNYTYLSKNVVDNWEK	300
	* :*****:***** * :*****:	
FVO_AJ277646_	VCPRKNLENAKFGLWVDGNCEDIPHVNEFSANDLFECNKLVFELSASDQPKQYEQHLDY	360
HB3_AMA1_PROT_	VCPRKNLENAKFGLWVDGNCEDIPHVNEFSANDLFECNKLVFELSASDQPKQYEQHLDY	360
W2MEF_PROT_	VCPRKNLENAKFGLWVDGNCEDIPHVNEFSANDLFECNKLVFELSASDQPKQYEQHLDY	360
3D7_AMA1_PROT_	VCPRKNLQNAKFGLWVDGNCEDIPHVNEFPAIDLFEKNLKFELSASDQPKQYEQHLDY	360
D10_AMA1_PROT_	VCPRKNLQNAKFGLWVDGNCEDIPHVNEFSAIDLFEKNLKFELSASDQPKQYEQHLDY	360
	*****:*****.* *****	
FVO_AJ277646_	EKIKEGFKNKNASMIKSAFLPTGAFKADRYKSHGKGYNWGNYNRETQKCEIFNVKPTCLI	420
HB3_AMA1_PROT_	EKIKEGFKNKNASMIKSAFLPTGAFKADRYKSRGKGYNWGNYNRETQKCEIFNVKPTCLI	420
W2MEF_PROT_	EKIKEGFKNKNASMIKSAFLPTGAFKADRYKSHGKGYNWGNYNRKTQKCEIFNVKPTCLI	420
3D7_AMA1_PROT_	EKIKEGFKNKNASMIKSAFLPTGAFKADRYKSHGKGYNWGNYNRETQKCEIFNVKPTCLI	420

```

D10_AMA1_PROT_      EKIKEGEKNKNASMIKSAFLPTGAFKADRYKSHGKGYNWGNYNNTETQKCEIFNVKPTCLI 420
*****:*****:*****

FVO_AJ277646_      NNSSYIATTALSHPIEVEHNFPCSLYKDEIKKEIERESKRIKLNNDNDEGNKEIIAPRIF 480
HB3_AMA1_PROT_      NNSSYIATTALSHPNVEENNFPCLYKDEIKKEIERESKRIKLNNDNDEGNKKIIAPRIF 480
W2MEF_PROT_         NNSSYIATTALSHPIEVEHNFPCSLYKDEIKKEIERESKRIKLNNDNDEGNKKIIAPRIF 480
3D7_AMA1_PROT_      NNSSYIATTALSHPIEVENNFPCSLYKDEIMKEIERESKRIKLNNDNDEGNKKIIAPRIF 480
D10_AMA1_PROT_      NNSSYIATTALSHPIEVEHNFPCSLYKNEIMKEIERESKRIKLNNDNDEGNKKIIAPRIF 480
***** ***:*****:* *****:*****

FVO_AJ277646_      ISD-KDSLKCPCDPEMVSNSTCRFFVCKCVERRAEVTSNNEVVVKEEYKDEYADIPEHKP 539
HB3_AMA1_PROT_      ISDDKDSLKCPCDPEIVSNSTCNFFVCKCVEKRAEVTSNNEVVVKEEYKDEYADIPEHKP 540
W2MEF_PROT_         ISDDIDSLKCPCDPEIVSNSTCNFFVCKCVEKRAEVTSNNEVVVKEEYKDEYADIPEHKP 540
3D7_AMA1_PROT_      ISDDKDSLKCPCDPEMVSNSTCRFFVCKCVERRAEVTSNNEVVVKEEYKDEYADIPEHKP 540
D10_AMA1_PROT_      ISDDKDSLKCPCDPEIVSNSTCNFFVCKCVERRAEVTSNNEVVVKEEYKDEYADIPEHKP 540
*** *****:*****.*****:*****

FVO_AJ277646_      TYDNMKIIIA---SSAAVAVLATILMVLYKRKGNAEKYDKMDQPQHYGKSTSRNDEMMLD 596
HB3_AMA1_PROT_      TYDNMKIIIA---SSAAVAVLATILMVLYKRKGNAEKYDKMDQPQHYGKSNSRNDEMMLD 597
W2MEF_PROT_         TYDKMKIIIA---SSAAVAVLATILMVLYKRKGNAEKYDKMDEPQHYGKSNSRNDEMMLD 597
3D7_AMA1_PROT_      TYDKMKIIIA---SSAAVAVLATILMVLYKRKGNAEKYDKMDEPQDYGKSNSRNDEMMLD 597
D10_AMA1_PROT_      TYDKMKIIIAPRTSSAAVAVLATILMVLYKRKGNAEKYDKMDEPQHYGKSNSRNDEMMLD 600
***:***** *****:*****.*.*****

FVO_AJ277646_      PEASFWGEEKRASHTTPVLMKEPYY 621
HB3_AMA1_PROT_      PEASFWG----- 604
W2MEF_PROT_         PEASFWGEEKRASHTTPVLMKEPYY 622
3D7_AMA1_PROT_      PEASFWGEEKRASHTTPVLMKEPYY 622
D10_AMA1_PROT_      PEASFWGEEKRASHTTPVLMKEPYY 625
*****

```

Figure AI.1. ClustalW (1.83) multiple alignment of AMA1 amino acid sequences from *P. falciparum* FVO, HB3, W2MEF, 3D7 and D10 strains. Hydrophobic cleft residues are highlighted with cyan and the dimorphic residues are underlined. The sequences were obtained from National Center for Biotechnology Information (NCBI).

Appendix II

The *N*-terminal conserved region of MSP2 is fully conserved across all strains of malaria, while the *C*-terminal conserved region contains a single point mutation (Asn to Ser, highlighted with green color in the sequence), as documented by Fenton *et al.* and Smythe *et al.*

AII.1 Alignment of *P.falciparum* 3D7 and FC27 MSP2 amino acid sequences

```

FC27MSP  IKNESKYSNTFINNAYNMSIRRSMANEGSNTNSVGANAPNADTIASGSQRSTNSASTSTT
3D7MSP2  IKNESKYSNTFINNAYNMSIRRSMAESKPSTGAGGSAGGSAGGSAGGSAGGSAGGSAGSG
*****
FC27MSP  NNG-----ESQTTTPTAADTIASGSQRSTNSASTSTTNNGESQTTTPTAADTPTATESNS
3D7MSP2  DNGGADAEGSSSTPATTTTTKTTTTTTTTNDAEASTSTSENPNHKNAETNPKGKGEVQE
...      *:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
FC27MSP  PSPPITTTESSSSGNAPNKTDG---KGESEKQNELNESTEEGPRAPQEPQTAENENPAA
3D7MSP2  PNQANKETQNNSNVQQDSQTKSNVPPTQDADTKSPTAQPEQAENSAPTAEQTESPELQSA
* . . . * : . . * . : . : * . . : : : . : : . : : . : : . : :
FC27MSP  PENKGTGQHGHMHGSRNNHPQNTSDSQKECTDGNKENC GAATSLLENSS
3D7MSP2  PENKGTGQHGHMHGSRNNHPQNTSDSQKECTDGNKENC GAATSLLENSS
*****

```

Figure II.1. Sequence alignment of *P.falciparum* 3D7 and FC27 MSP2. The residues highlighted in green are dimorphic within the *C*-terminal conserved region.

Appendix III

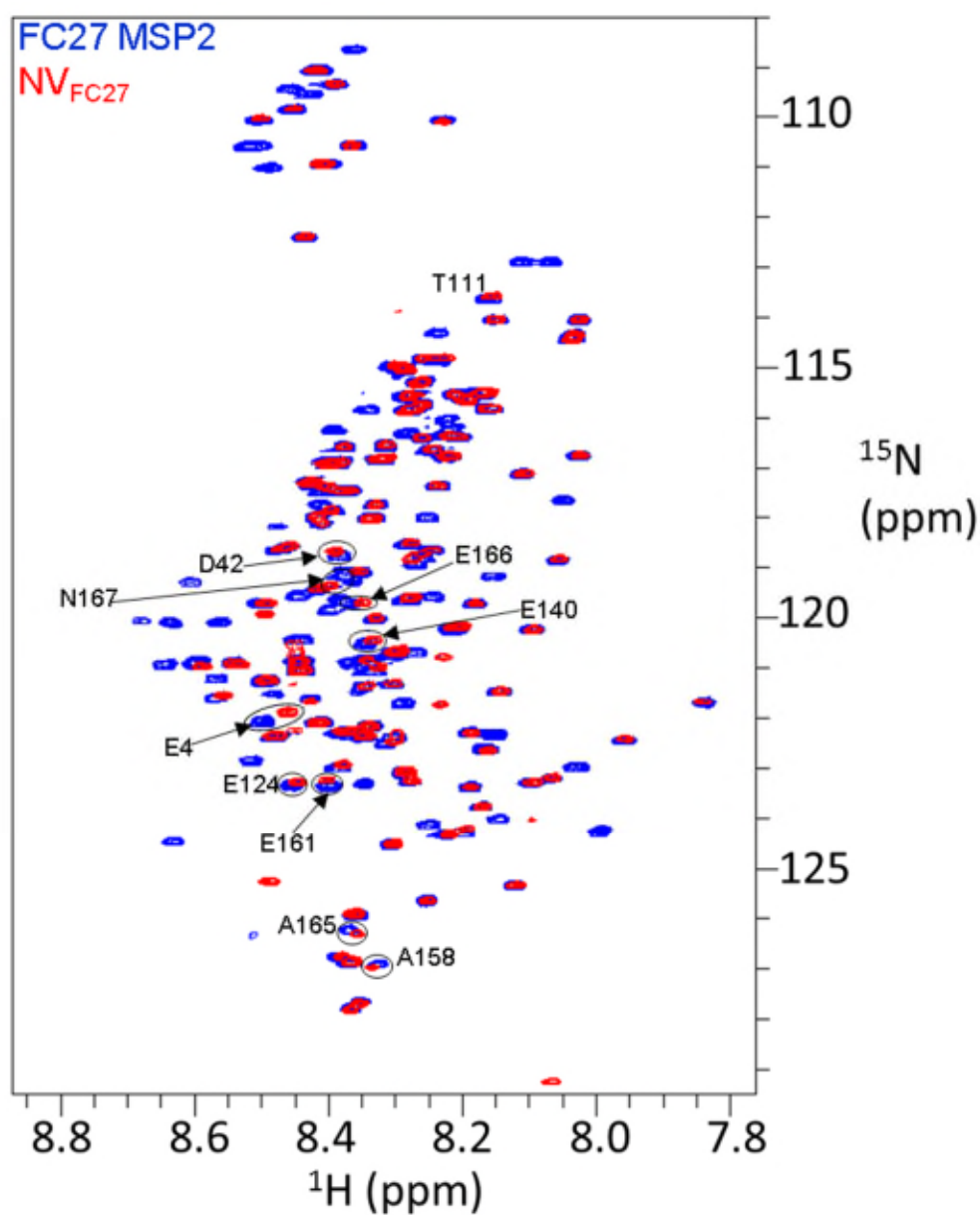


Figure III.1. Overlay of HSQC spectra of full-length FC27 MSP2 (blue) and NV_{FC27} (red). The chemical shift changes in the NV_{FC27} spectrum compared to the spectrum of full-length FC27 MSP2 are indicated by circles. Samples were prepared in sodium acetate buffer pH 4.5 and the experiments were at recorded 298 K.

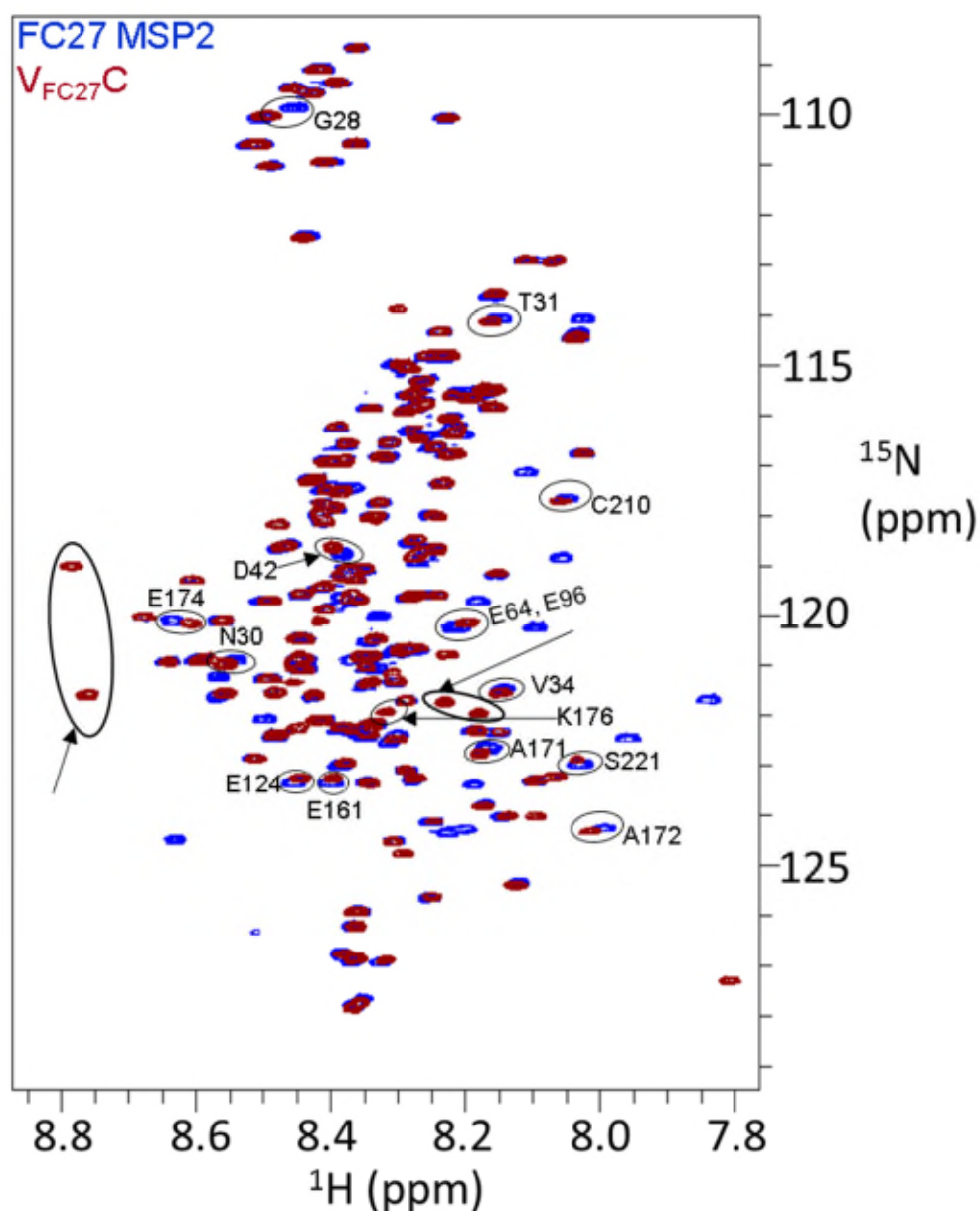


Figure III.2. Overlay of HSQC spectra of full-length FC27 MSP2 (blue) and V_{FC27C} (red). The chemical shift changes in the V_{FC27C} spectrum compared to the spectrum of full-length FC27 MSP2 are indicated by circles. The resonances that cannot be assigned by direct comparison with the spectrum of full-length FC27 MSP2 are indicated with arrows. Samples were prepared in sodium acetate buffer pH 4.5 and the experiments were at recorded 298 K.

Appendix IV

In general, T-cell independent antigens generate short-lived and low-affinity antibodies [232]. High-affinity antibodies can be generated by introducing strong T-cell epitopes of other antigens into these T-cell independent antigens. To improve the antigenic processing of MSP2 by the immune system, this antigen was linked to *Pf*AMA1 (which has T-cell epitopes [233]), through a *Pf*RON2 peptide (Figure IV.1). The C-terminus was chosen as an anchor point in linking the MSP2 to *Pf*AMA1 as a means of mimicking native MSP2, which is linked to the parasite membrane through a C-terminal GPI-anchor.

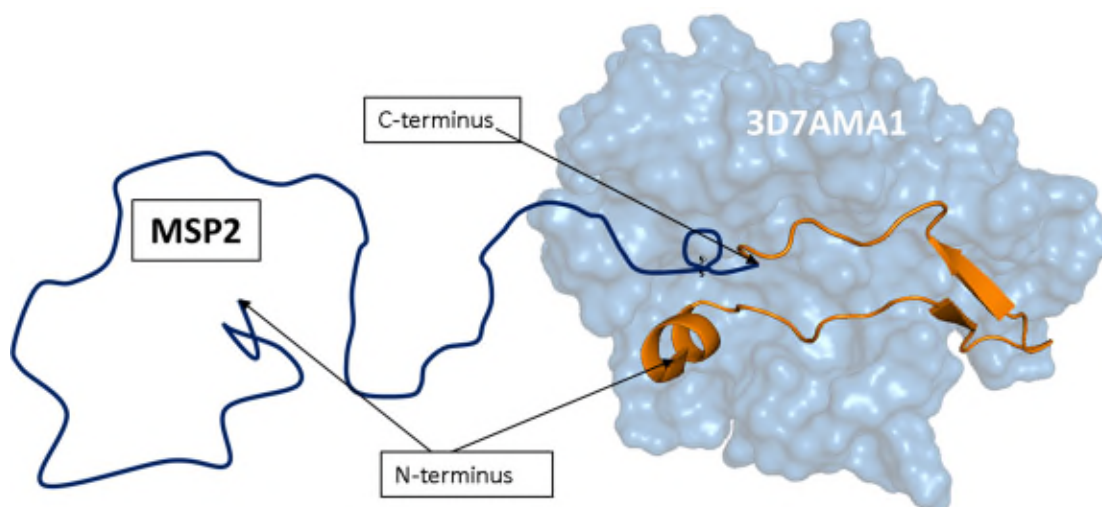


Figure IV.1. Schematic model of MSP2 linked to *Pf*AMA1 [59] through the *Pf*RON2 [59] peptide.

AIV.1 Preparation of plasmid constructs

Synthetic genes for both 3D7 MSP2-*Pf*RON2 and FC27 MSP2-*Pf*RON2 were purchased from GenScript (Piscataway, NJ, USA). Genes were sub-cloned into an expression pET32a vector between KpnI and NcoI restriction sites. The gene sequences were confirmed by sequencing the plasmids. Table 1 shows the protein sequences for MSP2-*Pf*RON2 chimeras.

Table IV.1. Amino acid sequences for 3D7MSP2-*Pf*RON2 and FC27MSP2-*Pf*RON2 chimeric constructs. The linker sequence between MSP2 and *Pf*RON2 is highlighted in red.

3D7MSP2- <i>Pf</i> RON2	GIKNE SKYSNTFINNAYNMSIRR SMAESKPSTGAGGSAGGSAGGSAGGSAGGS AGSGD GNGADAEGSSSTPATT TTTTKTT TTTT TTTTNDAEASTSTSS ENPNHKNAETNPK GKGEVQEPNQANKETQNN SNVQQDSQTKSNVPPTQDADTKSPTAQPEQAENSAPTAE
----------------------------	--

	QTESPELQSAPENKGTGQHGHMHGSRNNHPQNTSDSQKECTDGNKENC GAATSLLSN SSGGGGSGGGGSDITQQAKDIGAGPVASCFTTRMSPPQQICLNSVVNTALS
FC27MSP2- <i>Pf</i> RON2	GIKNESKYSNTFINNAYNMSIRRS MANEGSNTNSV GANAPNADTIASGSQRSTNSAS TSTTNNGESQTTTPTAADTIASGSQRSTNSASTSTTNNGESQTTTPTAADTPTATES NSPSPPIITTESSSSGNAPNKTDGKGEESEKQNELNESTEEGPRAPQEPQTAENENP AAPENKGTGQHGHMHGSRNNHPQNTSDSQKECTDGNKENC GAATSLLSNSSGGGGSG GGGSDITQQAKDIGAGPVASCFTTRMSPPQQICLNSVVNTALS

AIV.2 Protein expression

The plasmids were transformed into BL21 (DE3) cells and a single colony was inoculated into 1L LB media containing 100 mg/L ampicillin and cultured at 37 °C, 225 rpm. Protein was overexpressed for 4-6 h by adding IPTG at 1 mM concentration when the culture OD₆₀₀=0.6. Cells were pelleted by centrifugation at 6000 rpm, 4 °C, for 10 min.

AIV.3 Protein purification

AIV.3.1 Cell lysis

The bacterial cell pellets were resuspended in 20 mM Tris buffer (pH 8) containing 100 mM NaCl and protease inhibitors. The cells were lysed by incubating in boiling water (95 °C) for 10 min, carefully mixing every 3-4 min. Lysed cells were cooled immediately on ice for 15 min followed by incubation with DNase, 5 mM MgCl₂ for 20 min to clear the cellular DNA. Cell lysate was centrifuged at 50,000 g for 25 min and the supernatant containing the protein was collected.

AIV.3.2 His-Trap affinity chromatography

The supernatant was filtered through a 0.2 µm membrane filter and loaded onto a 5 mL His-Trap column equilibrated with buffer A (20 mM Tris, pH 8, 100 mM NaCl, 20 mM imidazole). Protein impurities bound to the His-Trap column through non-specific interactions were removed by thoroughly washing the column with 5-10 column volumes of 20 mM Tris (pH 8) containing 6 M Gdn-HCl and 2 mM β-mercaptoethanol. The column was re-equilibrated with 5 column volumes of buffer A. The protein was eluted using a linear gradient of the same buffer containing 500 mM imidazole (buffer B). All the buffers used in FPLC were filtered through 0.22 µm filter and degassed just before commencing the protein purification.

AIV.3.3 TEV cleavage and removal of thioredoxin

Purified protein was dialyzed against TEV cleavage buffer (50 mM Tris pH 8) at 4 °C for 16 h. The dialyzed protein solution was filtered through the 0.22 µm filter to remove any

aggregation or precipitation before cleavage reaction. The cleavage reaction was carried out at 34 °C for 12 h in the presence of 0.3 mM and 3 mM oxidized and reduced glutathione, respectively. TEV protease was used at a concentration of 1 mg per 100 mg of protein (1:100). The cleavage reaction was filtered through 0.22 µm filter to remove any precipitation or aggregation material. Both MSP2-*Pf*RON2 chimeras have an additional glycine residue at their *N*-terminus from the TEV cleavage site (ENLYFQ-G).

AIV.4 Preparation of MSP2-*Pf*RON2+*Pf*AMA1 complex

6His-tagged FVO and 3D7 *Pf*AMA1 proteins were purified as follows:

- 1) The Ni-NTA resin was equilibrated with 20 mM Tris buffer (pH 8).
- 2) Lyophilized *Pf*AMA1 protein was dissolved in the 20 mM Tris buffer (pH 8).
- 3) The dissolved protein was added to the Ni-NTA resin and incubated for 30 min at room temperature.
- 4) *Pf*AMA1 bound resin was washed 3-4 times with 20 mM Tris buffer (pH 8) to remove unbound/excess protein and incubated with the TEV cleaved reaction containing the 3D7MSP2-*Pf*RON2 chimera.
- 5) The MSP2-*Pf*RON2+*Pf*AMA1 complex was eluted with 20 mM Tris buffer (pH 8) containing 200 mM imidazole.
- 6) The samples were analysed by SDS-PAGE. Gels were loaded with 1.5 µg of protein per well and stained with Coomassie Blue.

The imidazole elution showed three clear protein bands corresponding to the 3D7 MSP2-*Pf*RON2 chimera, *Pf*AMA1 and thioredoxin (Figure IV.2).

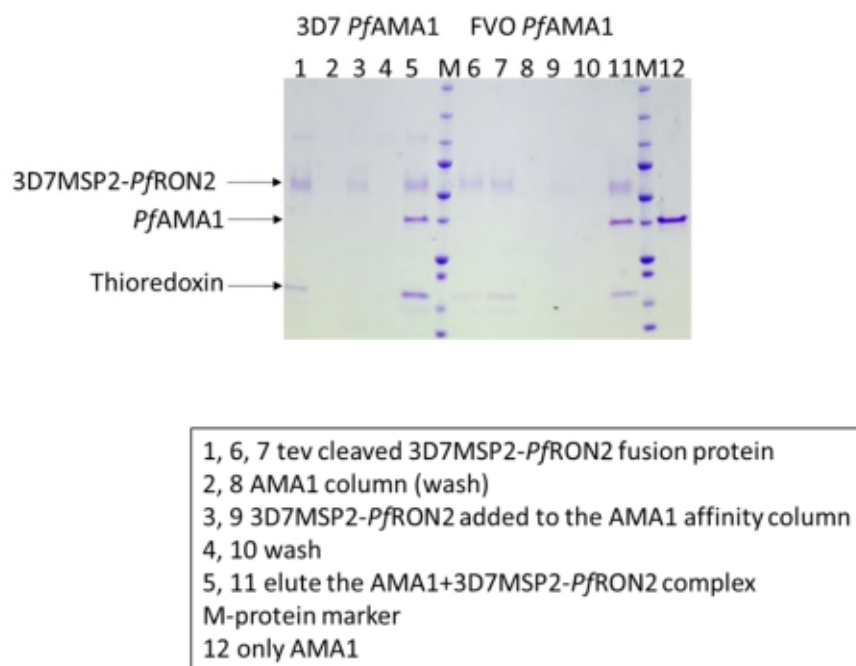


Figure IV.2. Analysis of 3D7 MSP2-*PfRON2*+*PfAMA1* complex formation by SDS-PAGE.

AIV.5 Size-exclusion chromatography (SEC)

Size-exclusion chromatography was carried out to purify the complex and the fractions were analysed by SDS-PAGE. MSP2-*PfRON2* chimera and *PfAMA1* were eluted at different time points (Figure IV.3) suggesting instability of the complex on the SEC column. Further optimization is required to stabilize this complex.

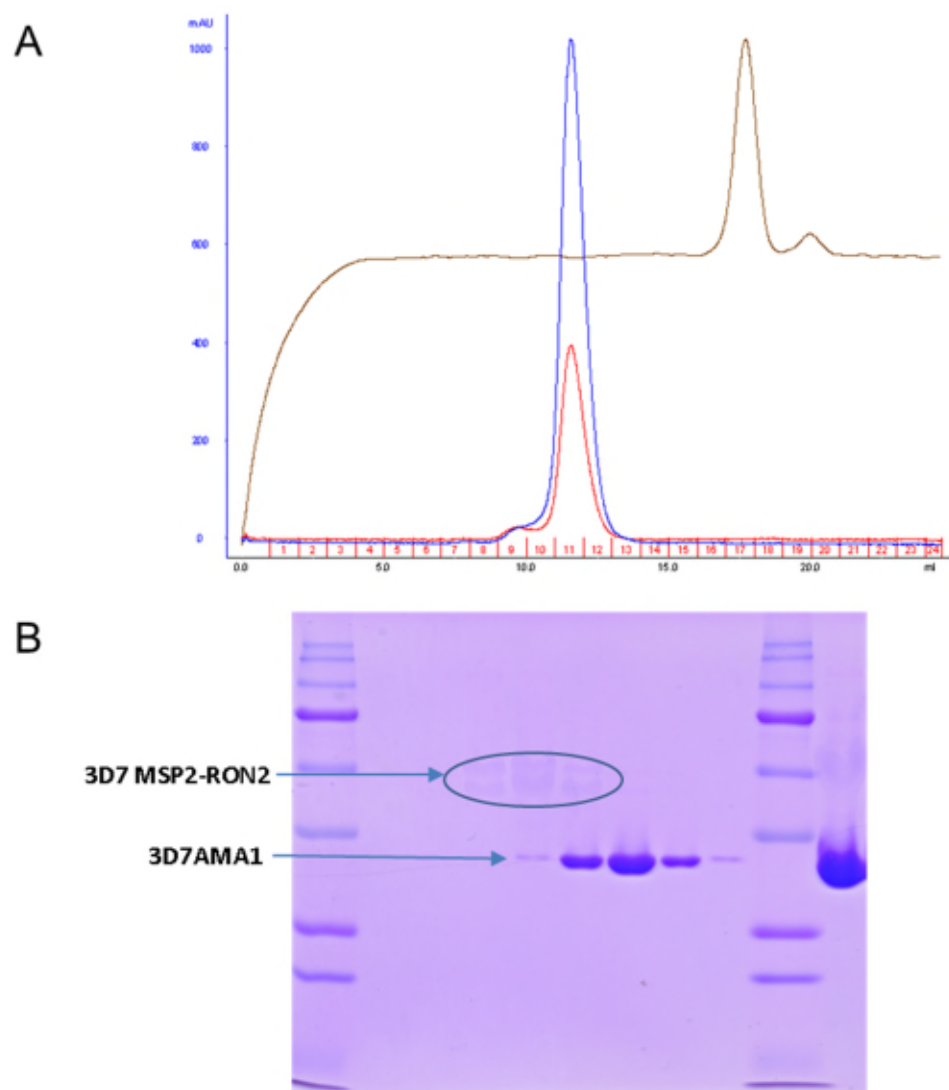


Figure IV.3. Purification and the analysis of 3D7 MSP2-*Pf*RON2+*Pf*AMA1 complex. A) SEC and B) SDS-PAGE analysis of the fractions. Blue and red peaks corresponding to absorbance at 280 and 214 nm, respectively. Orange peak indicates conductivity.

Appendix V

Results for Chapter 7

Table showing the backbone amide resonance assignments for 6D8 scFv. Blanks in the table indicates that assignments are unavailable.

		H	N	CA	CB
1	Gln	-	-	55.69	29.19
2	Val	7.42	123.55	61.92	31.39
3	Gln	8.35	124.68	54.86	32.10
4	Leu	8.88	127.10	55.73	41.75
5	Gln	9.05	122.43	54.35	30.05
6	Gln	9.87	129.00	55.57	31.37
7	Ser	7.85	114.68	58.49	63.86
8	Gly	8.57	106.91	45.07	-
9	Asp	-	-	55.82	41.07
10	Glu	8.54	119.98	55.30	33.39
11	Leu	8.50	127.74	54.41	42.39
12	Val	8.83	122.53	59.14	34.83
14	Pro	-	-	63.18	30.59
15	Gly	10.03	113.41	44.85	-
16	Ala	7.76	124.29	51.59	19.85
17	Ser	7.86	111.10	56.67	66.31
18	Val	8.41	117.59	59.71	35.05
19	Lys	8.47	126.39	55.46	33.69
20	Leu	9.13	129.75	54.00	41.92
21	Ser	9.10	113.86	57.11	67.67
22	Cys	8.82	124.36	56.30	46.89
23	Thr	9.14	125.95	63.05	70.10
24	Val	8.72	130.36	62.11	32.07
25	Ser	9.35	123.73	57.23	65.21
26	Gly	-	-	45.46	-
27	Phe	7.84	118.46	56.89	42.39
28	Asn	8.26	125.43	52.56	38.45
29	Ile	8.48	120.21	63.72	37.06
30	Lys	-	-	58.60	31.19
31	Asp	8.16	117.86	55.60	40.95
32	Asp	7.72	118.57	53.59	44.15

33	Phe	8.76	114.91	57.94	40.81
34	Ile	8.85	122.83	59.58	37.84
36	Trp	-	-	55.43	32.33
37	Met	9.43	120.94	53.98	35.94
38	Lys	9.31	122.24	54.08	36.96
39	Gln	9.39	125.88	54.96	31.95
40	Arg	8.81	130.02	54.62	30.19
41	Pro	-	-	64.84	31.18
42	Glu	8.83	116.17	58.27	27.70
43	Gln	7.95	117.99	55.55	29.95
44	Gly	8.17	106.97	44.93	-
45	Leu	7.18	121.00	54.86	43.27
46	Glu	-	-	54.99	34.16
47	Trp	9.65	126.59	58.99	30.77
48	Ile	8.68	127.42	62.44	38.00
49	Gly	5.94	102.18	45.21	-
50	Arg	8.25	112.23	53.02	32.68
53	Pro	-	-	63.29	32.91
54	Ala	8.32	115.56	54.66	18.23
55	Asn	6.99	109.97	52.29	39.48
56	Gly	8.74	110.18	45.72	-
57	Tyr	8.20	122.14	58.51	38.01
58	Thr	7.77	115.25	58.69	72.55
59	Lys	8.04	120.94	54.73	34.74
60	Tyr	8.64	122.02	56.17	40.52
61	Ala	8.94	127.78	50.94	17.71
64	Phe	-	-	58.50	40.87
65	Gln	7.45	121.23	58.68	28.11
66	Asp	8.58	117.41	55.49	40.38
67	Lys	7.46	119.64	56.01	33.24
68	Ala	7.74	120.58	49.97	21.76
69	Thr	8.60	113.17	61.84	71.48
70	Met	9.01	129.62	55.11	34.35
71	Thr	8.15	110.65	60.26	73.72
72	Ala	9.08	121.70	52.54	20.43
73	Asp	9.18	121.02	52.43	42.49
74	Thr	-	-	64.59	67.74
75	Ser	8.63	119.37	61.46	63.00

76	Ser	7.41	113.69	57.51	64.88
77	Asn	8.17	127.09	54.51	39.64
78	Thr	7.75	111.10	62.17	72.04
79	Ala	8.85	127.88	49.61	21.94
80	Tyr	8.81	117.49	56.61	41.53
81	Leu	9.51	120.92	57.80	40.82
82	Gln	8.81	127.73	53.45	30.33
83	Leu	8.59	127.01	52.90	44.76
84	Ser	8.22	116.78	57.06	65.88
85	Ser	8.23	115.75	57.23	61.93
86	Leu	8.39	119.68	56.80	41.60
87	Ala	9.78	125.80	49.91	22.29
90	Asp	-	-	55.21	42.75
91	Ala	7.59	125.33	53.32	18.76
92	Ala	8.55	125.46	51.53	21.62
93	Val	7.95	118.46	61.81	32.17
94	Tyr	8.90	124.23	56.52	41.04
95	Tyr	9.59	121.83	57.51	44.16
96	Cys	7.02	113.54	52.44	46.83
97	Ala	8.23	120.67	51.21	21.45
98	Thr	8.84	113.35	60.76	69.21
101	Val	-	-	61.34	30.60
102	Ala	7.61	120.21	53.74	20.07
103	Tyr	8.31	117.96	55.81	39.42
104	Trp	9.03	123.41	56.01	31.41
105	Gly	9.00	109.08	45.05	-
106	Gln	8.70	115.92	56.94	27.40
107	Gly	7.55	109.89	44.89	-
108	Thr	-	-	63.24	72.36
109	Leu	9.10	131.04	54.98	42.10
110	Val	8.92	128.64	61.50	33.31
111	Thr	8.75	122.60	61.44	70.07
112	Val	7.73	124.76	60.08	33.03
113	Ser	8.77	123.64	58.61	63.77
114	Ala	8.11	122.71	53.78	19.29
129	Ser	-	-	58.79	64.13
130	Thr	8.26	113.56	61.30	69.25
131	Asp	7.85	123.91	54.10	40.52

132	Ile	8.95	122.09	62.99	35.88
133	Val	8.52	132.07	62.27	33.04
134	Leu	9.17	130.77	52.98	41.87
135	Thr	9.41	119.92	61.62	69.86
136	Gln	9.74	130.23	54.59	30.24
137	Ser	8.78	117.93	55.50	65.48
140	Ser	7.79	111.71	56.93	65.01
141	Leu	8.99	125.64	54.75	44.97
142	Ala	8.52	128.43	50.14	19.56
143	Val	8.54	118.48	58.92	35.14
144	Ser	8.54	119.71	58.14	64.31
145	Leu	8.35	121.69	56.49	40.94
146	Gly	8.73	111.23	45.07	-
147	Gln	7.57	118.05	55.07	28.80
148	Arg	8.07	119.33	54.87	31.72
149	Ala	9.09	128.24	50.08	21.97
150	Thr	7.86	118.17	62.00	70.94
151	Ile	9.07	127.58	60.82	40.53
152	Ser	9.00	120.44	57.72	65.64
153	Cys	9.34	124.44	56.58	-
154	Lys	8.83	128.15	53.94	34.80
155	Ala	9.58	128.81	49.89	22.59
156	Ser	8.60	115.05	60.60	63.02
157	Gln	7.34	117.48	53.84	32.30
158	Ser	8.22	111.16	58.45	63.41
159	Val	8.20	118.99	59.83	30.18
160	Asp	7.90	121.63	53.54	41.71
161	His	8.97	117.87	56.86	30.08
162	Asp	-	-	55.57	39.81
163	Gly	8.63	105.98	45.47	-
164	Asp	7.73	120.34	53.16	43.92
165	Ser	8.74	116.52	57.93	65.16
166	Tyr	8.05	127.26	57.97	35.12
167	Met	6.58	119.58	52.07	33.53
168	Asn	9.41	125.37	50.67	41.64
169	Trp	8.68	115.92	55.90	33.70
170	Phe	9.82	118.69	56.53	44.00
171	Gln	9.34	121.89	53.74	34.40

172	Gln	9.58	127.84	55.07	31.50
174	Pro	-	-	64.19	31.04
175	Gly	8.69	113.04	45.55	-
176	Gln	8.03	118.90	53.81	31.47
177	Ser	8.42	115.41	56.68	62.38
178	Pro	-	-	62.74	30.64
179	Lys	8.77	121.26	53.95	35.03
180	Leu	9.03	129.44	55.97	41.43
181	Leu	-	-	55.08	45.05
182	Ile	6.94	116.36	57.51	43.01
183	Tyr	8.95	122.75	53.83	42.03
184	Ala	8.14	124.92	53.40	16.22
185	Ala	9.50	116.16	57.29	19.52
186	Ser	8.71	110.50	57.61	66.14
187	Asn	7.93	122.59	53.15	38.43
188	Leu	8.57	126.14	55.39	42.71
189	Glu	8.07	126.51	55.89	-
193	Pro	-	-	63.11	33.03
194	Ala	8.33	120.91	53.42	17.20
195	Arg	6.87	113.94	56.70	29.41
196	Phe	7.73	119.43	57.85	40.12
197	Ser	8.78	112.07	57.23	65.79
198	Gly	8.93	111.43	43.98	-
199	Ser	8.87	114.70	57.11	66.26
200	Gly	8.38	110.85	44.95	-
201	Ser	6.98	109.26	58.03	64.73
202	Gly	9.46	112.36,	47.98	-
203	Thr	8.36	116.05	61.16	70.86
204	Asp	7.30	122.83	54.75	43.38
205	Phe	8.56	121.34	57.47	42.38
206	Thr	9.00	115.99	61.09	72.61
207	Leu	8.75	127.16	53.08	40.95
208	Asn	9.00	124.51	51.63	43.10
209	Ile	8.68	121.11	59.98	39.17
210	His	8.61	123.56	53.46	33.04
211	Pro	-	-	62.46	33.60
212	Val	8.12	119.28	63.50	31.88
213	Glu	9.01	128.52	54.34	31.84

214	Glu	9.05	121.57	60.44	29.09
215	Glu	8.40	114.01	57.04	28.61
216	Asp	7.94	119.32	54.72	40.30
217	Ala	7.17	124.43	54.08	17.63
218	Ala	8.23	125.97	51.55	20.36
219	Thr	8.09	114.88	62.30	68.78
223	Gln	-	-	54.51	32.46
224	Gln	8.38	125.06	51.98	28.50
225	Thr	7.53	109.62	58.70	67.20
226	Asn	8.49	121.13	55.50	40.70
227	Glu	8.17	116.13	54.02	33.11
228	Asp	8.14	120.30	52.50	40.28
230	Tyr	-	-	54.77	34.70
231	Thr	7.18	109.18	59.48	73.38
232	Phe	8.81	116.82	56.80	43.42
233	Gly	9.15	108.95	44.66	-
234	Gly	8.40	104.14	46.35	-
235	Gly	7.08	106.39	44.75	-
236	Thr	8.11	118.52	61.73	72.90
237	Lys	8.44	129.97	56.46	32.64
238	Leu	9.17	131.06	54.52	44.39
239	Glu	9.16	127.25	54.19	32.69
240	Ile	8.58	122.19	58.13	36.32
241	Lys	8.45	128.84	55.65	33.16
242	Gly	8.59	111.97	44.95	-
243	Ser	8.30	115.90	58.43	64.06