Reconstructing ICAM-1-binding *var* genes of *Plasmodium falciparum*

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Eilidh Carrington

April 2016

Abstract

University of Liverpool, Eilidh Carrington for the Degree of Doctor of Philosophy, 2016.

Reconstructing ICAM-1-binding *var* genes of *Plasmodium falciparum*

Malaria is a major cause of morbidity and mortality globally. Severe malaria in patients infected with the most deadly species, *Plasmodium falciparum*, involves sequestration of *P. falciparum*-infected erythrocytes in the microvasculature, a process involving the interaction of numerous parasite and host factors. The highly variable gene family *var* encodes PfEMP1 proteins which are thought to be responsible for the majority of endothelial cell binding via receptors such as CD36, ICAM-1 and EPCR. To date, knowledge of the specific PfEMP1 domains that mediate binding to these receptors has emerged from the study of two reference parasite strains IT4 and 3D7.

The aim of this study is to identify and characterise new ICAM-1 binding *var* genes from several culture-adapted ICAM-1-selected patient isolates. Sequencing *var* genes is complicated by their high variability within the 50-60 gene copies per genome and the absence of overlap between *var* repertoires of different isolates. We test the predictive powers of a new *var* gene database of > 62,000 *var* genes generated by the Pf3k genome sequencing project. We find database matches to be a useful starting point for primer design in this targeted sequencing approach. However, key differences between database genes and those amplified from the patient isolates highlights the need for experimental validation. We perform sequence analysis of *var* genes from three patient isolates, BC12, J1 and PCM7, and identify their PfEMP1 domain structures. We extensively characterise the DBLβ domains, which are responsible for ICAM-1 binding, of these genes by phylogenetic and homology block analysis and compare them to known ICAM-1 binding DBLβ domains.

We express two of the newly identified DBLβ domains as recombinant proteins and analyse their binding to ICAM-1. Using SPR assays, we show that the recombinant DBLβ domains of J1a and BC12a bind ICAM-1^{D1D5} with nanomolar affinity. We compared their binding to the previously characterised IT4var13$^{DBL\beta}$ and studied the binding of all three DBLβ domains to four ICAM-1 mutant proteins. Residues K29, L42 and L44 are essential for ICAM-1 binding of all three DBLβ domains, whereas S22 is essential for J1a$^{DBL\beta}$ and BC12a$^{DBL\beta}$ binding but not IT4var13$^{DBL\beta}$ binding. These results have important implications for the design of anti-disease therapies targeting severe malaria.

To Granny, Gags and She

To Mum and Dad

To Jo and Eve

For everything

Acknowledgements

I would like to extend my sincere thanks to Professor Alister Craig for his supervision, guidance and advice. I particularly wish to thank him for his open-door policy, prompt e-mail responses, encouragement to attend conferences and his suggestion to enrol on the Biology of Parasitism course, a life- and career-changing experience. My thanks also go to Dr Britta Urban, Dr Simon Wagstaff and Dr Emily Adams for support and advice, with particular thanks to Simon for generating the seven genomes BLAST database.

I would like to thank Dr Yang Wu, Mr Tadge Szestak, Dr Aymen Madkhali and Dr Janet Storm for support, discussions and technical advice throughout my PhD. Special thanks go to Janet for proof reading the thesis and support during the write-up. Thanks also go to Dr Chris Moxon, Mr Basim Othman and Dr Laura Cruz and all of the above for friendship and providing a great working environment. Special thanks go to Mary Creegan for her incredible organisational skills and advice.

I would like to thank Dr Frank Lennartz for his invaluable advice on protein expression and SPR and for carrying out CD spectroscopy and one SPR experiment (IT4var13$^{DBL\beta}$ binding to ICAM-1^{S22A}). I thank Professor Matthew Higgins and his group for allowing me to work in their incredible facility and for providing the expression plasmid. Thanks go to Dr Thomas Otto for his database searches and Professor Chris Newbold and MalariaGEN for granting database access.

I would like to thank all my friends and colleagues at LSTM, in particular Glauber, Hanna, Eva, Darren, Angela, Ricky, Ghaith, Dave, Grazia, Adriana, Laura, Kayla, Paul Gavin, Jill and Alison for their friendship and support throughout my PhD. I thank the BoP class and staff of 2014 for a great summer of science, singing and dancing.

I thank my family for their unwavering support and encouragement, for long phone conversations, much needed holidays and for generally being amazing. Special thanks to my oldest friend Alison who always understands.

Finally, I thank Chris Spencer for an incredible three years in the Carrington-Spencer household. Anyone who can tolerate my singing for so long deserves a medal.

Contents

Abbreviations

ACT – artemisinin based combination therapy
APC – activated protein C
ASL – adenylosuccinate lyase
ATS – acidic terminal segment
BBB – blood brain barrier
BLAST – basic local alignment search tool
BSA – bovine serum albumin
CD – circular dichroism
CIDR – cysteine-rich interdomain region
CM – cerebral malaria
CR1 – complement receptor 1
CSA – chondroitin sulphate A
Ct – cycle threshold
cys/PoLV – cysteine/positions of limited variability
DBL – Duffy binding-like
DC – domain cassette
DEPC-H_2O – diethylpyrocarbonate treated water
dNTP – deoxyribonucleotide triphosphates
EBL – erythrocyte binding ligand
EPCR – endothelial protein C receptor
GLURP – glutamate-rich protein
HB – homology block
HDX MS – hydrogen/deuterium exchange mass spectroscopy
HMEC – human microvascular endothelial cells
ICAM-1 – intercellular adhesion molecule 1
IE – *Plasmodium falciparum* infected erythrocyte
Ig – immunoglobulin
IPTG – isopropyl β-D-1-thiogalactopyranoside
IRS – indoor residual spraying
ITN – insecticide treated bed net
KAHRP – knob-associated histidine-rich protein
k_a – association rate constant
k_d – dissociation rate constant
K_D – equilibrium dissociation rate constant (k_d/k_a)
kDa – kilodaltons
LFA-1 – lymphocyte function associated molecule 1
MAHRP1 – membrane-associated histidine-rich protein 1
MC – Maurer's cleft
MSP – merozoite surface protein

Ni-NTA – nickel nitrilotriacetic acid
NO – nitric oxide
no-RT – no reverse transcriptase
NTS – N-terminal segment
PAM – pregnancy associated malaria
PBS – phosphate buffered saline
PCR – polymerase chain reaction
PECAM-1 – platelet endothelial cell adhesion molecule 1
PEXEL/HT – *Plasmodium* export element/host targeting peptide
PfEMP1 – *Plasmodium falciparum* erythrocyte membrane protein 1
PfEMP3 – *Plasmodium falciparum* erythrocyte membrane protein 3
PfHDA2 – *Plasmodium falciparum* histone deacetylase 2
PfHP1 – *Plasmodium falciparum* heterochromatin protein 1
PfMC-2TM – *Plasmodium falciparum* Maurer's cleft two transmembrane protein
PfSBP1 – *Plasmodium falciparum* skeleton binding protein 1
Pf3k – *Plasmodium falciparum* 3000 genomes sequencing project
PHIST – *Plasmodium* helical interspersed subtelomeric protein
PNEP – PEXEL-negative exported protein
PoLV – positions of limited variability
PTEX – *Plasmodium* translocon complex
PV – parasitophorous vacuole
PVM – parasitophorous vacuole membrane
qPCR – quantitative polymerase chain reaction
RBC – red blood cell
RBL – reticulocyte binding ligand
RIFIN – repetitive interspersed family protein
RT – room temperature
RT-qPCR – reverse transcriptase quantitative polymerase chain reaction
SAXS – small angle x-ray scattering
SDS-PAGE – sodium dodecyl sulphate polyacrylamide gel electrophoresis
SM – severe malaria
SMA – severe malarial anaemia
SPR – surface plasmon resonance
STEVOR – subtelomeric variable open reading frame protein
STS – seryl tRNA synthetase
SURFIN – surface associated interspersed gene family protein
TM – transmembrane
UM – uncomplicated malaria
UPS – 5' upstream sequence
VCAM-1 – vascular cell adhesion molecule 1
VES – *var* expression site

VSA – variable surface antigen
WGS – whole genome sequencing
WHO – World Health Organisation

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Malaria prevalence and control

Malaria is a vector borne disease caused by species of the *Plasmodium* parasite. It occurs throughout the tropics and subtropics with distribution depending on suitable climatic conditions to support the *Anopheles* mosquito vectors and parasite development. In 2013, an estimated 3.3 billion people worldwide were at risk of malaria infection (WHO, 2014), with the percentage of at risk populations varying between countries (Figure 1.1). The latest figures place the estimated number of global malaria cases at 214 million (range: 149-303 million) and the number of estimated malaria deaths at 438,000 (range: 236,000-635,000) (WHO, 2015). This represents an 18% reduction in cases and a 48% reduction in deaths since 2000, largely due to widespread control programmes (WHO, 2015). Control strategies include distribution of insecticide-treated bed nets (ITN) which prevent mosquito bites at night, indoor residual spraying (IRS) which targets mosquitoes that land on indoor surfaces and treatment of clinical malaria cases using artemisinin based combination therapy (ACT) which has an enhanced impact on severe malaria (SM) and death compared to quinine-based drugs (Dondorp *et al.*, 2010). Of these control strategies, ITN has been the most successful with 68% of the decline in cases from 2000-2015 attributed to this strategy (Bhatt *et al.*, 2015). However, this is largely due to the early implementation of ITNs and the scale of distribution that has been achieved and does not necessarily identify ITNs as a more effective strategy than either IRS or ACT (Bhatt *et al.*, 2015). Implementation of IRS and ACT at a larger scale and continued ITN distribution will be required to continue the declining trend of malaria prevalence. However, threats to these programmes are increasing with the emergence of insecticide resistance (Ranson and Lissenden, 2016) and widespread artemisinin resistance in South East Asia (Fairhurst, 2015).

The majority of malaria cases (88%) and deaths (90%) occur in the WHO African Region (WHO, 2015). This region has the most unreliable malaria reporting system and incidence figures are predicted by mathematical modelling based on

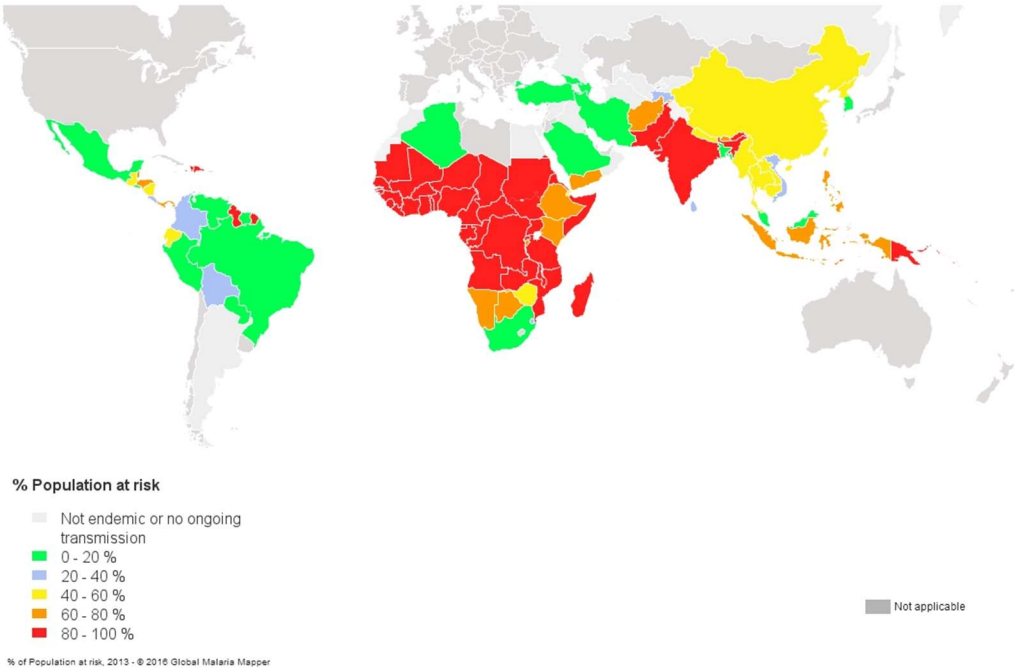


Figure 1.1. Global population at risk of malaria infection, 2013.
The percentage of the population at risk of malaria infection is shown for each country (WHO, 2014). The map was obtained from the WHO Global Malaria Mapper available at http://www.worldmalariareport.org/, accessed 14/04/2016.

prevalence data (WHO, 2015). Despite the increasing sophistication of these models, which can be calibrated to empirical data (Cameron *et al.*, 2015), they still rely on figures based on small areas which are then extrapolated to a country wide level and therefore remain unreliable. Despite this, these estimated figures can still be useful in planning control programmes and, as these become more widespread, reporting systems are also likely to improve. The majority of malaria deaths in sub-Saharan Africa occur in children under five, accounting for 10% of all child deaths in 2015 (WHO, 2015). Therefore, despite the positive effect of control programmes in reducing disease, malaria is still a major cause of morbidity and mortality, particularly in sub-Saharan Africa. New insecticides and drugs are required to counteract resistance but new strategies such as vaccines and improved diagnostics will be required to increase the likelihood of successful malaria control and elimination (Hemingway *et al.*, 2016).

1.2 Clinical symptoms of malaria

Malaria disease encompasses a range of clinical symptoms that can be divided into uncomplicated malaria (UM) and severe malaria (SM). UM, also referred to as mild malaria, is characterised initially with headaches, fatigue, muscle aches and abdominal pain, followed by fever, nausea, vomiting and postural hypotension (White *et al.*, 2014). Of the six *Plasmodium* species known to infect humans, *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale curtisi* and *P. ovale wallikeri* (Sutherland *et al.*, 2010) and *P. knowlesi* (Brock *et al.*, 2016), *P. falciparum* is responsible for the majority of cases of SM. This is related to the unique ability of *P. falciparum* infected erythrocytes (IEs) to sequester in the microcirculatory system of the human host causing obstruction of blood flow and interfering with endothelial cell function and metabolism (White *et al.*, 2014). IE sequestration results in the ability of later stage parasites to avoid the spleen, where they would otherwise be destroyed. SM consists of several disease phenotypes that are often grouped into severe malarial anaemia (SMA), respiratory distress and cerebral malaria (CM) but can overlap. SM symptoms vary between young children and adults with mainly the former described here.

SMA is common in high transmission areas where repeat infection often occurs. SMA is characterised by haemoglobin levels < 5 g per decilitre of blood in the presence of *P. falciparum* parasitaemia (Calis *et al.*, 2008). Both erythrocyte destruction upon parasite release and an increase in splenic clearance of erythrocytes, infected and uninfected, contribute to SMA (White *et al.*, 2014). In addition, erythropoiesis is affected resulting in a decrease in erythrocyte production (Dormer *et al.*, 1983). Platelet deficiency is also common and coagulation pathways are affected (White *et al.*, 2014).

Respiratory distress often results from metabolic acidosis. This refers to the build up of lactic acid, a by-product of anaerobic glycolysis, in the blood which occurs when insufficient oxygenated blood is available due to parasite sequestration (White *et al.*, 2014). Other causes of respiratory distress are pneumonia, fluid overload, acute lung injury and severe anaemia (Taylor *et al.*, 2012). Respiratory

distress is a key identifier of *P. falciparum* infected children at risk of death (Marsh *et al.*, 1995).

CM is classified as an unrousable coma, defined as a score of ≤ 2 on the Blantyre Coma Score in children (Molyneux *et al.*, 1989), in the presence of *P. falciparum* parasitaemia and no other apparent cause of coma. However, ruling out other causes of coma can be difficult and one study found that in 23% of CM patients the cause of death was not CM (Taylor *et al.*, 2004). One method of differentiation is detection of retinopathy which is present in CM but not other causes of coma, resulting in a more accurate clinical definition of CM (Taylor and Molyneux, 2015). Pathogenesis of CM is not entirely clear but is thought to be related to parasite sequestration in the brain, disruption of the blood brain barrier (BBB) and endothelial cell activation, processes that involve a complex interplay of parasite and host components. Brain swelling and the associated increase in intracranial pressure are associated with fatal outcomes of CM, with two distinct clinical groups identified suggesting that different mechanisms lead to brain swelling (Seydel *et al.*, 2015).

Sequestered parasites are found both in the brains of patients who died of CM and those who died of other causes but the number of parasitized vessels and number of sequestered parasites is higher in CM (MacPherson *et al.*, 1985, Dorovini-Zis *et al.*, 2011). However, there is also variation in pathology observed in those with true CM, namely brain swelling and parasitized vessels (CM1) and brain swelling, parasitized vessels and presence of haemorrhaging and thrombi (CM2) (Taylor *et al.*, 2004), suggesting either different mechanisms leading to these pathologies or the death of patients at different stages of the same process. BBB breakdown can be shown by detecting fibrinogen in the brain, which is excluded if BBB is intact (Dorovini-Zis *et al.*, 2011). The mechanisms involved are not fully understood but are thought to include parasite sequestration and cytokine release (Dorovini-Zis *et al.*, 2011). Endothelial activation is linked to BBB breakdown and is present in CM but also UM making it difficult to decipher cause and effect (Storm and Craig, 2014). Angiopoietein-1 and -2, which stabilise and antagonise endothelium respectively, von Willebrand factor, increase in expression of endothelial cell receptors and

soluble endothelial receptors are all associated with endothelial activation (Storm and Craig, 2014). The study of host endothelial cell-parasite interactions is a key focus of current research in the field and will provide insight into the molecular mechanisms leading to CM. Patients that recover from CM often display neurological symptoms, such as epilepsy, cerebral palsies, cognitive impairments and language deficits, many of which resolve within six months (Idro *et al.*, 2010, Birbeck *et al.*, 2010). However, long-lasting effects have been reported (Carter *et al.*, 2005), the full impact of which require further study.

Many clinical symptoms of severe malaria have been linked to IE sequestration in the microcirculatory system. We will now look more closely at the *P. falciparum* parasite and the molecular mechanisms surrounding this phenomenon.

1.3 *Plasmodium falciparum* life cycle

P. falciparum is an Apicomplexan parasite which has a complex life cycle requiring human and mosquito stages for completion (Figure 1.2). Sporozoites are stored in the salivary glands of the female *Anopheles* mosquito vector and injected into the skin of the human host upon blood feeding. Sporozoites then migrate via blood vessels to the liver where they first traverse (Tavares *et al.*, 2013) and then invade hepatic cells. A single sporozoite then undergoes multiple rounds of asexual replication within a vacuole to produce tens of thousands of merozoites. Merozoites are released in vesicles called merosomes, after parasite-induced cell death of the hepatocyte (Sturm *et al.*, 2006), which protect merozoites until they reach the microvasculature of the lung where they are released into the bloodstream (Baer *et al.*, 2007). Merozoites invade erythrocytes within seconds of release in a complex process involving many parasite-host protein interactions which can be divided into four key stages: initial attachment and re-orientation, tight junction formation, invasion via an actin-myosin motor and completion by membrane sealing and shedding of surface proteins (recently reviewed in (Beeson *et al.*, 2016)). The parasite then undergoes an asexual cycle within the resulting parasitophorous vacuole (PV) inside the erythrocyte consisting of early ring trophozoite, late trophozoite, developing schizont and developed schizont stages

culminating in merozoites ready for re-invasion. Release of merozoites and rupture of the IE occurs in synchronisation after every ~ 48 hour cycle and contributes to inflammatory response and clinical symptoms observed. A small proportion of parasites are committed to forming gametocytes which are required for onward transmission.

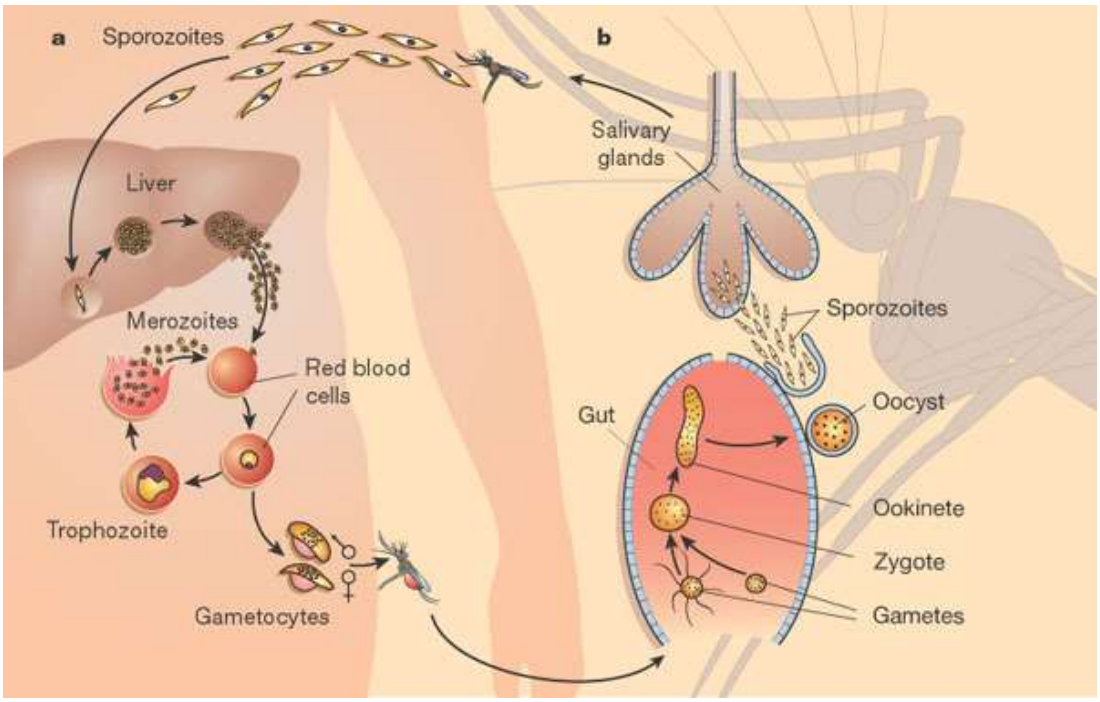


Figure 1.2. Life cycle of *Plasmodium falciparum*.
Life cycle stages in the human (a) and mosquito (b) hosts. See text for a detailed description. Figure is reprinted by permission from Macmillan Publishers Ltd: [Nature] (Wirth, 2002), copyright (2002).

Only the presence of mature gametocytes in a blood meal of the mosquito vector can facilitate the necessary sexual development which occurs in this host (Figure 1.2b). Once in the mosquito gut, gametocytes egress from the erythrocyte and form gametes, which for males includes an exflagellation process in which DNA is replicated and eight microgametes are produced. Fertilisation occurs to produce a zygote which undergoes DNA replication before differentiating into an ookinete which traverses the midgut epithelium by gliding motility. Once in the extracellular

matrix on the haemocoel side of the epithelium, the ookinete differentiates into an oocyst which produces thousands of sporozoites by sporogony. These then invade the salivary gland via the haemolymph where they reside until the mosquito takes a blood meal, completing transmission to the next human host (Smith *et al.*, 2014).

1.4 Modification of the host erythrocyte

Erythrocytes are highly specialised for gaseous exchange throughout the tissues and are unusual in that they do not contain organelles. As a result, they do not synthesise and traffic proteins. Therefore, in order for *P. falciparum* to survive within erythrocytes they extensively modify both the cytoplasm and surface of their host cells. These modifications begin during the early ring stage and result in the appearance of several structures including Maurer's clefts (MC) (Mundwiler-Pachlatko and Beck, 2013), the tubovesicular network (Elmendorf and Haldar, 1994), J-dots (Kulzer *et al.*, 2010) and knobs (Gruenberg *et al.*, 1983). These modifications are thought to collectively facilitate nutrient acquisition and localisation of parasite adherence proteins on the IE surface and reduce erythrocyte deformability as a consequence. The major adherence protein found on the IE surface is *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) which is a key protein involved in IE sequestration in the host tissues. Many of the modifications are involved in protein trafficking systems, although specific mechanisms are only just beginning to be elucidated.

Proteins exported into the erythrocyte fall into two categories, those which contain the *Plasmodium* export element/host-targeting (PEXEL/HT) targeting motif (Marti *et al.*, 2004, Hiller *et al.*, 2004) and those that do not, termed PEXEL-negative exported proteins (PNEPs) (Heiber *et al.*, 2013). PEXEL/HT proteins are processed by Plasmepsin V in the endoplasmic reticulum (Russo *et al.*, 2010, Boddey *et al.*, 2013) and both protein categories are transported across the PV membrane (PVM) into the erythrocyte cytoplasm by the *Plasmodium* translocon complex (PTEX) (de Koning-Ward *et al.*, 2009). From there, the exact mechanism of export to the IE cell surface is unclear but MCs are thought to act as intermediate sorting centres due to only transient localisation of PfEMP1 here (Mundwiler-Pachlatko and Beck, 2013).

MCs are not linked to the PVM or the erythrocyte membrane (McMillan *et al.*, 2013, Gruring *et al.*, 2011) and so a vesicular system has been proposed to traffic proteins to and from the MCs, supported by the identification of trafficking compartments (Hanssen *et al.*, 2010) and J-dots, which contain chaperone proteins (Kulzer *et al.*, 2012). PfEMP1 proteins are structurally similar to PNEPs but contain a modified PEXEL motif (Marti *et al.*, 2004). Two proteins that have been shown to be important in PfEMP1 trafficking, *P. falciparum* skeleton binding protein 1 (PfSBP1) (Blisnick *et al.*, 2000, Cooke *et al.*, 2006) and PfEMP1 trafficking protein 1 (PfPTP1) (Rug *et al.*, 2014) form a complex and are required for PfEMP1 trafficking to the MCs (Rug *et al.*, 2014). Membrane-associated histidine-rich protein 1 (MAHRP1) is also important in transfer to MCs with depletion of this protein resulting in PfEMP1 accumulation at the PVM (Spycher *et al.*, 2003, Spycher *et al.*, 2008). Several other proteins are important in PfEMP1 trafficking but have not been characterised in detail (Maier *et al.*, 2009). Once at the IE surface, PfEMP1 is found in clusters on knobs.

Knobs are protrusions of the IE membrane which provide contact points for adherence to host endothelium. Knob-associated histidine-rich protein (KAHRP) is a major knob component and is required for their formation (Crabb *et al.*, 1997). KAHRP has a role in anchoring knobs to the erythrocyte cytoskeleton via interactions with cytoskeleton components spectrin (Pei *et al.*, 2005) and ankyrin R (Weng *et al.*, 2014). PfEMP3 is another component of knobs whose exact role is unclear but both PfEMP3 and KAHRP are involved in PfEMP1 trafficking and functional display on the IE surface with disruption of either leading to reduction of cytoadherence (Waterkeyn *et al.*, 2000, Crabb *et al.*, 1997). KAHRP was originally thought to bind the acidic terminal sequence (ATS) of PfEMP1, linking it to the cytoskeleton (Kilejian *et al.*, 1991, Oh *et al.*, 2000, Waller *et al.*, 1999). However, recent evidence suggests that members of the *Plasmodium* helical interspersed subtelomeric (PHIST) protein family carry out this function (Mayer *et al.*, 2012, Oberli *et al.*, 2014, Proellocks *et al.*, 2014), with co-operation between PHIST proteins and a level of PfEMP1 type specificity suggested but not yet proven (Oberli *et al.*, 2014, Oberli *et al.*, 2016). Disruption of one of these proteins resulted in a

decrease in IE adhesion but did not affect knob formation or PfEMP1 trafficking (Proellocks *et al.*, 2014). There is still much to elucidate about the mechanisms that lead to PfEMP1 display on the IE surface. Once embedded in the membrane, the extracellular domains of PfEMP1 are exposed to the bloodstream where they can adhere to host receptors leading to sequestration of the parasite. Here, PfEMP1 also functions as the major parasite protein responsible for antigenic variation.

1.5 Antigenic variation

Many human pathogens are recognised and cleared by the immune system, while others have developed strategies to avoid recognition and establish long-lasting infections. One such strategy is to constantly change the proteins that are recognised by the host immune system, mainly surface exposed antigens, a process referred to as antigenic variation. Many pathogens achieve antigenic variation through sequential expression of multi-copy gene families, including bacteria (e.g. *Borrelia* and *Neisseria* spp), fungi (e.g. *Pneumocystis carinii* and *Candida* spp), and protozoan parasites (e.g. *Trypanosoma brucei* and *P. falciparum*) (Deitsch *et al.*, 2009). Antigenic variation in *P. falciparum* is mediated by surface proteins collectively known as variant surface antigen (VSA). PfEMP1 encoded by *var* genes, the most extensively studied VSA of *P. falciparum*, are the focus of this thesis and are introduced below. Other important VSA are repetitive interspersed family (RIFIN) proteins (Cheng *et al.*, 1998, Kyes *et al.*, 1999, Fernandez *et al.*, 1999), sub-telomeric variable open reading frame (STEVOR) proteins (Kaviratne *et al.*, 2002, Blythe *et al.*, 2008, Niang *et al.*, 2009) and possibly surface associated interspersed gene family (SURFIN) proteins (Winter *et al.*, 2005) and *P. falciparum* Maurer's cleft two transmembrane (PfMC-2TM) proteins (Sam-Yellowe *et al.*, 2004, Lavazec *et al.*, 2006).

RIFINs and STEVORs are encoded by *rif* and *stevor* genes, respectively, which are found in subtelomeric regions of all *P. falciparum* chromosomes (Gardner *et al.*, 2002). There are around 150-200 *rif* gene and 30-40 *stevor* gene copies per parasite genome of which there are both strain-specific and strain-transcending genes (Claessens *et al.*, 2011, Blythe *et al.*, 2008). RIFINs are small 30-45 kDa proteins that

are grouped into A- and B-types (Joannin *et al.*, 2008), both of which are expressed simultaneously with only A-types reportedly found at the IE surface (Petter *et al.*, 2007). They were originally thought to be involved in rosetting, the binding of IEs to uninfected erythrocytes (Kyes *et al.*, 1999, Fernandez *et al.*, 1999) but the identification of PfEMP1 as a rosette mediator (Rowe *et al.*, 1997) shifted research focus away from RIFINs. It is now thought that RIFINs mediate rosetting by binding to blood group A antigen on erythrocytes (Goel *et al.*, 2015). RIFINs are also expressed in sporozoite, merozoite and gametocyte stages suggesting they may play an important role throughout the parasite life cycle (Petter *et al.*, 2007, Wang *et al.*, 2010). Antibodies are generated against RIFINs that are associated with parasite clearance (Abdel-Latif *et al.*, 2002, Abdel-Latif *et al.*, 2003, Abdel-Latif *et al.*, 2004) and have been shown to be acquired rapidly in experimental infections in naive individuals (Turner *et al.*, 2011), supporting their role in antigenic variation.

STEVOR are similarly small, 30-40 kDa, proteins and have also been shown to have a role in rosetting (Niang *et al.*, 2014). STEVOR binds to uninfected erythrocytes via the glycophorin C receptor to form rosettes, a process that was shown to be protective against merozoite invasion blocking antibodies, contributing to parasite growth efficiency (Niang *et al.*, 2014). Additionally, STEVOR has a mechanical role at the IE membrane contributing to rigidity and enhancing IE sequestration (Sanyal *et al.*, 2012), a function which is also important in gametocyte sequestration (Tiburcio *et al.*, 2012). The role of anti-STEVOR antibodies in human infection is unclear as they were not found to be protective in children and are, therefore, speculated to be simply a marker of exposure (Schreiber *et al.*, 2008). More studies will be required to corroborate these findings. Overall, STEVORs and RIFINs appear to be secondary targets of the immune response against VSA with PfEMP1 remaining the predominant target (Chan *et al.*, 2012).

Little is known about the role of other VSAs. SURFINs are high molecular weight (280-300 kDa) proteins encoded by *surf* genes of which there are 10 gene copies (Winter *et al.*, 2005). *surf* genes are differentially expressed throughout the parasite life cycle (Mphande *et al.*, 2008) and SURFINs are reported to localise at the IE surface (Winter *et al.*, 2005). It is unknown if these proteins have a role in antigenic

variation. PfMC-2TM proteins are encoded by ~13 subtelomeric genes of which several are expressed at any one time (Sam-Yellowe *et al.*, 2004). They are relatively conserved with the exception of a short loop region which is exposed at the IE surface (Bachmann *et al.*, 2015). Whether this loop is the target of host antibodies has not been determined. In addition to parasite VSAs, *P. falciparum* infection induces a conformational change in an erythrocyte integral membrane protein, band 3 (Winograd and Sherman, 2004). Chemical modification of band 3 is reported to result in CD36 binding, thereby contributing to sequestration (Winograd *et al.*, 2004). A role in thrombospondin binding has also been reported (Eda *et al.*, 1999, Lucas and Sherman, 1998). However, the significance of modified band 3 is not understood.

1.6 *var* genes

The variant gene family encoding PfEMP1 proteins comprises 50-60 genes per genome (Gardner *et al.*, 2002, Rask *et al.*, 2010), collectively named *var* genes (Baruch *et al.*, 1995, Smith *et al.*, 1995, Su *et al.*, 1995). *var* genes are highly variable, not only within, but between parasite isolates resulting in vast global diversity (Rask *et al.*, 2010). They are classified by their 5' upstream sequence (UPS) into groups A, B, C and E with the relative abundance of each type similar between parasites (Gardner *et al.*, 2002, Lavstsen *et al.*, 2003, Kraemer and Smith, 2003, Rask *et al.*, 2010). Each parasite contains approximately 20% group A, 60% group B and 20% group C *var* genes (Rask *et al.*, 2010). Group E is unique in that it comprises only one gene, *var2csa*, whose protein product binds the chondroitin sulphate A (CSA) receptor of placental cells, resulting in pregnancy associated malaria (PAM) (Salanti *et al.*, 2003). *var2csa* is relatively conserved across isolates (Rask *et al.*, 2010) and appears to play an additional role in *var* gene regulation (Ukaegbu *et al.*, 2015). Group A and B genes are found in subtelomeric regions and are transcribed in opposite directions, while group C genes occupy a central chromosome position and are transcribed towards the telomere (Figure 1.3) (Gardner *et al.*, 2002). Two intermediate groups exist: B/A which are chimeric UPS B genes with UPS A gene

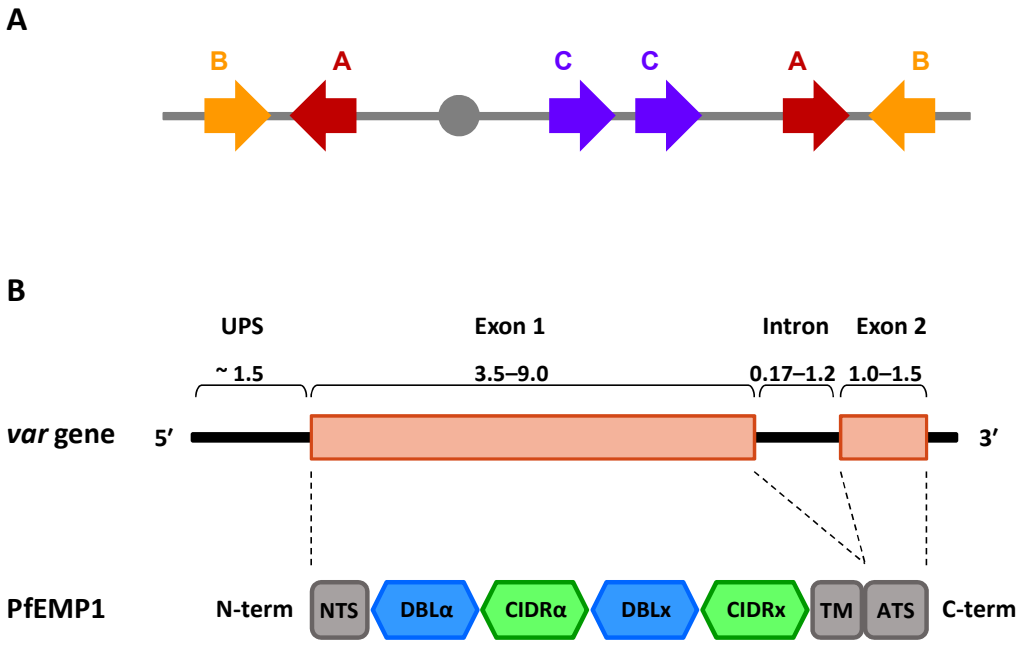


Figure 1.3. Chromosomal position and structure of *P. falciparum var* genes.
A, The chromosomal position and orientation of *var* gene subgroups. Group A and B *var* genes are subtelomeric and are transcribed in opposite directions as indicated. Group C *var* genes are found in central chromosome regions and are transcribed towards the telomere. The transitional group B/C *var* genes are upstream sequence (UPS) type B but are centrally located (not shown). **B**, The two exon structure of a *var* gene and the modular PfEMP1 protein it encodes. Numbers indicate length in kilobase pairs. Exon 1 encodes the extracellular region of PfEMP1 which consists of an N-terminal segment (NTS) and a variable number of Duffy binding-like (DBL) and cysteine-rich interdomain region (CIDR) domains. Exon 2 encodes the transmembrane (TM) and acidic terminal segment (ATS) domains.

properties and B/C which consist of centrally located UPS B genes (Lavstsen *et al.*, 2003). *var* genes consist of two exons: exon 1, which encodes the highly variable extracellular domains and the transmembrane (TM) domain, and exon 2 which encodes the semi-conserved intracellular acidic terminal sequence (ATS) (Figure 1.3).

1.7 *var* gene expression and regulation

Expression of *var* genes occurs via allelic exclusion whereby only one gene is transcribed and presented on the IE surface and switching of the expressed variant

occurs, resulting in clonal antigenic variation, a process regulated by epigenetic mechanisms (Scherf *et al.*, 1998). Expression is partly determined by the chromatin architecture with the active *var* gene in a euchromatic state and the remaining silenced *var* genes in areas of heterochromatin. Exact mechanisms of activation and switching between *var* genes are unknown but recent work implicates histone modifications, a *var* expression site (VES) and the *var* intron.

Silent *var* genes are marked by the histone modification H3K9me3 and *P. falciparum* heterochromatin protein 1 (PfHP1) (Flueck *et al.*, 2009, Lopez-Rubio *et al.*, 2007, Perez-Toledo *et al.*, 2009). The histone modifying enzymes involved are PfSIR2A which acts on UPS A, C and E *var* genes (Freitas-Junior *et al.*, 2005, Duraisingh *et al.*, 2005, Tonkin *et al.*, 2009, Merrick *et al.*, 2010), PfSIR2B which acts on UPS B genes (Tonkin *et al.*, 2009), and *P. falciparum* histone deacetylase 2 (PfHDA2) which acts on all genes (Coleman *et al.*, 2014). The active *var* gene is characterised by H3K9ac and H3K4me2/3 (Lopez-Rubio *et al.*, 2007) and the variant histones H2A.Z and H2B.Z (Petter *et al.*, 2013). The H3K36me3 modification, carried out by PfSET2, marks both the active *var* gene and silenced *var* genes but is depleted at the active gene during transcription, suggesting a role in regulation of expression of the active *var* gene within the 48 hour life cycle, i.e. keeping it in a poised state (Jiang *et al.*, 2013, Ukaegbu *et al.*, 2014). The histone methyltransferase PfSET10 has also been associated with the poised gene and is suggested to play a role in this transient repression (Volz *et al.*, 2012).

The expression of the active *var* gene is reported to occur at a specific site (*var* expression site, VES) and it is suggested that switching is mediated by transfer of a silent gene to and the active gene away from the VES (Lopez-Rubio *et al.*, 2009, Duraisingh *et al.*, 2005, Marty *et al.*, 2006, Ralph *et al.*, 2005, Voss *et al.*, 2006, Dzikowski *et al.*, 2007). This repositioning is mediated by actin which binds indirectly to the *var* gene intron (Zhang *et al.*, 2011a). The *var* intron plays a further role by pairing with the 5' upstream promoter region to carry out essential gene regulatory functions (Deitsch *et al.*, 2001, Voss *et al.*, 2006, Frank *et al.*, 2006, Swamy *et al.*, 2011). Additionally, the *var* intron transcribes both sense and

antisense non-coding RNA (Epp *et al.*, 2009) which has been shown to mediate UPS A gene silencing (Zhang *et al.*, 2014).

The exact molecular mechanisms involved in *var* gene switching remain unknown. However, switch rates between genes are variable (Fastman *et al.*, 2012, Horrocks *et al.*, 2004) with centrally located *var* genes more highly transcribed *in vitro* (Frank *et al.*, 2007, Peters *et al.*, 2007, Zhang *et al.*, 2011b, Fastman *et al.*, 2012). This is thought to be due to high activation bias towards centrally located *var* genes *in vitro* (Noble *et al.*, 2013). This activation bias supports theories that switching follows, to a certain extent, pre-programmed patterns (Recker *et al.*, 2011, Enderes *et al.*, 2011, Fastman *et al.*, 2012). A recent study of a newly culture-adapted field isolate supported the notion of variable switch rates between genes and showed that some switching patterns were maintained between clones but that switching from one dominant gene did not always lead to expression of another particular gene (Ye *et al.*, 2015). This presents a very complex view of *var* gene switching. It is likely that our study of *in vitro* cultures will not be sufficient to elucidate switching patterns as they may be influenced by a variety of host factors, such as immune pressure and receptor availability e.g. presence of CSA in pregnancy.

1.8 PfEMP1 structure and classification

PfEMP1 are large (200-450 kDa) modular proteins made up of an N-terminal segment (NTS), a variable number of Duffy binding-like (DBL) and cysteine-rich interdomain region (CIDR) domains, a TM domain and ATS region. DBL and CIDR domains mediate the binding capabilities of PfEMP1 and are classified into types DBLα, β, γ, δ, ε, ζ and x and CIDRα, β, γ and δ based on small regions of semi-conserved sequence (Smith *et al.*, 2000b, Lavstsen *et al.*, 2003, Rask *et al.*, 2010) and further into subtypes, e.g. DBLα1, also based on sequence similarity (Rask *et al.*, 2010). DBL domains belong to a wider family of *Plasmodium* adhesive domains which are mainly involved in erythrocyte binding and invasion (Adams *et al.*, 1992, Chitnis and Miller, 1994). In addition to classification by type, DBL domains can be separated into three structural subdomains, S1-S3, which are flanked by regions of sequence homology (homology blocks, HBs) (Smith *et al.*, 2000b, Rask *et al.*, 2010).

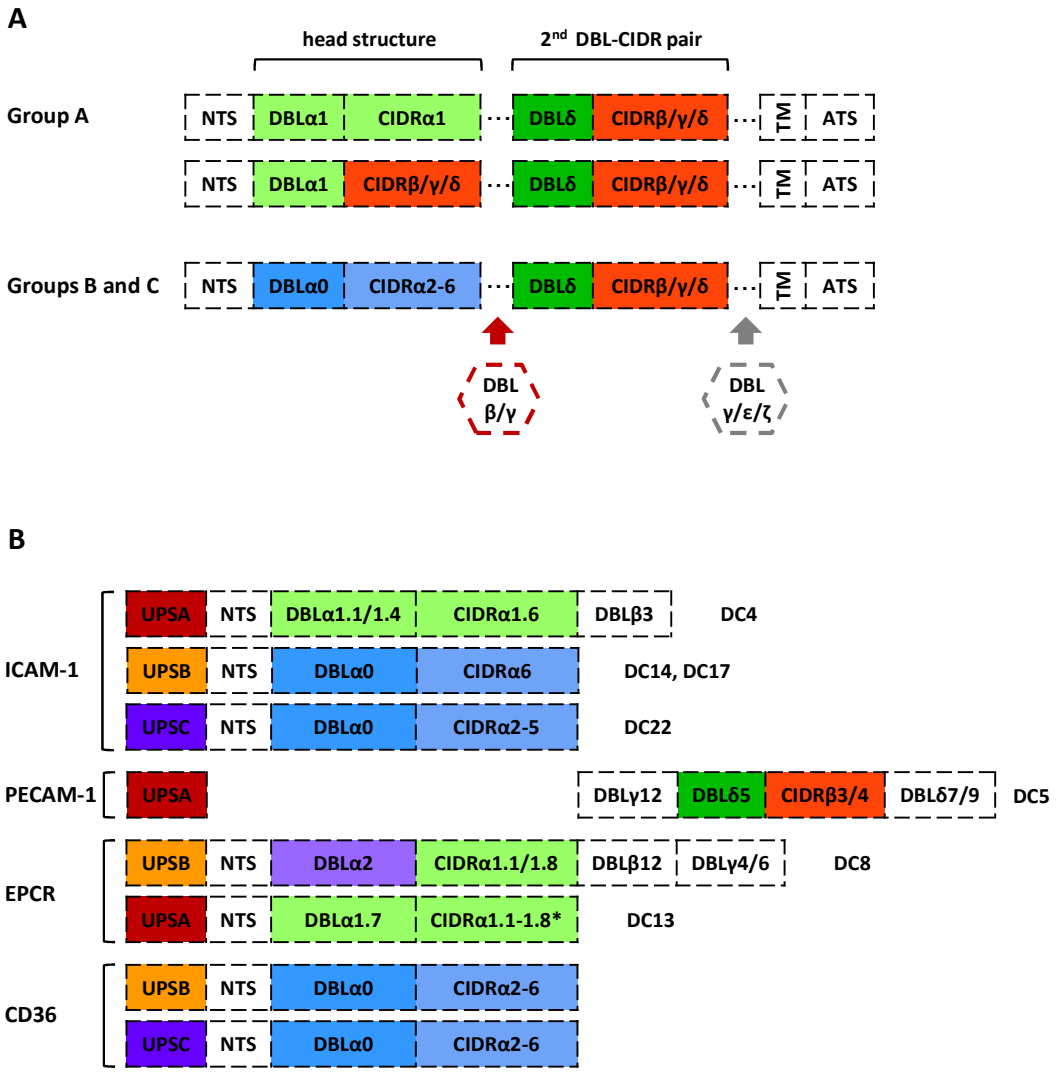


Figure 1.4. PfEMP1 domain architecture.

A, PfEMP1 proteins consist of a variable number of DBL and CIDR domains. Group B and C protein predominantly have a four domain structure made up of the N-terminal head structure (DBLα-CIDR) and a second DBL-CIDR pair. Group A PfEMP1 proteins (and some larger group B and C proteins) have addition DBL domains downstream of the head structure and/or the second DBL-CIDR pair. **B**, DBL and CIDR domains occur in tandem to form domain cassettes (DCs) which are linked to adhesion phenotypes, as indicated. DC8 is a chimeric B/A gene consisting of UPS type B with a structure typical of group A PfEMP1 proteins. * not CIDRα1.2 or 1.3. Figure modified from (Hviid and Jensen, 2015).

The domain composition varies between PfEMP1 proteins but the majority contain an N-terminal head structure, also known as component 1, consisting of DBLα- CIDR domains (Rask *et al.*, 2010) (summarised in Figure 1.4A). Most group B, B/C and C

proteins contain another DBL-CIDR pair downstream of the head structure, whereas most group A and B/A proteins contain this second DBL-CIDR pair plus additional DBL domains (Rask *et al.*, 2010). As a result, group A and B/A proteins are much longer than group B, B/C and C proteins. In a study of seven distinct genomes, combinations of DBL and CIDR domain types that frequently occur together were identified and are known as domain cassettes (DCs) (Rask *et al.*, 2010) (Figure 1.4B). There are two unique *var* genes that are semi-conserved across isolates, type 3 and VAR2CSA. The role of type 3 is unknown whereas VAR2CSA is important in PAM and the *var2csa* gene is thought to play a role in regulation of *var* gene expression (Salanti *et al.*, 2003, Ukaegbu *et al.*, 2015).

Crystal structures of several DBL (Higgins, 2008a, Khunrae *et al.*, 2009, Vigan-Womas *et al.*, 2012) and CIDR (Klein *et al.*, 2008, Lau *et al.*, 2015) domains have been solved and reveal them to consist mainly of α-helices joined by variable loops. This structure is conducive to sequence variation both in the loops, which lack structure, and in the variety of residues which can result in α-helices (Higgins and Carrington, 2014). Crystal structures, along with small angle x-ray scattering (SAXS) data, have identified two different overall quaternary PfEMP1 structures. VAR2CSA has a compact structure where the domains are folded together (Clausen *et al.*, 2012), whereas IT4var13 has a rigid, elongated structure (Brown *et al.*, 2013). The majority of PfEMP1 proteins are predicted to have a structure similar to IT4var13 based on the unique nature of VAR2CSA as a conserved protein responsible for a very particular case of receptor binding to placental syncytiotrophoblasts. However, the structures of additional full length PfEMP1 proteins will be required to confirm this theory.

1.9 PfEMP1 receptor binding and host disease

PfEMP1 proteins bind to a number of host receptors, with domain subtypes reported to determine specificity. However, determining the role of each PfEMP1-receptor interaction in development of severe disease is challenging due to the variety of SM phenotypes (see section 1.2) and the inconsistency of clinical definitions and control groups between studies. There are also difficulties in

relating a single interaction to a role in such a complex multi-faceted disease. However, a number of host receptors for PfEMP1 have been identified and their associations with malarial disease are discussed here. Specific domain subtypes that mediate binding are discussed extensively in Chapter 4 and so are not explicitly mentioned here.

1.9.1 CD36

CD36 is a glycoprotein expressed on endothelium cells, monocytes, macrophages, dendritic cells and platelets (Silverstein and Febbraio, 2009) and it was one of the first receptors identified as a mediator of IE binding (Ockenhouse *et al.*, 1989). CD36 binding is common in field isolates (Newbold *et al.*, 1997, Rogerson *et al.*, 1999) and is mediated by subtypes of CIDRα domains (Baruch *et al.*, 1997, Robinson *et al.*, 2003). However, the role of CD36 binding in disease is unclear with studies of patient isolates variously showing associations with UM (Udomsangpetch *et al.*, 1996, Ochola *et al.*, 2011), CM (Almelli *et al.*, 2014) or no associations (Newbold *et al.*, 1997, Marsh *et al.*, 1988, Rogerson *et al.*, 1999, Heddini *et al.*, 2001b). However, the majority of group B and C PfEMP1 contain CD36 binding domains and expression of these groups has been associated with UM (Kirchgatter and del Portillo, 2002, Kyriacou *et al.*, 2006, Rottmann *et al.*, 2006, Warimwe *et al.*, 2009). In addition, deficiency in CD36 does not protect against SM (Fry *et al.*, 2009). Overall, evidence supports an association of CD36 with UM (Cabrera *et al.*, 2014). However, some PfEMP1 proteins can bind to multiple receptors and a role for CD36 in co-operative binding has been observed with ICAM-1 (McCormick *et al.*, 1997, Yipp *et al.*, 2000, Gray *et al.*, 2003), an observation which alludes to the complexity of PfEMP1-host interactions.

The role of CD36 expressed on immune cells and platelets in IE adhesion is not clear. IE adherence to dendritic cells has been shown to inhibit their maturation with consequent reduced ability to stimulate T cells (Urban *et al.*, 1999). However, dendritic cell dysfunction was shown to be related to a high IE dose and to occur independently of PfEMP1 mediated adhesion to CD36 in another study (Elliott *et al.*, 2007b). Therefore, the relevance of dendritic cell CD36 binding in infection

remains disputed. However, functional dendritic cells have been shown to be essential for parasite clearance in animal models (Wykes *et al.*, 2007), highlighting their importance. A large scale association study found that monocyte CD36 expression is significantly higher in UM versus SM, suggesting a protective effect, whereas high platelet CD36 expression is associated with CM and lactic acidosis (Cserti-Gazdewich *et al.*, 2012). The mechanism by which platelet CD36 leads to disease is unknown but may involve IE adherence.

1.9.2 ICAM-1

Intercellular adhesion molecule-1 (ICAM-1), an immunoglobulin (Ig)-like molecule expressed on endothelium and leukocytes involved in transmigration of leukocytes to sites of inflammation (Lawson and Wolf, 2009), is another host receptor of PfEMP1 (Berendt *et al.*, 1989). Binding to ICAM-1 is mediated by DBLβ domains (Smith *et al.*, 2000a, Chattopadhyay *et al.*, 2004, Springer *et al.*, 2004) which bind at sites that are distinct but overlap with native binding of lymphocyte function associated molecule 1 (LFA-1) (Berendt *et al.*, 1992, Ockenhouse *et al.*, 1992a), suggesting that IE binding may interfere with native function. The link between ICAM-1 binding and disease severity has been studied in cultured field isolates with varying results. Some studies link ICAM-1 binding with severe malaria (Turner *et al.*, 2013) and cerebral malaria specifically (Newbold *et al.*, 1997, Ochola *et al.*, 2011), whereas others found no association with disease (Rogerson *et al.*, 1999, Heddini *et al.*, 2001b, Almelli *et al.*, 2014). However, ICAM-1 has been associated with sequestered parasites at autopsy (Turner *et al.*, 1994) and ICAM-1 expression levels are up-regulated in malaria infection (Turner *et al.*, 1998, Moxon *et al.*, 2013) as a consequence of endothelial activation. A recent study demonstrated that parasites from children with intestinal disease bound more strongly to ICAM-1 than those without, although sample sizes were small (Church *et al.*, 2016), implicating ICAM-1 binding in the development of organ-specific disease. Further studies will be required to investigate this link. It has also been shown that ICAM-1 binding can act in synergy with CD36 to contribute to sequestration (Yipp *et al.*, 2000, McCormick *et al.*, 1997, Gray *et al.*, 2003), suggesting that multi-receptor binding of a single PfEMP1 may contribute to disease progression.

The majority of ICAM-1 binding DBLβ domains occur in group B and C PfEMP1 which also contain CD36 binding domains (Howell *et al.*, 2008, Janes *et al.*, 2011, Rask *et al.*, 2010), a trait generally associated with UM (see section 1.9.1 above). One exception is a group A gene containing DC4 which binds ICAM-1 but not CD36 (Bengtsson *et al.*, 2013b). DC4 was identified when a gene from the 3D7 *P. falciparum* clone (PFD1235w) was specifically recognised by IgG from semi-immune African children and orthologues were found in other parasite isolates (Jensen *et al.*, 2004, Bengtsson *et al.*, 2013b). Antibodies against PFD123w are acquired in early life and are associated with protection against clinical disease (Jensen *et al.*, 2004, Lusingu *et al.*, 2006).The association of ICAM-1 binding with both UM and SM suggests that a single domain-receptor interaction is not informative of disease outcome but rather that the combination of domain binding characteristics of each PfEMP1 is important. Indeed, it is now thought that DCs are more informative of the clinical significance of PfEMP1 variants after DC8 and DC13 were found to be preferentially expressed by parasites selected for adhesion to brain endothelial cells (Avril *et al.*, 2012, Claessens *et al.*, 2012). However, very recent work on these DCs suggests that the situation is complex (see section 1.9.3 below). Interestingly, these particular DCs do not bind ICAM-1, suggesting the role of an additional receptor in brain endothelial cell binding.

1.9.3 EPCR

As mentioned above, panning of IEs on brain endothelial cells selected parasites expressing DC8 and DC13 containing PfEMP1 (Avril *et al.*, 2012, Claessens *et al.*, 2012). The transcription of these DCs in patient isolates is associated with CM (Lavstsen *et al.*, 2012, Bertin *et al.*, 2013, Almelli *et al.*, 2014) and SMA (Lavstsen *et al.*, 2012), suggesting a link to pathogenesis. The receptor for DC8 and DC13 was identified, by screening the full length IT4var20 protein (DC8) against a library of 2,500 human receptors, as the endothelial protein C receptor (EPCR) (Turner *et al.*, 2013). EPCR is an endothelial cell receptor which binds protein C and, in conjunction with thrombomodulin, produces activated protein C (APC) which mediates anticoagulant systems (Aird *et al.*, 2014). Additionally, EPCR is required for APC activation of PAR-1 which promotes cytoprotection, including activation of

anti-inflammatory responses and protection of the endothelial barrier (Aird *et al.*, 2014). Therefore, disruption of the native role of EPCR by IE binding suggests an important role in pathogenesis by promoting thrombi and by breakdown of the endothelial barrier. EPCR and thrombomodulin are expressed at low levels in brain endothelium (Laszik *et al.*, 1997) making them vulnerable to disruption. Interestingly, the loss of these receptors is observed in CM in association with adherent IEs (Moxon *et al.*, 2013). The mechanism by which EPCR is lost in CM is unknown and may well be separate from EPCR disruption by IE binding.

The presence of DC8 and DC13 in all but one of seven parasite genomes studied (Rask *et al.*, 2010) led to great optimism that EPCR binding is a common parasite feature that can be disrupted as a therapeutic method to prevent SM. The CIDRα domain mediates ECPR binding (Turner *et al.*, 2013) and the majority of non-CD36 binding CIDRα subtypes have been shown to bind EPCR with nanomolar affinity in recombinant protein assays (Lau *et al.*, 2015). Co-crystals of CIDRα bound to EPCR show an overlapping binding site with APC (Lau *et al.*, 2015), supporting IE blockage of native EPCR function. However, two recent papers challenge the relative importance of EPCR binding between different PfEMP1 proteins (Sampath *et al.*, 2015, Gillrie *et al.*, 2015). Several CIDRα domains were shown to bind EPCR with variable affinity but all within the nanomolar range (Sampath *et al.*, 2015), in keeping with previous results (Lau *et al.*, 2015). However, the ability of CIDRα domains to block APC binding varied, as did the effect of anti-EPCR antibodies, suggesting overlapping but distinct binding sites between CIDRα domains (Sampath *et al.*, 2015). These results are reminiscent of DBLβ::ICAM-1 interaction studies which show variable binding affinity and overlapping but distinct binding sites of different DBLβ domains (Brown *et al.*, 2013, Tse *et al.*, 2004, Madkhali *et al.*, 2014). Gillrie *et al.* (2015) assessed binding of parasite lines expressing a DC8 PfEMP1 (IT4var19) and a DC13 PfEMP1 (IT4var07) to human microvascular endothelial cells (HMEC) under flow conditions. The authors found that binding to transformed brain HMECs and lung HMECs was partially dependant on EPCR, whereas binding to primary brain HMECs and dermal HMECs was not dependant on EPCR, suggesting binding is mediated by different receptor(s) (Gillrie *et al.*, 2015). There is still much

to learn about EPCR, the most recently identified PfEMP1 receptor, and the contribution of this interaction to pathogenesis.

1.9.4 PECAM-1

Parasites, both laboratory-adapted (Treutiger *et al.*, 1997) and clinical isolates (Heddini *et al.*, 2001a), are known to bind platelet endothelial cell adhesion molecule 1 (PECAM-1). However, there appears to be no correlation to disease severity (Newbold *et al.*, 1997, Heddini *et al.*, 2001b). PfEMP1 is the parasite ligand responsible for binding and was originally mapped to CIDRα and DBLδ domains (Chen *et al.*, 2000). However, a recent study has shown PECAM-1 binding by group A PfEMP1 containing DC5, which consists of C-terminal domains (DBLγ12-DBLδ5-CIDRβ3/4-DBLβ7/9) (Berger *et al.*, 2013). The exact binding domain was not mapped conclusively (Berger *et al.*, 2013) and, therefore, the involvement of the N-terminal head structure has not been ruled out. A 3D7 parasite PfEMP1 containing DC5 was previously associated with SM (Jensen *et al.*, 2004) and antibodies against this PfEMP1 are developed early in infection (Cham *et al.*, 2009, Cham *et al.*, 2010) and are associated with protection against malarial disease (Magistrado *et al.*, 2007). It is therefore possible that PECAM-1 binding is mediated by distinct PfEMP1 that have differential roles in disease progression, similar to findings with ICAM-1 (see section 1.9.2).

1.9.5 Rosetting

Rosetting is the binding of multiple uninfected erythrocytes to a central IE, a process associated with SM (Carlson *et al.*, 1990, Rowe *et al.*, 1995, Kun *et al.*, 1998). Rosettes mediate pathogenesis by blocking microvessels and obstructing blood flow (Kaul *et al.*, 1991). Rosetting parasites have been shown to retain endothelial binding ability (Kaul *et al.*, 1991, Adams *et al.*, 2014) which could suggest a duel mechanism of disease progression. Several interactions are required to form rosettes and are mediated by the PfEMP1 head structure binding to host complement receptor 1 (CR1) (Rowe *et al.*, 1997, Rowe *et al.*, 2000), blood group A and B antigens (Vigan-Womas *et al.*, 2012, Barragan *et al.*, 2000) and heparin sulphate (Carlson and Wahlgren, 1992, Vogt *et al.*, 2003). In addition, certain

rosetting PfEMP1 proteins bind IgM via the C-terminal domains (Pleass *et al.*, 2016) which is thought to strengthen interactions between infected and uninfected erythrocytes (Clough *et al.*, 1998, Somner *et al.*, 2000). Recent studies have shown RIFIN and STEVOR proteins to mediate rosetting via blood group A antigen and glycophorin C, respectively (Goel *et al.*, 2015, Niang *et al.*, 2009) (discussed in section 1.5 above). Whether these VSAs act in tandem with, or independently of, PfEMP1 remains to be seen and will likely depend on the specific VSAs expressed on the IE surface at any one time.

1.9.6 Platelet mediated clumping

Platelet mediated clumping is similar to rosetting but involves binding of IEs to other IEs through platelets (Pain *et al.*, 2001). This phenotype is associated with SM (Pain *et al.*, 2001, Chotivanich *et al.*, 2004, Wassmer *et al.*, 2008) and has been proposed to cause disease by forming a bridge between IE and endothelium, allowing binding even in the absence of PfEMP1-specific receptors (Wassmer *et al.*, 2004), and through activation of platelets resulting in increased immune response (Srivastava *et al.*, 2008). The parasite factor that mediates clumping is unknown but PfEMP1 has been suggested. The platelet receptors involved are CD36 (Pain *et al.*, 2001), globular C1q receptor (Biswas *et al.*, 2007) and P-selectin (Wassmer *et al.*, 2008).

1.9.7 CSA and malaria in pregnancy

Pregnancy associated malaria (PAM) is associated with IE binding to CSA in the placenta (Fried and Duffy, 1996) and results in vessel blockage and pro-inflammatory response which leads to reduced placental efficiency, low birth weight and increased risk of fetal loss (Fried *et al.*, 1998, Rogerson *et al.*, 2007, Umbers *et al.*, 2011). The semi-conserved VAR2CSA PfEMP1 is responsible for specific binding to CSA expressed on placental cells (Salanti *et al.*, 2003). The second DBL domain has been shown to bind CSA with affinity comparable to the full length VAR2CSA (Clausen *et al.*, 2012, Srivastava *et al.*, 2011). However, previous studies suggested that the full length protein conferred much higher binding affinity (Khunrae *et al.*, 2010, Srivastava *et al.*, 2010) and a recent study conducted in a

physiologically relevant context corroborates this (Rieger *et al.*, 2015). PAM mainly affects women in their first pregnancy and antibodies against VAR2CSA increase with the number of pregnancies and are protective against PAM (Oleinikov *et al.*, 2007, Tuikue Ndam *et al.*, 2005, Tuikue Ndam *et al.*, 2006, Tutterrow *et al.*, 2012a, Tutterrow *et al.*, 2012b). Interestingly, VAR2CSA is rarely recognised (or is insignificant) by antibodies from men and children (Beeson *et al.*, 1999, Ricke *et al.*, 2000, Staalsoe *et al.*, 2001), suggesting VAR2CSA-expressing IEs are rapidly cleared in these patients (Oleinikov *et al.*, 2007). This very specific case of PfEMP1-ligand binding has led to optimism of an anti-disease vaccine, one of which (PlacMalVac) has completed Phase I clinical trials and is in preparation for Phase II clinical trials (PlacMalVac).

1.9.8 Other possible host receptors

A number of other host receptors have been shown to mediate IE binding with PfEMP1 suggested as the ligand but evidence is scarce. These include P-selectin, thrombospondin, E-selectin, VCAM-1, fractalkine, integrin αvβ3, fibronectin, NCAM and gC1qR-HABP1-p32 [reviewed by Rowe *et al.* (2009)].

1.10 Acquired immunity against VSAs

Immunity to symptomatic malaria is acquired gradually over the course of repeated *P. falciparum* infections (Marsh and Kinyanjui, 2006). In areas of endemic transmission, young children are most at risk of developing SM with one study reporting acquisition of immunity after just a few infections and the development of immune protection from non-cerebral SM by the age of 5 years (Gupta *et al.*, 1999). However, a recent study suggested a more gradual acquisition of protection against SM, based on models of hospital admission data (Griffin *et al.*, 2015). Immunity against UM symptoms is not acquired until adolescence or adulthood and sterilising immunity is thought to never, or very rarely, be achieved (McGregor, 1964), resulting in widespread asymptomatic infections. The acquired immunity achieved is very complex and consists of responses to a variety of antigens expressed at all parasite stages and is effected by both humoral and cell-mediated

responses (Beeson *et al.*, 2008). This poses a formidable challenge in the development of vaccines against malaria.

Antibodies are generated against surface exposed antigens, both VSAs on the IE surface and merozoite antigens. Important merozoite proteins include merozoite surface proteins (MSP) 1-3, erythrocyte binding antigens (EBAs), reticulocyte-binding protein homologues (RH) and apical membrane antigen 1 (AMA-1), all of which have been extensively studied and are considered vaccine candidates (Richards and Beeson, 2009). The antibodies acquired against VSAs display a high level of strain specificity due to the variable nature of these surface molecules (Newbold *et al.*, 1992, Biggs *et al.*, 1991, Marsh and Howard, 1986). Serum from adults in convalescence is able to recognise a diverse range of *P. falciparum* isolates (Chattopadhyay *et al.*, 2003), in keeping with exposure dependant development of immunity. However, acute infection of naive adults was shown to elicit cross-reactive antibodies in returned travellers (Elliott *et al.*, 2007a), suggesting the presence of shared epitopes between isolates. Anti-VSA antibodies are associated with protection against malarial disease in the majority (Marsh *et al.*, 1989, Tebo *et al.*, 2002, Dodoo *et al.*, 2001) but not all (Bull *et al.*, 1998) studies.

The relative contribution of anti-PfEMP1 antibodies to protection against disease is difficult to study due to the presence of other VSAs such as RIFINs and STEVORs. However, parasites genetically modified to inhibit surface PfEMP1 expression show markedly reduced recognition by immune serum compared to wild type parasites , suggesting only a minor role of other VSAs (Chan *et al.*, 2012). In addition, PfEMP1 alone was associated with protective immunity suggesting this protein family is the main target of anti-VSA immunity (Chan *et al.*, 2012). However, further study is required to establish the role of RIFIN, STEVOR and other VSAs in developing immunity. Anti-PfEMP1 response has been demonstrated to be variant-specific in children under 3 years but broadly cross-reactive in adults as measured by recognition of diverse DBLα domains, following the pattern observed for overall immune response to IEs (Barry *et al.*, 2011). The acquisition of antibodies against PfEMP1 appears to be structured in relation to the different groups A, B and C. Antibodies against PfEMP1 encoded by group A *var* genes are acquired first (Cham

et al., 2009, Cham *et al.*, 2010), supporting the role of particular PfEMP1 in development of SM (Rottmann *et al.*, 2006, Jensen *et al.*, 2004). The mechanisms by which anti-PfEMP1 antibodies mediate protection against SM include blocking adhesion, preventing rosette formation and opsonising IEs for phagocytosis, although specific mechanisms are not yet fully understood (Chan *et al.*, 2014).

1.11 PhD aims

- Identify the dominantly expressed *var* genes of several recently characterised ICAM-1-selected culture-adapted patient isolates.

- Perform targeted sequencing of these *var* genes by testing the predictive power of a new database of ~ 62,000 *var* genes.

- Obtain the full length sequence of two *var* genes identified specifically in the brains of children who died of cerebral malaria.

- Perform extensive sequence analysis of the *var* genes identified to elucidate their predicted binding phenotypes.

- Recombinantly express the DBLβ domains of the ICAM-1-selected culture-adapted patient isolates and characterise their ability to bind wild type ICAM-1 and four mutant ICAM-1 proteins.

- Relate the recombinant protein findings to the known IE binding profiles of these isolates and assess the implications for anti-adhesion therapies.

Chapter 2. General Methods

2.1 Parasite techniques

2.1.1 Parasite culture

Parasites were cultured according to standard techniques (Trager and Jensen, 1976) using RPMI-1640 medium (Sigma) supplemented with 37.5 mM HEPES, 7 mM D-glucose, 25 ng/ml gentamicin sulphate, 2 mM L-glutamine and 10% human serum (Haematology Department, Royal Liverpool University Hospital) at pH 7.2, hereafter, referred to as complete medium or serum-free medium when 10% human serum is omitted. O+ human erythrocytes were obtained from the National Blood and Transfusion Service and added at 2% haematocrit. Cultures were kept at a gas mixture of 96% nitrogen, 3% carbon dioxide and 1% oxygen (supplied by BOC). Parasitaemia was calculated by examination of 500 red blood cells on Giemsa stained (Giemsa's azur eosin methylene blue solution, Merck Millipore) thin blood films. Blood films were made by smearing a drop of parasite culture across a glass slide, which was left to air-dry. The smear was then fixed with absolute methanol for 10 seconds before staining with Giemsa (Giemsa's azur eosin methylene blue solution, Merck Millipore) diluted 1 in 10 with 10% phosphate-buffered water (20 mM Na_2HPO_4, 4 mM KH_2PO_4, pH 7.2) for at least 5 minutes. Slides were then washed with tap water and air-dried before for examination.

2.1.2 Reconstitution of frozen parasites

Parasites were removed from liquid nitrogen and placed at 37 °C to thaw quickly before being transferred to a 50 ml falcon tube. One fifth volume of pre-warmed 12% NaCl solution was added dropwise with continual mixing and allowed to stand for 5 minutes at room temperature. 5 ml of 1.8% NaCl solution was added dropwise over 5 minutes then allowed to stand for 5 minutes at room temperature. 5 ml of 0.9% NaCl solution was added dropwise over 5 minutes then allowed to stand for 5 minutes at room temperature. The solution was centrifuged at 600 x g for 5 minutes to pellet the cells. The supernatant was removed and the cells

resuspended in 25 ml pre-warmed serum-free medium (see 2.1.1). Cells were pelleted by centrifugation at 600 x g for 5 minutes. The supernatant was removed and the cells were resuspended in complete medium, transferred to a sterile tissue culture flask and cultured by standard techniques (see 2.1.1).

2.1.3 Cryopreservation of parasites

Parasite cultures at 3-5% ring stage parasitaemia were centrifuged at 600 x g for 5 minutes and the supernatant removed. Glycerolyte freezing solution (6.2 M glycerol, 140 mM sodium lactate, 130 mM sodium dihydrogen orthophosphate, 1.6 mM potassium chloride, pH 6.8) was added in a ratio of 5 volumes to 3 volumes of cell pellet with the first volume added dropwise over 5 minutes and left to stand for 5 minutes. The remaining 4 volumes were then added slowly with constant mixing. The resulting solution was transferred to sterile freezing vials and placed at - 80 °C for 24 hours before being transferred to liquid nitrogen.

2.1.4 Trophozoite enrichment by Plasmagel flotation

Plasmagel solution (Plasmion, Fresenius Kabi France) and serum-free medium was pre-warmed to 37 °C. Parasite cultures were centrifuged at 600 x g for 5 minutes and the supernatant removed. The cell pellet was resuspended such that the pellet made up 20%, serum-free medium 30% and Plasmagel 50% of the total volume. The solution was transferred to a 15 ml falcon tube and incubated for 20 minutes at 37 °C. The supernatant was removed and transferred to a new 15 ml falcon tube, centrifuged at 600 x g for 5 minutes and the pellet washed in 10 ml serum free medium. After centrifugation, the cell pellet was resuspended in either complete medium to continue the culture (see 2.1.1) or 1% Bovine Serum Albumin (BSA) in phosphate buffered saline (PBS; hereafter referred to as 1%BSA/PBS) for ICAM-1 selection (see 3.2.2). A small sample was taken for a Giemsa-stained thin blood film to confirm trophozoite enrichment.

2.2 Nucleic acid techniques

2.2.1 RNA extraction

Parasite RNA extraction was performed by a method adapted from that originally described by Kyes *et al.* (2000). Parasite cultures were pelleted by centrifugation at 600 x g for 5 minutes. The supernatant was aspirated and the pellet resuspended by tapping. Pre-warmed TRIzol® Reagent (Ambion) was added at 10 X pellet volume and mixed by pipetting until all clumps disappeared. Following incubation at 37 °C for at least 5 minutes, the TRIzol® lysed parasites were split into 1 ml aliquots and either frozen at -80 °C or RNA extraction continued immediately. To each 1 ml aliquot, 0.2 ml chloroform was added and the tube shaken vigorously for 15 seconds, left to stand for 2-3 minutes and centrifuged at 1,200 x g for 30 minutes at 4 °C. The upper aqueous layer was removed to a new 1.5 ml tube. 0.5 ml isopropanol was added and mixed by inverting the tube several times before incubation at 4 °C for at least 2 hours. The tube was mixed again by inverting several times and centrifuged at 16,200 x g for 30 minutes at 4 °C. The supernatant was carefully removed and the pellet washed in 0.5 ml 75% ethanol in DEPC-H_2O (diethyl pyrocarbonate treated H_2O, Sigma). After centrifugation at 16,200 x g for 5 minutes, the supernatant was removed and the pellet left to air dry at room temperature. 50 µl DEPC-H_2O was added to the pellet before incubation at 65 °C for 10 minutes and then placed on ice. The RNA was mixed by pipetting and the concentration and purity measured by NanoDrop™ 1000 Spectrophotometer (Thermo Scientific). RNA was stored at - 80 °C.

2.2.2 cDNA synthesis

An RNase free environment was maintained throughout the following reactions. 1.2 µg RNA was treated with DNase I (Amplification Grade, Sigma-Aldrich) according to manufacturer's instructions. The reaction comprised 2 µl 10 X Reaction Buffer, 2 µl DNase I and 1.2 µg RNA made up to a final volume of 20 µl with DEPC-H_2O. The reaction was incubated for 15 minutes at room temperature. 2 µl Stop Solution

(Sigma-Aldrich) was added and the reaction incubated at 70 °C for 10 minutes and then chilled on ice.

cDNA synthesis was performed with Tetro cDNA Synthesis Kit (Bioline) following the manufacturer's instructions. The reaction comprised 8 µl DNase treated RNA (~480 ng, above), 1 µl Oligo $(dT)_{18}$ primer, dNTP at 0.5 mM final concentration, 4 µl 5X RT Buffer, 10 U RiboSafe RNase Inhibitor, 200 U Tetro Reverse Transcriptase and DEPC-H_2O up to 20 µl. The reaction was incubated at 45 °C for 30 minutes, followed by 85 °C for 5 minutes to stop the reaction and chilled on ice. A control without reverse transcriptase (no-RT) was carried out to assess DNase I efficiency. cDNA was stored at -20 °C.

2.2.3 Primer design and polymerase chain reaction (PCR)

All novel primer sequences were designed with the aid of OligoCalc (Kibbe, 2007). Primer sequences from previously published work are reproduced with the appropriate citation. All primers are listed in Appendix A.

All PCR reactions were carried out with the proofreading *TaKaRa LA Taq*® DNA polymerase (Clontech, TaKaRa Bio Inc) according to the manufacturer's instructions. Each reaction comprised 10 X LA PCR Buffer II, magnesium chloride ($MgCl_2$) at 2.5 mM final concentration, deoxyribonucleotide triphosphates (dNTP) at 1 mM final concentration, primers at 0.3 mM final concentration each, 2.5 units of *TaKaRa LA Taq*®, 2 µl template and sterile water up to 50 µl final volume unless otherwise stated. Specific PCR cycling conditions are stated for each primer set in the appropriate chapter and annealing temperatures are listed in Appendix A.

2.2.4 PCR product purification

PCR products were visualised by gel electrophoresis on a 2% agarose gel in 1X TBE (89 mM Tris-borate, 2 mM EDTA, pH 8.3, Sigma) and stained with ethidium bromide at 3 µg/ml final concentration. Products were excised from the gel and purified using QIAquick Gel Extraction Kit (Qiagen) according to manufacturer's instructions. The PCR product was eluted into a sterile 1.5 ml tube in 50 µl sterile H_2O (Sigma).

The PCR product concentration was measured by NanoDrop™ 1000 Spectrophotometer (Thermo Scientific).

2.2.5 Plasmid sequencing

All plasmid sequencing was carried out by either the Core Genomic Facility, University of Sheffield or Source Bioscience Sequencing, Rochdale UK .

2.3 Protein techniques

2.3.1 SDS-PAGE

Proteins were separated by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) on pre-cast Precise™ Tris-HEPES 12% gels (ThermoFisher Scientific) in Pierce™ 20X Tris-HEPES-SDS Buffer diluted to 1X to contain 0.1 M Tris, 0.1 M HEPES, 3 mM SDS at pH 8. Samples were prepared by adding an equal volume of 2X sample buffer (65.8 mM Tris-HCl, pH 6.8, 26.3% glycerol, 2.1% SDS, 0.01% bromophenol blue, 5% 2-mercaptoethanol) and boiling for 10 minutes before being placed on ice and briefly spun down in a bench-top centrifuge. Amersham Low Molecular Weight Marker (GE Healthcare) was used as a protein standard on each gel. SDS-PAGE was carried out at 90-150 V for 60-90 minutes until the dye front was near the bottom of the gel.

2.3.2 Coomassie staining

Gels were stained with 0.16% Coomassie Brilliant Blue (ThermoFisher Scientific) in a solution of methanol, H_2O and acetic acid in the ratio 5:5:1 for ~10 minutes and washed with hot water several times before addition of destain solution (methanol, H_2O and acetic acid in the ratio 5:5:1). Gel images were taken using an InGenius Bioimager (Syngene).

2.3.3 Western blotting

N-terminal histidine tagged proteins were detected by Western blot. Samples were separated by SDS-PAGE (see 2.3.1) and transferred to an Amersham Protran 0.45 nitrocellulose membrane (GE Healthcare) in transfer buffer (25 mM Trizma-base, 190 mM glycine, 20% methanol) at 200 V for 1 hour. The membrane was stained with 0.1% Ponceau S (Sigma) in 5% acetic acid until bands became visible. The protein marker was marked with pen and the stain washed out with TNT buffer (10 mM Tris-HCl, 150 mM NaCl, 0.1% Tween-20). The membrane was blocked with 5% milk (Sigma) in TNT buffer for 1 hour, shaking at room temperature. Primary Mouse anti-his antibody (H1029, Sigma) was diluted 1:1000 in SignalBoost™ Immunoreaction Enhancer Kit (Merck Millipore) Solution 1 for primary antibodies and incubated on the membrane for 1 hour, shaking at room temperature. The membrane was washed 3 x 5 minutes with TNT buffer. The secondary Goat Anti-Mouse IgG-horseradish peroxidise (HRP) Conjugate antibody (Bio-Rad) was diluted 1:5000 in SignalBoost™ Immunoreaction Enhancer Kit (Merck Millipore) Solution 2 for secondary antibodies and incubated on the membrane overnight at 4 °C, shaking. The membrane was washed 3 x 5 minutes with TNT buffer and incubated for 5 minutes in SuperSignal™ West Pico Chemiluminescent Substrate (ThermoFisher Scientific) according to manufacturer's instructions. The membrane was then exposed to CL-XPosure™ Film (Thermo Scientific) and developed on a photon imaging systems machine (AFP Imaging Corp).

2.4 Software

2.4.1 Analysis of plasmid sequences

Plasmid sequences were initially visualised using Chromas Lite Version 2.1 (Technelysium Pty Ltd, latest version available at http://technelysium.com.au/). Sequence alignments were carried out using ClustalX 2.1 (Larkin *et al.*, 2007) unless otherwise stated.

2.4.2 Translation software

Nucleic acid sequences were translated into amino acid sequence using either ExPASy Translate tool (Swiss Institute of Bioinformatics (SIB), available at: http://web.expasy.org/translate/) or EMBOSS Transeq (European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI), available at: http://www.ebi.ac.uk/Tools/st/emboss_transeq/).

2.4.3 BLAST®

The National Centre for Biotechnology Information (NCBI) BLAST® resource (available at: http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to search for sequence similarities among the vast nucleotide database. The BLAST® align feature was also used to compare nucleotide sequences.

Chapter 3. Identifying dominantly expressed *var* genes of *Plasmodium falciparum*

3.1 Introduction

Sequestration of the malaria-causing parasite *Plasmodium falciparum* in the microvasculature of the human host is thought to be a key process in the pathogenesis of severe disease, including cerebral malaria (Newton *et al.*, 1998, Taylor *et al.*, 2004). Sequestration is mediated by the parasite protein family *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1), encoded by *var* genes (Baruch *et al.*, 1995, Smith *et al.*, 1995, Su *et al.*, 1995). The 50-60 *var* gene copies per genome are classified according to 5' upstream sequence (UPS), chromosomal location and transcriptional orientation into three main groups A, B and C (Voss *et al.*, 2000, Gardner *et al.*, 2002). They are subject to mutually exclusive expression regulated by epigenetic mechanisms (Scherf *et al.*, 1998). The large, highly variable, modular PfEMP1 proteins are made up of DBL (Duffy binding-like) and CIDR (cysteine-rich interdomain region) domains which are classified into types DBLα, β, γ, δ, ε, ζ and x and CIDRα, β, γ and δ based on sequence identity (Smith *et al.*, 2000b, Lavstsen *et al.*, 2003, Rask *et al.*, 2010). These domains mediate binding to a variety of host endothelial cell receptors (Rowe *et al.*, 2009, Turner *et al.*, 2013), including intercellular adhesion molecule 1 (ICAM-1) (Berendt *et al.*, 1989) which has been implicated in severe malaria.

The link between ICAM-1 binding and disease severity has been studied in cultured field isolates with varying results. Some studies link ICAM-1 binding with severe malaria (Turner *et al.*, 2013) and cerebral malaria specifically (Newbold *et al.*, 1997, Ochola *et al.*, 2011), whereas others found no association with disease (Rogerson *et al.*, 1999, Heddini *et al.*, 2001b). However, ICAM-1 has been associated with sequestered parasites at autopsy (Turner *et al.*, 1994) and ICAM-1 binding is likely to act in synergy with other receptors to contribute to sequestration and disease progression (Yipp *et al.*, 2000, McCormick *et al.*, 1997). The nature of such a

complex host-parasite interaction makes it difficult to pinpoint the exact role of ICAM-1 binding but it appears to play a significant role in disease pathogenesis.

The ICAM-1-PfEMP1 interaction has mainly been characterised in the reference strains of IT4 lineage and 3D7 (Howell *et al.*, 2008, Janes *et al.*, 2011, Oleinikov *et al.*, 2009, Bengtsson *et al.*, 2013b) and it is the DBLβ domain that mediates binding (Smith *et al.*, 2000a, Brown *et al.*, 2013). This focus on established reference strains is due to ease of culture, the knowledge of their *var* gene sequences and availability of qPCR primer sets targeting all *var* genes (Viebig *et al.*, 2007), a requirement for confirmation of a culture's dominantly expressed *var* gene. Sequencing the *var* genes of recently culture-adapted field isolates is a challenge. Primer sets designed to homology blocks within PfEMP1 domains are based on our knowledge from several *P. falciparum* genomes (Bull *et al.*, 2005, Rask *et al.*, 2010) and may not target all genes in the global population. Advances in whole genome sequencing (WGS) present a viable option but there is still difficulty in accurately assembling *var* genes due to low read coverage (Gardner *et al.*, 2002, Manske *et al.*, 2012). The Pf3k genome sequencing project (https://www.malariagen.net/projects/parasite/pf3k) aims to sequence 3,000 *P. falciparum* genomes worldwide and has generated a vast amount of data to date (Manske *et al.*, 2012, Miotto *et al.*, 2013), including a new database of ~62,000 *var* genes.

In this chapter, we investigate the practical applications of the Pf3k *var* database by assessing whether a full length *var* gene can be predicted from a short sequence tag. The DBLα tag in question is a 350-400 bp section of the N-terminal DBLα domain amplified with primers designed by Bull *et al.* (2005) to short regions of homology. Our group has recently characterised the binding profile of several culture-adapted patient isolates that have been selected for ICAM-1 binding (Madkhali *et al.*, 2014). This characterisation is extended to the molecular level by identifying the *var* genes responsible for ICAM-1 binding in these isolates and utilizing the Pf3k *var* database to obtain their full length sequences. We also test this approach on clinical samples from Malawi from which two *var* genes were

found to be expressed at a high level in the brains of multiple children who suffered fatal cerebral malaria (Tembo *et al.*, 2014).

3.2 Methods

3.2.1 Parasite isolates

Laboratory-adapted patient isolates PO69 and 8206, from Kenya, and BC12, J1 and PCM7, from UM patients in Thailand (Poyomtip *et al.*, 2012), were collected with consent in previous clinical studies. All patient-derived material, such as white blood cells, have been removed during culture. All five patient isolates were used for the *var* gene analysis described below. The laboratory line ItG derived from the IT 4/25/5 lineage (Ockenhouse *et al.*, 1992a) was used for DNA extraction and genotyping.

3.2.2 Parasite selection on ICAM-1

50 µl Protein A Dynabeads (Invitrogen) were washed three times with 500 µl 1% BSA/PBS, using a magnet to retain the beads each time. The beads were resuspended in 1% BSA/PBS and 2.5 µg ICAM-1^{D1D5} (an Fc fusion protein of ICAM-1 domains 1-5 (Tse *et al.*, 2004)) added to a final volume of 400 µl and incubated at room temperature with 15 rpm rotation using a stuart SB3 rotator (Bibby Scientific) for 1 hour. The beads were then washed three times with 500 µl 1% BSA/PBS using a magnet to retain the beads and resuspended in 200 µl 1% BSA/PBS. Parasite cultures were enriched for trophozoite stage using Plasmagel by standard techniques (see 2.1.4). Enriched parasites were resuspended in 200 µl 1% BSA/PBS and added to the ICAM-1 labelled Dynabead suspension. The mixture was rotated at 15 rpm using a stuart SB3 rotator (Bibby Scientific) for 45 minutes at room temperature. Two washes were carried out with 500 µl 1% BSA/PBS to remove unbound parasites. The beads were then transferred to a culture flask in complete medium with washed red blood cells (RBC) and cultured as standard (see 2.1.1).

3.2.3 DBLα tag polymerase chain reaction (PCR) of cDNA

RNA was extracted from ICAM-1 selected parasites at peak *var* gene expression, 6-12 hours post invasion (Kyes *et al.*, 2000), by a method modified slightly from Kyes *et al* (2000) and cDNA synthesised (see 2.2.1 and 2.2.2). The DBLα tag was amplified from parasite cDNA using the primers DBLαAF' and DBLαBR (Bull *et al.*, 2005) (Appendix A) with the proofreading *TaKaRa LA Taq*® DNA polymerase (Clontech, TaKaRa Bio Inc) as described (see 2.2.3). Reaction conditions were an initial denaturing step of 95 °C for 3 minutes, followed by 30 cycles of 95 °C for 30 seconds, 47 °C for 30 seconds, 65 °C for 30 seconds, and a final extension of 65 °C for 3 minutes. PCR products were purified as stated (see 2.2.4). A reaction was carried out with the no-RT control as template to show successful DNase treatment.

3.2.4 DBLα tag cloning, plasmid preparation and sequencing

DBLα tag PCR products were cloned into pCR^{TM}4-TOPO® vector (Invitrogen) at 10:1 insert to vector molar ratio. The reaction comprised 1 µl Salt Solution (1.2 M NaCl, 0.06 M $MgCl_2$), 1 µl pCR^{TM}4-TOPO® (10 ng) with ~ 10 ng insert in a final reaction volume of 6 µl. The reaction was incubated for 30 minutes at room temperature then placed on ice. The plasmid was transformed into One Shot® TOP10 *E. coli* competent cells (Invitrogen). 2 µl of ligation reaction were added to 50 µl competent cells, incubated on ice for 10 minutes, heat shocked at 42 °C for 30 seconds and kept on ice for 2 minutes before addition of 250 µl SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose). Cells were shaken for 1 hour at 37 °C before being spread on agar plates containing 100 µg/ml ampicillin and X-gal for selection of transformants. Two volumes were plated, 100 µl and 50 µl, and the plates were incubated at 37 °C overnight.

Positive colonies were picked and cultured overnight in LB broth (1% tryptone, 0.5% yeast extract, 85.6 mM NaCl; Sigma) with 100 µg/ml ampicillin. Plasmids were isolated using GenElute Plasmid Miniprep Kit (Sigma) or QIAprep Spin Miniprep Kit (QIAGEN) following the manufacturer's instructions. The culture was centrifuged at 1,877 x g for 10 minutes. The pellet was resuspended, lysed and precipitated using

the kit solutions. The resulting lysate was spun through a column, washed and eluted in 70 µl sterile water. Plasmids were restriction digested with *Eco*RI (New England Biolabs) to confirm the expected size of the insert. The reaction comprised of 2 µl NEBuffer I (10 mM Bis-Tris-Propane-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.0 at 25 °C), 20 U *Eco*RI, 5 µl plasmid and water up to 20 µl. The digest was incubated at 37 °C for at least one hour before being visualised by gel electrophoresis on a 2% agarose gel in TBE subsequently stained with ethidium bromide. Twenty to twenty-four plasmids per strain were sequenced as described (see 2.2.5). Sequences were visualised with Chromas Lite version 2.1 and alignments were carried out using ClustalX 2.1 on default settings (see 2.4.1).

3.2.5 Reverse transcriptase quantitative PCR (RT-qPCR)

RT-qPCR of the cDNA was carried out using Brilliant III Ultra-Fast SYBR® Green QPCR Master Mix (Agilent Technologies). Standard curves were generated for each new primer set (see Appendix A for primer sequences) by performing 10-fold serial dilutions of cDNA to produce five different concentrations of starting material. No-RT and DPEC-H_2O reactions were carried out as negative controls. All reactions were performed in triplicate. Each reaction comprised 10 µl 2X SYBR Green QPCR master mix, forward and reverse primers each at 0.5 µM final concentration, 2 µl template (cDNA, no-RT or DPEC-H_2O) and DPEC-H_2O to 20 µl final volume. Adenylosuccinate lyase (ASL) and seryl- tRNA synthetase (STS) were internal control genes (Viebig *et al.*, 2005). Reactions were carried out on an MxPro-Mx3005P machine under the following conditions: 95 °C for 3 minutes, followed by 40 cycles of 95 °C for 10 seconds, 60 °C for 10 seconds, and a final cycle of 95 °C for 1 minute, 55 °C for 30 seconds, 95 °C for 30 seconds.

Primer efficiency was calculated from the gradient of the standard curve using the equation:

$$E = 10^{1/-m}$$

where E is efficiency and m is the gradient of the curve.

RT-qPCR was repeated on cDNA at a single concentration with the same reagents and conditions (as above) using all primer pairs designed to particular isolates on the same plate. Ct values were normalised against the internal control gene using the equation:

$$2^{-\Delta Ct}$$

where ΔCt is the mean Ct value of the gene of interest minus the mean Ct value of the internal control gene. $2^{-\Delta Ct}$ values were transformed into percentage of the total for each parasite isolate.

3.2.6 Genomic DNA extraction

Genomic DNA (gDNA) was extracted from parasite cultures by saponin lysis of the RBC and the QIAamp® DNA Blood Mini Kit (Qiagen). Parasite cultures were centrifuged at 600 x g for 5 minutes and the resulting pellet was resuspended in 5 ml 0.15 % saponin and incubated for 5 minutes on ice, mixing each minute. Parasites were centrifuged at 600 x g for 5 minutes and washed in PBS 2-3 times. Following the final centrifugation, the parasite pellet was resuspended in 200 µl PBS and transferred to a 1.5 ml tube containing 20 µl Qiagen protease. The QIAamp® DNA Blood Mini Kit (Qiagen) spin protocol was followed with a final elution of the parasite DNA in 100 µl Buffer AE (10 mM Tris- HCl, 0.5 mM EDTA, pH 9).

3.2.7 Merozoite surface protein (MSP) and glutamate-rich protein (GLURP) genotyping

MSP primary PCR

Merozoite surface protein (MSP) genotyping was performed following Recommended Genotyping Procedures (Medicines for Malaria Venture, Amsterdam May 2007). An initial primary PCR reaction using both msp1 and msp2 primers comprised: 5 µl 10x NH_4 reaction buffer, $MgCl_2$ to 2 mM final concentration, dNTP to 0.2 mM final concentration, primers M1-OF, M2-OF, M1-OR, M2-OR (see Appendix A for primer sequences) to 0.3 µM final concentration, 2.5 U BIOTAQ (Bioline), 2 µl genomic DNA and sterile water up to 50 µl. Reaction conditions were an initial denaturing step of 94 °C for 2 minutes, followed by 30 cycles of 94 °C for

30 seconds, 54 °C for 1 minute, 72 °C for 1 minute, and a final extension of 72 °C for 5 minutes.

MSP2 secondary PCR

Secondary MSP2 PCR was carried out using 2 µl of the primary PCR as template. The reaction comprised the same concentrations as the primary PCR except with 1.5 mM final concentration of $MgCl_2$. Primer S1TAILFW was used in combination with N5-3D7-RVVIC (see Appendix A for primer sequences) each to 0.2 mM final concentration. The reaction was carried out under the following conditions: 94 °C for 2 minutes, followed by 30 cycles of 94 °C for 30 seconds, 50 °C for 45 seconds, 70 °C for 90 seconds, and a final extension of 70 °C for 5 minutes. All reactions were visualised by gel electrophoresis on a 3% agarose gel in TBE and stained with ethidium bromide. The reaction is referred to as "MSP2 N5 PCR" in Results 3.3.4.

Restriction digest of MSP2 PCR product

The MSP2 secondary PCR product underwent restriction digestion to further differentiate genotypes. The S1TAILFW/N5-3D7-RVVIC reaction product was digested with *Dde*I and *Scr*FI in separate reactions each comprising 7 µl PCR reaction product, 2 µl 10x NEB 2.1 buffer, 5 U of enzyme and sterile water up to 20 µl. Digestion products were visualised by gel electrophoresis on a 3% agarose gel in TBE and stained with ethidium bromide.

GLURP primary PCR

GLURP genotyping was performed following Recommended Genotyping Procedures (Medicines for Malaria Venture, Amsterdam May 2007). An initial primary PCR reaction comprised: 5 µl 10x NH_4 reaction buffer, $MgCl_2$ to 2 mM final concentration, dNTP to 0.2 mM final concentration, primers G-F3 and G-F4 (see Appendix A for primer sequences) to 0.3 µM final concentration, 2.5 U BIOTAQ (Bioline), 2 µl genomic DNA and sterile water up to 50 µl. Reaction conditions were an initial denaturing step of 94 °C for 2 minutes, followed by 30 cycles of 94 °C for 30 seconds, 54 °C for 1 minute, 72 °C for 1 minute, and a final extension of 72 °C for 5 minutes.

GLURP secondary PCR

Secondary GLURP PCR was carried out using 2 µl of the primary PCR as template. The reaction comprised the same components as the primary PCR but with primers G-NF and G-F4 (see Appendix A). Reaction conditions were the same but carried out for 25 cycles. PCR products were visualised by gel electrophoresis on a 2% agarose gel in TBE and stained with ethidium bromide.

3.2.8 Pf3k *var* gene database search, primer design and sequencing

Pf3k is a project set up as part of MalariaGEN which aims to sequence 3,000 *Plasmodium falciparum* genomes and is led by the University of Oxford, the Wellcome Trust Sanger Institute and the Broad Institute (https://www.malariagen.net/projects/parasite/pf3k). As a result, a database of over 62,000 *var* genes has been assembled. The short DBLα tags generated in this study were sent to Thomas D. Otto at the Wellcome Trust Sanger Institute where he performed a BLAST search of the database with parameters of at least 99% identity and 95% overlap. The returned sequence hits were sent back in FASTA format. The hits were used as reference genes to design primers (see 2.2.3) with overlapping products of ~1 kb along the length of the gene (see Appendix A for primer sequences). Predicted melting temperatures of 50-53 °C were aimed for to allow simultaneous PCR. The primers were tested on parasite gDNA (3.2.6) and the products cloned into pCR™4-TOPO® vector (Invitrogen) and sequenced as described (3.2.4). Overlapping sequences were assembled to reconstruct each *var* gene.

3.2.9 UPS and exon 2 PCR

5' UPS PCR was carried out using primers unique to the three main UPS types A, B and C (Mugasa *et al.*, 2012) (see Appendix A) in combination with the reverse DBLα specific primer designed for RT-qPCR (3.2.5). Reaction composition was as described (see 2.2.3) and reaction conditions were as follows: 95 °C for 3 minutes, followed by 30 cycles of 95 °C for 30 seconds, 52 °C for 30 seconds, 65 °C for 30 seconds, and a final extension of 65 °C for 3 minutes. Downstream PCR was carried out using a primer designed to the more conserved exon 2 (Lavstsen *et al.*, 2012) in

combination with an isolate specific forward primer (see Appendix A). Reaction composition was as described (see 2.2.3) and reaction conditions were as follows: 96 °C for 1 minute, followed by 40 cycles of 98 °C for 6 seconds, 48 °C for 15 seconds, 67 °C for 5 minutes, and a final extension of 68 °C for 10 minutes.

3.3 Results Part I: ICAM-1 binding patient isolates

3.3.1 Expressed DBLα tags of ICAM-1 binding patient isolates

Five ICAM-1 selected, culture-adapted patient isolates – BC12, J1, PCM7, PO69, 8206 – were re-selected on ICAM-1 and returned to culture to multiply before RNA extraction. As *var* genes are known to switch in culture (Recker *et al.*, 2011), the number of cycles between selection and RNA extraction was kept to a minimum and is indicated in the nomenclature chosen e.g. $BC12^{C2}$ refers to RNA extracted from isolate BC12 two cycles after ICAM-1 selection. "ukn" refers to parasite cultures that have previously been selected on ICAM-1 but have been cultured for an unknown number of cycles before and after cryopreservation. RNA quality and concentration was measured by NanoDrop™ before DNase I treatment and cDNA synthesis. Quality was defined as the absence of residual DNA, as detected by PCR of the no-reverse transcriptase (no-RT) cDNA synthesis reaction, suggesting successful DNase I treatment.

DBLα tags were PCR amplified using universal primers DBLαAF' and DBLαBR (Bull *et al.*, 2005) to the DBLα domain and cloned. Twenty four clones were sequenced per isolate with the exception of $PCM7^{ukn}$ for which 20 clones were sequenced. The number of unique sequences varied between isolates but all contained a prominent DBLα tag that was cloned at high frequency (Figure 3.1). DBLα tags were named after their patient isolate and lettered a, b, c, etc, in order of abundance. BC12 displayed a similar distribution of DBLα tags regardless of the number of cycles after selection (2, 8 and unknown) showing a clearly prominent tag, BC12a. $PCM7^{C3.1}$ and $PCM7^{C3.2}$ cDNA samples were reverse transcribed from the same RNA

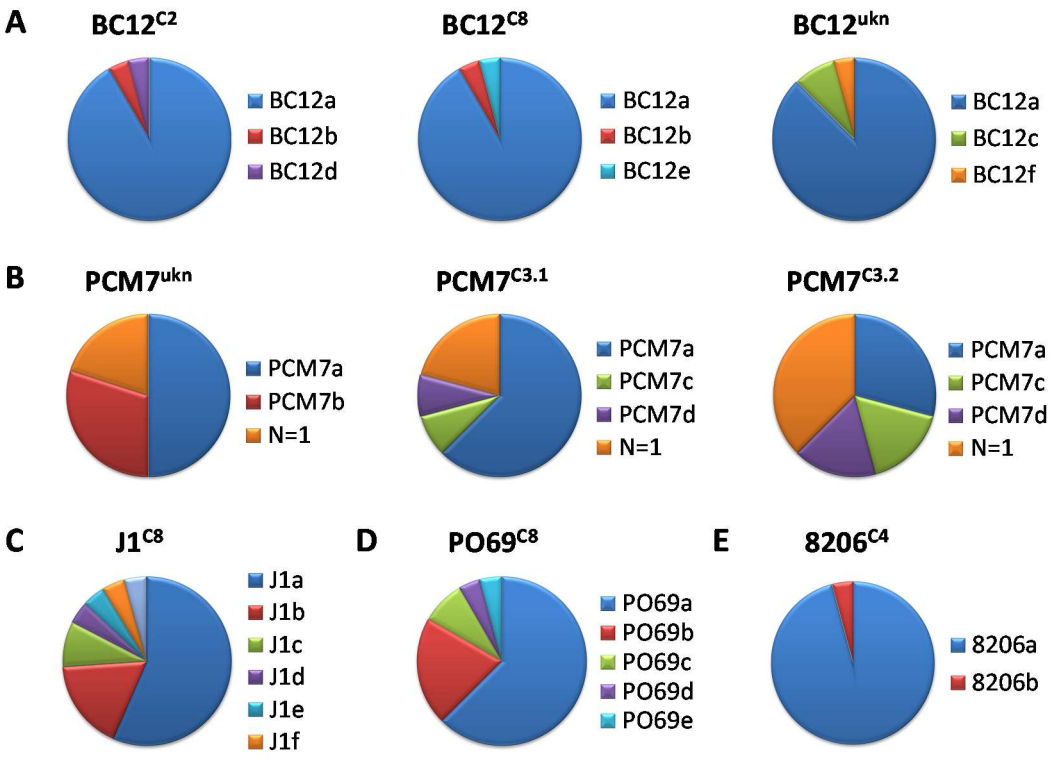


Figure 3.1. Distribution of expressed DBLα tags from ICAM-1 binding patient isolates.

Twenty four DBLα tags were amplified and sequenced from parasite cDNA. Each pie chart represents the distribution of DBLα tag sequence variants from parasite isolates BC12 **(A)**, PCM7 **(B)**, J1 **(C)**, PO69 **(D)** and 8206 **(E)**. The number of cycles between ICAM-1 selection and RNA extraction are indicated in superscript (C_). "ukn", unknown number of cycles. DBLα tags are named a, b, c, etc for each isolate. A high number of unique DBLα tags were sequenced from PCM7 and are referred to as N=1.

sample ($PCM7^{C3}$) and, although the same prominent DBLα tags were cloned from these two samples, the distribution varied. Most notable was the number of single DBLα tags detected: 9 in $PCM7^{C3.2}$, the highest detected in all samples, and 5 in $PCM7^{C3.1}$. $J1^{C8}$ and $PO69^{C8}$ showed a similar distribution of cloned DBLα tags, with the prominent tag cloned 13/24 and 15/24 times, respectively. Only two sequences were cloned from 8206^{C4}: 23 of 8206a and 1 of 8206b.

3.3.2 Dominantly expressed DBLα tags confirmed by RT-qPCR

Reverse transcriptase quantitative PCR (RT-qPCR) was performed to confirm that the most frequently cloned DBLα tags were in fact the dominantly expressed tags from each isolate. Specific primers to DBLα tags were designed (see Appendix A) and tested in a regular PCR reaction to confirm single band amplification (not shown). Primer efficiency was then tested by performing RT-qPCR on serially diluted cDNA to generate a standard curve (J1f example shown, Figure 3.2A and B, right panel) from which primer efficiency was calculated (see 3.2.5). Primer efficiencies were considered comparable if they were within 10% of each other (1.8-2.2) (Schmittgen and Livak, 2008) and any outside this range were re-designed. Two exceptions were primers to BC12e and BC12f for which expression was too low to generate a standard curve. However, each generated a single peak on the melt curve cycle of the RT-qPCR program (similar to dissociation curves shown in Figure 3.2A and B), suggestive of a single amplification product. ASL and seryl- tRNA synthetase (STS) were used as internal control genes (Viebig *et al.*, 2005). The no-RT and DPEC-H_2O controls were included on each plate (example shown, Figure 3.2C and D, respectively).

Having shown primer efficiencies to be comparable, RT-qPCR was repeated on each cDNA sample at a single concentration with all primer pairs designed to DBLα tags from that particular isolate on the same plate to minimise technical variation. Mean Ct values were normalised against the ASL internal control gene generating $2^{-\Delta Ct}$ values and individual values are expressed as a percentage of the sum of $2^{-\Delta Ct}$ values from each cDNA template (Figure 3.3). RT-qPCR data matches the expression patterns predicted from DBLα tag cloning of BC12, PO69 and 8206 (Figure 3.1), identifying BC12a, PO69a and 8206a as the dominantly expressed DBLα tags of these isolates (Figure 3.3). PO69b and PO69c are also expressed at high levels as predicted (Figure 3.1, Figure 3.3). PCM7, for which the predominant clone was PCM7a and for which clones PCM7b, PCM7c and PCM7d were represented at relatively high levels (Figure 3.1), in fact dominantly expresses PCM7a and

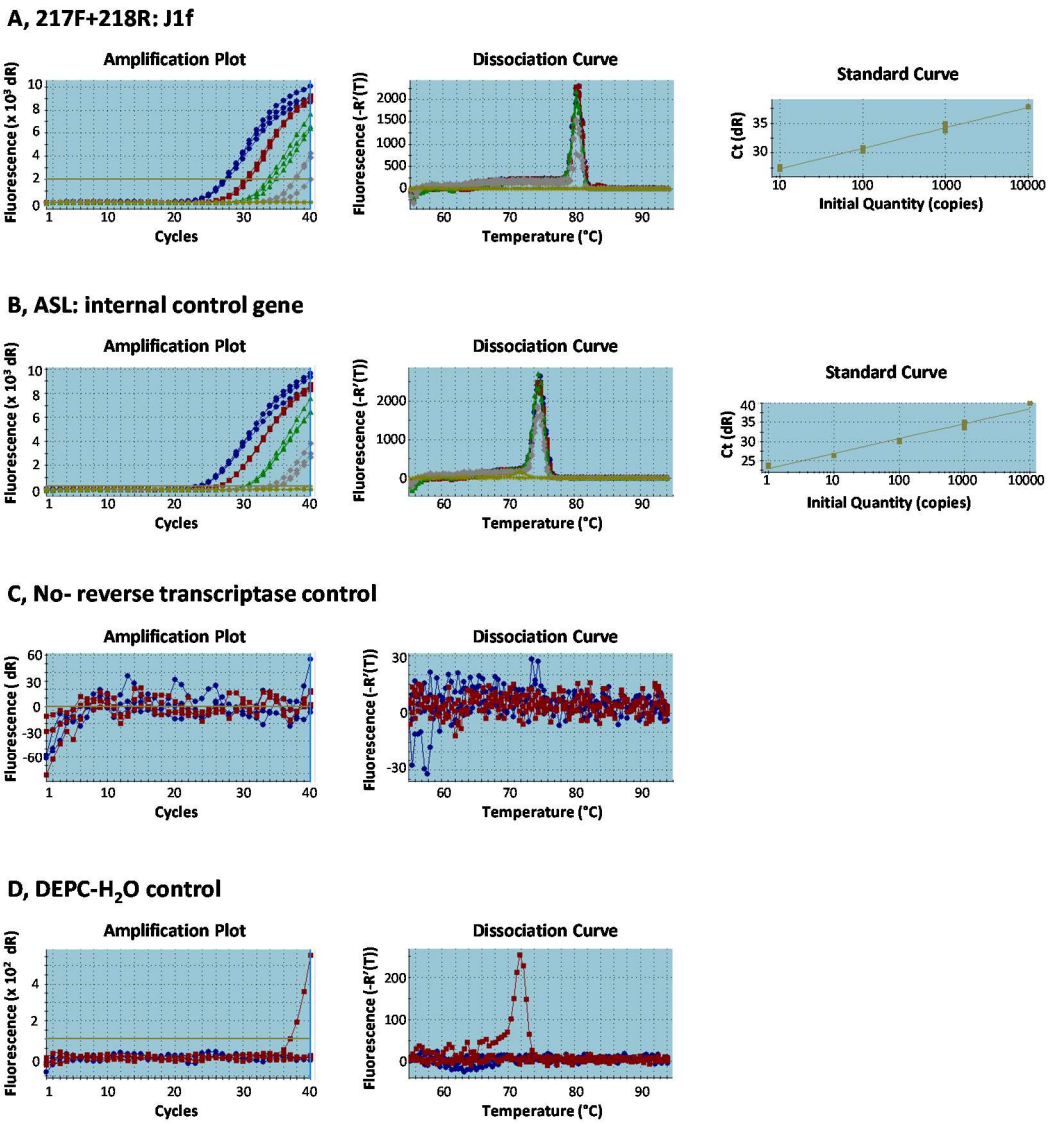


Figure 3.2. Example of RT-qPCR data output.

Primer efficiency was tested on serially diluted cDNA from $J1^{C8}$ using 217F+218R primers to J1f DBLα tag **(A)** and primers to internal control gene adenylosuccinate lyase (ASL, **B**). Amplification plots show the fluorescence signal detected over 40 cycles of RT-qPCR amplification with each cDNA concentration colour-coded (**A**, **B**). Dissociation curves show the melting temperature of the reaction products with a single peak indicating specific amplification of a single RT-qPCR product (**A**, **B**). Standard curves plot the average Ct value of triplicate wells against concentration of cDNA (10-fold serial dilutions represented by an arbitrary log scale, **A**, **B**). Amplification plots and dissociation curves of no-reverse transcriptase (**C**) and DEPC-H_2O (**D**) controls were carried out on the most concentrated cDNA to show a lack of specific amplification. Blue: triplicate wells of each control, red: triplicate wells with DEPC-H_2O (**C**, **D**).

expresses only PCM7d at a high level, with PCM7b and PCM7c hardly detected (Figure 3.3). It is interesting to note that PCM7d DBLα tag was not cloned from PCM7^{ukn} cDNA and yet is expressed at a relatively high level, highlighting a pitfall of this method of detection by cloning. Finally, J1 dominantly expresses J1a and has high levels of J1d and notable levels of J1b (> 10%, Figure 3.3).

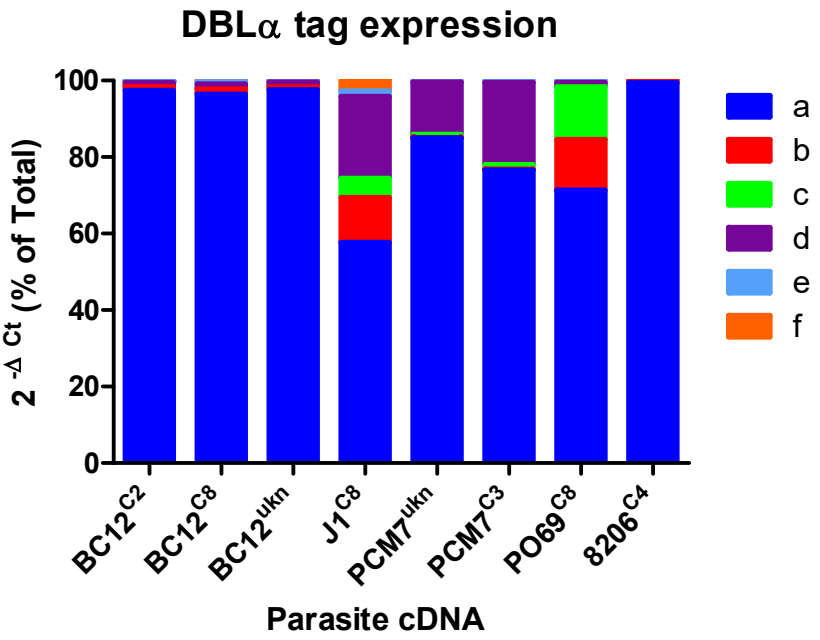


Figure 3.3. Relative DBLα tag expression as determined by RT-qPCR.
RT-qPCR was carried out with primers specific to each DBLα tag identified by cloning. Ct values were normalised against the ASL internal control gene to give $2^{-\Delta Ct}$ values and are shown as percentage of total for each cDNA. a, b, c, etc refer to DBLα tags sequenced from each individual isolate as shown in Figure 3.1.

3.3.3 Sequence comparison reveals PO69 and 8206 share DBLα tags with IT4 parasites

We performed a sequence alignment of all DBLα tags (not shown) and found no overlap between BC12, J1 and PCM7 sequence tags. However, the dominant DBLα tags PO69a and 8206a share 100% sequence identity. Likewise 8206b shares 100% sequence identity with PO69b. A BLAST search of the NCBI database identified all

DBLα tags cloned from PO69 and 8206 to share identity with *var* genes from the lineage IT (IT4/ItG), a common reference strain cultured in our lab. Through sequence alignments we identified PO69a and 8206a as IT4var16, PO69b and 8206b as IT4var12 and PO69c as IT4var41. This result prompted merozoite surface protein (MSP) and glutamate-rich protein (GLURP) genotyping.

3.3.4 MSP and GLURP genotyping of all isolates

MSP genotyping was carried out on gDNA from all five ICAM-1 selected patient isolates and the reference strain ItG (Figure 3.4). gDNA from an original, unselected sample of PO69 was included for comparison. An original sample of 8206 was unavailable. MSP2 N5 PCR shows ItG, PO69 ICAM-1 selected and 8206 share a band pattern which is consistent after both *Dde*I and *Scr*FI digest (Figure 3.4A, lanes 1, 3 and 4, respectively). This is corroborated by identical length of products in the GLURP PCR (Figure 3.4B, lanes 1, 3 and 4). PO69 unselected has a unique band pattern compared to PO69 ICAM-1 selected, ItG and 8206 suggesting that it began as a distinct genotype (Figure 3.4A and B, lane 2). PO69 unselected has a similar band pattern to PCM7 by both MSP2 N5 and GLURP PCR (Figure 3.4, lanes 2 and 7, respectively) but differs after MSP2 N5 restriction digests (Figure 3.4A, right gel, lanes 2 and 7). BC12 and J1 have a similar pattern in MSP2 N5 PCR and are only subtly different after digestion (Figure 3.4A, lanes 5 and 6, respectively). However, the size difference by GLURP PCR confirms they are unique genotypes (Figure 3.4B, lanes 5 and 6). No further work was carried out on PO69 ICAM-1 selected and 8206 as both cultures were shown to share the genotype of ItG parasites. Repeating ICAM-1 selection of PO69 unselected cultures could not be undertaken due to time constraints. Our objective to identify novel ICAM-1 binding *var* genes from patient samples continues with isolates BC12, J1 and PCM7.

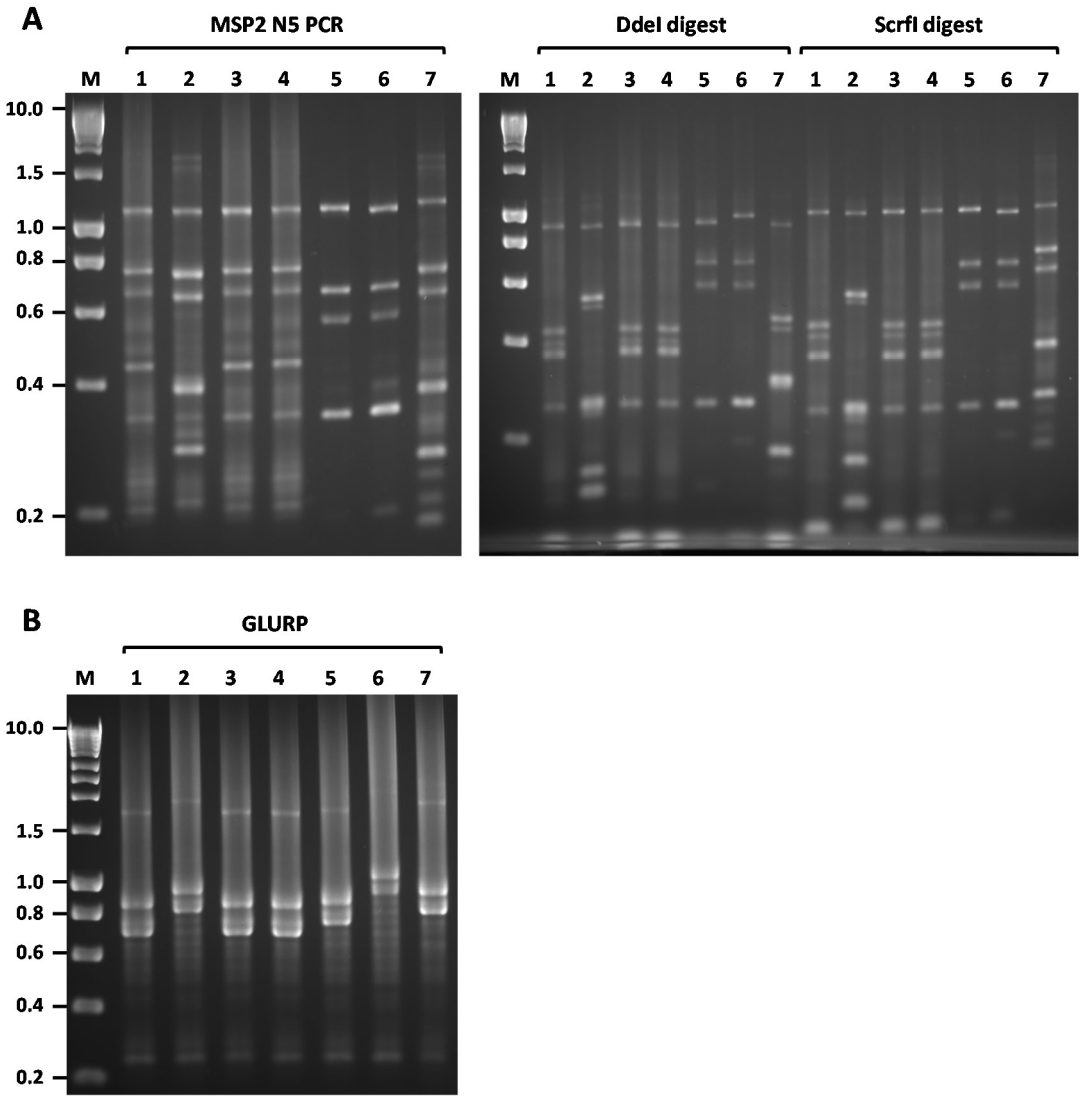


Figure 3.4. Genotyping of all parasite isolates.
MSP2 (**A**) and GLURP (**B**) genotyping was carried out as described in Methods section 3.2.7 on 1: ItG, 2: PO69 unselected, 3: PO69 ICAM-1 selected, 4: 8206, 5: BC12, 6: J1 and 7: PCM7 DNA. M: HyperLadder^{TM} 1 kb (Bioline), sizes indicated are in kilobases (kb).

3.3.5 Reconstructing full length *var* genes

This section now focuses on three Thai isolates – BC12, J1 and PCM7 – and their highly expressed DBLα tags. Six genes are analysed: the three dominantly expressed tags of each isolate, BC12a, J1a and PCM7a, and the secondary tags J1b, J1d and PCM7d which were expressed at a relatively high level. We first performed 5' upstream (UPS) typing of the six genes and made use of the new *var* gene database

generated by the Pf3k genome sequencing project to test whether you can predict a full length *var* gene sequence from a short DBLα tag. Sequence predictions of this kind that are easily tested experimentally could lead to faster sequencing of these highly variable genes, thus making it easier to analyse and test the predicted binding properties of the resulting PfEMP1 proteins.

5' UPS type classification

UPS typing performed using UPS primers A, B and C (Mugasa *et al.*, 2012) revealed BC12a, J1d, PCM7a and PCM7d to be UPS B type and J1a UPS C type. Two sequences were amplified from J1b with the UPS B primer but neither shared sequence identity with J1b DBLα. Another set of UPS primers (Kaestli *et al.*, 2006) tested on J1b DNA also failed to produce the correct sequence (data not shown). We were therefore unable to UPS type J1b (represented by red bar in Figure 3.6C).

Pf3k *var* gene database search yields variable number of hits

Highly expressed DBLα tags were BLAST searched against the Pf3k *var* gene database with parameters of at least 99% identity and 95% overlap. Five of the six DBLα tags resulted in ≤ 10 hit sequences each, whereas J1a resulted in 153 hits (Table 3.1).

Table 3.1. The number of hits to each DBLα tag in the Pf3k *var* gene database detected by BLAST search.

DBLα tag	# database hits
BC12a	**2**
J1a	**153**
J1b	**1**
J1d	**3**
PCM7a	**4**
PCM7d	**10**

The returned sequences to the DBLα tags with the fewest hits, BC12a, J1d and PCM7a, shared 99% sequence identity and are therefore essentially the same genes. Sequence hits to PCM7d (N=10) shared 97-99% identity, with one hit containing an additional ~200 bp insertion. The 153 hits to J1a share sequence similarity over ~3 kb at the 5' end but are variable downstream. Therefore, the more sequence hits that are returned from the DBLα tag search, the more likely it is that they will be of varying full length sequence, at least in our small sample set.

Novel primers to Pf3k hits successfully amplify products *in vitro*

Returned sequence hits to each DBLα tag were chosen as reference sequences for primer design based on length and consensus sequence. Primers were designed every ~1 kb to produce overlapping products (see Figure 3.5 for BC12a example) which were cloned and sequenced.

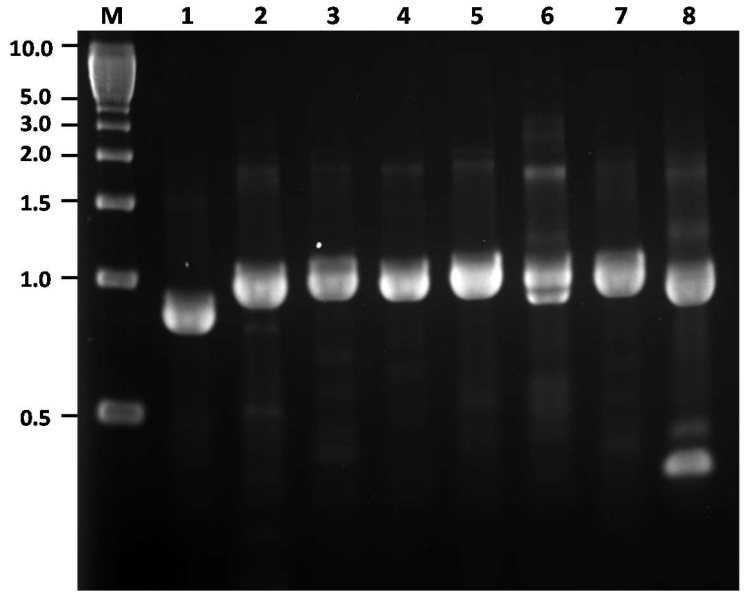


Figure 3.5. BC12a PCR fragments amplified with primers designed to XX0156-C.g40 Pf3k database hit.

Pf3k database sequence XX0156-C.g40 was returned to the BC12a DBLα tag. Primers were designed along the length of the returned sequence to produce overlapping products and used to PCR amplify BC12 gDNA. Primer pairs are 1: 227F-228R, 2: 229F-230R, 3: 231F-232R, 4: 233F-234R, 5: 235F-236R, 6: 237F-238R, 7: 239F-240R, 8: 241F-242R, see Appendix A for primer sequences. M: 1 kb DNA Ladder (NEB).

Schematics of the reference sequence match and the amplified PCR fragments are shown for each DBLα tag in Figure 3.6. Remarkably, BC12a, J1a and J1b fragments shared ≥ 99% identity with their respective reference sequences (Figure 3.6A, B, C). J1d fragments downstream of the DBLα tag shared ≥ 99% identity with its reference sequence but the upstream fragment had no significant similarity (Figure 3.6D, obtained by UPS PCR). PCM7a fragments amplified across three quarters of the length of the reference gene with ≥ 99% identity but the remaining fragments failed to amplify (Figure 3.6E). PCM7d fragments also shared ≥ 99% identity with their reference sequence with the exception of two 3' fragments whose sequences were highly AT rich, suggesting that priming occurred within the intron (Figure 3.6F, denoted by question marks).

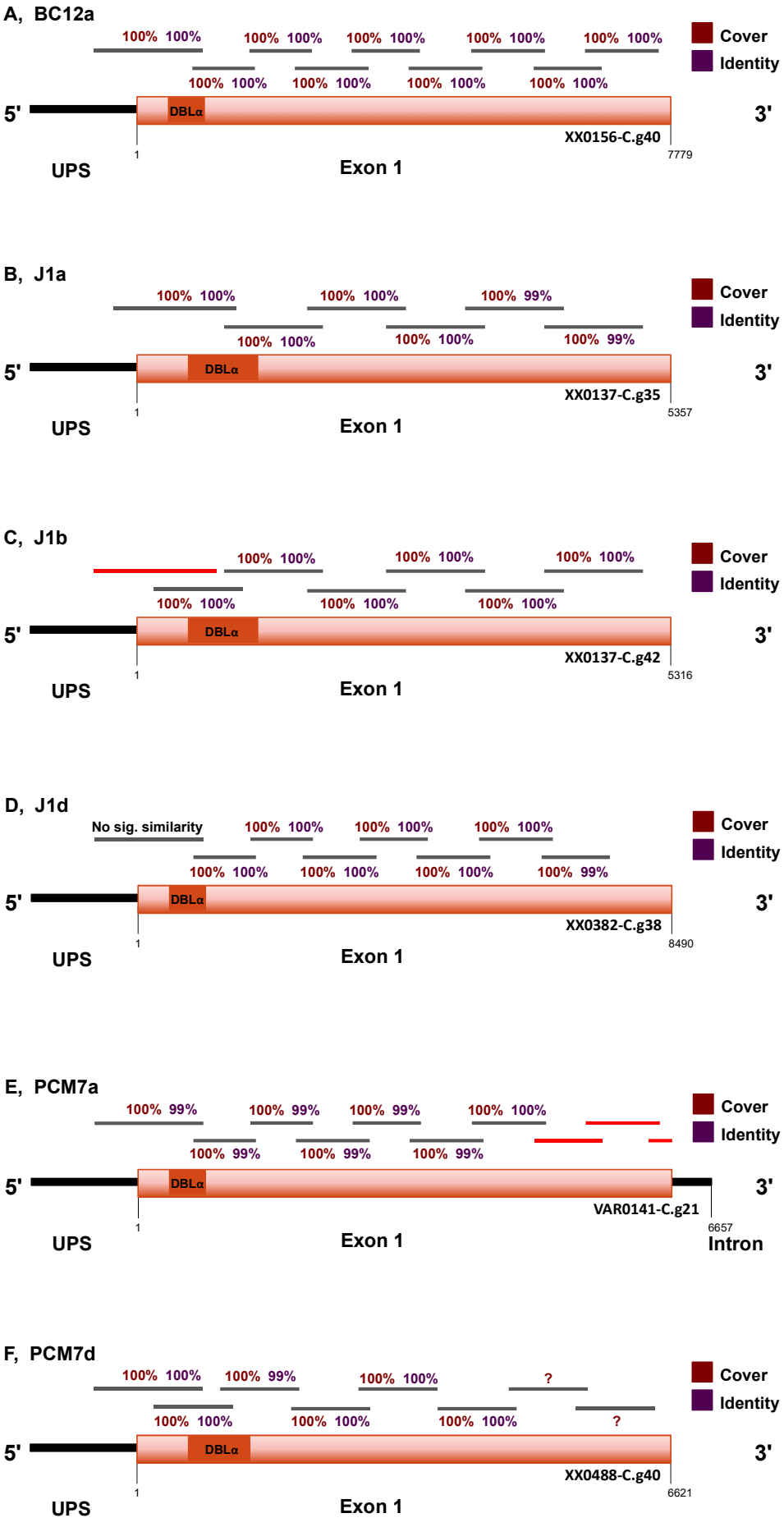

A, BC12a
100% 100%
100% 100%
100% 100%
100% 100%
100% 100%
100% 100%
100% 100%
100% 100%
100% 100%
Cover
Identity
5'
DBLα
3'
XX0156-C.g40
1
7779
UPS
Exon 1
B, J1a
100% 100%
100% 100%
100% 99%
100% 100%
100% 100%
100% 99%
Cover
Identity
5'
DBLα
3'
XX0137-C.g35
1
5357
UPS
Exon 1
C, J1b
100% 100%
100% 100%
100% 100%
100% 100%
100% 100%
100% 100%
Cover
Identity
5'
DBLα
3'
XX0137-C.g42
1
5316
UPS
Exon 1
D, J1d
No sig. similarity
100% 100%
100% 100%
100% 100%
100% 100%
100% 100%
100% 99%
Cover
Identity
5'
DBLα
3'
XX0382-C.g38
1
8490
UPS
Exon 1
E, PCM7a
100% 99%
100% 99%
100% 99%
100% 100%
100% 99%
100% 99%
100% 99%
Cover
Identity
5'
DBLα
3'
VAR0141-C.g21
1
6657
UPS
Exon 1
Intron
F, PCM7d
100% 100%
100% 99%
100% 100%
?
100% 100%
100% 100%
100% 100%
?
Cover
Identity
5'
DBLα
3'
XX0488-C.g40
1
6621
UPS
Exon 1

Figure 3.6. Schematic of DBLα tags and their reference gene matches from the Pf3k *var* database.

DBLα tags from each dominantly expressed *var* gene were searched against the Pf3k *var* database and the resulting match was treated as a reference gene for PCR primer design. Reference gene names are shown below each schematic. PCR fragments amplified from parasite gDNA are indicated by dark grey bars. Sequence coverage and identity to the reference are shown, as determined by BLAST Alignment. Upstream fragments were amplified using published primers to UPS A, B and C types (Mugasa *et al.*, 2012). Red bars indicate fragments that fail to amplify. "?" in **F** indicate that both fragments appeared to prime in intron sequence and could not be assembled.

3' sequences obtained/predicted by exon 2 PCR

A conserved primer to exon 2 (Lavstsen *et al.*, 2012) was used to fully sequence genes PCM7a and J1b (Figure 3.7A and B, respectively). We were unable to gain the complete sequence of genes BC12a, J1a, J1d and PCM7d due to either multiple products of the exon 2 PCR or cloning difficulty. The exon 2 PCR of these genes, shown in Figure 3.7C, provides some insight into the length of the genes yet to be sequenced. The amplification products are of size 1.9 kb for BC12a, 2.5 kb for J1a and 1.4 kb for J1d (Figure 3.7C, lanes 1, 2 and 3, respectively). PCM7d produced multiple products and therefore no size predictions can be made (Figure 3.7C, lane 4). This is likely due to the low annealing temperature of 48 °C, required for the reverse Ex2-reg primer, affecting the priming of the forward primer 297F (see Appendix A for primer sequences). Considering the introns of *var* genes are between 0.17-1.2 kb (Gardner *et al.*, 2002), the maximum coding sequence we are yet to obtain from BC12a and J1d is 1.73 kb and 1.23 kb, respectively. It is therefore likely that BC12a and J1d are sequenced up to the final extracellular domain with the average DBL domain between 0.96-1.5 kb and the average CIDR domain between 0.77-0.83 kb (Rask *et al.*, 2010). Additionally, the TM region of ATS which is encoded 5' prime to the splice site is 0.08 kb (Rask *et al.*, 2010). J1a has between 1.3-2.33 kb yet to be sequenced, likely encoding two or three additional 3' domains.

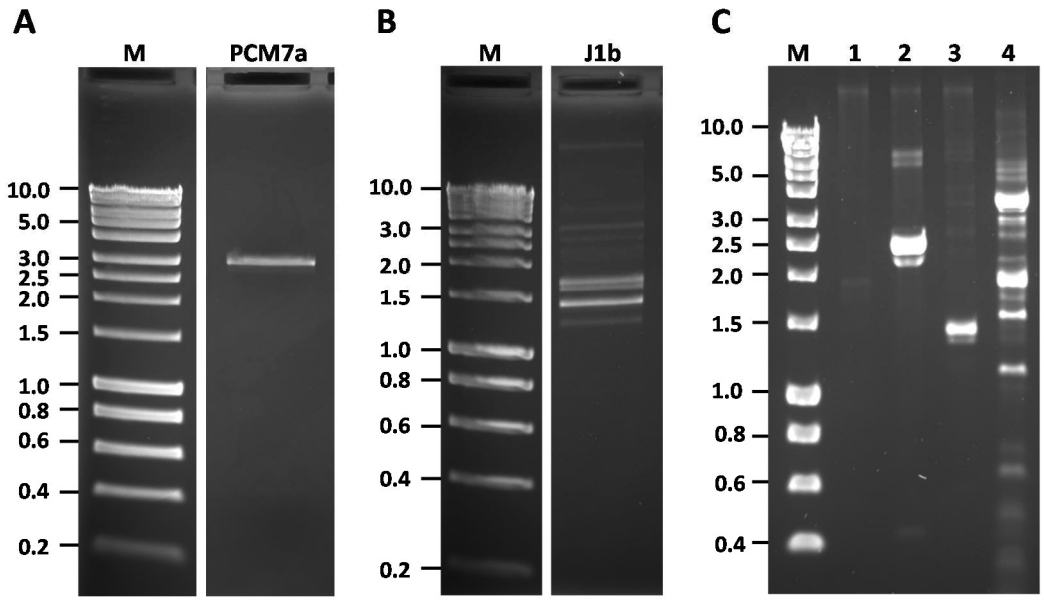


Figure 3.7. Exon 2 PCR of ICAM-1 binding genes.

PCR was carried out using a conserved primer to exon 2 (Ex2-reg) in combination with a gene specific primer (see Methods 3.2.9). PCM7a, primer 73F (A), and J1b, primer 253F (B), products were successfully cloned and sequenced. Unrelated lanes have been removed (A, B). C, DNA gel showing exon 2 PCR products of the ICAM-1 binding genes that could not be sequenced. 1: BC12a, primer 241F, 2: J1a, primer 310F, 3: J1d, primer 285F, 4: PCM7d, primer 297F. See Appendix A for primer sequences. M: HyperLadder^{TM} 1 kb (Bioline), sizes indicated are in kilobases (kb).

Sequence reconstruction

Overlapping sequence fragments were assembled to reconstruct these six highly expressed *var* genes from ICAM-1 binding parasite isolates: BC12a_1-7764, J1a_1-4254, J1b_1-5362, J1d_1-6884, PCM7a_1-6926 and PCM7d_1-4369.

J1d and PCM7a differential sequences present in Pf3k *var* database

The differential sequences obtained from J1d upstream of the DBLα tag (J1d_1-567) and PCM7a at the 3' end of the gene (PCM7a_4798-6762) were searched against the Pf3k *var* gene database for any sequence hits. J1d_1-567 sequence was present in the database within a gene that differed from J1d_1-6884 downstream from base 567. PCM7a_4798-6762 sequence was present in the database within a gene that differed from PCM7a_1- 6762 upstream from base 4798. Interestingly, PCM7a_4798-6762 was also present in a gene that matched the full length

PCM7a_1- 6762 completely, suggesting that in the time between our initial database search and carrying out the experiments, the sequence was added to the database.

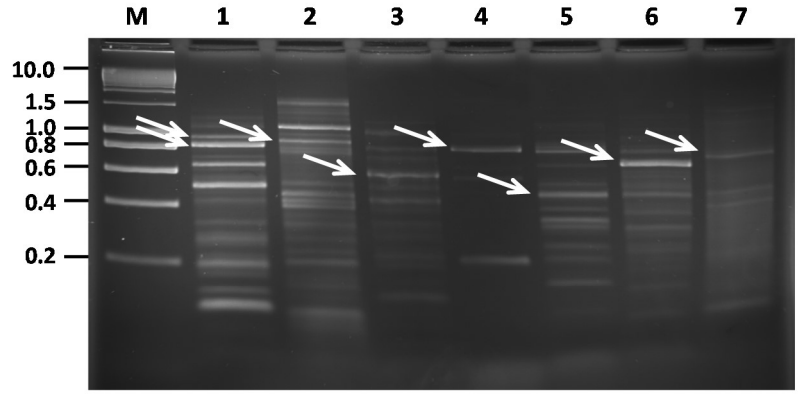
3.4 Results Part II: Brain-specific *var* genes in fatal cerebral malaria

Having shown proof of concept in Results section 3.3, we went on to test this method of searching the *var* gene database for hits with two sequences detected in the brains of multiple patients who had died of cerebral malaria (Tembo *et al.*, 2014). Sequences 28B1-1 (DBLα tag, 0.4 kb, accession number KC678110) and 62B1-1 (DBLα- DBLβ, 2.5 kb, accession number KC678109) were searched against the Pf3k *var* database by Thomas D. Otto, as above. A notable difference with these samples is that only gDNA extracted from tissue biopsies was available for testing.

3.4.1 Pf3k *var* database search and attempted reconstruction of 62B1-1

There were no hits to 28B1-1 sequence in the Pf3k *var* database. 62B1-1 resulted in two hits which share 99% identity with 62B1-1 and 96% identity with each other. The longer of the two hits, PF0311-C.g26 (10.3 kb) was used to design PCR primers along the length of the gene (see Appendix A). These primers were tested on gDNA samples from patients in which 62B1-1 sequence was dominantly expressed in the brain. PCR often resulted in multiple bands, making cloning difficult. An example of PCR products obtained is shown in Figure 3.8A with details of the resulting cloned and sequenced fragments shown in Figure 3.8B. Only minimal PCR optimisation could be carried out due to the low quantity of the sample.

A



B

Lane	Primers*	Expected size (bp)	# P.f/total seq.	Coverage c.f. PF0311-C.g26	Identity c.f. PF0311-C.g26
1	57F+58R	830	3/5, 0/5	32	83
2	59F+60R	810	0/5	-	-
3	63F+64R	589	4/5	100	87
4	65F+66R	796	-	-	-
5	66R+69F	541	0/5	-	-
6	69F+70R	719	5/5	84	90
7	71F+72R	805	2/2	21	93

Figure 3.8. Example of 62B1-1 PCR using primers designed to PF0311-C.g26.
A, Primers designed to the Pf3k *var* gene database hit PF0311-C.g26 were used to PCR 62B1-1 gDNA. White arrows indicated bands that were purified and cloned for sequencing. M: HyperLadder^{TM} 1 kb (Bioline), sizes indicated are in kilobases (kb). **B**, Details of PCRs 1-7. "*" primer details can be found in Appendix A. # P.f/total seq. refers to the number of *P. falciparum* fragments sequenced out of the total fragments successfully cloned. Coverage and identity of each fragment to PF0311-C.g26 are listed. A summary of all *P. falciparum* fragments sequenced is shown in Figure 3.9.

Overall, of the 37 primer combinations tested, only 7 primer pairs resulted in amplification of *Plasmodium* DNA and these displayed either low coverage or sequence identity of only 83% (summarised in Figure 3.9). Two exceptions are overlapping fragments which share 87 and 97% identity with the reference sequence and 93% when combined (Figure 3.9, black bar). However, these fragments are not in the overlap area with 62B1-1 and, without overlapping

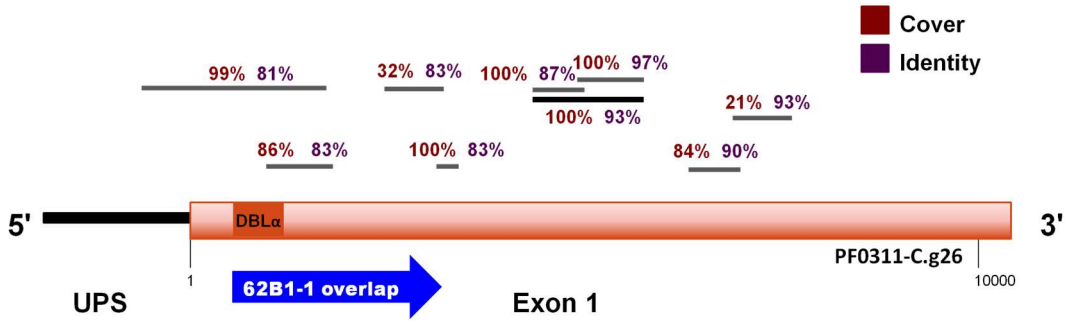


Figure 3.9. Schematic of 62B1-1 database match and successfully amplified PCR fragments from clinical gDNA samples.
PCR primers were designed to the 62B1-1 Pf3k *var* database match PF0311-C.g26 and tested on gDNA from tissue biopsies. Successfully amplified *P. falciparum* fragments are indicated by grey bars with their sequence coverage and identity to the reference shown, as determined by BLAST Alignment. Black bar indicates consensus in overlapping sequence between two fragments. The majority of fragments amplified were human host DNA (not shown).

upstream fragments, we cannot be sure that they originate from the same *var* gene. The remaining fragments sequenced were human host DNA, as confirmed by BLAST search.

3.4.2 Parasite DNA enrichment

In an attempt to isolate parasite DNA from the 62B1-1 gDNA sample, we used NEBNext® Microbiome DNA Enrichment Kit which separates microbial DNA from that of the human host. The kit control reaction successfully separated the target *E. coli* DNA from the human IMR-90 host DNA (Figure 3.10, faint band), whereas the 62B1-1 target DNA is not visible on the gel and the host DNA band is comparable to the input DNA (Figure 3.10). This result was confirmed by NanoDrop® measurement of 62B1-1 target DNA which was below accurate levels of quantitation. Further optimization of this method was not carried out due to low sample volume. Therefore, due to the very low abundance of parasite DNA and lack of material for reaction optimisation, we did no further work with these samples.

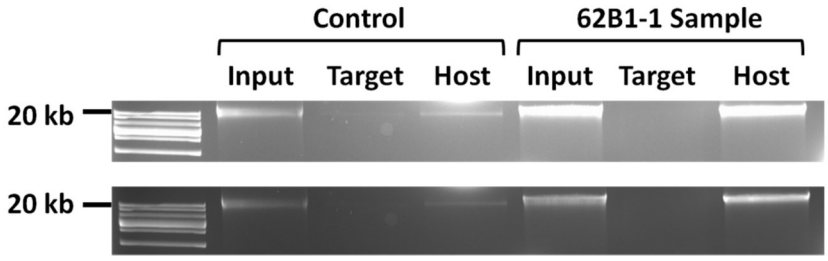


Figure 3.10. Parasite DNA enrichment of 62B1-1 gDNA sample.
NEBNext® Microbiome DNA Enrichment was carried out on kit Control, human IMR-90/*E.coli* DNA at 10:1 ratio, and 62B1-1 Sample at unknown ratio. Input: mixture of microbial and human DNA. Target: enriched microbial DNA. Host: human host DNA minus microbial DNA. The same gel is shown visualised with different exposure times.

3.5 Discussion

This Chapter describes the use of the new Pf3k *var* gene database of > 62,000 *var* genes to test whether a full length *var* gene can be predicted from a short DBLα tag. We tested this method on two different sample types: cDNA from ICAM-1 selected culture-adapted patient isolates whose binding profiles have been recently characterised (Madkhali *et al.*, 2014), and genomic DNA samples from a previous study that identified brain-specific *var* genes in Malawian children who died of cerebral malaria (Tembo *et al.*, 2014).

In the case of the ICAM-1 selected patient isolates, we first identified the expressed *var* genes using DBLα tag primers (Bull *et al.*, 2005) which target conserved homology blocks of the DBLα domain and are thought to amplify the majority of *var* genes. Indeed, such primers are frequently used in surveys of *var* gene diversity (Sulistyaningsih *et al.*, 2013, Mugasa *et al.*, 2012, Chen *et al.*, 2011, Albrecht *et al.*, 2010). However, universal amplification by these primers cannot be known for certain due to the high variability of *var* genes. This, combined with reliance on cloning, which can carry bias towards some sequences, results in uncertainty that all expressed DBLα tags have been identified. This is highlighted with PCM7^{ukn} cDNA

from which DBLα tag PCM7d was not cloned but was shown to be expressed at a high level by RT-qPCR (Figure 3.1, Figure 3.3). However, the most frequently cloned tag did prove to be the dominantly expressed tag in all isolates indicating that any bias only affects secondary tags, at least in these culture-adapted isolates. The over-representation of some minor tags identified by cloning has been observed previously (Gatton *et al.*, 2006) and highlights the importance of the more quantitative RT-qPCR method. Repeating ICAM-1 selection and RNA extraction increases confidence in this method as seen with the reproducible results of three different BC12 RNA extractions. Replicates can also identify DBLα tags that may have been missed as with PCM7d. However, this proved difficult with isolate J1 which tended to produce gametocytes. This was also the case with several other isolates from the initial characterisation study (Madkhali *et al.*, 2014) from which sufficient RNA could not be extracted due to gametocytogenesis (data not shown).

We performed ICAM-1 selection of these isolates to identify new ICAM-1 binding *var* genes. This is not necessarily selecting for a clonal population as multiple *var* genes can be expressed within a culture. Therefore, expression of secondary tags in J1, PCM7 and PO69 can be explained by multiple *var* genes within an isolate with the ability to bind ICAM-1, as seen with multiple IT4 *var* genes (Howell *et al.*, 2008). Alternatively, *var* gene switching may have occurred after selection which occurs in an ordered manner and involves initially switching to numerous genes before a single dominant gene emerges (Recker *et al.*, 2011). We minimised the effect of switching by extracting RNA as quickly as possible after ICAM-1 selection. Whether these secondary genes also have the ability to bind ICAM-1 is yet to be determined.

Our finding that PO69 and 8206 isolates are in fact ItG parasites highlights the importance of regularly checking genotypes of patient isolates and comparing them to any reference strain cultured in the lab. A low level contamination with such an established strain will likely outcompete any patient isolates which are less adapted to culture conditions. An additional disadvantage lies in the propensity for these patient isolates to switch to gametocytes in culture, a phenotype not displayed by ItG parasites.

Classification of *var* genes is based on a combination of upstream sequence, genomic location and direction of transcription (Voss *et al.*, 2000, Gardner *et al.*, 2002) (see Figure 1.3). Group A genes have been associated with severe disease (Jensen *et al.*, 2004, Kyriacou *et al.*, 2006, Rottmann *et al.*, 2006) as well as group B in some studies (Rottmann *et al.*, 2006, Kaestli *et al.*, 2006). The majority of ICAM-1 binding *var* genes identified in the reference strains IT4 and 3D7 are group B and C type (Janes *et al.*, 2011, Howell *et al.*, 2008) with only two group A ICAM-1 binders identified in 3D7: PF11_0521 (Oleinikov *et al.*, 2009, Gullingsrud *et al.*, 2013) and PFD1235w (Jensen *et al.*, 2004, Bengtsson *et al.*, 2013b). Both UPS A ICAM-1 binding genes have been shown to induce adhesion blocking (rat) antibodies to their recombinant domains (Gullingsrud *et al.*, 2013, Bengtsson *et al.*, 2013b), with PFD1235w specific antibodies additionally able to block binding of other DC4 expressing parasites, displaying cross-reactivity (Bengtsson *et al.*, 2013b). The dominantly expressed *var* genes of our ICAM-1 binding patient isolates are UPS B type and one UPS C type, an unsurprising result given that group A *var* genes are seldom expressed *in vitro* (Peters *et al.*, 2007, Janes *et al.*, 2011, Ye *et al.*, 2015). This is related to high switching bias towards centrally located short *var* genes (4 extracellular domains), all of which are UPS B or C, rather than the specific avoidance of UPS A gene expression (Noble *et al.*, 2013). Assuming IT4 and 3D7 strains are representative of all parasites then we expect only 0-2 group A ICAM-1 binding *var* genes per genome and their identification from these patient isolates would require alternative methods to those employed here. For example, whole genome sequencing would identify the full *var* repertoires of these isolates and allow analysis of their predicted binding functions. However, sequence predictions alone cannot predict ICAM-1 binding, particularly of the UPS A genes, discussed further in Chapter 4.

The *var* gene database, which is a result of the Pf3k genome sequencing project, is a new resource with the potential to provide valuable information on an unprecedented scale. However, the copious amount of 'omics' data now being generated presents the challenge of interpretation and of identifying practical applications. We wanted to test whether this vast resource can predict full length

var genes from their short ~400 bp DBLα tag. The results obtained for the ICAM-1 binding patient isolates were remarkable with the majority of successfully cloned fragments sharing ≥ 99% identity with the reference sequence (Figure 3.6). The two cases of differential sequence (J1d_1-567 and PCM7a_4798-6762) occurred at a single point in the sequence suggesting a gene recombination event or possibly a sequence assembly error. The former is likely as it has been reported that mitotic recombination frequently occurs in the *var* genes (Bopp *et al.*, 2013, Claessens *et al.*, 2014). Interestingly, recombination occurs between domains of the same type resulting in in-frame genes and domain structure preservation which produces viable genes (Claessens *et al.*, 2014), concordant with our results. Our finding that PCM7a_1-6762 full length sequence was present in the database in our second search is perhaps unsurprising as the Pf3k project is ongoing and the database is constantly expanding. It does, however, raise the interesting question of whether we are close to sequencing the worldwide *var* gene repertoire. If this is the case, it must be kept in mind that these genes are constantly undergoing recombination events, both mitotic (Bopp *et al.*, 2013, Claessens *et al.*, 2014) and meiotic in the mosquito vector (Ranford-Cartwright and Mwangi, 2012). Therefore, although the sequence might be known, the order in which they are combined in any one gene is still unpredictable.

Limitations of the database are highlighted by the varying number of hits to each DBLα tag, in particular 28B1-1 for which there was no hit and J1a for which there were 153 hits. The length of contiguous sequence also varies in the database and may only provide partial sequence. We utilised a primer to the conserved exon 2 (Lavstsen *et al.*, 2012) in an attempt to sequence the 3' end of the genes. However, this can result in long PCR products that are difficult to clone and sequence, a problem avoided in the original study by using the long PCR products to create a library for whole genome sequencing (Lavstsen *et al.*, 2012). We were able to sequence the full length genes of J1b and PCM7a using this primer. However, we were unable to clone and sequence exon 2 PCR products of BC12a, J1a, J1d or PCM7d but were able to make size predictions of all but PCM7d products. The un-sequenced region of J1a is the longest and is the result of having 153 returned

database hits which differed significantly in the 3' end making specific primer design impossible in this region.

Our attempts to apply this approach of targeted sequencing of a single *var* gene to gDNA collected from cases of fatal cerebral malaria were unsuccessful. This is mainly due to the vast amount of human host DNA in the samples (Figure 3.10). In the original study, DNA was extracted from half of a biopsy of approximately 0.4 x 0.4 x 1 cm tissue which contained a majority of human DNA which contributed to the failure of genotyping in many of the extractions, shown in Tembo *et al* Figure S1 (2014). Human DNA in the sample was also thought to be responsible for difficulties with reproducibility of the RT-qPCR (Tembo *et al.*, 2014). This is unsurprising when taking into account the respective sizes of the human and *P. falciparum* genomes (3,200 Mb and 22.9 Mb, respectively), the fact that the majority of biopsy cells are human and the uncertainty of the number of parasites in the biopsy. For example, consider that a ratio of 100: 1 human to parasite cells would consist of 320,000 Mb of human and 22.9 Mb *P. falciparum* genetic material, a ratio of 14,000: 1. This presents a challenge for even the best of PCR primers to gain access to the correct complementary sequence. Of the fragments that were successfully sequenced, only one shared > 90% identity with the database match (Figure 3.9) and we cannot be certain this fragment originates from the same gene as the original 62B1-1 sequence.

Therefore, having shown this method to be successful in section 3.3 with pure DNA samples, failure to amplify PCR fragments can be attributed to vast quantities of human DNA in the sample and a lack of material for either purification of parasite DNA or extensive optimisation of PCR conditions. As a result, we were unable to sequence the full length of these *var* genes that were found specifically in cerebral malaria. A major failing of the original study was in not collecting parasite blood samples from these patients on hospital admission. This could have resulted in purer gDNA samples, through well established nucleated cell depletion techniques followed by DNA extraction, or in establishing stable parasite lines upon successful culture adaptation, allowing their detailed characterisation. An alternative method would have been to perform laser capture microdissection (for review see Datta *et*

al. (2015)) of the tissue biopsies to isolate individual parasites, a technique that has been used to isolate liver stage parasites (Semblat *et al.*, 2005). However, the short sequences of 28B1-1 and 62B1-1 that are available, found specifically in the brains of fatal cerebral malaria cases (Tembo *et al.*, 2014), could still be used in surveillance of parasites causing CM to assess the distribution of these sequences.

3.6 Summary

We have successfully identified and sequenced the dominantly expressed *var* genes of patient isolates BC12, J1 and PCM7 which have been selected for binding to the human endothelial receptor ICAM-1. The sequences were identified using universal DBLα tag primers which result in products of around 400 bp. These tags were then searched against the new Pf3k *var* gene database for sequence matches. We found that these sequence matches were a useful starting point for primer design in this targeted sequencing approach but that some sequences amplified from the parasites differ from the database hits. Therefore, returned sequence matches to short DBLα tags are useful predictors of full length *var* gene sequences but require experimental validation. We were unable to successfully apply this approach to the 62B1-1 sequence identified specifically in the brains of patients who died of CM mainly due to low sample volume and the majority of the sample consisting of human DNA. Extensive analysis of sequences identified from the ICAM-1 selected isolates and the predicted 62B1-1 sequence is carried out in the next Chapter 4 to characterise their predicted binding abilities.

Chapter 4. Sequence analysis of *Plasmodium falciparum var* genes

4.1 Introduction

Plasmodium falciparum var genes are responsible for antigenic variation and cytoadhesion of the parasite IEs, both of which are major virulence factors that contribute to pathogenicity, resulting in the most severe form of human malaria. The 50-60 *var* gene copies per genome can be classified into three main groups A-C and two subgroups B/A and B/C (Gardner *et al.*, 2002, Lavstsen *et al.*, 2003, Kraemer and Smith, 2003). Groupings are based on the 5' upstream sequence (UPS), chromosomal location and direction of transcription (see Figure 1.3). In addition, group A genes are characterised by DBLα1 domains, while groups B and C are characterised by DBLα0 domains (Gardner *et al.*, 2002, Lavstsen *et al.*, 2003, Kraemer and Smith, 2003). Groups B/A and B/C represent transitional genes with criteria from more than one group (Gardner *et al.*, 2002, Lavstsen *et al.*, 2003, Kraemer and Smith, 2003). The conserved *var2csa* gene is unique among the *var* genes and is flanked by UPS E sequence (Salanti *et al.*, 2003). See Introduction section 1.6 for more details.

The *var* genes encode modular PfEMP1 proteins which are made up of DBL (Duffy binding-like) and CIDR (cysteine-rich interdomain region) domains which are classified into types DBLα, β, γ, δ, ε, ζ and x and CIDRα, β, γ and δ based on sequence identity (Smith *et al.*, 2000b, Lavstsen *et al.*, 2003, Rask *et al.*, 2010). The landmark analysis carried out by Rask *et al.* (2010) of *var* genes from seven distinct genomes redefined domain boundaries and identified domain subtypes based on sequence similarity, which is remarkably low at 31-45% identity among domains of the same type. Domain combinations that are frequently found together were identified and termed domain cassettes (DCs) (Rask *et al.*, 2010). Since their description, several of these DCs have been associated with specific binding capabilities (Avril *et al.*, 2012, Claessens *et al.*, 2012, Lavstsen *et al.*, 2012, Turner *et al.*, 2013, Bengtsson *et al.*, 2013b, Berger *et al.*, 2013). Despite the low sequence

identity between PfEMP1 proteins, they consists of short conserved sequence blocks, interspersed amongst stretches of highly variable sequence, which have been identified as homology blocks (HBs) (Rask *et al.*, 2010). HBs provide an alternative to sequence comparison by alignment and distance tree analysis which determine sequence relatedness based on the whole domain, instead focussing on short sequence similarities. The VarDom 1.0 Server is an online resource which stores the sequence information described above for the seven genomes dataset and can perform analysis of new PfEMP1 sequences (Rask *et al.*, 2010).

We identified and reconstructed the dominantly expressed *var* genes of the ICAM-1 binding patient isolates BC12, J1 and PCM7 in Chapter 3. This chapter aims to characterise these *var* gene sequences by identifying their domain structure and analysing their predicted binding phenotypes. We perform in-depth analysis of the newly identified DBLβ domains, which mediate ICAM-1 binding, and investigate their similarity to known ICAM-1 binding sequences. Analysis of the non-dominant genes identified in these isolates is also carried out along with the Pf3k *var* gene database matches to the brain-specific *var* genes, 28B1-1 and 62B1-1, identified in fatal cases of cerebral malaria. Due to the nature of the data in this chapter, the Results and Discussion sections have been combined to avoid repetition.

4.2 Methods

4.2.1 Characterising PfEMP1 domain structures

Amino acid sequences of *var* genes were entered into the VarDom 1.0 Server (Rask *et al.*, 2010) (available at http://www.cbs.dtu.dk/services/VarDom/) in FASTA format to define domain boundaries. The corresponding nucleic acid sequences were then separated into domains and BLAST searched against the entire 7 genomes dataset (Rask *et al.*, 2010). The top six hits were analysed and domain subtypes were identified by searching Figure S4 from Rask *et al.* (2010). Subtypes were included if a consensus between the six hits was reached or if the top hits had appropriate coverage. Where multiple subtypes were identified, they were included

if there were two of equal standing (separated by "/") or excluded if there were more than two matches with similar coverage.

4.2.2 DBLα tag cysteine/position of limited variability (cys/PoLV) analysis

DBLα tags were classified using a Perl script (kindly provided by Peter Bull) to identify positions of limited variability and the number of cysteines (cys/PoLV) present as described (Bull *et al.*, 2007). DBLα tag amino acid sequences beginning with DIGDI and ending with PQFLR were analysed based on 4 anchor points: a, b c and d. PoLV1 is identified as the 4 amino acids starting 10 amino acids downstream of anchor point a, the N-terminal D residue. PoLV2 and 3 are based around anchor point b which is a WW motif or anchor point c, a VW motif, in the absence of anchor point b. In relation to anchor point b, PoLV2 is defined as the 4 amino acids directly upstream of point b and PoLV3 as 4 amino acids starting 9 amino acids downstream of point b. In relation to anchor point c, PoLV2 is defined as the 4 amino acids beginning 12 amino acids upstream of point c and PoLV3 as the 4 amino acids directly downstream of point c. PoLV4 is fixed in relation to anchor point d, the C-terminal R residue, and is defined as 4 amino acids beginning 11 amino acids upstream of point d (Bull *et al.*, 2007). The output of the Perl script consists of the cys/PoLV group number and a distinct sequence identifier in the format "PoLV1-PoLV2-PoLV3-number of cysteines-PoLV4-sequence tag length" (Bull *et al.*, 2007).

4.2.3 Phylogenetic comparison by maximum likelihood method

Amino acid sequences were aligned using ClustalX 2.1 (Larkin *et al.*, 2007). Aligned files were opened in MEGA 6.06 (Tamura *et al.*, 2013), converted to MEGA files and analysed by maximum likelihood method with 100 bootstrap replications. The Jones-Taylor-Thornton (JTT) substitution model was used. The resulting tree was exported as a Newick file and uploaded to iTOL (interactive Tree of Life, available at: http://itol.embl.de/itol.cgi) (Letunic and Bork, 2007, Letunic and Bork, 2011) for aesthetic editing.

4.2.4 Homology block analysis

Homology blocks (HBs) of DBLβ domains were downloaded from the VarDom 1.0 Server (Rask *et al.*, 2010) and analysed in relation to a maximum likelihood tree generated as described (see section 4.2.3). Information on ICAM-1 binding ability was obtained from the following references: Howell *et al.* (2008), Janes *et al.* (2011) and Bengtsson *et al.* (2013b).

4.2.5 Sequence alignments

Initial sequence alignments were carried out with ClustalX 2.1 (Larkin *et al.*, 2007). Short peptide sequences were then aligned with T-Coffee, available at: http://tcoffee.crg.cat/apps/tcoffee/index.html (Di Tommaso *et al.*, 2011, Notredame *et al.*, 2000), and the resulting fasta_aln files were copied into the BoxShade Server, created by K. Hoffman and M. Baron and available at: http://www.ch.embnet.org/software/BOX_form.html.

4.3 Results and Discussion

4.3.1 PfEMP1 domain structures of ICAM-1 binding patient isolates

The dominantly expressed *var* gene sequences of three ICAM-1 binding patient isolates, BC12, J1 and PCM7, were characterised in detail. Domain boundaries and types were defined by the VarDom 1.0 Server (Rask *et al.*, 2010) and each domain was BLAST searched against the 7 genomes data set (Rask *et al.*, 2010) to identify subtypes. Domain structures are represented schematically in Table 4.1. Three of the six genes contain domain cassettes (DCs) as indicated by dashed lines. BC12a contains DC14 (DBLα0.6-CIDRα3.1-DBLβ5) and J1a and J1d each contain DC22 (DBLα0.4/18-CIDRα6-DBLβ5). All are of DBLα type 0 with varying sub-classifications (DBLα0.6, 0.4, etc). J1b and PCM7a DBLα0 domains generated diverse hits from the BLAST search and could not be further classified. DBLα0 is almost exclusively associated with UPS B and C *var* genes (Rask *et al.*, 2010) and is therefore not associated with rosetting, an UPS A virulence phenotype associated with DBLα

subtypes 1.5, 1.8, 1.6 and 2 (Ghumra *et al.*, 2012). The CIDRα domains from each gene are of a variety of subtypes, 2-4 and 6 (Table 4.1), all of which, along with CIDRα5, are predicted to bind CD36 (Robinson *et al.*, 2003, Turner *et al.*, 2013). This prediction is in agreement with the ability of antibodies against CD36 to block binding of these parasites to HDMEC cells by > 60% under flow conditions (Madkhali *et al.*, 2014). CD36 binding is mainly observed in UPS B and C *var* genes which are associated with uncomplicated malaria (Kirchgatter and del Portillo, 2002, Kyriacou *et al.*, 2006, Rottmann *et al.*, 2006, Warimwe *et al.*, 2009).

Table 4.1. Schematic representation of the domain structures of the dominantly expressed *var* genes of ICAM-1 binding patient isolates.

Gene name	NTS	PfEMP1 domain structure						DC #
BC12a	B4/3	DBLα0.6	CIDRα3.1	DBLβ5	DBLγ5	DBLδ1	CIDRβ1	14
J1a	B3	DBLα0.4/12	CIDRα6	DBLβ5				22
J1b	-	DBLα0	CIDRα3.1/2	DBLδ1	CIDRβ1			
J1d	ND	DBLα0.4	CIDRα6	DBLβ5	DBLγ11/15	DBLζ3/6	DBLε12	22
PCM7a	B3	DBLα0	CIDRα2.5/9	DBLβ5/8	DBLδ1	CIDRβ1/3		
PCM7d	B6/3	DBLα0.18	CIDRα4	DBLβ3/5				

ND – not defined, no hits in Rask dataset.
Dashed lines outline domain cassettes (DCs).
"/" indicates similarity to two subtypes.

BC12a, J1a and J1d contain DBLβ5 which is predicted to mediate binding to ICAM-1 (Howell *et al.*, 2008, Janes *et al.*, 2011). Indeed, all DBLβ5 domains tested to date bind to ICAM-1 (Janes *et al.*, 2011). The DBLβ domains of PCM7a and PCM7d had BLAST hits to two different domain subtypes: 5/8 for PCM7a and 3/5 for PCM7d. These could be predicted to bind ICAM-1 if they are more similar to DBLβ5 type. In addition, three DBLβ3 domains tested have been shown to bind ICAM-1 but several others do not (Howell *et al.*, 2008, Janes *et al.*, 2011, Bengtsson *et al.*, 2013b). Interestingly, J1b does not contain a DBLβ domain, suggesting that it is not involved in ICAM-1 binding. This gene was detected at notable levels of expression (12% of the overall $2^{-\Delta Ct}$ value, Figure 3.3) from the J1 isolate. A switching event may have

taken place in the culture after ICAM-1 selection as RNA was extracted from this isolate 8 cycles after selection. Given that the parasite isolates in this study have been selected on ICAM-1 (Chapter 3), the DBLβ subtypes identified though this sequence analysis provide confidence that these dominantly expressed *var* genes are in fact mediating ICAM-1 binding. However, experimental validation is required to confirm and to provide more data to increase confidence in these sequence predictions.

Three of the genes, BC12a, J1b and PCM7a, contain the C-terminal domain tandem DBLδ1-CIDRβ1. No role for this tandem has been identified to date. Indeed, specific roles of individual C-terminal domain sequences have not been extensively described. Domains DBLζ3/6 and DBLε12 of J1d could have a known role in binding and pathogenesis as these domains have been identified as IgM binding domains, often accompanied by an upstream DBLγ domain. However, IgM binding cannot be predicted by sequence and therefore the subtype of DBLζ/ε domain (Pleass *et al.*, 2016). This would require experimental testing. This domain structure in combination with DBLα types 1 and 2 is specifically associated with IgM binding mediated rosetting (Ghumra *et al.*, 2012). The other C-terminal domains may act as extenders to allow the N-terminal domains freer access to receptors away from the IE membrane. Alternatively, these C-terminal domains may act in tandem to bind as yet unidentified receptors as is the case with DC5 (DBLγ12-DBLδ5-CIDRβ3/4-DBLβ7/9) mediating PECAM-1 binding, for which no individual domain was clearly identified as responsible (Berger *et al.*, 2013). Other receptors that have been suggested to bind PfEMP1 but for which no specific binding domains have been identified, include VCAM-1, E-selectin (Ockenhouse *et al.*, 1992b) and P-selectin (Senczuk *et al.*, 2001).

Note that the full extracellular domain sequence of only J1b and PCM7a was obtained. BC12a, J1a, J1d and PCM7d may encode additional C-terminal domains that we were unable to sequence by exon 2 PCR (see Results section 3.3.5).

4.3.2 *In silico* analysis of non-dominant and brain specific *var* genes

In the previous chapter, a number of DBLα tags were identified from the ICAM-1 binding patient isolates BC12, J1 and PCM7 by cloning but were not expressed at a high level when tested by RT-qPCR (Figure 3.1, Figure 3.3). We will refer to these tags as the non-dominant *var* genes of these isolates. We searched these DBLα tags against the Pf3k *var* gene database and analysed the top four hit sequences in an attempt to extend our knowledge of the *var* repertoires of these isolates and to further test the prediction capabilities of the Pf3k database. We also repeated the search for a hit to the brain-specific sequence 28B1-1 with the hope that, similarly to the case of PCM7a, there may now be a match in the ever expanding database. There were two hits to 28B1-1 upon the second search. One hit is very short and consists of only the NTS and an incomplete DBLα, whereas the second hit is much longer. This sequence, along with the PF0311-C.g26 hit to 62B1-1, was included in this *in silico* analysis in an attempt to understand their importance in the pathogenesis of cerebral malaria.

Pf3k database hits to non-dominant DBLα tags of ICAM-1 binding patient isolates

The non-dominant DBLα tags were searched against the database with parameters of at least 99% identity and 95% coverage (Thomas D. Otto). Hits were obtained for 4/5 BC12 tags, 3/4 J1 tags and 15/18 PCM7 tags. The absence of hits in the database to 5 of the tags highlights that sequencing of the global *var* repertoire is yet to be achieved. The domain structure and subtypes of the top four hits to each tag was identified, as above, using the VarDom 1.0 Server and the 7 genomes dataset (Rask *et al.*, 2010). Schematics of the domain structures are shown in Table 4.2.

Table 4.2. Schematic representation of the domain structures of Pf3k database hits to non-dominant and brain-specific *var* genes.

See pages 71-73 below.

Query_match	NTS	PfEMP1 domain structure	ATS
28B1-1_XX0072-C.g6	B3	DBLα0/1 CIDRα1.1 DBLβ12 DBLγ6/4 DBLδ7/1 CIDRβ6/5	
28B1-1_XX0144-C.g197	B3	DBLα2/1.2	
62B1-1_PF0311-C.g26	A7	DBLα1 CIDRα1.7 DBLβ3/11 DBLγ2 DBLγ DBLγ2/10 DBLζ3 DBLε6	
BC12b_XX0174-C.g180		DBLα0.6/22 CIDRα3.1/4	
BC12b_XX0252-C.g84	B3	DBLα0.6/22 CIDRα3.1/4 DBLβ5/6	
BC12b_XX0118-C.g195	B3	DBLα0.6/9	
BC12b_XX0109-C.g60	B3	DBLα0.6/22 CIDRα3.1/4	
BC12c_NONE			
BC12d_XX0155-C.g25	B3	DBLα0 CIDRα2.5/10 DBLβ13 DBLδ1 CIDRβ1	
BC12d_XX0152-C.g16	B3	DBLα0.6 CIDRα3.2/1 DBLβ4 DBLγ5/16 DBLδ1 CIDRβ1	B
BC12d_XX0151-C.g13	B3	DBLα0.6 CIDRα3.2/1 DBLβ5/4 DBLδ CIDRγ1 DBLγ DBLζ4	
BC12d_XX0146-C.g41	B3	DBLα0.6 CIDRα3.2/1 DBLβ4/5 DBLγ5/16 DBLδ1 CIDRβ1	
BC12e_XX0115-C.g41	B5	DBLα2/0.9 CIDRα1.8/1 DBLβ12 DBLγ DBLδ1 CIDRβ1/5	
BC12e_XX0338-C.g73	B4/5	DBLα2/0.9 CIDRα1.8/1 DBLβ12 DBLγ6/4	
BC12e_XX0488-C.g19	B5/4	DBLα2/0.9 CIDRα1.8/1 DBLβ12 DBLγ6/4 DBLζ2 DBLε14/5	
BC12e_XX0110-C.g19	B5/4	DBLα2/0.9 CIDRα1.8/1	
BC12f_XX0162-C.g243	B3	DBLα0	
BC12f_XX0161-C.g102	B3	DBLα0 CIDRα2/4	
BC12f_XX0153-C.g127	B3	DBLα0 CIDRα2/4	
BC12f_XX0148-C.g7		DBLα0 CIDRα2/4	
J1c_XX0286-C.g17	B4	DBLα0.12 CIDRα2.1 DBLβ2/5 DBLγ DBLε14/3	
J1c_XX0285-C.g37	B4	DBLα0.12 CIDRα2.1 DBLβ2/5 DBLγ DBLε14/3	
J1c_XX0098-C.g21	B4	DBLα0.12 CIDRα2.1 DBLβ5 DBLζ4	
J1c_XX0055-C.g103		DBLα0.12/9 CIDRα2.1 DBLβ5	
J1e_XX0028-Cx.g65	B3	DBLα0.13/22 CIDRα3.1	
J1f_XX0174-C.g127	B3	DBLα0.1 CIDRα3.2/1	
J1f_XX0163-C.g123	B3	DBLα0	
J1f_XX0160-C.g29	B3	DBLα0.1 CIDRα3.2/1 DBLδ4 CIDRγ1 DBLζ6 DBLε9	
J1f_XX0151-C.g70	B3	DBLα0.1 CIDRα3.2/1 DBLδ1 CIDRβ6	

Query_match	NTS	PfEMP1 domain structure								ATS
J1g_NONE										
PCM7b_XX0174-C.g127	B3	DBLα0.1	CIDRα3.2/1							
PCM7b_XX0163-C.g123	B3	DBLα0								
PCM7b_XX0160-C.g29	B3	DBLα0.1	CIDRα3.2/1	DBLδ4	CIDRγ1	DBLζ6	DBLε9			
PCM7b_XX0151-C.g70	B3	DBLα0.1	CIDRα3.2/1	DBLδ1	CIDRβ6					
PCM7c_XX0288-C.g18	B3	DBLα0	CIDRα3.1	DBLβ5/10	DBLγ4	DBLζ2/4	DBLε4			
PCM7c_XX0099-C.g26	B3	DBLα0	CIDRα3.1	DBLβ10/5	DBLγ18/5	DBLδ1	CIDRβ1			B2/1
PCM7c_XX0362-C.g64	B3	DBLα0	CIDRα3.1	DBLβ10/5	DBLγ18/5					
PCM7c_XX0209-C.g48	B3	DBLα0	CIDRα3.1	DBLβ5/10	DBLγ	DBLζ2/4	DBLε4			
PCM7e_XX0103-C.g97	B3	DBLα0.18	CIDRα3.1	DBLβ5/8						
PCM7f_XX0392-C.g48	B3	DBLα0.18	CIDRα3.1	DBLβ5	DBLδ1	CIDRβ1				
PCM7f_XX0007-C.g454	B3	DBLα0.18	CIDRα3.1	DBLβ5						
PCM7f_XX0005-C.g378	B3	DBLα0.18	CIDRα3.1	DBLβ5	DBLδ1	CIDRβ4/3				
PCM7f_XX0004-C.g448	B3	DBLα0.18	CIDRα3.1	DBLβ5						
PCM7g_XX0460-C.g63	B3	DBLα0	CIDRα3.1	DBLδ1	CIDRβ1					
PCM7g_XX0127-C.g34	B3	DBLα0	CIDRα3.1	DBLδ1	CIDRβ1					
PCM7h_XX0155-C.g1	A1	DBLα1.1	CIDRα1.2	DBLβ11	DBLγ1	DBLε1	DBLγ8	DBLζ1	DBLε5	
PCM7h_XX0139-C.g32	A1	DBLα1.1	CIDRα1.2	DBLβ11	DBLγ1	DBLε1	DBLγ8	DBLζ1	DBLε5	
PCM7h_XX0360-C.g5	A1	DBLα1.1	CIDRα1.2	DBLβ11	DBLγ1	DBLε1	DBLγ8	DBLζ1	DBLε5	
PCM7h_XX0203-C.g5	A1	DBLα1.1	CIDRα1.2	DBLβ11	DBLγ1	DBLε1	DBLγ8	DBLζ1	DBLε5	var1
PCM7i_XX0089-C.g50	B3	DBLα0	CIDRα2.10/3	DBLδ1	CIDRγ4/10					
PCM7j_NONE										
PCM7k_XX0339-C.g87	B3	DBLα0.9/11	CIDRα2.4/1							
PCM7k_XX0485-C.g310		DBLα0.9/11	CIDRα2.4/1							
PCM7k_XX0480-C.g52	B3	DBLα0.9/11	CIDRα2.4/1	DBLδ1	CIDRβ1					
PCM7k_XX0475-C.g33	B3	DBLα0.9/11	CIDRα2.4/1	DBLδ1	CIDRβ1					
PCM7l_XX0248-C.g61	B3	DBLα0.11	CIDRα2.4	DBLβ8/5						
PCM7m_NONE										

Query_match	NTS	PfEMP1 domain structure						ATS
PCM7n_XX0172-C.g4	B3	DBLα0.9/6	CIDRα3.2/1	DBLβ8/12	DBLγ	DBLδ1	CIDRβ1	
PCM7n_XX0582-C.g107	B3	DBLα0.9/6						
PCM7n_XX0558-C.g38	B3	DBLα0.6/22	CIDRα3.2/1					
PCM7n_XX0548-C.g226	B3	DBLα0.6						
PCM7o_NONE								
PCM7p_XX0774-C.g285	B3	DBLα0						
PCM7p_XX0507-C.g52	B3	DBLα0	CIDRα2.6	DBLβ5/12	DBLδ1	CIDRβ6/1		
PCM7p_XX0496-C.g117	B3	DBLα0	CIDRα2.6	DBLβ5/12				
PCM7p_XX0489-C.g216	B3	DBLα0						
PCM7q_XX0154-C.g74	B3	DBLα0	CIDRα2.3/10					
PCM7q_XX0152-C.g24	B3	DBLα0	CIDRα2.3/10	DBLβ13/8	DBLδ1	CIDRβ1		B
PCM7q_XX0150-C.g243	B3	DBLα0						
PCM7q_XX0149-C.g37	B3	DBLα0	CIDRα2.3/10	DBLβ13/8	DBLδ1	CIDRβ1		
PCM7r_XX0356-C.g150	B3	DBLα0.13/5						
PCM7r_XX0099-C.g174	B3	DBLα0.13/5	CIDRα2.3					
PCM7r_XX0074-C.g17	B3	DBLα0.13/5	CIDRα2.3/10	DBLδ1	CIDRβ1			
PCM7r_XX0058-Cx.g64	B3	DBLα0.13/5	CIDRα2.3/10	DBLδ1	CIDRβ1			
PCM7s_XX0002-CW.g76	B3	DBLα0.20/1	CIDRα3.1/2	DBLδ1	CIDRγ			
PCM7s_XX0634-C.g83	B3	DBLα0.20	CIDRα3.1/2					
PCM7s_XX0582-C.g37	B3	DBLα0.20						
PCM7s_XX0558-C.g116	B3	DBLα0.20	CIDRα3.1/2	DBLδ1				
PCM7t_XX0172-C.g51	B3	DBLα0.17/1	CIDRα3	DBLδ1	CIDRβ1			
PCM7t_XX0139-C.g66	B3	DBLα0.17/1	CIDRα3	DBLδ1	CIDRβ1			
PCM7t_XX0138-C.g4200	B3	DBLα0.17/1						
PCM7t_XX0137-C.g207	B3	DBLα0.17/1						

The length of the hits was variable with the sequences of two tags, BC12f and J1e, encoding only the NTS-DBLα-CIDRα head structure. Of the non-dominant hits, only 4 sequences include the ATS region indicating the complete extracellular domain structure is present. The remaining hits are therefore partial sequences and cannot provide information on gene length which has previously been linked to sequence conservation (Buckee and Recker, 2012). Sixteen of the 22 non-dominant tags which had database hits had one apparent PfEMP1 domain structure (Table 4.2). However, of these, 4 had only one database hit (J1e, PCM7e, PCM7i and PCM7l). Five of the non-dominant tags had hits with 2 apparent domain structures with the differences occurring downstream of the NTS-DBLα-CIDRα head structure (BC12e, J1c, J1f, PCM7b and PCM7c). One tag had 3 different domain structures apparent in the database hits, two of which differed after the NTS-DBLα-CIDRα-DBLβ domains and one which differed downstream of the NTS-DBLα domains (BC12d, Table 4.2). The presence of genes in the database which share a DBLα tag but which have different PfEMP1 domain compositions reduces the reliability of the database as a predictive power and highlights the need for experimental validation of these predictions. This is perhaps unsurprising considering the constant recombination, both mitotic and meiotic, which occurs between *var* genes (Bopp *et al.*, 2013, Claessens *et al.*, 2014, Ranford-Cartwright and Mwangi, 2012).

The only DC present in the hits to the non-dominant tags is DC1 which is present in all hits analysed to PCM7h indicating this gene is the conserved VAR1 gene. Hits to BC12e have a domain structure similar to DC8 (BC12e: DBLα2/0.9-CIDRα1.8/1-DBLβ12-DBLγ6/4, DC8: DBLα2-CIDRα1.1-DBLβ12-DBLγ4/6) and are the only non-dominant tag hits to contain a predicted EPCR binding domain, CIDRα1.8/1 (Turner *et al.*, 2013, Lau *et al.*, 2015). The remaining sequences, with the exception of PCM7h VAR1 hits, contain CIDRα2-4 which are predicted to bind CD36 (Robinson *et al.*, 2003, Turner *et al.*, 2013).

Pf3k database hits to brain-specific *var* gene sequences 28B1-1 and 62B1-1

Two sequences were recently identified as brain-specific *var* genes in several cases of fatal cerebral malaria, 28B1-1 and 62B1-1 (Tembo *et al.*, 2014). In the previous chapter we attempted to reconstruct the 62B1-1 sequence from gDNA extracted

from tissue samples using primers to the Pf3k *var* gene database hit, PF0311-C.g26, to no avail (see Results section 3.4). Here, we identify the PfEMP1 domain structure of PF0311-C.g26 along with the newly identified XX0072-C.g6 database hit to 28B1-1 (Table 4.2) and analyse their binding predictions.

62B1-1_PF0311-C.g26 contains NTSA7, a class that is exclusively found in UPSA type *var* genes and contains DBLα1 which is also associated with UPS type A (Rask *et al.*, 2010). 28B1-1_XX0072-C.g6 contains NTSB3 which could be either UPS B or C type but has DBLα0/1 and so had BLAST matches of both type 0 and 1 suggesting a possible chimera (Rask *et al.*, 2010). Both genes contain CIDRα types that are predicted to bind EPCR (1.1 and 1.7) (Turner *et al.*, 2013, Lau *et al.*, 2015), a receptor that has recently been identified as the receptor for DC8 and DC13 (Turner *et al.*, 2013), cassettes which were found to be associated with severe malaria (Lavstsen *et al.*, 2012). 62B1-1_PF0311-C.g26 is unusual in that it contains three consecutive DBLγ domains. The function of these domains is unknown. The DBLζ3-DBLε6 domains of 62B1-1_PF0311-C.g26 may have the ability to bind IgM which is mediated by these domain types but cannot be predicted by sequence (Pleass *et al.*, 2016). Both genes are large with 62B1-1_PF0311-C.g26 containing 8 extracellular domains and 28B1-1_XX0072-C.g6 containing 6 extracellular domains. Neither hit includes an ATS region indicating they are not full length sequences.

The sequence predictions for these brain-specific *var* genes provide insight into possible mechanisms for disease pathogenesis. Both sequences are predicted to bind EPCR, an endothelial receptor involved in anti-coagulation and anti-inflammatory pathways (Gleeson *et al.*, 2012). It is thought that EPCR binding blocks the native function of this receptor, resulting in loss of cytoprotective effects (Turner *et al.*, 2013). In addition to EPCR binding, 62B1-1_PF0311-C.g26 has the structure of a rosette mediating *var* gene. Specifically, it contains a DBLα1 domain which has been shown to bind to uninfected erythrocytes via complement receptor 1 (CR1) (Rowe *et al.*, 1997, Rowe *et al.*, 2000), blood group A and B trisaccharides (Vigan-Womas *et al.*, 2012, Barragan *et al.*, 2000) and heparin sulphate (Carlson and Wahlgren, 1992, Vogt *et al.*, 2003). It also contains the C-terminal domain structure of IgM binding *var* genes (Ghumra *et al.*, 2012) which is thought to strengthen

interactions between infected and uninfected erythrocytes (Clough *et al.*, 1998, Somner *et al.*, 2000). However, these are predictions which could not be experimentally validated with the samples available. In depth study of the Pf3k *var* gene database will provide more information on *var* gene structure and sequence relationships and will likely identify new domain cassettes which could be useful in picking apart the pathogenic mechanisms observed experimentally.

4.3.3 DBLα tag analysis by cysteine/Positions of Limited Variability (cys/PoLV)

Many studies, including the present one (see Results section 3.3.1 and Discussion in 3.5), use the DBLα tag to identify the highly variable *var* genes within a population. This strategy is effective due to (i) all *var* genes, except the unique *var2csa* (Rask *et al.*, 2010), containing an N-terminal DBLα domain, (ii) suspected universal primers designed to relatively conserved homology blocks within the DBLα domain (Bull *et al.*, 2005), and (iii) the short length of the DBLα tag which can easily be cloned in large numbers from a single PCR per sample. The short DBLα tag has been classified using the number of cysteine residues present, positions of limited variability (PoLV) and length to generate distinct sequence identifiers which are reported to provide relevant biological information (Bull *et al.*, 2005, Bull *et al.*, 2007). We performed this analysis on the dominantly expressed DBLα tags of ICAM-1 binding patient isolates, the non-dominant tags from these isolates and the DBLα tag sequences of the brain-specific *var* genes 28B1-1 and 62B1-1.

DBLα tag sequences were checked for N-terminal DIGDI and C-terminal PQFLR motifs before analysing with a Perl script (kindly provided by Peter Bull, (Bull *et al.*, 2007)) which uses 4 anchor points within the sequence to identify the number of cysteines and PoLV1-4 (see Methods 4.2.2). These are described as belonging to one of six groups which are defined as follows: Group 1 – cys2, PoLV1 (MFK*); Group 2 – cys2, PoLV2 (*REY); Group 3 – cys2 not groups 1 or 2; Group 4 – cys4 not group 5; Group 5 – cys4, PoLV2 (*REY); Group 6 – cys1, 3, 5, >5 (Bull *et al.*, 2005). These groupings are based on findings that cys2, PoLV1 MFK* and PoLV2 *REY are

all independently associated with short DBLα tag length and that MFK* and *REY are mutually exclusive (Bull *et al.*, 2005).

All of the dominantly expressed *var* genes of the ICAM-1 binding patient isolates belong to PoLV Group 4, i.e. they are all cys4 and do not contain the *REY motif at PoLV2 (Table 4.3). The UPS type of each gene, as identified in the previous chapter, is included in Table 4.3 and is in keeping with observations that Group 4 DBLα tags are associated with group B/A, B, and B/C *var* genes but never group A in the reference parasite strains 3D7, HB3 and IT4 (Bull *et al.*, 2005, Bull *et al.*, 2007).

Table 4.3. DBLα tag cys/PoLV classification and UPS type of dominantly expressed *var* genes.

DBLα tag	PoLV group	Distinct ID sequence	UPS type
BC12a	4	LYIR-LRED-KAIT-4-PTYF-121	B
J1a	4	LFLG-LRND-NAIT-4-LTNF-119	C
J1b	4	LYLG-LRND-NAIT-4-LTNF-119	-
J1d	4	LYFG-LRED-KAIT-4-PTYF-124	B
PCM7a	4	LYRG-LRED-KAIT-4-PTYF-122	B
PCM7d	4	LYRG-LRED-KAII-4-PTNF-117	B

Groups 1-6 are all represented in the non-dominant and brain-specific *var* genes. 68B1-1 and PCM7h DBLα tags are the only PoLV Group 1 tags which are exclusive to group A *var* genes and those encoding more than 5 extracellular domains (Bull *et al.*, 2007). This is in agreement with the Pf3k database hits to these sequences predicting them to be UPS A type *var* genes, each with 8 domains (Table 4.2). The only tags belonging to Group 2 are 28B1-1 and BC12e. These are the only tags, other than 62B1-1, whose hits contain predicted EPCR binding domains. BC12e hits are those whose domain structure is similar to DC8 which has been associated with severe malaria (Lavstsen *et al.*, 2012). Group 2 tags are associated with rosetting (Bull *et al.*, 2005), another phenotype linked to disease pathogenesis. More generally, cys2 tags are found more frequently in CM than hyperparasitaemic patients (Kyriacou *et al.*, 2006) and are also found to be independently associated

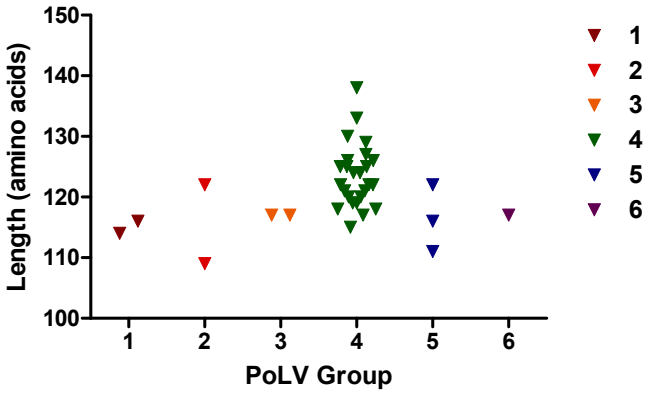
Table 4.4. DBLα tag cys/PoLV classification of non-dominant and brain-specific *var* genes.

DBLα tag	PoLV group	Distinct ID sequence
28B1-1	2	LYLD-VREY-KAIT-2-VPTN-109
62B1-1	1	MFKS-LRND-KAIT-2-PTYL-114
BC12b	4	MFKS-LREA-EAIT-4-PTYF-115
BC12c	4	LYLG-LRED-RAIT-4-PTYF-120
BC12d	4	LFYG-LRED-RAIT-4-PTYF-130
BC12e	2	LYLG-LREY-KAIT-2-QTYL-122
BC12f	4	LYLG-LRED-KAIT-4-PTYF-127
J1c	4	LFIG-LRED-KAIT-4-PTYF-126
J1e	4	LYLG-LRED-KAIT-4-PTYF-133
J1f	3	LYRR-LRED-KAIT-2-PTNL-117
J1g	4	LYRR-LRED-KAIT-4-PTYF-118
PCM7b	3	LYRR-LRED-KAIT-2-PTNL-117
PCM7c	5	LFLG-IREY-KAIT-4-PTNF-111
PCM7e	4	LYLG-LRED-KAIT-4-PTNL-120
PCM7f	4	LYLG-LRED-KAMI-4-PTNL-122
PCM7g	4	LYSG-LRED-KAIT-4-PTYF-125
PCM7h	1	MFKP-LREA-EAIT-2-PTNL-116
PCM7i	4	LYRG-LRED-KAIT-4-PTYF-124
PCM7j	4	LYRG-LRED-KAIT-4-PTYF-125
PCM7k	4	LHLG-LRED-KAMT-4-PTYF-122
PCM7l	4	LYLG-LRED-KAIT-4-PTYF-125
PCM7m	5	LFLG-IREY-KAIT-4-PTYF-116
PCM7n	4	LYRG-LRED-EAIT-4-PTYF-118
PCM7o	4	LFLG-LRED-KALT-4-PTYF-129
PCM7p	4	LFLG-LRED-KALT-4-PTYF-121
PCM7q	4	LFLG-LRED-KAIT-4-PTYF-126
PCM7r	5	IFRG-VREY-KALT-4-PTYF-122
PCM7s	6	LFLG-VREY-KAIT-5-PTYF-117
PCM7t	4	LYIG-LRED-NAIT-4-PTYF-138

with SM, young host age and parasite pathogenicity in multiple linear regression models (Warimwe *et al.*, 2009). Two additional cys2 tags were identified as belonging to Group 3, J1f and PCM7b, which upon closer inspection actually share identical database hits (Table 4.2, Table 4.4). A sequence alignment of these tags shows only two amino acid differences between them (not shown).

The remaining non-dominant DBLα tags are mainly PoLV Group 4 (N=19) and just three belonging to Group 5 (cys4, PoLV *REY) and one to Group 6 which has 5 cysteine residues (Table 4.4). The original reasoning behind this grouping according to tag length was that if sequences are recombining within a restricted subset, then the length of the sequence will stay similar after multiple recombination events, suggesting that the different functions of these distinct groups are desirable to be maintained within the *var* repertoire (Bull *et al.*, 2007). We compared the length of the tags from each Group in Figure 4.1A with the Group definitions included in Figure 4.1B. This generally follows the patterns observed previously (Bull *et al.*, 2005, Bull *et al.*, 2007) with cys2 tags and those containing either PoLV1 MFK* or PoLV2 *REY (Groups 1-3 and 5) having shorter tag sequence than Groups 4 and 6 (Figure 4.1). However, there are not enough sequences in our comparison to show meaningful associations. It will be interesting to perform this analysis on the entire Pf3k *var* gene database when it becomes publically available. This will not only provide many DBLα tags (which many studies have analysed, including those studies mentioned above (Bull *et al.*, 2007, Warimwe *et al.*, 2009)) but will also include longer sequence information, including domain structures, and sequences from a worldwide distribution, thus allowing a more comprehensive analysis.

A



B

Position of Limited Variability
Group 1: cys2, PoLV1 (MFK*)
Group 2: cys2, PoLV2 (*REY)
Group 3: cys2, not groups 1, 2
Group 4: cys4, not group 5
Group 5: cys4, PoLV2 (*REY)
Group 6: cys1, 3, 5, > 5

Figure 4.1. DBLα tag length comparison of PoLV Groups.
A, The cys/PoLV Group of all DBLα tags analysed is plotted against tag length. **B**, PoLV Group 1-6 definitions (Bull *et al.*, 2005).

The cys/PoLV method of sequence classification has been shown to predict some aspects of disease phenotypes which is remarkable for such a short sequence (350-400 bp) within a much larger gene (3.5-9 kb, exon 1). Here we have used two predictive methods: the Pf3k database search and the cys/PoLV classification of DBLα tags. The methods are in agreement in relation to *var* genes that are predicted to be associated with severe malaria, as discussed above. This is likely correct in relation to the 28B1-1 and 62B1-1 sequences important in cerebral malaria but we have no indication of the role of the non-dominant *var* genes as these were identified from the culture-adapted patient isolates after selection on ICAM-1. Therefore, their function *in vivo* is unknown. Classification methods to date can generally only predict *var* genes important in severe malaria which can aid efforts to produce anti-disease therapies. However, it is often assumed that the role

of those not associated with severe disease is less significant. In the absence of sterilising immunity (McGregor, 1964) parasites causing mild and asymptomatic malaria persist in the human population. Further study of *var* genes expressed in these infections will be required to understand the role of parasite sequestration in these patients. This may become increasingly important if we are to progress towards malaria elimination goals (Hemingway *et al.*, 2016).

4.3.4 Phylogenetic analysis of DBLβ domains predicted to bind ICAM-1

ICAM-1 binding by PfEMP1 proteins is mediated by a subset of DBLβ domains, namely all DBLβ5 domains tested, three DBLβ3 domains and one DBLβ1 domain have been identified to date (Howell *et al.*, 2008, Janes *et al.*, 2011, Bengtsson *et al.*, 2013b). We analysed the newly identified DBLβ domains of the ICAM-1 binding patient isolates in more detail by comparing them to known DBLβ sequences using the maximum likelihood method. A comparison with DBLβ domains of all subtypes identified in the Rask *et al* (2010) dataset shows that the dominantly expressed DBLβ domains from BC12a, J1a, J1d, PCM7a and PCM7d group relatively closely together but do not form a distinct cluster (Figure 4.2, red). Known ICAM-1 binding DBLβ domains (Howell *et al.*, 2008, Janes *et al.*, 2011, Bengtsson *et al.*, 2013b) are highlighted in orange to show their wide distribution. However, note that bootstrap support of ≥ 70% is present in relatively few nodes (black dots, Figure 4.2). Lower values indicate the unreliability of the groupings which reflects the recombination that we know to occur between the DBLβ domains (Bopp *et al.*, 2013, Claessens *et al.*, 2014). The DBLβ domains encoded by the Pf3k database hits to the non-dominant genes of these patient isolates are included.

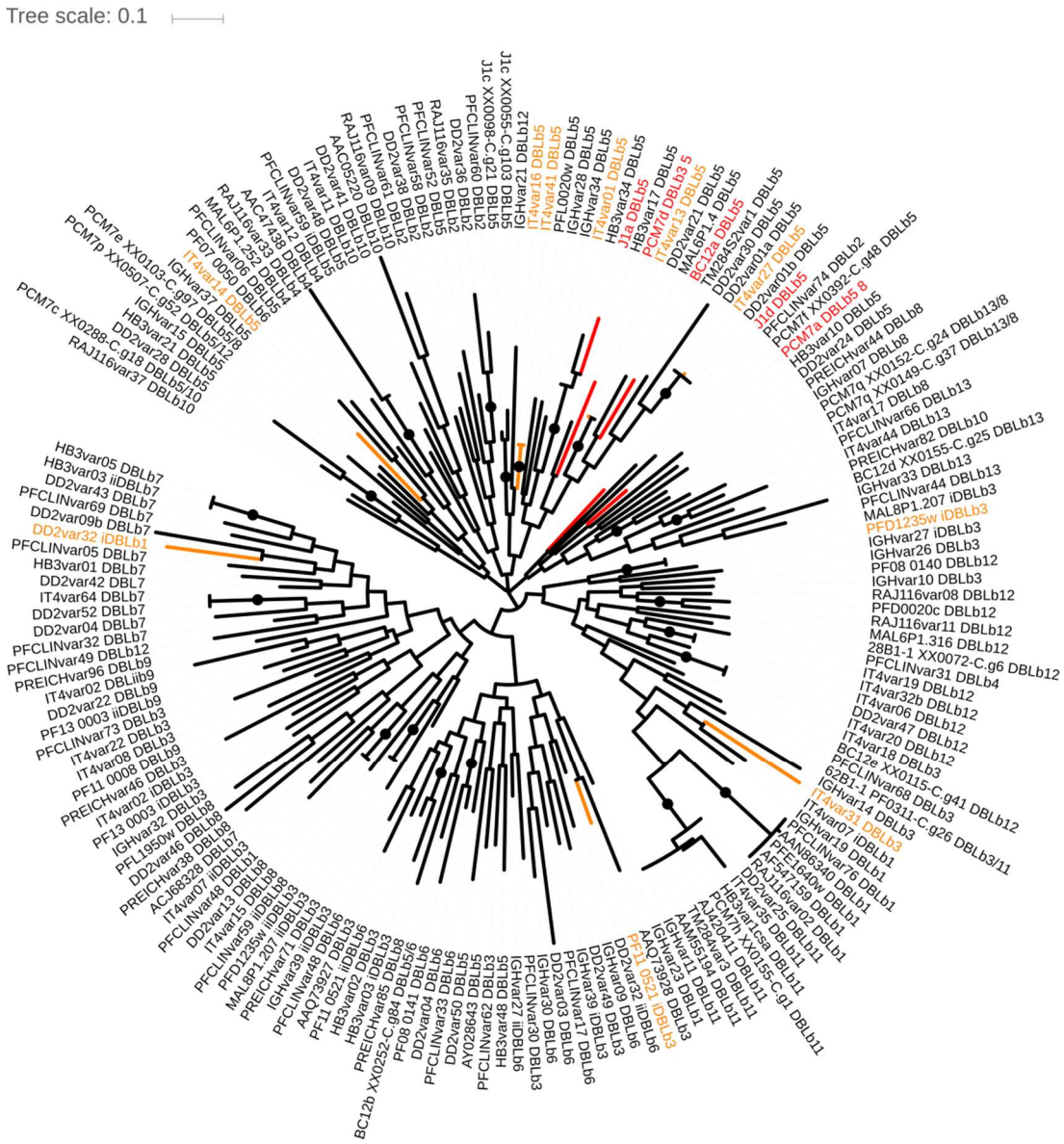


Figure 4.2. Phylogenetic comparison of all DBLβ domains in the VarDom 1.0 Server and BC12, J1 and PCM7 DBLβ domains.

A maximum likelihood tree was constructed from all DBLβ domains in the Rask *et al.* (2010) dataset, accessed via the VarDom 1.0 Server, and the DBLβ domains of the ICAM-1 binding patient isolates BC12, J1 and PCM7. Orange: known ICAM-1 binding DBLβ domains. Red: dominantly expressed *var* genes of BC12, J1 and PCM7. Black dots indicate bootstrap values ≥ 70%.

The dominantly expressed *var* genes of the ICAM-1 binding patient isolates encode DBLβ domains of subtype 5 (BC12a, J1a, J1d), 5/8 (PCM7a) and 3/5 (PCM7d). We therefore constructed a maximum likelihood tree of all DBLβ3, 5 and 8 domains

along with our newly identified DBLβ sequences (Figure 4.3). The majority of β3 sequences cluster together (purple) and are of UPS A type. Two exceptions (IGHvar10 and PFCLINvar68) fall within a cluster of β5 sequences (orange) of mainly UPS B type. PCM7d (DBLβ3/5) falls within the DBLβ3 group suggesting it is more similar to DBLβ3. DBLβ8 sequences (turquoise) form two clades in the midst of the DBLβ5 sequences. PCM7a (DBLβ5/8) falls in the middle of these two clades with the HB3var10 DBLβ5 sequence (Figure 4.3), suggesting it was accurately classified by BLAST analysis as falling between two types. The majority of DBLβ5 sequences are of UPS B type with a small cluster of UPS C sequences which include J1a (Figure 4.3). J1d forms a single branch in between the DBLβ3 clade and the other sequences, suggesting it is more distantly related to the other DBLβ5 sequences.

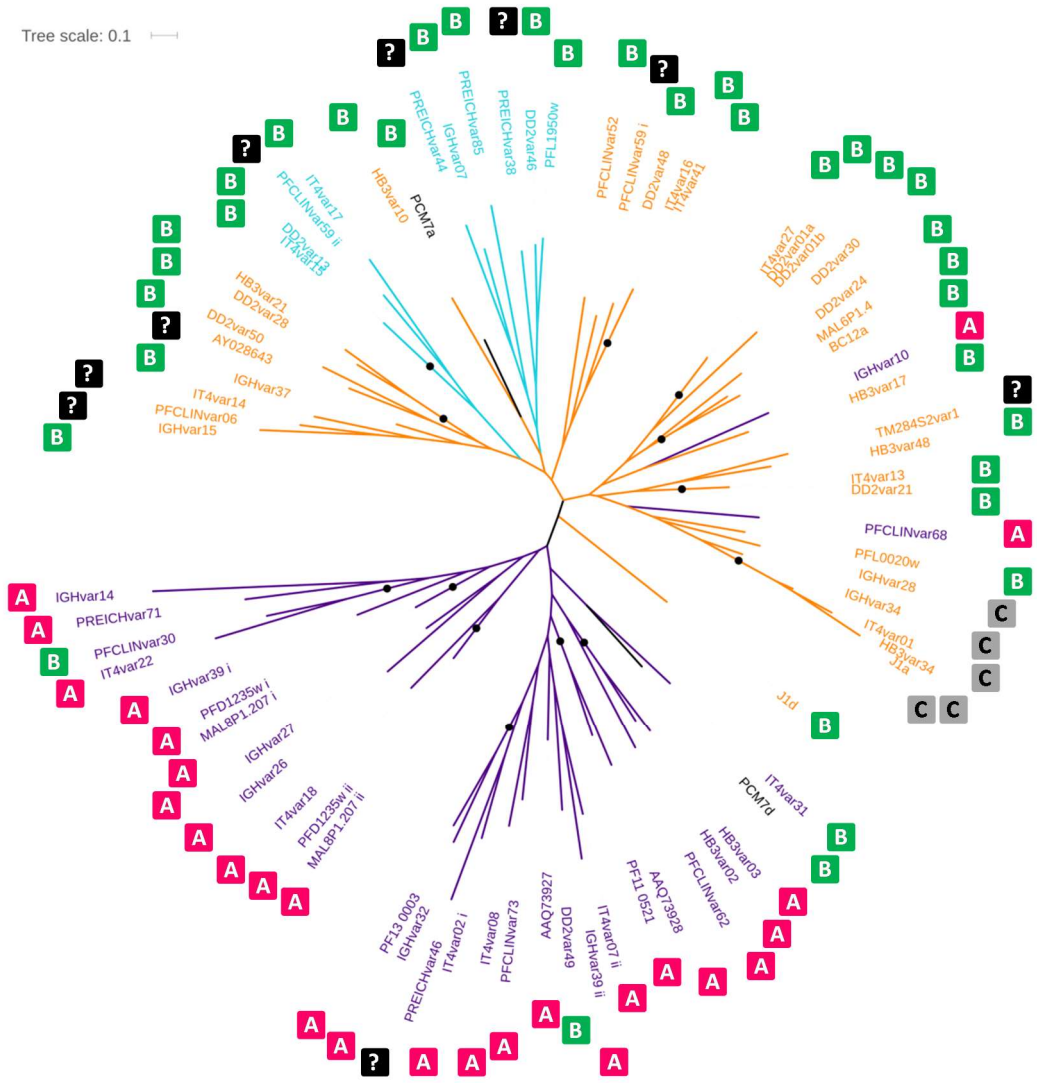


Figure 4.3. Phylogenetic comparison of type 3, 5 and 8 DBLβ domains.
A maximum likelihood tree was constructed from all DBLβ domains of subtype 3, 5 and 8 in the Rask *et al.* (2010) dataset and the DBLβ domains of the dominantly expressed *var* genes of the ICAM-1 binding patient isolates BC12, J1 and PCM7. Purple: DBLβ3, orange: DBLβ5, turquoise: DBLβ8, black: PCM7a and PCM7d which have similarities to two subtypes. The UPS type A-C or unknown (?) is indicated. Black dots indicate bootstrap values ≥ 70%.

The tree displayed in Figure 4.3 was re-analysed in relation to ICAM-1 binding, with ICAM-1 binding DBLβ in red, non-binders in blue and untested DBLβ domains in black (Figure 4.4). The sequences identified in this thesis are also in red as they are assumed to be responsible for ICAM-1 binding. There are seven DBLβ5 sequences that bind ICAM-1, ten if BC12a, J1a and J1d are included, and they are spread across

the DBLβ5 containing branches of the tree (Figure 4.4). Only IT4var16 and IT4var41 group very closely together. J1a groups with two other UPS C sequences of DBLβ5 type with a bootstrap value of 100% (indicated by black dot, Figure 4.4). One of these sequences is the ICAM-1 binder IT4var01 and the other is the untested sequence of HB3var34. It is possible that these very closely related sequences all have the ability to bind ICAM-1 specifically. However, the binding ability of HB3var34 would need to be tested to confirm this. There are three DBLβ3 ICAM-1 binders, two of which are of UPS A type and one UPS B. PCM7d groups with the UPS B IT4var31 DBLβ3 ICAM-1 binding domain (Figure 4.4). As described above, PCM7a falls between two DBLβ8 clusters, none of which are known to bind ICAM-1, although only two have been tested. Indeed, only 19 out of the 70 domains included in this tree have been tested for binding to ICAM-1 and the majority of those belong to the IT4 parasite strain.

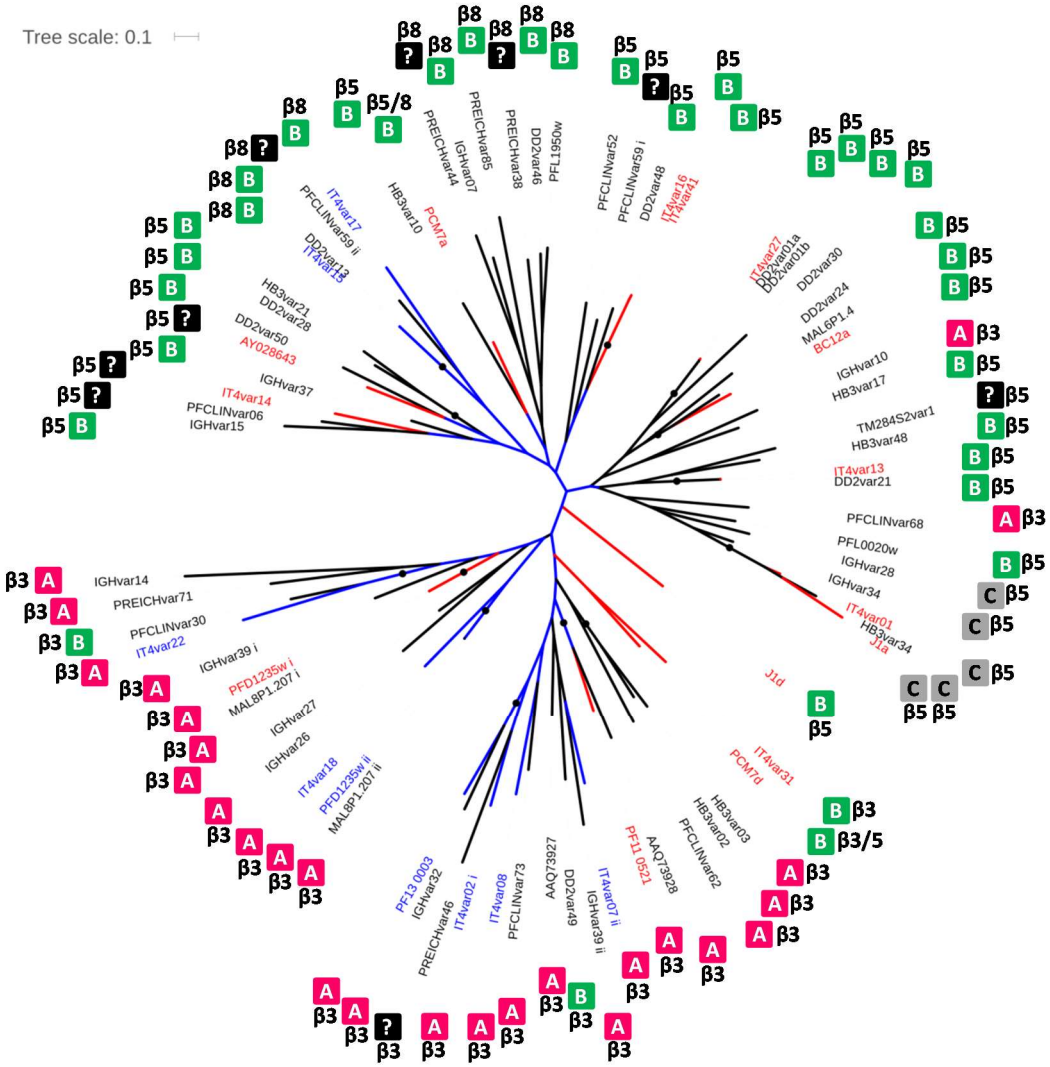


Figure 4.4. Phylogenetic comparison of type 3, 5 and 8 DBLβ domains in relation to ICAM-1 binding.

The maximum likelihood tree shown in Figure 4.3 (constructed from all DBLβ domains of subtype 3, 5 and 8 in the Rask *et al.* (2010) dataset and the DBLβ domains of the dominantly expressed *var* genes of the ICAM-1 binding patient isolates BC12, J1 and PCM7). DBLβ domains that bind ICAM-1 (red), do not bind ICAM-1 (blue) and have not been tested for ICAM-1 binding (black) are indicated. The UPS type (A-C or unknown (?)) and DBLβ subtype is indicated. Black dots indicate bootstrap values ≥ 70%.

Since the majority of work on ICAM-1 binding has been carried out on the IT4 parasite strain, we constructed a maximum likelihood tree of our DBLβ domains and all those from the IT4 sequences (Figure 4.5). The majority of the ICAM-1 binding DBLβ domains fall on one side of the tree, forming two clades. J1d groups

with an untested sequence, IT4var11, and PCM7a branches with two non-binders but is close to ICAM-1 binders IT4var13 and IT4var27. IT4var31 and PCM7d form a separate clade with a bootstrap value of 100%. As in Figure 4.4, J1a and IT4var01 group together with a bootstrap value of 100%. BC12a groups with IT4var27 with a bootstrap value of 83%. The relatedness of BC12a, J1a and PCM7d to known ICAM-1 binding DBLβ domains could be interpreted as further evidence that these domains are truly mediating ICAM-1 binding. This could also suggest that the untested sequences grouped with these ICAM-1 binding domains in Figure 4.4 can also mediate binding to ICAM-1. However, the fact that these sequences are known to recombine (Claessens *et al.*, 2014, Bopp *et al.*, 2013) weakens the reliability of phylogenetic analysis.

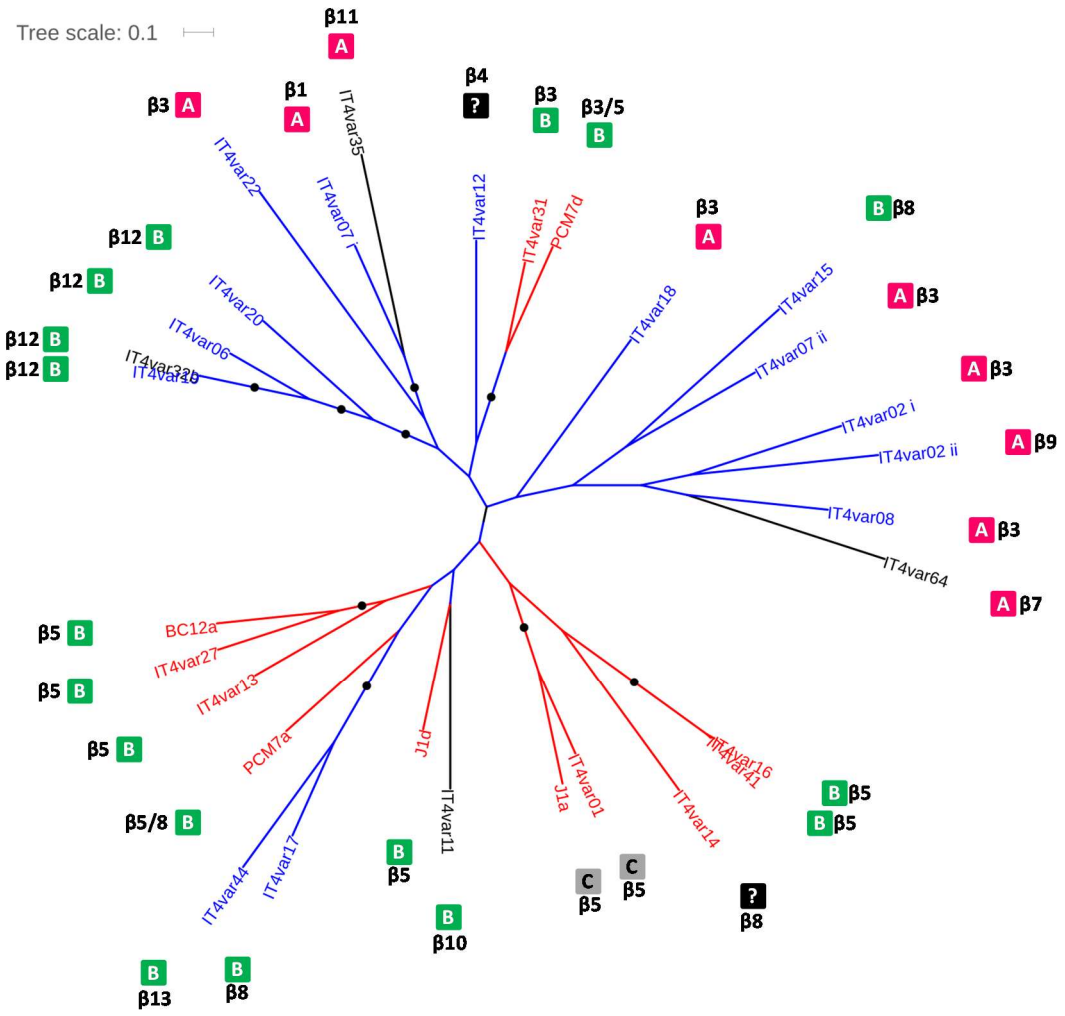


Figure 4.5. Phylogenetic comparison of IT4 and BC12, J1 and PCM7 DBLβ domains. A maximum likelihood tree constructed from all IT4 DBLβ domains and those of the dominantly expressed *var* genes of the ICAM-1 binding patient isolates BC12, J1 and PCM7. DBLβ domains that bind ICAM-1 (red), do not bind ICAM-1 (blue) and have not been tested for ICAM-1 binding (black) are indicated. The UPS type (A-C or unknown (?)) and DBLβ subtype is indicated. Black dots indicate bootstrap values ≥ 70%.

The results of our phylogenetic comparisons of the DBLβ domains show that there are some groupings of sequences known to mediate ICAM-1 binding. However, it is clear that there are also a number of exceptions which reduces the reliability of using sequence comparisons as predictors of domain function. Experimental validation is still necessary to test any predictions from sequence comparisons. In addition, the sequence of the entire DBLβ domain is compared when it may be

more informative to focus on key areas of the sequence such as homology blocks (Rask *et al.*, 2010) or areas specifically involved in contact with ICAM-1 (Lennartz *et al.*, 2015). These comparisons are addressed below.

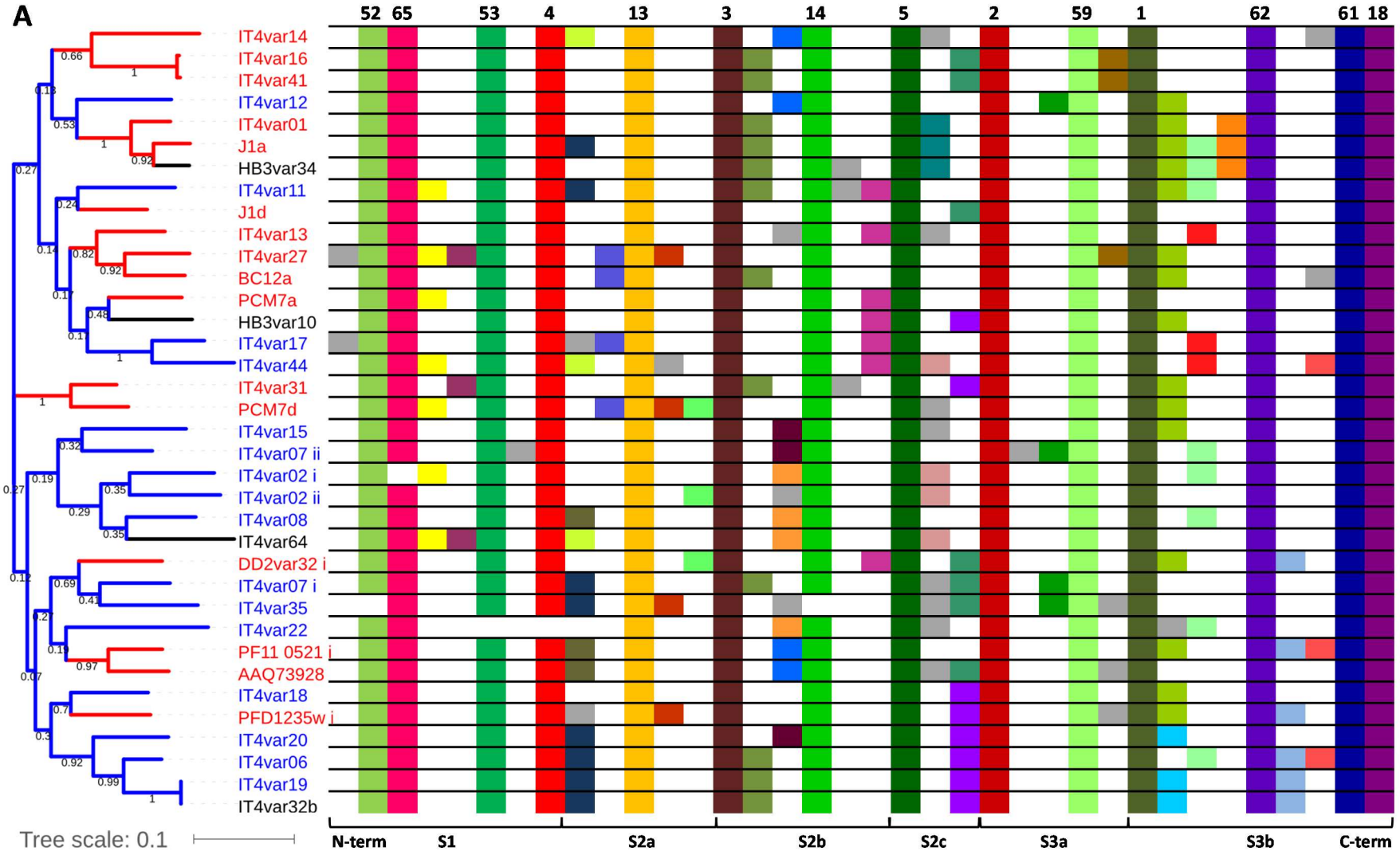
4.3.5 Homology block analysis of DBLβ domains predicted to bind ICAM-1

One downside of studying the relatedness of entire domain sequences is that small conserved sequences may be missed if the overall sequence groups elsewhere. This, along with the high frequency of recombination of the *var* genes, was the reasoning behind the definition of short sequence homology blocks (HBs) by Rask *et al.* (2010). We analysed the DBLβ domain HB structure of the ICAM-1 binding patient isolates, the IT4 *var* genes, and the individual *var* genes PF11_0521, PFD1235w, DD2var32, HB3var10, HB3var34 and AAQ73926, which either bind ICAM-1 or grouped closely with ICAM-1 binding sequences, in an attempt to identify shared HBs that indicate ICAM-1 binding ability.

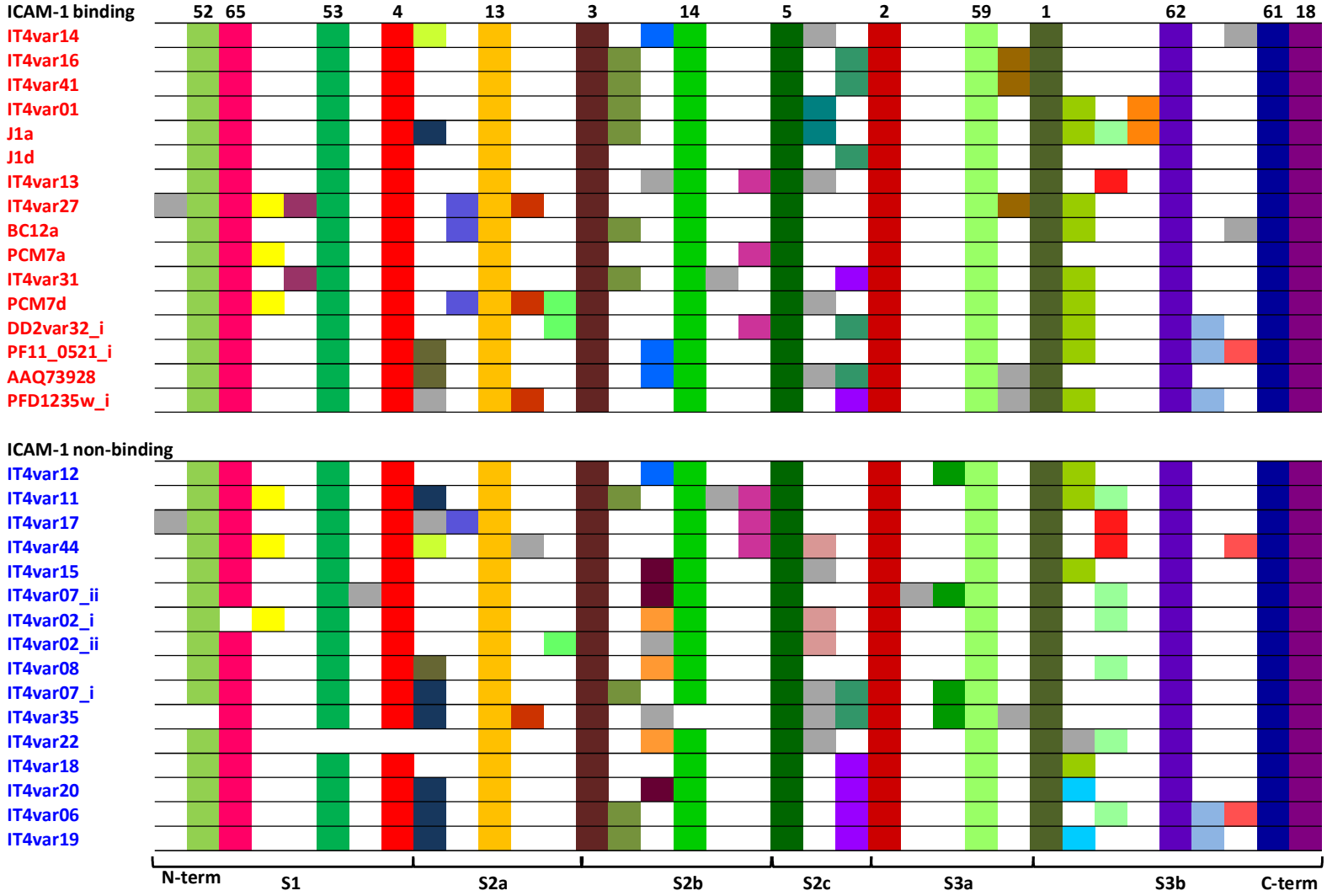
Figure 4.6A shows a maximum likelihood tree of these DBLβ domains and their HB composition. Note the low bootstrap values of many of the branches, indicating the unreliability of the groupings which reflects the recombination that occurs between the DBLβ domains (Claessens *et al.*, 2014, Bopp *et al.*, 2013). The core DBL HBs 1-5 are found in all the DBLβ domains with the exception of IT4var22 which does not contain HB4. HB13, which is found in all DBLα, β, and ζ domains, and HB18, which is found in all DBLβ and ζ domains (Rask *et al.*, 2010), are present in all domains (Figure 4.6). DBLβ specific HB53 and HB61 are also present in all domains except IT4var22 which does not contain HB53. The close proximity of HB53 to HB4, which is also missing in IT4var22, suggests that this section of IT4var22 DBLβ is significantly different from the other domains, likely as a result of a recombination event. HB59 is present in all the domains and HB52, HB65 and HB14 are present in all but one of the DBLβ domains (HB52 and HB14 absent from IT4var35 and HB53 absent from IT4var22). With the exception of the HBs mentioned above, no clear pattern emerges that could suggest specific HBs in ICAM-1 binding DBLβ domains. Only very closely related sequences share a HB pattern: IT4var16 identical to IT4var41 and IT4var19 identical to IT4var32b.

We separated the DBLβ domain HB compositions into ICAM-1 binding and non-binding groups in an attempt to see any significant pattern (Figure 4.6B). The majority of HBs appear in both groups with some exceptions. HB231 and HB408 appear only in the ICAM-1 binding DBLβ domains from IT4var01 and J1a. HB375 and HB382 also only appear in ICAM-1 binding domains. Four HBs are only found in non-binding DBLβ domains in this dataset: HB317, HB272, HB48 and HB614. However, these 'specific' HBs are only found in 2-4 domains and there are no clear associations between ICAM-1 binding and non-binding DBLβ domains (Figure 4.6B).

This lack of association is perhaps unsurprising when you consider that these defined HBs would need to include the specific contact residues necessary for ICAM-1 binding. The HB analysis was performed on linear sequences and HBs have an average length of 19 resides with the most common length 10 residues (Rask *et al.*, 2010). However, protein structures are not linear and binding sites can be formed from a number of secondary structures which may be some distance away in terms of primary structure. However, a recent study found that some HBs of the DBLα domain had a stronger association with certain disease phenotypes than the commonly used DBLα tag (Rorick *et al.*, 2013). The authors found HB expression to be significantly different in SM and UM and found that some HBs could distinguish different forms of severe disease. Specifically, they found that HB204 is associated with SM but is not associated with rosetting, suggesting that another mechanism for disease pathogenesis is responsible for the SM (Rorick *et al.*, 2013). IT4var13 contains HB204 and is known to bind CD36 (Robinson *et al.*, 2003) and ICAM-1 (Howell *et al.*, 2008) which is mediated by the CIDRα and DBLβ domains, respectively, resulting in its ability to bind endothelial cells (Madkhali, 2015). It is possible that this DBLα domain predictor reflects downstream sequences. A major impediment to our ability to associate specific sequences to disease phenotypes is that the underlying cause of many of those phenotypes is poorly understood. Indeed, disease phenotypes can overlap making it incredible difficult to disentangle cause and effect. A much larger dataset than that studied here is required to identify any specific associations between HBs, and indeed whole DBLβ domain sequences, and ICAM-1 binding.


A
52 65 53 4 13 3 14 5 2 59 1 62 61 18
IT4var14
IT4var16
IT4var41
IT4var12
IT4var01
J1a
HB3var34
IT4var11
J1d
IT4var13
IT4var27
BC12a
PCM7a
HB3var10
IT4var17
IT4var44
IT4var31
PCM7d
IT4var15
IT4var07 ii
IT4var02 i
IT4var02 ii
IT4var08
IT4var64
DD2var32 i
IT4var07 i
IT4var35
IT4var22
PF11 0521 i
AAQ73928
IT4var18
PFD1235w i
IT4var20
IT4var06
IT4var19
IT4var32b
Tree scale: 0.1
N-term
S1
S2a
S2b
S2c
S3a
S3b
C-term

B


ICAM-1 binding
52 65 53 4 13 3 14 5 2 59 1 62 61 18
IT4var14
IT4var16
IT4var41
IT4var01
J1a
J1d
IT4var13
IT4var27
BC12a
PCM7a
IT4var31
PCM7d
DD2var32_i
PF11_0521_i
AAQ73928
PFD1235w_i
ICAM-1 non-binding
IT4var12
IT4var11
IT4var17
IT4var44
IT4var15
IT4var07_ii
IT4var02_i
IT4var02_ii
IT4var08
IT4var07_i
IT4var35
IT4var22
IT4var18
IT4var20
IT4var06
IT4var19
N-term
S1
S2a
S2b
S2c
S3a
S3b
C-term

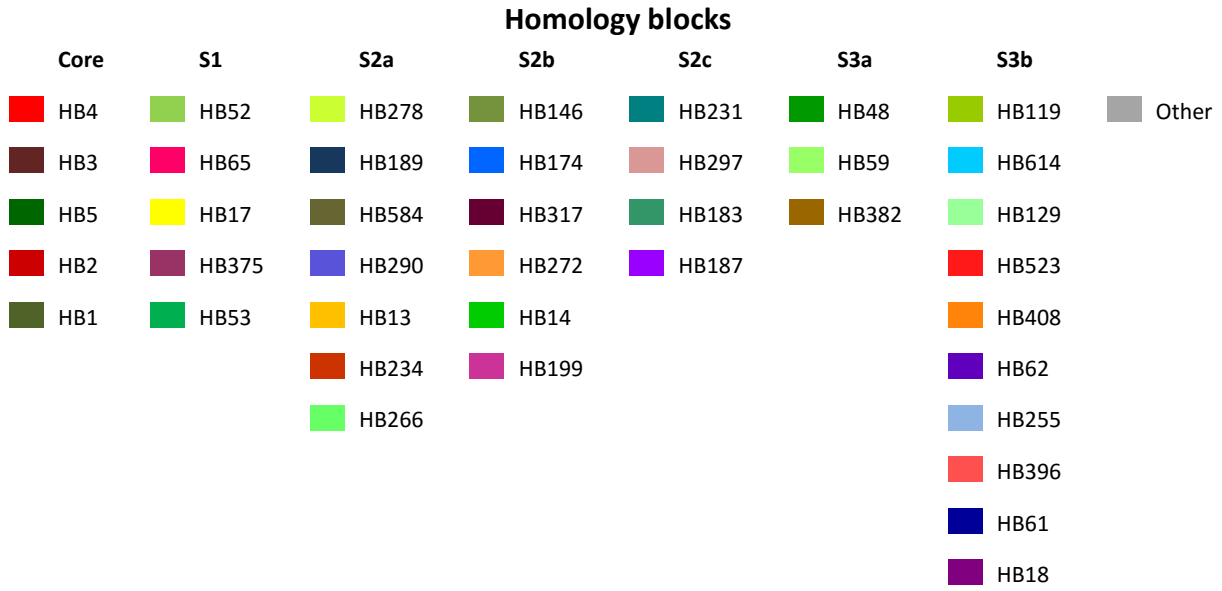


Figure 4.6. Homology block structure of DBLβ domains (Figure begins on page 91).

A (Page 91), left, maximum likelihood tree of ICAM-1 binding (red), non-binding (blue) and untested (black) DBLβ domains. Numbers indicate the bootstrap support. Right, homology block (HB) structure of each DBLβ sequence. Common HB numbers are indicated above the figure. **B (Page 92)**, HB structure of DBLβ domains grouped into ICAM-1 binding (red) and non-binding (blue). **Above**, colour key to HBs split into the following sections: Core: HBs shared with all DBLβ domains; S1: HBs N-terminal to HB4; S2a: HBs located between HB4 and HB3; S2b: HBs located between HB3 and HB5; S2c: HBs located between HB5 and HB2; S3a: HBs located between HB2 and HB1; S3b: HBs C-terminal to HB1; Other: HBs which appear >3 times in the sequences analysed. Sections are indicated below the HB structures with brackets.

4.3.6 Comparison of peptides important in DC4 DBLβ3::ICAM-1 binding

DC4-containing PfEMP1 are unusual among ICAM-1 binders in that they belong to group A (Bengtsson *et al.*, 2013b), suggesting that they may be more important in the development of SM. Evidence that this set of group A ICAM-1 binders may have a distinct binding site compared to group B and C binders is provided by differential blocking of binding by an antibody against domain 2 (D2) of ICAM-1 (Anja Jensen, personal communication). DC4 DBLβ binding is blocked by this antibody whereas group B and C DBLβ domains are not, suggesting the involvement of D2 in DC4 DBLβ ICAM-1 binding. A recent study has characterised a monoclonal antibody, 24E9, that blocks the interaction between the DBLβ3 domain of DC4 containing *var* genes and ICAM-1 and is effective against distinct parasite strains that also express DC4 (Lennartz *et al.*, 2015). The authors used 24E9 bound to PFD1235w DBLβ3 in hydrogen/deuterium exchange mass spectrometry (HDX MS) to identify key peptides that are masked by this antibody (Lennartz *et al.*, 2015). We performed a sequence comparison of these DBLβ peptides, P1-P3, to the dominantly expressed DBLβ domains of the ICAM-1 binding patient isolates (Figure 4.7) with particular interest in region P3a which was shown to be essential for 24E9 binding (Lennartz *et al.*, 2015).

P1 and P2a are relatively conserved which is in agreement with observations that these regions are conserved between both 24E9 binding (PFD1235w DBLβ3) and non-binding (Dd2var32 DBLβ1, IT4var13 DBLβ5 and IT4var16 DBLβ5) DBLβ domains and were likely identified due to steric hinderence (Lennartz *et al.*, 2015). There is very little conservation in the essential P3b region, although the DBLβ3 domains all share the A*NGG motif (Figure 4.7). However, this motif is also present in Dd2var32 DBLβ1 which does not bind 24E9 (Lennartz *et al.*, 2015). This sequence comparison suggests that the newly identified ICAM-1 binding DBLβ domains will not be inhibited by this DC4 specific antibody and that ICAM-1 binding is mediated by different peptides. They are, however, likely to be located on the same convex surface of the DBLβ domain (Lennartz *et al.*, 2015, Brown *et al.*, 2013, Howell *et al.*, 2008, Oleinikov *et al.*, 2009) which mainly consists of subdomain 3. Similar studies

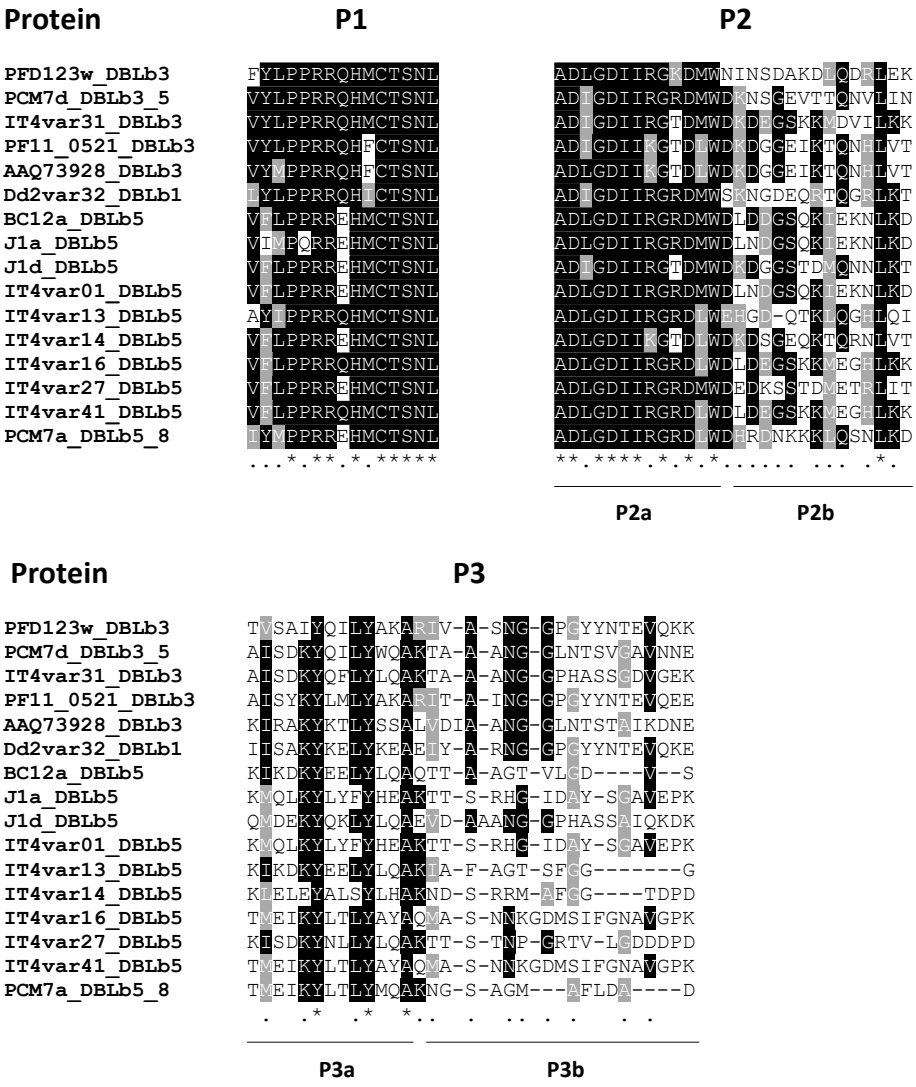


Figure 4.7. Sequence comparison of three peptides of DBLβ that are masked by a mAb specific to DC4 DBLβ3.

Sequences were aligned with T-Coffee (see Methods 4.2.5). P1-P3 were identified in HDX MS experiments with PFD1235w DBLβ3 as being masked by mAb binding and further characterised to show that region P3a is essential for this mAb to block ICAM-1 binding (Lennartz *et al.*, 2015).

are required to identify peptides from other DBLβ domains important in ICAM-1 binding.

The finding of remarkable sequence conservation between DC4 DBLβ3 domains (Bengtsson *et al.*, 2013b) could suggest that *var* gene classification at DC level is more informative than shorter sequence comparisons of single domains or tags. It may be that these common domain compositions represent desirable binding

functions that are maintained in the global *var* repertoire. This is supported by data on DC5 binding to PECAM-1 for which no single domain is clearly responsible (Berger *et al.*, 2013). However, the DC13 and DC8 binding EPCR story (Turner *et al.*, 2013, Lau *et al.*, 2015) is complicated by the inability of other research groups to replicate the results of binding experiments ((Gillrie *et al.*, 2015) and Alex Rowe, unpublished data). In particular, the findings of Gillrie *et al.* (2015) that the binding of DC13 and DC8 expressing parasites was not attributed to EPCR in all endothelial cell types tested, suggesting the involvement of other receptors. The highly variable nature of *var* genes, their constant recombination and their combined role in host receptor binding and antigenic variation make them very difficult to classify. This is exacerbated by the presence of exceptions in the majority of classification predictions which highlights the fact that sequence variability is on a continuous scale for which cut off points are required but are arbitrary.

4.3.7 Antigenic variation, receptor binding and the evolution of *var* genes

Pathogen-host interactions are the complex evolutionary result of centuries old battles for dominance. Many diverse pathogens, including viral, bacterial, protozoan and fungal examples, have independently developed mechanisms to avoid the host immune system with the ultimate goal of onward transmission. In this section we consider broader ideas of the evolution behind the *var* genes and the reason for their immense diversity at both the single parasite level and at the level of population genetics.

The dynamics of infection at population level follow different patterns which are largely dictated by host susceptibility and pathogen virulence and persistence within a population. Different strategies are implemented by different pathogens and lead to a range of population structures which can be modelled mathematically. An example of differential strategies and outcomes is found among the RNA viruses. Measles virus infects young hosts and emphasis is placed on transmission, resulting in a virus of one antigenic type for which an effective vaccine is available (Frank and Bush, 2007). Influenza, on the other hand, is a virus which infects a whole range of host ages and regularly produces new antigenic

types which can infect previously exposed individuals. Here, emphasis is on persistence in the population, rather than high transmissibility, resulting in the need for new vaccines each season (Frank and Bush, 2007). Cause and effect of such infection dynamics are inherently difficult to disentangle due to co-evolution and are not discussed here.

The *P. falciparum* parasite is a much more complex organism which utilises an antigenic variation strategy, as opposed to the random mutagenesis of RNA viruses, but parallels can be drawn with both situations. The different groups of *var* genes appear to act in different ways. The UPS A type *var* genes are similar to the measles example in that they are expressed in young children and immunity is quick to develop, whereas UPS B and C *var* genes follow the influenza pattern of persistence within the population. The presence of both strategies could be related to the trade off between what is beneficial in an individual infection versus what is beneficial to the population as a whole, as discussed by Holding and Recker (2015). Through modelling, they demonstrated that a parasite in an individual infection benefits from having multiple antigenic phenotypes which prolong infection by avoiding the immune response. However, population *var* repertoires would benefit from having a diverse set of parasites each with a single phenotype to reduce the chance of cross-reactivity between distinct infections (Holding and Recker, 2015). In reality, *P. falciparum* parasites have developed diverse individual *var* repertoires which have minimal overlap with other genetically distinct parasites, gaining the benefit of both scenarios.

The strategy of antigenic variation to avoid the host immune system is found in many pathogens. Bacterial examples include *Borrelia hermsii*, *Neisseria* spp. and *Trepomena pallidum* (Deitsch *et al.*, 2009). Protozoal examples include *Giardia lamblia* (Prucca *et al.*, 2011), *Babesia bovis* (Allred and Al-Khedery, 2004) and the extensively studied *Trypanosoma brucei* (Horn, 2014). An interesting aspect to consider for *P. falciparum* is the need to maintain receptor binding while producing antigenically distinct PfEMP1 proteins. This is thought to restrict the evolution and diversification of these genes (Holding and Recker, 2015), which is not the case in gene families with a role solely in antigenic variation, such as *T. brucei* and *G.*

lamblia. Clues as to how receptor binding ability is retained are found in *var* gene recombination mechanisms. Recombination mainly occurs between domains of the same type, is initiated in HBs and can involve multiple crossovers between the two genes in question (Claessens *et al.*, 2014, Kraemer *et al.*, 2007). HBs contain conserved amino acids important in the maintenance of protein structure (Kraemer *et al.*, 2007, Rask *et al.*, 2010) and they flank the three structural subdomains of DBL domains and the two structural subdomains of CIDR domains (Rask *et al.*, 2010), suggesting that their function in recombination initiation allows recombination while maintaining structure. In addition, the HBs shared between domain subtypes allow for multiple crossovers to occur, presumably maintaining receptor binding ability within a subtype. However, it is likely that the situation is, or was once, more complex to allow for generation of new domain types. The chromosomal location and orientation of transcription is thought to be responsible for the restriction of *var* gene recombination to mainly within the groups A, B and C (Gardner *et al.*, 2002, Kraemer and Smith, 2003, Kraemer *et al.*, 2007, Claessens *et al.*, 2014). Thus, resulting in specific adhesion phenotypes of gene groups e.g. CD36 binding of group B and C genes but not group A (Robinson *et al.*, 2003). In this manner, the groups are separate and preserve the specialisation of specific receptor binding, particularly between group A and C genes. However, the potential for recombination between groups is present, as evidenced by group B/A genes (Lavstsen *et al.*, 2003, Rask *et al.*, 2010), allowing the relatively rare generation of new gene types.

4.4 Summary

In this chapter, we further tested the predictive powers of the Pf3k *var* gene database by analysing the returned sequences to the non-dominant DBLα tags of the ICAM-1 binding patient isolates and the short sequences identified specifically in cases of fatal cerebral malaria, 28B1-1 and 62B1-1. We found that the domain structure and predicted group type of the database sequence hits were in agreement with predictions made from PoLV analysis of the DBLα tag. Together

these analyses identified group A and B/A virulence associated genes, which included the two genes identified as important in CM. We characterised the PfEMP1 domain structure of the dominantly expressed *var* genes of BC12, J1 and PCM7. The DBLβ domains of BC12a, J1a and J1d are all of subtype 5 and are predicted to bind ICAM-1. PCM7a and PCM7d DBLβ domains fell between two subtypes: 5/8 and 3/5, respectively. Phylogenetic comparisons showed that these DBLβ domains generally grouped with ICAM-1 binding domains but the presence of exceptions reinforces the need for experimental validation to confirm, and increase the reliability of, sequence predictions. HB analysis did not reveal any patterns specific to ICAM-1 binding DBLβ domains, although the relatively small number of sequences analysed may not detect such patterns. Finally, comparisons with DBLβ peptides identified as important in ICAM-1 binding of DC4-containing PfEMP1, suggest that binding of group A proteins and group B and C proteins to ICAM-1 involves distinct peptides. Mechanisms of *var* gene recombination provide clues as to how such differences in host receptor binding are maintained within *var* gene groups.

Chapter 5. Recombinant DBLβ::ICAM-1 molecular interactions

5.1 Introduction

ICAM-1 is an endothelial receptor involved in transmigration of leukocytes to sites of inflammation. Ligands for this receptor include the native lymphocyte function associated molecule 1 (LFA-1) (Marlin and Springer, 1987) and fibrinogen (Languino *et al.*, 1993), and the pathogens rhinovirus (Greve *et al.*, 1989) and *P. falciparum* IEs (Berendt *et al.*, 1989). ICAM-1 consists of five Ig domains (D1-D5) with the N-terminal D1 mediating binding to LFA-1 and *P. falciparum* IEs at distinct but overlapping sites (Berendt *et al.*, 1992, Ockenhouse *et al.*, 1992a). The crystal structure of D1 has been known for some time (Casasnovas *et al.*, 1998) and comprises a number of β-sheets named A-G (Figure 5.1). The *P. falciparum* binding site has been identified via mutagenesis studies as the BED side of ICAM-1 (Berendt *et al.*, 1992, Ockenhouse *et al.*, 1992a, Tse *et al.*, 2004), indicated in Figure 5.1.

Importantly, different parasite strains have been shown to be differentially affected by some of the mutations tested (Tse *et al.*, 2004), including the culture-adapted patient isolates that are the subject of this study (Madkhali *et al.*, 2014), indicating subtle contact residue differences. One such mutation is the naturally occurring homozygous K29M mutation known as ICAM-1^{Kilifi} which reduced binding of A4 parasites but had minimal effect on ItG (Tse *et al.*, 2004, Madkhali *et al.*, 2014). These results could help explain differences in studies which found ICAM-1^{Kilifi} to be either associated with CM (Fernandez-Reyes *et al.*, 1997), have no association to SM (Bellamy *et al.*, 1998, Ohashi *et al.*, 2001) or be protective of SM (Kun *et al.*, 1999), depending on the predominant PfEMP1s expressed in these geographically diverse populations.

PfEMP1 proteins are made up of a number of DBL and CIDR domains that mediate binding to a variety of host endothelial cell receptors. They belong to a wider family of *Plasmodium* proteins that also have DBL domains as their building blocks,

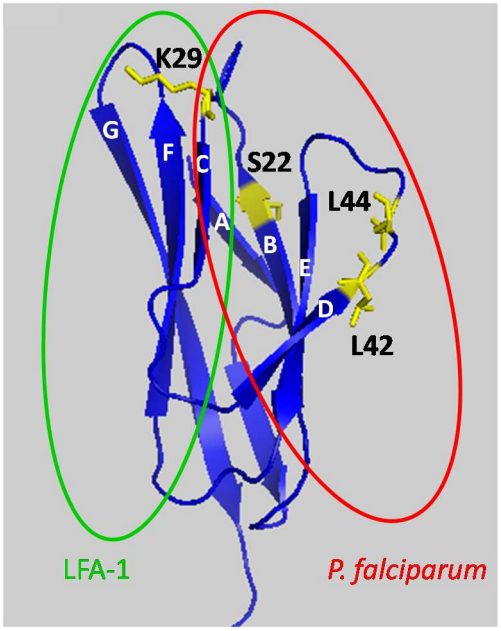


Figure 5.1. Crystal structure of the N-terminal domain 1 (D1) of ICAM-1.
The binding sites of LFA-1 and *P. falciparum*-infected erythrocytes are shown. β-sheets are labelled A-G and the four mutated residues assessed in this study are highlighted in yellow and labelled. Structure accessed via Protein Data Bank, deposited by Casasnovas *et al.* (1998).

including the EBL and RBL invasion proteins (Culleton and Kaneko, 2010). Although they are all α-helical proteins with a conserved overall fold, they share just 4% sequence identity and 15% similarity (Higgins and Carrington, 2014). Crystal structures of several PfEMP1 DBL (Higgins, 2008b, Khunrae *et al.*, 2009, Vigan-Womas *et al.*, 2012) and CIDR (Klein *et al.*, 2008, Lau *et al.*, 2015) domains have been solved. These, along with small angle x-ray scattering (SAXS) data, have revealed two different strategies of PfEMP1 organisation and receptor binding. VAR2CSA mediates IE binding to chondroitin sulphate A (CSA) on placental syncytiotrophoblasts and has been shown to have a compact structure where the domains are folded together, with the DBL2X domain mainly responsible for binding (Clausen *et al.*, 2012). In contrast, the ICAM-1 binding IT4var13 has been shown to have a rigid, elongated structure whereby binding is mediated by the DBLβ5 domain only (Brown *et al.*, 2013). This elongated, modular structure is predicted to be the norm, based on the uniqueness of VAR2CSA as a semi-conserved protein with a very specific binding phenotype (Salanti *et al.*, 2003) and evidence from an EPCR

binding gene that the CIDRα1.1 domain binds its receptor with affinity comparable to the full length protein (Turner *et al.*, 2013). There are no crystal structures of an ICAM-1 binding DBLβ domain to date. Mutagenesis of multiple residues of DBLβ domains have been shown to decrease ICAM-1 binding, however, it is unclear whether these are residues of the contact site or if they are important for protein folding (Howell *et al.*, 2008).

The majority of data on the DBLβ::ICAM-1 interaction is from parasites of the IT lineage. In the absence of structural information on which to model sequences, a more diverse set of DBLβ domains need to be characterised, especially since ICAM-1 binding cannot be solely predicted by sequence comparison (Howell *et al.*, 2008). This chapter aims to recombinantly express the DBLβ domains of ICAM-1-binding patient isolates identified in Chapter 3 and analysed in Chapter 4, and characterise their binding ability to wild type ICAM-1 and four mutant proteins (mutated residues indicated in Figure 5.1).

5.2 Methods

5.2.1 Expression plasmid preparation

Full length DBLβ domains were PCR amplified with the proofreading *TaKaRa LA Taq*® DNA polymerase (Clontech, TaKaRa Bio Inc), as described (see 2.2.3), from parasite gDNA using primers listed in Appendix A. Reaction conditions were an initial denaturing step of 95 °C for 3 minutes, followed by 30 cycles of 95 °C for 30 seconds, 52 °C for 30 seconds, 65 °C for 40 seconds, and a final extension of 65 °C for 3 minutes. PCR products were purified as stated (see 2.2.4). PCR products were restriction digested in a reaction comprising 2 µl CutSmart® Buffer (NEB), 20 U each of *Bam*HI and *Xho*I (NEB), 5 µl purified PCR product and made up to 20 µl with sterile water. The reaction was incubated overnight at 37 °C. A modified pET15b plasmid (kindly provided by the Higgins group, Oxford) was digested with the same reagents and incubated at 37 °C for 2 hours. Restriction digests of both vector and product were run on a 1% agarose gel, visualised with ethidium bromide and

purified as described (see 2.2.4). Purified digested PCR products were ligated into the pET15b vector using T4 DNA Ligase (NEB) in a reaction comprising 2 µl 10x T4 DNA Ligase Buffer, 10:1 ratio of product to vector, 1 µl T4 DNA Ligase (NEB) and nuclease- free water up to 20 µl. The ligation was incubated overnight at 16 °C and heat inactivated at 65 °C for 10 minutes.

5.2.2 Transformation of *E. coli*

Expression plasmids were transformed into One Shot® TOP10 *E. coli* competent cells (Invitrogen) as described (see 3.2.4). Overnight cultures were grown and Miniprep performed to isolate the plasmid as described (see 3.2.4). Plasmids were sent for sequencing (see 2.2.5) to confirm correct in frame sequence.

Expression plasmids were then transformed into SHuffle® T7 Express lysY competent *E. coli* (C3030) (NEB), a cell line optimized to form proteins containing disulphide bonds. Competent cells were thawed on ice for 10 minutes and 5 µl plasmid DNA added, mixed and incubated for 30 minutes on ice. Cells were heat-shocked at 42 °C for 30 seconds and placed on ice for 5 minutes. 950 µl SOC was added and the cells were incubated at 30 °C for 1 hour with agitation. 10- fold serial dilutions were carried out and 100 µl spread on agar plates containing 100 µg/ml ampicillin before incubating overnight at 30 °C. Ten colonies were picked and cultured overnight at 30 °C in 2-YT broth (1.6% tryptone, 1% yeast extract, 0.5% NaCl) before plasmid preparation as described (see 3.2.4). Plasmids were PCR amplified to confirm presence of the insert. Glycerol stocks were made by adding 800 µl 80% glycerol to 500 µl overnight culture and freezing at – 80 °C.

5.2.3 Recombinant protein expression

Optimisation

Expression clones were grown in 3 ml 2-YT broth (1.6% tryptone, 1% yeast extract, 0.5% NaCl) at 30 °C to an optical density at 600 nm (OD_{600}) of 0.8 or 1.2 before induction with 0.4 mM or 1 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside, Sigma). Following induction, cultures were grown at 25 °C for 16- 20 hours. Un-induced cultures at each OD_{600} were also grown. Cultures were centrifuged at 5000

x g for 20 minutes and the pellet resuspended in Lysis Buffer (20 mM Tris- HCl, 150 mM NaCl, 30 mM imidazole, 2% glycerol, 0.9% Triton X-100, pH 7.4) and placed at -80 °C overnight. Cells were thawed on ice and cOmplete EDTA-free protease inhibitors (Roche) were added before sonication with a VibraCell™ sonicator (Jencons Scientific Ltd) at 80 Amp, 10 second pulse, 10 second pause, repeated until cells are disrupted. Disrupted cells were spun down at 16,200 x g for 20 minutes in a bench-top centrifuge and samples of the supernatant and pellet were taken for SDS-PAGE and Western blot analysis (see 2.3.1, 2.3.2, 2.3.3). Optimal conditions for each DBLβ domain are show in Table 5.1.

Table 5.1. Optimal DBLβ expression conditions in SHuffle® C3030 cells.

DBLβ:pET15b	OD_{600}	[IPTG] (mM)	Temperature (°C)
BC12a clone #1	**1.2**	**1**	**25**
J1a clone #8	**1.2**	**1**	**25**
PCM7a clone #5	**1.2**	**0.4**	**25**

5.2.4 Recombinant protein purification

Affinity purification

N-terminal His-tagged DBLβ domains were purified by affinity chromatography using nickel-nitrilotriacetic acid agarose (Ni-NTA, Qiagen) under native conditions. 100 ml test cultures were induced in optimal conditions (Table 5.1) and cells disrupted by a freeze-thaw and sonication as described in section 5.2.3 above. Disrupted cells were spun down at 30,000 x g in a Beckman Avanti J25 centrifuge for 20 minutes. The supernatant was added to 1 ml Ni-NTA agarose pre-equilibrated in Lysis Buffer (20 mM Tris- HCl, 150 mM NaCl, 30 mM imidazole, 2% glycerol, 0.9% Triton X-100, pH 7.4) and incubated at 4 °C for 1 hour with gentle shaking. The mixture was loaded onto a polypropylene column (Qiagen) and the flow-through collected. The column was washed twice with 2.5 ml Wash Buffer (20 mM Tris-HCl, 150 mM NaCl, 30 mM imidazole, pH 7.4). Protein was eluted from the column with Elution Buffer (20 mM Tris-HCl, 150 mM NaCl, 200 mM imidazole, pH

7.4) in four elutes, followed by a final elute with elution buffer containing 500 mM imidazole. Samples were taken at each step for SDS-PAGE analysis (see 2.3.1).

Affinity purification of larger cultures followed the same protocol with volumes appropriately adjusted. A major difference was that the initial cell supernatant was passed over the Ni-NTA agarose 3 times and allowed to drip by gravitational flow. Protein was eluted in 10 fractions with Elution Buffer (20 mM Tris-HCl, 150 mM NaCl, 200 mM imidazole, pH 7.4). Samples were taken at each step for SDS-PAGE analysis (see 2.3.1).

Concentrating protein, buffer exchange and measurement of concentration

Elution fractions containing the desired protein were pooled and concentrated using an Amicon ultracell 30 kDa concentrator (Merck Millipore) to ~ 2.5 ml. Buffer exchange into 20 mM Tris-HCl, 150 mM NaCl, pH 7.4 was performed using a PD-10 column (GE Healthcare) by gravity protocol. Protein concentration was then measured by absorbance at 280 nm (A_{280}) on a NanoDrop® using the molar extinction coefficient and molecular weight of each domain.

Stability test

Affinity purified DBLβ domains were incubated at 4 °C or room temperature for 48 hours. Samples of each were analysed by SDS-PAGE (see 2.3.1, 2.3.2) along with a sample before incubation. All subsequent purification was carried out at 4 °C.

Gel filtration

Samples were further purified on a HiLoad 16/600 Superdex 75 prep grade column (GE healthcare). Samples were concentrated to approximately 3.5 ml before loading onto a column pre-equilibrated with running buffer (20 mM HEPES, 200 mM NaCl, pH 7.2, filtered (0.2 µm Filter), degassed). The running speed was 0.75 ml/minute and fractions were collected between 38 and 100 ml. Fractions were checked by SDS-PAGE (see 2.3.1, 2.3.2).

5.2.5 CD spectroscopy

Protein was concentrated to ~ 200 µl using an Amicon ultra 0.5 ml centrifugal filter 10 k (Merck Millipore) and dialyzed into 100 mM sodium phosphate buffer with 200

mM sodium fluoride, pH 7.4 using Slide-A-Lyzer 3.5 k Dialysis Cassettes (Thermo Scientific) overnight at 4 °C. The far-UV CD spectroscopy experiments were carried out (by Frank Lennartz) in a J-815 Spectropolarimeter (Jasco) equipped with a computer-controlled Peltier temperature control unit. Measurements were carried out at a protein concentration of 0.3 mg/ml using a 1 mM path cell. All measurements were taken at 20 °C between 195 and 260 nm wavelengths. Four spectra were recorded for each sample, averaged and corrected for buffer absorption.

5.2.6 Surface plasmon resonance (SPR)

SPR was carried out on a Biacore T200 machine (GE Healthcare). Protein A was immobilised onto a CM5 chip (GE Healthcare) by amine coupling. All experiments were carried out in buffer containing 10 mM HEPES, pH 7.2, 250 mM NaCl and 0.05% Tween- 20, filter sterilised and degassed. ICAM-1 proteins fused to the Fc region of human IgG1 (ICAM-1^{D1D5} (Lennartz *et al.*, 2015), ICAM-1^{Kilifi}, ICAM-1^{S22A}, ICAM-1^{L42A}, ICAM-1^{L44A} (Tse *et al.*, 2004)) were captured onto the Protein A surface. Coupling of the different ICAM-1 proteins to the chip was kept comparable at 1108 ± 108 response units (RU). DBLβ domains were flowed over the chip in a concentration series (1000 nM – 7.81 nM) at a flow rate of 30 µl/min with an association time of 240 seconds and a dissociation time of 600 seconds. The chip was regenerated after each concentration with the injection of 10 mM Glycine- HCl, pH 1.7 for 120 seconds at a flow rate of 10 µl/min which breaks the Protein A- Fc interaction. The signal from a flow channel without ICAM-1 was subtracted from all measurements. Single stranded DNA (ssDNA, Salmon Sperm DNA sodium salt, AppliChem) was added to BC12a$^{DBL\beta}$ to 1 mg/ml final concentration prior to injection to minimise background binding observed during test runs. Injection of ssDNA only had no effect on the signal (data not shown). Sensorgrams were fitted to a global 1:1 interaction model allowing calculation of kinetic values k_a, k_d and K_D using BIAevaluation software 2.0.3 (GE Healthcare).

SPR competition assay

Competition assays were carried out with ICAM-1^{D1D5} under the same conditions outlined above except that only one concentration of DBLβ was used (adjusted to give the same molar ratio of each protein). The first domain was injected at a flow rate of 30 µl/min with an association time of 240 seconds. The second domain was then injected at a flow rate of 30 µl/min with an association time of 240 seconds and a dissociation time of 600 seconds.

5.2.7 Flow adhesion assay

Flow adhesion assays were carried with the VenaFlux semi-automated microfluidic system and VenaFlux software (Cellix). The Vena8 Fluoro+^{TM} protein biochips (Cellix) used contain 8 channels which allow several proteins to be tested in immediate succession. Individual biochip channels were coated with 4 µl of 50 µg/ml of either ICAM-1^{D1D5}, ICAM-1^{S22A}, ICAM-1^{Kilifi}, ICAM-1^{L42A} or ICAM-1^{L44A} protein and incubated at 37 °C for 1 hour in a humidified petri dish. The protein was then aspirated and the channel blocked with 1% BSA/PBS for either 1 hour at 37 °C or overnight at 4 °C. Channels were warmed to 37 °C before use. IE were prepared in binding buffer (RMPI 1640 powder (Invitrogen) dissolved in H_2O, pH 7.2) at 3% parasitaemia and 2% haematocrit. The assay followed the Cellix protocol. The protein coated biochip was attached to a microscope stage within a plastic chamber whose temperature was 37 °C. The IE were drawn through the channel at a flow rate of 0.04 Pa for 8 minutes before washing with binding buffer. Bound parasites were counted in 7-10 microscope fields and the numbers adjusted to IE/ mm^2.

5.2.8 RT-qPCR

RNA was extracted from parasite cultures and cDNA synthesised as described (see 2.2.1, 2.2.2). Reverse transcriptase, quantitative PCR (RT-qPCR) was carried out as described using primers to DBLα tags of BC12 and J1 (see 3.2.5) and previously published primers to IT4 ICAM-1-binding *var* genes (Viebig *et al.*, 2007) , Appendix A.

5.3 Results

5.3.1 Optimisation of recombinant DBLβ domain expression

DBLβ domains from BC12a, J1a, J1d, PCM7a and PCM7d were tagged with an N-terminal hexa-histidine motif (His-tag) by cloning into a modified pET15b expression vector. Plasmids were sequenced to confirm the insert was correct and in-frame. J1d_DBLβ:pET15b and PCM7a_DBLβ:pET15b contain the correct sequence and BC12a_DBLβ:pET15b contained one silent (synonymous) mutation, not affecting the amino acid sequence (CTC > CTT, L166). J1a_DBLβ:pET15b contained one serine to threonine mutation near the N-terminus which we predict to have minimal effect on the fold of the protein and is not in the predicted binding region (Howell *et al.*, 2008). However, PCM7d_DBLβ:pET15b contained two mutations to amino acids with markedly different side chains which would have an unknown effect on the protein fold. Repeated efforts to clone this DBLβ sequence proved difficult and two additional clones also contained point mutations. We did not pursue this further due to time constraints.

Initial expression conditions were tested on PCM7a_DBLβ:pET15b clones in a range of bacterial strains, IPTG concentrations and temperatures (not shown). Based on these initial studies and previous experience with these domains (Frank Lennartz, personal communication), we tested clones of each DBLβ for inducibility and solubility at OD_{600} 0.8 or 1.2 with IPTG concentration of 0.4 or 1 mM (Figure 5.2) to determine the optimal expression conditions for each DBLβ domain. Shown in Figure 5.2 are examples of inducible clones for each construct and one example of a clone which constitutively expressed the recombinant DBLβ domain (Figure 5.2B, BC12a_DBLβ:pET15b clone #2). J1d$^{DBL\beta}$ protein was highly degraded (Figure 5.2D) in all clones tested. We reduced the induction time in an attempt to improve the stability of this domain but observed the same levels of degradation. Further optimisation was not undertaken due to time constraints. Some degradation of PCM7a$^{DBL\beta}$ (Figure 5.2E) at OD_{600} 0.8 was apparent with overexposure of the membrane (not shown). Optimal conditions for BC12a$^{DBL\beta}$, J1a$^{DBL\beta}$ and PCM7a$^{DBL\beta}$ are summarised in Table 5.1, Methods section 5.2.3.

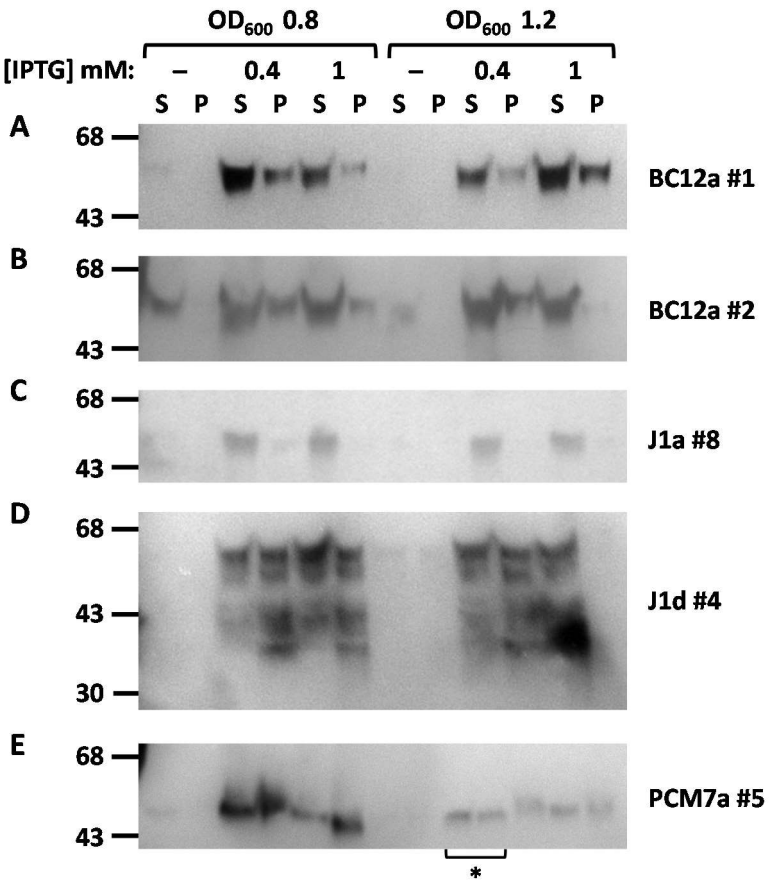


Figure 5.2. Examples of DBLβ expression optimisation.
SHuffle® C3030 were grown in 3 ml 2-YT broth at 37 °C to OD_{600} 0.8 or 1.2 and induced with 0.4 mM or 1 mM ITPG and grown at 25 °C for 16-20 hours. Cells were harvested, sonicated and samples taken of the supernatant (S) and pellet (P) and analysed by Western blot. DBLβ proteins were detected by α-His antibody. **A**, BC12a_DBLβ:pET15b clone #1, expected size 54.6 kDa. **B**, BC12a_DBLβ:pET15b clone #2, expected size 54.6 kDa. **C**, J1a_DBLβ:pET15b clone #8, expected size 54 kDa. **D**, J1d_DBLβ:pET15b clone #4, expected size 57 kDa. **E**, PCM7a_DBLβ:pET15b clone #5, expected size 44 kDa. "*" indicates a sample shared between two wells.

5.3.2 Optimisation of affinity purification and scale-up

DBLβ proteins were purified by affinity chromatography on Ni-NTA agarose, first in a small scale test from 100 ml culture, followed by scaling-up to 2 L culture. We observed varying results for each DBLβ domain. $BC12a^{DBL\beta}$ was successfully purified from both the small scale and large 2 L cultures with minimal carryover of non-

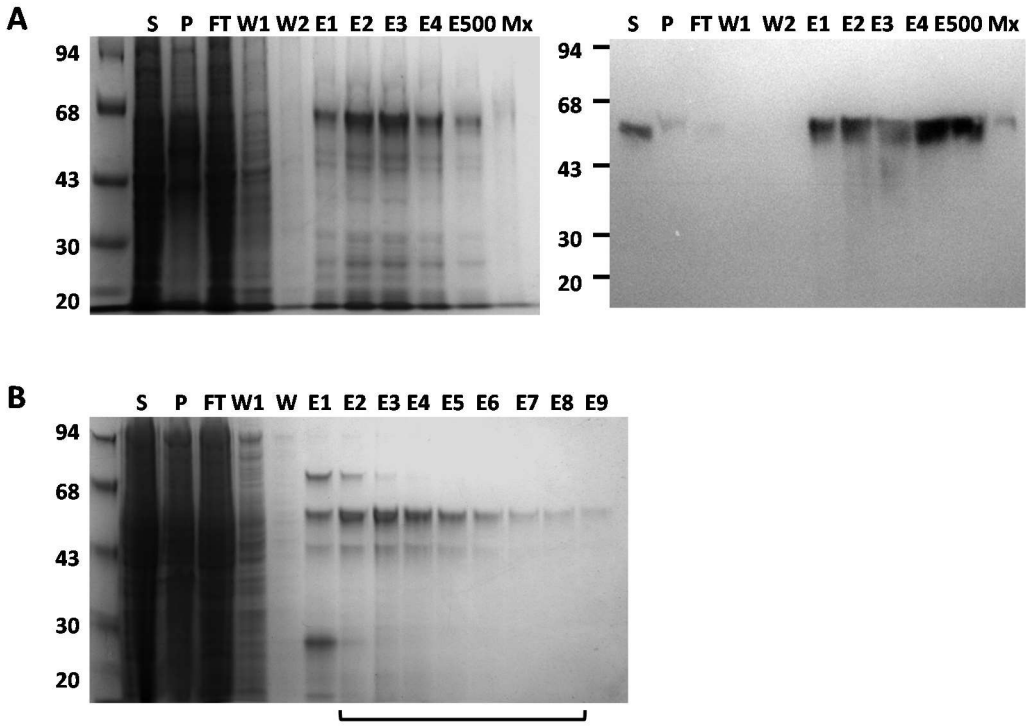


Figure 5.3. Ni-NTA purification of BC12a$^{DBL\beta}$.
A, Initial purification trial of 100 ml culture expressing BC12a$^{DBL\beta}$, visualised by coomassie (left) and Western blot (right). Protein was eluted with 200 mM imidazole in four fractions (E1-4) and finally with 500 mM imidazole (E500). **B**, Large scale 2 L culture expressing BC12a$^{DBL\beta}$ eluted with 200 mM imidazole, visualised by coomassie. Fractions pooled for further purification are shown (black bar) and consist of 7 mg total protein. S, supernatant. P, pellet. FT, flow-through. W, wash. E, elute. Mx, sample of Ni-NTA agarose matrix after elution.

specific proteins (Figure 5.3A and B, respectively). Seven elution fractions were pooled and contained ~ 7 mg total protein.

J1a$^{DBL\beta}$ purified in the small scale test culture (Figure 5.4A) but appeared as a double band in the 2 L culture with the majority of protein in the first elute (Figure 5.4B). A similar pattern was observed when purifying from 500 and 250 ml cultures (Figure 5.4C and D, respectively). Therefore, J1a$^{DBL\beta}$ was purified from four 100 ml cultures and pooled, producing ~ 3.7 mg total protein (one example shown, Figure 5.4E).

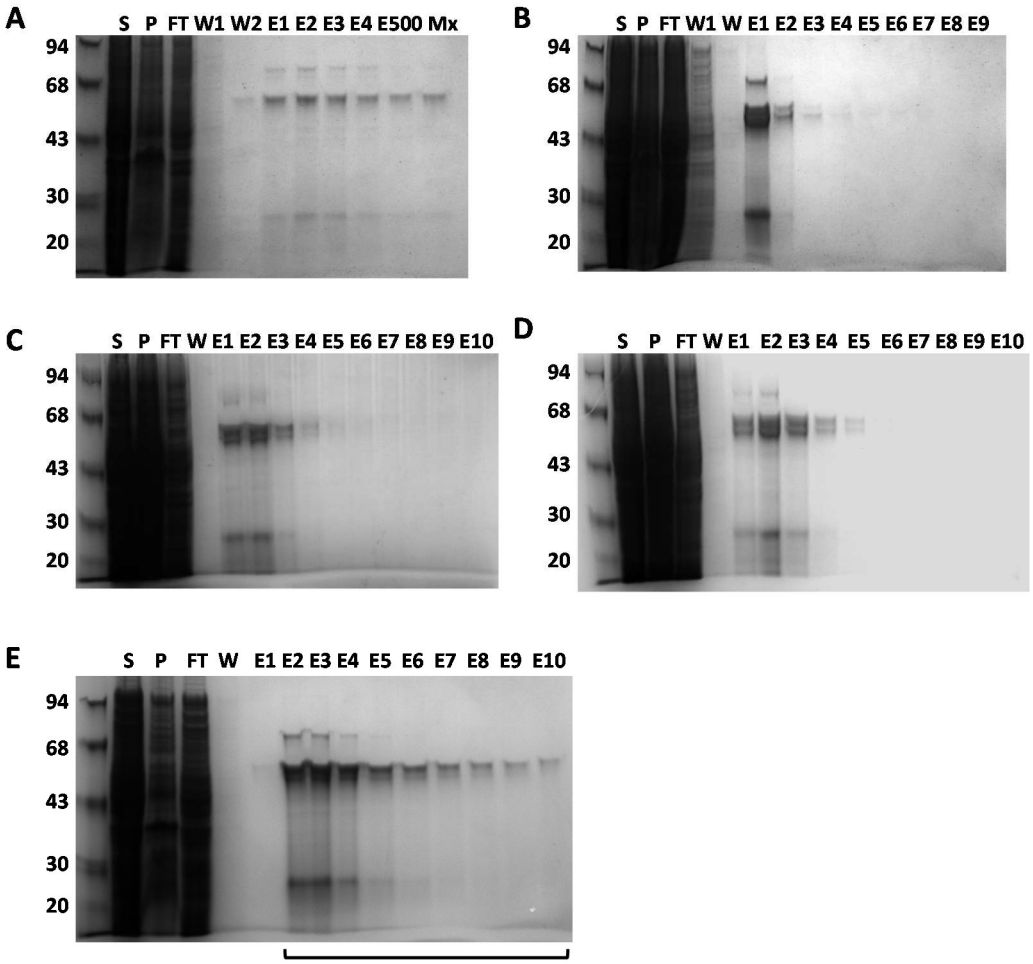


Figure 5.4. Ni-NTA purification of J1a$^{DBL\beta}$.
A, Initial purification trial of 100 ml culture expressing J1a$^{DBL\beta}$, visualised by coomassie. Protein was eluted with 200 mM imidazole in four fractions (E1-4) and finally with 500 mM imidazole (E500). Larger scale cultures expressing J1a$^{DBL\beta}$ of 2 L (**B**), 500 ml (**C**) and 250 ml (**D**) eluted with 200 mM imidazole. **E**, 1 of 4 purifications from 100 ml cultures that were pooled for further purification (black bar and data not shown). S, supernatant. P, pellet. FT, flow-through. W, wash. E, elute. Mx, sample of Ni-NTA agarose matrix after elution.

PCM7a$^{DBL\beta}$ purified from small scale culture appeared as a double band and slightly degraded, although it was unclear if this was simply a diffuse gel run (Figure 5.5A). Purification from a 2 L culture resulted in the majority of protein in the first elute which appears highly degraded (Figure 5.5B). A 250 ml culture produced very little protein which was similarly degraded (Figure 5.5C). A repeat of purification from a

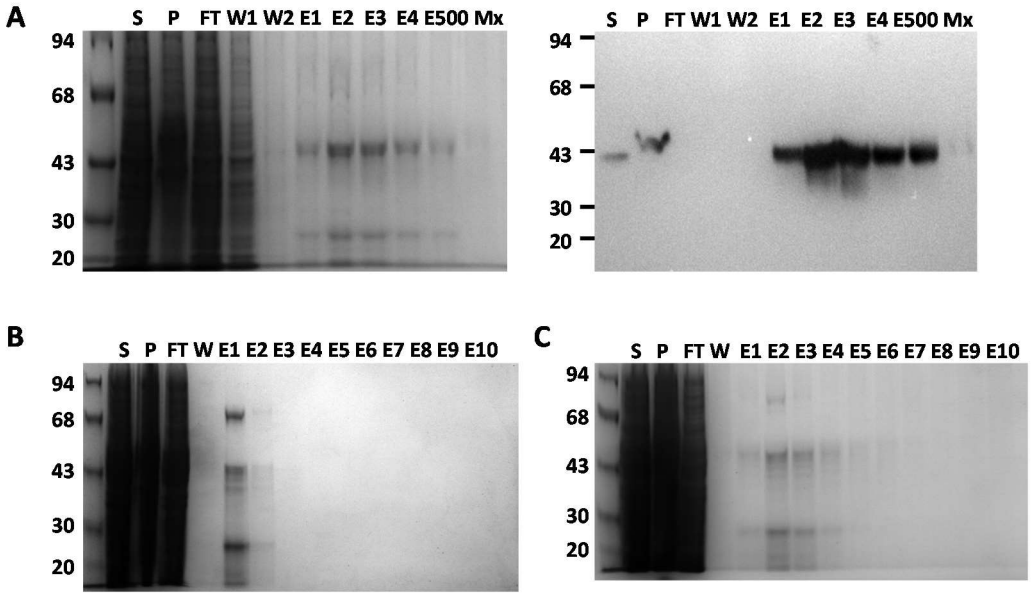


Figure 5.5. Ni-NTA purification of PCM7a$^{DBL\beta}$.
A, Initial purification trial of 100 ml culture expressing PCM7a$^{DBL\beta}$, visualised by coomassie (left) and Western blot (right). Protein was eluted with 200 mM imidazole in four fractions (E1-4) and finally with 500 mM imidazole (E500). Expression in larger scale 2 L (**B**) and 250 ml (**C**) cultures, visualised by coomassie. S, supernatant. P, pellet. FT, flow-through. W, wash. E, elute. Mx, sample of Ni-NTA agarose matrix after elution.

100 ml culture also resulted in degraded protein (data not shown). PCM7a$^{DBL\beta}$ production was not pursued further due to time constraints.

BC12a$^{DBL\beta}$ and J1a$^{DBL\beta}$ stability test

Due to the apparent sensitivity of the DBLβ domains, we performed a stability test. Samples of BC12a$^{DBL\beta}$ and J1a$^{DBL\beta}$ Ni-NTA purified proteins were incubated at 4 °C or room temperature (RT) for 48 hours. Extensive degradation of RT samples occurred and a slight shift in molecular weight of 4 °C samples indicates some instability (Figure 5.6). All steps in further purification were carried out at 4 °C to minimise protein degradation.

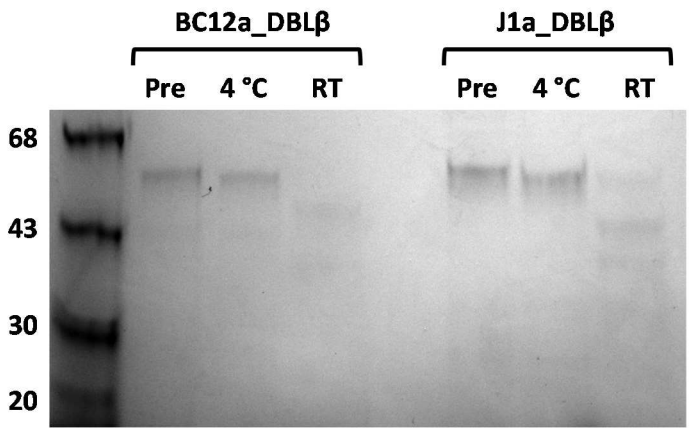


Figure 5.6. DBLβ domain stability test.
Ni-NTA purified BC12a^{DBLβ} and J1a^{DBLβ} were incubated at either 4 °C or room temperature (RT) for 48 hours. Samples are run alongside pre-incubated protein (Pre) in a coomassie stained SDS-PAGE gel.

5.3.3 Purification of BC12a^{DBLβ} and J1a^{DBLβ} by gel filtration

BC12a^{DBLβ} and J1a^{DBLβ} were further purified by gel filtration. Much of BC12a^{DBLβ} formed aggregates which eluted first, along with the majority of non-specific proteins (Figure 5.7A, left gel). However, sufficient protein subsequently eluted in fractions containing ~ 93% BC12a^{DBLβ} which were pooled as indicated in Figure 5.7A (right gel). A smaller proportion of J1a^{DBLβ} formed aggregates (Figure 5.7B, left gel), with the majority of protein eluting in subsequent fractions which were pooled as indicated to produce ~ 97% pure protein (Figure 5.7B, right gel).

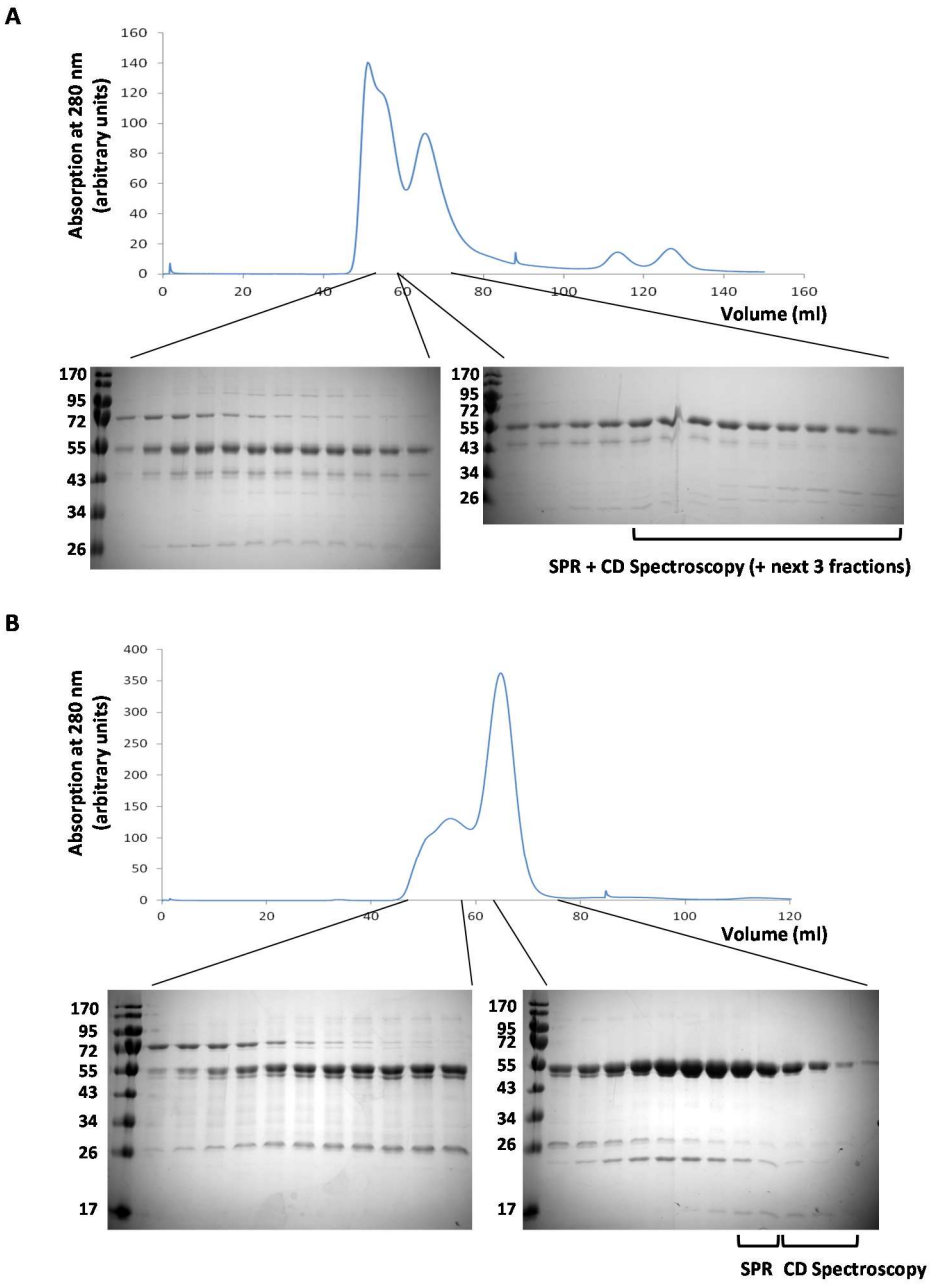


Figure 5.7. Purification of BC12a$^{DBL\beta}$ and J1a$^{DBL\beta}$ by gel filtration.
DBLβ domains were passed through a HiLoad 16/600 Superdex 75 prep grade column (GE healthcare) and 1 ml fractions collected between 38 and 100 ml. **A**, BC12a$^{DBL\beta}$ gel filtration and SDS-PAGE of samples from fractions 11-22 (left) and 23-35 (right). Fractions pooled for use in SPR and CD spectroscopy are indicated (black bar). **B**, J1a$^{DBL\beta}$ gel filtration and SDS-PAGE of samples from fractions 11-21 (left) and 22-34 (right). Fractions pooled for use in SPR and CD spectroscopy are indicated (black bars).

5.3.4 Circular dichroism (CD) spectroscopy of BC12a$^{DBL\beta}$ and J1a$^{DBL\beta}$

Circular dichroism (CD) spectroscopy was carried out (by Frank Lennartz, University of Oxford) to determine whether the purified DBLβ domains are folded correctly. Recombinant DBLβ domains from both BC12a and J1a resulted in spectra typical of α-helical proteins, as expected (for review and typical spectra see (Greenfield, 2006)).

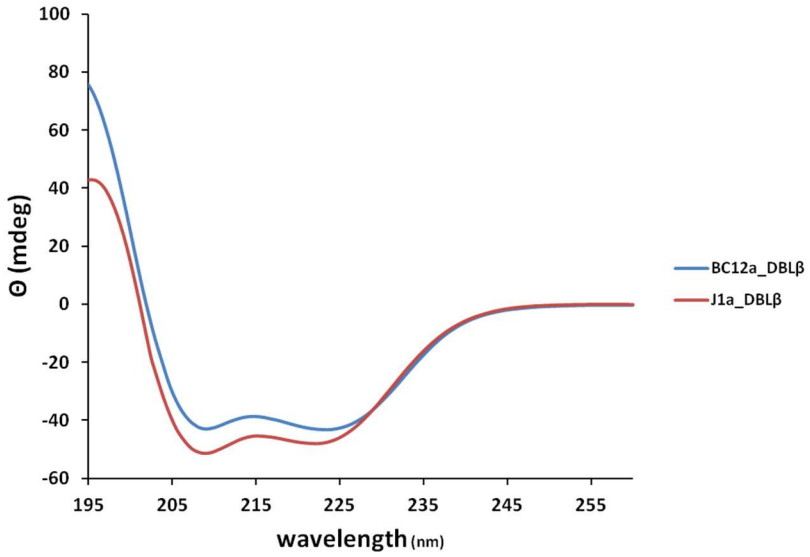


Figure 5.8. CD spectroscopy of BC12a$^{DBL\beta}$ and J1a$^{DBL\beta}$ recombinant domains.
CD spectra were recorded between 195 and 260 nm at 20 °C. For each protein, four measurements were averaged and corrected for buffer absorption.

5.3.5 Recombinant DBLβ interactions with ICAM-1

DBLβ binding kinetics to ICAM-1^{D1D5}

Surface plasmon resonance (SPR) was carried out to measure DBLβ::ICAM-1 binding kinetics. ICAM-1^{D1D5} was coupled to a Protein A chip via its Fc region, ensuring correct orientation and binding site availability on the chip surface. DBLβ domains were injected over the chip surface in a concentration series (flow channel 2, shown for J1a$^{DBL\beta}$ in Figure 5.9B). ICAM-1^{D1D5} was removed after each DBLβ concentration by low pH, regenerating the chip before an equivalent amount of ICAM-1^{D1D5} was

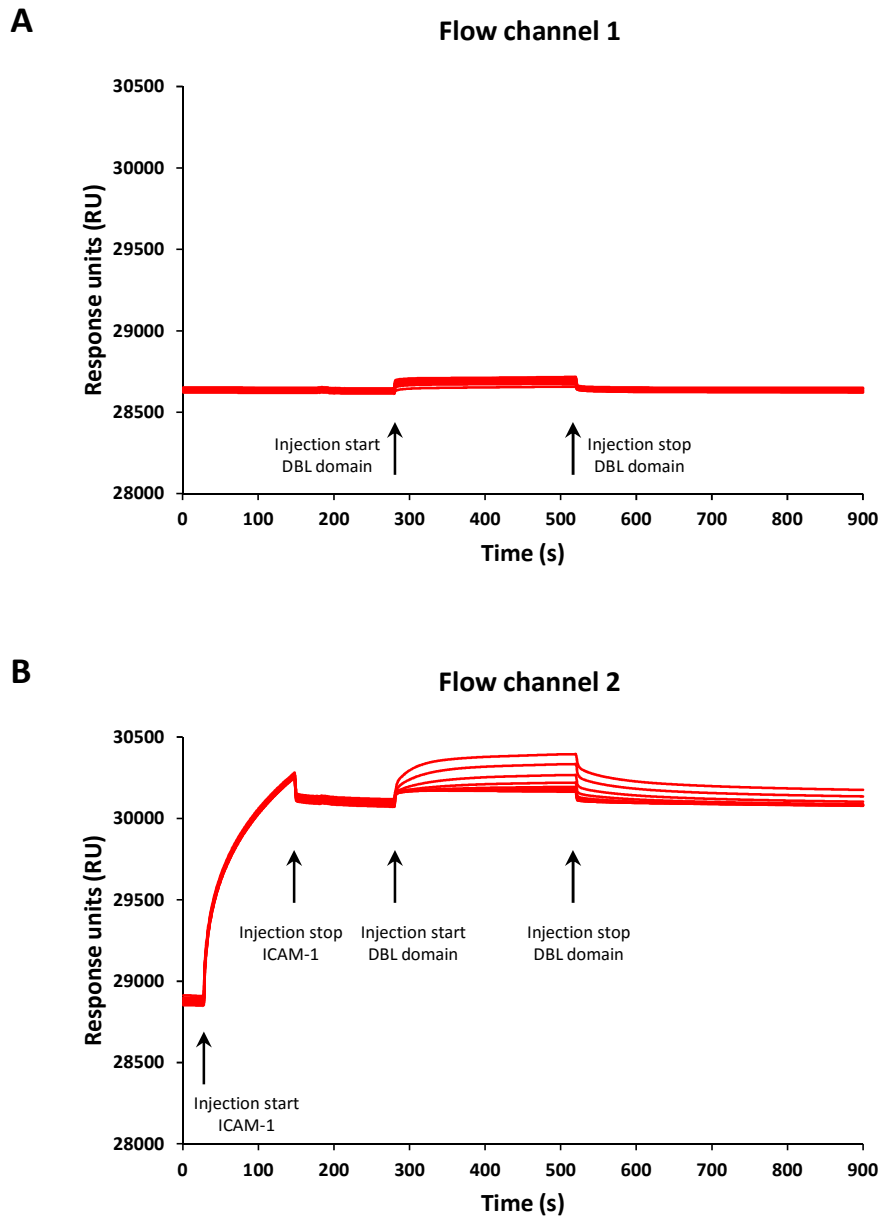


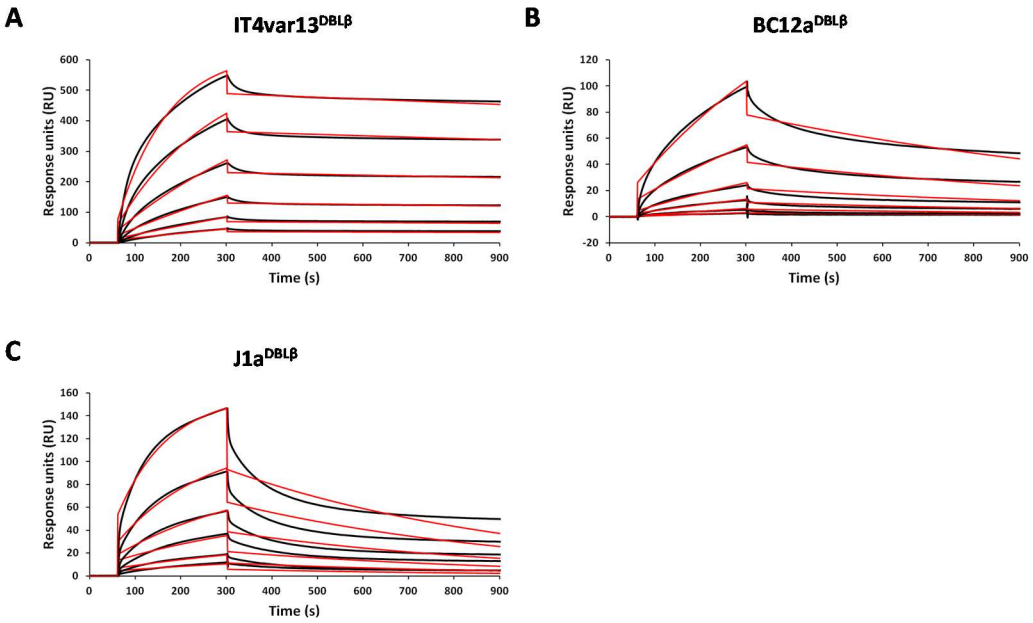
Figure 5.9. J1a SPR data from flow channels 1 and 2 showing that binding is specific to ICAM-1^{D1D5}.

A, J1a$^{DBL\beta}$ is injected over flow channel 1 in the absence of ICAM-1-^{D1D5}. There is a slight response due to a small change in the buffer containing the DBLβ domain which remains constant throughout the injection time (indicated with arrows). Multiple red lines represent the concentration series of J1a$^{DBL\beta}$ (see Methods 5.2.6 for details). **B**, ICAM-1^{D1D5} is injected over flow channel 2 where it binds throughout the injection time, indicated by the increase in response units (RU). RU remains constant until injection of the DBLβ domain, indicating stable binding of ICAM-1^{D1D5} to the channel. A concentration dependant response is observed when J1a$^{DBL\beta}$ is injected over flow channel 2, indicating specific binding to ICAM-1^{D1D5}. Flow channel

1 RU is subtracted from flow channel 2 RU to adjust for any background binding before analysis.

captured for the next concentration. DBLβ domains were also injected over a channel without ICAM-1 to test for unspecific binding (flow channel 1, see Figure 5.9A and legend for details). Data were fitted globally to a 1:1 interaction model based on previous studies of DBLβ::ICAM-1 interaction by SPR (Brown *et al.*, 2013, Lennartz *et al.*, 2015).

We performed SPR on the known ICAM-1^{D1D5} binder, IT4var13$^{DBL\beta}$, for comparison and as an experimental control. Binding kinetics of IT4var13$^{DBL\beta}$ to ICAM-1^{D1D5} display a fast association rate and a slow dissociation rate, with a K_D of 3.45 nM (Figure 5.10A, D), comparable to published values for this domain (Brown *et al.*, 2013). BC12a$^{DBL\beta}$ has a slow association rate and relatively fast dissociation rate, with the overall response much lower than IT4var13$^{DBL\beta}$ (100 RU compared to 550 RU, Figure 5.10B and A, respectively). BC12a$^{DBL\beta}$ has a high K_D of 197 nM indicating a low affinity for binding ICAM-1^{D1D5}. J1a$^{DBL\beta}$ also has an overall low response of 150 RU but has very interesting kinetics which are apparent in the shape of the sensorgram in Figure 5.10C. J1a$^{DBL\beta}$ association rate is fast and comparable to IT4var13$^{DBL\beta}$ at 2.96 X 10^4 M^{-1} s^{-1} but it has a very fast dissociation rate of 15.3 X 10^{-4} s^{-1} resulting in a higher K_D of 51.7 nM (Figure 5.10D).



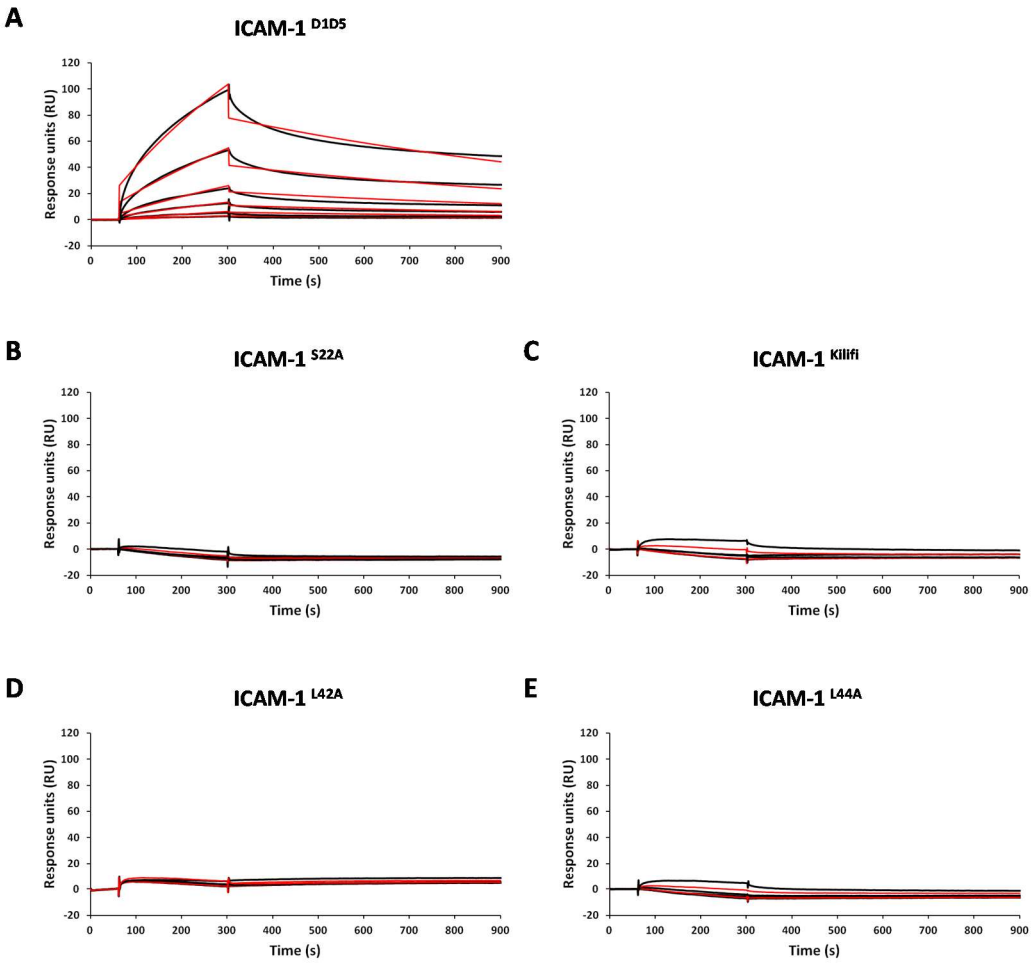
D

Interaction	k_a (X 10^4 M^{-1} s^{-1})	k_d (X 10^{-4} s^{-1})	K_D (nM)
ICAM-1 ^{D1D5}:: IT4var13^{DBLβ}	3.61	1.25	3.45
ICAM-1 ^{D1D5}:: BC12a^{DBLβ}	0.48	9.44	197
ICAM-1 ^{D1D5}:: J1a^{DBLβ}	2.96	15.3	51.7

Figure 5.10. DBLβ:ICAM-1^{D1D5} binding kinetics.
ICAM-1^{D1D5} was coupled to a sensor chip surface (1000 RU) and DBLβ domains were injected at 30 µl/min with an association time of 240 seconds and a dissociation time of 600 seconds. Shown are sensorgrams for the binding to ICAM-1^{D1D5} of IT4var13^{DBLβ} (**A**), BC12a^{DBLβ} (**B**) and J1a^{DBLβ} (**C**). Data (black lines) are modelled to a 1:1 global interaction model (red lines). Note the difference in scale of the y-axes. **D**, Kinetic parameters derived from SPR experiments.

DBLβ domains tested for binding to ICAM-1 mutants

We used SPR to test binding of recombinant BC12a^{DBLβ}, J1a^{DBLβ} and IT4var13^{DBLβ} to four ICAM-1 mutant proteins that have been shown to have differential effects on binding of different parasite isolates (Tse *et al.*, 2004, Madkhali *et al.*, 2014). ICAM-1^{S22A}, ICAM-1^{Kilifi}, ICAM-1^{L42A} and ICAM-1^{L44A} are all produced as Fc fusion proteins and were captured onto a Protein A chip following the protocol for ICAM-1^{D1D5} SPR experiments. Each of the four ICAM-1 mutations obliterated binding of BC12a^{DBLβ} with only minimal signal detected (Figure 5.11B-E. A, binding to ICAM-1^{D1D5}

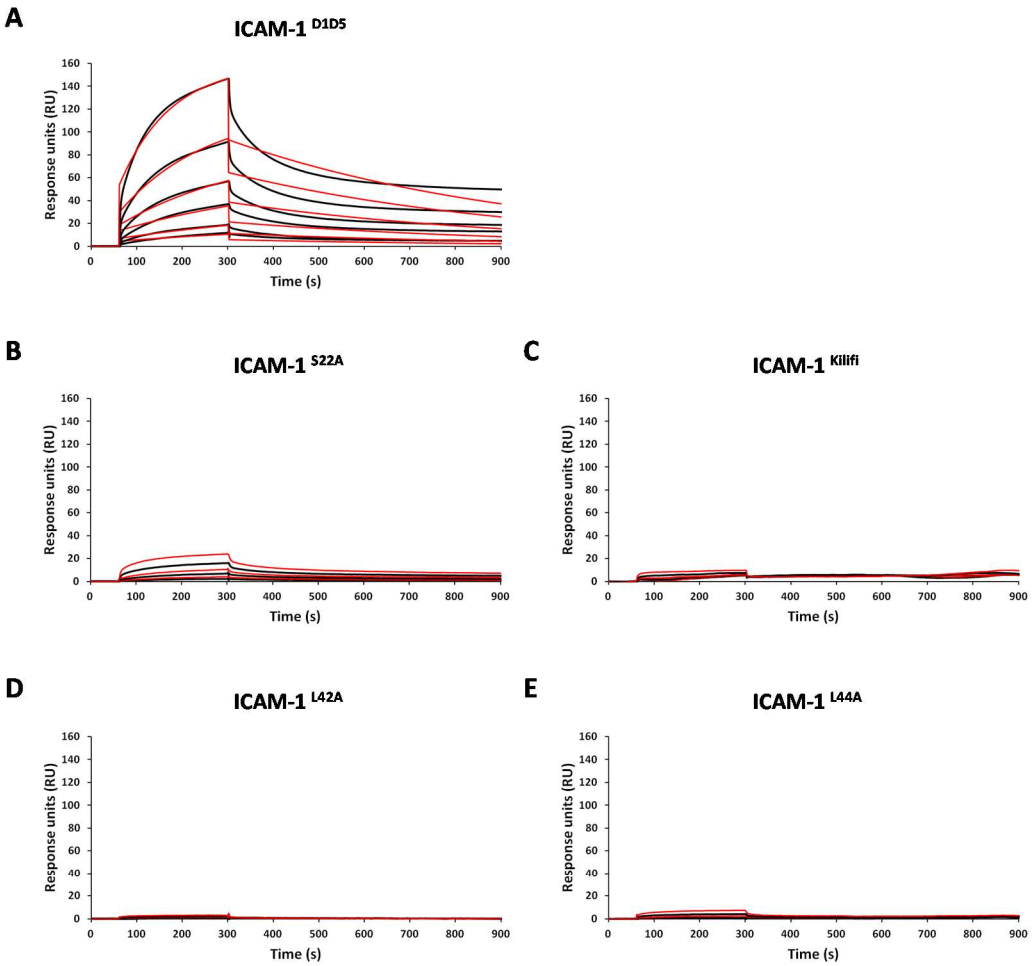


F

	ICAM-1 ^{S22A}	ICAM-1 ^{Kilifi}	ICAM-1 ^{L42A}	ICAM-1 ^{L44A}
% ICAM-1 ^{D1D5} binding	2.9	7.7	6.7	6.7

Figure 5.11. BC12a$^{DBL\beta}$ response to ICAM-1 mutant proteins.
ICAM-1 proteins were coupled to a sensor chip surface (1000 RU) and the DBLβ domain injected at 30 µl/min with an association time of 240 seconds and a dissociation time of 600 seconds. Shown are sensorgrams for the binding of BC12a$^{DBL\beta}$ to ICAM-1^{D1D5} (**A**), ICAM-1^{S22A} (**B**), ICAM-1^{Kilifi} (**C**), ICAM-1^{L42A} (**D**) and ICAM-1^{L44A} (**E**). Data (black lines) are modelled to a 1:1 global interaction model (red lines). The y-axis scale is kept constant for comparison. **F**, Response units measured to each ICAM-1 mutant protein expressed as a percentage of wild type ICAM-1^{D1D5} binding.

included for comparison). Binding to all mutants is < 8% of the ICAM-1^{D1D5} binding (Figure 5.11F), although it should be noted that signals are close to background



F

	ICAM-1 ^{S22A}	ICAM-1 ^{Kilifi}	ICAM-1 ^{L42A}	ICAM-1 ^{L44A}
% ICAM-1 ^{D1D5} binding	10.9	5.4	1.4	2.7

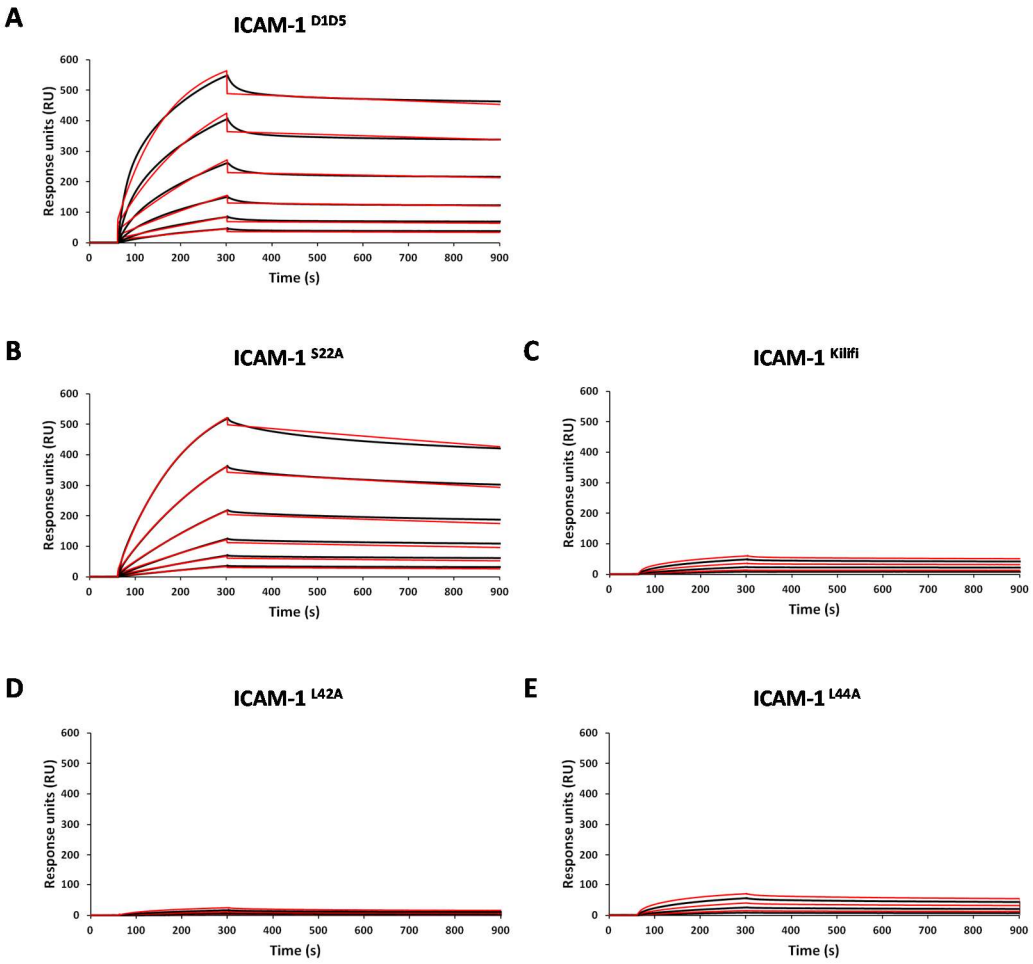
Figure 5.12. J1a$^{DBL\beta}$ response to ICAM-1 mutant proteins.
ICAM-1 proteins were coupled to a sensor chip surface (1000 RU) and the DBLβ domain injected at 30 µl/min with an association time of 240 seconds and a dissociation time of 600 seconds. Shown are sensorgrams for the binding of J1a$^{DBL\beta}$ to ICAM-1^{D1D5} (**A**), ICAM-1^{S22A} (**B**), ICAM-1^{Kilifi} (**C**), ICAM-1^{L42A} (**D**) and ICAM-1^{L44A} (**E**). Data (black lines) are modelled to a 1:1 global interaction model (red lines). The y-axis scale is kept constant for comparison. **F**, Response units measured to each ICAM-1 mutant protein expressed as a percentage of wild type ICAM-1^{D1D5} binding.

noise levels and, therefore, may not be accurate and are presented as representative values only.

Similarly, J1a$^{DBL\beta}$ only minimally interacted with the ICAM-1 mutant proteins (Figure 5.12B-E). Binding to ICAM-1^{Kilifi}, ICAM-1^{L42A} and ICAM-1^{L44A} is ≤ 5% of wild type ICAM-1^{D1D5} binding. ICAM-1^{S22A} binding is slightly higher at 11% of wild type binding (Figure 5.12F) but is still very low.

IT4var13$^{DBL\beta}$ also had minimal interaction with ICAM-1^{Kilifi}, ICAM-1^{L42A} and ICAM-1^{L44A} which is ≤ 10% of ICAM-1^{D1D5} binding (Figure 5.13C-F). However, IT4var13$^{DBL\beta}$ did bind to ICAM-1^{S22A} with 9.82 nM affinity (carried out by Frank Lennartz, University of Oxford, Figure 5.13B). A repeat of this experiment resulted in a similar sensorgram profile and a K_D of 9.96 nM, confirming the result (Frank Lennartz, data not shown). Therefore, the S22A mutation decreases IT4var13$^{DBL\beta}$ binding affinity for ICAM-1 but only by 5% (Figure 5.13F). This is a minor effect compared to BC12a$^{DBL\beta}$ and J1a$^{DBL\beta}$ for which binding is reduced by ≥ 90% (Figure 5.11B and Figure 5.12B, respectively), suggesting that this residue is not critical for IT4var13$^{DBL\beta}$ binding.

Taken together, the results of the mutant ICAM-1 binding assays suggest that the DBLβ domains share an overlapping but not identical binding surface in the area of the mutations. The Kilifi, L42A and L44A mutations dramatically reduce the binding ability of all three DBLβ domains. The S22A mutation has a similarly dramatic effect on BC12a$^{DBL\beta}$ and J1a$^{DBL\beta}$ binding but has minimal effect on IT4var13$^{DBL\beta}$ binding, suggesting a specific difference in contact residues in this area.



F

	ICAM-1 ^{S22A}	ICAM-1 ^{Kilifi}	ICAM-1 ^{L42A}	ICAM-1 ^{L44A}
% ICAM-1 ^{D1D5} binding	94.5	8.9	3.1	10

Figure 5.13. IT4var13$^{DBL\beta}$ response to ICAM-1 mutant proteins.
ICAM-1 proteins were coupled to a sensor chip surface (1000 RU) and the DBLβ domain injected at 30 µl/min with an association time of 240 seconds and a dissociation time of 600 seconds. Shown are sensorgrams for the binding of IT4var13$^{DBL\beta}$ to ICAM-1^{D1D5} (**A**), ICAM-1^{S22A} (**B**), ICAM-1^{Kilifi} (**C**), ICAM-1^{L42A} (**D**) and ICAM-1^{L44A} (**E**). Data (black lines) are modelled to a 1:1 global interaction model (red lines). The y-axis scale is kept constant for comparison. **F**, Response units measured to each ICAM-1 mutant protein expressed as a percentage of wild type ICAM-1^{D1D5} binding.

Competition assays confirm DBLβ domains share an overlapping binding surface

In addition to the mutant ICAM-1 data, we performed a competition SPR assay to confirm that the DBLβ domains share a binding surface. One DBLβ domain is injected over an ICAM-1^{D1D5} coated chip, followed immediately by a second DBLβ domain with the expectation that if the domains have separate binding sites, an additive response will be observed. Conversely, if the domains share a binding surface, the response of the second domain will not be higher than that of the first. We performed this experiment with the injection of IT4var$13^{DBLβ}$ followed by BC12a^{DBLβ} and, in a separate experiment, with the injection of IT4var$41^{DBLβ}$ followed by J1a^{DBLβ}. In each case, the injection of the second domain did not result in additional response units to that of the first domain (Figure 5.14), suggesting that the overall binding surface is shared between IT4var$13^{DBLβ}$ and BC12a^{DBLβ} and between IT4var$41^{DBLβ}$ and J1a^{DBLβ}. However, it should be noted that this technique will not pick up subtle differences in specific amino acid contact residues which do exist due to the high sequence variation between DBLβ domains. Therefore, we can conclude only that the overall binding surface is conserved between these domains.

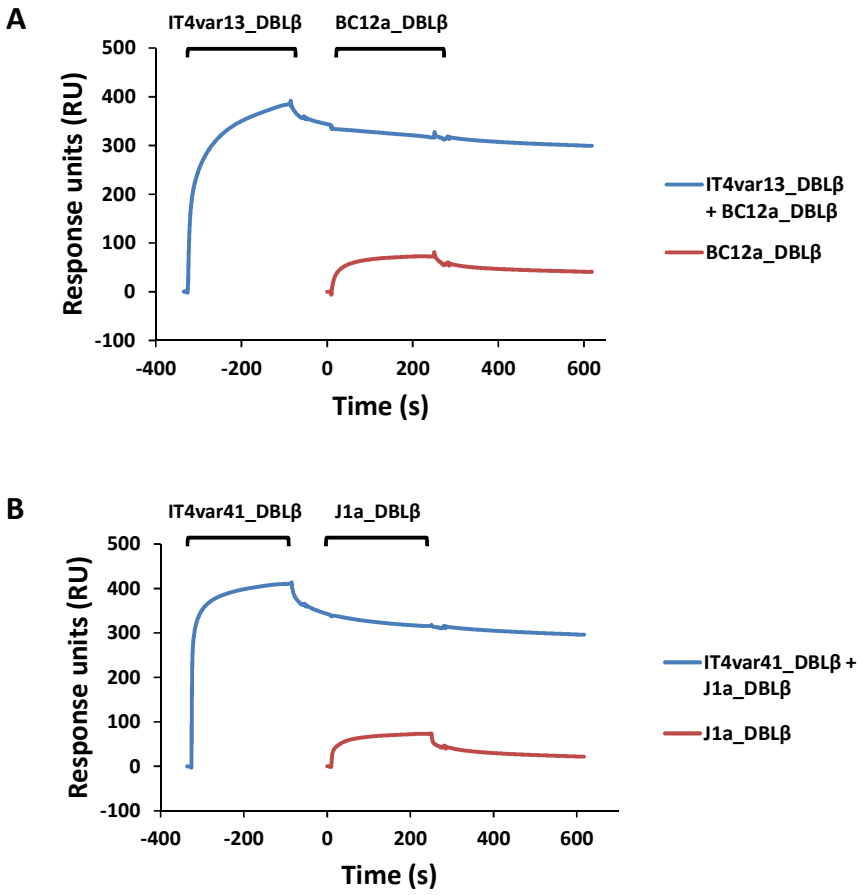


Figure 5.14. Competition assays confirm DBLβ domains share an overlapping binding site.

ICAM-1^{D1D5} protein was coupled to a sensor chip surface (1000 RU) and the first DBLβ domain injected at 30 µl/min with an association time of 240 seconds (indicated above the graph with bracket beginning at time -336 s), followed by injection of the second domain at 30 µl/min with an association time of 240 seconds (indicated above the graph with bracket beginning at time 0) and a dissociation time of 600 seconds. Shown are sensorgrams for the binding of **A**, IT4var13^{DBLβ} followed by BC12a^{DBLβ} to ICAM-1^{D1D5} and **B**, IT4var41^{DBLβ} followed by J1a^{DBLβ} to ICAM-1^{D1D5}, blue lines. The response to the second domain alone is shown in red.

5.3.6 Flow adhesion assay

The initial parasite binding study (Madkhali *et al.*, 2014) assessed the ability of different isolates to bind to ICAM-1 and the mutant ICAM-1 proteins in static adhesion assays (see discussion in 5.4 below for comparison). Our SPR experiments

with recombinant protein occur in channels with a constant flow of buffer. Although there are many differences between whole parasite and recombinant protein assays, we wanted to test parasite binding to the mutant ICAM-1 proteins under flow adhesion assay conditions which may be more comparable to those of SPR than the static adhesion assays. Results of parasite adhesion to ICAM-1^{D1D5} and the four ICAM-1 mutant proteins are presented in Figure 5.15A. Interestingly, BC12 and J1 parasites bind more strongly to the wild type ICAM-1^{D1D5} than IT4var13 parasites. In the majority of cases, the ICAM-1 mutations resulted in a 16-56% reduction in binding of all three parasite lines. Exceptions are the L42A and L44A mutations. ICAM-1^{L42A} resulted in a 78 and 76% reduction of J1 and IT4var13 binding, respectively, and ICAM-1^{L44A} resulted in an 88% reduction of IT4var13 binding (Figure 5.15A). RT-qPCR confirmed the dominantly expressed *var* gene of BC12 to be BC12a and IT4var13 to be *var13* (Figure 5.15B). The dominant *var* of J1 is J1a which represents 62% of expressed *vars* with a number of other *var* genes making up the remaining 38% (Figure 5.15B). This includes the previously identified secondary tags J1d and J1b and also includes J1f which is represented at a higher level than previously observed (Figure 3.3).

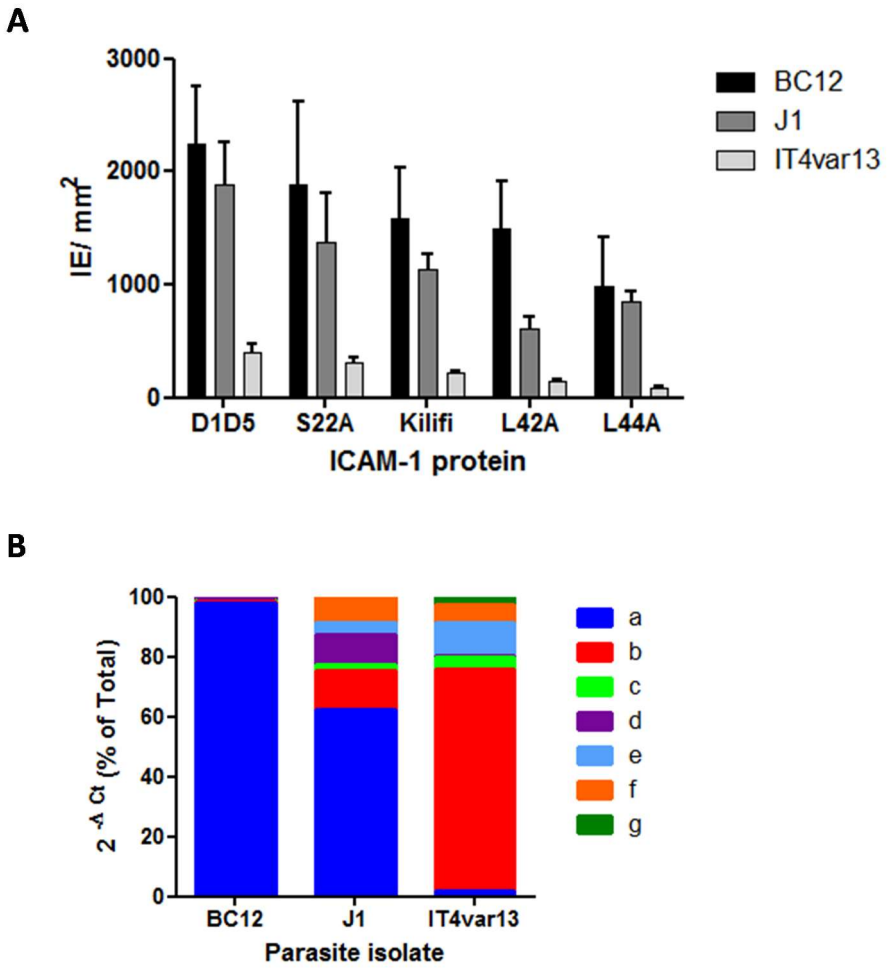


Figure 5.15. BC12, J1 and IT4var13 flow adhesion assays to ICAM-1 proteins. **A,** Channels were coated with 4 µl of 50 µg/ml ICAM-1^{D1D5}. Recently ICAM-1-selected IEs were passed through the channel at 3% parasitaemia, 2% haematocrit for 8 minutes. Bound parasites were counted in 7-10 fields and adjusted to IE/mm^2. Bars represent the mean of 3 (BC12, J1) or 4 (IT4var13) independent experiments with the standard error of the mean (SEM) shown (error bars). **B**, RT-qPCR of cDNA from each parasite culture. RT-qPCR was carried out using BC12 and J1 DBLα tag primers (a-e and a-f, respectively) and primers to IT4 ICAM-1 binding *var* genes (a: 01, b:13, c:14, d:16, e:27, f:31, g:41). Ct values were normalised against the ASL internal control gene to give $2^{-\Delta Ct}$ values and are shown as percentage of total for each isolate. The mean of 3 (BC12, J1) or 4 (IT4var13) independent experiments are shown.

5.4 Discussion

Here, we present data detailing the expression and purification of two newly identified DBLβ domains from the ICAM-1 selected patient isolates BC12 and J1. We have obtained kinetic information on their interaction with ICAM-1^{D1D5}, along with the previously studied DBLβ domain of IT4var13, and tested the binding of all three recombinant DBLβ domains to four ICAM-1 mutant proteins that have been previously shown to have differential effects on parasite binding assays (Tse *et al.*, 2004, Madkhali *et al.*, 2014). Preliminary data of flow adhesion assays to the wild type ICAM-1^{D1D5} and the four ICAM-1 mutants are also presented.

We successfully expressed and purified two of the five newly identified DBLβ domains to > 90% purity. BC12a^{DBLβ} and J1a^{DBLβ} are from the dominantly expressed *var* genes of the ICAM-1 selected patient isolates BC12 and J1, respectively. The remaining DBLβ domains are from the *var* genes J1d, PCM7a and PCM7d. We were unable to generate a plasmid containing the correct PCM7d^{DBLβ} sequence due to time constraints. J1d and PCM7a DBLβ domains appeared highly sensitive to degradation (Figure 5.2D and Figure 5.5, respectively). Alternative approaches to expression should be tested to improve stability of these domains. This could include testing of expression in alternative bacterial strains or expression in insect cell lines which has previously been successful for PfEMP1 domains (Brown *et al.*, 2013, Lau *et al.*, 2015). A possible explanation for failure to express *P. falciparum* proteins in *E. coli* is codon bias. This has previously been overcome by codon optimisation and subsequent synthesis of the resulting synthetic gene for the expression of the full length *var2csa* gene (Khunrae *et al.*, 2010). The current study was limited by time constraints and these strategies could not be pursued. However, the differential behaviour of these domains of the same type, and the same subtype in the case of BC12a, J1a, and J1d (DBLβ5), highlights the diversity of the PfEMP1 domains, reflected in their low sequence identity. It is unfortunate that neither J1d^{DBLβ} nor PCM7d^{DBLβ} could be expressed to test whether these secondary expressed genes can truly bind to ICAM-1. However, successful expression of

BC12a$^{DBL\beta}$ and J1a$^{DBL\beta}$ proteins allowed us to test two new DBLβ domains for ICAM-1 binding by SPR, along with the previously tested IT4var13$^{DBL\beta}$.

BC12a$^{DBL\beta}$, J1a$^{DBL\beta}$ and IT4var13$^{DBL\beta}$ all bind ICAM-1^{D1D5} with nanomolar affinity. IT4var13$^{DBL\beta}$ has the strongest interaction, followed by J1a$^{DBL\beta}$ and BC12a$^{DBL\beta}$ (K_D values: 3.45, 51.7, 197 nM, respectively, Figure 5.10). These K_D values are similar to those found for five IT4 DBLβ domains, tested in comparable SPR assays, which ranged from 2.6 nM for IT4var13$^{DBL\beta}$ (comparable to our 3.45 nM value for this domain) to 144 nM for IT4var31$^{DBL\beta}$ (Brown *et al.*, 2013). The UPS A type, DC4 DBLβ domain was also found to bind ICAM-1 with nanomolar affinity (7.9 nM) in an SPR assay where the DBLβ domain was immobilised on the chip and ICAM-1^{D1D2} passed over (Lennartz *et al.*, 2015). All three domains studied here are DBLβ5 type and these findings fit with previous observations that, to date, all domains of this type mediate ICAM-1 binding (Janes *et al.*, 2011). The difference in binding kinetics between IT4var13$^{DBL\beta}$ and J1a$^{DBL\beta}$ is very interesting. They have a similar fast association rate which can be thought of as necessary to overcome the flow rate of the capillaries to mediate IE binding. Indeed, ICAM-1 has been shown to improve efficiency of IE capture under flow conditions (Gray *et al.*, 2003). However, they differ in their dissociation rates. IT4var13$^{DBL\beta}$ has a slow dissociation rate which fits with stationary adhesion of the IE. J1a$^{DBL\beta}$, however, has a very fast dissociation rate which suggests that this domain could mediate rolling adhesion, in which the IE constantly attaches and detaches from the capillary wall in a rolling motion (Cooke *et al.*, 1994, Gray *et al.*, 2003).

The K_D value is calculated based on the association and dissociation rates and is therefore based on the shape of the sensorgram rather than the overall height of the response. In this case, the kinetic parameters fit with the response units recorded for each domain, i.e. the highest response was by IT4var13$^{DBL\beta}$, followed by J1a$^{DBL\beta}$ then BC12a$^{DBL\beta}$. The low responses recorded for J1a$^{DBL\beta}$ and BC12a$^{DBL\beta}$ are in agreement with the initial parasite binding study which found both parasite isolates to be low avidity ICAM-1-binders, with BC12 having a lower binding rate than J1 (1362 and 2391 IE/mm^2, respectively) (Madkhali *et al.*, 2014). Interestingly, the IT4var13 parasite line has a lower binding rate (900 IE/mm^2) than both BC12

and J1 in static adhesion assays (Madkhali, 2015), a result that is also apparent in our preliminary flow adhesion assays (Figure 5.15), highlighting differential results between IEs and recombinant protein. Great care must be taken when comparing data from, not only two different experimental techniques, but two different components, namely whole IEs and recombinant protein domains. There is no clear correlation between K_D values of DBLβ domains and the number of bound IEs in static adhesion assays. For example, IT4var16^{DBLβ} has the same K_D as J1a^{DBLβ} (51.1 and 51.7 nM, respectively) but parasites expressing *var16* have more than twice the binding rate of J1 parasites in static assays (5000 and 2391 IE/mm^2, respectively) (Brown *et al.*, 2013, Madkhali *et al.*, 2014, Madkhali, 2015). For further discussion of differential results between assays see pages 131 and 132 below.

The DBLβ binding site on ICAM-1 has been identified by mutagenesis studies and a number of ICAM-1 blocking antibodies as the BED side of D1 (Tse *et al.*, 2004, Berendt *et al.*, 1992, Ockenhouse *et al.*, 1992a). The four ICAM-1 mutants utilized in this study have been shown to have differential effects on the binding of different parasite strains (Tse *et al.*, 2004, Madkhali *et al.*, 2014). We studied the effects of these mutations on the ability of the recombinant DBLβ domains to bind ICAM-1 using SPR assays. The Kilifi (K29M), L42A and L44A mutations dramatically reduce the binding ability of all three recombinant DBLβ domains by ≥ 90%. The S22A mutation similarly reduces BC12a^{DBLβ} and J1a^{DBLβ} binding by ≥ 90% but has minimal effect on IT4var13^{DBLβ} binding, suggesting a specific difference in contact residues in this area (Figure 5.11, Figure 5.12, Figure 5.13). This reduction of recombinant protein binding is more marked than that observed in parasite adhesion assays (summarised in Figure 5.16). BC12 parasite static adhesion studies found a 43-66% reduction in binding to ICAM-1^{S22A}, ICAM-1^{Kilifi} and ICAM-1^{L42A} and actually saw an increase to 147% of ICAM-1^{D1D5} binding to ICAM-1^{L44A} (Madkhali *et al.*, 2014). J1 IE static binding to ICAM-1^{S22A}, ICAM-1^{Kilifi} and ICAM-1^{L44A} was 43, 58 and 86% of wild type binding, respectively, with only ICAM-1^{L42A} showing a reduction in binding comparable to the recombinant protein to 4% of wild type in static assays (Madkhali *et al.*, 2014). Indeed, reductions in binding to ICAM-1^{L42A} are the most consistent results across all assays (Figure 5.16). ICAM-1^{S22A} had a major effect on

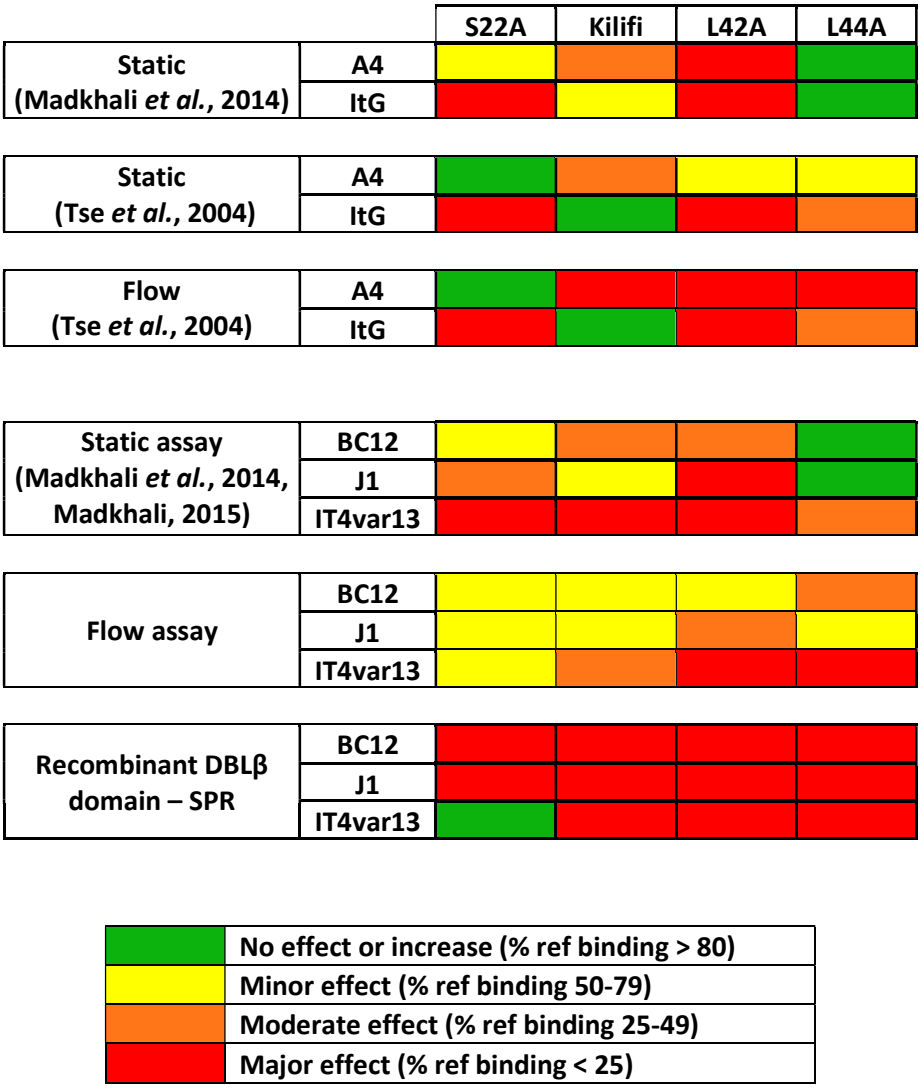


Figure 5.16. Summary of ICAM-1 mutant binding data across assays.
The effects of the ICAM-1 mutant proteins on binding are presented as a percentage of reference ICAM-1^{D1D5} binding. The results of the present study and previous studies (Tse *et al.*, 2004, Madkhali *et al.*, 2014, Madkhali, 2015) are summarised. The assay type is indicated.

IT4var13 binding in static assays, a minor effect in our preliminary flow assays and no effect on recombinant protein, as discussed. Any explanation for these differences is speculative and likely relates to assay set-up and the components being tested (discussed further below). Tse *et al.* (2004) proposed that any negative effect of the mutant proteins on binding is amplified in flow assays due to a greater strain on the interaction under flow and this is in keeping with their data on A4 and

ItG (Figure 5.16). However, different results were obtained for A4 and ItG static binding to L42A and L44A by Madkhali (2015), which brings into question the reproducibility of IE binding assays. It is worth noting that the expressed *var* genes of A4 and ItG were not confirmed in either study but are presumed to be predominantly IT4var14 and IT4var16, respectively. This raises the possibility that the expression of other genes within the culture may be influencing the outcome of binding assays and these could be variable between studies.

There are a number of factors to consider when comparing such data sets. The SPR assays allow us to study the specific protein-protein interaction of different DBLβ domains to ICAM-1, utilising the mutant ICAM-1 proteins to identify specific contact residues. This is a highly sensitive, automated technique. However, both proteins are taken out of the context of the full length protein which is membrane bound on the surface of a cell with numerous additional components. Static adhesion assays are performed by panning IEs over fixed recombinant ICAM-1, allowing testing of DBLβ domain binding in the context of the full length PfEMP1 on its native cell surface. However, assays rely on technical consistency and many replicates are required to reduce the standard error. Similarly, flow adhesion assays test binding of whole IEs to recombinant ICAM-1 but have the advantage of mimicking conditions of flow in the capillaries and are thought to be more physiologically relevant.

A potential area of variability in IE studies is the amount of PfEMP1 exported to the cell surface. This is inherently difficult to measure due to the highly variable nature of these proteins and, therefore, a lack of a universal PfEMP1 antibody. One method of detecting PfEMP1 is by Western blot using an antibody against the relatively conserved ATS region (Waterkeyn *et al.*, 2000, Maier *et al.*, 2007). Our study measures *var* gene expression to identify an isolate's dominantly expressed gene with an assumed correlation to PfEMP1 abundance but, with the exception of VAR2CSA (Bancells and Deitsch, 2013), post-transcriptional control is poorly understood. In addition, global *var* gene transcription in known to decrease upon parasite adaptation to culture (Peters *et al.*, 2007), which may affect surface protein abundance. The absence of knobs has also been shown to affect PfEMP1

expression (Horrocks *et al.*, 2005). However, all parasites in this study are amenable to Plasmagel floatation, indicating the presence of knobs.

Both static and flow assays spot ICAM-1 Fc fusion protein onto plastic surfaces in an unknown orientation, presumably with a consistent number of binding sites available between assays to allow comparison. However, SPR assays capture ICAM-1 via the Fc receptor, mediated by Protein A. Future studies should test the spotting of Protein A into flow channels, followed by ICAM-1 capture in an attempt to make SPR and flow assays more comparable. Recombinant ICAM-1 forms dimers via the Fc region (Bengtsson *et al.*, 2013a), the effect of which on parasite binding is unknown. ICAM-1 can also form dimers mediated by D4::D4 (Yang *et al.*, 2004) and D1::D1 (Casasnovas *et al.*, 1998) binding and has been suggested to exist in an intermediate monomeric-dimeric state on the endothelial cell surface (Oh *et al.*, 2011). However, the native state of ICAM-1 on the cell surface is unknown in the absence of direct visualisation. D1::D1 dimers form via the same area involved in DBLβ binding, the BED side, but the effect on DBLβ binding is unknown, as is the *in vivo* relevance of this dimer. Therefore, flow adhesion assays which test IE binding to induced endothelial cells (Gray *et al.*, 2003, Madkhali *et al.*, 2014) are perhaps the most physiologically relevant assays available since both proteins are in the context of their native cell. However, these assays are technically very difficult due to the need to combine two independently challenging cell cultures. Each assay provides important information and a combination of results obtained from all assays should be considered to build a picture of ICAM-1 binding. For example, IE binding to recombinant ICAM-1 informs us of their binding ability, protein-protein interactions are useful for analysing the kinetics of the binding and the sites involved, and IE binding to endothelial cells, in combination with receptor blocking antibodies, provide information on the contribution of different host receptors to overall IE binding.

Our SPR results show that J1a, BC12a and IT4var13 DBLβ domains bind ICAM-1^{D1D5} with nanomolar affinity, in agreement with previous studies (Brown *et al.*, 2013, Lennartz *et al.*, 2015). This has important implications for the design of antibody therapies which would require higher affinities to outcompete the DBLβ::ICAM-1

interaction. Binding to the ICAM-1 mutant proteins indicates subtle differences in contact residues between the different DBLβ domains. When the surface architecture of ICAM-1 is considered (Figure 5.17), the S22 residue is near the edge of the area identified as important for DBLβ binding. It is likely that IT4var13$^{DBL\beta}$ has a structural difference in the area that contacts here, resulting in the lack of effect of this mutation. This has implications for development of therapeutics such as drugs or antibody therapy, which should focus around the L42 and L44 residues and possibly extend to K29 (mutated in ICAM-1^{Kilifi}) to confer a cross-blocking effect. A co-crystal is required to fully understand the DBLβ::ICAM-1 bound structure and further guide the design of cross-acting therapeutics (discussed in Chapter 6). The difference in contact residues between DBLβ domains is also important when considering anti-disease vaccine development targeting ICAM-1 binding DBLβ domains.

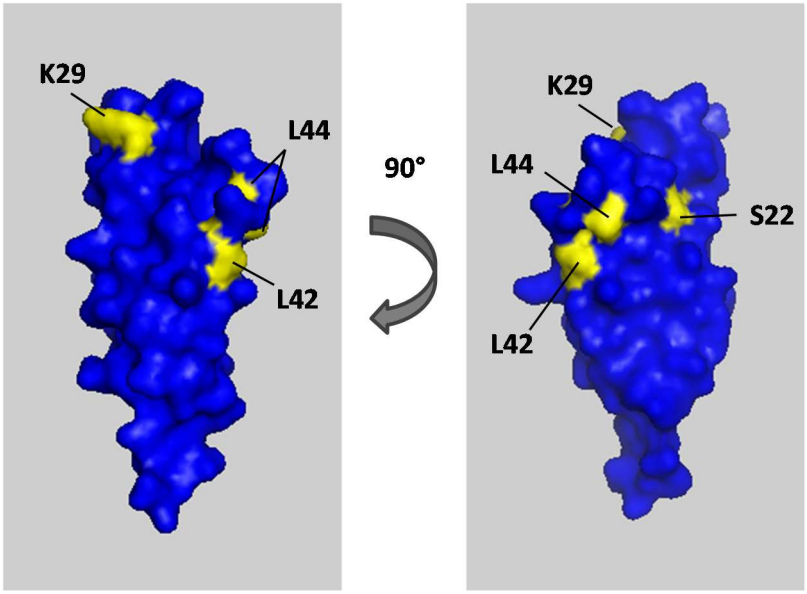


Figure 5.17. The crystal structure of ICAM-1 D1, showing the surface architecture. Residues of the mutated proteins tested in this study are labelled and coloured yellow. K29 residue is mutated to M in ICAM-1^{Kilifi}. Structure accessed via Protein Data Bank, deposited by Casasnovas *et al.* (1998).

5.5 Summary

In this chapter, we analysed the binding of two newly identified DBLβ domains to ICAM-1. Using SPR assays, we show that the recombinant DBLβ domains of J1a and BC12a bind ICAM-1^{D1D5} with nanomolar affinity. We compared their binding to the previously characterised IT4var$13^{DBL\beta}$ and studied the binding of all three DBLβ domains to four ICAM-1 mutant proteins. Residues K29, L42 and L44 are essential for ICAM-1 binding of all three DBLβ domains, whereas S22 is essential for J1a$^{DBL\beta}$ and BC12a$^{DBL\beta}$ binding but not IT4var$13^{DBL\beta}$ binding. This has important implications for the design of therapies targeting severe malaria, topics discussed in the final Chapter 6.

Chapter 6. Conclusions and implications for therapeutic targeting of PfEMP1

An estimated 438,000 deaths worldwide were attributed to severe malaria in 2015 (WHO, 2015). A key process in the development of SM is the sequestration of *P. falciparum* IE in the microvasculature which involves the interaction of numerous parasite and host factors. The most extensively studied interaction is that of the highly variable PfEMP1 protein family with various endothelial cell receptors. Evidence that SM is caused by a subset of PfEMP1 variants has led to hope of an anti-disease therapeutic targeting these interactions. The need for such therapeutic adjuncts to current drug treatment is evidenced in the ~ 10% mortality rate among SM patients in spite of effective anti-parasite treatment (Dondorp *et al.*, 2010). Extensive understanding of the mechanisms involved in IE sequestration is required to inform on the feasibility of adjunct therapies and vaccines targeted against PfEMP1 proteins.

The current study has extended our understanding of PfEMP1 interactions with the host endothelial receptor ICAM-1. The majority of studies of ICAM-1 binding have focussed on the parasite reference strains IT4 and 3D7 (Howell *et al.*, 2008, Janes *et al.*, 2011, Brown *et al.*, 2013, Bengtsson *et al.*, 2013b). Recent work characterised the ICAM-1 binding ability of a panel of ICAM-1-selected culture-adapted patient isolates, conducted at the cellular level of IEs (Madkhali *et al.*, 2014). The current study extends that characterisation to the molecular level, with the aim of identifying novel ICAM-1 binding PfEMP1 proteins and providing insight into a broader range of DBLβ::ICAM-1 interactions. We have identified and sequenced five new predicted ICAM-1 binding PfEMP1 proteins from three patient isolates, BC12, J1 and PCM7. Validation of ICAM-1 binding ability was carried out for two of these PfEMP1 proteins (BC12a and J1a) by recombinantly expressing their DBLβ domains and demonstrating ICAM-1^{D1D5} binding by SPR experiments. SPR experiments with the mutant ICAM-1 proteins ICAM-1^{S22A}, ICAM-1^{Kilifi}, ICAM-1^{L42A}

and ICAM-1^{L44A} showed an almost complete loss of binding of both BC12a$^{DBL\beta}$ and J1a$^{DBL\beta}$, confirming the importance of these residues in binding. The IT4var13$^{DBL\beta}$ domain bound ICAM-1^{D1D5} with high affinity, as previously described (Brown *et al.*, 2013), did not bind ICAM-1^{Kilifi}, ICAM-1^{L42A} or ICAM-1^{L44A}, in a similar result to BC12a$^{DBL\beta}$ and J1a$^{DBL\beta}$ domains, but did retain binding ability to ICAM-1^{S22A}. These results suggest overlapping but distinct ICAM-1 binding sites.

A key component of this study was to test the practical application of the new Pf3k *var* gene database in a targeted sequencing approach. The global *var* gene repertoire was once considered endless but, with advances in WGS methods, it is becoming apparent that the organisation of PfEMP1 proteins into DCs (Rask *et al.*, 2010) and the occurrence of recombination mainly within groups (Kraemer and Smith, 2003, Kraemer *et al.*, 2007, Claessens *et al.*, 2014), results in limitations on the extent of diversity compared to a situation of completely random domain associations. Therefore, if enough parasites genomes are sequenced globally, it is feasible that we may have access to the majority of possible *var* sequences. A number of the non-dominant DBLα tags queried against the database did not yield any matches, suggesting that this has yet to be achieved. Our limited analysis of the returned database sequences, in conjunction with PoLV analysis, suggests that they are an invaluable resource in predicting virulent group A *var* genes from short DBLα tag sequences. The wealth of information held in this database will likely lead to identification of new DCs, improve current classification techniques, aid our understanding of *var* gene recombination, and reveal the global distribution of *var* gene sequences. Together, this information is likely to enhance our understanding of these highly variable genes and their role in pathogenesis of SM. We look forward to the full analysis and public availability of this database.

The mechanism by which IE sequestration leads to disease is poorly understood. However, the finding of differential *var* gene expression in the microvasculature of different organs has led to the suggestion of organ-specific sequestration (Montgomery *et al.*, 2007, Tembo *et al.*, 2014). Our attempts to sequence the full length genes of the short 28B1-1 and 62B1-1 sequences, identified specifically in the brains of fatal CM (Tembo *et al.*, 2014), were unsuccessful due to a lack of

parasite material for experimentation. However, the Pf3k database hits to these sequences have properties of virulent *var* genes. Both hit sequences are long *var* genes that are predicted to bind EPCR. In addition, the 62B1-1 returned sequence has the domain structure associated with IgM-mediated rosetting. Future studies should test for the presence of these sequences to assess their prevalence within the population and their association with CM. In addition, the acquisition of parasites harbouring these genes could provide valuable information on host cell binding and allow the study of mechanisms of disease pathogenesis.

Our SPR results of recombinant DBLβ::ICAM-1 interactions suggest an overall shared binding site between the DBLβ domains tested with minor differences in contact residues. This is in agreement with previous studies of IE binding, although some key differences between assays were observed (Tse *et al.*, 2004, Madkhali *et al.*, 2014, Madkhali, 2015) (see Discussion 5.4). A comparison of our DBLβ sequences to key peptides in the interaction of a group A DC4 DBLβ3 domain with ICAM-1 showed that they are distinct (see section 4.3.6), suggesting differential binding peptides between group A and group B and C proteins. A co-crystal of DC4 DBLβ3 bound to ICAM-1 has been solved and confirms involvement of the D2 domain (Frank Lennartz, unpublished data). Current efforts are focussed on obtaining the crystal structures of the group B BC12a$^{DBL\beta}$ and the group C J1a$^{DBL\beta}$ bound to ICAM-1 for comparison (Frank Lennartz, current work). Crystal structures will provide insight into important binding residues on both ICAM-1 and the DBLβ domains and will reveal any epitopes shared between the different PfEMP1 groups. This information will inform on the design of therapeutic interventions such as drugs and vaccines and will reveal whether multiple targets will be required to prevent broad PfEMP1 binding to a single host cell receptor. The evidence to date suggests that key epitopes will vary between the PfEMP1 groups. The merits of such therapeutics against PfEMP1 are discussed below.

Potential therapeutic strategies can be broadly placed into three categories: vaccine development against PfEMP1, blocking and/or reversing PfEMP1 adhesion as an adjunct to current drugs, and blocking the pathology induced by IE binding. The development of such therapies requires considerable time and money and

must be justified. The main justification is that even with hospitalisation and drug treatment the mortality rate of SM is at least 8-11% (Dondorp *et al.*, 2010) and has been reported as high as 50% in patients with multiple SM disease phenotypes (Hora *et al.*, 2016). In such a complex disease, alleviation of one facet, such as parasitaemia with drug treatment, may not be enough to reverse disease progression, suggesting that multi-faceted treatment is required. Reducing the sequestered IE burden is one such strategy that could improve patient outcomes.

The first strategy of the development of vaccines against PfEMP1 is supported by findings that antibodies against PfEMP1 comprise the majority of anti-IE antibodies and are protective against disease (Chan *et al.*, 2012). However, the high level of sequence variation both within and between parasite PfEMP1 repertoires presents a formidable challenge in finding disease-protective targets. Antibodies protective against SM are developed in early life, in regions of high transmission, and have identified group A PfEMP1 proteins as important in anti-disease immunity (Jensen *et al.*, 2004). This knowledge, combined with an increase in understanding of particular group A interactions with host receptors (Bengtsson *et al.*, 2013b, Berger *et al.*, 2013, Turner *et al.*, 2013, Lau *et al.*, 2015), provides a starting point for vaccine design, although further studies will be required to determine the level of cross-reactivity that may be achieved. VAR2CSA is the most extensively studied PfEMP1 in relation to vaccine development targeting PAM. Great optimism surrounded this protein due to the specific receptor-ligand interaction involved and the relative conservation of the *var2csa* gene between isolates (Hviid, 2011). However, identification of vital epitopes has proven more difficult than anticipated, partly due to the compact quaternary structure of the protein (Khunrae *et al.*, 2010), and higher than expected sequence variation between isolates (Fried and Duffy, 2015). A modular protein structure is anticipated to be the norm among other PfEMP1 proteins (Brown *et al.*, 2013), suggesting that single domains mediate binding. In this case, sequence variation is the main challenge to vaccine design, although conservation of structure has been shown in spite of sequence variation (Lau *et al.*, 2015). An additional challenge lies in identifying which receptor-ligand interaction to target. EPCR is perhaps the most obvious due to a clear mechanism

for disease progression. In addition, the CIDRα binding site is simple and sequence analysis accurately predicts binding of recombinant CIDRα proteins (Lau *et al.*, 2015). However, the study of this receptor in relation to PfEMP1 binding is in its infancy and greater understanding of EPCR binding and its occurrence in clinical settings is required. An ideal vaccine would encompass multiple epitopes involved in several interactions with, for example, EPCR, ICAM-1 and PECAM-1, to provide the best chance of preventing IE binding. Alternatively, the confirmation of the importance of a single host receptor in pathogenesis would support the inclusion of the epitopes involved in a vaccine targeting multiple parasite life cycle stages.

The second strategy of blocking and/or reversing PfEMP1 adhesion could be achieved by drug or monoclonal antibody treatment. The high variability among PfEMP1 proteins is, again, the main challenge to interventions. However, rational drug design has been demonstrated against ICAM-1-binding PfEMP1 by mimicking the ICAM-1 binding site (Dormeyer *et al.*, 2006). The compound in question, (+)-epigalloyl-catechin-gallate, was shown to be effective at micromolar concentrations against two parasite strains of IT lineage and inhibition of binding ranged from 37-80% in a panel of culture-adapted patient isolates, providing proof of concept for rational drug design (Dormeyer *et al.*, 2006, Patil *et al.*, 2011). The structure was based on the ICAM-1 DE loop (Dormeyer *et al.*, 2006), which includes residue L44, suggesting that this drug would also inhibit BC12a$^{DBL\beta}$, J1a$^{DBL\beta}$ and IT4var13$^{DBL\beta}$ binding. The recent development of a high-throughput screening platform to identify drug candidates against IE-binding is hoped to identify more molecules which can be further developed into anti-adhesion drugs (Gullingsrud *et al.*, 2015). This is a useful initial screen but levels of cross-blocking ability between parasite strains will need to be tested to evaluate the effectiveness of such a drug. The use of a heparin with reduced anticoagulant activity, Sevuparin, has been shown to disrupt rosettes (Leitgeb *et al.*, 2011) and is currently in Phase II clinical trials. The results are highly anticipated and will provide insight into the effectiveness of targeting IE adhesion and the merits of pursuing additional anti-adhesion drugs. A monoclonal antibody raised against the ICAM-1 binding domain of the 3D7 DC4-containing PfEMP1 was shown to be cross-reactive with other DC4 PfEMP1 proteins

(Lennartz *et al.*, 2015), providing proof of principle for antibody therapy. However, such a treatment is likely to be too expensive for use in such a widespread disease affecting some of the world's poorest countries. Cross-reactive monoclonal antibodies instead provide valuable information on important epitopes that will be useful in drug and vaccine design.

The third strategy is to block the pathology induced by IE binding. This has the advantage of targeting host factors rather than the highly variable PfEMP1 proteins themselves and overcomes the existence of binding redundancy (e.g. the contribution of other VSAs to receptor binding, such as RIFINs and STEVORs in rosetting). However, extensive knowledge of the pathways involved is required to ensure the safety of such interventions. One such therapy currently under investigation is treatment with L-arginine (Brussee *et al.*, 2015). L-arginine is the substrate for nitric oxide (NO) synthase found in endothelial cells. Plasma levels of L-arginine are low in SM resulting in reduced NO availability which is linked to endothelial dysfunction. Testing of the efficacy of L-arginine treatment is on-going but early results show no improvement on lactate clearance or endothelial NO bioavailability (Yeo *et al.*, 2013, Brussee *et al.*, 2015). Identification of EPCR as a PfEMP1 receptor (Turner *et al.*, 2013) and the reported loss of EPCR in CM (Moxon *et al.*, 2013), has led to optimism for pathology blocking interventions. The involvement of EPCR in anti-coagulant and cytoprotective pathways provides mechanisms for disease progression which are currently under investigation as therapeutic targets.

All the therapeutic strategies described above require extensive understanding of PfEMP1-receptor interactions and their role in pathogenesis. Of particular interest is the extent of PfEMP1 sequence diversity which can maintain receptor binding, and the relationship between binding to a particular host receptor and disease progression. However, the expected outcome of interventions targeting PfEMP1 warrants consideration. Specifically, what is the target product profile (TPP) of therapeutics against SM? What are the desired features and characteristics of such interventions? How can the efficacy of therapeutics be demonstrated? Targeting IE sequestration aims to prevent the development of associated pathology and allow

parasite clearance by the spleen and the immune response. How would efficacy be measured and what level of efficacy would be deemed sufficient for widespread use? These are all challenging questions and highlight the difficulty of designing anti-disease therapies. Greater understanding of the complex interplay between *P. falciparum* IE and their human host is required for progression towards answering these key questions.

References

ABDEL-LATIF, M. S., CABRERA, G., KOHLER, C., KREMSNER, P. G., LUTY, A. J. & TEAM, C. S. 2004. Antibodies to rifin: a component of naturally acquired responses to *Plasmodium falciparum* variant surface antigens on infected erythrocytes. *Am J Trop Med Hyg,* 71**,** 179-86.

ABDEL-LATIF, M. S., DIETZ, K., ISSIFOU, S., KREMSNER, P. G. & KLINKERT, M. Q. 2003. Antibodies to *Plasmodium falciparum* rifin proteins are associated with rapid parasite clearance and asymptomatic infections. *Infect Immun,* 71**,** 6229-33.

ABDEL-LATIF, M. S., KHATTAB, A., LINDENTHAL, C., KREMSNER, P. G. & KLINKERT, M. Q. 2002. Recognition of variant Rifin antigens by human antibodies induced during natural *Plasmodium falciparum* infections. *Infect Immun,* 70**,** 7013-21.

ADAMS, J. H., SIM, B. K., DOLAN, S. A., FANG, X., KASLOW, D. C. & MILLER, L. H. 1992. A family of erythrocyte binding proteins of malaria parasites. *Proc Natl Acad Sci U S A,* 89**,** 7085-9.

ADAMS, Y., KUHNRAE, P., HIGGINS, M. K., GHUMRA, A. & ROWE, J. A. 2014. Rosetting *Plasmodium falciparum*-infected erythrocytes bind to human brain microvascular endothelial cells in vitro, demonstrating a dual adhesion phenotype mediated by distinct *P. falciparum* erythrocyte membrane protein 1 domains. *Infect Immun,* 82**,** 949-59.

AIRD, W. C., MOSNIER, L. O. & FAIRHURST, R. M. 2014. *Plasmodium falciparum* picks (on) EPCR. *Blood,* 123**,** 163-7.

ALBRECHT, L., CASTINEIRAS, C., CARVALHO, B. O., LADEIA-ANDRADE, S., SANTOS DA SILVA, N., HOFFMANN, E. H., DALLA MARTHA, R. C., COSTA, F. T. & WUNDERLICH, G. 2010. The South American *Plasmodium falciparum var* gene repertoire is limited, highly shared and possibly lacks several antigenic types. *Gene,* 453**,** 37-44.

ALLRED, D. R. & AL-KHEDERY, B. 2004. Antigenic variation and cytoadhesion in *Babesia bovis* and *Plasmodium falciparum*: different logics achieve the same goal. *Mol Biochem Parasitol,* 134**,** 27-35.

ALMELLI, T., NDAM, N. T., EZIMEGNON, S., ALAO, M. J., AHOUANSOU, C., SAGBO, G., AMOUSSOU, A., DELORON, P. & TAHAR, R. 2014. Cytoadherence phenotype of *Plasmodium falciparum*-infected erythrocytes is associated with specific pfemp-1 expression in parasites from children with cerebral malaria. *Malar J,* 13:333**,** 1-9.

AVRIL, M., TRIPATHI, A. K., BRAZIER, A. J., ANDISI, C., JANES, J. H., SOMA, V. L., SULLIVAN, D. J., JR., BULL, P. C., STINS, M. F. & SMITH, J. D. 2012. A restricted subset of *var* genes mediates adherence of *Plasmodium falciparum*-infected erythrocytes to brain endothelial cells. *Proc Natl Acad Sci U S A,* 109**,** E1782-90.

BACHMANN, A., SCHOLZ, J. A., JANSSEN, M., KLINKERT, M. Q., TANNICH, E., BRUCHHAUS, I. & PETTER, M. 2015. A comparative study of the localization and membrane topology of members of the RIFIN, STEVOR and PfMC-2TM protein families in *Plasmodium falciparum*-infected erythrocytes. *Malar J,* 14:274**,** 1-18.

BAER, K., KLOTZ, C., KAPPE, S. H., SCHNIEDER, T. & FREVERT, U. 2007. Release of hepatic *Plasmodium yoelii* merozoites into the pulmonary microvasculature. *PLoS Pathog,* 3**,** e171.

BANCELLS, C. & DEITSCH, K. W. 2013. A molecular switch in the efficiency of translation reinitiation controls expression of var2csa, a gene implicated in pregnancy-associated malaria. *Mol Microbiol,* 90**,** 472-88.

BARRAGAN, A., KREMSNER, P. G., WAHLGREN, M. & CARLSON, J. 2000. Blood group A antigen is a coreceptor in *Plasmodium falciparum* rosetting. *Infect Immun,* 68**,** 2971-5.

BARRY, A. E., TRIEU, A., FOWKES, F. J., PABLO, J., KALANTARI-DEHAGHI, M., JASINSKAS, A., TAN, X., KAYALA, M. A., TAVUL, L., SIBA, P. M., DAY, K. P., BALDI, P., FELGNER, P. L. & DOOLAN, D. L. 2011. The stability and complexity of antibody responses to the major surface antigen of *Plasmodium falciparum* are associated with age in a malaria endemic area. *Mol Cell Proteomics,* 10**,** M111 008326.

BARUCH, D. I., MA, X. C., SINGH, H. B., BI, X., PASLOSKE, B. L. & HOWARD, R. J. 1997. Identification of a region of PfEMP1 that mediates adherence of *Plasmodium falciparum* infected erythrocytes to CD36: conserved function with variant sequence. *Blood,* 90**,** 3766-75.

BARUCH, D. I., PASLOSKE, B. L., SINGH, H. B., BI, X., MA, X. C., FELDMAN, M., TARASCHI, T. F. & HOWARD, R. J. 1995. Cloning the *P. falciparum* gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. *Cell,* 82**,** 77-87.

BEESON, J. G., BROWN, G. V., MOLYNEUX, M. E., MHANGO, C., DZINJALAMALA, F. & ROGERSON, S. J. 1999. *Plasmodium falciparum* isolates from infected pregnant women and children are associated with distinct adhesive and antigenic properties. *J Infect Dis,* 180**,** 464-72.

BEESON, J. G., DREW, D. R., BOYLE, M. J., FENG, G., FOWKES, F. J. & RICHARDS, J. S. 2016. Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria. *FEMS Microbiol Rev,* 40**,** 343-72.

BEESON, J. G., OSIER, F. H. & ENGWERDA, C. R. 2008. Recent insights into humoral and cellular immune responses against malaria. *Trends Parasitol,* 24**,** 578-84.

BELLAMY, R., KWIATKOWSKI, D. & HILL, A. V. 1998. Absence of an association between intercellular adhesion molecule 1, complement receptor 1 and interleukin 1 receptor antagonist gene polymorphisms and severe malaria in a West African population. *Trans R Soc Trop Med Hyg,* 92**,** 312-6.

BENGTSSON, A., JOERGENSEN, L., BARBATI, Z. R., CRAIG, A., HVIID, L. & JENSEN, A. T. 2013a. Transfected HEK293 cells expressing functional recombinant intercellular adhesion molecule 1 (ICAM-1)--a receptor associated with severe *Plasmodium falciparum* malaria. *PLoS One,* 8**,** e69999.

BENGTSSON, A., JOERGENSEN, L., RASK, T. S., OLSEN, R. W., ANDERSEN, M. A., TURNER, L., THEANDER, T. G., HVIID, L., HIGGINS, M. K., CRAIG, A., BROWN, A. & JENSEN, A. T. 2013b. A novel domain cassette identifies *Plasmodium falciparum* PfEMP1 proteins binding ICAM-1 and is a target of cross-reactive, adhesion-inhibitory antibodies. *J Immunol,* 190**,** 240-9.

BERENDT, A. R., MCDOWALL, A., CRAIG, A. G., BATES, P. A., STERNBERG, M. J., MARSH, K., NEWBOLD, C. I. & HOGG, N. 1992. The binding site on ICAM-1 for *Plasmodium falciparum*-infected erythrocytes overlaps, but is distinct from, the LFA-1-binding site. *Cell,* 68**,** 71-81.

BERENDT, A. R., SIMMONS, D. L., TANSEY, J., NEWBOLD, C. I. & MARSH, K. 1989. Intercellular adhesion molecule-1 is an endothelial cell adhesion receptor for *Plasmodium falciparum*. *Nature,* 341**,** 57-59.

BERGER, S. S., TURNER, L., WANG, C. W., PETERSEN, J. E., KRAFT, M., LUSINGU, J. P., MMBANDO, B., MARQUARD, A. M., BENGTSSON, D. B., HVIID, L., NIELSEN, M. A., THEANDER, T. G. & LAVSTSEN, T. 2013. *Plasmodium falciparum* expressing domain cassette 5 type PfEMP1 (DC5-PfEMP1) bind PECAM1. *PLoS One,* 8**,** e69117.

BERTIN, G. I., LAVSTSEN, T., GUILLONNEAU, F., DORITCHAMOU, J., WANG, C. W., JESPERSEN, J. S., EZIMEGNON, S., FIEVET, N., ALAO, M. J., LALYA, F., MASSOUGBODJI, A., NDAM, N. T., THEANDER, T. G. & DELORON, P. 2013. Expression of the domain cassette 8 *Plasmodium falciparum* erythrocyte

membrane protein 1 is associated with cerebral malaria in Benin. *PLoS One,* 8**,** e68368.

BHATT, S., WEISS, D. J., CAMERON, E., BISANZIO, D., MAPPIN, B., DALRYMPLE, U., BATTLE, K. E., MOYES, C. L., HENRY, A., ECKHOFF, P. A., WENGER, E. A., BRIET, O., PENNY, M. A., SMITH, T. A., BENNETT, A., YUKICH, J., EISELE, T. P., GRIFFIN, J. T., FERGUS, C. A., LYNCH, M., LINDGREN, F., COHEN, J. M., MURRAY, C. L., SMITH, D. L., HAY, S. I., CIBULSKIS, R. E. & GETHING, P. W. 2015. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature,* 526**,** 207-11.

BIGGS, B. A., GOOZE, L., WYCHERLEY, K., WOLLISH, W., SOUTHWELL, B., LEECH, J. H. & BROWN, G. V. 1991. Antigenic variation in *Plasmodium falciparum*. *Proc Natl Acad Sci U S A,* 88**,** 9171-4.

BIRBECK, G. L., MOLYNEUX, M. E., KAPLAN, P. W., SEYDEL, K. B., CHIMALIZENI, Y. F., KAWAZA, K. & TAYLOR, T. E. 2010. Blantyre Malaria Project Epilepsy Study (BMPES) of neurological outcomes in retinopathy-positive paediatric cerebral malaria survivors: a prospective cohort study. *Lancet Neurol,* 9**,** 1173-81.

BISWAS, A. K., HAFIZ, A., BANERJEE, B., KIM, K. S., DATTA, K. & CHITNIS, C. E. 2007. *Plasmodium falciparum* uses gC1qR/HABP1/p32 as a receptor to bind to vascular endothelium and for platelet-mediated clumping. *PLoS Pathog,* 3**,** 1271-80.

BLISNICK, T., MORALES BETOULLE, M. E., BARALE, J. C., UZUREAU, P., BERRY, L., DESROSES, S., FUJIOKA, H., MATTEI, D. & BRAUN BRETON, C. 2000. Pfsbp1, a Maurer's cleft *Plasmodium falciparum* protein, is associated with the erythrocyte skeleton. *Mol Biochem Parasitol,* 111**,** 107-21.

BLYTHE, J. E., YAM, X. Y., KUSS, C., BOZDECH, Z., HOLDER, A. A., MARSH, K., LANGHORNE, J. & PREISER, P. R. 2008. *Plasmodium falciparum* STEVOR proteins are highly expressed in patient isolates and located in the surface membranes of infected red blood cells and the apical tips of merozoites. *Infect Immun,* 76**,** 3329-36.

BODDEY, J. A., CARVALHO, T. G., HODDER, A. N., SARGEANT, T. J., SLEEBS, B. E., MARAPANA, D., LOPATICKI, S., NEBL, T. & COWMAN, A. F. 2013. Role of plasmepsin V in export of diverse protein families from the *Plasmodium falciparum* exportome. *Traffic,* 14**,** 532-50.

BOPP, S. E., MANARY, M. J., BRIGHT, A. T., JOHNSTON, G. L., DHARIA, N. V., LUNA, F. L., MCCORMACK, S., PLOUFFE, D., MCNAMARA, C. W., WALKER, J. R., FIDOCK, D. A., DENCHI, E. L. & WINZELER, E. A. 2013. Mitotic evolution of *Plasmodium falciparum* shows a stable core genome but recombination in antigen families. *PLoS Genet,* 9**,** e1003293.

BROCK, P. M., FORNACE, K. M., PARMITER, M., COX, J., DRAKELEY, C. J., FERGUSON, H. M. & KAO, R. R. 2016. *Plasmodium knowlesi* transmission: integrating quantitative approaches from epidemiology and ecology to understand malaria as a zoonosis. *Parasitology,* 143**,** 389-400.

BROWN, A., TURNER, L., CHRISTOFFERSEN, S., ANDREWS, K. A., SZESTAK, T., ZHAO, Y., LARSEN, S., CRAIG, A. G. & HIGGINS, M. K. 2013. Molecular architecture of a complex between an adhesion protein from the malaria parasite and intracellular adhesion molecule 1. *J Biol Chem,* 288**,** 5992-6003.

BRUSSEE, J. M., YEO, T. W., LAMPAH, D. A., ANSTEY, N. M. & DUFFULL, S. B. 2015. Pharmacokinetic-Pharmacodynamic Model for the Effect of l-Arginine on Endothelial Function in Patients with Moderately Severe Falciparum Malaria. *Antimicrob Agents Chemother,* 60**,** 198-205.

BUCKEE, C. O. & RECKER, M. 2012. Evolution of the multi-domain structures of virulence genes in the human malaria parasite, *Plasmodium falciparum*. *PLoS Comput Biol,* 8**,** e1002451.

BULL, P. C., BERRIMAN, M., KYES, S., QUAIL, M. A., HALL, N., KORTOK, M. M., MARSH, K. & NEWBOLD, C. I. 2005. *Plasmodium falciparum* variant surface antigen expression patterns during malaria. *PLoS Pathog,* 1**,** e26.

BULL, P. C., KYES, S., BUCKEE, C. O., MONTGOMERY, J., KORTOK, M. M., NEWBOLD, C. I. & MARSH, K. 2007. An approach to classifying sequence tags sampled from *Plasmodium falciparum var* genes. *Mol Biochem Parasitol,* 154**,** 98-102.

BULL, P. C., LOWE, B. S., KORTOK, M., MOLYNEUX, C. S., NEWBOLD, C. I. & MARSH, K. 1998. Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. *Nat Med,* 4**,** 358-60.

CABRERA, A., NECULAI, D. & KAIN, K. C. 2014. CD36 and malaria: friends or foes? A decade of data provides some answers. *Trends Parasitol,* 30**,** 436-44.

CALIS, J. C., PHIRI, K. S., FARAGHER, E. B., BRABIN, B. J., BATES, I., CUEVAS, L. E., DE HAAN, R. J., PHIRI, A. I., MALANGE, P., KHOKA, M., HULSHOF, P. J., VAN LIESHOUT, L., BELD, M. G., TEO, Y. Y., ROCKETT, K. A., RICHARDSON, A., KWIATKOWSKI, D. P., MOLYNEUX, M. E. & VAN HENSBROEK, M. B. 2008. Severe anemia in Malawian children. *N Engl J Med,* 358**,** 888-99.

CAMERON, E., BATTLE, K. E., BHATT, S., WEISS, D. J., BISANZIO, D., MAPPIN, B., DALRYMPLE, U., HAY, S. I., SMITH, D. L., GRIFFIN, J. T., WENGER, E. A., ECKHOFF, P. A., SMITH, T. A., PENNY, M. A. & GETHING, P. W. 2015. Defining the relationship between infection prevalence and clinical incidence of *Plasmodium falciparum* malaria. *Nat Commun,* 6**,** 8170.

CARLSON, J., HELMBY, H., HILL, A. V., BREWSTER, D., GREENWOOD, B. M. & WAHLGREN, M. 1990. Human cerebral malaria: association with erythrocyte rosetting and lack of anti-rosetting antibodies. *Lancet,* 336**,** 1457-60.

CARLSON, J. & WAHLGREN, M. 1992. *Plasmodium falciparum* erythrocyte rosetting is mediated by promiscuous lectin-like interactions. *J Exp Med,* 176**,** 1311-7.

CARTER, J. A., MUNG'ALA-ODERA, V., NEVILLE, B. G., MURIRA, G., MTURI, N., MUSUMBA, C. & NEWTON, C. R. 2005. Persistent neurocognitive impairments associated with severe falciparum malaria in Kenyan children. *J Neurol Neurosurg Psychiatry,* 76**,** 476-81.

CASASNOVAS, J. M., BICKFORD, J. K. & SPRINGER, T. A. 1998. The domain structure of ICAM-1 and the kinetics of binding to rhinovirus. *J Virol,* 72**,** 6244-6.

CHAM, G. K., TURNER, L., KURTIS, J. D., MUTABINGWA, T., FRIED, M., JENSEN, A. T., LAVSTSEN, T., HVIID, L., DUFFY, P. E. & THEANDER, T. G. 2010. Hierarchical, domain type-specific acquisition of antibodies to *Plasmodium falciparum* erythrocyte membrane protein 1 in Tanzanian children. *Infect Immun,* 78**,** 4653-9.

CHAM, G. K., TURNER, L., LUSINGU, J., VESTERGAARD, L., MMBANDO, B. P., KURTIS, J. D., JENSEN, A. T., SALANTI, A., LAVSTSEN, T. & THEANDER, T. G. 2009. Sequential, ordered acquisition of antibodies to *Plasmodium falciparum* erythrocyte membrane protein 1 domains. *J Immunol,* 183**,** 3356-63.

CHAN, J. A., FOWKES, F. J. & BEESON, J. G. 2014. Surface antigens of *Plasmodium falciparum*-infected erythrocytes as immune targets and malaria vaccine candidates. *Cell Mol Life Sci,* 71**,** 3633-57.

CHAN, J. A., HOWELL, K. B., REILING, L., ATAIDE, R., MACKINTOSH, C. L., FOWKES, F. J., PETTER, M., CHESSON, J. M., LANGER, C., WARIMWE, G. M., DUFFY, M. F., ROGERSON, S. J., BULL, P. C., COWMAN, A. F., MARSH, K. & BEESON, J. G. 2012. Targets of antibodies against *Plasmodium falciparum*-infected erythrocytes in malaria immunity. *J Clin Invest,* 122**,** 3227-38.

CHATTOPADHYAY, R., SHARMA, A., SRIVASTAVA, V. K., PATI, S. S., SHARMA, S. K., DAS, B. S. & CHITNIS, C. E. 2003. *Plasmodium falciparum* infection elicits both variant-specific

and cross-reactive antibodies against variant surface antigens. *Infect Immun,* 71**,** 597-604.

CHATTOPADHYAY, R., TANEJA, T., CHAKRABARTI, K., PILLAI, C. R. & CHITNIS, C. E. 2004. Molecular analysis of the cytoadherence phenotype of a *Plasmodium falciparum* field isolate that binds intercellular adhesion molecule-1. *Mol Biochem Parasitol,* 133**,** 255-65.

CHEN, D. S., BARRY, A. E., LELIWA-SYTEK, A., SMITH, T. A., PETERSON, I., BROWN, S. M., MIGOT-NABIAS, F., DELORON, P., KORTOK, M. M., MARSH, K., DAILY, J. P., NDIAYE, D., SARR, O., MBOUP, S. & DAY, K. P. 2011. A molecular epidemiological study of *var* gene diversity to characterize the reservoir of *Plasmodium falciparum* in humans in Africa. *PLoS One,* 6**,** e16629.

CHEN, Q., HEDDINI, A., BARRAGAN, A., FERNANDEZ, V., PEARCE, S. F. & WAHLGREN, M. 2000. The semiconserved head structure of *Plasmodium falciparum* erythrocyte membrane protein 1 mediates binding to multiple independent host receptors. *J Exp Med,* 192**,** 1-10.

CHENG, Q., CLOONAN, N., FISCHER, K., THOMPSON, J., WAINE, G., LANZER, M. & SAUL, A. 1998. *stevor* and *rif* are *Plasmodium falciparum* multicopy gene families which potentially encode variant antigens. *Mol Biochem Parasitol,* 97**,** 161-76.

CHITNIS, C. E. & MILLER, L. H. 1994. Identification of the erythrocyte binding domains of *Plasmodium vivax* and *Plasmodium knowlesi* proteins involved in erythrocyte invasion. *J Exp Med,* 180**,** 497-506.

CHOTIVANICH, K., SRITABAL, J., UDOMSANGPETCH, R., NEWTON, P., STEPNIEWSKA, K. A., RUANGVEERAYUTH, R., LOOAREESUWAN, S., ROBERTS, D. J. & WHITE, N. J. 2004. Platelet-induced autoagglutination of *Plasmodium falciparum*-infected red blood cells and disease severity in Thailand. *J Infect Dis,* 189**,** 1052-5.

CHURCH, J. A., NYAMAKO, L., OLUPOT-OLUPOT, P., MAITLAND, K. & URBAN, B. C. 2016. Increased adhesion of *Plasmodium falciparum* infected erythrocytes to ICAM-1 in children with acute intestinal injury. *Malar J,* 15**,** 1-6.

CLAESSENS, A., ADAMS, Y., GHUMRA, A., LINDERGARD, G., BUCHAN, C. C., ANDISI, C., BULL, P. C., MOK, S., GUPTA, A. P., WANG, C. W., TURNER, L., ARMAN, M., RAZA, A., BOZDECH, Z. & ROWE, J. A. 2012. A subset of group A-like *var* genes encodes the malaria parasite ligands for binding to human brain endothelial cells. *Proc Natl Acad Sci U S A,* 109**,** E1772-81.

CLAESSENS, A., GHUMRA, A., GUPTA, A. P., MOK, S., BOZDECH, Z. & ROWE, J. A. 2011. Design of a variant surface antigen-supplemented microarray chip for whole transcriptome analysis of multiple *Plasmodium falciparum* cytoadherent strains, and identification of strain-transcendent *rif* and *stevor* genes. *Malar J,* 10:180**,** 1-16.

CLAESSENS, A., HAMILTON, W. L., KEKRE, M., OTTO, T. D., FAIZULLABHOY, A., RAYNER, J. C. & KWIATKOWSKI, D. 2014. Generation of antigenic diversity in *Plasmodium falciparum* by structured rearrangement of *Var* genes during mitosis. *PLoS Genet,* 10**,** e1004812.

CLAUSEN, T. M., CHRISTOFFERSEN, S., DAHLBACK, M., LANGKILDE, A. E., JENSEN, K. E., RESENDE, M., AGERBAEK, M. O., ANDERSEN, D., BERISHA, B., DITLEV, S. B., PINTO, V. V., NIELSEN, M. A., THEANDER, T. G., LARSEN, S. & SALANTI, A. 2012. Structural and functional insight into how the *Plasmodium falciparum* VAR2CSA protein mediates binding to chondroitin sulfate A in placental malaria. *J Biol Chem,* 287**,** 23332-45.

CLOUGH, B., ATILOLA, F. A., BLACK, J. & PASVOL, G. 1998. *Plasmodium falciparum*: the importance of IgM in the rosetting of parasite-infected erythrocytes. *Exp Parasitol,* 89**,** 129-32.

COLEMAN, B. I., SKILLMAN, K. M., JIANG, R. H., CHILDS, L. M., ALTENHOFEN, L. M., GANTER, M., LEUNG, Y., GOLDOWITZ, I., KAFSACK, B. F., MARTI, M., LLINAS, M., BUCKEE, C. O. & DURAISINGH, M. T. 2014. A *Plasmodium falciparum* histone deacetylase regulates antigenic variation and gametocyte conversion. *Cell Host Microbe,* 16**,** 177-86.

COOKE, B. M., BERENDT, A. R., CRAIG, A. G., MACGREGOR, J., NEWBOLD, C. I. & NASH, G. B. 1994. Rolling and stationary cytoadhesion of red blood cells parasitized by *Plasmodium falciparum*: separate roles for ICAM-1, CD36 and thrombospondin. *Br J Haematol,* 87**,** 162-70.

COOKE, B. M., BUCKINGHAM, D. W., GLENISTER, F. K., FERNANDEZ, K. M., BANNISTER, L. H., MARTI, M., MOHANDAS, N. & COPPEL, R. L. 2006. A Maurer's cleft-associated protein is essential for expression of the major malaria virulence antigen on the surface of infected red blood cells. *J Cell Biol,* 172**,** 899-908.

CRABB, B. S., COOKE, B. M., REEDER, J. C., WALLER, R. F., CARUANA, S. R., DAVERN, K. M., WICKHAM, M. E., BROWN, G. V., COPPEL, R. L. & COWMAN, A. F. 1997. Targeted gene disruption shows that knobs enable malaria-infected red cells to cytoadhere under physiological shear stress. *Cell,* 89**,** 287-96.

CSERTI-GAZDEWICH, C. M., DHABANGI, A., MUSOKE, C., SSEWANYANA, I., DDUNGU, H., NAKIBONEKA-SSENABULYA, D., NABUKEERA-BARUNGI, N., MPIMBAZA, A. & DZIK, W. H. 2012. Cytoadherence in paediatric malaria: ABO blood group, CD36, and ICAM1 expression and severe *Plasmodium falciparum* infection. *Br J Haematol,* 159**,** 223-36.

CULLETON, R. & KANEKO, O. 2010. Erythrocyte binding ligands in malaria parasites: intracellular trafficking and parasite virulence. *Acta Trop,* 114**,** 131-7.

DATTA, S., MALHOTRA, L., DICKERSON, R., CHAFFEE, S., SEN, C. K. & ROY, S. 2015. Laser capture microdissection: Big data from small samples. *Histol Histopathol,* 30**,** 1255-69.

DE KONING-WARD, T. F., GILSON, P. R., BODDEY, J. A., RUG, M., SMITH, B. J., PAPENFUSS, A. T., SANDERS, P. R., LUNDIE, R. J., MAIER, A. G., COWMAN, A. F. & CRABB, B. S. 2009. A newly discovered protein export machine in malaria parasites. *Nature,* 459**,** 945-9.

DEITSCH, K. W., CALDERWOOD, M. S. & WELLEMS, T. E. 2001. Malaria. Cooperative silencing elements in *var* genes. *Nature,* 412**,** 875-6.

DEITSCH, K. W., LUKEHART, S. A. & STRINGER, J. R. 2009. Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nat Rev Microbiol,* 7**,** 493-503.

DI TOMMASO, P., MORETTI, S., XENARIOS, I., OROBITG, M., MONTANYOLA, A., CHANG, J. M., TALY, J. F. & NOTREDAME, C. 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res,* 39**,** 13-7.

DODOO, D., STAALSOE, T., GIHA, H., KURTZHALS, J. A., AKANMORI, B. D., KORAM, K., DUNYO, S., NKRUMAH, F. K., HVIID, L. & THEANDER, T. G. 2001. Antibodies to variant antigens on the surfaces of infected erythrocytes are associated with protection from malaria in Ghanaian children. *Infect Immun,* 69**,** 3713-8.

DONDORP, A. M., FANELLO, C. I., HENDRIKSEN, I. C., GOMES, E., SENI, A., CHHAGANLAL, K. D., BOJANG, K., OLAOSEBIKAN, R., ANUNOBI, N., MAITLAND, K., KIVAYA, E., AGBENYEGA, T., NGUAH, S. B., EVANS, J., GESASE, S., KAHABUKA, C., MTOVE, G., NADJM, B., DEEN, J., MWANGA-AMUMPAIRE, J., NANSUMBA, M., KAREMA, C., UMULISA, N., UWIMANA, A., MOKUOLU, O. A., ADEDOYIN, O. T., JOHNSON, W. B., TSHEFU, A. K., ONYAMBOKO, M. A., SAKULTHAEW, T., NGUM, W. P., SILAMUT, K., STEPNIEWSKA, K., WOODROW, C. J., BETHELL, D., WILLS, B., ONEKO, M., PETO, T.

E., VON SEIDLEIN, L., DAY, N. P., WHITE, N. J. & GROUP, A. 2010. Artesunate versus quinine in the treatment of severe falciparum malaria in African children (AQUAMAT): an open-label, randomised trial. *Lancet,* 376**,** 1647-57.

DORMER, P., DIETRICH, M., KERN, P. & HORSTMANN, R. D. 1983. Ineffective erythropoiesis in acute human *P. falciparum* malaria. *Blut,* 46**,** 279-88.

DORMEYER, M., ADAMS, Y., KRAMER, B., CHAKRAVORTY, S., TSE, M. T., PEGORARO, S., WHITTAKER, L., LANZER, M. & CRAIG, A. 2006. Rational design of anticytoadherence inhibitors for *Plasmodium falciparum* based on the crystal structure of human intercellular adhesion molecule 1. *Antimicrob Agents Chemother,* 50**,** 724-30.

DOROVINI-ZIS, K., SCHMIDT, K., HUYNH, H., FU, W., WHITTEN, R. O., MILNER, D., KAMIZA, S., MOLYNEUX, M. & TAYLOR, T. E. 2011. The neuropathology of fatal cerebral malaria in malawian children. *Am J Pathol,* 178**,** 2146-58.

DURAISINGH, M. T., VOSS, T. S., MARTY, A. J., DUFFY, M. F., GOOD, R. T., THOMPSON, J. K., FREITAS-JUNIOR, L. H., SCHERF, A., CRABB, B. S. & COWMAN, A. F. 2005. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell,* 121**,** 13-24.

DZIKOWSKI, R., LI, F., AMULIC, B., EISBERG, A., FRANK, M., PATEL, S., WELLEMS, T. E. & DEITSCH, K. W. 2007. Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. *EMBO Rep,* 8**,** 959-65.

EDA, S., LAWLER, J. & SHERMAN, I. W. 1999. *Plasmodium falciparum*-infected erythrocyte adhesion to the type 3 repeat domain of thrombospondin-1 is mediated by a modified band 3 protein. *Mol Biochem Parasitol,* 100**,** 195-205.

ELLIOTT, S. R., PAYNE, P. D., DUFFY, M. F., BYRNE, T. J., THAM, W. H., ROGERSON, S. J., BROWN, G. V. & EISEN, D. P. 2007a. Antibody recognition of heterologous variant surface antigens after a single *Plasmodium falciparum* infection in previously naive adults. *Am J Trop Med Hyg,* 76**,** 860-4.

ELLIOTT, S. R., SPURCK, T. P., DODIN, J. M., MAIER, A. G., VOSS, T. S., YOSAATMADJA, F., PAYNE, P. D., MCFADDEN, G. I., COWMAN, A. F., ROGERSON, S. J., SCHOFIELD, L. & BROWN, G. V. 2007b. Inhibition of dendritic cell maturation by malaria is dose dependent and does not require *Plasmodium falciparum* erythrocyte membrane protein 1. *Infect Immun,* 75**,** 3621-32.

ELMENDORF, H. G. & HALDAR, K. 1994. *Plasmodium falciparum* exports the Golgi marker sphingomyelin synthase into a tubovesicular network in the cytoplasm of mature erythrocytes. *J Cell Biol,* 124**,** 449-62.

ENDERES, C., KOMBILA, D., DAL-BIANCO, M., DZIKOWSKI, R., KREMSNER, P. & FRANK, M. 2011. *var* gene promoter activation in clonal *Plasmodium falciparum* isolates follows a hierarchy and suggests a conserved switching program that is independent of genetic background. *J Infect Dis,* 204**,** 1620-31.

EPP, C., LI, F., HOWITT, C. A., CHOOKAJORN, T. & DEITSCH, K. W. 2009. Chromatin associated sense and antisense noncoding RNAs are transcribed from the *var* gene family of virulence genes of the malaria parasite *Plasmodium falciparum*. *RNA,* 15**,** 116-27.

FAIRHURST, R. M. 2015. Understanding artemisinin-resistant malaria: what a difference a year makes. *Curr Opin Infect Dis,* 28**,** 417-25.

FASTMAN, Y., NOBLE, R., RECKER, M. & DZIKOWSKI, R. 2012. Erasing the epigenetic memory and beginning to switch--the onset of antigenic switching of *var* genes in *Plasmodium falciparum*. *PLoS One,* 7**,** e34168.

FERNANDEZ-REYES, D., CRAIG, A. G., KYES, S. A., PESHU, N., SNOW, R. W., BERENDT, A. R., MARSH, K. & NEWBOLD, C. I. 1997. A high frequency African coding polymorphism

in the N-terminal domain of ICAM-1 predisposing to cerebral malaria in Kenya. *Hum Mol Genet,* 6**,** 1357-60.

FERNANDEZ, V., HOMMEL, M., CHEN, Q., HAGBLOM, P. & WAHLGREN, M. 1999. Small, clonally variant antigens expressed on the surface of the *Plasmodium falciparum*-infected erythrocyte are encoded by the *rif* gene family and are the target of human immune responses. *J Exp Med,* 190**,** 1393-404.

FLUECK, C., BARTFAI, R., VOLZ, J., NIEDERWIESER, I., SALCEDO-AMAYA, A. M., ALAKO, B. T., EHLGEN, F., RALPH, S. A., COWMAN, A. F., BOZDECH, Z., STUNNENBERG, H. G. & VOSS, T. S. 2009. *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog,* 5**,** e1000569.

FRANK, M., DZIKOWSKI, R., AMULIC, B. & DEITSCH, K. 2007. Variable switching rates of malaria virulence genes are associated with chromosomal position. *Mol Microbiol,* 64**,** 1486-98.

FRANK, M., DZIKOWSKI, R., COSTANTINI, D., AMULIC, B., BERDOUGO, E. & DEITSCH, K. 2006. Strict pairing of *var* promoters and introns is required for *var* gene silencing in the malaria parasite *Plasmodium falciparum*. *J Biol Chem,* 281**,** 9942-52.

FRANK, S. A. & BUSH, R. M. 2007. Barriers to antigenic escape by pathogens: trade-off between reproductive rate and antigenic mutability. *BMC Evol Biol,* 7:229**,** 1-13.

FREITAS-JUNIOR, L. H., HERNANDEZ-RIVAS, R., RALPH, S. A., MONTIEL-CONDADO, D., RUVALCABA-SALAZAR, O. K., ROJAS-MEZA, A. P., MANCIO-SILVA, L., LEAL-SILVESTRE, R. J., GONTIJO, A. M., SHORTE, S. & SCHERF, A. 2005. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell,* 121**,** 25-36.

FRIED, M. & DUFFY, P. E. 1996. Adherence of *Plasmodium falciparum* to chondroitin sulfate A in the human placenta. *Science,* 272**,** 1502-4.

FRIED, M. & DUFFY, P. E. 2015. Designing a VAR2CSA-based vaccine to prevent placental malaria. *Vaccine,* 33**,** 7483-8.

FRIED, M., MUGA, R. O., MISORE, A. O. & DUFFY, P. E. 1998. Malaria elicits type 1 cytokines in the human placenta: IFN-gamma and TNF-alpha associated with pregnancy outcomes. *J Immunol,* 160**,** 2523-30.

FRY, A. E., GHANSA, A., SMALL, K. S., PALMA, A., AUBURN, S., DIAKITE, M., GREEN, A., CAMPINO, S., TEO, Y. Y., CLARK, T. G., JEFFREYS, A. E., WILSON, J., JALLOW, M., SISAY-JOOF, F., PINDER, M., GRIFFITHS, M. J., PESHU, N., WILLIAMS, T. N., NEWTON, C. R., MARSH, K., MOLYNEUX, M. E., TAYLOR, T. E., KORAM, K. A., ODURO, A. R., ROGERS, W. O., ROCKETT, K. A., SABETI, P. C. & KWIATKOWSKI, D. P. 2009. Positive selection of a CD36 nonsense variant in sub-Saharan Africa, but no association with severe malaria phenotypes. *Hum Mol Genet,* 18**,** 2683-92.

GARDNER, M. J., HALL, N., FUNG, E., WHITE, O., BERRIMAN, M., HYMAN, R. W., CARLTON, J. M., PAIN, A., NELSON, K. E., BOWMAN, S., PAULSEN, I. T., JAMES, K., EISEN, J. A., RUTHERFORD, K., SALZBERG, S. L., CRAIG, A., KYES, S., CHAN, M. S., NENE, V., SHALLOM, S. J., SUH, B., PETERSON, J., ANGIUOLI, S., PERTEA, M., ALLEN, J., SELENGUT, J., HAFT, D., MATHER, M. W., VAIDYA, A. B., MARTIN, D. M., FAIRLAMB, A. H., FRAUNHOLZ, M. J., ROOS, D. S., RALPH, S. A., MCFADDEN, G. I., CUMMINGS, L. M., SUBRAMANIAN, G. M., MUNGALL, C., VENTER, J. C., CARUCCI, D. J., HOFFMAN, S. L., NEWBOLD, C., DAVIS, R. W., FRASER, C. M. & BARRELL, B. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature,* 419**,** 498-511.

GATTON, M. L., PETERS, J. M., GRESTY, K., FOWLER, E. V., CHEN, N. & CHENG, Q. 2006. Detection sensitivity and quantitation of *Plasmodium falciparum var* gene

transcripts by real-time RT-PCR in comparison with conventional RT-PCR. *Am J Trop Med Hyg,* 75**,** 212-8.

GHUMRA, A., SEMBLAT, J. P., ATAIDE, R., KIFUDE, C., ADAMS, Y., CLAESSENS, A., ANONG, D. N., BULL, P. C., FENNELL, C., ARMAN, M., AMAMBUA-NGWA, A., WALTHER, M., CONWAY, D. J., KASSAMBARA, L., DOUMBO, O. K., RAZA, A. & ROWE, J. A. 2012. Induction of strain-transcending antibodies against Group A PfEMP1 surface antigens from virulent malaria parasites. *PLoS Pathog,* 8**,** e1002665.

GILLRIE, M. R., AVRIL, M., BRAZIER, A. J., DAVIS, S. P., STINS, M. F., SMITH, J. D. & HO, M. 2015. Diverse functional outcomes of *Plasmodium falciparum* ligation of EPCR: potential implications for malarial pathogenesis. *Cell Microbiol,* 17**,** 1883-99.

GLEESON, E. M., O'DONNELL, J. S. & PRESTON, R. J. 2012. The endothelial cell protein C receptor: cell surface conductor of cytoprotective coagulation factor signaling. *Cell Mol Life Sci,* 69**,** 717-26.

GOEL, S., PALMKVIST, M., MOLL, K., JOANNIN, N., LARA, P., AKHOURI, R. R., MORADI, N., OJEMALM, K., WESTMAN, M., ANGELETTI, D., KJELLIN, H., LEHTIO, J., BLIXT, O., IDESTROM, L., GAHMBERG, C. G., STORRY, J. R., HULT, A. K., OLSSON, M. L., VON HEIJNE, G., NILSSON, I. & WAHLGREN, M. 2015. RIFINs are adhesins implicated in severe *Plasmodium falciparum* malaria. *Nat Med,* 21**,** 314-7.

GRAY, C., MCCORMICK, C., TURNER, G. & CRAIG, A. 2003. ICAM-1 can play a major role in mediating *P. falciparum* adhesion to endothelium under flow. *Mol Biochem Parasitol,* 128**,** 187-93.

GREENFIELD, N. J. 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc,* 1**,** 2876-90.

GREVE, J. M., DAVIS, G., MEYER, A. M., FORTE, C. P., YOST, S. C., MARLOR, C. W., KAMARCK, M. E. & MCCLELLAND, A. 1989. The major human rhinovirus receptor is ICAM-1. *Cell,* 56**,** 839-47.

GRIFFIN, J. T., HOLLINGSWORTH, T. D., REYBURN, H., DRAKELEY, C. J., RILEY, E. M. & GHANI, A. C. 2015. Gradual acquisition of immunity to severe malaria with increasing exposure. *Proc Biol Sci,* 282**,** 20142657.

GRUENBERG, J., ALLRED, D. R. & SHERMAN, I. W. 1983. Scanning electron microscope-analysis of the protrusions (knobs) present on the surface of *Plasmodium falciparum*-infected erythrocytes. *J Cell Biol,* 97**,** 795-802.

GRURING, C., HEIBER, A., KRUSE, F., UNGEFEHR, J., GILBERGER, T. W. & SPIELMANN, T. 2011. Development and host cell modifications of *Plasmodium falciparum* blood stages in four dimensions. *Nat Commun,* 2:165**,** 1-11.

GULLINGSRUD, J., MILMAN, N., SAVERIA, T., CHESNOKOV, O., WILLIAMSON, K., SRIVASTAVA, A., GAMAIN, B., DUFFY, P. E. & OLEINIKOV, A. V. 2015. High-throughput screening platform identifies small molecules that prevent sequestration of *Plasmodium falciparum*-infected erythrocytes. *J Infect Dis,* 211**,** 1134-43.

GULLINGSRUD, J., SAVERIA, T., AMOS, E., DUFFY, P. E. & OLEINIKOV, A. V. 2013. Structure-function-immunogenicity studies of PfEMP1 domain DBL2betaPF11_0521, a malaria parasite ligand for ICAM-1. *PLoS One,* 8**,** e61323.

GUPTA, S., SNOW, R. W., DONNELLY, C. A., MARSH, K. & NEWBOLD, C. 1999. Immunity to non-cerebral severe malaria is acquired after one or two infections. *Nat Med,* 5**,** 340-3.

HANSSEN, E., CARLTON, P., DEED, S., KLONIS, N., SEDAT, J., DERISI, J. & TILLEY, L. 2010. Whole cell imaging reveals novel modular features of the exomembrane system of the malaria parasite, *Plasmodium falciparum*. *Int J Parasitol,* 40**,** 123-34.

HEDDINI, A., CHEN, Q., OBIERO, J., KAI, O., FERNANDEZ, V., MARSH, K., MULLER, W. A. & WAHLGREN, M. 2001a. Binding of *Plasmodium falciparum*-infected erythrocytes to

soluble platelet endothelial cell adhesion molecule-1 (PECAM-1/CD31): frequent recognition by clinical isolates. *Am J Trop Med Hyg,* 65**,** 47-51.

HEDDINI, A., PETTERSSON, F., KAI, O., SHAFI, J., OBIERO, J., CHEN, Q., BARRAGAN, A., WAHLGREN, M. & MARSH, K. 2001b. Fresh isolates from children with severe *Plasmodium falciparum* malaria bind to multiple receptors. *Infect Immun,* 69**,** 5849-56.

HEIBER, A., KRUSE, F., PICK, C., GRURING, C., FLEMMING, S., OBERLI, A., SCHOELER, H., RETZLAFF, S., MESEN-RAMIREZ, P., HISS, J. A., KADEKOPPALA, M., HECHT, L., HOLDER, A. A., GILBERGER, T. W. & SPIELMANN, T. 2013. Identification of new PNEPs indicates a substantial non-PEXEL exportome and underpins common features in *Plasmodium falciparum* protein export. *PLoS Pathog,* 9**,** e1003546.

HEMINGWAY, J., SHRETTA, R., WELLS, T. N., BELL, D., DJIMDE, A. A., ACHEE, N. & QI, G. 2016. Tools and Strategies for Malaria Control and Elimination: What Do We Need to Achieve a Grand Convergence in Malaria? *PLoS Biol,* 14**,** e1002380.

HIGGINS, M. K. 2008a. Overproduction, purification and crystallization of a chondroitin sulfate A-binding DBL domain from a *Plasmodium falciparum var2csa*-encoded PfEMP1 protein. *Acta Crystallogr Sect F Struct Biol Cryst Commun,* 64**,** 221-3.

HIGGINS, M. K. 2008b. The structure of a chondroitin sulfate-binding domain important in placental malaria. *J Biol Chem,* 283**,** 21842-6.

HIGGINS, M. K. & CARRINGTON, M. 2014. Sequence variation and structural conservation allows development of novel function and immune evasion in parasite surface protein families. *Protein Sci,* 23**,** 354-65.

HILLER, N. L., BHATTACHARJEE, S., VAN OOIJ, C., LIOLIOS, K., HARRISON, T., LOPEZ-ESTRANO, C. & HALDAR, K. 2004. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science,* 306**,** 1934-7.

HOLDING, T. & RECKER, M. 2015. Maintenance of phenotypic diversity within a set of virulence encoding genes of the malaria parasite *Plasmodium falciparum*. *J R Soc Interface,* 12**,** 20150848.

HORA, R., KAPOOR, P., THIND, K. K. & MISHRA, P. C. 2016. Cerebral malaria - clinical manifestations and pathogenesis. *Metab Brain Dis,* 31**,** 225-37.

HORN, D. 2014. Antigenic variation in African trypanosomes. *Mol Biochem Parasitol,* 195**,** 123-9.

HORROCKS, P., PINCHES, R., CHRISTODOULOU, Z., KYES, S. A. & NEWBOLD, C. I. 2004. Variable *var* transition rates underlie antigenic variation in malaria. *Proc Natl Acad Sci U S A,* 101**,** 11129-34.

HORROCKS, P., PINCHES, R. A., CHAKRAVORTY, S. J., PAPAKRIVOS, J., CHRISTODOULOU, Z., KYES, S. A., URBAN, B. C., FERGUSON, D. J. & NEWBOLD, C. I. 2005. PfEMP1 expression is reduced on the surface of knobless *Plasmodium falciparum* infected erythrocytes. *J Cell Sci,* 118**,** 2507-18.

HOWELL, D. P., LEVIN, E. A., SPRINGER, A. L., KRAEMER, S. M., PHIPPARD, D. J., SCHIEF, W. R. & SMITH, J. D. 2008. Mapping a common interaction site used by *Plasmodium falciparum* Duffy binding-like domains to bind diverse host receptors. *Mol Microbiol,* 67**,** 78-87.

HVIID, L. 2011. The case for PfEMP1-based vaccines to protect pregnant women against *Plasmodium falciparum* malaria. *Expert Rev Vaccines,* 10**,** 1405-14.

HVIID, L. & JENSEN, A. T. 2015. PfEMP1 - A Parasite Protein Family of Key Importance in *Plasmodium falciparum* Malaria Immunity and Pathogenesis. *Adv Parasitol,* 88**,** 51-84.

IDRO, R., KAKOOZA-MWESIGE, A., BALYEJJUSSA, S., MIREMBE, G., MUGASHA, C., TUGUMISIRIZE, J. & BYARUGABA, J. 2010. Severe neurological sequelae and

behaviour problems after cerebral malaria in Ugandan children. *BMC Res Notes,* 3:104**,** 1-6.

JANES, J. H., WANG, C. P., LEVIN-EDENS, E., VIGAN-WOMAS, I., GUILLOTTE, M., MELCHER, M., MERCEREAU-PUIJALON, O. & SMITH, J. D. 2011. Investigating the host binding signature on the *Plasmodium falciparum* PfEMP1 protein family. *PLoS Pathog,* 7**,** e1002032.

JENSEN, A. T., MAGISTRADO, P., SHARP, S., JOERGENSEN, L., LAVSTSEN, T., CHIUCCHIUINI, A., SALANTI, A., VESTERGAARD, L. S., LUSINGU, J. P., HERMSEN, R., SAUERWEIN, R., CHRISTENSEN, J., NIELSEN, M. A., HVIID, L., SUTHERLAND, C., STAALSOE, T. & THEANDER, T. G. 2004. *Plasmodium falciparum* associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A *var* genes. *J Exp Med,* 199**,** 1179-90.

JIANG, L., MU, J., ZHANG, Q., NI, T., SRINIVASAN, P., RAYAVARA, K., YANG, W., TURNER, L., LAVSTSEN, T., THEANDER, T. G., PENG, W., WEI, G., JING, Q., WAKABAYASHI, Y., BANSAL, A., LUO, Y., RIBEIRO, J. M., SCHERF, A., ARAVIND, L., ZHU, J., ZHAO, K. & MILLER, L. H. 2013. PfSETvs methylation of histone H3K36 represses virulence genes in *Plasmodium falciparum*. *Nature,* 499**,** 223-7.

JOANNIN, N., ABHIMAN, S., SONNHAMMER, E. L. & WAHLGREN, M. 2008. Sub-grouping and sub-functionalization of the RIFIN multi-copy protein family. *BMC Genomics,* 9:19**,** 1-14.

KAESTLI, M., COCKBURN, I. A., CORTES, A., BAEA, K., ROWE, J. A. & BECK, H. P. 2006. Virulence of malaria is associated with differential expression of *Plasmodium falciparum var* gene subgroups in a case-control study. *J Infect Dis,* 193**,** 1567-74.

KAUL, D. K., ROTH, E. F., JR., NAGEL, R. L., HOWARD, R. J. & HANDUNNETTI, S. M. 1991. Rosetting of *Plasmodium falciparum*-infected red blood cells with uninfected red blood cells enhances microvascular obstruction under flow conditions. *Blood,* 78**,** 812-9.

KAVIRATNE, M., KHAN, S. M., JARRA, W. & PREISER, P. R. 2002. Small variant STEVOR antigen is uniquely located within Maurer's clefts in *Plasmodium falciparum*-infected red blood cells. *Eukaryot Cell,* 1**,** 926-35.

KHUNRAE, P., DAHLBACK, M., NIELSEN, M. A., ANDERSEN, G., DITLEV, S. B., RESENDE, M., PINTO, V. V., THEANDER, T. G., HIGGINS, M. K. & SALANTI, A. 2010. Full-length recombinant *Plasmodium falciparum* VAR2CSA binds specifically to CSPG and induces potent parasite adhesion-blocking antibodies. *J Mol Biol,* 397**,** 826-34.

KHUNRAE, P., PHILIP, J. M., BULL, D. R. & HIGGINS, M. K. 2009. Structural comparison of two CSPG-binding DBL domains from the VAR2CSA protein important in malaria during pregnancy. *J Mol Biol,* 393**,** 202-13.

KIBBE, W. A. 2007. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res,* 35.

KILEJIAN, A., RASHID, M. A., AIKAWA, M., AJI, T. & YANG, Y. F. 1991. Selective association of a fragment of the knob protein with spectrin, actin and the red cell membrane. *Mol Biochem Parasitol,* 44**,** 175-81.

KIRCHGATTER, K. & DEL PORTILLO, H. A. 2002. Association of severe noncerebral *Plasmodium falciparum* malaria in Brazil with expressed PfEMP1 DBL1 alpha sequences lacking cysteine residues. *Mol Med,* 8**,** 16-23.

KLEIN, M. M., GITTIS, A. G., SU, H. P., MAKOBONGO, M. O., MOORE, J. M., SINGH, S., MILLER, L. H. & GARBOCZI, D. N. 2008. The cysteine-rich interdomain region from the highly variable *Plasmodium falciparum* erythrocyte membrane protein-1 exhibits a conserved structure. *PLoS Pathog,* 4**,** e1000147.

KRAEMER, S. M., KYES, S. A., AGGARWAL, G., SPRINGER, A. L., NELSON, S. O., CHRISTODOULOU, Z., SMITH, L. M., WANG, W., LEVIN, E., NEWBOLD, C. I., MYLER,

P. J. & SMITH, J. D. 2007. Patterns of gene recombination shape *var* gene repertoires in *Plasmodium falciparum*: comparisons of geographically diverse isolates. *BMC Genomics,* 8**,** 45.

KRAEMER, S. M. & SMITH, J. D. 2003. Evidence for the importance of genetic structuring to the structural and functional specialization of the *Plasmodium falciparum var* gene family. *Mol Microbiol,* 50**,** 1527-38.

KULZER, S., CHARNAUD, S., DAGAN, T., RIEDEL, J., MANDAL, P., PESCE, E. R., BLATCH, G. L., CRABB, B. S., GILSON, P. R. & PRZYBORSKI, J. M. 2012. *Plasmodium falciparum*-encoded exported hsp70/hsp40 chaperone/co-chaperone complexes within the host erythrocyte. *Cell Microbiol,* 14**,** 1784-95.

KULZER, S., RUG, M., BRINKMANN, K., CANNON, P., COWMAN, A., LINGELBACH, K., BLATCH, G. L., MAIER, A. G. & PRZYBORSKI, J. M. 2010. Parasite-encoded Hsp40 proteins define novel mobile structures in the cytosol of the *P. falciparum*-infected erythrocyte. *Cell Microbiol,* 12**,** 1398-420.

KUN, J. F., KLABUNDE, J., LELL, B., LUCKNER, D., ALPERS, M., MAY, J., MEYER, C. & KREMSNER, P. G. 1999. Association of the ICAM-1Kilifi mutation with protection against severe malaria in Lambarene, Gabon. *Am J Trop Med Hyg,* 61**,** 776-9.

KUN, J. F., SCHMIDT-OTT, R. J., LEHMAN, L. G., LELL, B., LUCKNER, D., GREVE, B., MATOUSEK, P. & KREMSNER, P. G. 1998. Merozoite surface antigen 1 and 2 genotypes and rosetting of *Plasmodium falciparum* in severe and mild malaria in Lambarene, Gabon. *Trans R Soc Trop Med Hyg,* 92**,** 110-4.

KYES, S., PINCHES, R. & NEWBOLD, C. 2000. A simple RNA analysis method shows *var* and *rif* multigene family expression patterns in *Plasmodium falciparum*. *Mol Biochem Parasitol,* 105**,** 311-5.

KYES, S. A., ROWE, J. A., KRIEK, N. & NEWBOLD, C. I. 1999. Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with *Plasmodium falciparum*. *Proc Natl Acad Sci U S A,* 96**,** 9333-8.

KYRIACOU, H. M., STONE, G. N., CHALLIS, R. J., RAZA, A., LYKE, K. E., THERA, M. A., KONE, A. K., DOUMBO, O. K., PLOWE, C. V. & ROWE, J. A. 2006. Differential *var* gene transcription in *Plasmodium falciparum* isolates from patients with cerebral malaria compared to hyperparasitaemia. *Mol Biochem Parasitol,* 150**,** 211-8.

LANGUINO, L. R., PLESCIA, J., DUPERRAY, A., BRIAN, A. A., PLOW, E. F., GELTOSKY, J. E. & ALTIERI, D. C. 1993. Fibrinogen mediates leukocyte adhesion to vascular endothelium through an ICAM-1-dependent pathway. *Cell,* 73**,** 1423-34.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and clustal X version 2.0. *Bioinformatics,* 23**,** 2947-2948.

LASZIK, Z., MITRO, A., TAYLOR, F. B., JR., FERRELL, G. & ESMON, C. T. 1997. Human protein C receptor is present primarily on endothelium of large blood vessels: implications for the control of the protein C pathway. *Circulation,* 96**,** 3633-40.

LAU, C. K., TURNER, L., JESPERSEN, J. S., LOWE, E. D., PETERSEN, B., WANG, C. W., PETERSEN, J. E., LUSINGU, J., THEANDER, T. G., LAVSTSEN, T. & HIGGINS, M. K. 2015. Structural conservation despite huge sequence diversity allows EPCR binding by the PfEMP1 family implicated in severe childhood malaria. *Cell Host Microbe,* 17**,** 118-29.

LAVAZEC, C., SANYAL, S. & TEMPLETON, T. J. 2006. Hypervariability within the Rifin, Stevor and Pfmc-2TM superfamilies in *Plasmodium falciparum*. *Nucleic Acids Res,* 34**,** 6696-707.

LAVSTSEN, T., SALANTI, A., JENSEN, A. T., ARNOT, D. E. & THEANDER, T. G. 2003. Sub-grouping of *Plasmodium falciparum* 3D7 *var* genes based on sequence analysis of coding and non-coding regions. *Malar J,* 2:27**,** 1-14.

LAVSTSEN, T., TURNER, L., SAGUTI, F., MAGISTRADO, P., RASK, T. S., JESPERSEN, J. S., WANG, C. W., BERGER, S. S., BARAKA, V., MARQUARD, A. M., SEGUIN-ORLANDO, A., WILLERSLEV, E., GILBERT, M. T., LUSINGU, J. & THEANDER, T. G. 2012. *Plasmodium falciparum* erythrocyte membrane protein 1 domain cassettes 8 and 13 are associated with severe malaria in children. *Proc Natl Acad Sci U S A,* 109**,** E1791-800.

LAWSON, C. & WOLF, S. 2009. ICAM-1 signaling in endothelial cells. *Pharmacol Rep,* 61**,** 22-32.

LEITGEB, A. M., BLOMQVIST, K., CHO-NGWA, F., SAMJE, M., NDE, P., TITANJI, V. & WAHLGREN, M. 2011. Low anticoagulant heparin disrupts *Plasmodium falciparum* rosettes in fresh clinical isolates. *Am J Trop Med Hyg,* 84**,** 390-6.

LENNARTZ, F., BENGTSSON, A., OLSEN, R. W., JOERGENSEN, L., BROWN, A., REMY, L., MAN, P., FOREST, E., BARFOD, L. K., ADAMS, Y., HIGGINS, M. K. & JENSEN, A. T. 2015. Mapping the Binding Site of a Cross-Reactive *Plasmodium falciparum* PfEMP1 Monoclonal Antibody Inhibitory of ICAM-1 Binding. *J Immunol,* 195**,** 3273-83.

LETUNIC, I. & BORK, P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics,* 23**,** 127-128.

LETUNIC, I. & BORK, P. 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Research,* 39**,** 475-478.

LOPEZ-RUBIO, J. J., GONTIJO, A. M., NUNES, M. C., ISSAR, N., HERNANDEZ RIVAS, R. & SCHERF, A. 2007. 5' flanking region of *var* genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol Microbiol,* 66**,** 1296-305.

LOPEZ-RUBIO, J. J., MANCIO-SILVA, L. & SCHERF, A. 2009. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe,* 5**,** 179-90.

LUCAS, J. Z. & SHERMAN, I. W. 1998. *Plasmodium falciparum*: thrombospondin mediates parasitized erythrocyte band 3-related adhesin binding. *Exp Parasitol,* 89**,** 78-85.

LUSINGU, J. P., JENSEN, A. T., VESTERGAARD, L. S., MINJA, D. T., DALGAARD, M. B., GESASE, S., MMBANDO, B. P., KITUA, A. Y., LEMNGE, M. M., CAVANAGH, D., HVIID, L. & THEANDER, T. G. 2006. Levels of plasma immunoglobulin G with specificity against the cysteine-rich interdomain regions of a semiconserved *Plasmodium falciparum* erythrocyte membrane protein 1, VAR4, predict protection against malarial anemia and febrile episodes. *Infect Immun,* 74**,** 2867-75.

MACPHERSON, G. G., WARRELL, M. J., WHITE, N. J., LOOAREESUWAN, S. & WARRELL, D. A. 1985. Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. *Am J Pathol,* 119**,** 385-401.

MADKHALI, A. M. 2015. *Adhesion Analysis of Different PfEMP1 Variants to CD36, ICAM-1 and Primary Endothelial Cells.* PhD, University of Liverpool.

MADKHALI, A. M., ALKURBI, M. O., SZESTAK, T., BENGTSSON, A., PATIL, P. R., WU, Y., ALHARTHI, S., JENSEN, A. T., PLEASS, R. & CRAIG, A. G. 2014. An analysis of the binding characteristics of a panel of recently selected ICAM-1 binding *Plasmodium falciparum* patient isolates. *PLoS One,* 9**,** e111518.

MAGISTRADO, P. A., LUSINGU, J., VESTERGAARD, L. S., LEMNGE, M., LAVSTSEN, T., TURNER, L., HVIID, L., JENSEN, A. T. & THEANDER, T. G. 2007. Immunoglobulin G antibody reactivity to a group A *Plasmodium falciparum* erythrocyte membrane protein 1 and protection from *P. falciparum* malaria. *Infect Immun,* 75**,** 2415-20.

MAIER, A. G., COOKE, B. M., COWMAN, A. F. & TILLEY, L. 2009. Malaria parasite proteins that remodel the host erythrocyte. *Nat Rev Microbiol,* 7**,** 341-54.

MAIER, A. G., RUG, M., O'NEILL, M. T., BEESON, J. G., MARTI, M., REEDER, J. & COWMAN, A. F. 2007. Skeleton-binding protein 1 functions at the parasitophorous vacuole membrane to traffic PfEMP1 to the *Plasmodium falciparum*-infected erythrocyte surface. *Blood,* 109**,** 1289-97.

MANSKE, M., MIOTTO, O., CAMPINO, S., AUBURN, S., ALMAGRO-GARCIA, J., MASLEN, G., O'BRIEN, J., DJIMDE, A., DOUMBO, O., ZONGO, I., OUEDRAOGO, J. B., MICHON, P., MUELLER, I., SIBA, P., NZILA, A., BORRMANN, S., KIARA, S. M., MARSH, K., JIANG, H., SU, X. Z., AMARATUNGA, C., FAIRHURST, R., SOCHEAT, D., NOSTEN, F., IMWONG, M., WHITE, N. J., SANDERS, M., ANASTASI, E., ALCOCK, D., DRURY, E., OYOLA, S., QUAIL, M. A., TURNER, D. J., RUANO-RUBIO, V., JYOTHI, D., AMENGA-ETEGO, L., HUBBART, C., JEFFREYS, A., ROWLANDS, K., SUTHERLAND, C., ROPER, C., MANGANO, V., MODIANO, D., TAN, J. C., FERDIG, M. T., AMAMBUA-NGWA, A., CONWAY, D. J., TAKALA-HARRISON, S., PLOWE, C. V., RAYNER, J. C., ROCKETT, K. A., CLARK, T. G., NEWBOLD, C. I., BERRIMAN, M., MACINNIS, B. & KWIATKOWSKI, D. P. 2012. Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. *Nature,* 487**,** 375-9.

MARLIN, S. D. & SPRINGER, T. A. 1987. Purified intercellular adhesion molecule-1 (ICAM-1) is a ligand for lymphocyte function-associated antigen 1 (LFA-1). *Cell,* 51**,** 813-9.

MARSH, K., FORSTER, D., WARUIRU, C., MWANGI, I., WINSTANLEY, M., MARSH, V., NEWTON, C., WINSTANLEY, P., WARN, P., PESHU, N. & ET AL. 1995. Indicators of life-threatening malaria in African children. *N Engl J Med,* 332**,** 1399-404.

MARSH, K. & HOWARD, R. J. 1986. Antigens induced on erythrocytes by *P. falciparum*: expression of diverse and conserved determinants. *Science,* 231**,** 150-3.

MARSH, K. & KINYANJUI, S. 2006. Immune effector mechanisms in malaria. *Parasite Immunol,* 28**,** 51-60.

MARSH, K., MARSH, V. M., BROWN, J., WHITTLE, H. C. & GREENWOOD, B. M. 1988. *Plasmodium falciparum*: the behavior of clinical isolates in an *in vitro* model of infected red blood cell sequestration. *Exp Parasitol,* 65**,** 202-8.

MARSH, K., OTOO, L., HAYES, R. J., CARSON, D. C. & GREENWOOD, B. M. 1989. Antibodies to blood stage antigens of *Plasmodium falciparum* in rural Gambians and their relation to protection against infection. *Trans R Soc Trop Med Hyg,* 83**,** 293-303.

MARTI, M., GOOD, R. T., RUG, M., KNUEPFER, E. & COWMAN, A. F. 2004. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science,* 306**,** 1930-3.

MARTY, A. J., THOMPSON, J. K., DUFFY, M. F., VOSS, T. S., COWMAN, A. F. & CRABB, B. S. 2006. Evidence that *Plasmodium falciparum* chromosome end clusters are cross-linked by protein and are the sites of both virulence gene silencing and activation. *Mol Microbiol,* 62**,** 72-83.

MAYER, C., SLATER, L., ERAT, M. C., KONRAT, R. & VAKONAKIS, I. 2012. Structural analysis of the *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) intracellular domain reveals a conserved interaction epitope. *J Biol Chem,* 287**,** 7182-9.

MCCORMICK, C. J., CRAIG, A., ROBERTS, D., NEWBOLD, C. I. & BERENDT, A. R. 1997. Intercellular adhesion molecule-1 and CD36 synergize to mediate adherence of *Plasmodium falciparum*-infected erythrocytes to cultured human microvascular endothelial cells. *J Clin Invest,* 100**,** 2521-9.

MCGREGOR, I. A. 1964. Studies in the Acquisition of Immunity of *Plasmodium Falciparum* Infections in Africa. *Trans R Soc Trop Med Hyg,* 58**,** 80-92.

MCMILLAN, P. J., MILLET, C., BATINOVIC, S., MAIORCA, M., HANSSEN, E., KENNY, S., MUHLE, R. A., MELCHER, M., FIDOCK, D. A., SMITH, J. D., DIXON, M. W. & TILLEY, L.

2013. Spatial and temporal mapping of the PfEMP1 export pathway in *Plasmodium falciparum*. *Cell Microbiol,* 15**,** 1401-18.

MERRICK, C. J., DZIKOWSKI, R., IMAMURA, H., CHUANG, J., DEITSCH, K. & DURAISINGH, M. T. 2010. The effect of *Plasmodium falciparum* Sir2a histone deacetylase on clonal and longitudinal variation in expression of the *var* family of virulence genes. *Int J Parasitol,* 40**,** 35-43.

MIOTTO, O., ALMAGRO-GARCIA, J., MANSKE, M., MACINNIS, B., CAMPINO, S., ROCKETT, K. A., AMARATUNGA, C., LIM, P., SUON, S., SRENG, S., ANDERSON, J. M., DUONG, S., NGUON, C., CHUOR, C. M., SAUNDERS, D., SE, Y., LON, C., FUKUDA, M. M., AMENGA-ETEGO, L., HODGSON, A. V., ASOALA, V., IMWONG, M., TAKALA-HARRISON, S., NOSTEN, F., SU, X. Z., RINGWALD, P., ARIEY, F., DOLECEK, C., HIEN, T. T., BONI, M. F., THAI, C. Q., AMAMBUA-NGWA, A., CONWAY, D. J., DJIMDE, A. A., DOUMBO, O. K., ZONGO, I., OUEDRAOGO, J. B., ALCOCK, D., DRURY, E., AUBURN, S., KOCH, O., SANDERS, M., HUBBART, C., MASLEN, G., RUANO-RUBIO, V., JYOTHI, D., MILES, A., O'BRIEN, J., GAMBLE, C., OYOLA, S. O., RAYNER, J. C., NEWBOLD, C. I., BERRIMAN, M., SPENCER, C. C., MCVEAN, G., DAY, N. P., WHITE, N. J., BETHELL, D., DONDORP, A. M., PLOWE, C. V., FAIRHURST, R. M. & KWIATKOWSKI, D. P. 2013. Multiple populations of artemisinin-resistant *Plasmodium falciparum* in Cambodia. *Nat Genet,* 45**,** 648-55.

MOLYNEUX, M. E., TAYLOR, T. E., WIRIMA, J. J. & BORGSTEIN, A. 1989. Clinical features and prognostic indicators in paediatric cerebral malaria: a study of 131 comatose Malawian children. *Q J Med,* 71**,** 441-59.

MONTGOMERY, J., MPHANDE, F. A., BERRIMAN, M., PAIN, A., ROGERSON, S. J., TAYLOR, T. E., MOLYNEUX, M. E. & CRAIG, A. 2007. Differential *var* gene expression in the organs of patients dying of falciparum malaria. *Mol Microbiol,* 65**,** 959-67.

MOXON, C. A., WASSMER, S. C., MILNER, D. A., JR., CHISALA, N. V., TAYLOR, T. E., SEYDEL, K. B., MOLYNEUX, M. E., FARAGHER, B., ESMON, C. T., DOWNEY, C., TOH, C. H., CRAIG, A. G. & HEYDERMAN, R. S. 2013. Loss of endothelial protein C receptors links coagulation and inflammation to parasite sequestration in cerebral malaria in African children. *Blood,* 122**,** 842-51.

MPHANDE, F. A., RIBACKE, U., KANEKO, O., KIRONDE, F., WINTER, G. & WAHLGREN, M. 2008. SURFIN4.1, a schizont-merozoite associated protein in the SURFIN family of *Plasmodium falciparum*. *Malar J,* 7:116**,** 1-12.

MUGASA, J., QI, W., RUSCH, S., ROTTMANN, M. & BECK, H. P. 2012. Genetic diversity of expressed *Plasmodium falciparum var* genes from Tanzanian children with severe malaria. *Malar J,* 11**,** 230.

MUNDWILER-PACHLATKO, E. & BECK, H. P. 2013. Maurer's clefts, the enigma of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A,* 110**,** 19987-94.

NEWBOLD, C., WARN, P., BLACK, G., BERENDT, A., CRAIG, A., SNOW, B., MSOBO, M., PESHU, N. & MARSH, K. 1997. Receptor-specific adhesion and clinical disease in *Plasmodium falciparum*. *Am J Trop Med Hyg,* 57**,** 389-98.

NEWBOLD, C. I., PINCHES, R., ROBERTS, D. J. & MARSH, K. 1992. *Plasmodium falciparum*: the human agglutinating antibody response to the infected red cell surface is predominantly variant specific. *Exp Parasitol,* 75**,** 281-92.

NEWTON, C. R., TAYLOR, T. E. & WHITTEN, R. O. 1998. Pathophysiology of fatal falciparum malaria in African children. *Am J Trop Med Hyg,* 58**,** 673-83.

NIANG, M., BEI, A. K., MADNANI, K. G., PELLY, S., DANKWA, S., KANJEE, U., GUNALAN, K., AMALADOSS, A., YEO, K. P., BOB, N. S., MALLERET, B., DURAISINGH, M. T. & PREISER, P. R. 2014. STEVOR is a *Plasmodium falciparum* erythrocyte binding protein that mediates merozoite invasion and rosetting. *Cell Host Microbe,* 16**,** 81-93.

NIANG, M., YAN YAM, X. & PREISER, P. R. 2009. The *Plasmodium falciparum* STEVOR multigene family mediates antigenic variation of the infected erythrocyte. *PLoS Pathog,* 5**,** e1000307.

NOBLE, R., CHRISTODOULOU, Z., KYES, S., PINCHES, R., NEWBOLD, C. I. & RECKER, M. 2013. The antigenic switching network of *Plasmodium falciparum* and its implications for the immuno-epidemiology of malaria. *Elife,* 2**,** e01074.

NOTREDAME, C., HIGGINS, D. G. & HERINGA, J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol,* 302**,** 205-17.

OBERLI, A., SLATER, L. M., CUTTS, E., BRAND, F., MUNDWILER-PACHLATKO, E., RUSCH, S., MASIK, M. F., ERAT, M. C., BECK, H. P. & VAKONAKIS, I. 2014. A *Plasmodium falciparum* PHIST protein binds the virulence factor PfEMP1 and comigrates to knobs on the host cell surface. *FASEB J,* 28**,** 4420-33.

OBERLI, A., ZURBRUGG, L., RUSCH, S., BRAND, F., BUTLER, M. E., DAY, J. L., CUTTS, E. E., LAVSTSEN, T., VAKONAKIS, I. & BECK, H. P. 2016. *Plasmodium falciparum* PHIST Proteins Contribute to Cytoadherence and Anchor PfEMP1 to the Host Cell Cytoskeleton. *Cell Microbiol,* [Epub ahead of print].

OCHOLA, L. B., SIDDONDO, B. R., OCHOLLA, H., NKYA, S., KIMANI, E. N., WILLIAMS, T. N., MAKALE, J. O., LILJANDER, A., URBAN, B. C., BULL, P. C., SZESTAK, T., MARSH, K. & CRAIG, A. G. 2011. Specific receptor usage in *Plasmodium falciparum* cytoadherence is associated with disease outcome. *PLoS One,* 6**,** e14741.

OCKENHOUSE, C. F., BETAGERI, R., SPRINGER, T. A. & STAUNTON, D. E. 1992a. *Plasmodium falciparum*-infected erythrocytes bind ICAM-1 at a site distinct from LFA-1, Mac-1, and human rhinovirus. *Cell,* 68**,** 63-9.

OCKENHOUSE, C. F., TANDON, N. N., MAGOWAN, C., JAMIESON, G. A. & CHULAY, J. D. 1989. Identification of a platelet membrane glycoprotein as a falciparum malaria sequestration receptor. *Science,* 243**,** 1469-71.

OCKENHOUSE, C. F., TEGOSHI, T., MAENO, Y., BENJAMIN, C., HO, M., KAN, K. E., THWAY, Y., WIN, K., AIKAWA, M. & LOBB, R. R. 1992b. Human vascular endothelial cell adhesion receptors for *Plasmodium falciparum*-infected erythrocytes: roles for endothelial leukocyte adhesion molecule 1 and vascular cell adhesion molecule 1. *J Exp Med,* 176**,** 1183-9.

OH, H. M., KWON, M. S., KIM, H. J., JEON, B. H., KIM, H. R., CHOI, H. O., NA, B. R., EOM, S. H., SONG, N. W. & JUN, C. D. 2011. Intermediate monomer-dimer equilibrium structure of native ICAM-1: implication for enhanced cell adhesion. *Exp Cell Res,* 317**,** 163-72.

OH, S. S., VOIGT, S., FISHER, D., YI, S. J., LEROY, P. J., DERICK, L. H., LIU, S. & CHISHTI, A. H. 2000. *Plasmodium falciparum* erythrocyte membrane protein 1 is anchored to the actin-spectrin junction and knob-associated histidine-rich protein in the erythrocyte skeleton. *Mol Biochem Parasitol,* 108**,** 237-47.

OHASHI, J., NAKA, I., PATARAPOTIKUL, J., HANANANTACHAI, H., LOOAREESUWAN, S. & TOKUNAGA, K. 2001. Absence of association between the allele coding methionine at position 29 in the N-terminal domain of ICAM-1 (ICAM-1(Kilifi)) and severe malaria in the northwest of Thailand. *Jpn J Infect Dis,* 54**,** 114-6.

OLEINIKOV, A. V., AMOS, E., FRYE, I. T., ROSSNAGLE, E., MUTABINGWA, T. K., FRIED, M. & DUFFY, P. E. 2009. High throughput functional assays of the variant antigen PfEMP1 reveal a single domain in the 3D7 *Plasmodium falciparum* genome that binds ICAM1 with high affinity and is targeted by naturally acquired neutralizing antibodies. *PLoS Pathog,* 5**,** e1000386.

OLEINIKOV, A. V., ROSSNAGLE, E., FRANCIS, S., MUTABINGWA, T. K., FRIED, M. & DUFFY, P. E. 2007. Effects of sex, parity, and sequence variation on seroreactivity to candidate pregnancy malaria vaccine antigens. *J Infect Dis,* 196**,** 155-64.

PAIN, A., FERGUSON, D. J., KAI, O., URBAN, B. C., LOWE, B., MARSH, K. & ROBERTS, D. J. 2001. Platelet-mediated clumping of *Plasmodium falciparum*-infected erythrocytes is a common adhesive phenotype and is associated with severe malaria. *Proc Natl Acad Sci U S A,* 98**,** 1805-10.

PATIL, P. R., GEMMA, S., CAMPIANI, G. & CRAIG, A. G. 2011. Broad inhibition of *Plasmodium falciparum* cytoadherence by (+)-epigallocatechin gallate. *Malar J,* 10:348**,** 1-4.

PEI, X., AN, X., GUO, X., TARNAWSKI, M., COPPEL, R. & MOHANDAS, N. 2005. Structural and functional studies of interaction between *Plasmodium falciparum* knob-associated histidine-rich protein (KAHRP) and erythrocyte spectrin. *J Biol Chem,* 280**,** 31166-71.

PEREZ-TOLEDO, K., ROJAS-MEZA, A. P., MANCIO-SILVA, L., HERNANDEZ-CUEVAS, N. A., DELGADILLO, D. M., VARGAS, M., MARTINEZ-CALVILLO, S., SCHERF, A. & HERNANDEZ-RIVAS, R. 2009. *Plasmodium falciparum* heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of *var* genes. *Nucleic Acids Res,* 37**,** 2596-606.

PETERS, J. M., FOWLER, E. V., KRAUSE, D. R., CHENG, Q. & GATTON, M. L. 2007. Differential changes in *Plasmodium falciparum var* transcription during adaptation to culture. *J Infect Dis,* 195**,** 748-55.

PETTER, M., HAEGGSTROM, M., KHATTAB, A., FERNANDEZ, V., KLINKERT, M. Q. & WAHLGREN, M. 2007. Variant proteins of the *Plasmodium falciparum* RIFIN family show distinct subcellular localization and developmental expression patterns. *Mol Biochem Parasitol,* 156**,** 51-61.

PETTER, M., SELVARAJAH, S. A., LEE, C. C., CHIN, W. H., GUPTA, A. P., BOZDECH, Z., BROWN, G. V. & DUFFY, M. F. 2013. H2A.Z and H2B.Z double-variant nucleosomes define intergenic regions and dynamically occupy *var* gene promoters in the malaria parasite *Plasmodium falciparum*. *Mol Microbiol,* 87**,** 1167-82.

PLACMALVAC. Available: http://cmp.ku.dk/collaborations/placmalvac/.

PLEASS, R. J., MOORE, S. C., STEVENSON, L. & HVIID, L. 2016. Immunoglobulin M: Restrainer of Inflammation and Mediator of Immune Evasion by *Plasmodium falciparum* Malaria. *Trends Parasitol,* 32**,** 108-19.

POYOMTIP, T., SUWANDITTAKUL, N., SITTHICHOT, N., KHOSITNITHIKUL, R., TAN-ARIYA, P. & MUNGTHIN, M. 2012. Polymorphisms of the *pfmdr1* but not the *pfnhe-1* gene is associated with *in vitro* quinine sensitivity in Thai isolates of *Plasmodium falciparum*. *Malar J,* 11:7**,** 1-7.

PROELLOCKS, N. I., HERRMANN, S., BUCKINGHAM, D. W., HANSSEN, E., HODGES, E. K., ELSWORTH, B., MORAHAN, B. J., COPPEL, R. L. & COOKE, B. M. 2014. A lysine-rich membrane-associated PHISTb protein involved in alteration of the cytoadhesive properties of *Plasmodium falciparum*-infected red blood cells. *FASEB J,* 28**,** 3103-13.

PRUCCA, C. G., RIVERO, F. D. & LUJAN, H. D. 2011. Regulation of antigenic variation in *Giardia lamblia*. *Annu Rev Microbiol,* 65**,** 611-30.

RALPH, S. A., SCHEIDIG-BENATAR, C. & SCHERF, A. 2005. Antigenic variation in *Plasmodium falciparum* is associated with movement of *var* loci between subnuclear locations. *Proc Natl Acad Sci U S A,* 102**,** 5414-9.

RANFORD-CARTWRIGHT, L. C. & MWANGI, J. M. 2012. Analysis of malaria parasite phenotypes using experimental genetic crosses of *Plasmodium falciparum*. *Int J Parasitol,* 42**,** 529-34.

RANSON, H. & LISSENDEN, N. 2016. Insecticide resistance in African anopheles mosquitoes: A worsening situation that needs urgent action to maintain malaria control. *Trends Parasitol,* 32**,** 187-96.

RASK, T. S., HANSEN, D. A., THEANDER, T. G., GORM PEDERSEN, A. & LAVSTSEN, T. 2010. *Plasmodium falciparum* erythrocyte membrane protein 1 diversity in seven genomes--divide and conquer. *PLoS Comput Biol,* 6**,** 1-23.

RECKER, M., BUCKEE, C. O., SERAZIN, A., KYES, S., PINCHES, R., CHRISTODOULOU, Z., SPRINGER, A. L., GUPTA, S. & NEWBOLD, C. I. 2011. Antigenic variation in *Plasmodium falciparum* malaria involves a highly structured switching pattern. *PLoS Pathog,* 7**,** e1001306.

RICHARDS, J. S. & BEESON, J. G. 2009. The future for blood-stage vaccines against malaria. *Immunol Cell Biol,* 87**,** 377-90.

RICKE, C. H., STAALSOE, T., KORAM, K., AKANMORI, B. D., RILEY, E. M., THEANDER, T. G. & HVIID, L. 2000. Plasma antibodies from malaria-exposed pregnant women recognize variant surface antigens on *Plasmodium falciparum*-infected erythrocytes in a parity-dependent manner and block parasite adhesion to chondroitin sulfate A. *J Immunol,* 165**,** 3309-16.

RIEGER, H., YOSHIKAWA, H. Y., QUADT, K., NIELSEN, M. A., SANCHEZ, C. P., SALANTI, A., TANAKA, M. & LANZER, M. 2015. Cytoadhesion of *Plasmodium falciparum*-infected erythrocytes to chondroitin-4-sulfate is cooperative and shear enhanced. *Blood,* 125**,** 383-91.

ROBINSON, B. A., WELCH, T. L. & SMITH, J. D. 2003. Widespread functional specialization of *Plasmodium falciparum* erythrocyte membrane protein 1 family members to bind CD36 analysed across a parasite genome. *Mol Microbiol,* 47**,** 1265-78.

ROGERSON, S. J., HVIID, L., DUFFY, P. E., LEKE, R. F. & TAYLOR, D. W. 2007. Malaria in pregnancy: pathogenesis and immunity. *Lancet Infect Dis,* 7**,** 105-17.

ROGERSON, S. J., TEMBENU, R., DOBANO, C., PLITT, S., TAYLOR, T. E. & MOLYNEUX, M. E. 1999. Cytoadherence characteristics of *Plasmodium falciparum*-infected erythrocytes from Malawian children with severe and uncomplicated malaria. *Am J Trop Med Hyg,* 61**,** 467-72.

RORICK, M. M., RASK, T. S., BASKERVILLE, E. B., DAY, K. P. & PASCUAL, M. 2013. Homology blocks of *Plasmodium falciparum var* genes and clinically distinct forms of severe malaria in a local population. *BMC Microbiol,* 13:244**,** 1-14.

ROTTMANN, M., LAVSTSEN, T., MUGASA, J. P., KAESTLI, M., JENSEN, A. T., MULLER, D., THEANDER, T. & BECK, H. P. 2006. Differential expression of *var* gene groups is associated with morbidity caused by *Plasmodium falciparum* infection in Tanzanian children. *Infect Immun,* 74**,** 3904-11.

ROWE, A., OBEIRO, J., NEWBOLD, C. I. & MARSH, K. 1995. *Plasmodium falciparum* rosetting is associated with malaria severity in Kenya. *Infect Immun,* 63**,** 2323-6.

ROWE, J. A., CLAESSENS, A., CORRIGAN, R. A. & ARMAN, M. 2009. Adhesion of *Plasmodium falciparum*-infected erythrocytes to human cells: molecular mechanisms and therapeutic implications. *Expert Rev Mol Med,* 11**,** e16.

ROWE, J. A., MOULDS, J. M., NEWBOLD, C. I. & MILLER, L. H. 1997. *P. falciparum* rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature,* 388**,** 292-5.

ROWE, J. A., ROGERSON, S. J., RAZA, A., MOULDS, J. M., KAZATCHKINE, M. D., MARSH, K., NEWBOLD, C. I., ATKINSON, J. P. & MILLER, L. H. 2000. Mapping of the region of complement receptor (CR) 1 required for *Plasmodium falciparum* rosetting and demonstration of the importance of CR1 in rosetting in field isolates. *J Immunol,* 165**,** 6341-6.

RUG, M., CYRKLAFF, M., MIKKONEN, A., LEMGRUBER, L., KUELZER, S., SANCHEZ, C. P., THOMPSON, J., HANSSEN, E., O'NEILL, M., LANGER, C., LANZER, M., FRISCHKNECHT, F., MAIER, A. G. & COWMAN, A. F. 2014. Export of virulence proteins by malaria-

infected erythrocytes involves remodeling of host actin cytoskeleton. *Blood,* 124**,** 3459-68.

RUSSO, I., BABBITT, S., MURALIDHARAN, V., BUTLER, T., OKSMAN, A. & GOLDBERG, D. E. 2010. Plasmepsin V licenses *Plasmodium* proteins for export into the host erythrocyte. *Nature,* 463**,** 632-6.

SALANTI, A., STAALSOE, T., LAVSTSEN, T., JENSEN, A. T., SOWA, M. P., ARNOT, D. E., HVIID, L. & THEANDER, T. G. 2003. Selective upregulation of a single distinctly structured *var* gene in chondroitin sulphate A-adhering *Plasmodium falciparum* involved in pregnancy-associated malaria. *Mol Microbiol,* 49**,** 179-91.

SAM-YELLOWE, T. Y., FLORENS, L., JOHNSON, J. R., WANG, T., DRAZBA, J. A., LE ROCH, K. G., ZHOU, Y., BATALOV, S., CARUCCI, D. J., WINZELER, E. A. & YATES, J. R., 3RD 2004. A *Plasmodium* gene family encoding Maurer's cleft membrane proteins: structural properties and expression profiling. *Genome Res,* 14**,** 1052-9.

SAMPATH, S., BRAZIER, A. J., AVRIL, M., BERNABEU, M., VIGDOROVICH, V., MASCARENHAS, A., GOMES, E., SATHER, D. N., ESMON, C. T. & SMITH, J. D. 2015. *Plasmodium falciparum* adhesion domains linked to severe malaria differ in blockade of endothelial protein C receptor. *Cell Microbiol,* 17**,** 1868-82.

SANYAL, S., EGEE, S., BOUYER, G., PERROT, S., SAFEUKUI, I., BISCHOFF, E., BUFFET, P., DEITSCH, K. W., MERCEREAU-PUIJALON, O., DAVID, P. H., TEMPLETON, T. J. & LAVAZEC, C. 2012. *Plasmodium falciparum* STEVOR proteins impact erythrocyte mechanical properties. *Blood,* 119**,** e1-8.

SCHERF, A., HERNANDEZ-RIVAS, R., BUFFET, P., BOTTIUS, E., BENATAR, C., POUVELLE, B., GYSIN, J. & LANZER, M. 1998. Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of *var* genes during intra-erythrocytic development in *Plasmodium falciparum*. *EMBO J,* 17**,** 5418-26.

SCHMITTGEN, T. D. & LIVAK, K. J. 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc,* 3**,** 1101-8.

SCHREIBER, N., KHATTAB, A., PETTER, M., MARKS, F., ADJEI, S., KOBBE, R., MAY, J. & KLINKERT, M. Q. 2008. Expression of *Plasmodium falciparum* 3D7 STEVOR proteins for evaluation of antibody responses following malaria infections in naive infants. *Parasitology,* 135**,** 155-67.

SEMBLAT, J. P., SILVIE, O., FRANETICH, J. F. & MAZIER, D. 2005. Laser capture microdissection of hepatic stages of the human parasite *Plasmodium falciparum* for molecular analysis. *Methods Mol Biol,* 293**,** 301-7.

SENCZUK, A. M., REEDER, J. C., KOSMALA, M. M. & HO, M. 2001. *Plasmodium falciparum* erythrocyte membrane protein 1 functions as a ligand for P-selectin. *Blood,* 98**,** 3132-5.

SEYDEL, K. B., KAMPONDENI, S. D., VALIM, C., POTCHEN, M. J., MILNER, D. A., MUWALO, F. W., BIRBECK, G. L., BRADLEY, W. G., FOX, L. L., GLOVER, S. J., HAMMOND, C. A., HEYDERMAN, R. S., CHILINGULO, C. A., MOLYNEUX, M. E. & TAYLOR, T. E. 2015. Brain swelling and death in children with cerebral malaria. *N Engl J Med,* 372**,** 1126-37.

SILVERSTEIN, R. L. & FEBBRAIO, M. 2009. CD36, a scavenger receptor involved in immunity, metabolism, angiogenesis, and behavior. *Sci Signal,* 2**,** re3.

SMITH, J. D., CHITNIS, C. E., CRAIG, A. G., ROBERTS, D. J., HUDSON-TAYLOR, D. E., PETERSON, D. S., PINCHES, R., NEWBOLD, C. I. & MILLER, L. H. 1995. Switches in expression of *Plasmodium falciparum var* genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell,* 82**,** 101-10.

SMITH, J. D., CRAIG, A. G., KRIEK, N., HUDSON-TAYLOR, D., KYES, S., FAGAN, T., PINCHES, R., BARUCH, D. I., NEWBOLD, C. I. & MILLER, L. H. 2000a. Identification of a *Plasmodium falciparum* intercellular adhesion molecule-1 binding domain: a

parasite adhesion trait implicated in cerebral malaria. *Proc Natl Acad Sci U S A,* 97**,** 1766-71.

SMITH, J. D., SUBRAMANIAN, G., GAMAIN, B., BARUCH, D. I. & MILLER, L. H. 2000b. Classification of adhesive domains in the *Plasmodium falciparum* erythrocyte membrane protein 1 family. *Mol Biochem Parasitol,* 110**,** 293-310.

SMITH, R. C., VEGA-RODRIGUEZ, J. & JACOBS-LORENA, M. 2014. The *Plasmodium* bottleneck: malaria parasite losses in the mosquito vector. *Mem Inst Oswaldo Cruz,* 109**,** 644-61.

SOMNER, E. A., BLACK, J. & PASVOL, G. 2000. Multiple human serum components act as bridging molecules in rosette formation by *Plasmodium falciparum*-infected erythrocytes. *Blood,* 95**,** 674-82.

SPRINGER, A. L., SMITH, L. M., MACKAY, D. Q., NELSON, S. O. & SMITH, J. D. 2004. Functional interdependence of the DBLbeta domain and c2 region for binding of the *Plasmodium falciparum* variant antigen to ICAM-1. *Mol Biochem Parasitol,* 137**,** 55-64.

SPYCHER, C., KLONIS, N., SPIELMANN, T., KUMP, E., STEIGER, S., TILLEY, L. & BECK, H. P. 2003. MAHRP-1, a novel *Plasmodium falciparum* histidine-rich protein, binds ferriprotoporphyrin IX and localizes to the Maurer's clefts. *J Biol Chem,* 278**,** 35373-83.

SPYCHER, C., RUG, M., PACHLATKO, E., HANSSEN, E., FERGUSON, D., COWMAN, A. F., TILLEY, L. & BECK, H. P. 2008. The Maurer's cleft protein MAHRP1 is essential for trafficking of PfEMP1 to the surface of *Plasmodium falciparum*-infected erythrocytes. *Mol Microbiol,* 68**,** 1300-14.

SRIVASTAVA, A., GANGNARD, S., DECHAVANNE, S., AMIRAT, F., LEWIT BENTLEY, A., BENTLEY, G. A. & GAMAIN, B. 2011. Var2CSA minimal CSA binding region is located within the N-terminal region. *PLoS One,* 6**,** e20270.

SRIVASTAVA, A., GANGNARD, S., ROUND, A., DECHAVANNE, S., JUILLERAT, A., RAYNAL, B., FAURE, G., BARON, B., RAMBOARINA, S., SINGH, S. K., BELRHALI, H., ENGLAND, P., LEWIT-BENTLEY, A., SCHERF, A., BENTLEY, G. A. & GAMAIN, B. 2010. Full-length extracellular region of the var2CSA variant of PfEMP1 is required for specific, high-affinity binding to CSA. *Proc Natl Acad Sci U S A,* 107**,** 4884-9.

SRIVASTAVA, K., COCKBURN, I. A., SWAIM, A., THOMPSON, L. E., TRIPATHI, A., FLETCHER, C. A., SHIRK, E. M., SUN, H., KOWALSKA, M. A., FOX-TALBOT, K., SULLIVAN, D., ZAVALA, F. & MORRELL, C. N. 2008. Platelet factor 4 mediates inflammation in experimental cerebral malaria. *Cell Host Microbe,* 4**,** 179-87.

STAALSOE, T., MEGNEKOU, R., FIEVET, N., RICKE, C. H., ZORNIG, H. D., LEKE, R., TAYLOR, D. W., DELORON, P. & HVIID, L. 2001. Acquisition and decay of antibodies to pregnancy-associated variant antigens on the surface of *Plasmodium falciparum*-infected erythrocytes that protect against placental parasitemia. *J Infect Dis,* 184**,** 618-26.

STORM, J. & CRAIG, A. G. 2014. Pathogenesis of cerebral malaria--inflammation and cytoadherence. *Front Cell Infect Microbiol,* 4:100**,** 1-8.

STURM, A., AMINO, R., VAN DE SAND, C., REGEN, T., RETZLAFF, S., RENNENBERG, A., KRUEGER, A., POLLOK, J. M., MENARD, R. & HEUSSLER, V. T. 2006. Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science,* 313**,** 1287-90.

SU, X. Z., HEATWOLE, V. M., WERTHEIMER, S. P., GUINET, F., HERRFELDT, J. A., PETERSON, D. S., RAVETCH, J. A. & WELLEMS, T. E. 1995. The large diverse gene family *var* encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell,* 82**,** 89-100.

SULISTYANINGSIH, E., FITRI, L. E., LOSCHER, T. & BERENS-RIHA, N. 2013. Diversity of the *var* gene family of Indonesian *Plasmodium falciparum* isolates. *Malar J,* 12:80**,** 1-13.

SUTHERLAND, C. J., TANOMSING, N., NOLDER, D., OGUIKE, M., JENNISON, C., PUKRITTAYAKAMEE, S., DOLECEK, C., HIEN, T. T., DO ROSARIO, V. E., AREZ, A. P., PINTO, J., MICHON, P., ESCALANTE, A. A., NOSTEN, F., BURKE, M., LEE, R., BLAZE, M., OTTO, T. D., BARNWELL, J. W., PAIN, A., WILLIAMS, J., WHITE, N. J., DAY, N. P., SNOUNOU, G., LOCKHART, P. J., CHIODINI, P. L., IMWONG, M. & POLLEY, S. D. 2010. Two nonrecombining sympatric forms of the human malaria parasite *Plasmodium ovale* occur globally. *J Infect Dis,* 201**,** 1544-50.

SWAMY, L., AMULIC, B. & DEITSCH, K. W. 2011. *Plasmodium falciparum var* gene silencing is determined by cis DNA elements that form stable and heritable interactions. *Eukaryot Cell,* 10**,** 530-9.

TAMURA, K., STECHER, G., PETERSON, D., FILIPSKI, A. & KUMAR, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution,* 30**,** 2725-2729.

TAVARES, J., FORMAGLIO, P., THIBERGE, S., MORDELET, E., VAN ROOIJEN, N., MEDVINSKY, A., MENARD, R. & AMINO, R. 2013. Role of host cell traversal by the malaria sporozoite during liver infection. *J Exp Med,* 210**,** 905-15.

TAYLOR, T. E., FU, W. J., CARR, R. A., WHITTEN, R. O., MUELLER, J. S., FOSIKO, N. G., LEWALLEN, S., LIOMBA, N. G. & MOLYNEUX, M. E. 2004. Differentiating the pathologies of cerebral malaria by postmortem parasite counts. *Nat Med,* 10**,** 143-5.

TAYLOR, T. E. & MOLYNEUX, M. E. 2015. The pathogenesis of pediatric cerebral malaria: eye exams, autopsies, and neuroimaging. *Ann N Y Acad Sci,* 1342**,** 44-52.

TAYLOR, W. R., HANSON, J., TURNER, G. D., WHITE, N. J. & DONDORP, A. M. 2012. Respiratory manifestations of malaria. *Chest,* 142**,** 492-505.

TEBO, A. E., KREMSNER, P. G., PIPER, K. P. & LUTY, A. J. 2002. Low antibody responses to variant surface antigens of *Plasmodium falciparum* are associated with severe malaria and increased susceptibility to malaria attacks in Gabonese children. *Am J Trop Med Hyg,* 67**,** 597-603.

TEMBO, D. L., NYONI, B., MURIKOLI, R. V., MUKAKA, M., MILNER, D. A., BERRIMAN, M., ROGERSON, S. J., TAYLOR, T. E., MOLYNEUX, M. E., MANDALA, W. L., CRAIG, A. G. & MONTGOMERY, J. 2014. Differential PfEMP1 expression is associated with cerebral malaria pathology. *PLoS Pathog,* 10**,** e1004537.

TIBURCIO, M., NIANG, M., DEPLAINE, G., PERROT, S., BISCHOFF, E., NDOUR, P. A., SILVESTRINI, F., KHATTAB, A., MILON, G., DAVID, P. H., HARDEMAN, M., VERNICK, K. D., SAUERWEIN, R. W., PREISER, P. R., MERCEREAU-PUIJALON, O., BUFFET, P., ALANO, P. & LAVAZEC, C. 2012. A switch in infected erythrocyte deformability at the maturation and blood circulation of *Plasmodium falciparum* transmission stages. *Blood,* 119**,** e172-80.

TONKIN, C. J., CARRET, C. K., DURAISINGH, M. T., VOSS, T. S., RALPH, S. A., HOMMEL, M., DUFFY, M. F., SILVA, L. M., SCHERF, A., IVENS, A., SPEED, T. P., BEESON, J. G. & COWMAN, A. F. 2009. Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in *Plasmodium falciparum*. *PLoS Biol,* 7**,** e84.

TRAGER, W. & JENSEN, J. B. 1976. Human malaria parasites in continuous culture. *Science,* 193**,** 673-5.

TREUTIGER, C. J., HEDDINI, A., FERNANDEZ, V., MULLER, W. A. & WAHLGREN, M. 1997. PECAM-1/CD31, an endothelial receptor for binding *Plasmodium falciparum*-infected erythrocytes. *Nat Med,* 3**,** 1405-8.

TSE, M. T., CHAKRABARTI, K., GRAY, C., CHITNIS, C. E. & CRAIG, A. 2004. Divergent binding sites on intercellular adhesion molecule-1 (ICAM-1) for variant *Plasmodium falciparum* isolates. *Mol Microbiol,* 51**,** 1039-49.

TUIKUE NDAM, N. G., SALANTI, A., BERTIN, G., DAHLBACK, M., FIEVET, N., TURNER, L., GAYE, A., THEANDER, T. & DELORON, P. 2005. High level of *var2csa* transcription by *Plasmodium falciparum* isolated from the placenta. *J Infect Dis,* 192**,** 331-5.

TUIKUE NDAM, N. G., SALANTI, A., LE-HESRAN, J. Y., COTTRELL, G., FIEVET, N., TURNER, L., SOW, S., DANGOU, J. M., THEANDER, T. & DELORON, P. 2006. Dynamics of anti-VAR2CSA immunoglobulin G response in a cohort of senegalese pregnant women. *J Infect Dis,* 193**,** 713-20.

TURNER, G. D., LY, V. C., NGUYEN, T. H., TRAN, T. H., NGUYEN, H. P., BETHELL, D., WYLLIE, S., LOUWRIER, K., FOX, S. B., GATTER, K. C., DAY, N. P., TRAN, T. H., WHITE, N. J. & BERENDT, A. R. 1998. Systemic endothelial activation occurs in both mild and severe malaria. Correlating dermal microvascular endothelial cell phenotype and soluble cell adhesion molecules with disease severity. *Am J Pathol,* 152**,** 1477-87.

TURNER, G. D., MORRISON, H., JONES, M., DAVIS, T. M., LOOAREESUWAN, S., BULEY, I. D., GATTER, K. C., NEWBOLD, C. I., PUKRITAYAKAMEE, S., NAGACHINTA, B. & ET AL. 1994. An immunohistochemical study of the pathology of fatal malaria. Evidence for widespread endothelial activation and a potential role for intercellular adhesion molecule-1 in cerebral sequestration. *Am J Pathol,* 145**,** 1057-69.

TURNER, L., LAVSTSEN, T., BERGER, S. S., WANG, C. W., PETERSEN, J. E., AVRIL, M., BRAZIER, A. J., FREETH, J., JESPERSEN, J. S., NIELSEN, M. A., MAGISTRADO, P., LUSINGU, J., SMITH, J. D., HIGGINS, M. K. & THEANDER, T. G. 2013. Severe malaria is associated with parasite binding to endothelial protein C receptor. *Nature,* 498**,** 502-5.

TURNER, L., WANG, C. W., LAVSTSEN, T., MWAKALINGA, S. B., SAUERWEIN, R. W., HERMSEN, C. C. & THEANDER, T. G. 2011. Antibodies against PfEMP1, RIFIN, MSP3 and GLURP are acquired during controlled *Plasmodium falciparum* malaria infections in naive volunteers. *PLoS One,* 6**,** e29025.

TUTTERROW, Y. L., AVRIL, M., SINGH, K., LONG, C. A., LEKE, R. J., SAMA, G., SALANTI, A., SMITH, J. D., LEKE, R. G. & TAYLOR, D. W. 2012a. High levels of antibodies to multiple domains and strains of VAR2CSA correlate with the absence of placental malaria in Cameroonian women living in an area of high *Plasmodium falciparum* transmission. *Infect Immun,* 80**,** 1479-90.

TUTTERROW, Y. L., SALANTI, A., AVRIL, M., SMITH, J. D., PAGANO, I. S., AKO, S., FOGAKO, J., LEKE, R. G. & TAYLOR, D. W. 2012b. High avidity antibodies to full-length VAR2CSA correlate with absence of placental malaria. *PLoS One,* 7**,** e40049.

UDOMSANGPETCH, R., TAYLOR, B. J., LOOAREESUWAN, S., WHITE, N. J., ELLIOTT, J. F. & HO, M. 1996. Receptor specificity of clinical *Plasmodium falciparum* isolates: nonadherence to cell-bound E-selectin and vascular cell adhesion molecule-1. *Blood,* 88**,** 2754-60.

UKAEGBU, U. E., KISHORE, S. P., KWIATKOWSKI, D. L., PANDARINATH, C., DAHAN-PASTERNAK, N., DZIKOWSKI, R. & DEITSCH, K. W. 2014. Recruitment of PfSET2 by RNA polymerase II to variant antigen encoding loci contributes to antigenic variation in *P. falciparum*. *PLoS Pathog,* 10**,** e1003854.

UKAEGBU, U. E., ZHANG, X., HEINBERG, A. R., WELE, M., CHEN, Q. & DEITSCH, K. W. 2015. A Unique Virulence Gene Occupies a Principal Position in Immune Evasion by the Malaria Parasite *Plasmodium falciparum*. *PLoS Genet,* 11**,** e1005234.

UMBERS, A. J., AITKEN, E. H. & ROGERSON, S. J. 2011. Malaria in pregnancy: small babies, big problem. *Trends Parasitol,* 27**,** 168-75.

URBAN, B. C., FERGUSON, D. J., PAIN, A., WILLCOX, N., PLEBANSKI, M., AUSTYN, J. M. & ROBERTS, D. J. 1999. *Plasmodium falciparum*-infected erythrocytes modulate the maturation of dendritic cells. *Nature,* 400**,** 73-7.

VIEBIG, N. K., GAMAIN, B., SCHEIDIG, C., LEPOLARD, C., PRZYBORSKI, J., LANZER, M., GYSIN, J. & SCHERF, A. 2005. A single member of the *Plasmodium falciparum var* multigene family determines cytoadhesion to the placental receptor chondroitin sulphate A. *EMBO Rep,* 6**,** 775-81.

VIEBIG, N. K., LEVIN, E., DECHAVANNE, S., ROGERSON, S. J., GYSIN, J., SMITH, J. D., SCHERF, A. & GAMAIN, B. 2007. Disruption of *var2csa* gene impairs placental malaria associated adhesion phenotype. *PLoS One,* 2**,** e910.

VIGAN-WOMAS, I., GUILLOTTE, M., JUILLERAT, A., HESSEL, A., RAYNAL, B., ENGLAND, P., COHEN, J. H., BERTRAND, O., PEYRARD, T., BENTLEY, G. A., LEWIT-BENTLEY, A. & MERCEREAU-PUIJALON, O. 2012. Structural basis for the ABO blood-group dependence of *Plasmodium falciparum* rosetting. *PLoS Pathog,* 8**,** e1002781.

VOGT, A. M., BARRAGAN, A., CHEN, Q., KIRONDE, F., SPILLMANN, D. & WAHLGREN, M. 2003. Heparan sulfate on endothelial cells mediates the binding of *Plasmodium falciparum*-infected erythrocytes via the DBL1alpha domain of PfEMP1. *Blood,* 101**,** 2405-11.

VOLZ, J. C., BARTFAI, R., PETTER, M., LANGER, C., JOSLING, G. A., TSUBOI, T., SCHWACH, F., BAUM, J., RAYNER, J. C., STUNNENBERG, H. G., DUFFY, M. F. & COWMAN, A. F. 2012. PfSET10, a *Plasmodium falciparum* methyltransferase, maintains the active *var* gene in a poised state during parasite division. *Cell Host Microbe,* 11**,** 7-18.

VOSS, T. S., HEALER, J., MARTY, A. J., DUFFY, M. F., THOMPSON, J. K., BEESON, J. G., REEDER, J. C., CRABB, B. S. & COWMAN, A. F. 2006. A *var* gene promoter controls allelic exclusion of virulence genes in *Plasmodium falciparum* malaria. *Nature,* 439**,** 1004-1008.

VOSS, T. S., THOMPSON, J. K., WATERKEYN, J., FELGER, I., WEISS, N., COWMAN, A. F. & BECK, H. P. 2000. Genomic distribution and functional characterisation of two distinct and conserved *Plasmodium falciparum var* gene 5' flanking sequences. *Mol Biochem Parasitol,* 107**,** 103-15.

WALLER, K. L., COOKE, B. M., NUNOMURA, W., MOHANDAS, N. & COPPEL, R. L. 1999. Mapping the binding domains involved in the interaction between the *Plasmodium falciparum* knob-associated histidine-rich protein (KAHRP) and the cytoadherence ligand *P. falciparum* erythrocyte membrane protein 1 (PfEMP1). *J Biol Chem,* 274**,** 23808-13.

WANG, C. W., MWAKALINGA, S. B., SUTHERLAND, C. J., SCHWANK, S., SHARP, S., HERMSEN, C. C., SAUERWEIN, R. W., THEANDER, T. G. & LAVSTSEN, T. 2010. Identification of a major *rif* transcript common to gametocytes and sporozoites of *Plasmodium falciparum*. *Malar J,* 9:147**,** 1-6.

WARIMWE, G. M., KEANE, T. M., FEGAN, G., MUSYOKI, J. N., NEWTON, C. R., PAIN, A., BERRIMAN, M., MARSH, K. & BULL, P. C. 2009. *Plasmodium falciparum var* gene expression is modified by host immunity. *Proc Natl Acad Sci U S A,* 106**,** 21801-6.

WASSMER, S. C., LEPOLARD, C., TRAORE, B., POUVELLE, B., GYSIN, J. & GRAU, G. E. 2004. Platelets reorient *Plasmodium falciparum*-infected erythrocyte cytoadhesion to activated endothelial cells. *J Infect Dis,* 189**,** 180-9.

WASSMER, S. C., TAYLOR, T., MACLENNAN, C. A., KANJALA, M., MUKAKA, M., MOLYNEUX, M. E. & GRAU, G. E. 2008. Platelet-induced clumping of *Plasmodium falciparum*-infected erythrocytes from Malawian patients with cerebral malaria-possible modulation *in vivo* by thrombocytopenia. *J Infect Dis,* 197**,** 72-8.

WATERKEYN, J. G., WICKHAM, M. E., DAVERN, K. M., COOKE, B. M., COPPEL, R. L., REEDER, J. C., CULVENOR, J. G., WALLER, R. F. & COWMAN, A. F. 2000. Targeted

mutagenesis of *Plasmodium falciparum* erythrocyte membrane protein 3 (PfEMP3) disrupts cytoadherence of malaria-infected red blood cells. *EMBO J,* 19**,** 2813-23.
WENG, H., GUO, X., PAPOIN, J., WANG, J., COPPEL, R., MOHANDAS, N. & AN, X. 2014. Interaction of *Plasmodium falciparum* knob-associated histidine-rich protein (KAHRP) with erythrocyte ankyrin R is required for its attachment to the erythrocyte membrane. *Biochim Biophys Acta,* 1838**,** 185-92.
WHITE, N. J., PUKRITTAYAKAMEE, S., HIEN, T. T., FAIZ, M. A., MOKUOLU, O. A. & DONDORP, A. M. 2014. Malaria. *Lancet,* 383**,** 723-35.
WHO. 2014. World Malaria Report 2014. Available: http://www.who.int/malaria/publications/world_malaria_report_2014/en/ [Accessed 14/04/2016].
WHO. 2015. World Malaria Report 2015. Available: http://www.who.int/malaria/publications/world-malaria-report-2015/report/en/ [Accessed 13/04/2016].
WINOGRAD, E., EDA, S. & SHERMAN, I. W. 2004. Chemical modifications of band 3 protein affect the adhesion of *Plasmodium falciparum*-infected erythrocytes to CD36. *Mol Biochem Parasitol,* 136**,** 243-8.
WINOGRAD, E. & SHERMAN, I. W. 2004. Malaria infection induces a conformational change in erythrocyte band 3 protein. *Mol Biochem Parasitol,* 138**,** 83-7.
WINTER, G., KAWAI, S., HAEGGSTROM, M., KANEKO, O., VON EULER, A., KAWAZU, S., PALM, D., FERNANDEZ, V. & WAHLGREN, M. 2005. SURFIN is a polymorphic antigen expressed on *Plasmodium falciparum* merozoites and infected erythrocytes. *J Exp Med,* 201**,** 1853-63.
WIRTH, D. F. 2002. Biological revelations. *Nature,* 419**,** 495-6.
WYKES, M. N., LIU, X. Q., BEATTIE, L., STANISIC, D. I., STACEY, K. J., SMYTH, M. J., THOMAS, R. & GOOD, M. F. 2007. *Plasmodium* strain determines dendritic cell function essential for survival from malaria. *PLoS Pathog,* 3**,** e96.
YANG, Y., JUN, C. D., LIU, J. H., ZHANG, R., JOACHIMIAK, A., SPRINGER, T. A. & WANG, J. H. 2004. Structural basis for dimerization of ICAM-1 on the cell surface. *Mol Cell,* 14**,** 269-76.
YE, R., ZHANG, D., CHEN, B., ZHU, Y., ZHANG, Y., WANG, S. & PAN, W. 2015. Transcription of the *var* genes from a freshly-obtained field isolate of *Plasmodium falciparum* shows more variable switching patterns than long laboratory-adapted isolates. *Malar J,* 14:66**,** 1-10.
YEO, T. W., LAMPAH, D. A., ROOSLAMIATI, I., GITAWATI, R., TJITRA, E., KENANGALEM, E., PRICE, R. N., DUFFULL, S. B. & ANSTEY, N. M. 2013. A randomized pilot study of L-arginine infusion in severe falciparum malaria: preliminary safety, efficacy and pharmacokinetics. *PLoS One,* 8**,** e69587.
YIPP, B. G., ANAND, S., SCHOLLAARDT, T., PATEL, K. D., LOOAREESUWAN, S. & HO, M. 2000. Synergism of multiple adhesion molecules in mediating cytoadherence of *Plasmodium falciparum*-infected erythrocytes to microvascular endothelial cells under flow. *Blood,* 96**,** 2292-8.
ZHANG, Q., HUANG, Y., ZHANG, Y., FANG, X., CLAES, A., DUCHATEAU, M., NAMANE, A., LOPEZ-RUBIO, J. J., PAN, W. & SCHERF, A. 2011a. A critical role of perinuclear filamentous actin in spatial repositioning and mutually exclusive expression of virulence genes in malaria parasites. *Cell Host Microbe,* 10**,** 451-63.
ZHANG, Q., SIEGEL, T. N., MARTINS, R. M., WANG, F., CAO, J., GAO, Q., CHENG, X., JIANG, L., HON, C. C., SCHEIDIG-BENATAR, C., SAKAMOTO, H., TURNER, L., JENSEN, A. T., CLAES, A., GUIZETTI, J., MALMQUIST, N. A. & SCHERF, A. 2014. Exonuclease-mediated degradation of nascent RNA silences genes linked to severe malaria. *Nature,* 513**,** 431-5.

ZHANG, Q., ZHANG, Y., HUANG, Y., XUE, X., YAN, H., SUN, X., WANG, J., MCCUTCHAN, T. F. & PAN, W. 2011b. From *in vivo* to *in vitro*: dynamic analysis of *Plasmodium falciparum var* gene expression patterns of patient isolates during adaptation to culture. *PLoS One,* 6**,** e20591.

Appendix A

PCR primer sequences

DBLα tag, UPS and Exon 2 previously published primers

Target	T_A (°C)	Primer name	Sequence 5'-3'
DBLα tag †	47	DBLαAF'	GCACG(A/C)AGTTT(C*/T)GC
DBLα tag †	47	DBLαBR	GCCCATTC(G/C)TCGAACCA
UPS A ‡	52	upsA-5'	ATTAYATTTGTTGTAGGTGA
UPS B ‡	52	17DBLa-5'	ATGTAATTGTTGTTTTTTTTTTTGTTAGAATATTTAAA
UPS C ‡	52	5B1-5'	CACATATARTACGACTAAGAAACA
Exon 2 ᵻ	48	Ex2-reg	TCTTCATAYTCRCTTTC

† (Bull *et al.*, 2005), ‡ (Mugasa *et al.*, 2012), ᵻ (Lavstsen *et al.*, 2012)

ICAM-1 binding isolate RT-qPCR primers

Target	T_A (°C)	Primer name	Sequence 5'-3'
DBLα tag BC12a	60	E-195F	AAGCGGAAAAACACTACGAAGAT
DBLα tag BC12a	60	E-196R	TAGCCATAGATGGACTTTCACTA
DBLα tag BC12b	60	E-197F	AAGCGCAACACAAAAGCATTATGCA
DBLα tag BC12b	60	E-198R	TTTGTTATATGTTATATTTCCTCCATTACA
DBLα tag BC12d	60	E-201F	GACGAGAGGGAGGACGAA
DBLα tag BC12d	60	E-202R	ATTATCACCACCACATGTTTTTCGA
DBLα tag BC12e	60	E-203F	AGATCTCGCTACAAAAAAGACGGT
DBLα tag BC12e	60	E-204R	TAGCAGTCGTACCGTATGTGAT
DBLα tag BC12f	60	E-205F	ACGGATCCCAAAGCGAAAGA
DBLα tag BC12f	60	E-206R	TTGGGCTCCACTTTGTCCAT
DBLα tag J1a	60	E-183F	TCGTACTACAAAAATGATAATGACCG
DBLα tag J1a	60	E-184R	ATGCCCACCCTTAATGGAAGA
DBLα tag J1b	60	E-185F	AGACGAATGGGAAGGTACCAGA
DBLα tag J1b	60	E-186R	AGCCGAACTTCCTCCTTCAGA
DBLα tag J1c	60	E-211F	AGGCAGCRAAAGACCACTAC
DBLα tag J1c	60	E-212R	TTCTCCTGAACCACATGTATTTCTA
DBLα tag J1d	60	E-213F	AAGAAGAACAATAAGAATCCGGCAA
DBLα tag J1d	60	E-214R	TCTAAAATATTCAGCATTATCTGTTACA
DBLα tag J1e	60	E-215F	AGACGAATGTGAAGACGAATGT
DBLα tag J1e	60	E-216R	TGCACACGCTCTTTGTCGAA
DBLα tag J1f	60	E-217F	AGAGGGAAGAAAGGCGCAA
DBLα tag J1f	60	E-218R	AGAGTCATACCATCATCTGCGAT
DBLα tag PCM7a	60	αE-3F	TGGATTGATGAACGGCGCACA

DBLα tag PCM7a	60	αE-4R	TACCACTCTTAACGTCACACG
DBLα tag PCM7b	60	αE-1F	AGATCGCTACCAAGATACTGA
DBLα tag PCM7b	60	αE-2R	CATTCCAAAGAGTCATACCAT
DBLα tag PCM7c	60	E-269F	AAGATGAACAAATTAGGGAATACTGGT
DBLα tag PCM7c	60	E-270R	TGGCATTTATCCAGACTCCAAGT
DBLα tag PCM7d	60	E-271F	ACAATGATGATACTGACAAAAACTATTACA
DBLα tag PCM7d	60	E-272R	AGGCTCTCCATTGACACATCTA
DBLα tag PCM7e	60	E-273F	TACAAACTCGCTACGGAAGTGAT
DBLα tag PCM7e	60	E-274R	AGTTCCATCGATACACTGGCA
DBLα tag PO69a	60	E-189F	GAAGAATGTGGCGCTAAGAGAT
DBLα tag PO69a	60	E-190R	TCATTTTGCTCCGCCTTACATGTAAT
DBLα tag PO69b	60	E-191F	ACAAGCAGCAGGAATGGGAAGA
DBLα tag PO69b	60	E-192R	TCGCCTGTAGGTTTTTCTCCGT
DBLα tag PO69c	60	E-193F	TCGCTACAATGGTGATAAAGATCCAA
DBLα tag PO69c	60	E-194R	TCCTTTGGCGCTTCGCATGTTA
DBLα tag PO69d	60	E-221F	AATATATGAAAAATTGAATGGCGCAGAA
DBLα tag PO69d	60	E-222R	TCGTTCTCCTCTATTGCACGT
DBLα tag PO69e	60	E-223F	TGACAAATTGAAGGATAAGGAGGAA
DBLα tag PO69e	60	E-224R	ATGGATTTAGTGCATCACCACAT
DBLα tag 8206a	60	E-207F	TGGCGCTAAGAGATCGCTA
DBLα tag 8206a	60	E-208R	TTGTTAGAACATCTCCAGTAGCA
DBLα tag 8206b	60	E-209F	AAGCAGCAGGAATGGGAAGA
DBLα tag 8206b	60	E-210R	CTGTAGGTTTTTCTCCGTTAGA

Genotyping primers published in Recommended Genotyping Procedures (Medicines for Malaria Venture, Amsterdam May 2007)

Target	T_A (°C)	Primer name	Sequence 5'-3'
MSP primary (p)	54	M1-OF	CTAGAAGCTTTAGAAGATGCAGTATTG
MSP p	54	M1-OR	CTTAAATAGTATTCTAATTCAAGTGGATCA
MSP p	54	M2-OF	ATGAAGGTAATTAAAACATTGTCTATTATA
MSP p	54	M2-OR	CTTTGTTACCATCGGTACATTCTT
MSP secondary (s)	50	S1TAILFW	GCTTATAATATGAGTATAAGGAGAA
MSP s	50	N5-3D7-RVVIC	CTGAAGAGGTACTGGTAGA
GLURP p	54	G-F3	ACATGCAAGTGTTGATCCTGAAG
GLURP p + s	54	G-F4	TGTAGGTACCACGGGTTCTTGTGG
GLURP s	54	G-NF	TGTTCACACTGAACAATTAGATTTAGATCA

Primers designed against *var* database hits to ICAM-1 binding isolates

Target	T_A (°C)	Primer name	Sequence 5'-3'
XX0156-C.g40 (BC12a)	52	E-227F	TAGTGAAAGTCCATCTATGGCTA
XX0156-C.g40 (BC12a)	52	E-228R	TCCATTCATCATACAGTTTCTGACT
XX0156-C.g40 (BC12a)	52	E-229F	TGGTGATGGACAAACAGAAATTGAA
XX0156-C.g40 (BC12a)	52	E-230R	ACGCCTTTCTACCACCAGCA
XX0156-C.g40 (BC12a)	52	E-231F	ACCTATTATGAGATCCAATCCATGT
XX0156-C.g40 (BC12a)	52	E-232R	TTTCTCATATTTATCGCAGGCGTTT
XX0156-C.g40 (BC12a)	52	E-233F	TGCAATCACAGGAGTATGAGACA
XX0156-C.g40 (BC12a)	52	E-234R	TTCTTTAGGTTTTGGGAATATTGTATCT
XX0156-C.g40 (BC12a)	52	E-235F	AGGTCCCTCCAGAATTTTTGC
XX0156-C.g40 (BC12a)	52	E-236R	TGGGTGGCACACAAATACTAC
XX0156-C.g40 (BC12a)	52	E-237F	AGTCTCAATGCCGCCTGCT
XX0156-C.g40 (BC12a)	52	E-238R	TTGACCCCATTCTTCAAGGTAAC
XX0156-C.g40 (BC12a)	52	E-239F	TGTCACACTTAAAGATGAAAGTAGT
XX0156-C.g40 (BC12a)	52	E-240R	ATGGGGATTCTTCACGATTTTTCATA
XX0156-C.g40 (BC12a)	52	E-241F	ATGAGGAAAAAGTCAGTGGGAAA
XX0156-C.g40 (BC12a)	52	E-242R	AGTGAACGCAGCAAAACCGAT
XX0156-C.g40 (BC12a)	52	E-242F	ATCGGTTTTGCTGCGTTCACT
XX0004-C.g20 (J1a)	52	E-257F	TCTTCCATTAAGGGTGGGCAT
XX0004-C.g20 (J1a)	52	E-258R	ACATTTACTAATTTCTTTCCTCCATTCT
XX0004-C.g20 (J1a)	52	E-259F	TGATGAAGATAATGACTGTGAAACGA
XX0004-C.g20 (J1a)	52	E-260R	ACATATGTTCTCTTCTTTGTGGCA
XX0004-C.g20 (J1a)	52	E-261F	TGAAAATCATTCCAATCGTAATCCTAA
XX0004-C.g20 (J1a)	52	E-262R	ACTATTTGGAGGGAGTAATTCTTGTA
XX0004-C.g20 (J1a)	52	E-263F	ATGGGACAAAATGCAACTGAAATACT
XX0004-C.g20 (J1a)	52	E-264R	ATTATCGCCACTAAATGTAACGGTTT
XX0137-C.g35 (J1a)	52	E-309F	AAAGGTGACCAAACAGATGAAAAGAA
XX0137-C.g35 (J1a)	52	E-310R	AGTGATAAATATCTATCTCCATCTGTTGT
XX0137-C.g42 (J1b)	52	E-243F	ATGCCAACTTAATCACTCTCTTCAT
XX0137-C.g42 (J1b)	52	E-243R	ATGAAGAGAGTGATTAAGTTGGCAT
XX0137-C.g42 (J1b)	52	E-244R	TGTTTCAGCACAGAACGCGTT
XX0137-C.g42 (J1b)	52	E-245F	AGTGGTACTAATGATAAAGAAAAAGGAA
XX0137-C.g42 (J1b)	52	E-246R	TGTGGGTCTGGGCACTTGT
XX0137-C.g42 (J1b)	52	E-247F	AAGAAACAAAACGCATTAAGGACAT
XX0137-C.g42 (J1b)	52	E-248R	TGTGGGAGGAGGTTTGGAA
XX0137-C.g42 (J1b)	52	E-249F	ATATAGTAGTACGTGGTGTTGCT
XX0137-C.g42 (J1b)	52	E-250R	AGTTTTGCAGTCCTTTTCTTTACCAT
XX0137-C.g42 (J1b)	52	E-251F	AGGTAGTGAGATAACTTTTGATGATA
XX0137-C.g42 (J1b)	52	E-252R	TTCATCACTATTATCTGCTGGTTCA

XX0137-C.g42 (J1b)	52	E-253F	TGAACCAGCAGATAATAGTGATGAA
XX0382-C.g38 (J1d)	52	E-254R	TGTAACCTCATAATAAGAACATTGCCA
XX0382-C.g38 (J1d)	52	E-255F	AGTAATAAAGAGCGTGAGAATAATCCT
XX0382-C.g38 (J1d)	52	E-256R	ACCTCACTTGACGGCCATCT
XX0352-C.g23 (J1d)	52	E-275F	ACAGAAAAGGGGGTAAAGCAGAT
XX0352-C.g23 (J1d)	52	E-276R	TCAATCGCATGCTTGTCTTTATCTTT
XX0352-C.g23 (J1d)	52	E-277F	AAAATCAATGGAAACAAATGGACGAA
XX0352-C.g23 (J1d)	52	E-278R	TGTGATAGACCACATAACATTCCT
XX0352-C.g23 (J1d)	52	E-279F	AGGTAGTGATGTGAGTGATGTAGA
XX0352-C.g23 (J1d)	52	E-280R	ACCTGTTTTCAAACGATTCATACCT
XX0352-C.g23 (J1d)	52	E-281F	AAGCATACATGAAGGATCAGAAGATT
XX0352-C.g23 (J1d)	52	E-282R	CTGTCGCGGCCTTGCATTTA
XX0352-C.g23 (J1d)	52	E-283F	TAAAGAATACAGCATACTTGTTAGTAA
XX0352-C.g23 (J1d)	52	E-284R	TCTGTAAAGTGTTTCTCACCCATA
VAR0141-C.g21 (PCM7a)	55	E-15F	TGTGGTGGAGGAAATCTAAC
VAR0141-C.g21 (PCM7a)	55	E-16R	CTTCTCCGCAGGTTTCGTGACC
VAR0141-C.g21 (PCM7a)	55	E-17F	CATATTGCGAAGCATGTCCGTG
VAR0141-C.g21 (PCM7a)	55	E-18R	GCATCCACATGTGTATCTTC
VAR0141-C.g21 (PCM7a)	55	E-19F	GGCAACACATGTATAACAGG
VAR0141-C.g21 (PCM7a)	55	E-20R	GGTGGCATGTAGATATCTTCGGC
VAR0141-C.g21 (PCM7a)	55	E-21F	GCAAGATTGACACAACGTATTCC
VAR0141-C.g21 (PCM7a)	55	E-22R	CTTACACGTACCACTAACACCCG
VAR0141-C.g21 (PCM7a)	55	E-23F	GTCCTGGTATGCCAGTTGACG
VAR0141-C.g21 (PCM7a)	55	E-24R	GCCATATTTGAGGTTGCATGC
VAR0141-C.g21 (PCM7a)	55	E-25F	GAGGAGGAAGTCACTGACGAC
VAR0141-C.g21 (PCM7a)	55	E-26R	CACCAGTCAACACGCTTGTCAC
VAR0141-C.g21 (PCM7a)	55	E-27F	GATTAAGGCGTTAGAAGCTAGTGG
VAR0141-C.g21 (PCM7a)	55	E-28R	GCAGCGTCAGTGCACGTATCTAG
VAR0141-C.g21 (PCM7a)	55	E-29F	GTGCCATTTCTTGTAGATC
VAR0141-C.g21 (PCM7a)	55	E-30R	CAGTTCTCTGTTCGAATTGCTC
VAR0141-C.g21 (PCM7a)	55	E-31F	GGCCATGTATAGAAAATGG
VAR0141-C.g21 (PCM7a)	55	E-32R	GGAAGCGGACATAACATCGC
XX0488-C.g40 (PCM7d)	52	E-289F	AACCCCTTTCAGAACTTTGTCAT
XX0488-C.g40 (PCM7d)	52	E-289R	ATGACAAAGTTCTGAAAGGGGTT
XX0488-C.g40 (PCM7d)	52	E-290R	TACACGCTTCATTTCCGTCATCT
XX0488-C.g40 (PCM7d)	52	E-291F	TCAAGTGGTGAAGAACATGAAGATA
XX0488-C.g40 (PCM7d)	52	E-292R	TTGACGTCGGGGAGGTAAATA
XX0488-C.g40 (PCM7d)	52	E-293F	ATGAGCGTAATACAGGTGAACCA
XX0488-C.g40 (PCM7d)	52	E-294R	TGCGGTACTATAAACGGTGTTACT
XX0488-C.g40 (PCM7d)	52	E-295F	ACCTGTGGTTAACTTCTTGTTCGA

XX0488-C.g40 (PCM7d)	52	E-296R	ACGCAAACTCCTCGAGATATTCT
XX0488-C.g40 (PCM7d)	52	E-297F	TATGGCAAGGAATGTTATGTGCTTT
XX0488-C.g40 (PCM7d)	52	E-298R	ACTTTTGGGTATTTGCAGTATTTGGAA
XX0488-C.g40 (PCM7d)	52	E-299F	TTATGTTCGATAGCTTTTCTGTTCATGA
XX0488-C.g40 (PCM7d)	52	E-300R	TTGTCCAGATCCTCCAAGATAGT
XX0488-C.g40 (PCM7d)	52	E-300F	ACTATCTTGGAGGATCTGGACAA

Primers designed against PF0311-C.g26 hit to clinical sequence 62B1-1

Target	T_A (°C)	Primer name	Sequence 5'-3'
PF0311-C.g26 (62B1-1)	57	E-51F	CATGTGATGCTCCAAGGGATG
PF0311-C.g26 (62B1-1)	57	E-52R	GATCGGCACTATAGAGGACAG
PF0311-C.g26 (62B1-1)	57	E-53F	GACATGTGACACAGTGGAC
PF0311-C.g26 (62B1-1)	57	E-54R	GCATCTTGTAAATCTTTGGTACCTG
PF0311-C.g26 (62B1-1)	57	E-55F	CCGCGGATTGTATTGATG
PF0311-C.g26 (62B1-1)	57	E-56R	GGCTGCGAGAAGCACATGAACC
PF0311-C.g26 (62B1-1)	57	E-57F	CTAATCGTGATACTAGGCGTTC
PF0311-C.g26 (62B1-1)	57	E-58R	CATGCTGGTTTGCACTTTTCAC
PF0311-C.g26 (62B1-1)	57	E-59F	GCGGATATAGTGATCATACCC
PF0311-C.g26 (62B1-1)	57	E-60R	CTTCTAGGAGGCATACATGG
PF0311-C.g26 (62B1-1)	57	E-61F	GAATGAACGACCAGTGCCAG
PF0311-C.g26 (62B1-1)	57	E-62R	CCATCACACTCTGTCACC
PF0311-C.g26 (62B1-1)	57	E-63F	GATGTGCGCGCGAAACTTATG
PF0311-C.g26 (62B1-1)	57	E-64R	CACTATCCTCACAGGTACAC
PF0311-C.g26 (62B1-1)	57	E-65F	GCTAACACGTATGATAGCGC
PF0311-C.g26 (62B1-1)	57	E-66R	GATTCTCGTATGAGTAGTC
PF0311-C.g26 (62B1-1)	57	E-67F	GACTGGTGGAAAGAACATGGTC
PF0311-C.g26 (62B1-1)	57	E-68R	CGTTTCCACAGCAGCGCATTG
PF0311-C.g26 (62B1-1)	57	E-69F	CCTGATGCGTGCAAAATAGTGG
PF0311-C.g26 (62B1-1)	57	E-70R	GCAAATTCCTCCAAGGTGG
PF0311-C.g26 (62B1-1)	57	E-71F	GGTAGTGATGTGAGTGAAGTAG
PF0311-C.g26 (62B1-1)	57	E-72R	GAGGTCTCGCTGGTTGAGGTG
PF0311-C.g26 (62B1-1)	55	E419-79F	AAGTGTAATGTTGCTGATGGA
PF0311-C.g26 (62B1-1)	55	E419-80R	ATGAACCCCGATTATAACAAC
PF0311-C.g26 (62B1-1)	55	E419-81F	TGCGCCGTAACAAATAGC
PF0311-C.g26 (62B1-1)	55	E419-82R	ACAGATGTAATAGAAGAGTGC
PF0311-C.g26 (62B1-1)	55	E419-83R	ACCTACTCCCGTGTCACC
PF0311-C.g26 (62B1-1)	55	E419-84R	CTTACCGGCCTCAACAGCAAA
PF0311-C.g26 (62B1-1)	52	E-91F	TTAGATTACGTCCCTCAATATTTA
PF0311-C.g26 (62B1-1)	52	E-93F	ATTACCTTTACTAACCCTGATGAA

PF0311-C.g26 (62B1-1)	52	E-93R	TTCATCAGGGTTAGTAAAGGTAAT
PF0311-C.g26 (62B1-1)	52	E-94F	AAGCCTGATGGTAATTGTGGA
PF0311-C.g26 (62B1-1)	52	E-94R	TCCACAATTACCATCAGGCTT
PF0311-C.g26 (62B1-1)	52	E-95R	ATCTTTCCACGTTATAGTTTCCGT
PF0311-C.g26 (62B1-1)	52	E-97F	AAACGGAAGGCATACGAGGAA
PF0311-C.g26 (62B1-1)	52	E-97R	TTCCTCGTATGCCTTCCGTTT
PF0311-C.g26 (62B1-1)	52	E-98F	TTGCCTCCAAGAAGACAACAT
PF0311-C.g26 (62B1-1)	52	E-99F	TGTCGTGCTATCCGTTACAGT
PF0311-C.g26 (62B1-1)	52	E-100F	TAATGATACATCGGTCACCACA
PF0311-C.g26 (62B1-1)	52	E-101F	TACGAACATGCACGAGTCGAT
PF0311-C.g26 (62B1-1)	52	E-101R	ATCGACTCGTGCATGTTCGTA
PF0311-C.g26 (62B1-1)	52	E-105F	AACACCTACCCCAATGTCAAA
PF0311-C.g26 (62B1-1)	52	E-105R	TTTGACATTGGGGTAGGTGTT
PF0311-C.g26 (62B1-1)	52	E-110F	AATGGAGAGGAAAGTCAGAAAACA
PF0311-C.g26 (62B1-1)	52	E-110R	TGTTTTCTGACTTTCCTCTCCATT
PF0311-C.g26 (62B1-1)	52	E-115F	AGACCTCAATGGACACAACAA
PF0311-C.g26 (62B1-1)	52	E-115R	TTGTTGTGTCCATTGAGGTCT
PF0311-C.g26 (62B1-1)	52	E-120F	ATCTGGTCATGTGGAAAGAATATA
PF0311-C.g26 (62B1-1)	52	E-120R	TATATTCTTTCCACATGACCAGAT
PF0311-C.g26 (62B1-1)	52	E-124F	AAGGAACAACCAGAAGGATATGAA
PF0311-C.g26 (62B1-1)	52	E-124R	TTCATATCCTTCTGGTTGTTCCTT
PF0311-C.g26 (62B1-1)	52	E-128F	GCATTGGCCTTGACTTCGATT
PF0311-C.g26 (62B1-1)	52	E-128R	AATCGAAGTCAAGGCCAATGC

Gene-specific forward primers used in combination with Ex-reg for Exon 2 PCR

Target	T_A (°C)	Primer name	Sequence 5’-3’
BC12a	48	241F	ATGAGGAAAAAGTCAGTGGGAAA
J1a	48	310F	ACAACAGATGGAGATAGATATTTATCACT
J1b	48	253F	TGAACCAGCAGATAATAGTGATGAA
J1d	48	285F	TGGTTAGTTGAATGGGGTAAAGAA
PCM7a	48	73F	ATGAAAGAAATATCAGATAAAATAG
PCM7d	48	297F	TATGGCAAGGAATGTTATGTGCTTT

DBLβ expression construct primers (contain BamHI (F) and XhoI (R) digestion sites)

Target	T_A (°C)	Primer name	Sequence 5'-3'
BC12a_DBLβ	52	E-301F	ggtggaggatccCCATGTGCTAAACCCAGTGGT
BC12a_DBLβ	52	E-302R	tccaccctcgagttaACAATTACATGGTGTATCGTGATCAT
J1a_DBLβ	52	E-303F	ggtggaggatccGCTTGTAGTGGAGACCCCA
J1a_DBLβ	52	E-304R	tccaccctcgagttaACACTTACATACATCACCATACCCA
J1d_DBLβ	52	E-305F	ggtggaggatccCCGTGTGCTACTCCTAGTG
J1d_DBLβ	52	E-306R	tccaccctcgagttaACACTCACAAGCATTAGCATACT
PCM7a_DBLβ *	50	E-179F	ggtggagctagcCCATGTACAGGCAAAGATAACG
PCM7a_DBLβ *	50	E-182R	tccaccggtaccttaACATTTACACGCTTTATCGTACC
PCM7d_DBLβ	52	E-307F	ggtggaggatccCCGTGTGGCAAAACAGATGGT
PCM7d_DBLβ	52	E-308R	tccaccctcgagttaACAAGCACAAGCTTTTTTATGGTCAT

***contain digestion sites NheI (E-179F) and KpnI (E-182R)**

IT4 RT-qPCR primers (Viebig *et al.*, 2007)

Target	T_A (°C)	Primer name	Sequence 5'-3'
IT4var01	60	IT4var01F	GATCCGCCAGCAAAAGAAG
IT4var01	60	IT4var01R	CCCCCTTTATATTTTTGTCTGC
IT4var13	60	IT4var13F	GTAAACATCAGGCGTGTAAGG
IT4var13	60	IT4var13R	TGTTCCTCTCCGCTGAAGA
IT4var14	60	IT4var14F	CAAGATGGAAGCGGTAAAG
IT4var14	60	IT4var14R	CATGCATTATCCCAAAGAT
IT4var16	60	IT4var16F	ATGGTAGACAAGCTGTTCGTTT
IT4var16	60	IT4var16R	AGCACAGGCTCCTACTGAATT
IT4var27	60	IT4var27F	CAATAACGACAACCCTGGCA
IT4var27	60	IT4var27R	TGGTGTCTTCGTCGGTTTTT
IT4var31	60	IT4var31F	ACTGGTCGTAAAGGTGCACA
IT4var31	60	IT4var31R	CTCCCTTCAAATCACTTCCC
IT4var41	60	IT4var41F	AACATATGTTTGATAGAATTG
IT4var41	60	IT4var41R	TGGCATCTGTAGGCACGAA