University of Oxford

Structure of the *Pf*EMP1:EPCR complex and inhibition of the interaction

Clinton Kwun Yau Lau

A thesis presented for the degree of
Doctor of Philosophy

Department of Biochemistry
Lady Margaret Hall
Trinity Term 2015

Abstract

Structure of the *Pf*EMP1:EPCR complex and inhibition of the interaction

Clinton Kwun Yau Lau

Doctor of Philosophy

Lady Margaret Hall, University of Oxford

Trinity Term 2015

The *Plasmodium falciparum* erythrocyte membrane protein 1 (*Pf*EMP1) family is central for *Plasmodium falciparum* infection. This family mediates interactions between the infected erythrocyte and the host endothelial surface, making it an obvious vaccine target. However this family is very diverse, with sixty different variants in the genome of each parasite. Despite the diversity within the *Pf*EMP1 family, only around ten ligands have been identified. This suggests that there must be conservation of ligand binding amongst the sequence variety, which could be targeted in vaccine design strategies.

This thesis explores this possibility, using the interaction between the CIDRα1 domain within the *Pf*EMP1 and endothelial protein C receptor (EPCR) as an example, as this interaction has been associated with severe childhood malaria. This thesis details the solution of the first *Pf*EMP1:ligand crystal structure. The two CIDR:EPCR complex structures show the binding surface of the CIDR domains to consist of a central hydrophobic residue which protrudes into a hydrophobic groove in EPCR, surrounded by a ring of hydrophilic residues. This surface mimics features of the natural EPCR ligand, protein C, and can block this ligand interaction. Combination of this structure with sequence data from 885 CIDRα1 domains show that the EPCR-binding surfaces of CIDRα1 domains are conserved in shape and bonding potential, despite dramatic sequence diversity. Having identified the conserved features of this interaction, three synthetic immunogens were designed to elicit cross-inhibitory antibodies against this interaction site. The work within this thesis shows for the first time how the malarial *Pf*EMP1 proteins manage to simultaneously diverge whilst maintaining the capacity to bind to their ligand, and explores vaccine design strategies to raise a cross-inhibitory response.

Acknowlegements

I am very thankful for the many collaborators, colleagues, friends and family who have supported me through my DPhil.

I would like to acknowledge my supervisor, Prof. Matt Higgins. Since my first day as a Part II project student and through my DPhil, Matt has kept me focussed on answering the big questions in my project. From crystal freezing to conferences, planning to publishing, even in the rare medical emergency, Matt has always been there to help, offering sage advice and expert technical help. Matt, I am eternally grateful for the time you have spent with me on this project.

Much of the work detailed in this thesis was completed in collaboration with the Lavstsen group at the Centre of Medical Parasitology, University of Copenhagen. Our meeting to unite the structure of the CIDR:EPCR complex with their formidable database of sequences was one of the highlights of my DPhil. Thank you for all the advice and the vast amount of protein, baculovirus and sera you have given me. I thank Dan Alanine, Joe Illingworth and Simon Draper at the Jenner Institute, who spent a lot of time maintaining mice for hybridoma production, not included in this thesis. I am grateful to my examiners, Prof. Alister Craig and Prof. Radu Aricescu, for reading this sizeable thesis and for their discussion in my *viva voce* examination. I also acknowledge my funding bodies, the Medical Research Council and Lady Margaret Hall.

I have had a lot of help from people within the Department of Biochemistry, here in Oxford. I have really appreciated having a thesis committee, consisting of Elspeth Garman and Sylvia McLain, who have set aside many hours to discuss my project over the course of three years. I am also very grateful to Ed Lowe, who manages the crystallography facilities, and David Staunton, who manages the biophysics suite. As well as maintaining excellent facilities in the department, they have always been available to explain techniques and help with data collection. I'd also like to thank Peter Judge and Gill McLure for giving me the opportunity to teach undergraduates, which I have enjoyed immensely.

The ever-growing Higgins Lab has provided a great scientific and social environment during this DPhil. I have enjoyed the company of everyone who has passed through the lab, from postdocs to summer students, and I am thankful for the never-ending supply of cake, tea and football that you have provided over the years. Within the lab, I'd like to thank Jon, who has worked tirelessly to ensure the smooth operation of the lab, and Gareth and Jack, two students who have both made contributions in this thesis. I am especially appreciative of my desk-mates Harriet, Frank and Kate, who I now count amongst my closest friends.

My friends inside and outside the department have kept me well occupied in my spare time. There are really too many people to mention and to thank properly here, but I'd like to thank Chris, Paul, Verity, Erin, Richard, Jasmin, Jo, Peter, Katie, Nikki, as well as the people at Oxford University Badminton Club and in the Oxford Shinty Team. Grace, thank you for putting up with me over the last few years.

Finally, thanks to my family, my brother Des and my mum and dad. I cannot express my gratitude to you for your guidance and help through my formative years. Without your never-ending love and support, none of this would have been possible.

Contents

List of Figures

List of Tables

List of abbreviations

AMA-1	Apical membrane antigen 1
ATS	Acidic terminal segment
AUC	Analytical ultracentrifugation
CD36	Cluster of differentiation 36
CelTOS	Cell-traversal protein for ookinetes and sporozoites
CIDR	Cysteine-rich inter-domain region
CSA	Chondroitin sulfate A
CSP	Circumsporozoite protein
DBL	Duffy-binding like
DC	Domain cassette
EBL	Erythrocyte-binding like
EPCR	Endothelial protein C receptor
ICAM-1	Intercellular adhesion molecule-1
MALLS	Multi-angle laser light scattering
MSP	Merozoite surface protein
NTS	N-terminal segment
Pf	*Plasmodium falciparum*
*Pf*EMP1	*Plasmodium falciparum* erythrocyte membrane protein 1
RBP	Reticulocyte binding protein
Rh	Reticulocyte binding homologue
SAXS	Small angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SPR	Surface plasmon resonance
TRAP	Thrombospondin-related anonymous protein

Chapter 1

Introduction

1.1 Malaria, the disease

The worldwide treatment and prevention of malaria remains one of the biggest challenges for mankind. Caused by parasites of the *Plasmodium* genus, malaria is spread by mosquitoes that can infect humans with a single bite. The World Health Organisation (WHO) has an extensive treatment and prevention program, with the ultimate goal of the eradication of malaria. However, despite a steady decline in the number of reported cases and deaths since the turn of the millenium [1], malaria is still one of the top five causes of death in children under the age of five [2].

1.1.1 Global burden

Malaria is a worldwide disease, with cases of indigenously-caused malaria reported in all continents with the exception of Antarctica. The areas with the highest number of malaria cases are sub-Saharan Africa and South-East Asia (Figure 1.1), with 82% of reported cases occuring in sub-Saharan Africa alone in 2013 [1]. Some of the world's poorest countries are hardest hit by malaria, in part because the economic cost of malaria is high [3]. Fortunately international initiatives such as the Bill and Melinda Gates Foundation and Medicines for Malaria Venture, and more recently pharmaceutical companies such as GlaxoSmithKline,

have dedicated a substantial amount of resources to treat and prevent malaria in these countries, and to research better intervention strategies [4].

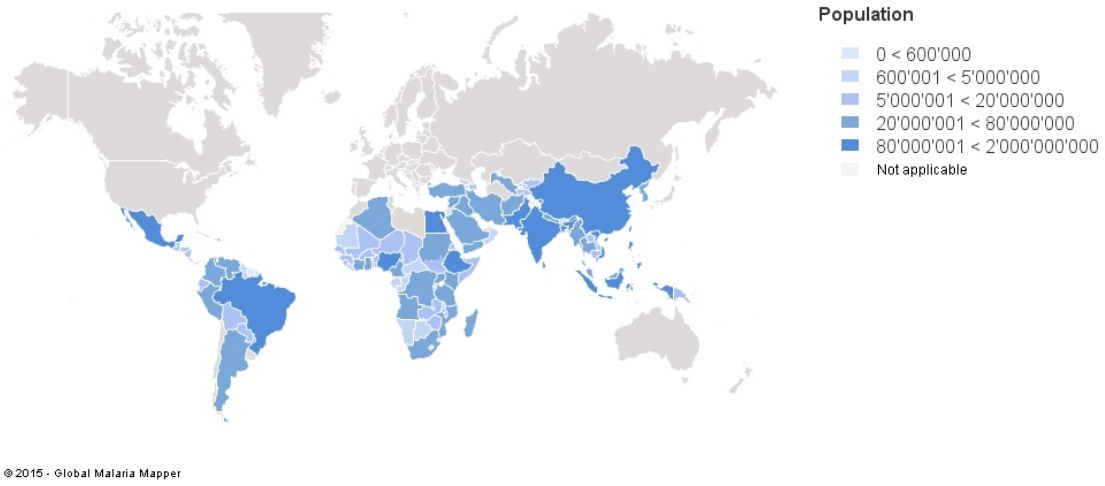


Figure 1.1: World map showing the number of global incidences of malaria by country per year. Map produced using the Global Malaria Mapper (www.world-malariareport.org).

The number of cases and deaths caused by malaria is difficult to estimate precisely, due to the inherent uncertainties in extrapolating the actual number of cases from the number of reported cases, especially in regions where medical coverage is sparse or poorly organised. The World Health Organisation estimated that there were 198 million cases in 2013 (95% confidence limits 124-283 million), with 600,000 deaths (95% confidence limits 367,000-755,000) [1]. However there are suggestions that this is a gross underestimation, with one study concluding malaria deaths in 2010 were underestimated by a factor of two [5] and a second that combined narratives of deaths in India with national measures of malaria prevalence, suggesting that WHO estimates of malaria mortality in India may be misrepresented by up to a factor of ten [6].

Malaria is most dangerous to naïve populations, with immunity acquired after one or two infections [7]. In malaria endemic countries the largest population group at risk is young children. Indeed, of the 600,000 deaths in 2013, 78% were children under 5 years old [1]. The second largest group at risk is first-time mothers, as the placenta offers a naïve environment where malaria can take hold and threaten the development unborn child [8]. Travellers are also a relatively small but significant group at risk [9].

1.1.2 *Plasmodium* species, the causative agent of malaria

The link between blood-borne parasites and malaria was first made by Laveran, a medical officer in the French Army treating patients with malarial fevers [10, 11]. These parasites, of the *Plasmodium* genus, are part of the phylum, Apicomplexa, related to other intracellular parasitic organisms such as *Babesia*, *Cryptosporidia* and *Toxoplasma*, which can all cause disease in humans. To date, six species of *Plasmodium* have been shown to cause malaria in humans. *P. falciparum*, *P. vivax* and *P. malariae* were shown to cause malaria by 1916 [12, 13]; *P. ovale* was shown to cause malaria in 1930 [14] and have since been split into two species, *P. ovale curtisi* and *P. ovale wallikeri*; and *P. knowlesi*, a simian-infective species has recently been found to cause malaria in humans [15, 16, 17]. These parasites are transmitted from human host to human host via mosquito vectors of the *Anopheles* genus, facilitating the spread of the *Plasmodium* parasites through malaria-endemic regions.

Of these six species *Plasmodium falciparum* and *Plasmodium vivax* are responsible for the large majority of malaria cases (>90%) [1]. The geographical dispersion of these species is similar, but not identical (Figure 1.2). While *P. vivax* is responsible for a mere 8% of total world malaria cases, *P. vivax* causes 47% of the cases of malaria outside of Africa. Therefore much of the research on malaria has focussed on *P. falciparum* and *P. vivax* as the two main causative agents of malaria.

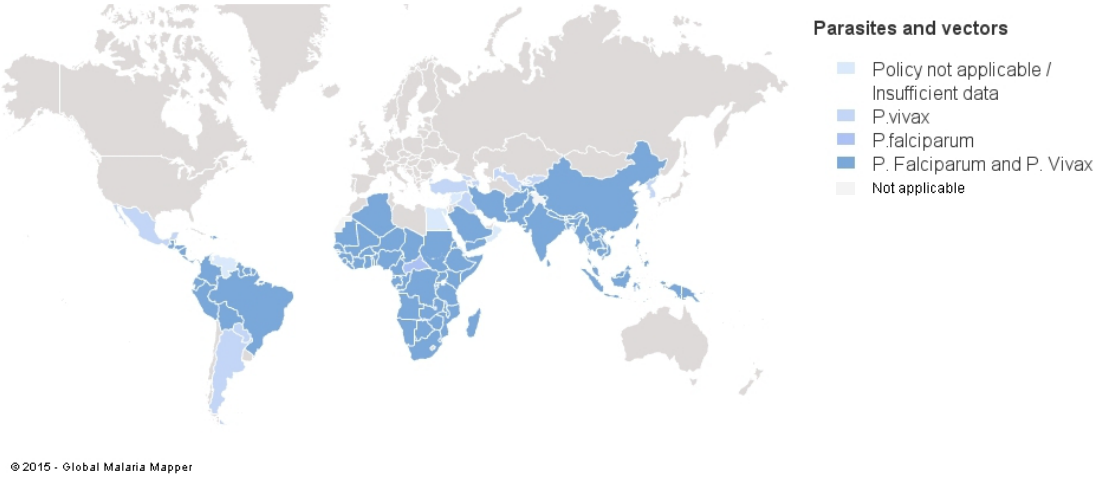


Figure 1.2: World map showing global distribution of *Plasmodium falciparum* and *Plasmodium vivax* by country. Map produced using the Global Malaria Mapper (www.worldmalariareport.org).

1.1.2.1 The life cycle of *Plasmodium* parasites

Plasmodia are obligate intracellular parasites that undergo part of their life cycle in mosquitos, and part in a mammalian host (Figure 1.3). When an anopheline mosquito bites a host, $20 - 120$ *Plasmodium* parasites are injected as sporozoites into the dermis [18, 19] (Figure 1.3, Step 1). These sporozoites make their way into the bloodstream, either directly into the capillaries or via the lymphatic system [20].

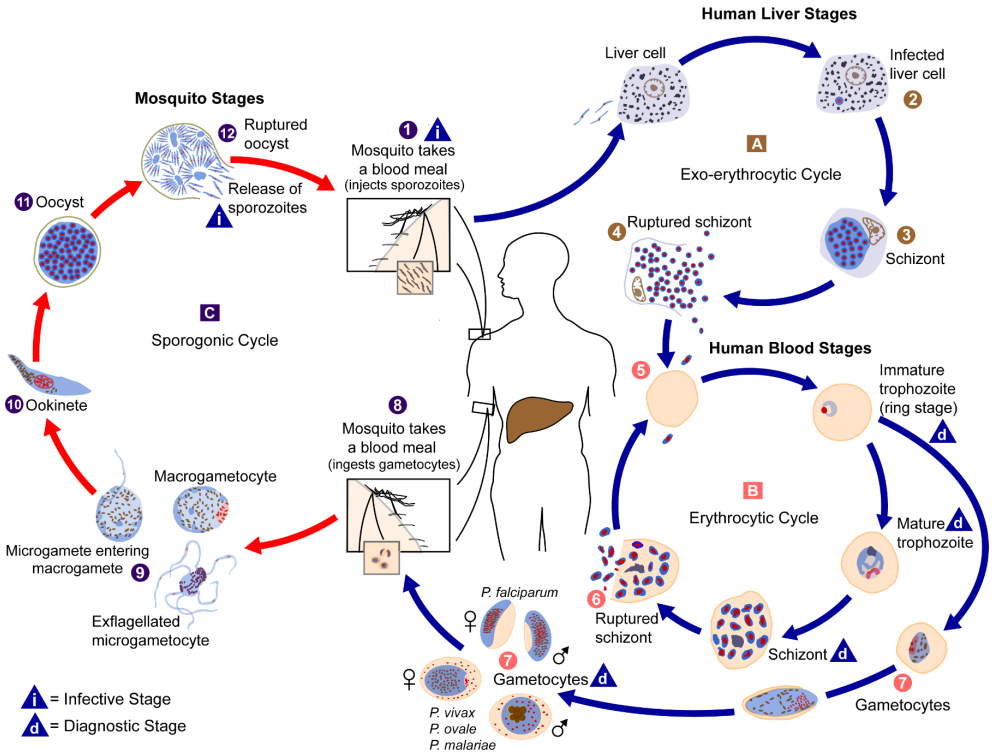


Figure 1.3: The life cycle of a *Plasmodium* parasite. Life cycle steps are numbered for reference in text. Adapted from the CDC Public Health Image Library.

Once in the bloodstream, sporozoites migrate quickly to the liver [21]. Here, the parasite sequesters in the liver sinusoidal lumen. This sequestration is thought to be mediated by the *Plasmodium* circumsporozoite protein (CSP), which coats the sporozoite, interacting with heparan sulfate proteoglycans in the lumen [22]. This initial interaction is followed by invasive and non-invasive movement via Kupffer cells or endothelial cells [23] to reach the hepatocytes. Upon reaching the hepatocytes, the sporozoites may pass through several hepatocyte cells, using proteins such as CelTOS [24], before developing a parasitophorous vacuole in a final hepatocyte invasion step [25]. In the hepatocytes the parasite develops,

before undergoing rounds of replication, developing into a schizont containing around two thousand parasites. These merozoite-form parasites, when mature, bud from hepatocytes in membrane structures known as merosomes. These round vesicles, varied in size, package the merozoites in a non-immunogenic membrane, devoid of phosphatidyl-serine, a classic sign of a dying cell [26, 27]. This is hypothesised to enable the merozoites to safely traverse the space of Disse back into the bloodstream, without detection by immune cells such the Kupffer cells that would result in phagocytosis (Steps 2-4).

The length of time the parasite spends in the hepatic stage of the life cycle can vary dramatically, depending on the *Plasmodium* species and the host species. While the majority of parasites are released from the liver between two and sixteen days after infection [27], subpopulations of some *Plasmodium* species, such as the human-infective *Plasmodium vivax* and *Plasmodium ovale*, can establish latent populations, termed hypnozoites, in the liver [28, 29]. These quiescent cells can reactivate many days later, resuming their life cycle in the host.

Once released by the liver the parasites enter a red blood cell-dependent cycle of invasion, multiplication and egression, multiplying twenty-fold in each cycle (Steps 5-6). This cycle is initiated by the disruption of the merosomes releasing their merozoite contents in the pulmonary microvessels [30]. These merozoites quickly invade red blood cells in the bloodstream by an ordered sequence of events. Initially the merozoite binds to red blood cell through low affinity interactions, such between merozoite surface protein 1 (MSP1) and Band 3 on the red blood cell surface [31, 32, 33]. This initial contact is quickly followed by reorientation of the merozoite, so its apical end is proximal to the red blood cell membrane. Here, interactions with a higher affinity take place, mediated by the EBL and Rh *Plasmodium* invasion protein families [34], preceding an ordered release of proteins and lipids from the rhoptries [35]. This sequence both facilitates the translation of the merozoite into the cell via a moving tight junction, and provides a parasitophorous vacuole for the merozoite to enter [36].

As the merozoite enters the red blood cell it sheds the majority of the surface coat proteins [37, 38], entering completely into the cell sheathed in the parasitophorous vacuole,

sealing the red blood cell surface once it has invaded. Once inside the red blood cell, the parasite modifies the red blood cell, transforming it into an environment suitable for parasite growth. The parasitophorous vacuole is extended into a network for the transport of proteins from the parasite to the erythrocyte and the nutrients back into the parasite [39]. Other modifications are made, specific to the *Plasmodium* species. Indeed, the two major species of *Plasmodia* markedly differ in their effect on the host cell. *Plasmodium falciparum*, which invades erythrocytes, extends the parasitophorous membrane into a network, exporting proteins through Maurer's clefts to modify the cytoskeleton at the surface of the erythrocyte [40]. These proteins rigidify the membrane, forming knob-like structure at the surface, housing proteins such as *Pf*EMP1 [41]. These proteins allow the infected erythrocytes to cytoadhere to the walls of the bloodstream, sequestering red blood cells in the microvasculature of organs such as the brain or the placenta [42], allowing them to avoid splenic clearance. While *Plasmodium vivax* also forms an extensive membrane system like *P. falciparum* [43], it preferentially invades reticulocytes rather than erythrocytes. These reticulocytes, once invaded by *P. vivax*, have been observed to be more deformable and without knobs, rather than becoming rigid and covered in knobs like *P. falciparum*-infected erythrocytes [44]. These differential characteristics allow the infected erythrocyte evade the immune system. Since the red blood cell stage of the disease is accompanied by a massive increase in parasite load, due to its cyclical nature, this immune evasion important to prevent clearance by the host.

In the red blood cell, the parasite matures and multiplies into $16 - 32$ schizonts. These schizonts egress from the red blood cell by disrupting both the parasitophorous vacuole and the plasma membrane, erupting from the blood cell into the bloodstream [45]. The large majority of these parasites egress as merozoites, ready for a new round of the invasion. A proportion of the parasites, within the erythrocytes, develop into male and female gametocytes in each erythrocytic cycle (Step 7). This proportion can vary depending on the conditions in the host [46, 47], depending on conditions that may negatively impact asexual reproduction such as drug treatment [48].

These gametocytes flow in the peripheral circulation, waiting for a mosquito to take a bloodmeal. Once taken up into the mosquito midgut, the gametocytes mature into ga-

metes, then fuse to form a zygote [49] (Step 8-9). Meiosis and genetic recombination soon occur, with the resultant ookinete gliding to the midgut lumen endothelium [50] (Step 10). Here, the ookinete migrates through the epithelial layer of cells. This migration uses some of the same proteins, such as CelTOS (cell-traversal protein for ookinetes and sporozoites), as the sporozoite invasion of the liver endothelium, and so the mechanism of invasion is hypothesised to be similar [24]. However instead of invading another cell type, here ookinetes develop into oocysts (Step 11). Thousands of sporozoite-form parasites can develop inside each oocyst, which escape the oocyst when mature in a protease-dependent manner [51] (Step 12). These sporozoites migrate to the salivary glands, where they can invade, ready for another round of human infection.

1.1.3 Pathogenesis varies by *Plasmodium* species

The symptoms of *Plasmodium* infection in non-immune humans present between 6-14 days after the initial mosquito bite. This matches well with the time it takes from sporozoite injection to the cyclical blood stage growth phase. Indeed, some of the hallmark symptoms of severe malaria correlate with the invasion and depletion of healthy blood cells in the circulatory system [52].

Initial malaria infection presents with a fever (>90% of cases), chills (80%), headache (70%) and diaphoresis, excessive sweating (64%) [53]. These symptoms resemble viral infection and can resolve if the parasite load does not increase, or if treatment is administered quickly. If the parasite load increases sufficiently, more serious symptoms can develop. The onset of severe symptoms is most common in *Plasmodium falciparum* infections, though cases have been reported in other parasite strains, with growing attention to severe malaria caused by *Plasmodium vivax* [54].

Severe symptoms overlap between malaria caused by *P. falciparum*, and malaria caused by other species but are not identical. Typical severe manifestations of malaria include cerebral malaria, pulmonary or renal failure, severe anaemia and hyperlactatemia [53], with severe anaemia and cerebral malaria the most common presentations. These manifestations can lead to hospitalisation due to life-threatening symptoms such as coma, vomiting

and respiratory distress in patients, with death following from hours to days, depending on the state of the patient at admission [55]. *Plasmodia* species differ in terms of how severe symptoms present due to their differing biology. In *Plasmodium falciparum*-malaria, around 10% of cases progress to the severe stage, classified according to WHO guidelines, with children presenting with severe anaemia and cerebral malaria, while adults more often present with pulmonary and renal failure [55]. This correlates with a very high *P. falciparum* load in the system, with infected red blood cells sequestering in the microvasculature of organs [56].

Plasmodium vivax-based severe malaria, on the other hand, has a much lower presentation rate of cerebral malaria, with patients often presenting instead with severe anaemia. This correlates with a much lower level of sequestration by *P. vivax*-infected reticulocytes, with a lower parasite load required to cause severe anaemia than for *P. falciparum* [54]. In addition to this acute phase, malaria caused by *Plasmodium vivax* and *Plasmodium ovale* has a chance of relapse, due to reactivation of quiescent hypnozoites in the liver [54].

The precise mechanisms behind these severe manifestations are still not completely understood, with most evidence coming from correlations between disease manifestation and indirect data, such as insight gained from autopsy and circulating cytokine levels [56, 42, 57]. Severe anaemia is, on first glance, the most straight-forward to explain, with increased parasite loads infecting and destroying more red blood cells, leaving fewer cells to carry oxygen around the body. However, there are gaps in our pathological knowledge. The decrease in haemoglobin concentration in *falciparum* malaria is larger than can be accounted for by the calculated parasite load [58, 59]. Moreover, malaria caused by *Plasmodium vivax* is associated with a lower parasite load than in *falciparum*-caused malaria, yet causes a larger decrease in haemoglobin concentration [60].

Similarly, our understanding of the pathological basis behind other manifestations of organ failure and coma due to cerebral malaria is incomplete. Sequestration of *falciparum* parasites has long been correlated to these manifestations of malaria, as post-mortem studies found infected erythrocytes adhered to the microvasculature of organs such as the brain [61, 62, 63]. This agrees with the observation that *P. vivax*, which does not sequester to

the same extent as *P.falciparum*, does not commonly cause cerebral malaria [54]. However, the difficulty in studying the sequestered parasite population within the human host means that direct evidence for a link between sequestration and severe malaria symptoms is lacking [64]. The best estimate of the sequestered parasite load to date comes from measuring *Pf*HRP2 concentration level and extrapolating to estimate the total parasite biomass, taking into account the amount of *Pf*HRP2 released at different points in the erythrocytic cycle, and comparing this biomass with the number of circulating parasites to calculate the sequestered biomass [65]. Two uses of this assay give differing results: one study in Thailand correlated an increase in sequestered biomass with an increased disease severity [65], while another study in Gambia concluded little change in sequestered biomass between children with uncomplicated and severe malaria [66].

1.2 Towards the global eradication of malaria

Governments and charitable organisations around the world have made large steps towards the eradication of malaria through the application of vector and drug-based prevention methods, faster and more accurate diagnosis, and treatment of malaria cases using drug combination therapies. The World Health Assembly set a target in the year 2000 of reducing the number of malaria cases by 75% by 2015. As of the beginning of 2014, 55 of the 66 countries with sufficient data for analysis by the WHO are on course to meet this target. Though this is encouraging, most of these 55 countries did not have a high number of cases in 2000. The majority of countries where malaria is most prevalent, in sub-Saharan Africa did not submit enough data to be analysed [1]. In these countries it has been observed that the infection prevalence, estimated from national surveys, is in decline, with a recent study estimating a disease prevalence decrease of 50% since 2000 [67].

Current intervention strategies focus on protecting vulnerable groups in malaria-endemic countries and providing fast and effective treatment for cases of malaria, bearing in mind the need to slow the emergence of resistance to current therapies. Interventions centre both around the vector, anopheline mosquitoes, and the parasite, with different therapies tailored to the location and the group at risk. The World Health Organisation publishes

guideline intervention strategies that are adopted by countries in malaria-endemic regions, supported by charitable organisations.

1.2.1 Vector controls

Historically large-scale spraying of mosquito habitats using insecticides such as dichlorodiphenyltrichloroethane (DDT), were used to control and eradicate mosquitoes in malaria endemic areas. Indeed, between the 1940s and 1950s, DDT was used to great effect, especially in USA and Northern Europe where malaria was eradicated. However the emergence of resistance to DDT, in part due to the often indiscriminate application led to its decline in use in the 1960s [68, 69].

Today, WHO-advised vector control strategies consist of insecticide-treated nets and indoor residual spraying. These strategies are considered to have less environmental impact whilst still providing effective protection for the population groups at risk (children and first-time mothers) [1]. Both of these local controls have been proven effective, affecting not only to the protected house, but to any nearby unprotected houses by acting as a mosquito-sink, reducing the overall vector level [70, 71, 72]. Both of these strategies mainly use pyrethroid-based insecticides, which bind to mosquito sodium channels resulting in repetitive axonic excitation [73, 74]. Other non-pyrethroid insecticides are also being promoted, with DDT being reintroduced for some indoor residual spray regimens [1].

1.2.2 Drugs

The latest WHO report recommends intermittent presumptive treatment using sulfadoxine-pyrimethamine for protection of at risk populations in high-risk areas. This drug combination inhibits parasite metabolism enzymes dihydrofolate reductase and dihydropteroate synthase, synergistically impacting folate metabolism necessary for nucleic acid synthesis [75]. This is supplemented in young children (3-59 months) with atovaquone, a selective inhibitor of the parasite electron transport chain. These drugs are taken every few months while the person is at risk, providing a basic level of protection against malaria [76, 77, 1].

Studies have shown a decrease in the number of hospitalised cases caused by malaria, though the effect on mortality is uncertain due to ethical limitations [77]. This prophylactic strategy is designed as a cost-effective way of reducing the incidences of malaria, supported by vector control and good access to treatment.

Travellers from the UK are advised to take mefloquine (trade name Lariam) or an atovaquone-proguanil (trade name Malarone) before and during visits to malaria-endemic countries, which are both considerably more expensive than sulfadoxine-pyrimethamine, but more effective due to sulfadoxine-pyrimethamine resistance in many malaria-endemic regions. Mefloquine interferes with the detoxification of free haem in infected red blood cells, causing it to instead crystallise and kill the blood stage parasite [75]. Atovaquone-proguanil is considered a safer drug combination, with less severe side effects. Proguanil acts in a similar way to pyrimethamine, disrupting thymine synthesis by inhibiting deoxythymidilate synthase [75].

While the prophylactic treatments are useful to treat low-level parasitaemia, treatment of higher level parasitaemia requires the use of fast-acting drugs to effectively curb the infection. Chloroquine, which is related to mefloquine, and sulfadoxine-pyrimethamine have both declined in use due to the rise of resistance [78, 79]. Artemisinin, derived from the Chinese plant qinghao (Latin name *Artemisia annua*) [80], remains the most effective treatment for malaria. Artemisinin is thought to act through the breakage of an endoperoxide bridge [81], but the targets are still not fully understood [75], with the *Plasmodium* Ca^{2+}-ATPase [82] and haemoglobin degradation [83] postulated as targets. Recent research has uncovered a *P. falciparum* gene that confers artemisinin resistance, K13-propeller (also known as kelch) [84].

1.2.3 Current challenges for the malaria eradication programme

Despite the progress made against malaria there are two major challenges to the ultimate goal of malaria elimination. Firstly, the targeting and implementation anti-malarial strategies in a coordinated and sustained fashion can be difficult. This is primarily due to the requirement of a high level of funding and a good infrastructure in the target lo-

cation, which are not always available [85]. Indeed, the WHO estimated 278 million of the 840 million people at risk of malaria live in houses without an insecticide treated net [1]. This challenge is part economical and part social in nature, and is being dealt with by organisations such as the Roll Back Malaria Partnership, working in collaboration with the governments in the affected areas to coordinate action against malaria.

The second major challenge is to stay ahead of emerging parasite resistance to existing malaria interventions through the discovery of novel therapies and prevention methods. Man and *Plasmodia* have been locked in an drug versus resistance arms race since the widespread introduction of drugs such as proguanil and chloroquine and the subsequent resistance in *Plasmodium falciparum* ten to fifteen years later [86]. Today, drug resistance is widespread to many front-line treatments [75]. There are even reports of the failures of combination therapies, such as the commonly used atovaquone-proguanil (Malarone), through multiple mutations in the *Plasmodium* [87, 88, 89]. The emergence of artemesinin resistance in South-East Asia in the past 6 years [65, 90, 91] now means resistance has been observed for every antimalarial drug class.

Drug resistance is also on the rise in the vectors that carry the parasites. Resistance against the commonly used pyrethroid-class insecticide has been reported since the 1990s [92, 93, 94, 95]. Indeed a recent study, conducted in Mozambique, shows that malaria vector *Anopheles funestus* has gained resistance to the pyrethroid, permethrin, since 2002, to the point where 90% of a test population remained alive after an hour-long exposure to the insecticide [95].

New drug classes are desperately sought after to circumvent the resistance that the *Plasmodium* has evolved. Drug identification programmes screen many compounds for antimalarial activity. In these programmes alternatives to artemisinin are being developed, centred around the active endoperoxide bond. Two of these, OZ277 and OZ439, have been shown to effectively clear artemisinin-resistance parasites *ex vivo* [96, 97]. New drug classes are also being explored, though progress is slow. Recently a new class, which targets eEF-2 (eukaryotic elongation factor 2) of *Plasmodium falciparum*, has been synthesised and tested *ex vivo* [98]. This candidate shows good promise to become the next class of antimalarial

drug.

Though these drug discovery programmes are generating new drug classes that could treat malaria for years to come, they do not overcome the problem of drug resistance. These new drug classes could suffer from similar problems of resistance, as parasites could evolve ways to circumvent these new drugs. This had led to a search for critical targets that a parasite cannot evolve away from, which can be targeted by drugs or vaccines.

1.2.4 The search for a vaccine

Vaccines are in some ways preferable to drugs, as a vaccine ideally promises long-lasting protection against the disease in question through the priming of the immune system. In the malarial field, targets for vaccination have been identified and explored from all human stages of the *Plasmodium* life cycle focussing on *Plasmodium falciparum*, since this parasite causes over 90% of malaria deaths worldwide. These targets are proteins that are presented at the surface of the parasite or parasite-infected cell, and so are vulnerable to a primed immune system. The most extensively explored targets are CSP from the sporozoite stage, with recent attention drawn to *Pf*CelTOS and TRAP [99, 100, 101]; MSP1, AMA1, and recently Rh5 from the merezoite stage [102, 103, 104]; *Pf*EMP1 from the erythrocytic stage [7]; and *Pfs*48/45 and *Pfs*230 from the sexual stage [105]. From this library of targets, the current most promising candidates are CSP and Rh5.

Though sporozoite infection is in itself asymptomatic, an effective sporozoite vaccine is an attractive proposition, as it would promise protection against new infections of sporozoites before they caused symptoms as merozoites. The majority of research into sporozoite vaccines has focussed on CSP [99, 100], with one CSP-based vaccine candidate through Phase III trials. This vaccine, RTS,S, produced by the Walter Reed Army Institute of Research in partnership with GlaxoSmithKline, consists of the thrombospondin domain and central repeat of CSP fused to hepatitis B surface antigen. The results from the Phase III trial show that, with a booster dose (20 months after the initial vaccination), this vaccine has an efficacy around 30% [106]. This compares unfavourably with the WHO target of 90% vaccine efficacy, but is the best vaccine against malaria to date. Another

approach garnering much attention is the use of sporozoites, attenuated by radiation or genetic alteration, as a vaccine. Irradiated sporozoite Phase I trials have shown promise in these small scale trials, whilst antibodies raised using genetically attenuated sporozoites can block sporozoite infection [107, 108].

Anti-merozoite targets have mostly focussed on highly expressed proteins at the surface of the parasite, such as MSPs and AMA1. However, though antibodies raised against these MSP1 block invasion [102, 109] *in vitro*, and vaccination with AMA1 confers some protection against *P. falciparum* in monkeys [103], phase II trials using either antigen show little protection [110, 111]. The reason for this is unclear, but additional trials are ongoing using different vaccine formulations. A more recent target that has garnered much attention is Rh5, which lacks the polymorphisms found in other parasite surface proteins. This protein is essential for the successful invasion of erythrocytes [104], and so a vaccine targeting this protein should prevent the parasite from entering the erythrocytic replication cycle. This protein is well conserved between different *P. falciparum* strains, and antibodies raised against Rh5 can retard growth of homologous and heterologous parasite strains through binding Rh5 itself or host receptor basigin [112, 113]. Additionally, a *Pf*Rh5-based vaccine has recently shown promise, inducing protection against heterologous *Plasmodium falciparum* strains in *Aotus* monkeys [114].

1.2.4.1 Mimicking natural acquired immunity by targeting *Pf*EMP1

Natural acquired partial immunity exists in humans as adults in malaria endemic areas do not suffer severe cases of malaria to the same extent as children. Indeed children seem to acquire protection in an age-dependent manner [115]. One component of this natural immunity has been shown to be antibodies against *Plasmodium falciparum* erythrocyte membrane protein 1 (*Pf*EMP1), which is expressed by *Plasmodium falciparum* inside the erythrocyte and shuttled to the erythrocyte surface [7]. Malaria-exposed children exhibit strain-specific reactivity against this protein, developed after a couple of years [7, 116]. Moreover, the sickle cell trait carries a level of resistance against severe malaria, mediated through the impaired cytoadherence of infected red blood cells containing the abnormal

haemoglobin variant [117]. These observations have led to a substantial body of research focussed on using *Pf*EMP1 as a vaccine.

1.3 *Pf*EMP1 as a molecular determinant of cytoadherence and antigenic variation

*Pf*EMP1 was initially identified as a parasite protein on the surface of infected erythrocytes, that was strain-specific and a source of antigenic variation of the surface [118, 119]. This protein, varying in size between 160 and 300 kDa [119, 41] is transcribed by the parasite in the intra-erythrocytic ring and trophozoite stages [120] and trafficked to the erythrocyte surface independent of the parasite PEXEL trafficking machinery, despite containing a PEXEL-like signal sequence [40]. Here, they are held in knob structures on the surface of the infected erythrocyte, where they can mediate interactions with host ligands [41].

After the identification of *Pf*EMP1 on the surface of the infected erythrocyte, the link between this protein and sequestration of infected erythrocyte was explored. A link between the two was quickly established, with *Pf*EMP1 expression and cytoadherence modulated in a similar fashion by selection on melanoma cells [121]. A contemporary study showed that sera from *Aotus* monkeys, that blocked *in vitro* cytoadherence, reacted with *Pf*EMP1 [122]. This serum was shown to localise to the electron dense knobs, that had previously been shown as the site of cytoadherence [123, 124]. Drawing together these lines of evidence led to the association of *Pf*EMP1 with the knobs of infected erythrocytes. Both the presence of *Pf*EMP1 and the presentation of these proteins in knob structures are necessary to the cytoadherence of the blood cell [41, 125].

Cytoadherence is important for the parasite, as it anchors infected red blood cells in the microvasculature of organs [56], preventing these cells from circulating to the spleen. This is advantageous for the parasite, as *falciparum*-infected erythrocytes are rigid compared to their uninfected counterparts. These rigid infected erythrocytes would otherwise get caught in the red pulp of the spleen, which acts as a cellular filter [126]. Here, infected

cells would be cleared by resident macrophages [126].

As well as adhering to the endothelial membrane, infected red blood cells can also adhere to both uninfected and infected red blood cells (respectively termed rosetting and clumping). Rosetting was originally observed as an infected red blood cell binding uninfected red blood cell spontaneously *in vitro* [127], and was linked to *Pf*EMP1 through the study of a rosetting and non-rosetting subpopulation of the lab strain R29 [128], followed by the characterisation of a rosetting *Pf*EMP1 [129] . While rarely seen in histological samples [130], this phenomenon has been observed in *ex vivo* circulation perfusion studies [131]. This phenomenon may be another method of immune evasion, shielding the infected parasite from components of the immune system, and would also hold uninfected red blood cells in close proximity, for the next generation of merozoites to invade. Clumping of infected red blood cells has been identified as a CD36-mediated phenomenon where platelets cross-link multiple CD36-binding red blood cells [132].

Molecular characterisation of the *Pf*EMP1 began in the 1990s with three papers studying different aspects of *Pf*EMP1s [133, 41, 134]. These three papers found that there were multiple, sequentially-distinct *Pf*EMP1 genes (termed *var* genes) in the parasite, but identified a conserved architecture for *Pf*EMP1 proteins. The extracellular domain consists of combinations of two domain types both rich in cysteines. The cysteine-rich Duffy-binding-like (DBL) domains show sequence similarity to portions of other malarial surface proteins, notably EBA-175, a *P. falciparum* merozoite invasion protein [135], and the Duffy antigen-binding proteins of *P. vivax* and *P. knowlesi* [136]. The second domain type, termed the cysteine-rich interdomain region (CIDR), contains some features of the C-terminal half of the DBL domain [137], notably some of the placing of cysteine and tryptophan residues [133]. At the N-terminus of the *Pf*EMP1 there is a semi-conserved head structure, consisting of a DBL and a CIDR domain. This is followed by combinations of between two and seven DBL and CIDR domains. Individual DBL and CIDR domains have been described to bind a range of ligands, including ICAM-1 [138], CD36 [139], EPCR [140], and placental chondroitin sulfate A [141, 142]. It is these ligand binding domains which mediate the interactions between the infected erythrocyte and other host cells [143]. A single *Pf*EMP1 can contain multiple ligand binding domains, allowing it to bind two receptors

mediating different interactions [144], although simultaneous ligand binding has not been demonstrated in the previous literature. The protein is anchored into the membrane by a single transmembrane helix. The intracellular side of this helix is followed by a conserved acidic terminal segment, so-called due to its net negative charge [133]. This acidic terminal segment is involved in knob presentation and in interactions with proteins in the cytosol such as the positively charged KAHRP (knob-associated histidine-rich protein) [145, 146].

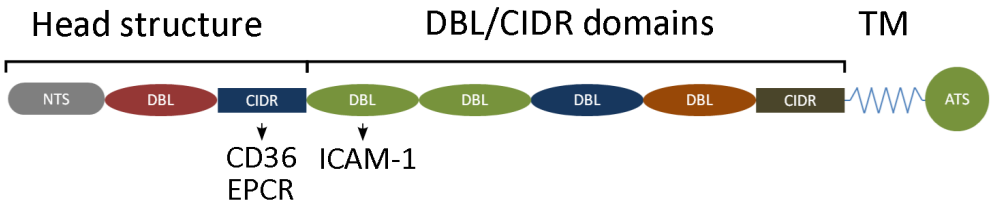


Figure 1.4: The typical domain architecture of a *Pf*EMP1 protein, showing the semi-conserved head structure, the following combination of DBL and CIDR domains, the transmembrane helix (TM) and the acidic terminal segment (ATS). Typical ligands are labelled underneath their binding domains.

Despite having a conserved architecture, *Pf*EMP1 proteins have extremely diverse sequences [133]. Indeed the genome of the 3D7 strain contains 59 different *var* genes [147], with a range estimated between 50-150 *var* genes per genome [133, 148, 149, 150]. These *var* genes are very similarly organised, consisting of two exons, split by one intron, with the first exon encoding the extracellular domain and the transmembrane helix, and the second exon encoding the intracellular domain. These *var* genes are located at the telomeres and centromeres of the *P. falciparum* chromosomes, arranged in head-to-tail, head-to-head and tail-to-tail tandem arrangements in the telomeres, while only being arranged in an head-to-tail orientation in the centromeres [147]. This large library of proteins was linked to the antigenic variation, previously observed on the surface of infected erythrocytes [151, 118], with parasites expressing one *var* gene by the trophozoite stage [152, 134, 153]. This monoclonal expression is tightly regulated at the transcriptional level [154, 155], with epigenetic silencing marks [156, 157], gene locus [158], the *var* intron and the 5' upstream promoter [159, 160] all linked to *var* gene expression.

Between generations of parasites, epigenetic marks transferred to the daughter cells can cause the same *var* gene to be expressed as the parent [161]. However, recent research

points to fast and slow coordinated switching between variants between one generation and the next, suggesting that switching is a more complex process than previously thought [154]. Clearly antigenic variation is a complex, highly evolved process, which allows the parasite to evade the adaptive immune system.

1.3.1 Identifying conservation in *Pf*EMP1 proteins to predict severity

While there is an evolutionary pressure to diversify the array of *Pf*EMP1 genes within the parasite to avoid the immune system, there is a counteracting pressure to conserve the adhering nature of the *Pf*EMP1 family, in order to ensure the survival of the parasite to replicate. In addition to this selection pressure on *Pf*EMP1s, certain *Pf*EMP1s seem to correlate with severe malaria or mild malaria, presumably due to their adhesive properties and the properties of their ligand. Naturally acquired anti-*Pf*EMP1 antibodies have been shown to bind to parasites from patients with severe malaria better than parasites from patients with mild malaria [162, 163, 116]. Whether this confers a selective advantage to the parasite is unclear. Perhaps there are two ways for the parasite to effectively multiply and spread, with parallels to be drawn with the viral world. It has been observed that different viruses ensure their survival by two contrasting mechanisms - either by a "hit and run" strategy where a high multiplication rate and high infectivity allow the virus to quickly spread whilst overwhelming host (influenza, measles), or by a "hit and stay" strategy where the virus does not destroy the host as quickly, rather aiming for survival and establishment of a stable population (hepatitis B) [164]. Parasites expressing *Pf*EMP1 variants associated with severe and mild malaria may represent a manifestation of these two phenomena. Gametocytogenesis is correlated with high parasitemia level observed in severe malaria [47, 165]. Other side effects of a high parasitemia have also been linked to a high rate of gametocytogenesis such as decreasing haemoglobin concentrations [165]. High rates of gametocytogenesis would allow a "hit and run" parasite to maximise the chances of transfer to the mosquito vector before the host dies. Mild malaria, with a lower parasite load, increases the chances of host survival to spread the parasite over a longer period of time. Regardless of whether this is true, the finding that certain *Pf*EMP1s correlate with severity of disease raises the possibility of conserved molecular determinants

in *Pf*EMP1 that could be targetted in therapies. Due to this, multiple efforts have been made to try and characterise the diversity observed between different *Pf*EMP1 proteins and differentially link subgroups of *Pf*EMP1s with disease severity or ligand binding.

1.3.1.1 *var* genes have a conserved upstream sequence

With the release of the complete 3D7 *P. falciparum* genome, a sequence was identified 1.5 kilobases upstream of each *var* gene, that could be assigned to one of three main classes, upsA, upsB or upsC, with upsD and E also found [147, 166]. Intermediate groups, in terms of chromosomal placement and orientation, B/A and B/C were also described. UpsA and B *var* genes were found at the telomeres of the *Pf*EMP1 in opposite orientations, and upsC *var* genes were located at the centromeres [147, 166]. This organisation has been suggested to facilitate recombination between members of each ups class, especially within the upsA and B subclasses, as the close proximity of the genes on the same chromosome would prevent alignment two *var* genes in opposite directions [166, 143]. Indeed members of the upsA class are closely related to one another, whilst upsB and upsC classes are related to one another, suggesting recombination, but not to upsA. UpsA *Pf*EMP1 proteins have been linked to severe malaria, with children acquiring antibodies to upsA proteins before B and C [116], and upsA are upregulated by selection by binding to endothelial surfaces [167]. However, their predictive power is limited due to their broad characterisation; not all upsA *var* genes are upregulated in severe disease [167]. Moreover, ICAM-1 binding has been observed to both upsA and upsB type *Pf*EMP1 [138, 168].

1.3.1.2 Domains are conserved between different *Pf*EMP1

The sequence variation within the individual extracellular domains has also been studied [169]. This study took the amino acid sequences from the twenty *Pf*EMP1 proteins sequenced at that point, and grouped the DBL and CIDR domains using the alignment program clustalW. The results of this study showed DBLs to fall into one of five classes named DBLα to ε, while CIDRs fell into one of three classes, named CIDRα to γ. Within these domains, there are areas of sequence similarity. These homology blocks have been

suggested as segments of protein sequence that are similar between domains [170]. Though useful, these blocks do not allow the identification of ligand binding sites.

Progress has been made for ICAM-1 binding prediction, with the DBLβ subclass identified as ICAM-1 binding [171]. Molecular indicators of CD36-binding have been identified using domain classification, since all identified CD36-binders are of the CIDRα subclass [172]. However, it is not true that all domains of one class bind to one ligand, with some CIDRα domains also shown to bind EPCR [140]. These sequence based approaches revealed the true diversity of individual domains, with ICAM-1 binders, for example, with sequence identity of around 45% [171].

1.3.1.3 Domain cassettes correlate with severity of disease

Another attempt to explain the remarkable diversity in *Pf*EMP1 sequence was introduced in 2010 [170]. Annotation of 399 *Pf*EMP1 proteins allowed the identification of domain cassettes, defined in this paper as "two or more consecutive domains belonging to particular subclasses" that are present in 3 or more of the 7 sequenced genomes. This data suggested twenty-two domain cassettes. Three of these domain cassettes had been previously identified as *var1*, *var2csa* and *var3* and are well conserved between parasites [173, 170] (Figure 1.5). Domain cassette 1, that encompassed all *var1* genes, is transcribed, but commonly truncated or mutated and seems not to be expressed at the surface [174, 173], and hence is thought to be a pseudo-var gene. Domain cassette 2, consisting of *var2csa* variants from different genomes, is well conserved and well-characterised as an important molecular determinant of pregnancy-associated malaria, binding to CSA, and upregulated in placental malaria [175]. Domain cassete 3 encodes the small *var3* *Pf*EMP1 proteins. These proteins only have two DBL domains on the extracellular side of the protein, though the physiological consequence of this is not known [176].

Four other domain casettes have since been identified to contain ligand binding domains, with three of these correlating with severe malaria (Figure 1.5). These ligand-binding interactions are conserved between instances of the same domain cassette. Domain cassette 4, consisting of DBLα1.4/DBLα1.1-CIDRα1.6-DBLβ3, has recently been identified in a sub-

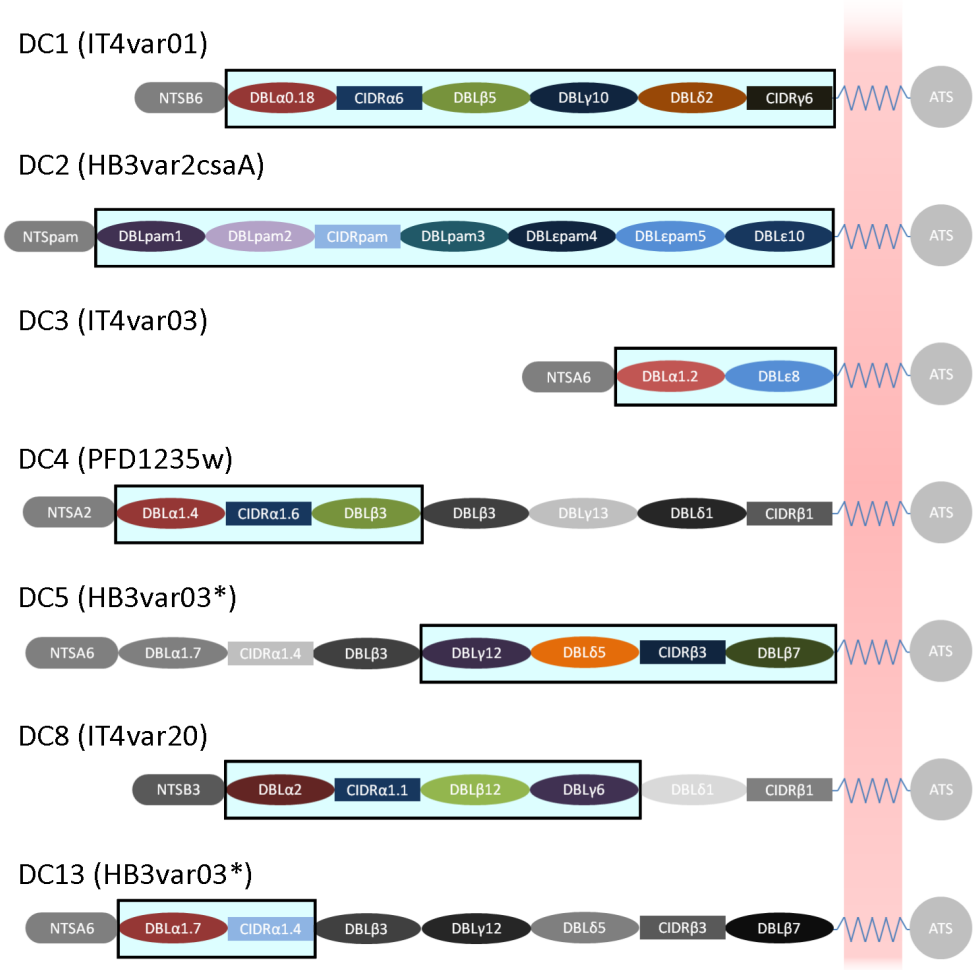


Figure 1.5: Domain cassettes, that have been linked to ligand binding or malaria severity are shown here. Each domain cassette is outlined in the context of an example *Pf*EMP1. *Note that HB3var03 has two domain cassettes within its architecture.

set of upsA-type *Pf*EMP1 proteins [168]. This domain cassette contains a subset ICAM-1 binding domains, and also contains a CIDRα1 domain subclasses. While CIDRα1 domain subclasses 1.4 and 1.1 have been shown to bind EPCR, CIDRα1.6 has not yet been tested for EPCR binding. Domain cassettes 5, 8 and 13 were studied in a field study in Tanzania in 2012, where these domain cassettes were found to correlate, to different extents, with severe malaria [177]. Though not conclusively correlated in this study, domain cassette 5 (DBLγ12-DBLδ5-CIDRβ3/4-DBLβ7/9) was transcribed to higher levels in patients with severe malaria than in patients with uncomplicated malaria. A follow-up study showed that anti-DC5 IgG was present in children living in malaria-endemic areas, with the ligand identified to be PECAM-1 [178]. Domain cassettes 8 and 13 have been conclusively associated with severe malaria both in the afore mentioned Tanzanian field study [177], and a second study on the opposite side of Africa in Benin [179]. In addition to this, domain cassettes 8 and 13 are upregulated when parasites are selected on binding to brain microvascular endothelial cells [180, 181].

These sequence-based approaches to characterise *Pf*EMP1 proteins have been partially successful, identifying domains that are conserved between different *Pf*EMP1. Domain cassettes could be of some use to predicting severity of malaria, though it is perfectly possible that some domains that correlate with severe malaria could fall outside the domain cassette classification system or that domains could undergo homologous recombination between different cassettes, reducing the usefulness of domain cassettes as a predictor of malaria severity.

1.3.1.4 Structural conservation of DBL and CIDR domains

Despite the limited level of sequence conservation, structures of DBL and CIDR domains have revealed there to be limited structural conservation between domains [182]. Ten DBL domains have thus far been crystallised, with five domains from invasion proteins and five from *Pf*EMP1 proteins [183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193]. These DBL domains have 8% sequence identity, and 18% sequence similarity. Analysis of these structures reveals two core helical segments that are conserved between different

DBL domains. The 14 completely conserved residues from these domains point inwards into the structure, helping to maintain this fold [182] (Figure 1.6).

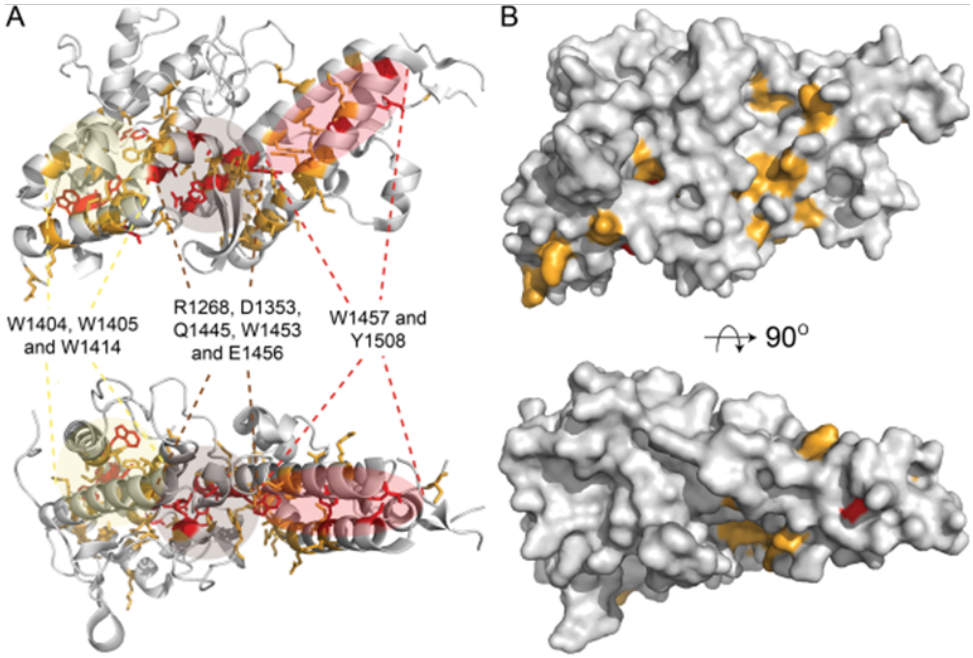


Figure 1.6: The conserved DBL fold is shown in cartoon representation (A) and surface representation (B). Completely conserved residues (red, labelled) point inwards into the structure, with few patches of conservation apparent on the surface of the protein. Figure from reference [182].

At the start of this DPhil, there were only two crystal structures of the CIDR domain [137, 189]. While these domains have a conserved set of helices, identified by their conserved cysteine and tryptophan residues, they differ considerably in overall conformation, with the MC179 adopting an open conformation between the two sets of helices, while the two bundles in the var0 structure are packed closer together (Figure 1.7).

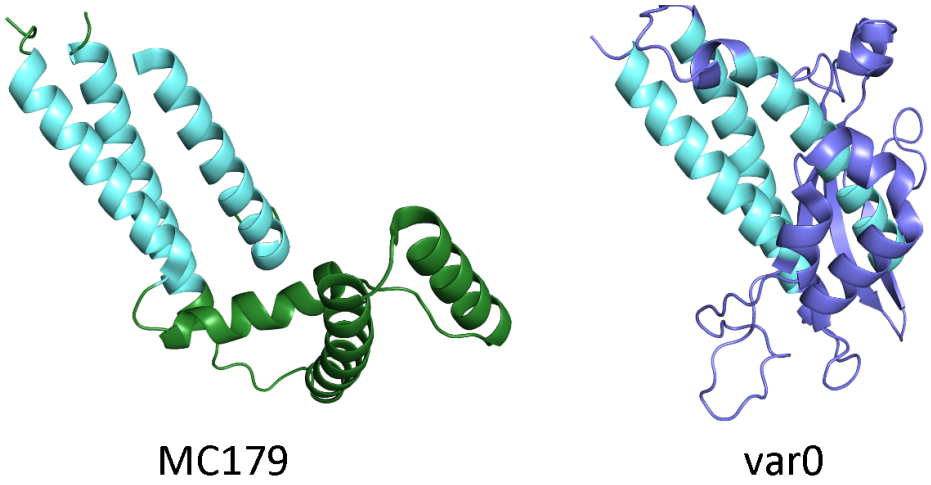


Figure 1.7: The conserved CIDR fold is shown in cartoon representation for the two solved CIDR structures, with the putative conserved helices in light blue.

1.3.2 The correlation of *Pf*EMP1 ligand binding with disease severity

As well as attempting to link the extremely diverse *Pf*EMP1 proteins with disease severity, the ligands to which *Pf*EMP1 proteins bind have been extensively studied. The hypothesis underlining this line of inquiry is that since sequestration is highly correlated with malarial severity, and since *Pf*EMP1 variants that correlate with severe and uncomplicated malaria are somewhat distinct, the location, expression level and identity of the ligand to which the sequestered red blood cells bind is likely to influence the severity of the infection.

Many *Pf*EMP1 ligands have been identified, with the majority either responsible for cytoadherence of the infected erythrocyte to the membrane, clumping of infected red blood cells or rosetting to uninfected red blood cells. CD36 [194], ICAM-1 [194, 138], CSA [195], PECAM-1 [178], CR1 [128], HABP1 [196], IgM [197], thrombospondin-1 [194], P-selectin [198] and EPCR [140] have all been suggested to bind to *Pf*EMP1, with fractalkine [199], E-selectin [200], VCAM-1 [200] also suggested to mediate interactions between infected erythrocytes and host cells. However, despite this large diversity of ligands, certain *Pf*EMP1:ligand interactions are seen more often than others in patients. CD36 and ICAM-1 are often identified as the main ligands of *Pf*EMP1 [201], with CSA identified as the most important ligand in pregnancy-associated malaria [175], and recent attention focussed on EPCR, identified as a *Pf*EMP1 ligand in 2012 [140].

1.3.2.1 The interaction of *Pf*EMP1 with CD36

CD36-binding is the most common adhesive phenotype [202], and is expressed in a diverse range of tissues. This ligand was originally identified as a human molecular determinant for *P. falciparum*-infected erythrocyte cytoadherence [203, 204]. It has since been characterised as a class B scavenger receptor, with a diverse range of functions, implicated in angiogenesis, lipid metabolism and pathogen recognition [205]. As CD36 has immunological functions, *Pf*EMP1 binding to CD36 may have immunological consequences. Cells over-expressing CD36 have been shown increase the uptake of LPS to mediate bacterial phagocytosis. These cells are also known to activate dendritic cells and macrophages to stimulate immune system activation and phagocytosis [206]. This pathway is known to be important for the regulation of malaria infection [207], as CD36-/- mice are more susceptible to high *Plasmodium chabaudi* parasitaemia levels than CD36+/+ [208]. However, CD36-expressing dendritic cells, after incubation with parasitised erythrocytes, did not stimulate T-cells as effectively as immature or mature dendritic cells [209]. While the precise molecular mechanisms to explain these rather contradictory phenomena are not clear, this may represent the "hit and stay" niche that allows *P. falciparum* to inhabit the human host in an almost commensal manner. Indeed, CD36 binding does not correlate with severity of disease [210].

1.3.2.2 The interaction of *Pf*EMP1 with ICAM-1

Unlike CD36, binding to ICAM-1, which is observed less frequently as a binding phenotype, does correlate with severe disease [210]. ICAM-1 was discovered as a molecule that mediated lymphocyte-fibroblast adhesion, upregulated on the surface of dermal fibroblast in response to pro-inflammatory cytokines, interferon-γ and interleukin-1 [211]. This protein, consisting of five immunoglobulin domains followed a transmembrane helix and a short cytoplasmic segment, is involved in the "postcode principle", whereby local secretion of chemokines can cause bloodstream endothelial cells to express receptors to bind to passing lymphocytes [212]. ICAM-1 has been shown to be expressed in higher quantities in patients with severe malaria than non-severe malaria [213, 210, 214], with upregulation

observed in several organs, particularly the brain. This, combined with the fact that CD36 is not expressed in the brain has led to the association of ICAM-1 with severe malaria. ICAM-1 binds to DBLβ domains from upsA, upsB and ups C *Pf*EMP1s [171, 168]. A binding site on the DBLβ has been identified for upsA DBLβ [215], and upsB DBLβ [216]. On the ICAM-1 surface the DBL binding site has been suggested to be located on a putative dimerisation interface, the BED face [217, 218]. It is not currently clear whether the dimeric ICAM-1 interface must be broken for *Pf*EMP1 to bind.

1.3.2.3 The interaction of *Pf*EMP1 with CSA

Chondroitin sulfate A is somewhat of a special case. It is strongly linked with pregnancy associated malaria [219], with var2CSA identified as its cognate *Pf*EMP1. Indeed this *Pf*EMP1 is upregulated upon panning with CSA in both laboratory strains [175], and is expressed by parasite isolates from pregnant women [220, 221]. Unlike other *Pf*EMP1 proteins, var2CSA is well conserved between different families [175]. In addition to this multiple domains from the same var2CSA *Pf*EMP1 sequence have been shown to have CSA-binding ability [222, 142]. Due to the identification of this variant as the determinant of pregnancy-associated malaria, this *Pf*EMP1 is currently in vaccine trials.

1.3.2.4 The interaction of *Pf*EMP1 with EPCR

EPCR has gained much attention in the last two years since its association with severe malaria in 2013 [140]. In this study, a domain cassette 8-containing *Pf*EMP1, IT4var20, was screened against 2505 human plasma membrane proteins using the Retrogenix Cell Microarray (Retrogenix Ltd., UK). This screen identified EPCR as the sole receptor for both domain cassette 8, associating EPCR with severe malaria. Indeed, some gene variants of EPCR correlate with protection from severe malaria [223], and aberrant brain endothelium expression of EPCR associates with severe malaria [224, 225]. Subsequent studies with different length *Pf*EMP1 constructs from both domain cassette 8 and domain cassette 13 found that the minimal binding portion of the *Pf*EMP1 that could bind to EPCR was the CIDRα1 domain contained within both domain cassettes. This domain bound to EPCR

with a similar affinity to the whole protein, with an observed K_D from surface plasmon resonance data of 29 nM for the CIDR versus 10 nM for the whole protein.

EPCR is an endothelial receptor, in the MHC class superfamily, which is ubiquitously expressed on the bloodstream endothelium [226]. Previous structural studies show this protein to have a lipid bound in its central cavity, rather than a peptide, as in other MHC family members [227] (Figure 1.8). This protein interacts with protein C in the bloodstream [228, 227] (Figure 1.8), which is generated by the thrombin-thrombomodulin complex responding to tissue damage [229]. Upon interaction, protein C becomes activated, binding to protease activated receptor 1 (PAR-1) *in cis* [230]. PAR-1 transduces the signal inside the cell to downstream effectors of this signal, which mediate anti-inflammatory and cyto-protective effects [231] (Figure 1.9, top panel).

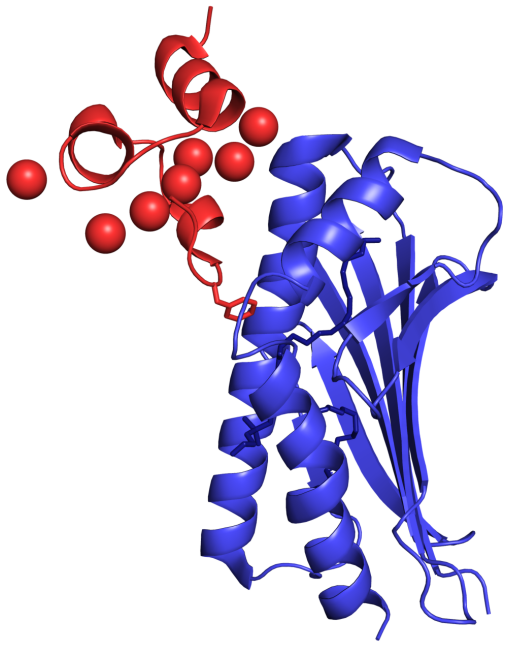


Figure 1.8: The crystal structure of EPCR (blue) bound to protein C (red) [227]. The bound lipid in the EPCR cavity is in deep blue, stick representation, while the protein C bound calcium ions are in sphere representation.

After the initial discovery of EPCR as a *Pf*EMP1 ligand, further SPR experiments showed that protein C could be prevented from binding to EPCR by CIDR domains from *Pf*EMP1 proteins encoding domain cassette 8 [140]. This observation led to the proposal of an immuno-modulatory function of *Pf*EMP1 binding. This model proposes that blocking of EPCR prevents protein C from binding, causing thrombin to bind to PAR-1 and mediate

pro-inflammatory effects [232], rather than anti-inflammatory effects (Figure 1.9). Since there is a lower amount of EPCR on the brain capillary endothelia than elsewhere in the bloodstream [226], these vessels may be especially sensitive to this inhibition of protein C by sequestering infected erythrocytes.

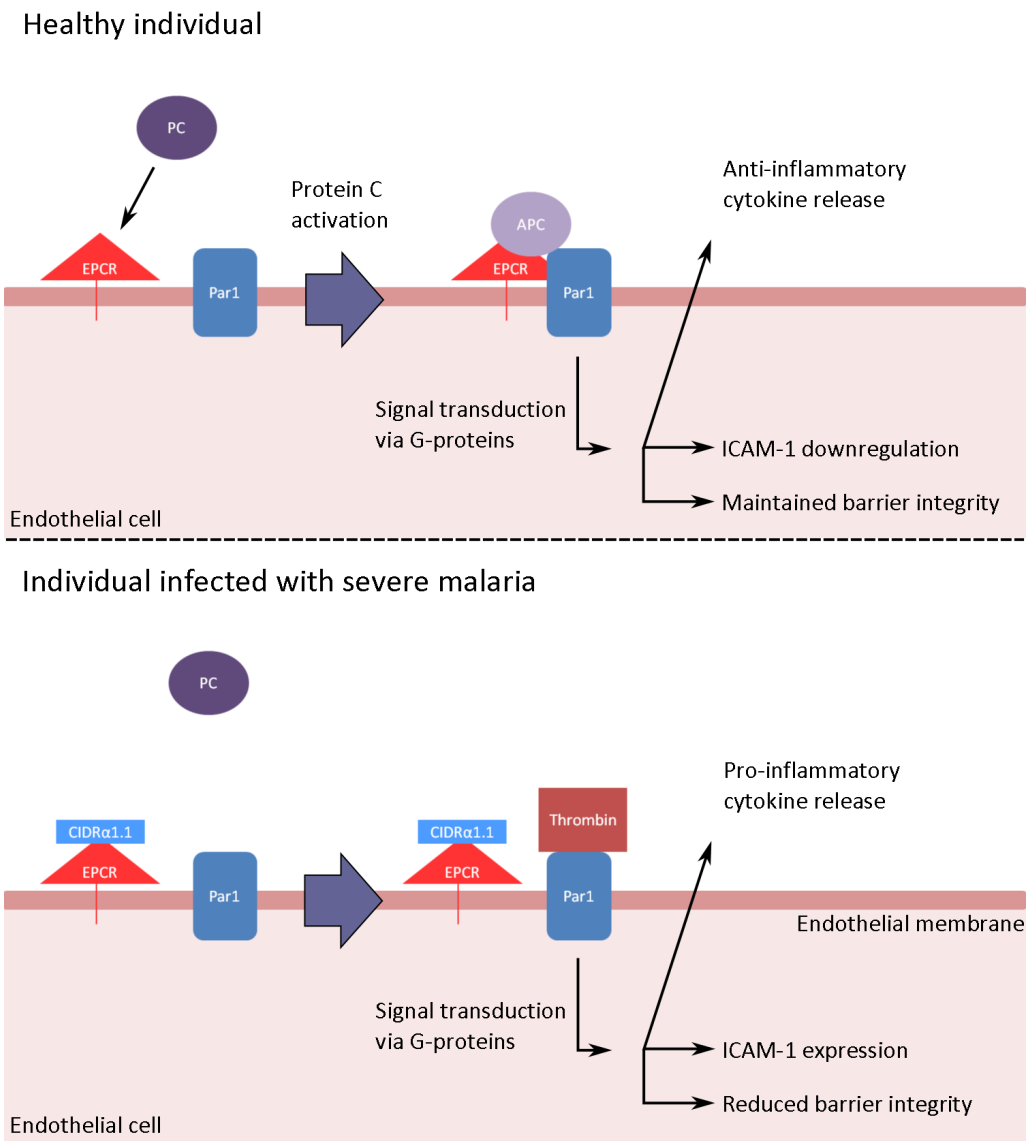


Figure 1.9: The proposed model depicting how EPCR-binding *Pf*EMP1 proteins could modulate the immune response, depicting the PAR-1 signal pathway in the absence (top panel) and presence (lower panel) of EPCR-binding *Pf*EMP1 proteins.

This model is tantalising, not least because inflammation of the capillaries is associated with cerebral malaria, observed in the brains of suffering patients [42], but also because this model has links to ICAM-1 expression. Thrombin signalling, through PAR-1, in

the absence of protein C signalling, induced ICAM-1 expression in human umbilical vein endothelial cells as one of the inflammatory effectors [233]. This could allow EPCR-binding *Pf*EMP1 proteins to act to express ICAM-1 proteins, available for any ICAM-1 binding *Pf*EMP1s, potentially enhancing the level of sequestration of infected erythrocytes.

1.4 The suitability of *Pf*EMP1 as a vaccine candidate

*Pf*EMP1 is an attractive target for vaccine design as natural acquired immunity, developed in patients after only a couple of infections with *Plasmodium falciparum*, is at least partially targetted against *Pf*EMP1 [7]. However, the major challenge that needs to be surmounted is the identification of conservation to target in a vaccine. The antibodies that were responsible for immunity, described in early studies, were strain-specific. This is not surprising since sequence conservation between *Pf*EMP1 proteins is extremely low. Indeed between two domains that bind the same ligand, sequence identity can be as low as 40%. A successful vaccine could not hope to contain all *Pf*EMP1 proteins from all strains, so a conserved target on the *Pf*EMP1 needs to be identified, that would produce cross-reactive antibodies. This vaccine may need to inhibit multiple *Pf*EMP1 interactions to fully prevent sequestration of infected erythrocytes in the host.

The only current *Pf*EMP1 vaccine candidate to date is the var2csa *Pf*EMP1, responsible for pregnancy-associated malaria, due to the limited diversity in var2csa sequence. Cross-inhibitory antibodies have been raised, with the extent of cross-inhibition dependent on subunit composition of the immunogen [234, 235]. Currently, studies are ongoing to identify the ideal vaccine immunogen [236].

1.4.1 Methods to elicit broadly inhibiting antibodies

If a conserved target can be identified, several immunogen design methods could be used to focus the immune response against this epitope. This is currently a major goal in viral vaccine design, especially in influenza and HIV-1 [237, 238]. In this area, vaccine design strategies such as glycan masking, stapled helical peptides and epitope grafting are being

used to try and select an epitope.

1.4.1.1 Glycan masking

Glycosylation has been suggested as a method of masking immunodominant, non-inhibitory viral epitopes, and attempts have been made to use artificially induced N-linked glycans to selectively mask parts of a potential immunogen. Indeed glycosylation of naturally exposed sections of the V3 loop in the HIV-1 protein gp120 result in dampening of the previously immunodominant epitope [239]. Subsequent experiments have tested the use of glycans to both mask epitopes and specify epitopes [240, 241]. A glycan-masked gp120 mutant has been created where an inhibitory epitope is left exposed, which no longer binds non-neutralising antibodies, whilst preserving antibodies to the inhibitory epitope [240, 242]. In another study, sera raised against a glycan-engineered haemagglutanin contained inhibitory antibodies that could not only inhibit the same strain virus, but also inhibited two viruses of other strains [243].

This strategy has been attempted in *Plasmodium vivax*, where the Duffy binding protein was engineered with glycosylations, attempting to mask epitopes distinct from a sulfotyrosine binding pocket and dimerisation domain, which are thought to be important to invasion of reticulocytes [244]. However this study was not successful in eliciting inhibitory antibodies in greater proportions than wild-type.

1.4.1.2 Stapled helical peptides

Removing the epitope from its natural environment is an attractive prospect. This strategy allows the circumvention of the problem of immunodominant epitopes, and facilitates the specification of the epitope. However, presenting an epitope by itself can present challenges if the epitope is expected to have secondary structure. Short peptides may not fold into secondary structure due to a lack of the cooperativity required to form an α-helix, for example. Stapled peptides promise the solution to this problem. By introducing synthetic amino acids, capable of cross-linking between either the positions i & $i+4$, and i & $i+7$,

the peptide is driven into a helical conformation [245]. This approach has previously been used to recreate a HIV-1 epitope using a small stapled peptide [246]. This peptide bound to broadly neutralising antibodies, but has not yet been used as an immunogen. Care must be taken if this immunogen is taken forward as peptides are poor immunogens and may be processed differently by antigen-presenting cells [247].

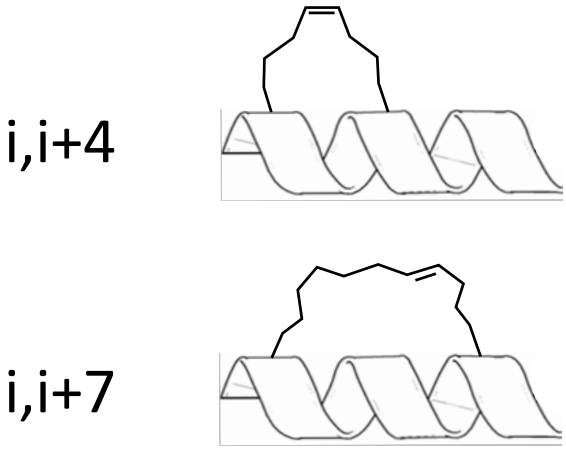


Figure 1.10: Illustration depicting two different covalent staples between residues (i,i+4) and (i,i+7), designed to drive helical peptide formation.

1.4.1.3 Epitope grafting

An alternative approach has arisen with the advances in *in silico* design programs, facilitating both the prediction of protein conformation using energy scoring functions and the manipulation of both structure and sequence. The Rosetta software suite has made some large advances in the field of vaccine design [248]. This software has facilitated the recreation of a target epitope on a scaffold protein in several different methods, depending on the complexity of the epitope.

Various options are currently available to recreate a target epitope. Sidechain grafting is the simplest method to graft an epitope onto a scaffold. In this process the backbone polypeptide is first matched between the epitope and a scaffold. This allows the sidechains on the scaffold to be mutated to match the epitope sidechains. This method has been used to transplant an inhibitory epitope from respiratory syncytial virus (RSV) onto a

scaffold from *Staphylococcus aureus* [249]. This technique deals well with commonly found secondary structural elements such as helices, but can struggle with loops or distorted helices. To graft epitopes consisting of non-standard secondary structure, backbone grafting can sometimes provide a solution. Instead of identifying a scaffold backbone to mutate the sidechains, the entire polypeptide, including the backbone is grafted from the epitope to the scaffold [250, 251]. This was used to try and replicate HIV gp120 and gp41 epitopes, that elicit broadly neutralising antibodies. The most complex grafting procedure is called fold from loops (FFL) [252]. This recent approach, used to recreate an RSV epitope on a small scaffold, is similar to backbone grafting, but here the scaffold is folded around the epitope, which is held rigid. This is advantageous, as it can force the scaffold to accommodate the epitope, rather than the epitope changing to accommodate the scaffold. The scaffold is then mutated to stabilise the artificial conformation. This approach holds great promise, as long as a suitable starting scaffold is found.

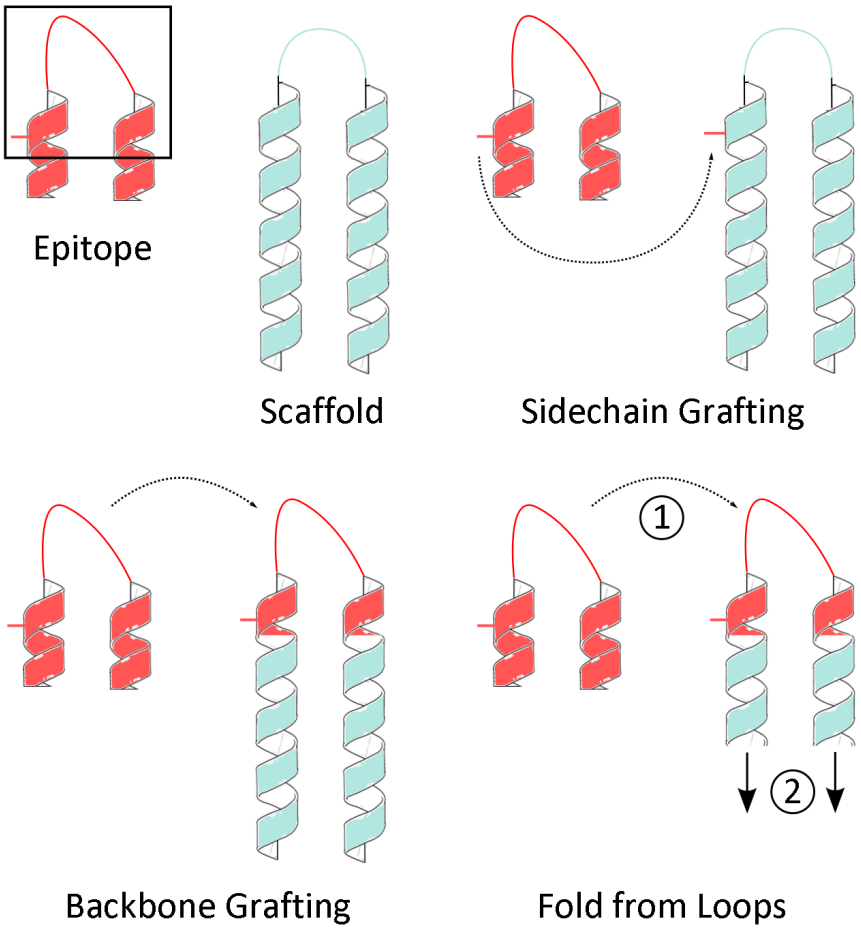


Figure 1.11: Illustration depicting three epitope grafting methods to graft an epitope (red, boxed) onto a scaffold (teal): sidechain grafting (side chain depicted by a stick); backbone grafting; fold from loops, where (1) the epitope is locked in place then (2) the scaffold is folded around the epitope.

1.5 Project Aims

At the start of my DPhil, I aimed to solve the first structure a *Pf*EMP1:ligand complex. I chose to focus on the interaction between the CIDR domain of the *Pf*EMP1 protein and EPCR, as this recently identified interaction correlates with severe cerebral malaria.

Using this structure, I wanted to identify the molecular determinants of the *Pf*EMP1:-EPCR interaction, and see how these determinants are conserved between different EPCR-binding CIDR domains. This would be achieved by mapping the sequence of diverse EPCR-binding CIDR variants onto the structure, to identify regions of conservation. This would allow the identification of a suitable epitope to target in immunogen design.

Since EPCR and ICAM-1 are linked in the current model of EPCR-mediated pathogenesis of severe malaria, I wanted to investigate the possibility of simultaneous interaction of *Pf*EMP1 constructs with both ligands. Biophysical and structural information would reveal whether this needs to be taken into account or could even be exploited when designing an immunogen. *Pf*EMP1 domains that contain adjacent EPCR and ICAM-1 binding domains had been identified by our collaborators. Given that the current, albeit speculative, model of the binding includes ICAM-1 expression as a downstream effector, this dual interaction is of great interest.

Having identified conserved determinants of the *Pf*EMP1:EPCR interaction, I would then attempt to create novel immunogens, designed to raise inhibitory antibodies to EPCR. Glycan masking, stapled helical peptides and epitope grafting would be attempted to focus the response to the binding site. These novel immunogens would be tested by immunising rats and assessing the polyclonal sera, with the aim to produce a higher proportion of inhibitory antibodies than injection of the CIDR alone.

Chapter 2

Methods

2.1 Cloning

2.1.1 Bioinformatic analysis

FUGUE [253] was used as guidance to select construct domain boundaries. This program compares the input amino-acid sequence first to sequence homologues, then to secondary structure profiles of proteins from the protein databank (PDB) to predict structure-sequence homology. The algorithm compares local segments of primary sequence to secondary structure profiles, taking into account the amino acid identity and typical placement in structures to predict secondary structure. Clustal Omega was used for multiple sequence alignments, to search for sequence homology within a set of related proteins. This program uses a progressive alignment system to compile the complete alignments [254]. ESPript was used to visualise the alignments [255]. Glycosylation sites within the protein sequence were predicted using the NetNGlyc server (http://www.cbs.dtu.dk/services/NetNGlyc/). Structures from the protein databank were viewed using Pymol [256].

2.1.2 Polymerase Chain Reaction

Polymerase chain reactions were carried out using KOD Hot-Start DNA polymerase (Millipore-Merck, Germany) on an MJ Mini Personal Thermal Cycler (Bio-Rad, USA), according to the manufacturer's protocol. A typical cycle would consist of 95°C, 3 minutes, to activate the polymerase; then 25 cycles consisting of 95°C for 1 minute to denature the DNA, 55°C for 1 minute for primer annealing (Appendix A) contains a list of primers used) and 68°C for 1 minute.

2.1.3 Restriction Digests

Restriction digests were performed using restriction enzymes, according to the table in Appendix A (New England Biolabs, USA). DNA samples were incubated with 1 µl of each enzyme in CutSmart Buffer (NEB) at 37 °for 15 h. Cut DNA was separated from other components using the PCR purification protocol (QIAGEN, the Netherlands) for small inserts or the QIAquick gel extraction kit (QIAGEN) for vectors, according to the manufacturer's protocol.

2.1.4 Ligation

In the case of ligation of an insert into a vector backbone, 10 µg of cut vector DNA was mixed with 10µg or 70µg of cut insert. These reactions were incubated with 1 µl of T4 DNA ligase (NEB) in the presence of T4 DNA ligase buffer for 15 h at 16 °C. In the case of ligation of two vector ends, 10 µg of vector was incubated with 1 µl of T4 DNA ligase in the presence of T4 DNA ligase buffer for 15 h at 16°C. 1 µl of ligation mixture was then transformed into DH5α *Escherichia coli* cells (according to manufacturer's protocol, Thermo Fisher, USA), and plated onto agar containing the appropriate antibiotic. These plates were incubated at 37°C for 15 h. Colonies that grew were individually picked and grown up in 3 ml Lysogeny Broth (LB) for 15 h at 37 °C. Vector DNA then was extracted from these bacteria using a QIAprep miniprep kit (QIAGEN). This DNA was sequenced using a Sanger sequencing service (Source Bioscience, UK).

2.1.5 Site Directed Mutagenesis

Sets of complementary primers were designed, flanking the mutated DNA sequence with at least 15 base pairs, complementary to the vector (Appendix A). These primers were used to amplify the gene and vector via PCR (as in section 2.1.2, with an extended extension time (5 minutes) to accommodate the extra length of DNA to be polymerised). This amplified DNA was treated with DpnI (NEB, USA) for 6 h at 37°C to cleave methylated template DNA. The mutated DNA was transformed into XL-10 Gold Ultracompetent cells (according to manufacturer's protocol, Agilent Technologies, USA) and plated on agar plates with an appropriate antibiotic. These plates were incubated at 37°C for 15 h. Colonies that grew were individually picked and grown up in 3 ml LB for 15 h at 37 °C. Vector DNA then was extracted from these bacteria using a QIAprep miniprep kit (QIAGEN). This DNA was sequenced using a Sanger sequencing service (Source Bioscience, UK).

2.1.6 Agarose Gel Electrophoresis

DNA samples were mixed with 6x gel loading dye (NEB, USA) and run on a 0.1% (v/v) agarose gel in TBE buffer (89 mM Tris pH 8.0; 89 mM boric acid, 2 mM EDTA) at 100 mA for one hour. Samples were imaged using a G:Box Chemi-XRQ imaging system (Synoptics Ltd, UK). Bands were compared against the 100 bp or 1 kb DNA Hyperladder plus (Bioline, UK), or 100bp or 1 kb DNA ladder (NEB, USA) to infer their size. If bands were excised, a Dark Reader® (Clare Chemical Research, USA) was used to illuminate the bands, then bands were excised using a clean razor blade. DNA could then be extracted using a QIAquick Gel Extraction Kit (QIAGEN).

2.2 Expression

2.2.1 *E. coli* expression

Escherichia coli is the simplest and quickest expression system in regular use. There are many different strains for overexpression of protein, mostly based on the T7 system, which harnesses the power of the bacteriophage T7 RNA polymerase to express large quantities of recombinant protein [257]. All *E. coli* expression described in this thesis used the pET15b expression vector backbone (Novagen, Merck, Germany). This backbone places the gene of interest behind an IPTG-inducible promoter, with a lactose repressor gene to prevent expression before induction. This vector has been converted by Prof. Higgins to include an insert encoding a hexahistidine tag, followed by a tobacco etch virus (TEV) protease cleavage site upstream of the multiple cloning site (pEt15bcon, Figure 2.1). The two additions to the pEt15b backbone would allow for affinity purification of the recombinant protein using nickel-NTA agarose, with the option to cleave the tag using TEV protease.

2.2.1.1 Preparation of chemically competent *E. coli*

When commercial stocks of chemically competent cells were not used, cells were made competent to DNA uptake using 100 mM $CaCl_2$. 3 ml cultures of the *E. coli* strain, containing the appropriate antibiotic were grown for 15 h, then centrifuged at $14,000 \times g$ for 1 minute to pellet the bacterial cells. This pellet was resuspended in 1 ml 100 mM $CaCl_2$ at 4°C and once again centrifuged at $14,000 \times g$ for 1 minute. Cells were resuspended in 100 µl of 100 mM $CaCl_2$ at 4°C and frozen using a dry ice/ethanol bath if not immediately required.

2.2.1.2 Small scale expression trials

Constructs trialled for expression in *E. coli* were grown up in LB to an optical density (600 nm) 1 at 37°C. Aliquots of these cells were induced with 1 mM (final concentration) isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated at three different tempera-

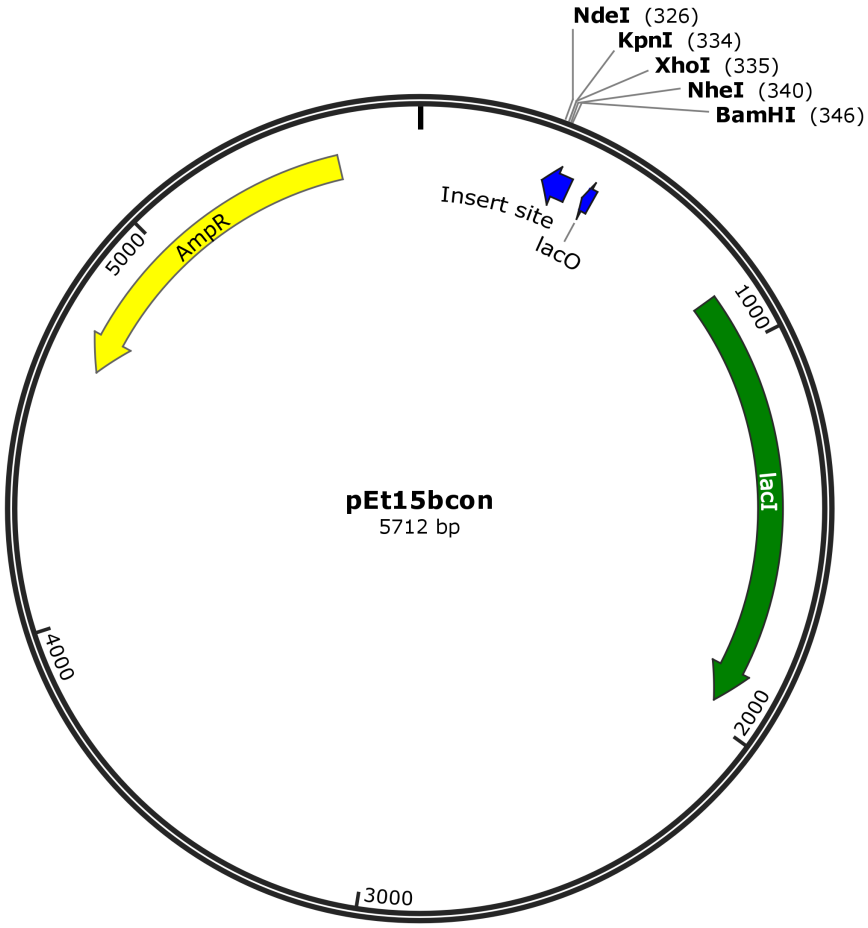


Figure 2.1: pEt15bcon vector map, showing the nucleotide position of various features.

tures (typically 25°C, 30°C, 37°C), with samples taken after 3 hours and after 15 hours. Cells were pelleted by centrifugation at $14000 \times g$, lysed in PBS by sonication for 4 seconds and centrifuge at $14000 \times g$ for 5 minutes. Whole cell samples were taken before the final centrifugation to assess the amount of protein in the whole cell, and after to assess amount of soluble protein. These samples were analysed by Western blot, using an anti-His antibody to detect the amount of protein that had been expressed with a histidine tag (Section 2.3.2).

2.2.1.3 Large scale expression

To express proteins on a larger scale in *E. coli*, overnight cultures were diluted into 1 L of 2xYT broth (Melford Laboratories Ltd, UK) in 2 L plastic Tunair® flasks (Sigma-Aldrich, Germany). Bacterial cultures were grown at 37 °C to an optical density at 600 nm of 1. These cultures were then induced with 1 mM (final concentration) IPTG, before expression at temperatures and times determined by expression trials. After expression, cells were pelleted at $4{,}000 \times g$ for 20 minutes.

2.2.2 Lepidopteran cell/baculovirus expression

The lepidopteran cell/baculovirus expression system was first harnessed in 1983 to overexpress human interferon-β [258]. It has since become a very powerful expression platform, primarily for the expression of eukaryotic proteins, which can require post-translational modifications and folding machinery not provided by prokaryotic systems [259]. Constructs for lepidopteran/baculovirus expression were cloned into the pAcGP67-A vector (BD Biosciences, USA), which has been converted to include an insert encoding a hexahistidine tag, followed by a tobacco etch virus (TEV) protease cleavage site upstream of the multiple cloning site (pAcGP67-Acon, Figure 2.2). This vector is co-transfected with linearised AcNPV baculovirus DNA into insect cells, where the vector integrates with the baculovirus DNA via homologous recombination, which can be transcribed to produce recombinant viral particles. This vector contains a polyhedrin promoter that is harnessed for constitutive expression of the gene of interest, with a GP67 secretion signal facilitating

secretion of recombinant proteins, containing an N-terminal his-tag, into the media (Figure 2.2). Recombinant virus was made in *Sf*9 cells, derived from *Spodoptera frugiperda* ovarian tissue, and recombinant protein expressed in *Sf*9 cells or High Five cells, derived from *Trichoplusia ni* ovarian cells.

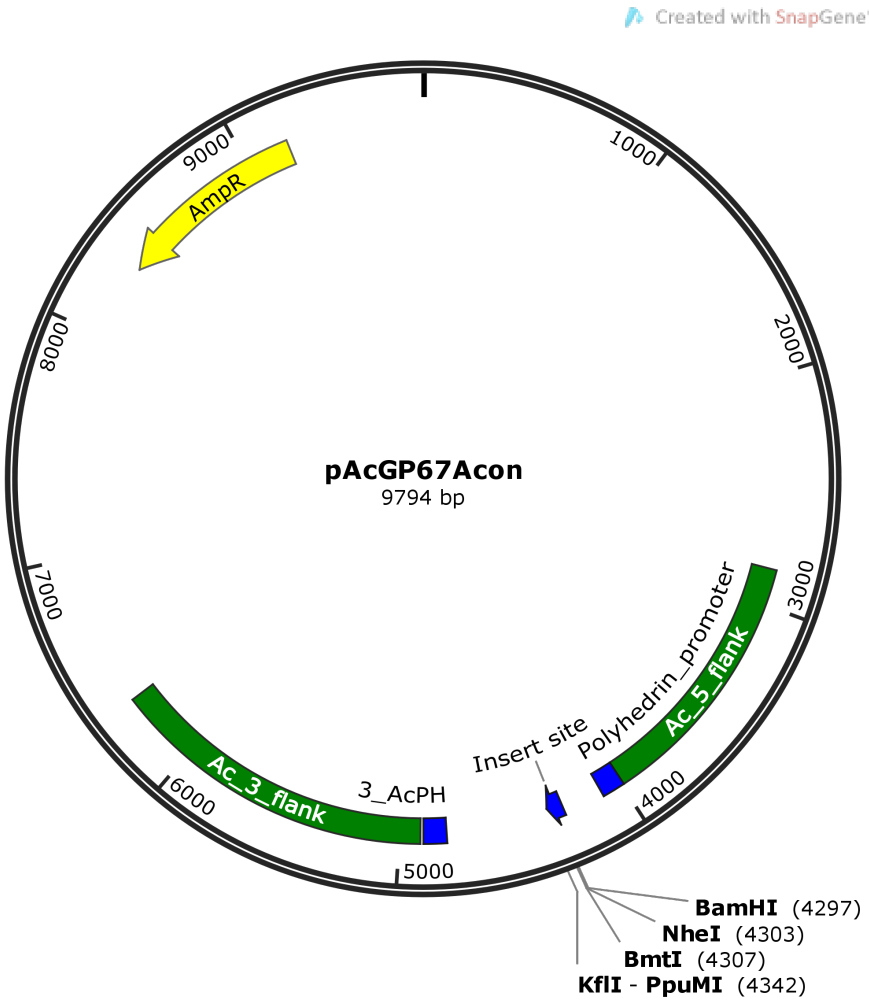


Figure 2.2: pAcGP67Acon vector map showing the nucleotide position of various features.

2.2.2.1 Maintenance of insect cell lines

*Sf*9 and High Five cells were both grown in suspension culture at 26°C in Erlenmeyer flasks (Corning Inc., USA) shaking at 150 rpm. Cells were maintained in Sf-900™ II SFM (ThermoFisher Scientific, USA) for *Sf*9 cells, or Express Five® SFM (ThermoFisher Scientific, USA) supplemented with L-glutamine (not included in the media formulation,

Sigma-Aldrich). Cells were diluted to 5×10^5 cells/ml every two days.

2.2.2.2 Transfection of *Sf*9 cells

Constructs in the pAcGP67con vector were co-transfected into 7×10^5 *Sf*9 cells with linearised AcNPV baculovirus DNA (BaculoGold transfection kit, BD Biosciences, USA) according to manufacturer's protocol, harvesting virus six days later. This virus mixture was plaque purified to select single viral lines for expression. This was achieved using plaque assays: monolayers of 5×10^5 cells were incubated with 10^2-10^5 dilutions of the initial viral stock for one hour, then the viral stock was removed and replaced with 0.4% Agarose; 1.3x Sf9 media; 10% penicillin-streptomycin mixture (100 units/ml Penecillin; 100 µg Streptomycin; Sigma-Aldrich). These plates were incubated at 26°C for 7 days, then single plaques were identified and picked.

2.2.2.3 Viral amplification

Single plaques were incubated with 1.5×10^6 *Sf*9 cells in 5 ml Sf-900™ II SFM (ThermoFisher Scientific, USA) in a T25 flask (Corning Inc., USA) in static culture at 26°C. P1 virus was harvested after 6 days by spinning down cells at $500 \times g$ and harvesting the virus-containing supernatent. P1 virus was amplified to P2 virus, incubating 0.25 ml P1 virus with 5×10^6 *Sf*9 cells in 10 ml Sf-900™ II SFM in a T75 flask (Corning Inc., USA) in static culture at 26°C; P2 to P3 virus by incubating 2 ml P2 virus with 2.5×10^8 *Sf*9 cells in suspension in 500 ml Sf-900™ II SFM in a 2 L flask (Corning Inc. USA) at 26°C, shaking at 150 rpm.

2.2.2.4 Small scale expression trials

Constructs were trialled for expression in *Sf*9 cells as P1 virus, infecting 5×10^4 *Sf*9 cells in a 96 well plate with 50 µl P1 virus. Protein was left to express at 26°C for 3 days, then the supernatant and the whole cell sample (resuspended in PBS) were taken for Western Blot, using an anti-His antibody to detect the amount of protein that had been expressed

with a histidine tag (Section 2.3.2).

2.2.2.5 High Five infection and expression

High Five cells at 1×10^6 cells/ml were infected with P3 virus, 50 ml P3 stock for 500 ml cells, at a multiplicity of around 5 (virus particles:insect cells). Infected cells were then incubated in 2 L roller bottles (Corning Inc., USA) at 26°C, shaking at 150 rpm. Cells were harvested after 72 hours, with cells spun down at $4{,}000\times g$ for 15 minutes. The supernatent fraction, containing secreted recombinant protein, was filtered through a 0.22 µm filter (Stericup® , Merck) before subsequent processing.

2.2.3 S2 expression

The Schneider 2 cell line from *Drosophila melanogaster* has been used for around 20 years for recombinant protein expression. Compared to Lepidopteran cells, such as High Five cells, S2 cells can grow at much higher densities, allowing for more protein expression in a given volume of media. The ExpreS^{2}ion S2 system uses a proprietary lipofection-based transfection reagent (ExpreS^2 Insect-TR, ExpreS^2 Biotechnologies) to transfect the ExpreS^2 vector, pExpreS2-1, containing the constructs into S2 cells. The ExpreS^2 vector contains a Zeocin resistance gene to facilitate antibiotic selection using Zeocin in both *E. coli* cells, used to subclone inserts into the vector, and S2 cells. This vector is designed to integrate into the S2 genome, then constitutively express recombinant protein by linking the gene to an actin promoter (Figure 2.3).

2.2.3.1 Maintenance of S2 cells

S2 cell stocks were grown in Ex-Cell420 media (Sigma-Aldrich); 10% (v/v) foetal bovine serum; 10% penicillin-streptomycin mixture (Sigma-Aldrich) in Erlenmeyer flasks at 25°C, shaking at 110 rpm. Cells split twice per week; once per week by dilution to 5×10^6 cells/ml; once per week by resuspension in fresh media, centrifuging the cells at $1000 \times g$ for 5 minutes, then resuspending the cells in fresh media at 5×10^6 cells/ml.

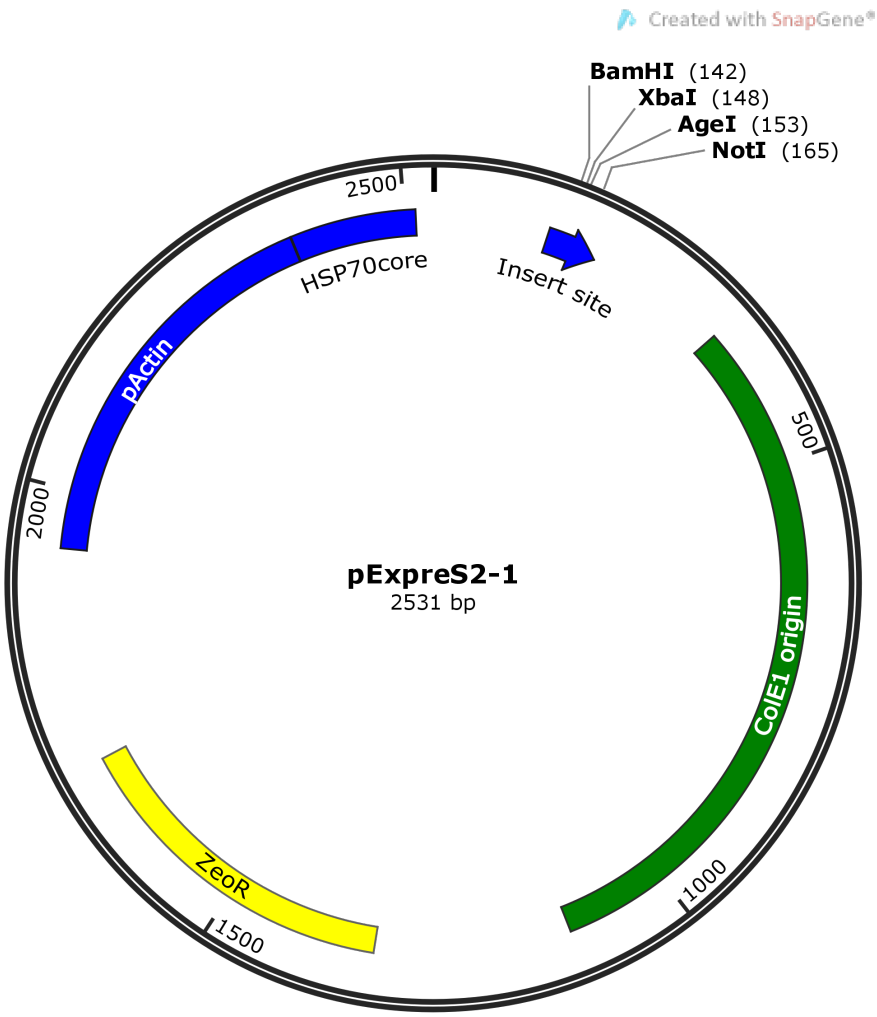


Figure 2.3: pExpreS2-1 vector map showing the nucleotide position of various features.

2.2.3.2 S2 transfection

1×10^6 S2 cells were suspended in Ex-Cell420 media at a density of 2×10^6 cells/ml. 50 µl were added to the cell suspension, followed by 12.5 µg of vector DNA. After inversion of solution, contents were transferred to a T25 flask (Corning Inc.). Foetal bovine serum was added to a concentration of 10% (v/v) after 4 hours. Zeocin was added after 1 day, to select for successfully transfected cells. These transfections were incubated for three weeks for selection, replacing the Zeocin-containing media twice per week. After three weeks of selection, cells were analysed for protein expression.

2.2.3.3 S2 large scale expression

Large scale S2 cultures were grown at 25°C in 2 L cylindrical flasks (Corning), with 500 ml cells in each flask. At the penultimate split, the media was changed to Ex-Cell420 without foetal calf serum. The ultimate split was by dilution, harvesting four days later. Cells were spun down at 4000 × g for 15 minutes, with the supernatent filtered through a 0.22 µm filter before subsequent purification.

2.3 Protein analysis

2.3.1 Polyacrylamide gel electrophoresis

Samples were run on a biphasic sodium-dodecyl sulfate polyacrylamide gel (Bio-Rad Laboratories Inc., USA) to assess purity of sample. Samples were boiled in loading buffer (0.02 M Tris-HCl, pH 8; 10% SDS (w/v); 20% glycerol; 0.05% bromophenol blue) to denature proteins and coat them in the negative-charged SDS. The boiled samples were loaded into sample lanes of the gel, then run at 180 mV, first through a 4% polyacrylamide stacking gel to concentrate samples, then through a resolving gel (between 10% and 20% polyacrylamide depending on desired resolution range), to separate proteins by size. Gels were stained with a fast, Coomassie-based stain (InstantBlue™, Expedeon, UK), with band size compared to markers (PageRuler prestained protein marker, Thermo-Fisher).

2.3.2 Western blot

Proteins were detected using a Western blot, using the horseradish peroxidase system. Samples were first run on an SDS biphasic polyacrylamide gel at 180 mV for 1 hour (Section 2.3.1). The gel was then blotted onto an Immobilon® PVDF transfer membrane (Merck Millipore) using a semi-dry Western blotting system (Bio-Rad Laboratories Inc.) for 1 hour at 50 mA. The membrane was then washed in Tris buffered saline (TBS; 20 mM Tris-HCl pH 8; 500 mM NaCl), then blocked for 1 hour in 10 ml Pierce protein-free (TBS) blocking buffer (Life Technologies, UK). After further washing with TBS; 0.2% Tween-20; 0.05% TRITON X-100, horseradish peroxidase (HRP)-conjugated antibody was added at a dilution of 1 in 10,000 in 10 ml blocking buffer and incubated for another hour. Membranes treated with HRP-conjugated antibody were developed using developer reagents (GE Healthcare), exposing the membrane to Amersham Hyperfilm ECL (GE Healthcare) for between 3 and 10 seconds, before developing on an automatic film processor (Eastman Kodak, USA).

2.4 Protein purification

2.4.1 Sonication

To lyse *E. coli* cells (Section 2.2.1.3), pellets were resuspended in 20 mM Tris-HCl pH 8; 300 mM NaCl; 15 mM imidazole; 0.05% TRITON X-100; 10% glycerol (lysis buffer). Bacterial pellets were lysed using a Sonicator S-4000 (Misonix, USA) at a power of 60 W, 3 s on; 7 s off, until 10,000 Joules were passed through the pellets. The lysate was then spun down at 50,000$\times g$ for 20 minutes.

2.4.2 Tangential flow filtration

Most eukaryotic media formulations contain additives that can prevent protein binding to Ni-NTA. To efficiently remove these components, a Centramate™ tangential flow filtration (TFF) system (Pall Corp., USA) was used to change the media from eukaryotic systems

(Sections 2.2.2.5, 2.2.3.3) for a suitable buffer for further purification, whilst concentrating the sample from litres to 150 ml. Sample was loaded onto the TFF system, and washed with at least 1 L of buffer, before elution. EPCR was exchanged into 20 mM Tris pH 8; 500 mM NaCl, while other proteins were exchanged into 20 mM Tris pH 8; 300 mM NaCl; 20 mM imidazole.

2.4.3 Nickel-NTA affinity purification

6-His tagged proteins were purified from soluble cell lysate (the supernatent from Section 2.4.1) using Nickel-NTA beads (QIAGEN). Beads were equilibrated with lysis buffer. Soluble cell lysate was incubated with the Ni-NTA beads for 10 minutes on a gravity flow column, before letting the lysate run through. Protein, bound to the beads, was washed with 5 column volumes of 20 mM Tris-HCl pH 8; 500 mM NaCl; 15 mM imidazole, then with 10 column volumes of 20 mM Tris-HCl pH 8; 300 mM NaCl; 15 mM imidazole. Protein was eluted from these beads in five 10 ml fractions in 20 mM Tris-HCl pH 8; 300 mM NaCl; 200 mM imidazole.

To purify His-tagged proteins from tangential flow filtration eluant (Section 2.4.2), 3 ml Ni-NTA beads were washed into buffer used in tangential flow filtration, then incubated with the eluant overnight at 4°C. These beads were then loaded onto a gravity flow column, proceeding as above.

2.4.4 Insoluble preparations

To purify protein from insoluble inclusion bodies, the lysate pellet was used (Section 2.4.1), discarding the soluble lysate. The membrane fraction was removed from the inclusion bodies by washing the pellet with 1 ml lysis buffer (Section 2.4.1 for formulation). The inclusion bodies were then dissolved in 6 M guanidine-hydrochloride; 20 mM Tris pH 8; 300 mM NaCl; 15 mM imidazole; 10 mM β-mercaptoethanol, aided by Dounce homogenisation. This mixture was left rolling at 33 rpm on a SRT6 tube roller (Bibby Scientific Ltd, UK) for 15 h, before centrifugation at $40,000 \times g$ for 20 minutes.

The supernatant from this centrifugation was bound to pre-equilibrated Ni-NTA beads on a gravity flow column. Bound protein was washed with 6 M guanidine-hydrochloride; 20 mM Tris pH 8; 300 mM NaCl; 15 mM imidazole with the addition of a glutathione redox buffer (300 µM oxidised glutathione; 3 mM reduced glutathione, Sigma-Aldrich) to allow for disulfide shuffling as the chaotropic agent was removed. After this initial wash step, guanidine hydrochloride was removed using a step-wise gradient in steps of 10%. At each step, 10 colume volumes of buffer was added, including an incubation of 15 minutes. Once the guanidine hydrochloride had been removed in this fashion, protein was eluted in five 10 ml fractions in 20 mM Tris-HCl pH 8; 300 mM NaCl; 200 mM imidazole.

2.4.5 TEV protease cleavage

His-tagged proteins were buffer exchanged into phosphate buffered saline (ThermoFisher Scientific); 300 µM oxidised glutathione; 3 mM reduced glutathione (Sigma-Aldrich). TEV protease (laboratory construct, including a histidine tag) was then added at 1 mg per 10 ml, incubating the reaction for 15 h at room temperature. The His-tagged TEV was subsequently separated from the cleaved protein using a Ni-NTA column, collecting the cleaved product in the flow through.

2.4.5.1 TEV protease purification

TEV protease was grown from a laboratory construct in Origami B cells (Merck Millipore). 2 litres of cells were grown as in Section 2.2.1.3, inducing at an OD_{280nm} of 0.8, and expressing protein for 3 h at 18°C. Protein was purified from the soluble fraction using Ni-NTA affinity purification (Section 2.4.3), dialysing the elution fractions into PBS; 300 µM oxidised glutathione; 3 mM reduced glutathione.

2.4.6 Protein deglycosylation

Proteins expressed in eukaryotic systems were deglycosylated using a mixture of endoHf (NEB) and endoF3 (kindly provide by Chan Li, Higgins lab) at a 50:1 protein:enzyme ratio

in 50 mM MES pH 6.5, 150 mM NaCl, incubating the mixture for 15 h at 25 °C.

2.4.7 Protein biotinylation

Biotin acceptor peptide (BAP) tagged proteins were biotinylated using the enzyme BirA. 20 µg of BirA (Sigma-Aldrich) was added to 1 mg of protein at 30 µM with 0.3 mM biotin and 5 mM ATP. This was incubated at 25°C for 15 h, then buffer-exchanged using PD10 column (GE Healthcare, USA) into PBS.

2.4.8 Protein fluorescent labelling

Protein was labelled using AlexaFluor dyes (488 nm ThermoFisher Scientific), via a TFP-ester, coupling to primary amines in the protein. 0.5 mg of protein was labelled at 2 mg/ml, incubating the protein with the dye for 15 h at 4°C. The reaction was quenched by adding 1 mM ethanolamine, removing the unreacted fluorophore by dialysis into PBS.

2.4.9 Size exclusion chromatography

Size exclusion chromatography (SEC) was used as a final purification step to separate monodisperse protein from other contaminants, included aggregates. SEC carried out by gel filtration on either an ÄKTA pure or an ÄKTA purifier (both GE Healthcare, UK). Protein samples were concentrated to 2 ml, loaded into a sample loop, and run on an appropriate column: a HiLoad Superdex 75 16/60 (prep grade, GE Healthcare) for proteins under 70 kDa; a HiLoad Superdex 200 16/60 (prep grade, GE Healthcare) for proteins over 70 kDa. The elution volume of any peaks in Abs_{280nm} were compared to calibration values (provided by manufacturer) to assess the approximate molecular weight the species contained within the peak. 2 ml elution fractions were collected from the void volume and analysed by SDS-PAGE to confirm size and assess purity.

2.4.10 Polyclonal antibody purification

Antibodies were purified against a target protein using the columns from Section 2.4.10.1. To purify the antibodies, the column was first washed into PBS. Rat sera samples were then flown over the column five times, incubating for five minutes each time. The beads were washed with 5 column volumes of PBS, then antibody was eluted using 0.1 M glycine pH 3, immediately neutralising the eluant with 0.1 M Tris-HCl pH 8. Eluant was then dialysed into PBS overnight at 4°C.

2.4.10.1 Column preparation

0.5 mg of the target protein was covalently coupled to cyanogen bromide-activated-sepharose beads (Sigma-Aldrich) to purify antibodies from rat sera. Beads were rehydrated and washed for 30 minutes with 1 mM HCl, then washed first with 10 colume volumes of distilled water, then 5 column volumes of 100 mM $NaHCO_3$ pH 8.5; 500 mM NaCl (coupling buffer). 5 mg protein per ml gel were incubated with the beads for 2 hours at room temperature, mixing the solution end-over-end using a LD 79 test-tube-rotator (Labinco B.V., Netherlands). The beads were washed using coupling buffer, then blocked for 2 hours using 1 M ethanolamine pH 8 at room temperature, again mixing end-over-end. The blocked beads the underwent 5 wash cycles of 10 column volumes of coupling buffer followed by 10 column volumes of 0.1 M acetate buffer pH 4; 500 mM NaCl to get rid of the blocking solution and any residual unreacted protein, before storing the column in 1 M NaCl at 4°C.

2.5 Biophysical Methods

2.5.1 Analytical gel filtration

Analytical size exclusion chromatography (SEC) is a quick method to assess the shape-dependent size of species in a mixture. Analytical SEC experiments were run on an ÄKTA-purifier (GE Healthcare) in 20 mM HEPES pH 7.5, 150 mM NaCl unless stated otherwise.

100 µl of protein at around 1 Abs_{280nm} was loaded into a 100 µl loop and injected onto a Superdex 75 10/30 column (analytical grade, GE Healthcare), a Superdex 200 10/30 column (analytical grade, GE Healthcare). The elution volume of peaks in the Abs_{280nm} trace were compared to calibration values (supplied by the manufacturer) to estimate the molecular weight of each species.

2.5.2 Circular Dichroism

Circular Dichroism allows the estimation of secondary structure in a sample. In this technique, electomagnetic radiation is plane polarised, and shone through the sample. Different secondary structural elements and unfolded polypeptide absorb the L-polarised and R-polarised components of to different extents, due to the chirality of the molecules held in a particular secondary structure [260]. Analysis of the resulting spectra (usually taken between 190 nm and 260 nm) can lead to the identity of secondary structure content of the protein, either by eye, comparing to ideal spectra, or by input of the data into a program, such as K2D3 [261], that compares the spectrum to a reference set of proteins. The inclusion of a Peltier water bath allows the measurement of circular dichroism at different temperatures, allowing the study of protein stability.

Measurements were taken using a J-815 spectropolarimeter (JASCO, Japan), with attached Peltier water bath. Single circular dichroism spectra were measured between 190 nm and 260 nm at 25°C. For thermal melt experiments, CD spectra were measured between 200 nm and 250 nm at 2°C intervals between 20°C and 90°C. Protein samples were held in a quartz cuvette (Starna Scientific Ltd, UK) with proteins at 0.1 mg/ml in 10 mM phosphate buffer, pH 7.5; 150 mM NaF. 4 spectra were averaged per sample, subtracting the circular dichroism of the buffer. Measurements were discarded if the high tension of the solution was over 1000 V at 190 nm.

2.5.3 Surface Plasmon Resonance

Surface plasmon resonance allows for the calculation of K_D, $k_{association}$ and $k_{dissociation}$ for a given interaction. One component of the interaction is coupled to a gold-coated chip (component 1), whist the other is flown over in solution (component 2). The mass change on the chip is measured by observing the change in the critical angle where plasmons are produced. This angle is dependent on the local environment of the chip surface, so can detect when a species binds. Measurements are taken over a concentration range of component 2, analysing using the BIAevaluation software. Evaluation either measures bound concentration of component 2 at the steady state, for each concentration in the series to determine K_D, or fits the kinetic binding of component 2 at different concentrations to determine $k_{association}$ and $k_{dissociation}$, inferring K_D using the equation, $K_D = \frac{k_{dissociation}}{k_{association}}$.

Surface plasmon resonance was performed on a BIAcoreT200 (GE Healthcare). Two coupling methods were used during this thesis. EPCR was coupled to around 2000 response units (RU) via lysine residues to CM5 chips (GE Healthcare) according to manufacturer's protocol. A concentration series of the second component was flown over the flowpath in 10 mM HEPES pH 7.5; 150 mM NaCl; 0.005% Tween-20 at 30 µl/min, 240 s association, 300 s dissociation. The chip was regenerated between each sample using a 30 s pulse of 10 mM NaOH.

Biotinylated EPCR was coupled to CAP chips to around 200 RU using the Biotin CAPture kit (Series S, GE Healthcare), according to manufacturer's protocol. A concentration series of the second component was flown over the flowpath in 10 mM HEPES pH 7.5; 150 mM NaCl; 0.005% Tween-20; 1 mg/ml salmon sperm ssDNA (Panreac AppliChem, Germany) at 30 µl/min, 240 s association, 300 s dissociation. Regeneration of the chip surface was accomplished using a 120 s pulse of 6 M Guanidine, 0.25 M NaOH.

The ICAM-1:head structure and dual-binding experiments were run over a Protein A- or Protein G-immobilised CM5 chip (immobilised by Frank Lennartz) in 10 mM HEPES pH 8; 300 mM NaCl; 0.05% Tween-20. ICAM-1 was flown over the chip at 1 µM for 120 s at 3 µl/min. For the ICAM-1:head structure, a head structure dilution series was then flown over the chip for at 40 µl/min, 240 s association, 600 s dissociation. For the dual-binding

experiments, head structure was flown over at 250 nM, 20 µl/min, 240 s, then a dilution series of EPCR collected at 30 µl/min, 240 s association, 720 s dissociation. 10 mM Glycine, pH 2 was used as a regeneration solution

2.5.4 Analytical ultracentrifugation

Analytical ultracentrifugation was carried out using a Beckman Optima XL-1 analytical ultracentrifuge (Beckman Coulter, USA). This technique facilitates the calculation of shape-dependent molecular weight.

2.5.4.1 Sedimentation velocity analytical ultracentrifugation

Sedimentation velocity AUC studies how proteins travel down the length of a centrifuge cell. by observing the concentration of sedimenting boundaries over time we can calculate the sedimentation and diffusion coefficients by solving the Lamm equation (Equation 2.1), where c is the concentration of species by weight; t is time; D is the diffusion coefficient; r is the distance from the centre of the rotor; s is the sedimentation coefficient and ω is the rotor speed. This allows the identification of D and s, leading to the calculation of the shape dependent molecular weight (the buoyant molecular weight), M_b, using the Svedberg equation (Equation 2.2), where R is the gas constant and T is the temperature [262].

$$\frac{\delta c}{\delta t} = D\big[\frac{\delta^2 c}{\delta r^2} + \frac{1}{r}\frac{\delta c}{\delta r}\big] - s\omega^2\big[r\frac{\delta c}{\delta r} + 2c\big] \tag{2.1}$$

$$\frac{s}{D} = \frac{M_b}{RT} \tag{2.2}$$

Protein samples at 0.5 Abs_{280nm} were loaded into sapphire cells, then spun at 40,000 rpm. The interference caused by the boundary between solute and depleted sample was recorded every 4 minutes. These data were plotted in sedfit [263], and fit using a continuous sedimentation coefficient distribution model, defining the initial meniscus of the solution

and bottom of the cell, allowing the frictional ratio to vary in solving the Lamm equation.

2.5.4.2 Equilibrium analytical ultracentrifugation

Equilibrium analytical ultracentrifugation looks at the sedimentation position of a particle in the cell. For a single species in solution, this allows the solution of the Svedberg equation (Equation 2.3) [262].

$$c(r) = c_0 \exp\left[\frac{M_b\omega^2}{RT}\left(\frac{r^2 - r_0^2}{2}\right)\right] \tag{2.3}$$

Experiments were run until equilibrium at 20°C, with protein at 8 µM. Absorbance data were taken at a wavelength of 280 nm at equilibrium at 10,000, 12,000, 14,000 and 16,000 rpm analysed the data using sedphat [264], globally calculating the molecular weight modelling a single species of an interacting system.

2.5.5 Isothermal titration calorimetry

Isothermal titration calorimetry is a very accurate way of calculating the thermodynamic contributions to an interaction between two species. One component, initially in a syringe is titrated into a second component in a cell. This titration is accompanied by a heat change, produced due to the proteins interacting, that is measured by comparison to a reference cell. The heat change is monitored over the course of the experiment, until all of the cell species is bound. This allows the calculation of K_a, the thermodynamic association constant (and n, interaction stoichiometry), which in turn allows the calcultion of ΔG, the Gibbs free energy change. ΔH, the enthalpy change, can be calculated from the heat change, and this allows the inference of ΔS, the entropy change [265].

Isothermal titration calorimetry (ITC) measurements were carried out on a MicroCal iTC200 System (GE Healthcare). Samples were dialysed for 15 h into 20 mM HEPES pH 7.5; 150 mM NaCl at 4°C. Experiments were performed at 25°C with the syringe component (60 µl) at ten times higher in concentration than the cell component (300

µl). Concentrations of both components were checked using a BCA assay kit, according to manufacturer's protocol (Merck Millipore). Data were checked for outliers and subsequently integrated and fit by nonlinear least-squares fitting using Origin ITC Software (GE Healthcare).

2.5.6 Size exclusion chromatography multi-angle laser light scattering

Size exclusion chromatography can be linked to multi-angle laser light scattering (SEC-MALLS) to allow the shape-independent determination of species separated by SEC. Monochromatic light is shone through the eluting sample, collecting the intensity of scattered light at multiple angles. The molecular weight of the sample then can be calculated from the Zimm equation (Equation 2.4), where K is a constant for light scattering, dependent on the refractive index of the solution; c_X is the concentration of sample X; R_c is the Rayleigh scattering at angle θ at concentration c; M_X is the molecular weight of sample X; A_2 is the second virial coefficient, a correction constant for concentrated particle-particle interactions. Simultaneous equations at multiple angles can solve this equation for A_2 and the molecular weight, M_X [266].

$$\frac{Kc_X}{R_c(\theta)} = \frac{1}{M_X} + 2A_2c_X \tag{2.4}$$

Protein samples, at 0.1 mg/ml, were purified by size exclusion chromatography on a Superdex 75 10/30 column (analytical grade, GE Healthcare) in 20 mM HEPES pH 7.5; 150 mM NaCl, then analysed using laser light scattering, detected at 662 nm wavelength at 8 scattering angles between 20.6°and 149.1°using a Heleos 8 instrument (Wyatt Technology). Molecular weights were calculated using the Zimm equation by ASTRA 6.1 (Wyatt Technology).

2.5.7 Small angle X-ray scattering

Small angle X-ray scattering (SAXS) is a solution scattering technique that gives information about particle size and shape beyond techniques such as AUC and SEC-MALLS. The resolution obtained from SAXS experiments is substantially lower than that from X-ray crystallography, but SAXS can be easier to perform, as it is a solution method.

One dimensional scattering data were collected on beamline P12 at DESY (Hamburg, Germany) at a wavelength of 1.24 Ånm using a Pilatus 2M pixel X-ray detector, or at BM29 at the ESRF (Grenoble, France) at a wavelength of 0.99 /AA using a Pilatus 1M pixel X-ray detector. For each protein sample, a series of data sets were collected for a 2-fold dilution series from 4 mg/ml to 0.125 mg/ml, with buffer measurements taken in between each sample measurement. All samples were in 20 mM HEPES pH 7.5; 150 mM NaCl, with complexes mixed on the beamline unless stated otherwise. Samples were filter directly before measurement.

Concentration series for each sample, after normalisation and buffer subtraction, were extrapolated to zero to check for concentration dependent aggregation in Primus [267]. The radius of gyration, R_g, and zero angle scattering, $I(0)$, were calculated using AutoRg. Data points from the lowest scattering angle q, as well as at those at higher q with a low signal to noise ratio, were removed, depending on AutoRg results and Kratky plots. Inverse Fourier calculations of $I(q)$ were used to yield $P(r)$ functions, $I(0)$, R_g and the maximum dimension (D_{max}) using GNOM [268].

Initial bead models for these data were created *ab initio* using DAMMIF [269]. These were averaged using DAMAVER [270], and refined from the averaged model, based on the original $P(r)$ function, using DAMMIN [271]. The resulting model was converted into an envelope using Situs [272], before docking crystal structures into the envelope using Sculptor [273].

2.6 Assays

2.6.1 Enzyme-linked immunosorbent assays

100 ng of protein was coated on Nunc Maxisorp® 96 well clear plates at a concentration of 2 µg/ml in PBS, incubating for 15 h at 25°C. Plates were then washed 3 times with 150 µl PBS; 0.05% Tween-20 (PBS/Tween), before blocking with Pierce protein-free (TBS) blocking buffer (Life Technologies, UK) for 1 h. After six more washes with PBS/Tween, antibody samples were applied for 1 h. Plates were washed six times with PBS/Tween, and then probed with an anti-IgG antibody conjugated to horseradish peroxidase (Sigma). Unbound antibody was washed away (six washes with PBS/Tween) after a one hour incubation, with bound antibody detected using TMB substrate (Sigma-Aldrich), incubating for 15 minutes, and quenching with 2 M sulfuric acid. Absorbance was detected using a SpectraMax M3 plate reader (Molecular Devices, UK) at 450 nm.

2.6.2 Inhibition assays

Nunc Maxisorp® 96 well clear plates were coated with 100 ng streptavidin/well at a concentration of 2 µg/ml in PBS, incubating for 15 h at 25°C. Plates were then washed 3 times with 150 µl PBS; 0.05% Tween-20, before blocking with Pierce protein-free (TBS) blocking buffer (Life Technologies). After six more washes with PBS/Tween, 200 ng EPCR was incubated at 2 µg/ml in each well in PBS for 2 h. During this period, 200 ng of fluorescent CIDR (Section 2.4.8) at 4 µg/ml was incubated with 50 µl of antibody solution for 1 h. Plates were washed six times after the EPCR incubation with PBS/Tween, then incubated with the antibody/CIDR solution for a further hour. The plate was washed six times in PBS/Tween, then 100 µl PBS was added. Fluorescence was measured using a SpectraMax M3 plate reader, with an excitation wavelength of 494 nm and an emission wavelength of 519 nm; medium photon multiplier tube (PMT) sensitivity; 20 readings per well.

2.7 Crystallographic Methods

2.7.1 Protein crystallisation

Since the first protein crystal structure, of myoglobin, was solved in 1958 [274], X-ray crystallography has become a key technique in the field of structural biology, used to get high resolution structures of proteins to gain insight into their function. The first step is to crystallise the protein into an ordered lattice that will ensure a clean diffraction pattern to high resolution. This is general accomplished by vapour diffusion, where a protein is slowly concentrated in a sitting or hanging drop as water diffuses from this drop to equilibrate with a well solution. As the protein concentrates towards the saturation point, ideally nucleation events occurs, followed by crystal growth as the protein concentration drops (Figure 2.4). The precise conditions for this to occur depends on the protein, with temperature, crystallisation agent, additives and initial concentration all factors in whether a protein will crystallise or precipitate.

Purified protein was concentrated to 10 mg/ml, and centrifuged at $14,000 \times g$ to remove aggregates. Sitting-drop broad screens were set up in MRC 2-well crystallisation plates (Swissci, Hampton research, USA) to maximise the breadth of crystallisation conditions initially trialled. 50 µl of each commercial condition was dispensed into each well, with 100 nl of the condition being mixed with 100 nl of protein solution by a mosquito® Crystal (TTP Labtech, UK). Typically, JCSG+, Pact Premier, Morpheus and MIDAS (all Molecular Dimensions, UK) were initially used, with ProPlex (Molecular Dimensions) added for complexes. These broad screens offer a range of salt, precipitant and buffer conditions and concentrations in which the protein may crystallise.

Any condition which produced crystals was optimised. There were a number of ways to empirically optimise a crystal condition, depending on the initial crystal obtained. Typically, a condition was optimised using the Silver Bullets additive screen (50 nl, Molecular Dimensions), added to the drop in a new 96 well plate. This set of 96 small molecules and additives have been empirically chosen, based on previous success. Fine screens were set up in some cases, varying the concentrations and pH of components in the original

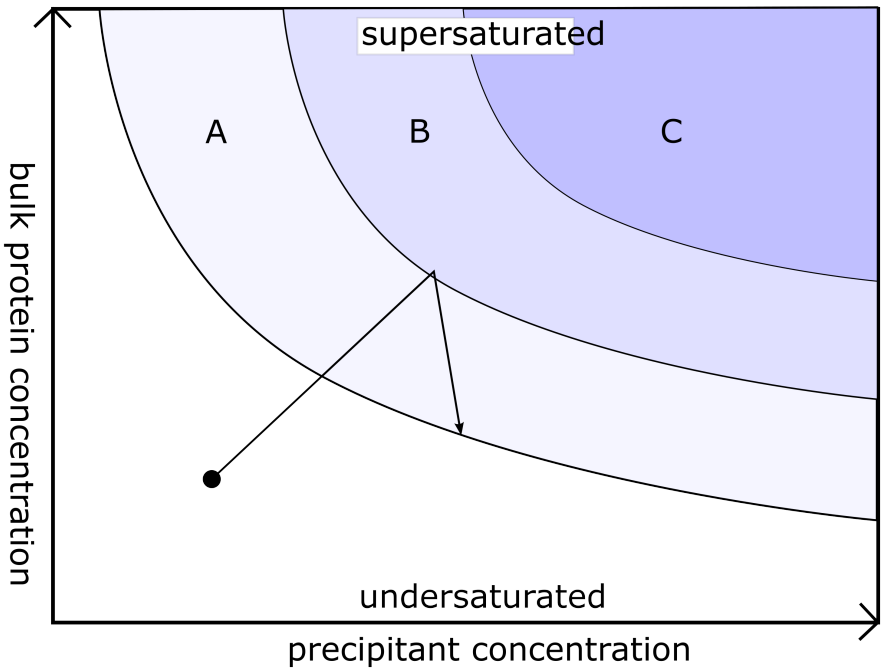


Figure 2.4: A crystallisation phase diagram showing the ideal progression leading to crystal growth. **A** represents the labile phase where crystal growth can occur, **B** represents the nucleation phase where protein nucleates and **C** represents precipitation. The protein starts in an undersaturated solution (circle), and increases in concentration as water diffuses out of the drop. As the protein and precipitant concentration increase, the solution becomes supersaturated causing nucleation and crystal growth to occur. This decreases the bulk concentration, while the crystal grows, until the solution is undersaturated once more.

condition to try and improve crystal morphology. Seeding of crystals was trialled in cases of overnucleation. Providing micronuclei can allow protein to crystallise at lower concentrations instead of nucleating [275]. Micronuclei were made either through the use of a Seed-Bead (according to manufacturer's protocols; Hampton Research, USA), or by using a 200 µl pipette to break down crystals in the drop.

2.7.2 Crystal cryocooling

It is necessary to protect protein crystals as the doses of X-ray radiation used to collect a dataset can damage the protein in the crystal, reducing diffraction at high resolution and potentially introducing structural artifacts in the protein. To protect the crystal, data is routinely collected at very low temperatures (100 K) to slow the rate of damage. This required the crystals to be cryocooled on harvesting for storage using liquid nitrogen. Crystals were looped from their drops, into the same crystallisation condition, adding a cryoprotectant to prevent ice formation[1]. This protectant, usually 25% glycerol, ethylene glycol or 3-methyl-pentan-1,5-diol (MPD), causes water to form a glass upon cooling, rather than freezing into ice, which could disrupt the crystal lattice. Once in the cryoprotected condition, the crystal can then be frozen in liquid nitrogen and stored in a cooled Dewar.

2.7.3 Data collection

Diffraction data for this thesis were collected at the Diamond Light Source, UK. Loops, containing crystals were loaded onto a goniometer, continuously cooled by a stream of gaseous nitrogen (100 K). X-rays with wavelengths of around 0.9 - 1 Å were shone at the crystals to generate a diffraction pattern. Data were initially collected on a Pilatus detector for 3 angles (0°, 45°and 90°) to assess the diffraction of the crystal and to identify the space group. Quality was assessed firstly by visual inspection of the diffraction pattern, looking at spot quality and the resolution to which the crystal diffracted. Autoprocessing by FastDP [276] and xia2 [277] allowed the determination of the crystallographic spacegroup and the estimated resolution limit. Crystals which diffracted with clean, sharp spots to a resolution

[1]The exception to this is crystals from the Morpheus screen, which includes cryoprotectant in the well solution.

below 4 Ångströms were subjected to comprehensive data collection, using EDNA [278] and Mosflm [279] as a basis for a collection strategy. This strategy calculates the ideal way to collect all unique reflections based the spacegroup of the crystal. Data were collected by multiple 0.1° oscillations, collecting all reflections over a wedge defined by the strategy.

2.7.4 Data indexing and refinement

Complete datasets were initially processed in iMosflm [279]. Using a small subset of data (typically 3 images), this program first finds all spots, then indexes found spots, calculating the crystal lattice and unit cell dimensions. iMosflm also gives an estimate of the space group, based on the unit cell dimensions. The unit cell is refined by comparing sets of images from different wedges around the crystal. All images within the dataset were then integrated, to find the intensity for each *hkl* coordinate. At this point the resolution cut off was determined by looking at the signal:noise ratio, $I/\sigma(I)$, with a cut off value of 2. Pointless [280] was then used to unambiguously calculate the space group, by comparing symmetry-related intensities. Checking for systematic absences allows the distinction between enantiomorphic spacegroups.

Data from the set of images were then scaled and merged to form one complete dataset. This was accomplished by the program, Scala [281]. This program scales and merges datasets from different images to produce one complete *hkl* dataset with intensities, but without phase data.

At this point data were checked for quality using R_{merge}, which is a measure of internal consistency within the dataset. This quality control compares the spread of n independent measurements of the intensity at each spot at reciprocal coordinates *hkl* around their average within a given shell (Equation 2.5). Typically resolution shells are cut off if $R_{merge} > 0.6$, though in outer shells and at higher multiplicity, this cut off is extended to around 0.8. At this point 5% of the refections are set aside for later calculation of R_{free}

$$R_{merge} = \frac{\sum_{hkl} \sum_{i=1}^{n} |I_i(hkl) - \overline{I}(hkl)|}{\sum_{hkl} \sum_{i=1}^{n} I_i(hkl)} \tag{2.5}$$

2.7.5 Molecular replacement

Electron density is calculated by summing the structure factors and phases of all diffractions points in reciprocal space (Equation 2.6). Collecting X-ray diffraction data gives us the reciprocal diffraction pattern with the intensities at each point. However phase information, ϕ, is not detected; it must be obtained through other means.

$$\rho(xyz) = \frac{1}{V}\sum_{hkl}|F(hkl)|e^{-2\pi i[hx+ky+lz-\phi(hkl)]} \tag{2.6}$$

Strategies using anomalous diffraction, or heavy metal ions are often used to provide phase information. However it is possible to use phase information from solved structures as a starting point. This is followed by iterative cycles of rebuilding and refinement to improve the phase data.

Molecular replacement was achieved using Phaser-MR [282]. Instead of using the traditional combination of Patterson functions and least-squares analysis used by earlier programs to search reciprocal space [283, 284], Phaser uses maximal likelihood methods in translation and rotation functions [285], combining it with multivariate statistical analyses to find possible solutions.

Scaled and merged data were input into Phaser-MR with polypeptide models from the PDB pruned of any non-protein moeities. In addition to this the spacegroup, identified by Pointless, and the solvent content, calculated by Matthews_coef, were supplied. Output solutions were examined by their log likelihood gain (LLG) value. This value measures the likelihood of the data, given the input model, compared to a model of random atoms. Z-scores were also calculated, comparing LLG values calculated with the LLG from a random set of rotation and translation functions, judged as the number of standard deviations the solution LLG is above or below the mean value of the random set. The best solution was taken on to model rebuilding and refinement.

2.7.6 Model rebuilding and refinement

The solution from molecular replacement contains extra electron density compared to the electron density calculated using the molecular replacement model. This can be visualised in Coot [286]. Coot loads two electron density maps; the $2F_{obs} - F_{calc}$ map allows a good visualisation of all of the observed electron density from the data, while the $F_{obs} - F_{calc}$ shows the difference (positive or negative) between the electron density from the data and the electron density calculated from the model. Electron density not accounted for by the model can then be built into, either manually or by using the program Buccaneer [287]. This program builds polyalanine chain into unmodelled electron density, inserting idealised secondary structural features where appropriate.

Rounds of rebuilding were interspersed with refinement steps. Refinement was accomplished by two programs, refmac [288] and autoBUSTER [289]. Both of these programs use maximal likelihood methods to refine the atomic model, minimising the difference between the experimental data and the calculated data from the model. The programs then output the refined model, and the electron density maps (observed and calculated) in the light of the refined model. autoBUSTER was used in particular to elucidate possible loop density, as it contains algorithms to model missing structure. Output density from autoBUSTER hence reveals more extra density which can be modelled into.

The quality of each refinement step was assessed primarily using R factors (R_{work} and R_{free}), which should drop during each refinement step. These R factors assess the difference between the observed and calculated structure factors using equation 2.7. R_{work} is calculated from the working set of reflections, whilst R_{free} uses a small subset of reflections (around 5%) not used in refinement to calculate the R-factor.

$$R = \frac{\sum_{hkl} |F_{hkl}^{obs} - F_{hkl}^{calc}|}{\sum_{hkl} F_{hkl}^{obs}} \tag{2.7}$$

Successive refinement and rebuilding allow the construction of a good model, which explains the observed data. The final R_{work} is generally between 0.15 and 0.25 at a resolution of 2.5 Å, signifying a model that fits the observed data well [290]. The final R_{free} value should

be at most 5% above R_{work}, to show that the model does not overinterpret the data [290].

2.8 Immunogen design using Rosetta

The Rosetta software suite has made a great deal of progress in the last twenty years. Its algorithm revolves around the observation that short segments of three amino acids can be found in the same set of conformations in different structures [291]. The current iterations of Rosetta constrain the conformation of the input polypeptide using the conformations of overlapping three and nine amino acid segments. This is combined with the Rosetta energy function, score12, which describes the inter- and intra-molecular forces in the system, to drive the input polypeptide conformation to its lowest energy conformation. Additional modifications and programs allow Rosetta to perform a wide variety of functions including epitope grafting.

2.8.1 RosettaScore

RosettaScore was used to extract various information from Rosetta's silent format, an output format which contains all structures in a compressed format. This program allows the extraction of individual models from this file as pdb files. Each model is also scored using Rosetta's score function.

2.8.2 Epitope-scaffold folding

A script, available in the Rosetta package, was used with Rosetta's structure prediction to fold the polypeptide chain around the epitope (Appendix B). Rosetta treats the epitope as a rigid body and folds the rest of the sequence around this rigid body, using a 3-mer and 9-mer library of amino acid conformations to predict a coarse conformation. After this first step, a full atom relaxation step was performed using the Rosetta energy function. C_α-C_α scaffold constraints were generated using a custom Python script (Appendix B) and added to the run to guide polypeptide folding to the scaffold topology. 10,000 structures were

generated using this protocol in Rosetta's proprietary *silent* format.

2.8.3 Sequence design

The Rosetta scripting interface [292] was used to design a custom script to search for mutations to stabilise the epitope topology. This script mutates residues in the polypeptide sequence to alanine except designated residues. Random residues are then assigned identities based on residue position, in the hydrophobic core, at the surface or at the boundary. After each round of mutations the structure was allowed to move, then was scored using the packstat filter [293]. Four rounds of mutation were succeeded by a couple of filters, which picked structures with an improved score. 10,000 runs were performed using this script, with successful models output in the *silent* format.

Chapter 3

The structure of the CIDR:EPCR complex

3.1 Outline

Both the CIDR domains and EPCR needed to be cloned into appropriate overexpression systems to produce enough protein to study using biophysical and structural techniques. Bioinformatic analyses were used to identify appropriate domain boundaries for recombinant constructs. I selected sequence diverse CIDR domains identified in [140], cloning the genes into an *Escherichia coli* expression system. This system expressed recombinant protein in insoluble inclusion bodies. A Ni-NTA bound, step-wise gradient refolding technique allowed CIDR domains from these inclusion bodies to be successfully refolded. *Drosophila* Schneider 2 cells, an insect cell line, were used to express EPCR, with protein purified from the media by affinity purification on a Ni-NTA column, followed by gel filtration. Purification tags and glycans were successfully removed from proteins destined for crystallography. The set of CIDRs were tested for secondary structure using circular dichroism, and for functional activity against the recombinant EPCR using surface plasmon resonance. These experiments showed the proteins to interact, suggesting they were folded correctly. Crystal trials were set up with using a selection of CIDRs with and without EPCR, with two CIDR:EPCR complexes crystallising, producing crystals that diffracted to a resolution

better than 3 Ångströms. The two structures were solved, gaining phase information by molecular replacement using a previously solved EPCR crystal structure [227]. Analysis of these structures reveals the molecular interaction site of the CIDR:EPCR complex, and shows how CIDR domains inhibit the interaction between protein C and EPCR.

3.2 CIDR cloning and expression

3.2.1 CIDR construct design

Five different CIDRα1 domains from *Pf*EMP1 variants containing domain cassette 8 and 13 were selected for overexpression. These CIDR domains had been shown to bind EPCR previously [140]. HB3var03 contains a DC13 CIDR from the HB3 strain, thought to originate in Honduras [294]; IT4var06 (DC8), IT4var07 (DC13) and IT4var20 (DC8) are from IT4/FCR3 strain, representing South East Asia [295]; PFDb0020c (DC8) is from the 3D7 strain, which is thought to derive from sub-Saharan Africa [295, 296]. CIDR constructs were designed using secondary structural information from the two previously published CIDR structures, the Palo Alto 89F5varO (varO) NTSDBL-CIDR head structure [189] and MC179 [137]. These two crystal structures, as well as previous sequence comparisons suggested a conserved three helical bundle [170]. In addition to this, varO has an additional β-sheet portion at the N terminus of the CIDR, which looks to form part of the CIDR domain; the MC179 construct was truncated at the N-terminus, so the equivalent portion was not crystallised. FUGUE predicted similar secondary structure elements in the five selected CIDRα1 domains, including the same conserved helices in the other CIDR structure and the β-sheet segment in the var0 head structure (Figure 3.1). In addition, conserved cysteines suggest five disulfide bonds in the CIDR structures. On the basis of these analyses, I designed the CIDR constructs to start at the beginning of this β-sheet segment with a stop codon at the end of the last conserved helix.

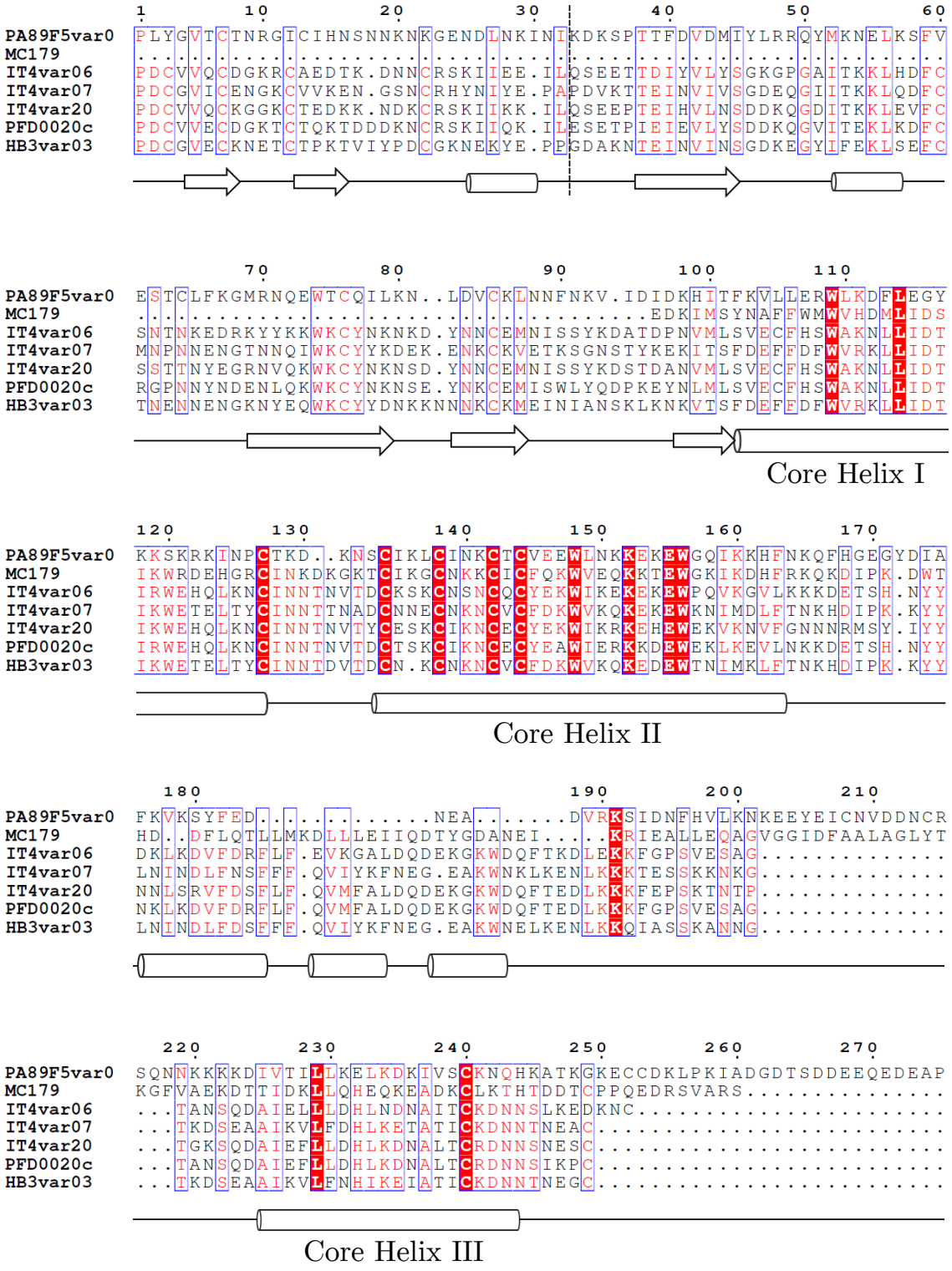


Figure 3.1: FUGUE secondary structure predictions for are shown below a sequence alignment between the five chosen CIDRs and the two CIDRs with solved structures. The three conserved helices are marked Core Helix I, II and III.

3.2.2 CIDR construct expression

Using these predicted domain boundaries, I cloned the five CIDR domains from their genomic DNA (kindly provided by the Lavstsen group at the University of Copenhagen) into the pEt15bcon vector backbone, using the primers specified in Appendix A. This vector contains the DNA sequence for an N-terminal His-tag for purification of the recombinant protein. Since the CIDR domains have multiple disulfide bonds predicted by comparison to previous structures, pFH225 and pJM59 strains *E. coli* were chosen to trial CIDR expression. These two strains, designed by the Oxford Protein Production Facility (OPPF) are based on the B834 strain, modified by the addition of vectors containing sulfhydryl oxidases to aid disulfide bond formation. Soluble and insoluble fractions from these small scale trials were analysed by Western blot, detecting using an anti-His antibody, conjugated to HRP (QIAGEN, The Netherlands). They revealed the optimum conditions for CIDR expression, summarised in the table below. These expression trials show that though only two CIDRs expressed protein in the soluble fraction, all CIDRs express at high levels in *E. coli* inclusion bodies.

CIDRα	Chosen expression conditions			
	E. coli strain	Temperature (°C)	Time (h)	solubility
IT4var06	pJM59	30	3	i
IT4var07	pFH225	30	3	s/i
IT4var20	pJM59	25	3	i
PFD0020c	pJM59	25	3	i
HB3var03	pFH225	30	3	s/i

Table 3.1: Summary table of optimal expression conditions for the CIDR domains. All CIDRs had insoluble expression (i), while only two had soluble expression(s).

3.2.3 Refolding and purification of CIDRs

Since a significant amount of all the CIDR constructs was present in the insoluble fraction, I purified the insoluble inclusion bodies from a 2 L culture of IT4var20CIDR to trial two refolding protocols for the CIDRs. These inclusion bodies were first broken apart using a

Dounce homogeniser, then solubilised using the chaotropic agent, guanidine hydrochloride, and the reducing agent β-mercaptoethanol, to reduce any disulfide bonds. These inclusion bodies were first tested by drop-wise dilution into a set of refolding buffers (AthenaES, USA). This collection of buffers contain various additives which have been empirically shown to aid protein refolding. Subsequent analysis of these trials revealed a low yield of soluble protein, with a lot of precipitate in the refolding solution. This suggests that fast dilution was not an appropriate method for refolding.

An alternative to fast dilution is on-column gradient refolding. Here, the unfolded CIDR domains are first bound to a nickel-NTA agarose beads in a column, then refolded using a stepwise decreasing gradient of guanidine hydrochloride in the presence of a glutathione redox buffer. This method allows a protein to refold slowly, while decreasing the potential for aggregation by binding the protein to the nickel-NTA agarose. This refolding procedure yielded better results, allowing for protein to be refolded into a soluble, monodisperse form, as shown by subsequent gel filtration (Figure 3.2). Yields varied from 50 µg/L for IT4var06CIDR, to 250 µg/L for the HB3var03CIDR. The yields of IT4var06CIDR and PFD0020cCIDR were particularly low, so these two constructs were discontinued.

3.3 EPCR cloning and expression

3.3.1 EPCR construct design

EPCR constructs were based on two different sets of construct boundaries. The first, termed EPCR long (EPCRL), matched the boundaries of the previously crystallised construct [227]. The second construct, termed EPCR short (EPCRS), was designed based on the resolved residues in the two crystal structures available (PDB code 1L8J, 1LQV). The previously solved crystal structure shows a lipid in the central cavity, and 3 glycan residues on the surface of the protein, which may be important for the structure of the protein. As the crystallised EPCR construct was previously expressed in the Chinese hamster ovary cells, I trialled these constructs in eukaryotic expression systems as well as *E. coli*.

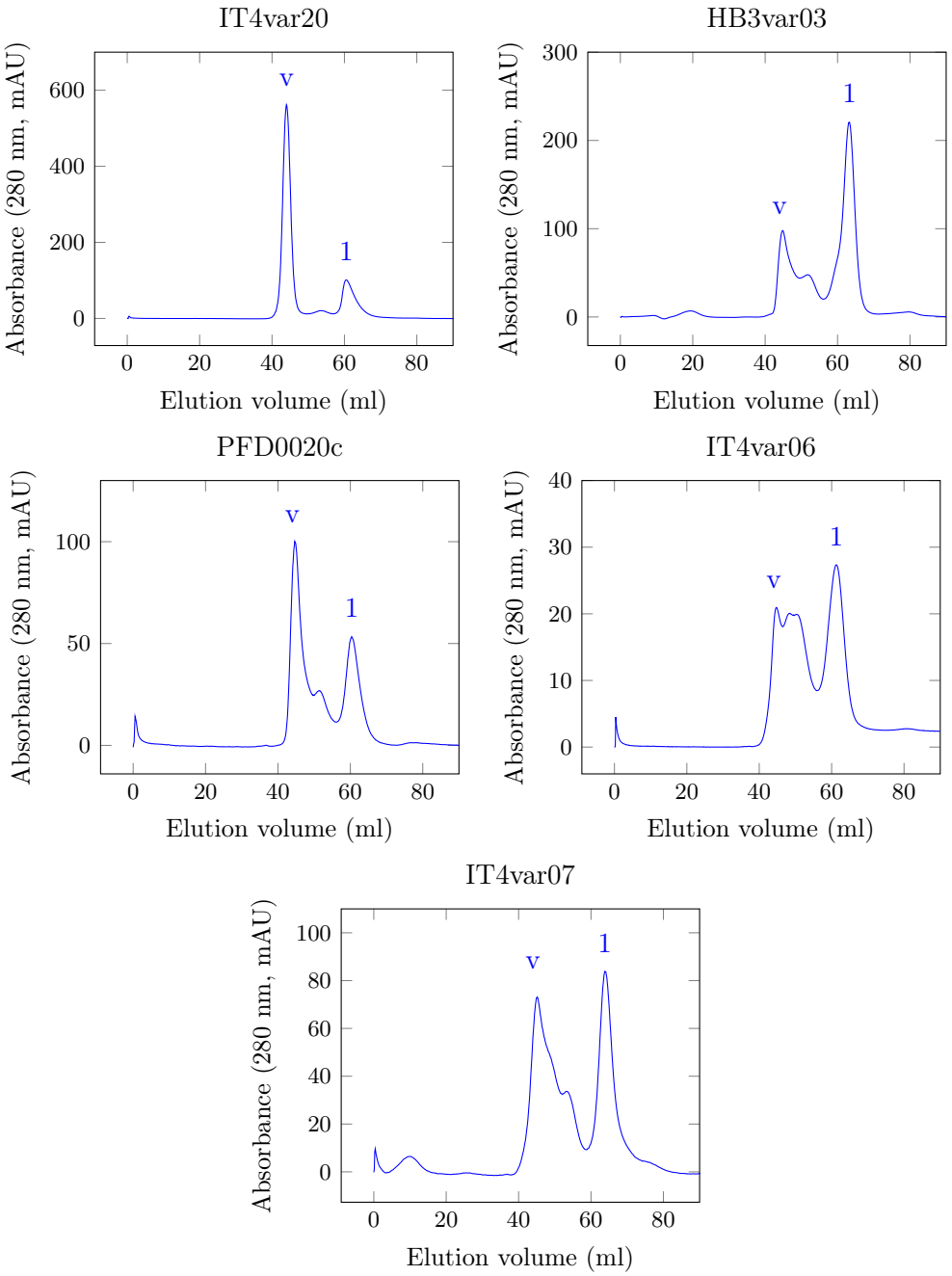


Figure 3.2: Preparative gel filtration traces for the five refolded CIDR domains. The peak corresponding to the monodisperse protein at the correct size is marked by a 1, with the void peak at 45 ml (v).

3.3.2 EPCR expression trials in *E. coli*

These constructs were initially cloned into pEt15bcon for *E. coli* using primer pairs in Appendix A. I tried to express both EPCRL and EPCRS in strains Origami B and Shuffle cells. Expression trial samples of insoluble and soluble fractions were probed via Western blot for the presence of the histidine-tagged EPCR. EPCR expressed in the insoluble fraction, with no detectable EPCR in the soluble fraction. One reason EPCR may not have folded is due to the absence of the lipid in the central cavity. I hence attempted to refold the protein from soluble inclusion bodies by drop-wise fast dilution into the set of refolding buffers (AthenaES), in the presence of exogenous phosphotidylethanolamine. Concentration of the refolded protein and subsequent gel filtration revealed very little soluble protein, signified by the small sample peak at 20 ml in Figure 3.3. As a result of this low yield I opted to trial a eukaryotic system.

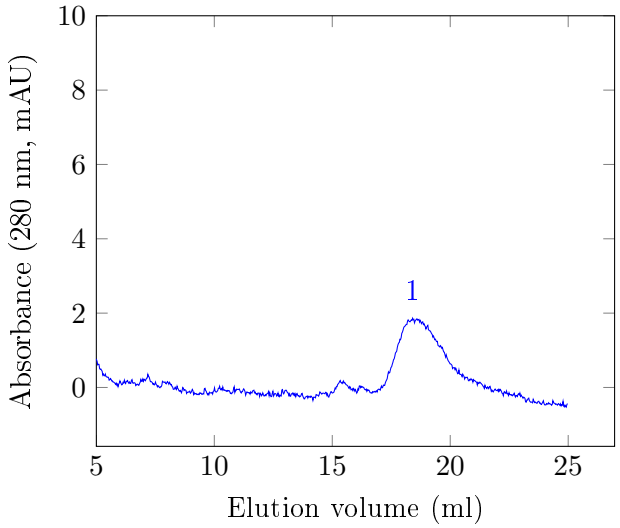


Figure 3.3: Analytical gel filtration trace for EPCR, expressed in *E. Coli* and refolded with exogenous lipid. The peak at the correct size for EPCR is marked with a 1.

3.3.3 EPCR expression in *Sf*9/baculovirus expression system

Eukaryotic cell expression systems have the requisite machinery to modify proteins post-translation. This would facilitate the addition of lipid and glycans to the sites identified in the previous EPCR crystal structures, and the formation of the disulfide bond in EPCR,

in the ER of the cell [297]. I cloned both EPCRS and EPCRL into the pAcGP67A vector backbone using primer pairs in Appendix A, converted to add an N-terminal 6-histidine tag and TEV protease cleavage site to the recombinant protein (pAcGP67Acon, created by Prof. Higgins). This vector was co-transfected with linearised baculovirus DNA into *Sf*9 cells. Virus was produced by homologous recombination between the linearised DNA and the vector was purified using a plaque assay. Five purified viruses per construct were amplified to a P1 virus stock, which was tested for protein expression in small-scale expression trials. Samples of media from these trials were probed for expressed and secreted EPCR using an anti-penta-his antibody, directly conjugated to HRP (QIAGEN). This blot showed expression of both constructs in the *Sf*9/baculovirus system, with EPCRL expressing to a higher level than EPCRS.

EPCRL, which expressed to a higher level than EPCRS, was taken on to larger scale expression. P1 virus was amplified twice to produce P3 virus, then 1 L of *Sf*9 cells was infected with P3 virus. After 72 h the protein was harvested from the media by centrifuging the cells. This media, containing the recombinant protein, was buffer exchanged and concentrated using tangential flow filtration into 20 mM Tris pH 8; 500 mM NaCl, then incubated with 3 ml Ni-NTA overnight, followed by a Ni-NTA purification. Elution fractions were subsequently concentrated and gel filtered on an Superdex 75 column. The absorbance trace at 280 nm showing pure EPCRL in the supernatant at an elution volume consistent with a monomer species of the correct size. The final yield of this preparation was 2 mg/L *Sf*9 cells.

3.3.4 EPCR expression in S2 cells

Though the yield from the *Sf*9/baculovirus system was acceptable, I opted to trial expression in *Drosophila* S2 cells, due to previous success in the lab [113]. This system used a proprietary expression vector, pExpreS^2-1 (ExpreS^{2}ion Biotechnologies), to constitutively express the protein of interest in S2 cells, secreting the protein in the media. In order to maximise yield, the EPCR gene was resynthesised, optimising the codon usage for *Drosophila* cells. The EPCRL domain boundaries were used, as this construct produced

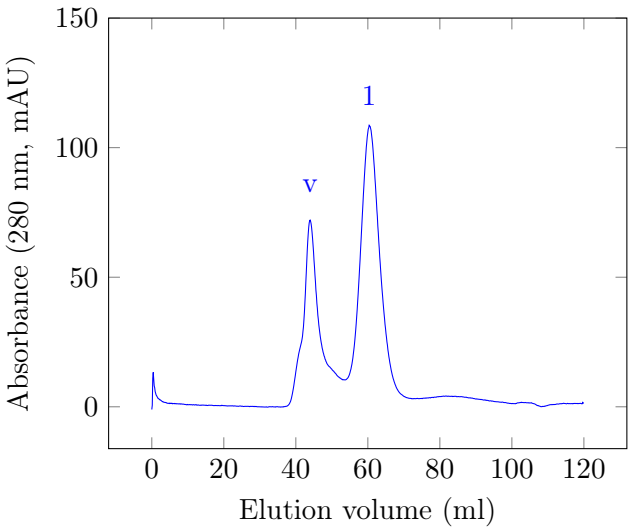


Figure 3.4: Preparative gel filtration trace for EPCR from Sf9 cells. The peak corresponding to soluble, monodisperse EPCR is marked with a 1, and the void peak is marked with a v.

a higher protein yield in *Sf*9 cells. For purification and biophysics purposes, I added four N-terminal elements: a BiP leader sequence for protein secretion [298]; a biotin acceptor peptide (BAP) tag; a hexa-histidine tag and a TEV protease cleavage site (GeneArt, ThermoFisher Scientific). These tags would allow for secretion and purification of the protein and would be useful for biophysical experiments. I sub-cloned this gene into the pExpreS^2 vector, then transfected this gene into S2 cells. I simultaneously transfected a GFP construct in the pExpreS^2 vector, as a positive control.

After three weeks of selection using Zeocin, the large majority of the positive control cell line were positive for GFP [1]. At this point expression trials for the EPCR-expressing S2 cells were then set up. After 72 h samples of media were analysed by Western blot. These trials showed that EPCR expressed in S2 cells, and as a result EPCR was grown up on a larger scale.

EPCR-containing supernatent was buffer exchanged and concentrated using tangential flow filtration into 20 mM Tris pH 8, 500 mM NaCl. Ni-NTA was then used to purify EPCR from the supernatent via its His-tag. Elution fractions from this Ni-NTA were concentrated to 2 ml for size exclusion gel chromatography. This gel filtration of these fractions shows

[1] Cells were checked for GFP fluorescence with the help of Sapan Gandhi.

EPCR to be monodisperse in nature (Figure 3.5). EPCR expression in S2 cells gave a final yield of 30 mg/L, clearly superior to other expression systems.

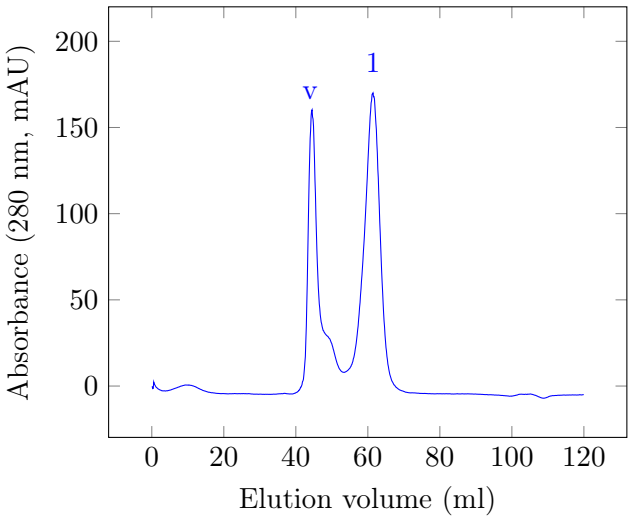
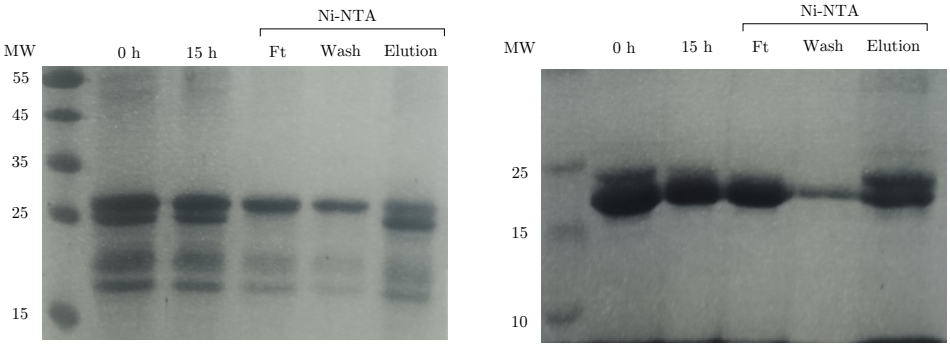


Figure 3.5: Preparative gel filtration trace for EPCR from S2 cells. The peak corresponding to monomeric EPCR is marked with a 1, and the void peak is marked with a v.

3.3.5 Preparing proteins for crystallography

Flexibility in proteins can prevent the crystallisation of individual monomers into the regular arrays that are required to produce a good diffraction pattern. I hence attempted to cleave the N-terminal purification tags introduced in the CIDR and EPCR constructs and the glycosylations added by the S2 expression system. I trialled TEV protease cleavage, utilising the TEV protease cleavage site introduced into the constructs, to cleave the N-terminal tags on the CIDR and EPCR constructs. CIDR domains were incubated overnight with TEV protease, removing uncleaved protein and the TEV protease, which had a histidine tag, the following morning using Ni-NTA agarose. A subsequent SDS-PAGE gel shows that the majority of the CIDR is cleaved after the overnight incubation, flowing through the Ni-NTA column (Ft) due to the lack of His-tag (Figure 3.6a). This step also removed a contaminating species at 24 kDa, which remained bound to the Ni column in the elution fraction. Not all of the CIDR was cleaved, with some CIDR present in the elution fraction.

EPCR from S2 cells was incubated overnight with TEV protease, again removing uncleaved protein and TEV protease using Ni-NTA. The subsequent SDS-PAGE gel demonstrates that EPCR, once incubated with TEV protease, no longer binds to the Ni-NTA column due to the lack of a His-tag (Figure 3.6b). The EPCR construct had four predicted glycosylation sites, which could be glycosylated by the S2 expression system. Glycans added by the system are typical of insect glycans: paucimannose with or without the addition of fucose, not trimmed or sialated like their mammalian counterparts [299]. These glycans were removed using a combination of endodeglycosidases Hf and F3, with overnight reactions showing glycan removal by SDS-PAGE gel (Figure 3.7, left). Analytical gel filtration of these samples reinforces this success, as deglycosylated EPCR elutes at a later volume than glycosylated EPCR (Figure 3.7, right).



(a) IT4var07 CIDR **(b)** EPCR

Figure 3.6: SDSPAGE gels of recombinant IT4var07 CIDR, typical of CIDR TEV cleavage (a) and EPCR (b). Samples were taken before and after incubation (0 h and 15 h), and from the flowthrough, wash and elution fractions from the subsequent Ni-NTA purification.

Subsequently large preparations of recombinant protein were purified using the following standard purification procedure. Since the CIDR domains were not produced in large amounts in the soluble fraction from the OPPF strains, CIDRs were instead expressed in the parent strain, B834. Protein was purified from *E. coli* inclusion bodies using Ni-NTA, then refolded using a decreasing gradient of guanidine hydrochloride. EPCR was affinity-purified from S2 supernatent using Ni-NTA. The His-tags from both proteins was removed using TEV protease, with glycans on EPCR removed using a combination of deglycosidase enzymes.

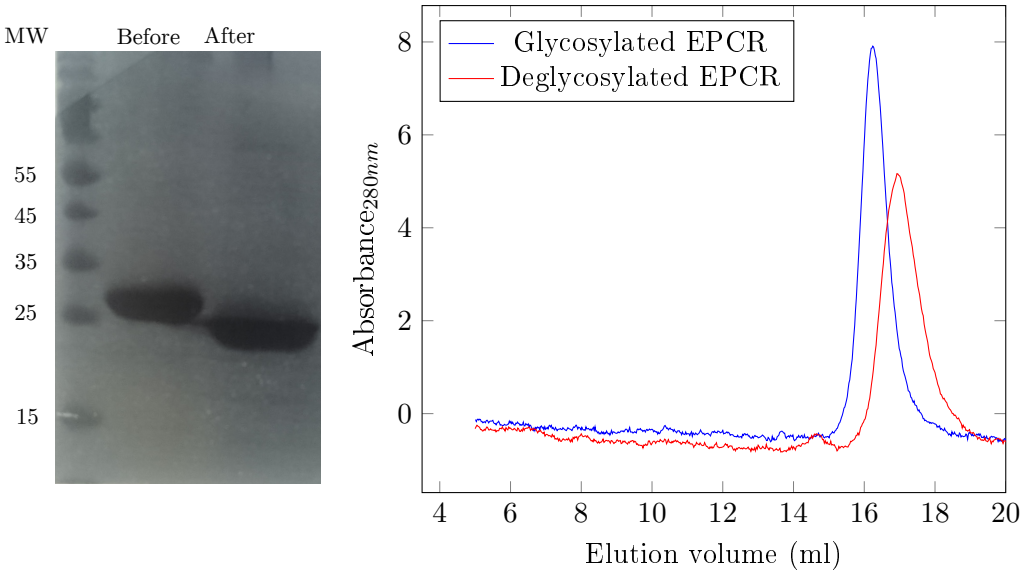


Figure 3.7: Left: SDSPAGE analysis of EPCR, before and after deglycosylation, showing a band shift signifying deglycosylation. Right: Analytical gel filtration of glycosylated and deglycosylated EPCR. The peak shift to a later elution volume signifies deglycosylation of EPCR, which elutes as a species of lower molecular weight.

3.4 Functional validation of recombinant proteins

Given that both CIDR and EPCR were recombinantly produced, it was important to make sure that recombinant protein is folded and functionally active before proceeding with experiments. The following section describes the characterisation of the recombinant proteins.

3.4.1 Assessing the fold of CIDRs

Since the CIDR constructs had been refolded from inclusion bodies, it was important to check that the purified protein isolated from gel filtration contained secondary structure. To check this, I took circular dichroism spectra of the three expressed CIDRs. A CD spectrum was taken between 190 nm and 260 nm at room temperature for each CIDR, correcting for the buffer contribution. The spectra show that all three CIDR domains contain a high α-helix content, comparing the spectra to standards [260]. This result is consistent with the secondary structure predictions, which suggest CIDR domains are

predominantly α-helical domains.

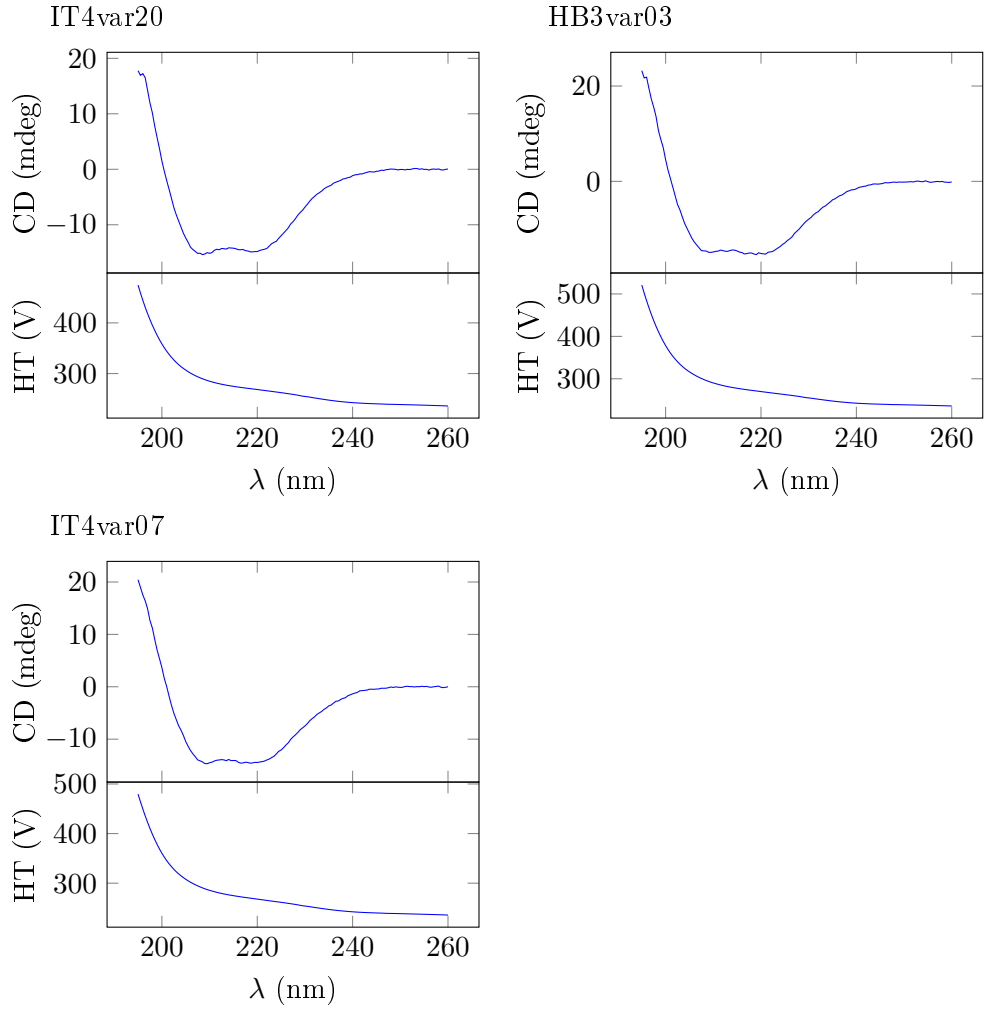


Figure 3.8: Circular Dichroism traces of the refolded CIDRs. Spectra were compared to typical spectra for an alpha helical protein to assess secondary structure. Data were disregarded if high tension (HT) voltage was over 1000 V.

3.4.2 Binding of recombinant CIDRs to recombinant EPCR using surface plasmon resonance

To check the function of both the CIDRs and EPCR, I used surface plasmon resonance to compare the binding affinity of the proteins to control proteins[2], used in [140]. In the first set of experiments I tested recombinant EPCR for activity against a previous HB3var03 CIDR construct. Two flowpaths on a CM5 chip were coated with EPCR - one

[2]With thanks to the Lavstsen group at the University of Copenhagen

with recombinant EPCR and one with a control EPCR(Sino Biological, P.R. China). A control HB3var03 CIDR sample was flown over the flowpaths at different concentrations, subtracting the baseline to give each dataset. Though the dataset suffered badly from baseline drift, the observed curves from these experiments show that both the recombinant EPCR and the control EPCR interact with the control CIDR with similar affinities (Figure 3.9). The affinities were calculated from equilibria bound concentrations.

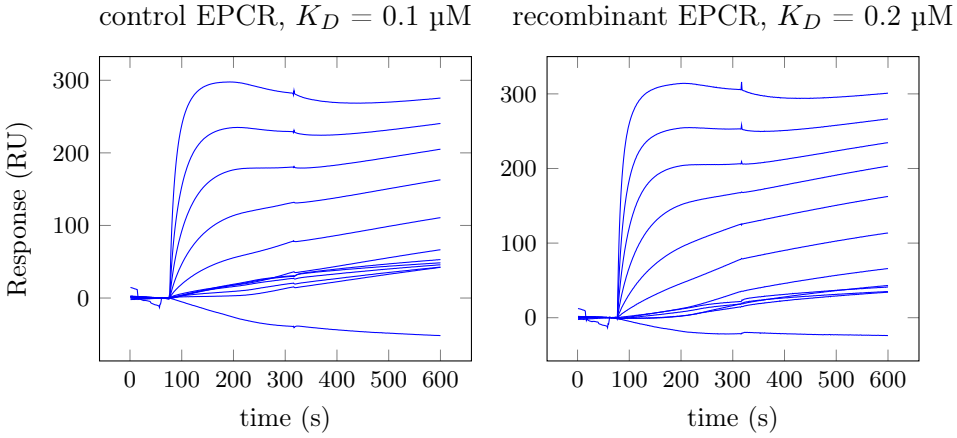


Figure 3.9: SPR curves of control HB3var03 CIDR binding to the control EPCR and recombinant EPCR respectively. Each dataset consists of binding data for a 12 sample 2-fold dilution series from 1 µM, with the blank-subtracted data shown in blue. Putative affinities are calculated using affinity fitting.

After confirming the recombinant EPCR was functionally active, I tested the panel of expressed CIDRs for functional activity by coupling the EPCR to a CM5 chip and flowing a concentration series of each CIDR over the flow cell (Figure 3.10). The panel of CIDRs gave relatively poor data, with kinetic models fitting poorly to the data. Thermodynamic dissociation constants, K_D, were calculated instead using affinity fitting. The putative K_D values are shown in table 3.2. These values are likely an underestimation of the actual K_D, as the curves at lower concentration do not saturate, a requirement for affinity fitting. The IT4var20CIDR dataset had to be curtailed to give a realistic fit, as at higher concentration the R_{max} decreased.

Due to the poor quality of data, especially in testing the recombinant CIDRs, I opted to try a different SPR coupling method, coupling using the biotinylated biotin acceptor peptide (BAP), included in the S2 EPCR construct, rather than using lysine coupling. The CAP chip system facilitates this, consisting of an SPR chip coated in single-stranded

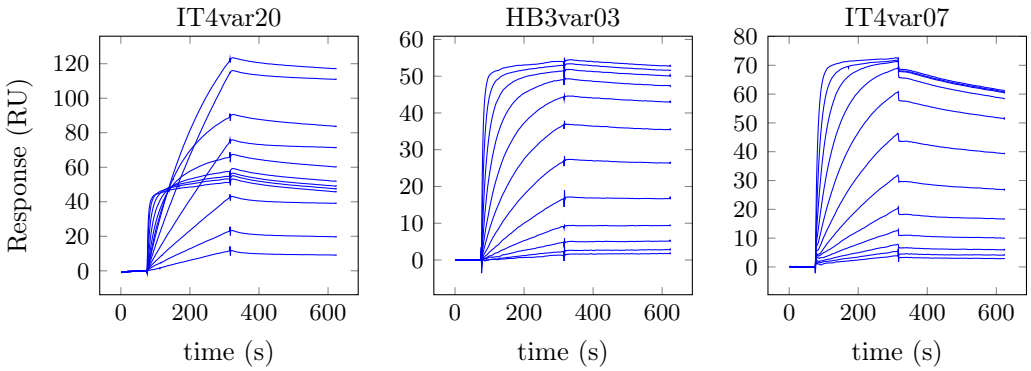


Figure 3.10: Surface plasmon resonance of the panel of refolded CIDRα1 domains. Each dataset consists of binding data for a 12 sample 2-fold dilution series from 1 µM, with the blank-subtracted data shown in blue.

CIDR	K_D (nM)	R_{max} (RU)
IT4var20	2000	n/a
IT4var20 (curtailed)	9	n/a
HB3var03	16	56
IT4var07	20	75

Table 3.2: SPR summary table for the panel of refolded CIDRα1 domains binding to EPCR. The curtailed dataset took the lowest 5 concentration curves to calculate constants.

DNA over which complementary ssDNA, coupled to strepatavidin, is flown. This results in a regeneratable streptavidin surface to which a biotinylated protein can bind. This system allows the presentation of a surface of EPCR in the same orientation as *in vivo*, as the BAP tag was designed on the N terminus of the EPCR construct, which is proximal to the membrane *in vivo*. This also facilitates the accurate calculation of kinetic parameters as all EPCR binding sites are available for the CIDR.

To facilitate this I biotinylated the BAP-tagged EPCR and coupled it to a CAP chip. I then flowed over concentration series of CIDRs (Figure 3.11). Unlike with the CM5 method, kinetic models fit the data well (χ^2 values of 9, 16 and 0.7 respectively), allowing for an accurate calculation of kinetic and thermodynamic constants. The results show that if EPCR is presented in the correct orientation, CIDRs bind to EPCR at sub-nanomolar affinities, with fast association rate and a slow dissociation rate.

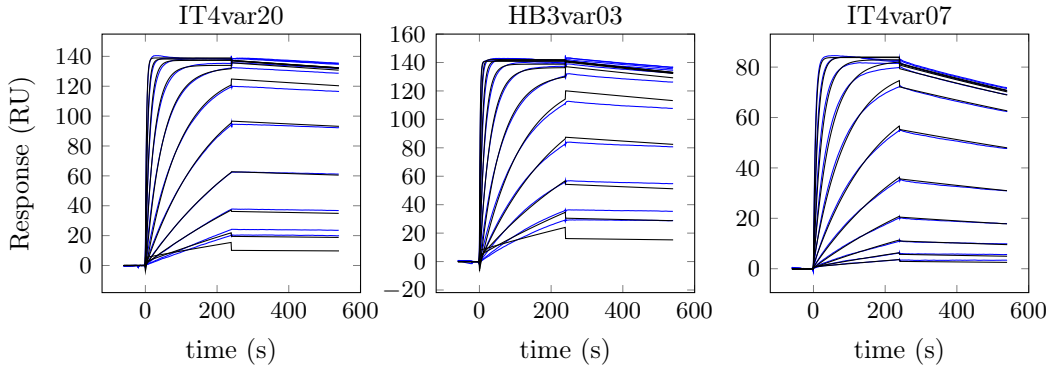


Figure 3.11: Biotinylated SPR curves for IT4var20, HB3var03 and IT4var07 binding to EPCR. Each set of curves is a series of two-fold dilutions from 1 µM, with blank-subtracted data shown in blue and kinetic fits shown in black.

CIDRα1	K_D (nM)	k_a (M^{-1} s^{-1})	k_d (s^{-1})	χ^2
IT4var20	0.37	3.3×10^5	1.2×10^{-4}	9
HB3var03	0.37	5.3×10^5	2.0×10^{-4}	16
IT4var07	2.9	1.6×10^5	4.7×10^{-4}	0.7

Table 3.3: A summary of SPR parameters from experiments testing the binding of the CIDRs to EPCR using the CAP chip method. All parameters are calculated using kinetic 1:1 binding models.

3.4.3 Assessing glycosylation-dependence of the CIDR:EPCR interaction

EPCR was deglycosylated for crystallography, as glycans can sometimes hinder the formation of an ordered lattice. It was hence important to check whether the interaction was glycosylation dependent. Glycosylated EPCR and deglycosylated EPCR were coupled to the CM5 chip on different flowpaths, and control HB3var03CIDR was flown over both surfaces (Figure 3.12). The similar dissociation constants for these interactions suggest that the interaction is not glycosylation-dependent, calculated using kinetic 1:1 modelling.

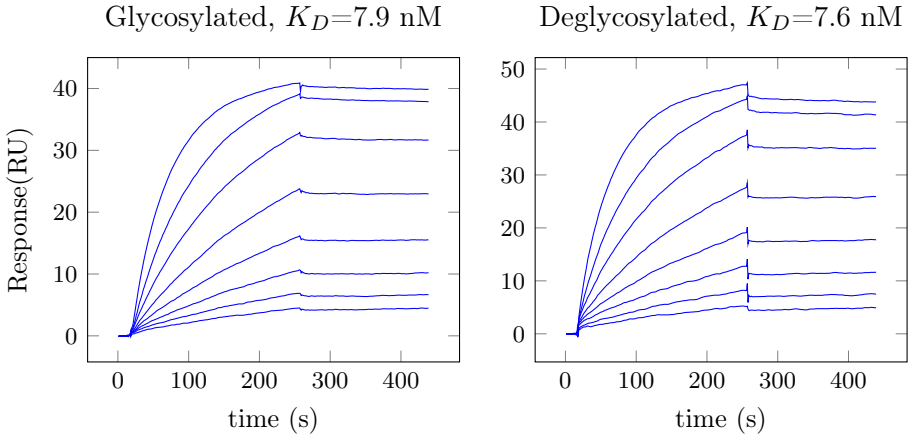


Figure 3.12: SPR curves demonstrating the glycosylation-independent nature of the CIDR:EPCR interaction. Each set of curves is a two-fold dilution series from 125 µM, resulting in similar K_D values, with the blank-subtracted data shown in blue.

3.4.4 Size-exclusion chromatography multi-angle laser light scattering of EPCR:CIDR

The CIDR domains and EPCR were also checked for interaction using gel filtration to test the purity of the complex. SEC-MALLS was run for EPCR and the HB3var03 CIDR domain individually and a 1:1 mixture of the two proteins. This experiment runs the protein through a size exclusion gel chromatography column, in this case a Superdex 75 column, before shining laser light at multiple angles through the eluant to ascertain shape-independent size of the components. The results of these experiments show that the EPCR:CIDR complex is stable over the lifetime of a gel filtration experiment and pure, with a consistent

molecular weight (48 kDa) across the peak suggesting a 1:1 complex (Figure 3.13). Both components also elute as single peaks with a consistent molecular weight (26 kDa for the CIDR, 22 kDa for EPCR) across the peak suggesting monodisperse species.

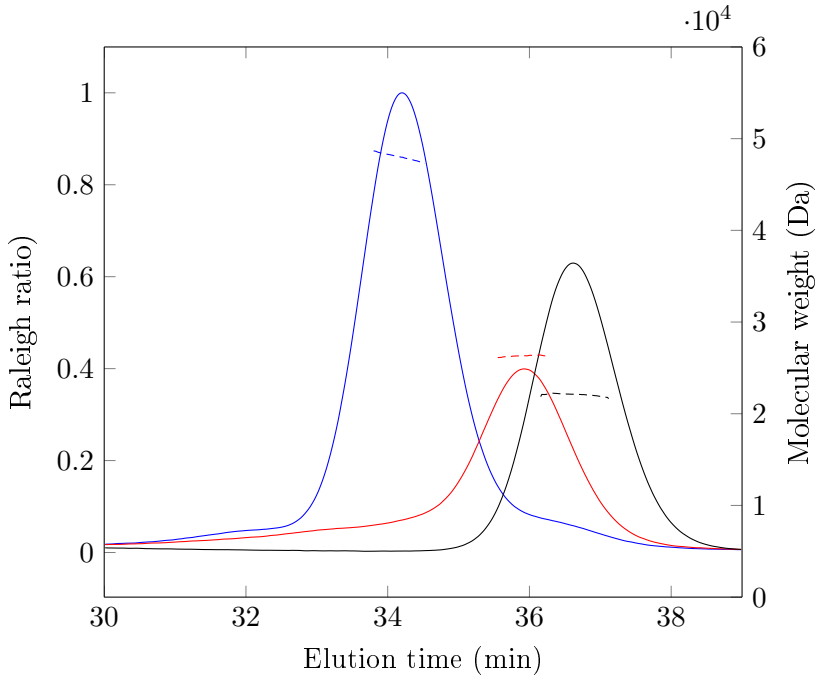


Figure 3.13: SECMALLS of the CIDR:EPCR complex. EPCR is shown in blue, the HB3var03 CIDR in red and the complex in black. Molecular weights are shown as dashed lines, while Rayleigh ratio is shown as solid lines.

3.4.5 Analytical Ultracentrifugation of EPCR:CIDR

To confirm this 1:1 stoichiometry, both velocity and equilibrium analytical ultracentrifugation (AUC) experiments were run. Velocity AUC experiments were run using both the components individually and in a 1:1 mixture. Fitting the data to a continuous c(S) model shows that EPCR sediments as a singular, monomeric peak, while the CIDR domain sediments with a very small (<5%) amount of dimer, as well as a monomeric species. The complex sediments as a singular species at the correct molecular weight, suggesting all protein in the solution is bound in complex (Figure 3.14). The equilibrium AUC confirmed that a 1:1 complex is observed, with an observed molecular weight of 48 kDa matching well to the calculated molecular weight of 46.8 kDa (Figure 3.16).

As well as size parameters, we can get semi-quantitive shape parameters from the velocity

AUC data, in the form of the frictional ratio, $\frac{f}{f_0}$. This ratio compares the frictional coefficient of the sample species, f, to the frictional coefficient of a spherical object of similar size, f_0. The CIDR domain from HB3var03 and EPCR sediment with frictional ratios of 1.37 and 1.34 respectively, with 1.3 being a typical value for a globular protein [300], while the complex sediments with a frictional ratio of 1.5, which is consistent with two adjacent spheroids forming a complex (Table 3.4).

Species	f/f_0	MW_{app} (kDa)	MW_{calc} (kDa)	Sedimentation Coefficient (S)	Comment
HB3var03 CIDR	1.37	28.4	26.1	2.42	monomer
		54.7	52.2	3.78	dimer
EPCR	1.34	22.2	20.7	2.12	monomer
Complex	1.5	47.9	46.8	3.17	complex

Table 3.4: A summary table of frictional coefficients and apparent molecular weights from the three datasets.

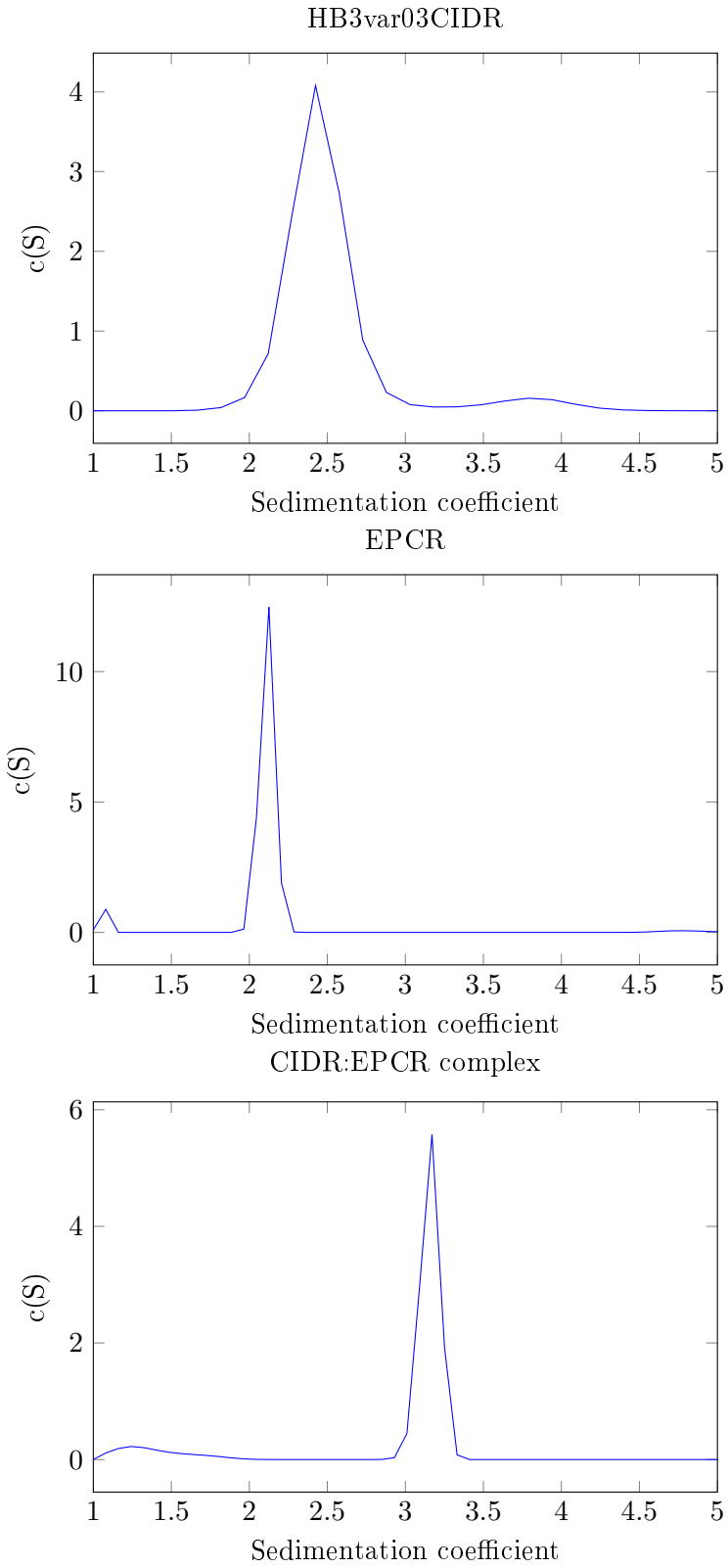


Figure 3.14: Velocity analytical ultracentrifugation c(S) fits, generated by sedfit. Residuals shown in 3.15.

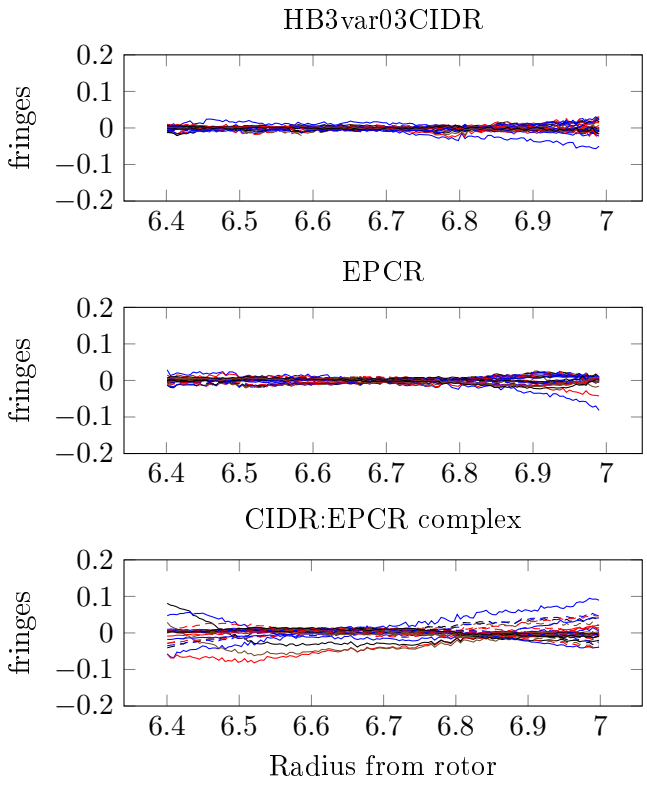


Figure 3.15: Velocity analytical ultracentrifugation residuals for the fits shown in 3.14.

3.5 Solving the structure of the CIDR:EPCR complexes

3.5.1 Crystallisation of the CIDR:EPCR complex

Crystal plates were set up for both the CIDRs alone and CIDR:EPCR complexes to maximise the chance that one would crystallise. Components of each complex were incubated together for thirty minutes, before gel filtration of the mixture (Figure 3.17). The complex peak was selected in each case and concentrated to 10 mg/ml. The CIDR domains and complexes were dispensed into 96-well sitting drop plates, using the commercial screens JCSG+, Pact Premier, Morpheus and Midas, with the addition of ProPlex in the case of complexes (all Molecular Dimensions). Plates were set up at 4 °C and 18 °C. Of these species only two complexes crystallised, with no CIDR domains forming crystals alone.

Crystals of HB3var03 CIDRα1:EPCR grew as needle clusters in many PACT conditions, varying in length from 10 to 100 µm. The largest, with the most distinct needles grew with a reservoir solution of 0.2 M NaNO3; 0.1 M BTP pH 8.5; 20% PEG 3350. SDS-PAGE analysis

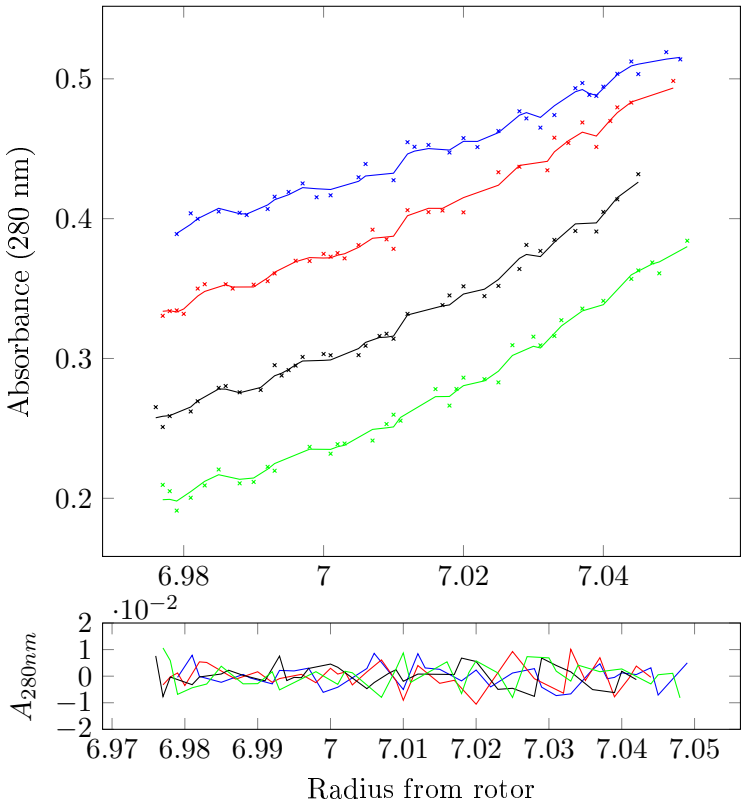


Figure 3.16: Equilibrium analytical ultracentrifuge data of the CIDR:EPCR complex. Complexes were run at 10,000 rpm, 12,000 rpm, 14,000 rpm and 16,000 rpm (blue, red, black and green plots) with the fit at each speed shown as a line and the residuals shown in the bottom panel.

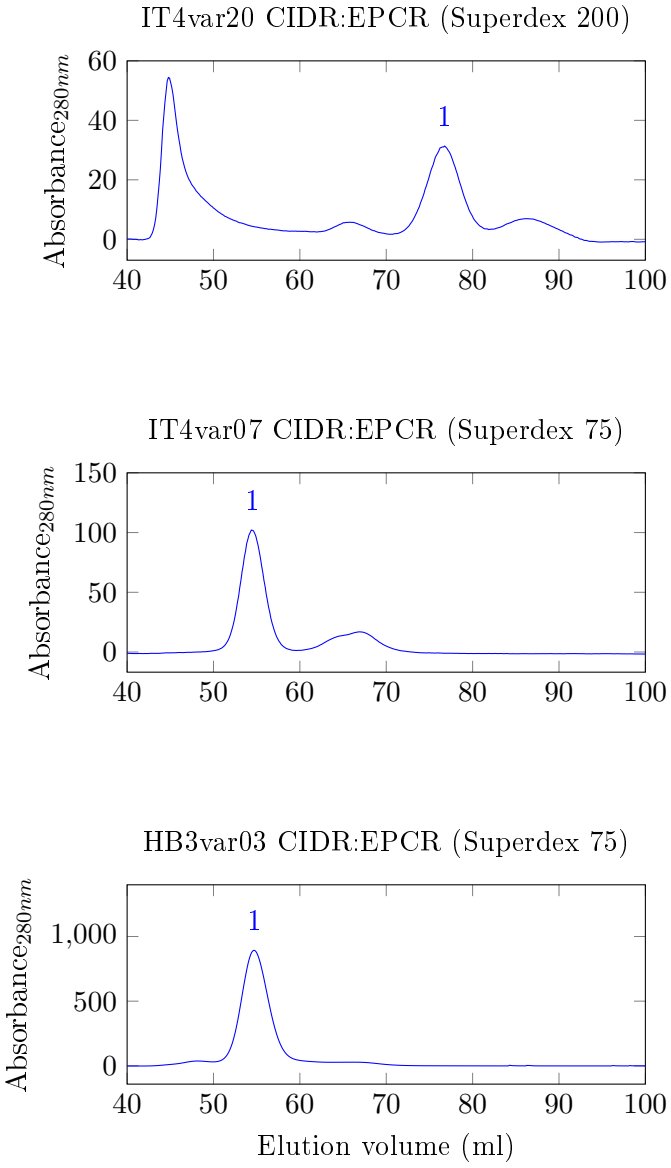


Figure 3.17: Preperative gel filtrations for the three complexes set up in crystal plates. The IT4var20 CIDR:EPCR complex was run on a Superdex 200, while other species were run on a Superdex 75. Peaks labelled 1 elute at the right size for a 1:1 complex and contain both proteins on an SDS-PAGE gel.

confirmed these crystals to contain both components. Needle clusters were transferred into a drop of well solution containing 25% ethylene glycol and single needles detached. These needles were then cryo-cooled in liquid nitrogen for storage and data collection.

Multiple crystal forms of the IT4var07 CIDRα1:EPCR complex grew in the Proplex and PACT premier crystallisation screens over 1 week. The predominant forms, crossed needles and hexagonal prisms, grew in PEG-based conditions containing sodium and/or chloride, at pH of between 7 and 8. The hexagonal prisms were larger and clearer than the crosses, reaching between 60 and 100 µm in diameter, so these were selected for optimisation. SDS-PAGE analysis confirmed these hexagonal crystals to contain both components.

Optimisation was achieved using the Silver Bullets additive screen. The largest crystals grew best with a PACT reservoir solution of 0.2 M NaNO3; 0.1 M BTP pH 7.5; 20% PEG 3350, with additives from Silver Bullets well condition H9 included[3]. Crystals were harvested after 7 days. I initially attempted to loop these crystals into drops of mother liquor containing 25% glycerol, ethylene glycol or MPD, but these crystals immediately fractured, presumably due to the sudden change in condition. Instead, I gradually added MPD to the crystals in a drop of mother liquor to a final concentration of 25%. These crystals did not crack and could be cryo-cooled in liquid nitrogen for storage and data collection.

3.5.2 Data collection and solution of HB3var03CIDR:EPCR

Data were collected for the HB3var03 CIDR:EPCR complex at the Diamond Light Source on beamline I-04. Initially a series of cryocooled needles were mounted on the goniometer and shot at three angles to check for diffraction quality and resolution limits. This identified two crystals which diffracted to 3 Ångströms. These datasets were processed automatically by FastDP and Xia2, which both suggested a space group of $C222_1$. I collected full datasets from these two crystals, using a collection strategy suggested in the outputs of EDNA and Mosflm (Table 3.5). The dataset which diffracted to a higher resolution was processed using iMosflm to integrate the reflections. The data collection statistics are summarised

[3]0.13% 2,5-pyridinedicarboxylic acid; 0.13% pyromellitic acid; 0.13% salicylic acid; 0.13% trans–1,2-cyclohexanedicarboxylic acid; 0.13% trans-cinnamic acid; 0.013 HEPES pH 6.9

below (Table 3.6). This crystal diffracted to 2.65 Å, with a signal-to-noise ratio in the outer shell: $\frac{I}{\sigma I} = 2.3$. The R_{merge} value in the outer shell is 0.58, within the generally accepted limits for data quality.

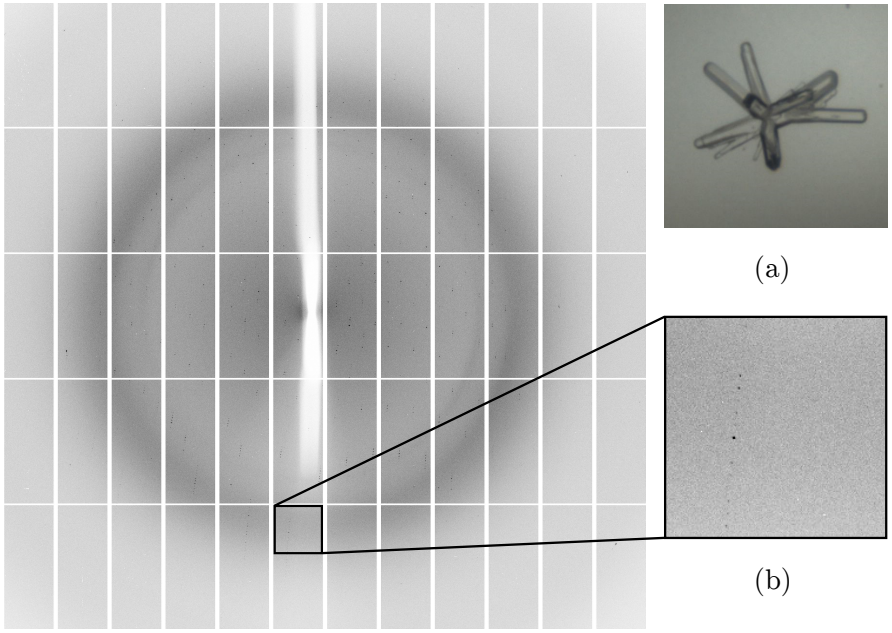


Figure 3.18: Diffraction data for HB3var03 CIDR:EPCR. Single needles were broken from needle cluster crystals (a). A sample diffraction image is shown on the left, a portion of which has been magnified (b).

I attempted to use molecular replacement to get phase information for this dataset in the first instance. I used EPCR, which makes up 43% of the protein mass in the complex as the search model. The top result for this search showed two EPCR molecules in the unit cell, which was consistent with the Matthew's coefficient estimation. As well as the EPCR molecules, with additional helical density present in the electron density map. Building short helices into this density, and refining the dataset, yielded additional density. Buccaneer identified more possible protein backbone in the electron density maps, which I curated, before successive rounds of refinement and rebuilding allowed the solution of the core three α-helical bundle, a binding arm close to EPCR, and a β-sheet segment on the opposite face. At this point, I started to assign sequence to the main chain, using bulky tryptophans in the core helical bundle as a starting point. Another round of refinement and rebuilding then allowed modelling of the lipid in the central EPCR cavity and a couple of sugar residues which had not been fully cleaved in the preparation. Finally the most of

the loops were resolved and modelled, apart from the loop consisting of residues 683-692, which could not be resolved. Final solution parameters are shown in Table 3.7.

An R_{work} of 0.22 suggests a good fit of model to data at 2.65 Å, while the slightly higher R_{free} value of 0.25 shows that independent, unprocessed reflections can be explained by the model and suggests that the model does not overinterpret the data. The Ramachandran plot shows that 92.8% of the residues are in the allowed region, with 2 residues from EPCR in the generously allowed regions. These two residues are two threonines, one from each copy EPCR (T88). This particular residue is strained in all solved EPCR structures [227], suggesting it is modelled correctly.

3.5.3 Data collection and solution of IT4var07CIDR:EPCR

Data for the second complex were collected at Diamond Light Source on beamline I-02. Fragments of the hexagonal prisms were mounted on the goniometer and shot in the same method as for the HB3var03 CIDR:EPCR complex. Initial scans revealed one crystal which diffracted to 3 Ångströms. FastDP and Xia2 processed this initial scan to the space group $P3_121$. A complete dataset was collected based on strategies identified by EDNA and Mosflm (Table 3.5). Subsequent processing by iMosflm gave data collection statistics below (Table 3.6). The signal-to-noise ratio of 2.3 in the outer shell justifies the resolution cut-off of 2.9 Å. The R_{merge} value for the outer shell is high compared to accepted values of around 0.8, but this is not unusual for datasets with a high multiplicity.

Initial attempts to gain phase information used the HB3var03 CIDR:EPCR model as a search model. This initial molecular replacement solution was poor, with poorly resolved electron density. Use of Pointless confirmed that Mosflm had assigned the spacegroup incorrectly - the spacegroup was an enantiomorphic spacegroup, $P3_221$. I reprocessed the data with this new spacegroup and used this new dataset in molecular replacement with the same search model. This resulted in a well resolved EPCR molecule, and a well resolved CIDR core. Cutting back CIDR loops and mutating the remaining CIDR residues to the IT4var07CIDR sequence resulted in a much better resolved structure, facilitating the rebuilding of the some of the IT4var07CIDR loops in the correct conformations. Electron

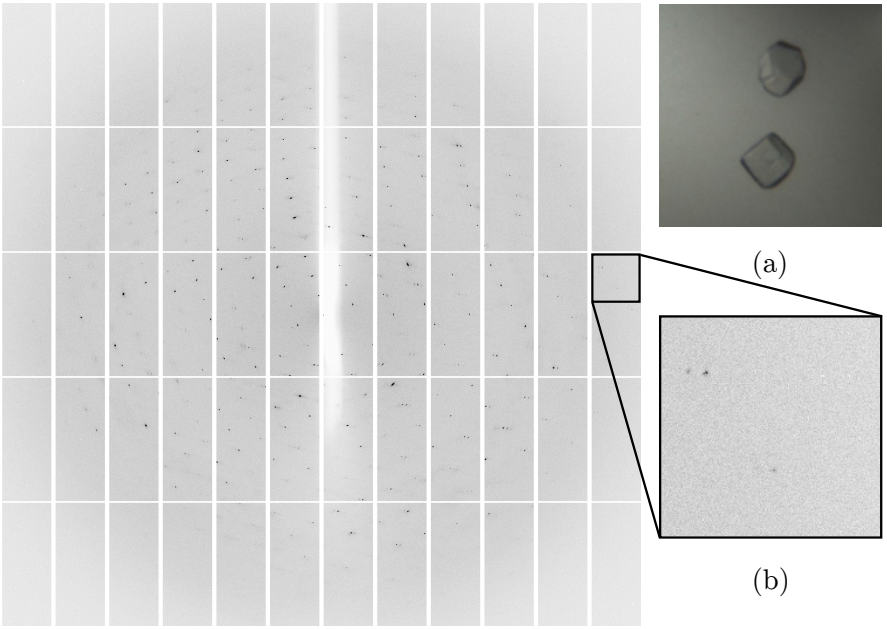


Figure 3.19: Diffraction data for IT4var07 CIDR:EPCR. This complex grew as hexagonal prisms (a). A sample diffraction image is shown on the left, a portion of which has been magnified (b).

density for the loops consisting of residues 552-562, 594-599, 678-691 was of low quality, so the polypeptide segments in these loops were not built. The solution was checked for model bias by refining using either component of the complex. Density maps after these refinements showed clear density for the omitted component, suggesting model bias was not a factor in our solution. Final solution parameters are summarised below (Table 3.7).

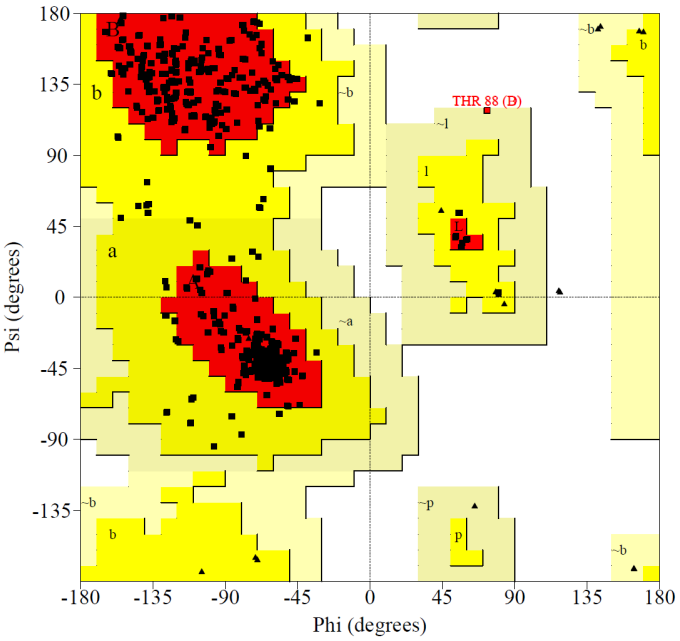
Given the slightly lower resolution of this structure compared to the HB3var03 CIDR:-EPCR complex (2.9 Å versus 2.65 Å), the higher R_{work} and R_{free} values represent a good fit of model to data. The Ramachandran plot shows a high number of residues in the allowed region (92%), with 0.3% residues (1 residue) in the generously allowed region and no residues in the disallowed region. The residue, T88, in the generously allowed region is the same residue in EPCR as in the other structure.

CIDR:EPCR complex	HB3var03	IT4var07
Beamline	I-04	I-02
Wavelength (Å)	0.98	0.9787
Aperture (µ)	200	200
Transmission	50%	100%
Omega start	0°	51°
Oscillation	0.2°	0.2
Exposure (s)	0.2	0.2
Total oscillation	360°	180°

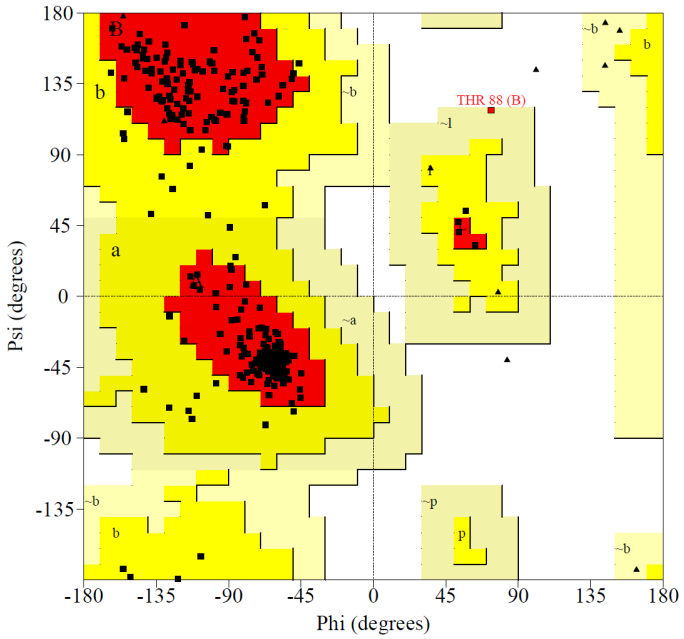
Table 3.5: Summary table of X-ray diffraction data collection strategies for the CIDR:-EPCR complexes.

	HB3var03 CIDRα1:EPCR			IT4var07 CIDRα1:EPCR		
Beamline	Diamond I-04			Diamond I-02		
Wavelength (Å)	0.98			0.9787		
Space group	$C222_1$			$P3_221$		
Cell parameters (Å)	a = 66.14, b = 94.72, c = 290.84; α= β= γ= 90°			a = b = 56.02, c = 250.50; α= β= 90°, γ= 120°		
Shell	Overall	Inner	Outer	Overall	Inner	Outer
Resoution(Å)	54.23-2.65	54.23-8.38	2.79-2.65	125.25-2.90	125.25-9.17	3.06-2.90
R_{mrg}	0.093	0.033	0.58	0.061	0.034	0.882
$I/\sigma(I)$	9.2	24.2	2.3	16.4	40	2.3
Completeness (%)	95.3	90.8	95.4	99.9	99.9	99.1
Multiplicity	3.6	3.5	3.6	8.4	7.1	6.9

Table 3.6: Data collection statistics for both complexes. IT4var07 statistics are displayed as calculated for the correct spacegroup, $P3_221$



(a) Ramachandran plot for the HB3var03 CIDR:EPCR complex



(b) Ramachandran plot for the IT4var07 CIDR:EPCR complex

Figure 3.20: Ramachandran plots for the HB3var03 CIDR:EPCR (a) and the IT4var07 CIDR:EPCR (b) complex structures, generated by Procheck [301]. Triangles represent glycine residues, while other residues are represented by squares. One residue in each complex is in a disallowed region (T88 from EPCR).

	HB3var03 CIDRα1:EPCR	IT4var07 CIDRα1:EPCR
Resolution (Å)	2.65	2.90
R_{work}	0.222	0.247
R_{free}	0.255	0.276
No. of protein residues in model	733	359
rmsd bond lengths (Å)	0.009	0.011
rmsd bond angles (°)	1.25	1.28
Ramachandran plot		
Allowed region (%)	92.8	92.0
Additional allowed region (%)	6.9	7.7
Generously allowed region (%)	0.3	0.3
Disallowed region (%)	0	0

Table 3.7: Refinement statistics for both CIDR:EPCR complexes.

3.6 Analysis of the CIDR:EPCR complexes

The two CIDR:EPCR complexes crystallise in different spacegroups, in different conditions. Despite this, the three different copies of the complexes (two in the HB3var03 CIDR:EPCR structure, one in the IT4var07 CIDR:EPCR structure) look very similar. The EPCR molecules are similar to one another, aligning with an RMSD of 0.18 Å. Though the CIDRs from the two complexes have a different primary sequence (78% identity) they share a common fold and interaction site. Indeed overlaying the polypeptide backbone gives an RMSD of 0.3 Å, when overlaying the two domains. When complexes are aligned by the EPCR molecule, the CIDR-CIDR RMSD is 2 Å(Figure 3.21).

The CIDR domains comprise of a long three-helical core, with a β-sheet segment at the N-terminus of the core bundle, and a kinked helix followed by another helix between the second and third core helices, termed the binding arm, which lay roughly perpendicular to the core helices. The binding arm contains the majority of the residues which directly contact the EPCR (7 of 9 residues, Figure 3.24) and, along with the proximal portion of

the EPCR, is well resolved, with B-factors of 20-40 in the structure (Figure 3.22). Away from the interaction site, the majority of the surface comprises of loops. Portions of these loops could not be resolved in our datasets, while other portions have higher B-factors. Both of these factors suggest the loops in the CIDRs to be highly flexible.

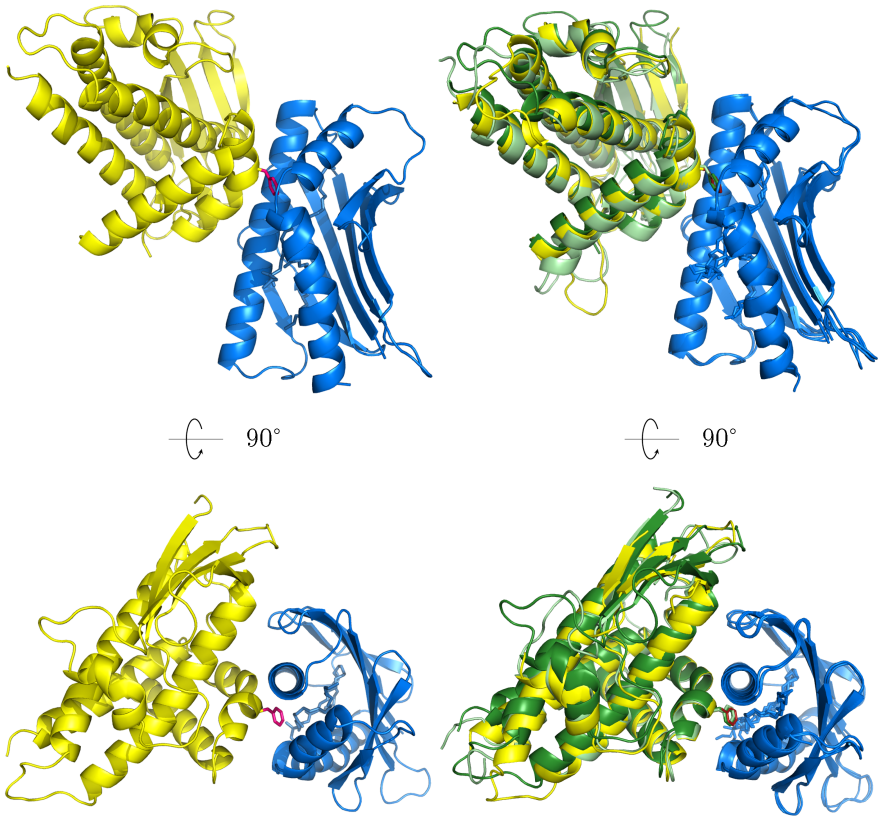


Figure 3.21: The CIDR:EPCR crystal structure. Left: Backbone representation of the HB3var03 CIDR (yellow) in complex with EPCR (blue). F656, the protruding central phenylalanine is highlighted in pink. Right: The two HB3var03 CIDRs (yellow, dark green) and the IT4var07 CIDR (light green) superimpose with a RMSD of 1.4 Å.

3.6.1 Analysis of the interaction site

The HB3var03 CIDR:EPCR interface has a surface area of 978 Å^2, while the IT4var07 CIDR:EPCR interface has a surface area of 950 Å^2. In both structures there are nine residues that are closer than three Ångströms to EPCR residues (Table 3.8). These nine residues consist a pair of phenylalanines, forming a central hydrophobic patch, and seven hydrophilic residues which form a ring surrounding the central patch. One of the central

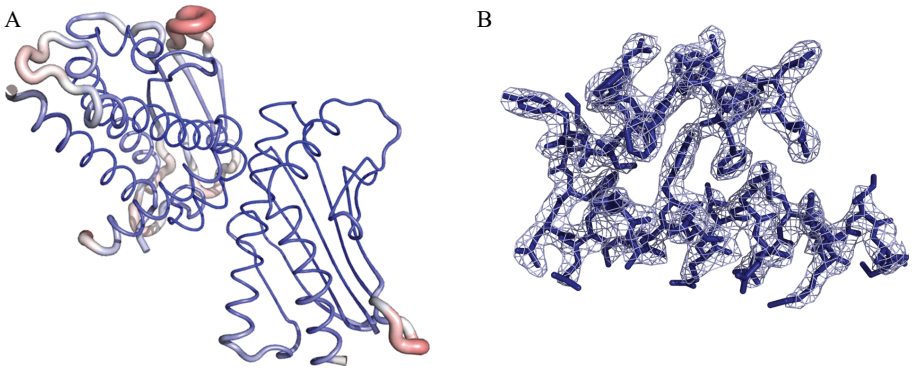


Figure 3.22: Left: A representation of HB3var03 CIDR:EPCR crystal structure (molecules A and B), with thicker backbone representation denoting higher B factors. Right: The electron density, contoured at 1.0 σ() for the residues from HB3var03 CIDR that form the helix and kinked helix at the heart of the EPCR binding site.

phenylalanines protrudes into a hydrophobic box in EPCR consisting of hydrophobic portions of 8 amino acid side chains, plus a small portion of the bound lipid (coloured pink in Figure 3.21). This protruding phenylalanine is presented at the tip of a kinked helix, which stabilised by a phenylalanine, a tryptophan and a valine (F651, F658 and W669 in the HB3var03 structure), which pack together on the opposite side to the kink (Figure 3.23). The rest of the binding residues on the CIDR interact with sidechains from EPCR residues at the interface (Figure 3.24). These residues form either hydrogen bonds or salt bridges to stabilise the interaction.

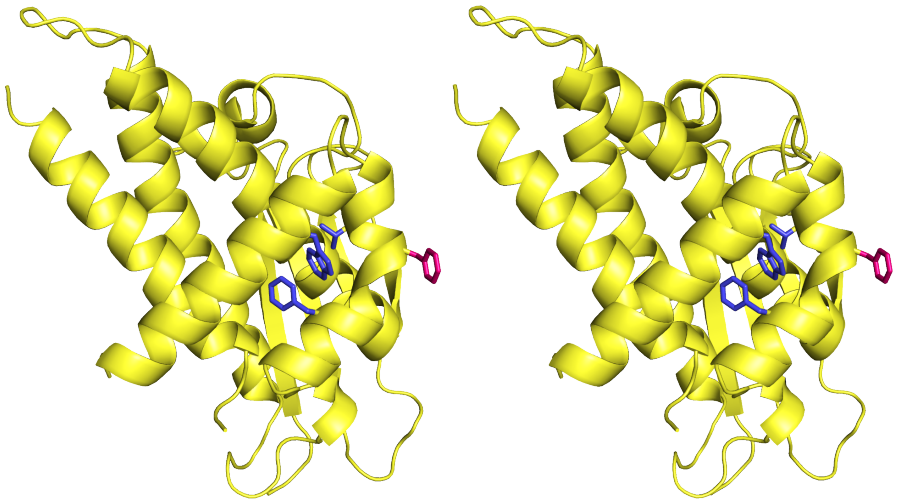


Figure 3.23: Stereo view of the protruding phenylalanine (pink, F656). F651, F658 and W669 (blue), point inwards between the two binding helices.

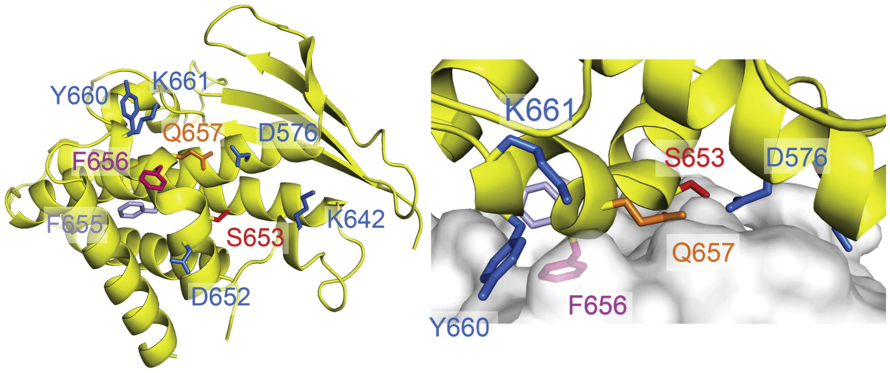


Figure 3.24: Left: The EPCR-binding surface on the HB3var03 CIDR (shown in yellow). Right: The interaction site; EPCR is represented as a translucent surface. CIDR residues that interact with EPCR are shown in stick representation and are named in labels of the same colour.

CIDR residue		EPCR residue	Interaction type
HB3var03	IT4var07		
D576	D573	R81	Hydrogen bond
		R85	Hydrogen bond
K642	K640	Y23 (backbone carboxyl)	Hydrogen bond
		D45	Hydrogen bond
D652	N650	S71	Hydrogen bond
S653	S651	R81	Hydrogen bond
F655	F653	Y154	Hydrophobic
		R156	Hydrophobic
F656	F654	Hydrophobic box	Hydrophobic
Q657	Q655	R81	Hydrogen bond
		Q85	Hydrogen bond
Y660	Y660	E86	Hydrogen bond
K661	K659	E86	Hydrogen bond

Table 3.8: A table of the interacting residues. The hydrophobic box consists of EPCR residues Q75, L79, L82, Q150, L151, Y154, R156 and T157.

3.6.2 Solution characterisation of the CIDR:EPCR complex structure

In order to check the consistency of the complex crystal structure with the complex in solution, I carried out small angle X-ray scattering experiments with both components alone and in complex. These experiments are carried out in solution and so are able to gain low-resolution structural information of the complex, free of any potential crystallisation artefacts. SAXS data were collected for the CIDR from HB3var03, EPCR and the complex. Linear Guinier plots show no aggregation in the samples used. The reciprocal-space datasets were transformed to real-space data, using DAMMIF and DAMMIN to create molecular envelopes that fit the data. Molecular weights can be calculated from the Porod volume of each species 3.9. The molecular weights observed agree for both the components alone and the complex, and the molecular envelopes are consistent with the structures, as shown by the fits below (Figure 3.25). The envelopes generated from these data are shown in Figure 3.26 and are consistent with the crystal structure.

Species	R_g (nm)	Volume (nm^3)	MW_{app} (kDa)	MW_{calc} (kDa)	D_{max} (nm)	Chi
CIDR	2.34	42.57	25.0	26.1	6.5	1.1
EPCR	1.99	38.89	22.9	20.7	5.5	1.2
Complex	2.63	76.45	45.0	46.1	7.4	2.7

Table 3.9: Summary table of the CIDR:EPCR SAXS.

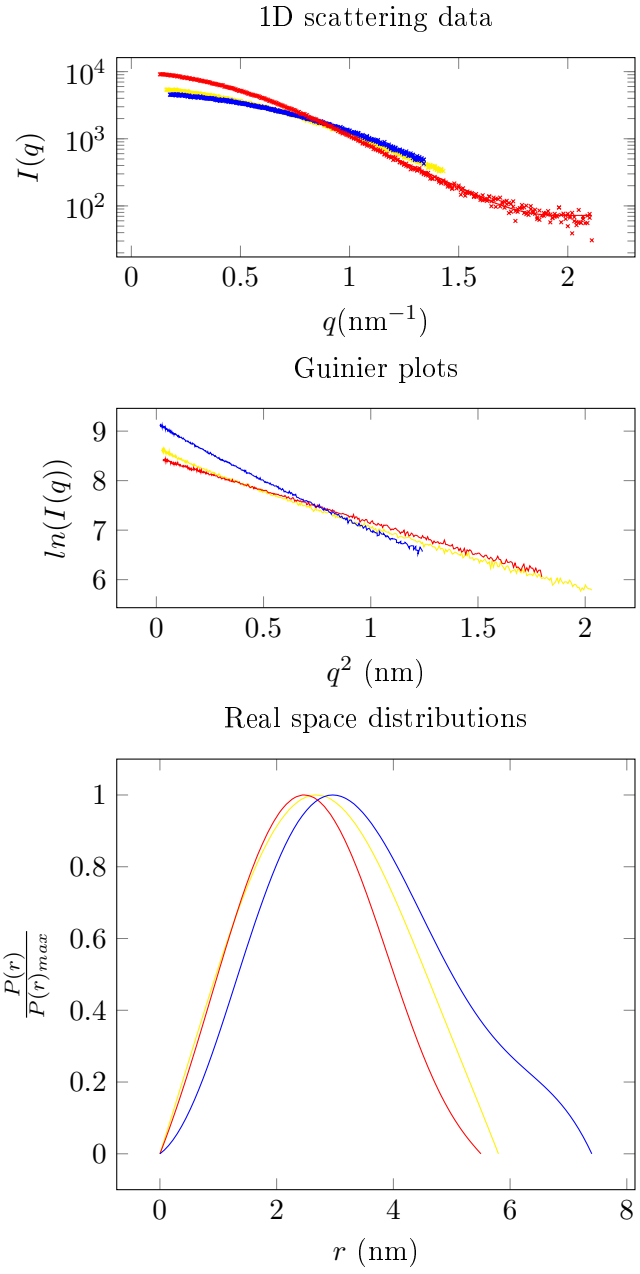
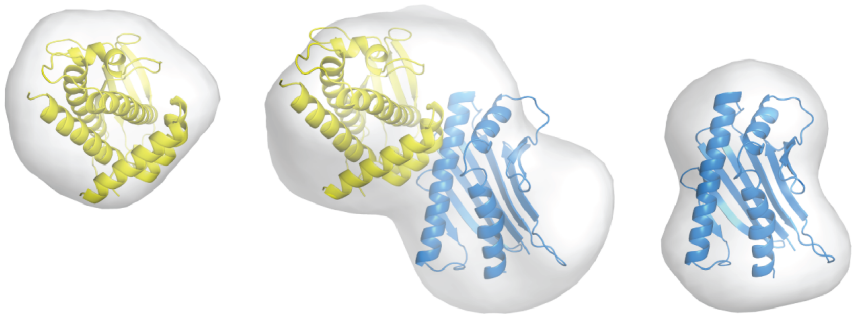


Figure 3.25: HB3var03 CIDR:EPCR SAXS data. In each plot EPCR is in red, HB3var03 CIDR in yellow and the complex in blue.



HB3var03 CIDR HB3var03 CIDR:EPCR EPCR

Figure 3.26: SAXS envelopes of the HB3var03 CIDR:EPCR complex and its components.

3.7 Discussion

CIDR construct design was guided by previous crystal structures of PA 89F5varO [189] and MC179 [137], which do not bind EPCR. Despite considerable sequence diversity (22% identity), the domain boundaries identified by FUGUE led to well defined constructs, tractable to crystallisation. The proteins expressed well in *E. coli*, but predominantly in the insoluble fraction. This was not entirely unexpected, given the number of disulfide bonds predicted by the primary sequence. These disulfide bonds could cross-link different molecules or pair incorrectly within a polypeptide chain, preventing the protein from folding properly. The on-column refolding procedure allowed for an acceptable amount of three of the five CIDR domains to be purified for crystallography. The yields varied between the different CIDRs, with DC13 CIDRs providing better yields than DC8 CIDRs. This could be due to an extra cysteine in the DC8 CIDRs. This cysteine will not be paired in these domains, and could potentially crosslink with other proteins leading to aggregation. Despite this, recombinant, monodisperse protein from these preparations was subsequently shown to be functional and folded. EPCR construct design was much more straight forward. Attempts to refold EPCR around an exogenously added lipid were not successful, suggesting perhaps that the lipid is necessary for correct folding. Indeed another MHC family member, MHC-II, is folded in the presence of a chaperone, the invariant chain, which associates instead of a lipid in the groove between the two α-helices [302], which are extracted as the protein is transported to the membrane. Both *Sf*9 cells and S2 cells contain the machinery for disul-

fide bond formation and post-translational modification, and expressed soluble, functional EPCR. The S2 expression system in particular gave an high yield of pure protein that was correctly folded.

The previous data on this interaction showed the two proteins to interact with nanomolar affinities with a slow dissociation rate [140]. Due to this I expected the two proteins to form a stable complex, suitable for crystallisation studies. Once the recombinant proteins had been expressed, I carried out a series of functional studies to show that the proteins were folded and suitable for crystallisation. The CD experiments were particular important, since the CIDR domains were refolded. They reveal a high degree of α-helicity for all the CIDR domains expressed. This matches the structural predictions made by FUGUE, based on the previous crystal structures. The initial SPR experiments showed that both the recombinant EPCR and the CIDRs bound with affinities comparable to their respective controls. In addition, these experiments showed that the interaction between EPCR and the CIDRs was not glycosylation dependent. This allowed me to deglycosylate the insect cell-produced EPCR before crystallisation trials, potentially eliminating a source of flexibility that could prevent crystal contacts forming.

The data produced by these first SPR experiments was poor, with data fitting poorly to kinetic models. As a result I tried using the CAPture chip system to couple EPCR to the chip. This system has two advantages over the CM5 system. Firstly, it allows the coupling of EPCR to the chip in the same orientation as it is found on an endothelial cell membrane. This presents all the binding sites on EPCR to the flow cell, as opposed to in the CM5 system, where EPCR is likely to be coupled in multiple orientations on the chip. Some binding sites in the CM5 system may have been partially available, potentially obscuring the true dissociation constant. The second advantage of the CAP system is that the chip surface regenerates completely. In the CAP chip regeneration strips all components from the chip, leaving the single-stranded DNA surface behind. In the CM5 chip system, some CIDR may not be stripped from the chip in the regeneration step, which could affect the samples in the next sample, or the EPCR surface could be damaged by the regeneration solution. I observed the CIDR domains binding to EPCR with sub-nanomolar affinities using the CAP system. These affinities are 20-fold tighter than those previously reported,

most likely due to the difference in coupling method. These tight affinities comprise of a fast on rate and a very slow off rate with the large majority of CIDR still bound to the chip after the 300 second dissociation time. These values are likely correct for the CIDR:EPCR interaction, as the kinetic fits, using a 1:1 binding model, are good. The extent to which these tight binding parameters are conserved is explored in Chapter 4. Due to the good quality of this data, this coupling method was used as standard in other chapters in this thesis.

This 1:1 binding model used in the SPR experiments was justified in subsequent SEC-MALLS and AUC experiments. These experiments all show that the CIDR domains and EPCR bind unambiguously as a 1:1 complex, with no higher order complexes observed. The complexes formed were stable over the lifetime of a SEC-MALLS experiment. This suggested that the recombinant proteins would remain a complex during crystallisation trials and that complexes should be set up at a ratio of 1:1.

Two CIDRs crystallised in complex with EPCR in different space groups, with different crystal contacts. Despite this, the fold of the CIDRs and the CIDR:EPCR interaction site was very similar in both structures. The crystal structures also fit into the envelopes generated in the SAXS data (Figure 3.26). This makes it likely that the complex structure is not an artefact of crystallisation, but rather represents the true structure of the complex in solution. The two crystallised CIDR domains contain the three conserved helices, as predicted in [170], completed with an N-terminal β-sheet segment and the binding helices in the variable portion between core helices 2 and 3.

The three conserved helical core is extremely similar to those in the varO [189] and MC179 [137] structures, with the HB3var03CIDR (chain A) core overlaying with these domains with RMSDs of 1.2 Å and 0.7 Å respectively (Figure 3.27). The varO domain also contains the β-sheet subdomain, which is in a different conformation to that of the EPCR-binding CIDRs, while the MC179 construct does not contain this sequence. The varO CIDR domain does not bind EPCR, belonging to a different CIDR subclass (CIDRγ). This is reflected in the crystal structures, as the variable portion between the second and third conserved helices is different in the varO structure. While the core helices align well, the conformation

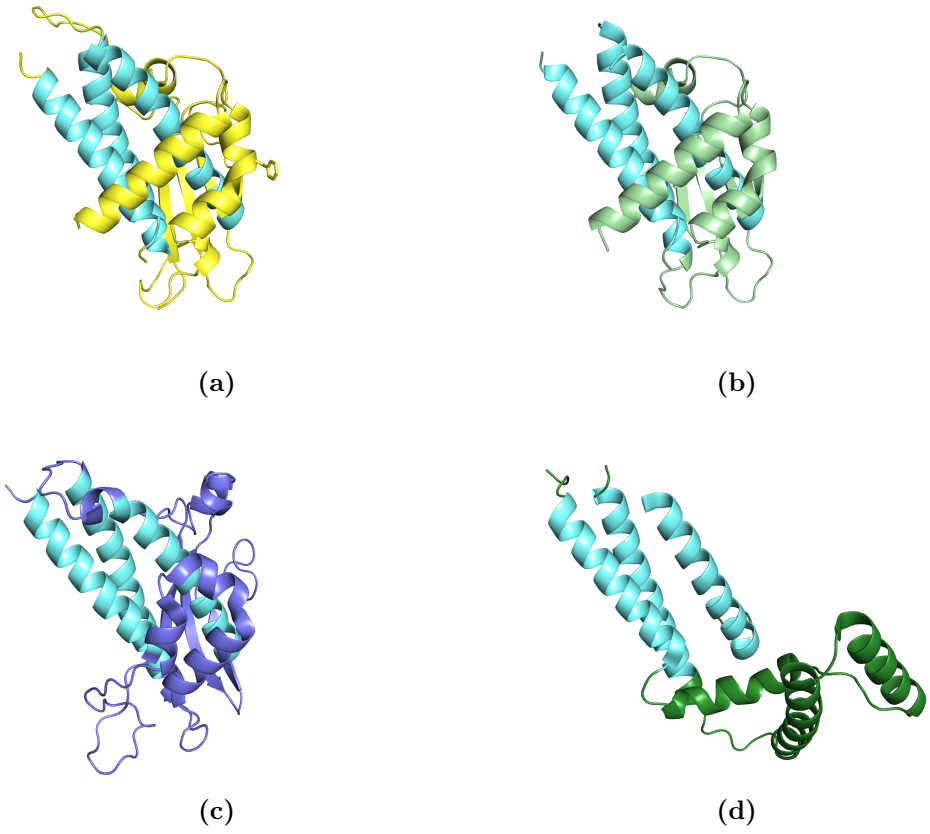


Figure 3.27: A comparison of the four published CIDR domain structures, aligned by their three conserved helices (shown in aquamarine). (a) The HB3var03 CIDR (chain A). (b) The IT4var07 CIDR. (c)The Palo Alto 87F5 CIDR. (d) The MC179 CIDR.

of the second helical segment of MC179 is completely different from the other crystallised CIDRs [137]. The MC179 construct binds to CD36 rather than EPCR, with the binding site yet to be identified. The structures reported here bring up the intriguing possibility that CIDR binding is at least partly mediated between the loop between the second and third core helices, which can be interchanged between different CIDRs through homologous recombination. Recombination in the *var* gene family that encodes the *Pf*EMP1 proteins is currently being explored [303], but evidence for recombination within CIDR domains has not yet been discovered.

The binding surface on both CIDRs consists on a hydrophobic pair of phenylalanines surrounded by a ring of hydrophilic residues, which is common in protein:protein interaction sites [304]. This interaction interface has a small surface area of 978 $Å^2$. Given this small size, the subnanomolar affinity is high [305]. This interaction site is validated and

characterised in mutagenesis studies in Chapter 4 (Section 4.4.2).

It is evident from the crystal structures solved that, while there is no density for the His-tags, which have been removed, EPCR is not completely deglycosylated. There is clear density on N119 in both structures for three consecutive sugar moieties, which correspond to two N-acetyl glucosamines and one mannose monomer from core glycan structure added by the S2 system. Since both endoHf and endoF3 cleave after the proximal glycan, this suggests that this particular site has not been deglycosylated in every EPCR in the crystal. In the density maps for the two complexes there is space for the rest of the paucimannose glycosylation though very little is resolved. This perhaps suggests flexibility in the conformation of the rest of the glycan. Other glycosylation sites are occupied with at least one well resolved N-acetyl glucosamine, but other sugar monomers may be present. As expected, no glycans are involved in the interaction, confirming the SPR data that showed the interaction between the two proteins is glycosylation-independent.

The two complex structures also explain the previous observation that Protein C cannot bind to EPCR once the CIDR has bound [140]. The previous crystal structure of EPCR to the Gla domain from Protein C [227] shows that Protein C binds to EPCR in a similar way to the CIDR domains. In their structure Protein C occupies 780 $Å^2$ of the surface of EPCR, including the hydrophobic box. Protein C interacts with EPCR through a phenyalanine, F4, which protrudes into the same hydrophobic box as in the CIDR:EPCR structures. This central interaction between Protein C and EPCR is supplemented by another two hydrophobic residues, and other hydrophilic contacts from amino acid side chains and calcium ions in Protein C. These interactions mainly occupy the segment of EPCR distal to the membrane.

The data presented here is consistent with the previous model of EPCR-related pathology presented in [140]. In this model the CIDR domain binding to EPCR prevents signals from Protein C from transducing into the cell via EPCR. The consequences of this include ICAM-1 expression, pro-inflammatory cytokine release and a decrease in endothelium barrier integrity. This model is yet to be extensively tested, but it does offer a tantalising glimpse of how the parasite could be manipulating the host to its advantage. Indeed other

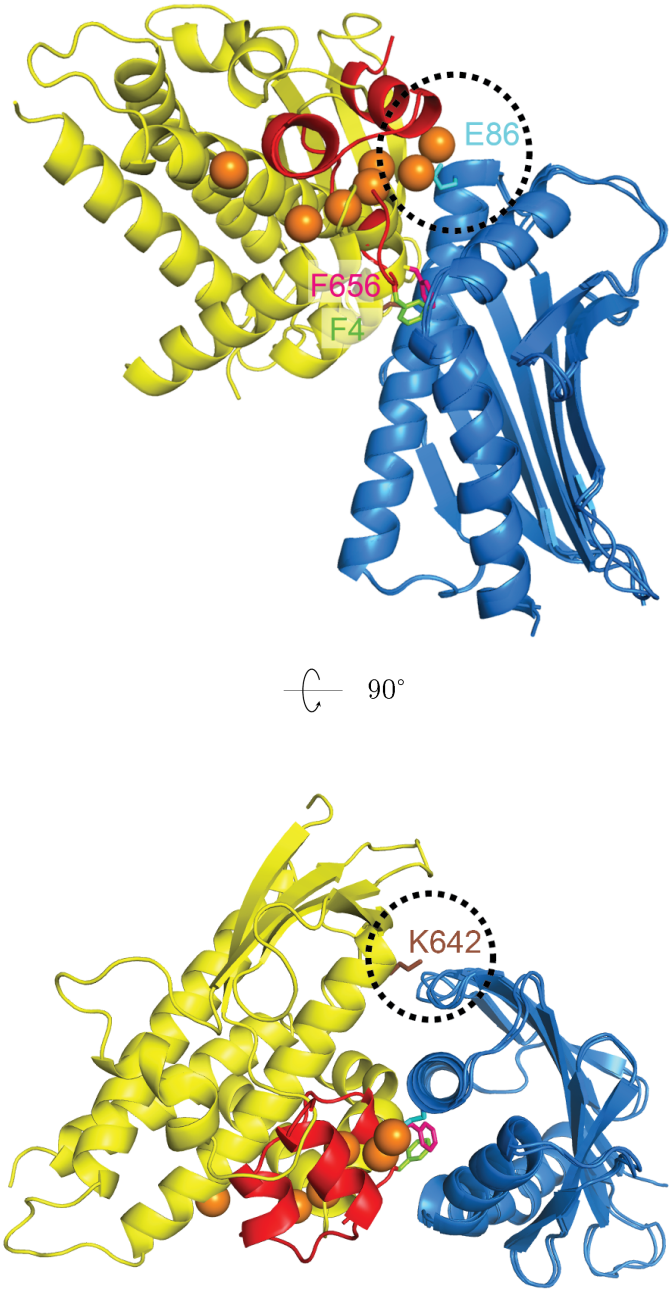


Figure 3.28: A comparison of Protein C and CIDR binding to EPCR. Protein C(red, calcium ions as orange spheres) is overlayed onto the CIDR (yellow) to show the overlapping binding sites.

*Pf*EMP1 domains bind ICAM-1, so the CIDR:EPCR interaction could either synergise or facilitate the DBL:ICAM-1 interaction. This possibility is explored in Chapter 5. The structures presented here show how the malaria parasite has evolved to bind in precisely the place as the natural ligand Protein C, preventing the interaction between Protein C and EPCR. If the model holds true, it suggests that preventing the CIDR:EPCR interaction could alleviate at least some of the symptoms of severe malaria.

The two ligands, Protein C and the CIDR domains, share 625 $Å^2$ of the binding surface on EPCR. This may be advantageous for the malaria parasite, as evolving to use an important existing binding site makes it difficult for the host to mutate EPCR to prevent binding of the parasite protein without adversely affecting interactions with host proteins. From a therapeutic point of view, designing a molecule that binds to EPCR to prevent CIDR binding without affecting the binding of Protein C would be difficult. The alternative, to design a molecule that can bind to the CIDR to block the interaction without also binding Protein C is explored in this thesis. This is as the interactions between Protein C and EPCR excluding the hydrophobic box are a mixture of protein-protein and protein-calcium interactions, substantially different from those in the CIDR structure. This raises hope that it is possible to block the CIDR interaction without also blocking the Protein C interaction.

Chapter 4

Conservation of fold and function in EPCR-binding CIDRα1 domains

4.1 Outline

Following the structural characterisation of the CIDR:EPCR complex, I sought to apply the knowledge gained to the vast sequence diversity that exists amongst EPCR-binding CIDRα1 domains. This was accomplished in collaboration with the Lavstsen group, University of Copenhagen, who had assembled a sequence database of 885 CIDRα1 domains. We initially tested representatives of the eleven identified CIDRα1 subclasses for EPCR-binding activity by SPR, showing that two of these subclasses did not bind EPCR. We then combined our data with the EPCR-binding CIDRα1 domain sequences to identify conserved features that allow the majority of these domains to bind EPCR. Mapping sequence diversity onto the structure showed that the most conserved residues point inwards into the structure, responsible for maintaining the fold of the CIDR domain. The surface of the CIDR is much less well conserved, with the sequence diversity suggesting that the CIDRα1 domains have diversified in an effort to evade the immune system.

In parallel to this, I analysed the binding of the representative set of CIDRα1 domains to assess the biophysical conservation of the interaction with EPCR. The resulting data

showed that all EPCR-binding domains tested bind EPCR with an conserved slow dissociation rate but with a variable association rate, suggesting that the slow dissociation rate is important to the interaction *in vivo*. Combining our knowledge of the binding interface with sequence alignments also allowed us to mutate residues at the interaction interface, and residues that stabilised the kinked helix structure to try and rationalise the sequence diversity we observed in the context of interaction strength. These studies reveal that the a conserved hydrophobic protrusion is important for the slow dissociation rate and that rest of the interaction site is conserved through chemical property rather than sequence.

4.1.1 Declaration

Sequence alignments and Shannon entropy scores were generated by the Lavstsen group, University of Copenhagen, using a database of *Pf*EMP1 sequences from laboratory strains [170], field isolates [177] and data from the MalariaGEN community project on *Plasmodium falciparum* population genomics [91]. In addition natural CIDRα1 domain variants, expressed in insect cells were provided by the Lavstsen group. These were for ELISA studies, carried out by the Lavstsen group, and SPR studies, carried out by myself.

4.2 The identification of EPCR-binding CIDRα1 domains

The library of 885 CIDRα1 sequences, assembled and analysed by the Lavstsen group, splits into eight subclasses, 1.1-1.8, with CIDRα1.5, 1.6 and 1.8 splitting into two further subclasses (1.5a, 1.5b, 1.6a, 1.6b, 1.8a & 1.8b; Figure 4.1). Representative members of each subclass were tested for binding to EPCR using two methods. Louise Turner, from the Lavstsen group, used ELISA to assess the binding of EPCR to members of each subclass (Figure 4.1, circles). These representative CIDR domains were a mixture of laboratory strains, parasite isolates and predicted domains from genome sequences. In parallel I tested the binding of the CIDRα1 domains to EPCR by surface plasmon resonance, using the biotin/streptavidin coupling method that I had optimised in Section 3.4.2. In the majority of cases, concentration series from 1 µM to 1 nM of the CIDR were flown over the

immobilised EPCR. In some cases, where weaker binding was observed, the concentration range was adjusted to 8 µM to 8 nM in order to calculate the kinetic constants. The binding curves are shown below, split by subclass and summarised in table 4.1.

Both the SPR and the ELISA data show that all CIDRα1 subclasses can bind to EPCR with the exception of the CIDRα1.2 and 1.3 subclasses. The K_D values ranged from sub-nanomolar to over 100 µM, with most of the CIDRα1 domains (20 out of 24 EPCR-binding CIDRs) binding with a K_D of under 100 µM.

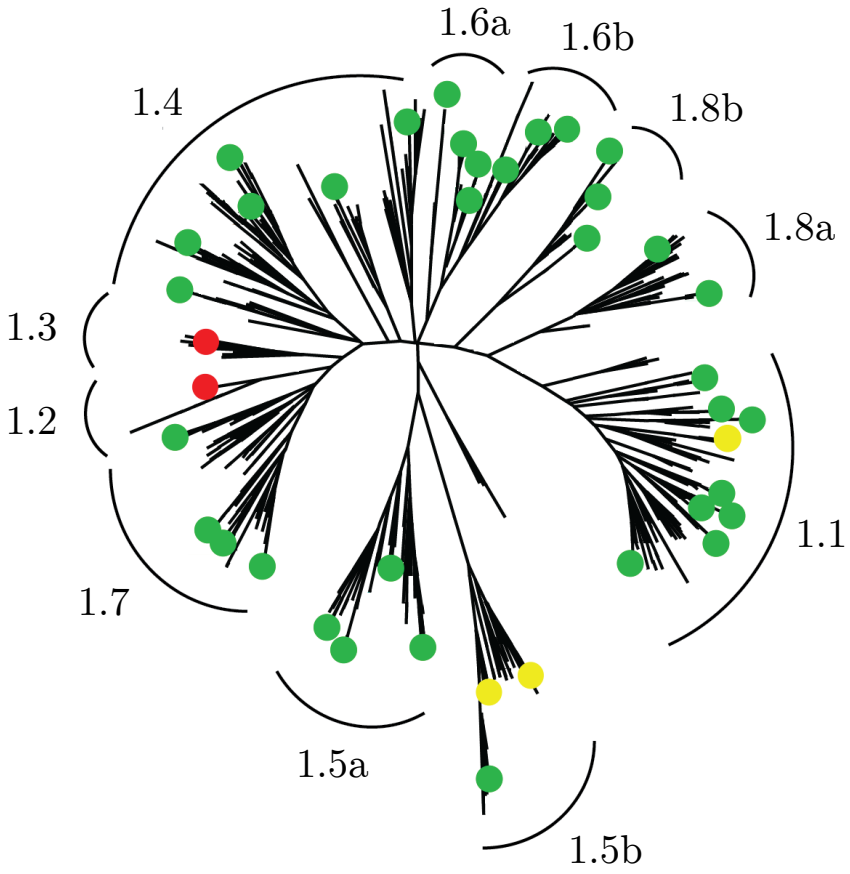


Figure 4.1: A maximum likelihood tree (bootstrap n=50) showing the sequence diversity of 885 identified CIDRα1 domains. Subclasses are grouped and labelled, with dots representing the degree of EPCR binding by ELISA (red = negative; yellow = weak; green = positive).

CIDRα1	Subclass	Range	K_D (nM)	k_{assoc} (M^{-1} s^{-1})	k_{dissoc} (s^{-1})
IT4var19	1.1	0.5 nM - 0.5 µM	16	2.27×10^4	3.64×10^{-4}
IT4var20	1.1	0.9 nM - 1 µM	0.37	3.32×10^5	1.23×10^{-4}
2110-3	1.1	0.9 nM - 1 µM	32	1.57×10^4	5.11×10^{-4}
RAJ116var08	1.1	0.9 nM - 1 µM	1.4	2.79×10^4	3.85×10^{-5}
ERS009959	1.1	7.8 nM - 8 µM	182	3.89×10^2	7.24×10^{-5}
HB3var1csa	1.2	0.9 nM - 1 µM	-	-	-
PFE1640w	1.3	0.9 nM - 1 µM	-	-	-
HB3var03	1.4	0.9 nM - 1 µM	0.37	5.32×10^5	1.97×10^{-4}
IT4var07	1.4	0.9 nM - 1 µM	1.3	3.70×10^5	4.68×10^{-4}
DD2var32	1.4	0.9 nM - 1 µM	0.3	1.06×10^6	3.43×10^{-4}
1965-2	1.5a	0.9 nM - 1 µM	54	1.89×10^4	1.03×10^{-3}
ERS010323	1.5a	0.9 nM - 1 µM	57	1.47×10^4	8.32×10^{-4}
1918-5	1.5b	7.8 nM - 8 µM	229	4.71×10^2	1.08×10^{-4}
1983-13	1.5b	7.8 nM - 8 µM	168700	2.80	4.70×10^{-4}
HB3var02	1.6a	7.8 nM - 8 µM	1400	9.75×10^2	1.39×10^{-3}
PFD1235w	1.6a	0.9 nM - 1 µM	8.8	7.55×10^3	6.62×10^{-5}
IT4var18	1.6b	1.8 nM - 2 µM	56	5.62×10^3	3.16×10^{-4}
ERS010570	1.6b	0.9 nM - 1 µM	15	3.51×10^4	5.31×10^{-4}
ERS010031	1.6b	0.9 nM - 1 µM	8.8	3.70×10^4	3.28×10^{-4}
IT4var22	1.7	0.9 nM - 1 µM	17	6.13×10^4	1.05×10^{-3}
1965-8	1.7	0.9 nM - 1 µM	3.5	5.92×10^4	2.10×10^{-4}
ERS010178	1.8a	0.9 nM - 1 µM	48	1.77×10^4	8.42×10^{-4}
MAL6P1.316	1.8a	0.9 nM - 1 µM	50	1.67×10^4	8.28×10^{-4}
2053-3	1.8b	0.9 nM - 1 µM	4	2.02×10^4	7.65×10^{-5}
1702-3	1.8b	0.9 nM - 1 µM	6	3.64×10^3	2.28×10^{-5}
ERS010532	1.8b	0.9 nM - 1 µM	24.2	6.55×10^3	1.58×10^{-4}

Table 4.1: SPR data summary table for CIDRα1 domains, tested for EPCR binding. No binding was observed for the CIDRα1.2 and 1.3 subclasses.

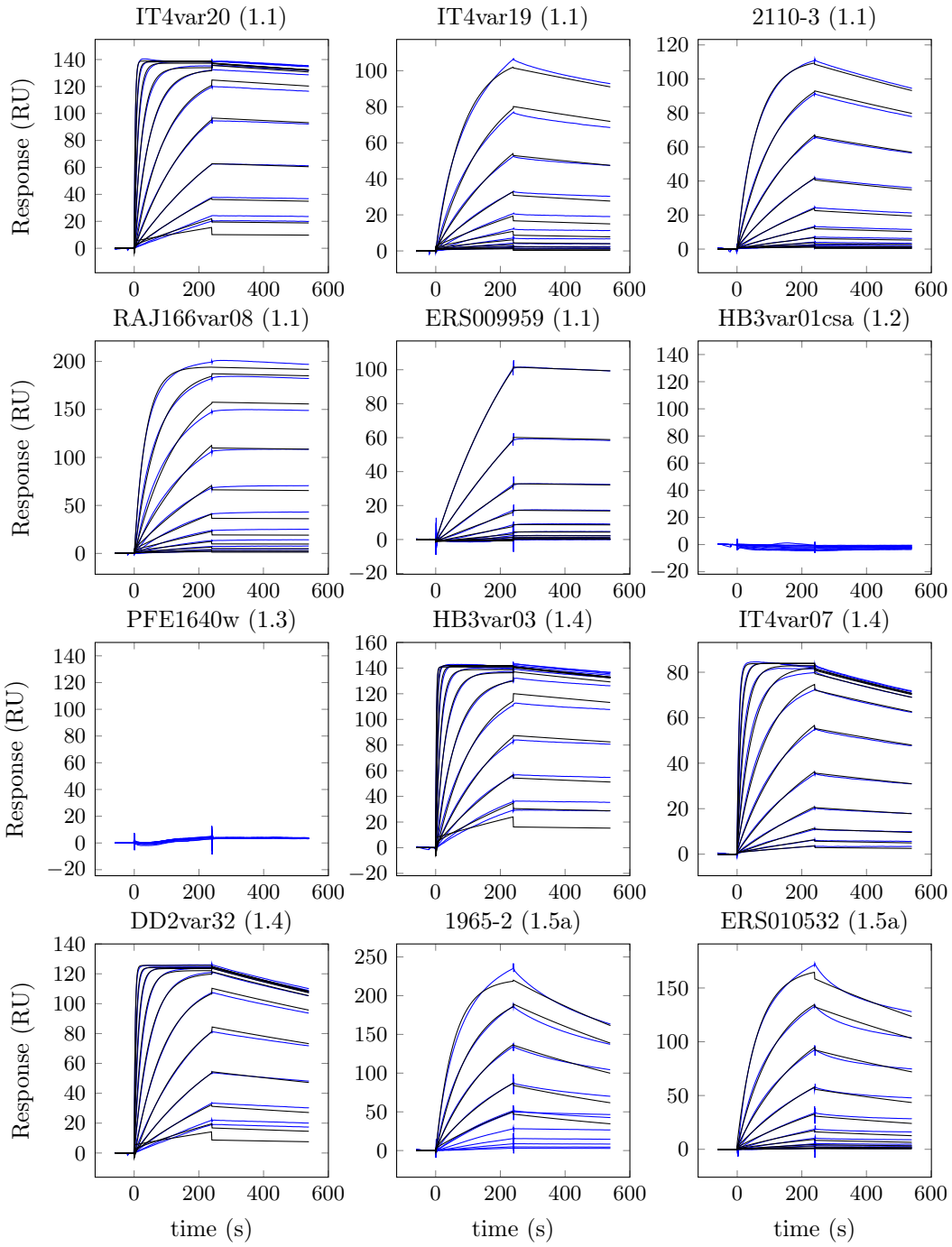


Figure 4.2: Surface plasmon resonance traces studying the variation in EPCR-binding by natural variants. This selection covers variants from CIDRα subclasses 1.1, 1.2, 1.3, 1.4 and 1.5a. Each dataset consists of binding data for a 12 sample 2-fold dilution series from 1 µM, with the blank-subtracted data shown in blue and kinetic fits shown in black.

4.2.1 Analysis of the natural variants

The SPR experiments in Section 4.2 showed that all CIDRs tested bind to EPCR, with the exception of CIDRα1 subclasses 1.2 and 1.3. Upon first examination there seems to be a considerable variation in binding affinity, with the thermodynamic dissociation constant K_D varying by up to four orders of magnitudes from 0.37 nM to 1.4 µM in the majority of CIDRs tested. In addition to this one CIDR, from field isolate 1983-13, showed very low binding activity (K_D = 170 µM). Driven by the observation of a consistently slow dissociation rate in these experiments plotted the spread of K_D values, $k_{association}$ and $k_{dissociation}$ values from these CIDR variants. I chose to plot these on a logarithmic axis (Figure 4.5), given that the kinetic and thermodynamic constants are related to the Gibb's free energy of interaction, ΔG, by the equation below, where $K_D = \frac{k_{dissociation}}{k_{association}}$.

$$\Delta G^{\circ} = -RT \ln K_D \tag{4.1}$$

This analysis of the 26 CIDR domains shows a skewed distribution of K_D and $k_{dissociation}$, whilst a normal distribution for $k_{association}$. The skew can be measured using the Fisher-Pearson standardised moment coefficient, G_1 (Equation 4.2, [306]). This measure of skewness looks at the difference between each datapoint and the mean value, using the average third central moment ($\frac{\sum_{i=1}^{n}(x_i-\bar{x})^3}{n}$) to measure the skew of the dataset, standardising the value and correcting for sample size. This value is compared to a set of critical values which show expected G_1 values for a given sample size. The G_1 value for the K_D values is 1.5, which is comfortably above the value for a statistically significant skew for this sample size (0.72, [306]). The corresponding Fisher-Pearson coefficients for $k_{association}$, 0.04, and $k_{dissociation}$, −0.6, suggest that this drive towards tighter affinities is mediated by $k_{dissociation}$, and not by $k_{association}$.

$$G_1 = \frac{\sqrt{n(n-1)}}{n-1} \frac{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^3}{n}}{s^3} \tag{4.2}$$

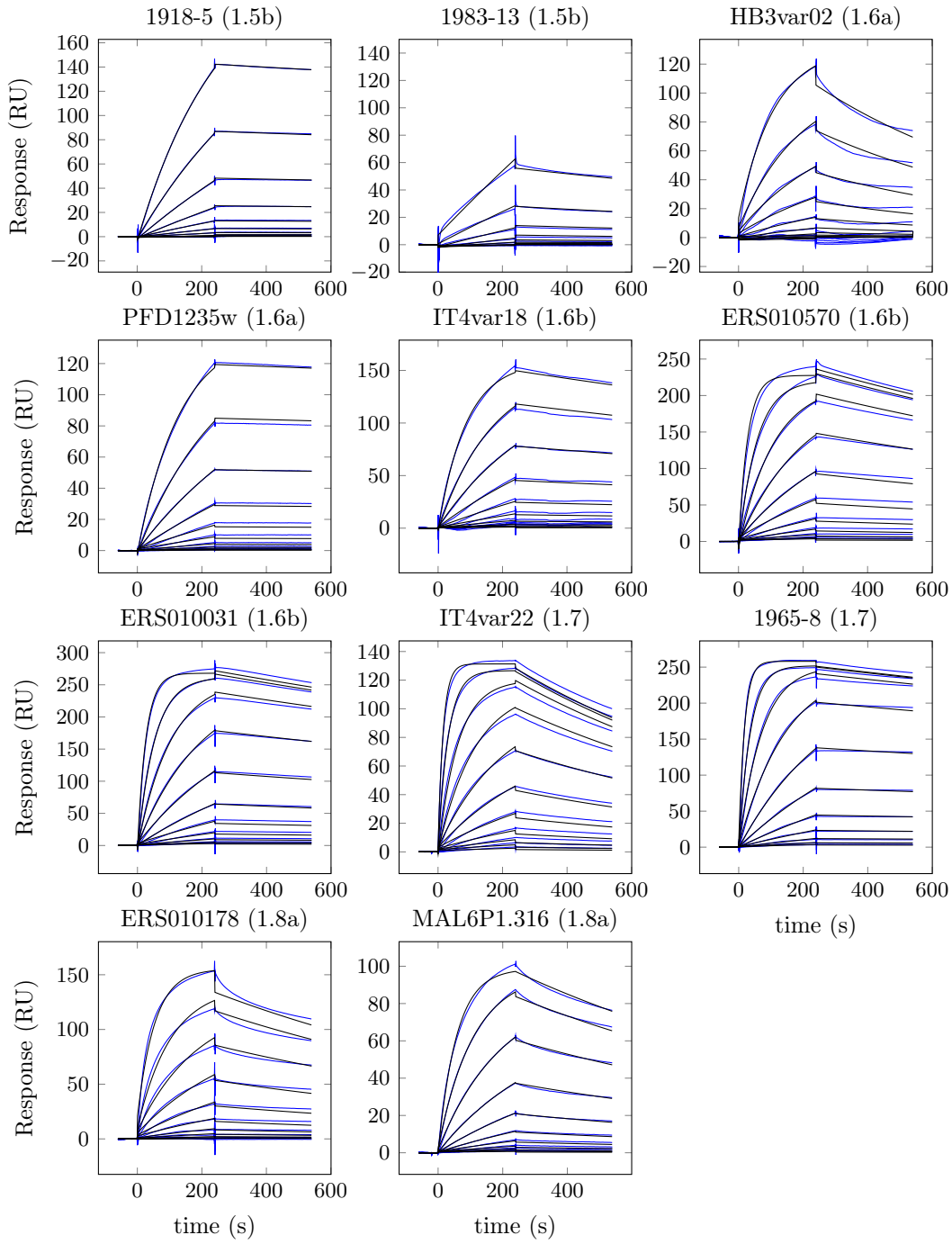


Figure 4.3: Surface plasmon resonance traces studying the variation in EPCR-binding by natural variants. This selection covers variants from CIDRα subclasses 1.5b, 1.6a, 1.6b, 1.7 and 1.8a. Each dataset consists of binding data for a 12 sample 2-fold dilution series from 1 µM, with the blank-subtracted data shown in blue and kinetic fits shown in black.

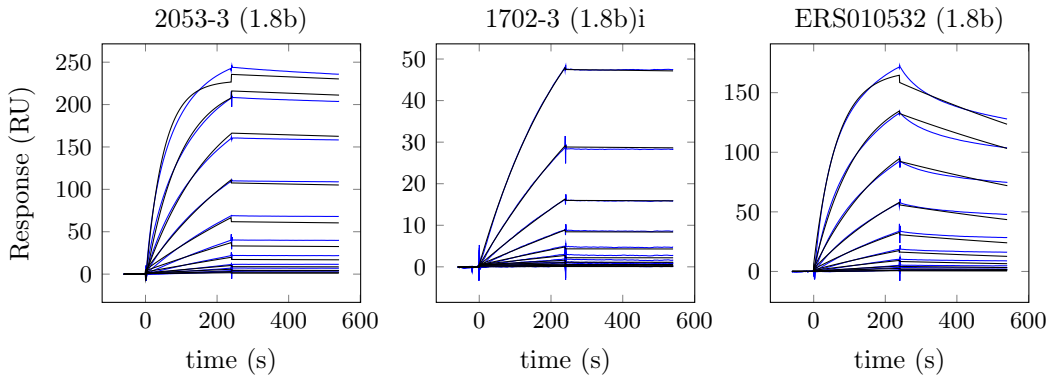


Figure 4.4: Surface plasmon resonance traces studying the variation in EPCR-binding by natural variants. This selection covers variants from the CIDRα1.8b subclass. Each dataset consists of binding data for a 12 sample 2-fold dilution series from 1 µM, with the blank-subtracted data shown in blue and kinetic fits shown in black.

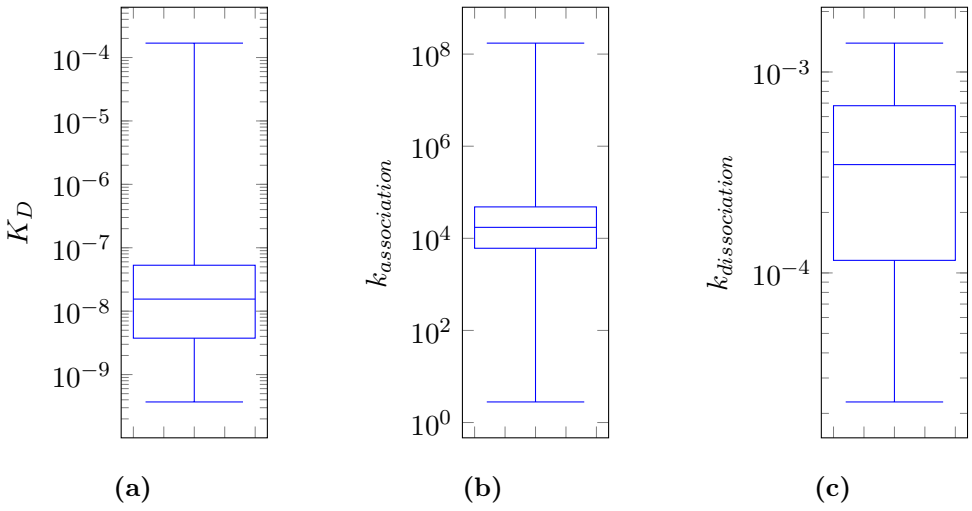


Figure 4.5: Box plots to summarise the K_D, (a), $k_{association}$, (b), and $k_{dissociation}$, (c) for the 24 EPCR-binding variants tested by SPR. This box plot shows the range of the data (upper and lower lines), the interquartile range (boundaries of the box) and the median value (central line, within the box).

4.2.2 ITC measurements of the CIDR:EPCR interaction

I chose to thermodynamically characterise the CIDR:EPCR interaction using isothermal titration calorimetry to gain insight into the entropy and enthalpy contributions to the interaction strength. I injected EPCR at 50 µM into a cell containing 5 µM HB3var03 CIDR (Figure 4.6). The calculated K_D is similar to that calculated by SPR (1.5 ± 0.8 nM). The interaction is driven by a negative enthalpy change (-52.3 ± kJmol^{-1}); the strength of approximately ten hydrogen bonds. This negative enthalpy is accompanied by an unfavourable entropy change (-6.7 ± 1.7 Jmol^{-1}K^{-1}), principally due to the loss of entropy due to two species binding together. This gives a final Gibbs free energy change of -10.5 kJmol^{-1}. The interaction stoichiometry, as expected, was 1:1 ($n = 1$).

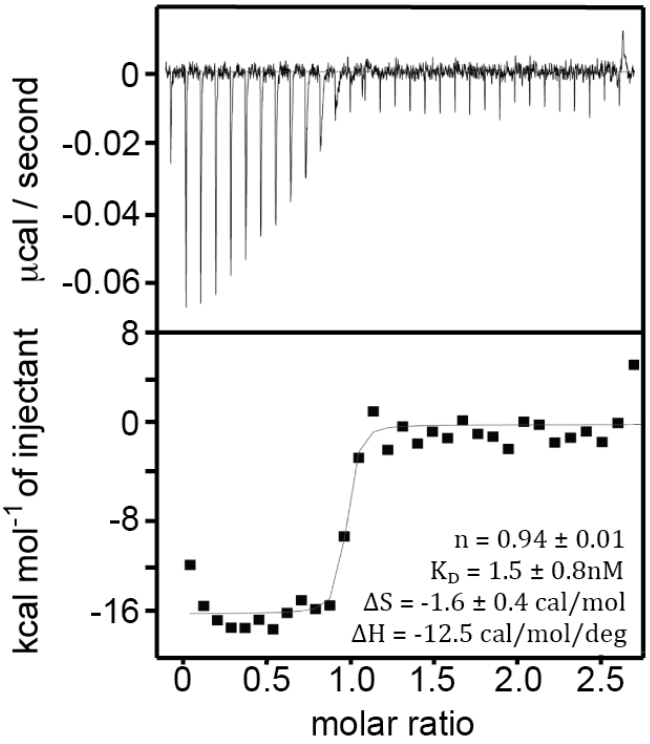


Figure 4.6: Isothermal calorimetry for the HB3var03 CIDR:EPCR interaction.

4.3 Examining sequence-structure conservation of the CIDRα1 domains

Once the EPCR-binding CIDRα1 subclasses were identified, sequence alignments were generated for the 737 CIDRs from these subclasses and mapped these alignments onto the elements of secondary structure, identified in the HB3var03 CIDR:EPCR crystal structure (Figure 4.7). Alignment of the CIDRs to a secondary structure map shows that the majority of the conserved residues are within secondary structural elements, with lower sequence conservation in loops. Moreover some CIDRα1 domains contain extra insertions or deletions within these loops. This analysis was of limited use as the conservation of the tertiary structure could not easily be identified using these alignments. In order to assess the conservation of the tertiary structure of the protein, we hence quantified this sequence conservation and plotted these values onto the HB3var03 CIDR structure.

4.3.1 Mapping sequence diversity onto the CIDR:EPCR structure

We decided to use the Shannon entropy (SE) of residue property to investigate the conservation of the residues in the CIDRα1 domain [308] on the HB3var03 structure. This equation (Equation 4.3) studies the unpredictability of a residue type, i at a given position l. There are 6 residue classes considered, aliphatic (AVLIMC), aromatic (FWYH), polar (STNQ), positive (KR), negative (DE) and special (GP), and an extra category for gaps in sequence. A Shannon entropy score of 0 signifies absolute conservation, and a score of 0.84 signifying an equal chance of any of the seven groups, the theoretical maximum score. The resulting entropy value implies an evolutionary pressure to keep a particular class of residue at a given position. The Shannon entropy scores reveal that 45 of the 218 residues in the structure are conserved in their chemical property, with a Shannon entropy score of below 0.1. The rest of the residues fall between 0.1 and 0.9 in the distribution shown in figure 4.8.

$$s(l) = -\sum_{i=1}^{7} p_i(l) \log p_i(l) \tag{4.3}$$

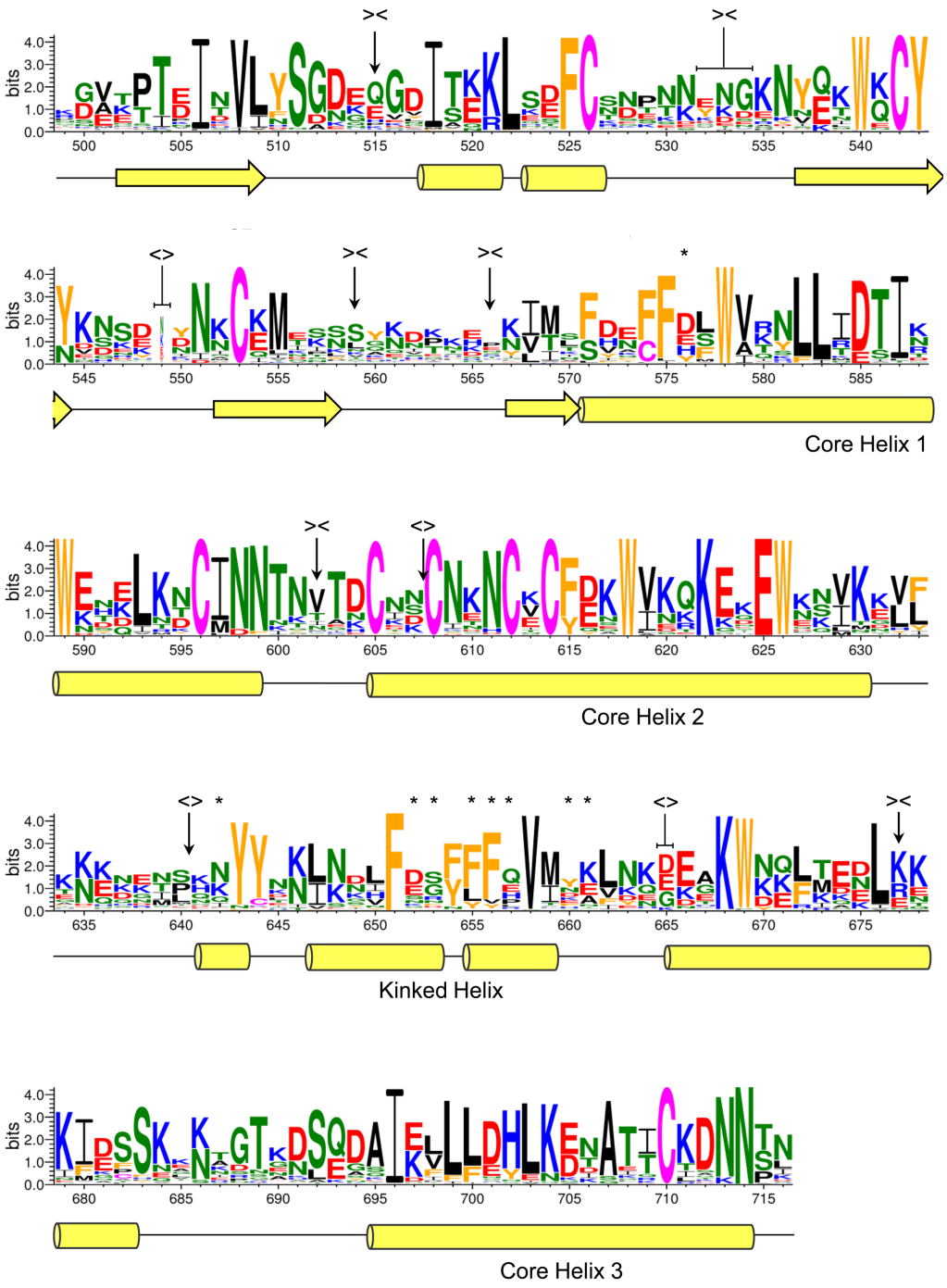


Figure 4.7: Weblogo [307] showing sequence conservation at each position in the CIDR domain. Numbering shown is from the HB3var03 sequence, with HB3var03 CIDRα1 structural elements shown underneath. Insertions and deletions are marked by <> and >< respectively. Residues identified as binding residues are starred.

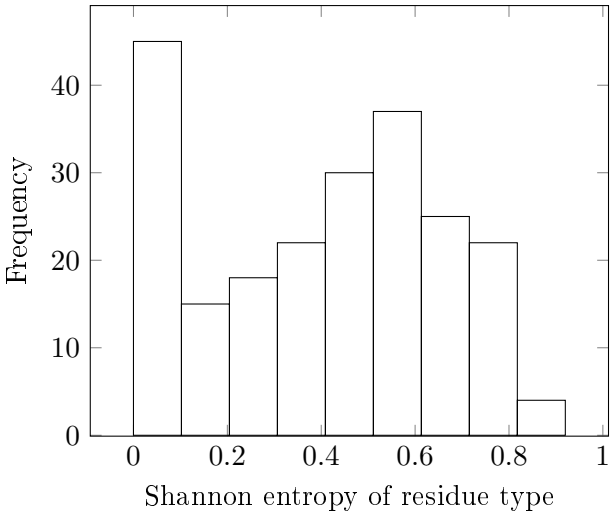


Figure 4.8: Histogram showing Shannon entropy of residue property scores for the EPCR-binding CIDRα1 domains.

Mapping the Shannon entropy of residue property scores onto the structure reveals that the inwards-facing residues are well conserved, with Shannon entropy scores below 0.12. All but one of the completely conserved residues in the structure point inwards into the structure (Figure 4.9). This set consists of 3 large bulky tryptophans and 7 cysteines[1], forming the hydrophobic core of the CIDR fold; K622[2] and E625, which form a buried salt bridge; N714 located at the C-terminal domain, which is surface exposed; and F651, which could help to pack the binding arm aginst the conserved three-helical bundle. 47 residues have a Shannon entropy below 0.12; this corresponds to a probability for residue type i at position l of 0.95. 39 of the 47 residues belong to the aromatic and aliphatic groups, contributing to the stability of the CIDR fold. The other residues mostly contribute to salt bridges within the structure, with only two residues, K668 and N714 pointing outwards from the structure. K668 is located on the back of the binding arm at the boundary between the hydrophobic core and the surface, where it could make transient interactions with E671 and T591. N714 is completely surface exposed, located at the end of the third core helix.

Outwards-facing residues have higher Shannon entropy scores, with the majority of the surface (83% of the solvent accessible surface area) having a score over 0.3. This means that these residues are less well conserved in chemical property. Residues with a score between

[1] one of these cysteines pairs with C542, which has a Shannon entropy score of 0.005
[2] NB: Unless stated, all numbering is using the HB3var03 sequence

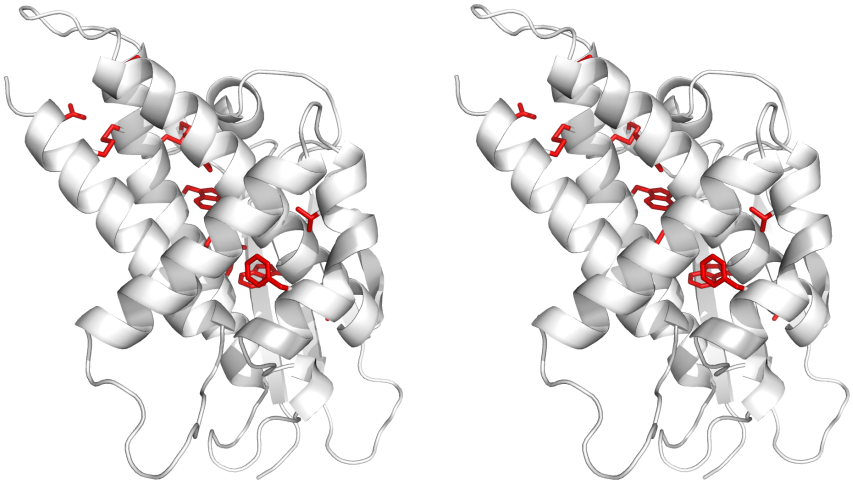


Figure 4.9: Stereo view of the HB3var03 structure showing the totally conserved residues (red). These residues face inwards into the CIDR fold.

0.3 and 0.5 generally have two predominant residue types at that particular positions. Surface residues with a score higher than 0.5 make up 62% of the surface of the CIDR. This score indicates a low level of conservation; a score of 0.5 is consistent with an equal chance of a polar, positive or negative residue at a solvent-exposed position. A mere 16% of the surface of the CIDR has a SE score under 0.3. The structure reveals that these residues are generally located in cavities and clefts in the surface, with the exception of the phenylalanine that is central to the EPCR interaction, and a couple of patches near the C-terminus of the protein (Figure 4.10).

4.3.2 Conservation of the binding site residues

The Shannon entropy score of the binding residues is summarised in table 4.2. The central phenylalanine, which protrudes into the groove on EPCR is well conserved with a Shannon entropy score of 0.17. This is due, in part to a large occupancy of this position by phenylalanine, and to a lesser extent tyrosine, which are both aromatic residues. One CIDRα1 subclass, 1.5b, has a 60% occupancy of valine at this position, in the aliphatic residue type group. If we exclude this subclass from Shannon entropy calculation, the property entropy score decreases to 0.04.

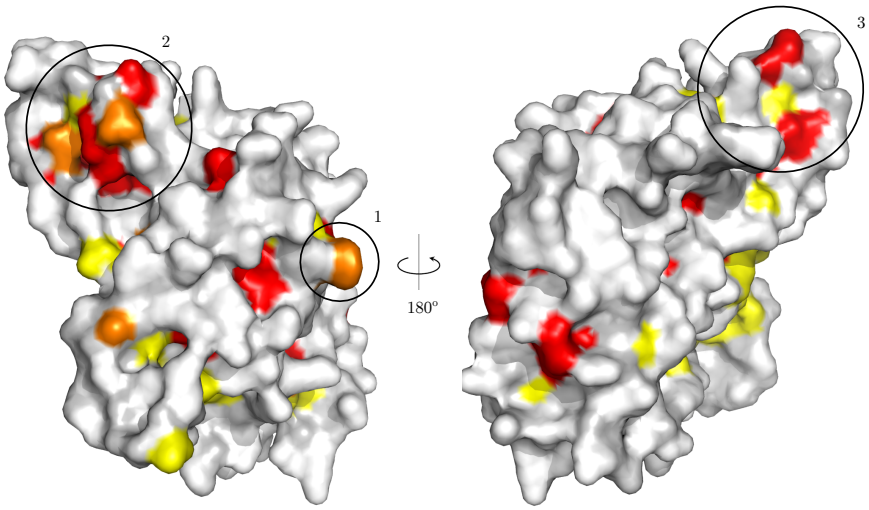


Figure 4.10: Surface representation of the HB3var03 CIDR Shannon entropy scores. The majority of residues have a score of above 0.3 (white). More highly conserved residues are coloured in yellow ($0.2 < SE < 0.3$), orange ($0.1 < SE < 0.2$) and red ($0 < SE < 0.1$). The protruding phenylalanine (1) and two conserved C-terminal conserved patches (2, 3) are circled.

The adjacent phenylalanine is less well conserved, with a property entropy score of 0.31. This is largely due to a high occupancy of phenylalanine and leucine at this position. Since both of these residues make hydrophobic contacts with the surface of EPCR at this position, this substitution is unsurprising.

Slightly more surprising is the lack of conservation of the rest of the binding site. The hydrophilic residues that supplement the hydrophobic patch are not well conserved. Of the seven hydrophilic residues that directly contact EPCR, only one has a property entropy score below 0.5 (D576, 0.4), suggesting these residues are not well conserved. All the hydrophilic residues positions are occupied by several different classes of residue with varying size and charge, as shown in the alignment in figure 4.11. D576, for example, is occupied by a negatively charged residue in 62% of CIDRα1 domains, but a neutral histidine in 22% and the bulky, aromatic tyrosine in 14% of cases. The only characteristic that is conserved amongst these residues is the ability to form hydrogen bonds.

HB3var03 residue	Property entropy score
D576	0.40
K642	0.55
D652	0.53
S653	0.74
F655	0.31
F656	0.17
Q657	0.81
Y660	0.74
K661	0.65

Table 4.2: Property entropy scores for the residues responsible for interacting with EPCR. Residues and residue numbering are from HB3var03.

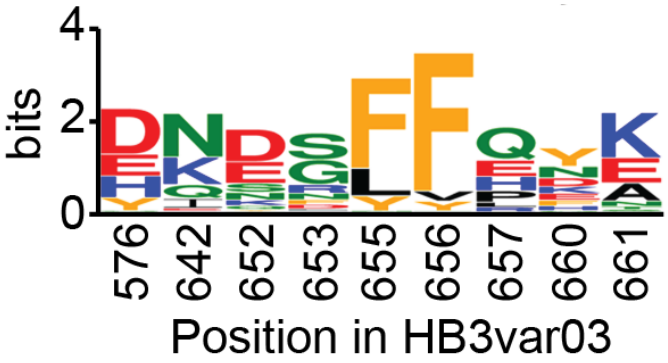


Figure 4.11: Weblogo [307] showing the sequence diversity of binding residues in the EPCR-binding CIDRα1 domains

4.4 Probing the interaction strength between CIDR domains and EPCR

I probed the interacting residues identified above to assess their importance in maintaining the high affinity of the interaction and to probe whether different mutants contribute differentially to the off and on rates. Using these mutants I explored some of the sequence diversity observed in the natural variants.

4.4.1 Site-directed mutagenesis of the CIDRα1 domain

I mutated residues involved in the binding site, guided by the sequences of EPCR-binding CIDRs, to test their importance in the interaction between the CIDRα1 domain and EPCR, and to validate the structure. Residues mutated included those that directly contacted EPCR, or were involved in the structure of the kinked helix.

For each mutant, I designed appropriate primer pairs to change the parent vector at the mutation site (Appendix A). These primers were used in a Quikchange PCR, with the resulting nicked vectors transformed into ultra-competent *E. coli*. Correctly mutated constructs were expressed in B834 cells and purified using the same protocol used to express and purify their parent constructs. The rationale behind this was that I did not expect any of the constructs to have large effects on the overall CIDR fold. This was justified as all of the following constructs expressed to comparable yields.

4.4.2 Assessing the binding of CIDRα1 mutants to EPCR

A 2-fold dilution series of each mutant from 1 µM was flown over the EPCR-covered chip surface; data collected from these experiments is summarised in table 4.3.

HB3var03	Range	K_D (nM)	k_{assoc} (M^{-1} s^{-1})	k_{dissoc} (s^{-1})
wt	0.9 nM - 1 µM	0.37	5.32×10^5	1.97×10^{-4}
F651A	15.6 nM - 250 nM	2.5	1.60×10^5	3.94×10^{-4}
S653F	0.9 nM - 1 µM	0.6	1.2×10^6	7.8×10^{-4}
S653FQ657R	0.9 nM - 1 µM	27	1.7×10^5	4.6×10^{-3}
F655L	0.9 nM - 1 µM	43	1.45×10^5	6.30×10^{-4}
F655Y	1.8 nM - 500 nM	0.5	1.10×10^6	5.72×10^{-4}
F656A	0.9 nM - 1 µM	129	1.16×10^5	1.5×10^{-2}
F656R	0.9 nM - 1 µM	61	5.23×10^3	3.20×10^{-4}
F656V	0.9 nM - 1 µM	1.5	1.45×10^5	2.23×10^{-4}
F656Y	0.9 nM - 1 µM	1.9	3.29×10^5	3.79×10^{-4}
Q657K	0.9 nM - 1 µM	72	6.89×10^4	4.99×10^{-3}
Q657R	0.9 nM - 1 µM	90	2.30×10^5	2.08×10^{-2}
V658A	0.9 nM - 1 µM	37	1.94×10^4	7.19×10^{-4}
W669S	15.6 nM - 1 µM	11	4.01×10^4	4.49×10^{-4}

Table 4.3: SPR summary table for HB3var03 mutants.

4.4.2.1 Mutation of the central phenylalanines

The majority of mutants made changed the central phenylalanine (F656) to other residues found in those positions in naturally occurring EPCR-binding CIDRs, as well as to alanine. Changing the phenylalanine to a tyrosine has little effect on the binding constants of the interaction. This is not unexpected, as this change maintains the hydrophobic ring that can interact with the hydrophobic box. The F656V mutant also had little effect on the binding - the K_D increases by a factor of 5 (Figure 4.12). This mutant would have a lower surface area to interact with the hydrophobic box, so it is surprising that the kinetics and thermodynamics are so similar.

Mutating F656 to alanine had by far the largest effect on the interaction strength, increasing the K_D by a factor of 500. This increase is mainly due to the large change in the kinetic dissociation constant, which increases by a factor of 100 compared to a kinetic association constant drop by a factor of 5 (Figure 4.12). This suggests that a residue at this position must have a large, bulky hydrophobic side chain to bind to EPCR. Mutation of F656 to arginine, which is positively charged, increases the K_D by a factor of 200, mainly due to a slower association rate, as the dissociation rate is still slow. The slow dissociation rate is consistent with the hydrophobic portion of the arginine chain interacting to a limited extent with hydrophobic groove, while the slower association rate could be due to the charge on the arginine head group.

Mutating the second of the hydrophobic pair to leucine or tyrosine, commonly found in other domain subclasses, did not significantly change the K_D. This is expected as subclasses with these variations bind EPCR with high affinity, and as neither of these residue substitutions constitutes a big change in properties.

4.4.2.2 Mutation of Glutamine 657 to explain the non-binding CIDRα1 domains

The two subclasses that do not bind EPCR (CIDRα1.2 and 1.3) had fairly typical binding sites, with the exception of a conserved lysine at the position equivalent to Q657 (Figure

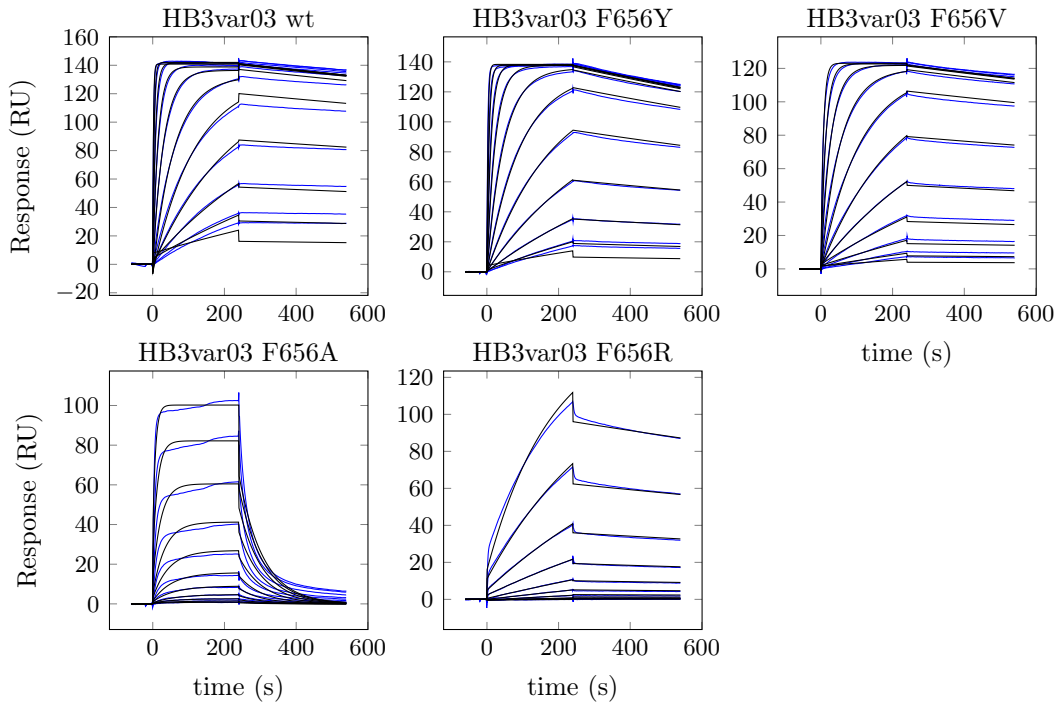


Figure 4.12: Surface plasmon resonance traces comparing phenylalanine mutants with wildtype. Each dataset consists of binding data for a 12 sample 2-fold dilution series from 1 µM, with the blank-subtracted data shown in blue and kinetic fits shown in black.

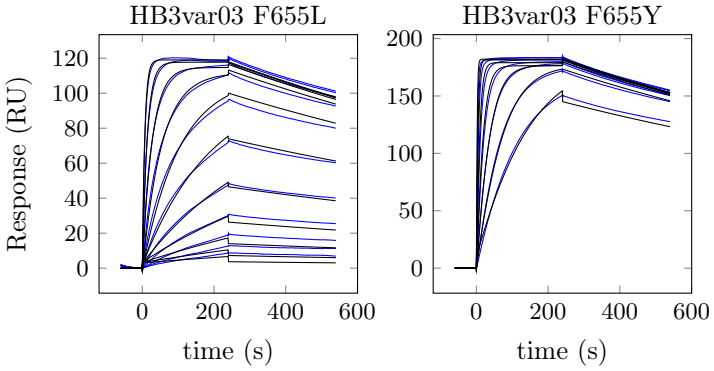


Figure 4.13: Surface plasmon resonance traces comparing phenylalanine mutants with wildtype. Each dataset consists of binding data for a 12 sample 2-fold dilution series from 1 µM, with the blank-subtracted data shown in blue and kinetic fits shown in black.

4.14). Sequence alignments of the binding residues show that this lysine represents the only major difference between these two subclasses and other CIDRα1 domains. I hypothesised that mutating this residue in HB3var03 would decrease the affinity of the interaction with EPCR. This mutant did bind with a lower affinity than the wildtype, with a K_D value of 72 nM. This increase in K_D was mediated through a 10-fold drop in the $k_{association}$ and a 20-fold increase in the $k_{dissocation}$. However, the interaction was not completely abolished, suggesting a role for other residues.

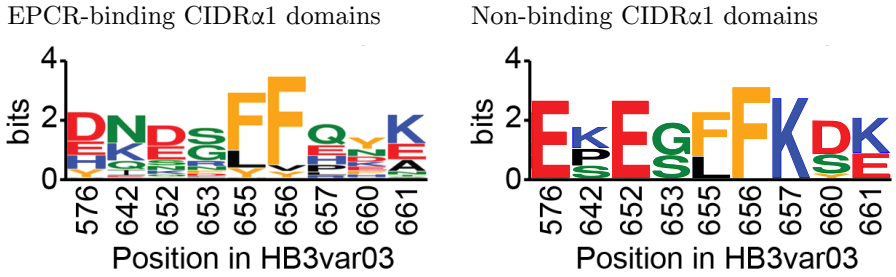


Figure 4.14: Weblogo showing the sequence diversity of residues positions equivalent to the HB3var03 binding residues in binding (left) and non-binding (right) CIDRα1 domains.

A similar mutation, Q657R, similarly increased the K_D, mainly mediated through an increase in $k_{dissociation}$. This residue variation is present in the subclass CIDRα1.6b, which binds EPCR with nanomolar affinities. Examination of the 1.6b subclass revealed second substitution at position 653. At this position over half of the 1.6b domains had another large, hydrophobic residue (typically phenylalanine or isoleucine). I hypothesised that an equivalent mutation in the Q657R mutant, S653F, could partially restore binding functionality of the Q657 mutant by increasing the size of the hydrophobic patch in the centre of the interaction site. Indeed the S653FQ657R mutant bound with a lower kinetic dissociation constant than the single Q657R mutant alone (4.6×10^{-3} versus 2.08×10^{-2}). Testing the S653F mutant in isolation reveals a lower K_D and a higher $k_{dissociation}$ than the wild type, suggesting that this mutation is only advantageous to partially restore binding function to the Q657R mutant, rather than to increase the affinity in general.

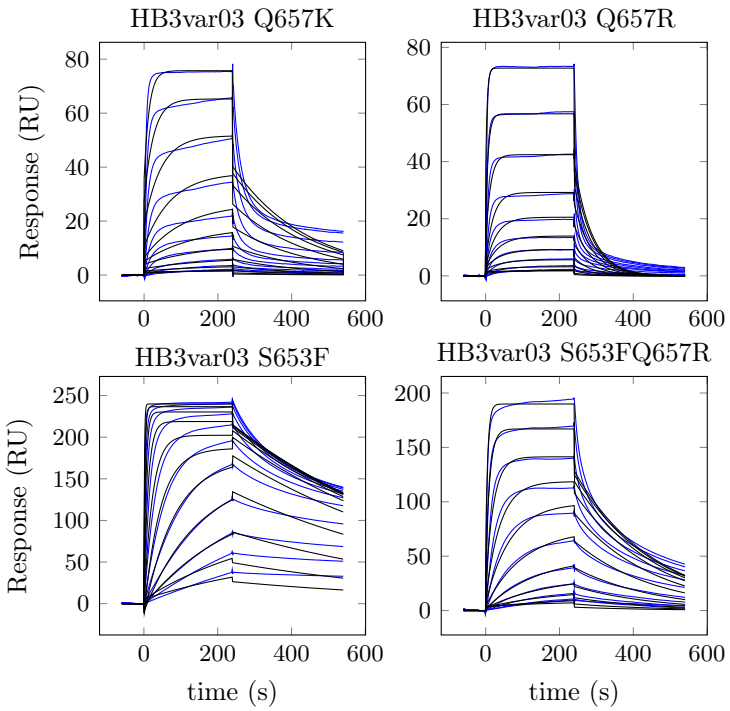


Figure 4.15: Surface plasmon resonance traces showing the effect of Q657 mutations (top panels), and the effect of the rescue mutation, S653F (bottom panels). Each dataset consists of binding data for a 12 sample 2-fold dilution series from 1 µM, with the blank-subtracted data shown in blue and kinetic fits shown in black.

4.4.2.3 Mutation of residues that stabilise the kinked helix

The structure of the kinked helix at the binding interface seems to be stabilised by residues F651, V658 and W669, which are well conserved in the CIDRα1 domain class. We chose to mutate the F and V to an alanine, while mutating the W to serine, as found in the CIDRα1.1 domain from ERS009959, which binds EPCR with a K_D of 182 nM. These mutants bound slightly less well to EPCR, with K_D values of between 1.5 nM and 11 nM. In all three structural mutants, the change in $k_{association}$ is greater than the change in $k_{dissociation}$.

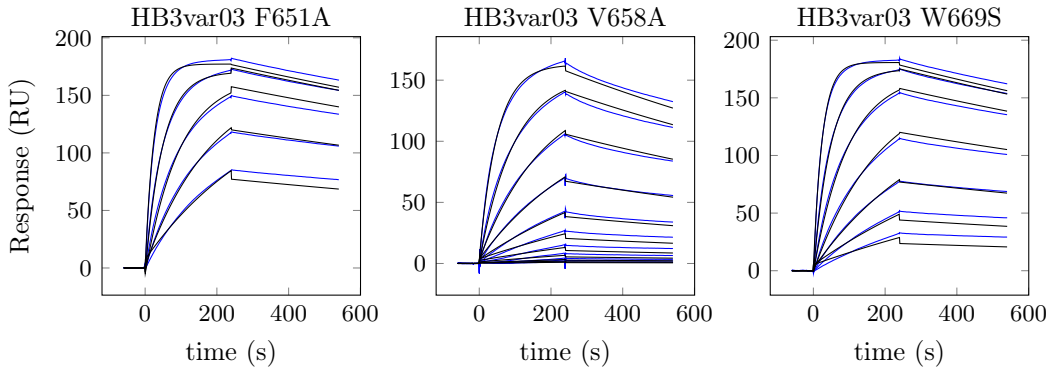


Figure 4.16: Surface plasmon resonance traces showing the effect of mutations of the three kink-forming residues to small amino acids. Each dataset consists of binding data for a 12 sample 2-fold dilution series from 1 µM, with the blank-subtracted data shown in blue and kinetic fits shown in black.

4.5 Discussion

As *Pf*EMP1 is exposed to the immune system, *Plasmodium falciparum* parasites contain around sixty *var* genes, switching between variants to evade the immune system. In order to maintain interactions with host ligands, different *Pf*EMP1 proteins from the same parasite can contain diverse copies of domains that can bind to the same ligand; the EPCR-binding CIDRs from IT4 variants 7 and 20 have 44% sequence identity. Understanding the features that facilitate the conservation of ligand binding between variants will be key to determining whether a therapy can be designed to block the interaction.

The sequence diversity in the CIDRα1 domain class is remarkable given that the large

majority of these CIDR subclasses bind EPCR. This diversity is driven by the need for the *Plasmodium falciparum* parasite to evade the adaptive immune system, but it is restricted by the molecular determinants required to bind EPCR. Our ELISA and SPR studies on natural variants show that all of the CIDRα1 subclasses bind to EPCR, with the exception of the CIDRα1.2 and 1.3 subclasses. These two subclasses are associated with *Pf*EMP1 *var1* genes. These genes are considered pseudo-var genes, as they seem not to be expressed at the surface of infected erythrocytes [174, 173].

Mapping the diversity of the CIDRα1 domains onto the structures provided insight into both the CIDR fold and the surface residues, showing the degree to which the binding surface on the CIDR is conserved. The conservation was plotted using the Shannon entropy of residue property, which categorises residues into classes based on property. This method is very simple compared to other algorithms, which can include differential penalties based on an amino acid substitution matrix, such as the BLOSUM series. Given the low sequence identity observed within the CIDRα1 domain class, this use of Shannon entropy is justified as other algorithms are designed to work with homologues with a higher sequence identity, while Shannon entropy of residue property is designed to look for chemical conservation. To account for this the binding site residues were carefully examined when drawing conclusions, looking at residue substitutions at each position.

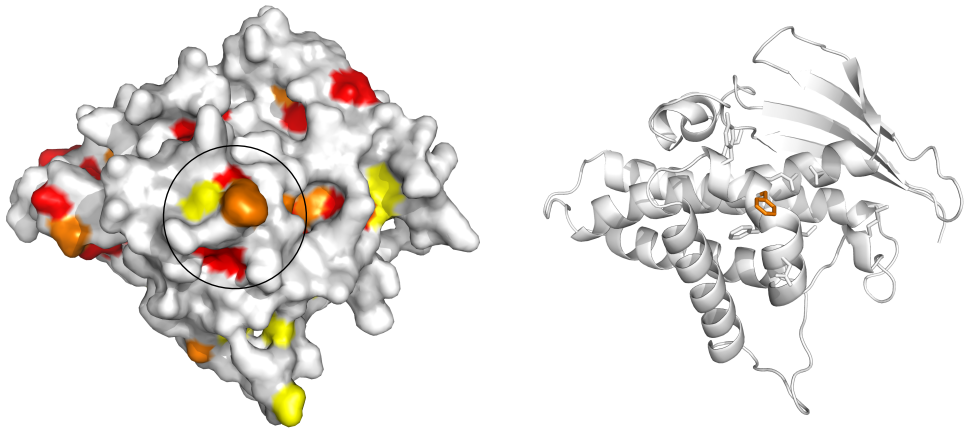
The inwards pointing residues, which are responsible for the overall CIDR fold, are well conserved. This is perhaps because there is no selection pressure to diversify these residues - they are not exposed to the immune system, and any major mutation may disrupt the CIDR fold, preventing correct folding and EPCR binding. These residues tended to be hydrophobic in character. The majority of the well conserved residues (Shannon entropy score below 0.12) were involved in the central three-helical bundle that is conserved between all CIDRs. Indeed conserved residues pointing into the three-helical bundle core are also conserved in property in the two other crystallised CIDR domains, var0 and MC179. In the CIDRα1 subclass the kink-forming residues, F653, V658 and W669, were all well conserved, preserving the presentation of the residues that interact with EPCR.

Our sequence-structure analysis suggests that the surface of the domain is the location of

the diversity seen between different CIDRα1 domains. This is consistent with a constant pressure, exerted by the host adaptive immune system, on the outwards-facing residues to diversify. We found that the majority of the surface residues have a Shannon entropy score of over 0.5, which means that they are highly diversified (Figure 4.17, left). Antibodies binding to these diverse surfaces on one CIDR will likely not recognise all variants. This diversity on the surface of the CIDR explains the finding that natural acquired immunity is specific to the *Pf*EMP1 variant [7]. Most of the residues with Shannon entropy scores below 0.2 are located in clefts and cavities in the surface of the CIDR. These clefts and cavities are unlikely to be accessed by antibodies, with most antibody:protein sites being roughly planar in nature [309].

The only well conserved binding residue found was the phenylalanine that dips into the hydrophobic cleft of EPCR (SE score of 0.17), with the second phenylalanine, F655, intermediately conserved (SE score of 0.31). The only binding residues that showed significant levels of conservation were the two bulky hydrophobic residues, F655 and F656. Other binding residues had low conservation scores above 0.4 (Figure 4.17, right). The Weblogo confirms this diversity - all of the hydrophilic residues vary in shape, size and charge. The only characteristic conserved between the majority of these hydrophilic is the ability to form hydrogen bonds. This suggests a conservation of hydrophobic or hydrophilic character to form the interaction with EPCR.

The majority of the EPCR-binding CIDRα1 domains tested bound EPCR with K_D value under 100 nM. The two weaker binding 1.5b CIDR domains bound with affinities of 229 nM and 168 µM respectively, and one 1.6a member bound with a weak affinity of 1.4 µM. The large range of K_D values was not entirely unexpected, due to the diversity in sequence, driven by evolution. However, given the reported dissociation coefficient of 30 nM [228] for the interaction between Protein C and EPCR, it is surprising that some of the domains tested, which are expressed by parasites *in vivo*, have CIDRs binding to EPCR with much lower affinities. This could be explained by the presentation of multiple *Pf*EMP1 domains on the surface of an infected erythrocyte, leading to avidity effects increasing the affinity between the infected erythrocyte and the endothlial surface. Alternatively, these particular *Pf*EMP1 variants may contain other ligand-binding domains



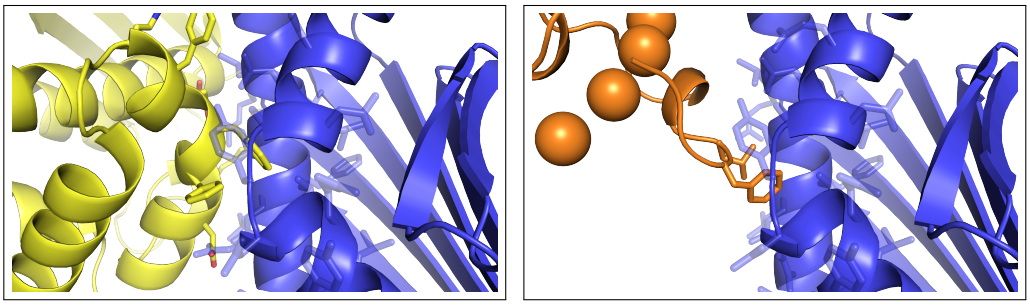
(a) **(b)**

Figure 4.17: (a) Surface representation of the CIDR showing the diversity of the surface. Only the phenylalanine, F656, circled, is well conserved. (b) Other binding residues are shown in stick representation. Well conserved residues are coloured in yellow ($0.2 < SE < 0.3$), orange ($0.1 < SE < 0.2$) and red ($0 < SE < 0.1$).

to adhere the infected erythrocyte to the endothelium. In these cases the EPCR interaction could still modulate the immune system by preventing the protein C signalling pathway.

Due to this large range of affinities, I looked at the variation in kinetic association and dissociation rates. My analyses of the two kinetic parameters shows differences in their variation and skew. The variation of $k_{association}$ can be represented as a normal distribution, centring around 2×10^4 M^{-1} s^{-1}. The $k_{dissocation}$ values, on the other hand vary in a skewed manner, signified by the Fisher-Pearson standardised moment coefficient of -0.6. This suggests that while $k_{association}$ is not under evolutionary selection, there is a selection pressure for a low $k_{dissociation}$, which leads to the slow dissociation rate observed in all natural CIDRα1 variants tested. Previous data suggests a $k_{association}$ of 5×10^5 M^{-1} s^{-1} for the Protein C:EPCR interaction, which is comparable to the average CIDR:EPCR $k_{association}$ value. On the other hand, all of the CIDRs had a lower $k_{dissociation}$ compared to the value for Protein C EPCR, 5×10^{-2} s^{-1} [310, 311]. A larger interface surface area and a deeper protruding phenylalanine (Figure 4.18) may contribute to the CIDR:EPCR having a slower dissociation rate. This slow dissociation rate means that the CIDR:EPCR complex is very stable once formed; once the CIDRs bind to EPCR, they stay bound. In the environment of the bloodstream this would be important to keep infected erythrocytes

in place against the shearing force exerted by the flow of blood. This would enable them to avoid splenic clearance. This stable *Pf*EMP1:EPCR interaction could also prevent Protein C from binding EPCR molecules already bound by *Pf*EMP1s, leading to the inflammation seen in patients with cerebral malaria [42, 140].



(a) **(b)**

Figure 4.18: The central interaction site between CIDR (yellow) and EPCR (blue) (a), and Protein C (orange) and EPCR (b) is shown with binding residues from both EPCR ligands shown as sticks. The central phenylalanine in the CIDR structure protrudes further into the EPCR groove (residues shown as transparent sticks) than the corresponding residue in the Protein C structure.

In parallel to this diversity based approach of characterising the conservation of interaction, I designed nineteen mutants, changing residues involved in the interaction site to investigate the properties of the binding site. I focussed on the hydrophobic patch, especially the protruding phenylalanine, as this residues at the centre of both the CIDR:EPCR and Protein C:EPCR complexes. I also mutated the Q657 position to try and explain why the CIDRα1.2 and 1.3 subclasses did not bind EPCR.

Mutating the central pair of phenylalanines to residues found in other CIDRα1 domains had little effect on the binding kinetics. This is in line with our SPR studies on the variants that contain those particular mutations. For example, the F656V mutant reflects sequence diversity in the CIDRα1.5b subclass. This subclass is more distantly related to the other CIDRα1 subclasses (Figure 4.1). Representatives of this subclass bind EPCR with a much lower affinity in SPR, but maintain an extremely slow off rate. Parasites containing this subclass of CIDR have other CIDR subclasses that bind EPCR encoded in their genome, so this could perhaps be an alternative, sequence diverse variants that can bind EPCR without a large bulky hydrophobic residue.

Though mutating the protruding phenylalanine to residues found in the other CIDRα1 subclasses had little effect on the binding to EPCR, mutating the residue to extreme alternatives revealed a little on how the interaction is conserved. Given the mutation from phenylalnine to valine had little effect on the interaction, a F656A mutant was created, decreasing the hydrophobic interaction area further. This mutation was not represented in the sequences that the Lavstsen group had analysed, and SPR experiments showed that though the $k_{association}$ was not much affected, the $k_{dissociation}$ was one hundred times higher than the wildtype. This reflects a much faster dissociation rate. This mutation, combined with the conservation of a low $k_{dissociation}$ in the natural variants tested, suggests that residues in this position must be hydrophobic with a surface area at least the size of valine.

In trying to rationalise the lack of binding activity of the CIDRα1.2 and 1.3 domains, we identified Q657 as an important position, due to the presence of a lysine in this position in both subclasses that do not bind EPCR. Both Q657K and Q657R mutations increased the K_D of the interaction by around 200 times. This mutation alone cannot fully explain the lack of binding by the 1.2 and 1.3 domains, but may be a major contributor to this, given the need for the conservation of a slow dissociation rate. The $k_{dissociation}$ of the Q657R mutant could be partially rescued by a second mutation, S653F. This could give a glimpse into how the 1.6b subclass can still bind EPCR with a slow dissociation rate, despite having the detrimental Q657R mutation.

We identified F651, V658 and W669 as residues that could stabilise the kinked helix that presents the two central hydrophobic residues, as in the crystal structure these face inwards in the binding arm. These residues are almost completely conserved in the CIDRα1 domain class; the only variation in any of these residues in the natural variants studied is in the CIDRα1.1 domain from ERS009959, which has a serine in place of the tryptophan. This CIDR bound weaker than the other domains in its subclass (K_D = 182 nM), with a substantially slower association rate, and a dissociation rate typical of this interaction. This could reflect a collapsed kinked helix, unable to present the bulky hydrophobic residue correctly to interact with the hydrophobic box in EPCR. I made a W669S mutant to reflect this, and alanine mutants of the other two residues. Though none of these mutations abolished the interaction, they all bound EPCR with between 10 and 100-fold weaker

affinities. This shows that the exact presentation of the interfacial residues is important to the CIDR:EPCR interaction.

I characterised the interaction from a thermodynamic point of view to look at whether the interaction was entropically or enthalpically driven. The results of the ITC experiments show that the interaction is enthalpically driven by a ΔH of -52.3 kJmol^{-1}. This binding interaction is entropically unfavourable; this is expected due to the large loss in entropy for two interacting molecules. Given that protein-protein hydrogen bonds have energies of around 5 kJmol^{-1} [312] and that 7 hydrophilic residues are involved in the interaction, the observed enthalpy is consistent with these residues making hydrogen bonds, plus van der Waal's interactions from the large, bulky hydrophobic residue pair.

Looking at the mutant data and the ITC data together, we can start to infer roles of some residues in the CIDRα1 interaction with EPCR. The central phenylalanine, F656, is responsible for much of the slow dissociation of the interaction, while not contributing greatly to the association rate. Given that this interaction is enthalpy-driven, it seems likely that the phenylalanine contributes to the stability of the complex through van der Waals forces between the phenylalanine and the hydrophobic box of EPCR. The association rate is likely to be mediated through the hydrophilic residues which are stronger at greater distances than van der Waals forces. The presentation of the binding surface is also important for association, suggested by the kink-forming residue mutants. The hydrophilic contacts could hence initiate complex formation, supplemented by van der Waals forces once the two proteins are in close proximity. This would lead to a strong, stable complex as observed in SPR.

These experiments reveal that CIDRα1 domains bind EPCR with a conserved slow dissociation rate. The structural core of the CIDR domain to be well conserved, with sequence variability at the surface of the domain. Of the nine binding residues, only the central phenylalanine, which protrudes into the groove of EPCR, is well conserved. The rest of the residues only maintain their hydrophobic or hydrophilic character, with much variability, especially in the hydrophilic residues. This identifies the molecular determinants of the CIDR:EPCR interaction, that will guide immunogen design.

Chapter 5

Characterising EPCR and ICAM-1 dual binding by PfEMP1

5.1 Outline

The extracellular portion of *Pf*EMP1 proteins consists of between two and seven DBL or CIDR domains, some of which can bind to host ligands. In this chapter I investigate how EPCR binds multidomain *Pf*EMP1 constructs. In collaboration with Frank Lennartz, I also ask whether a single *Pf*EMP1 construct can simultaneously bind two ligands, EPCR and ICAM-1. I optimised expression for *Pf*EMP1 head constructs in the *Sf*9/baculovirus system, producing two different three-domain head structures to test. These domains were tested for EPCR and ICAM-1 binding by SPR and gel filtration, which showed that simultaneous binding can occur. Furthermore binding of one ligand does not affect the affinity of the partner interaction. SAXS data was collected for one head structure in complex with each ligand individually. These data show that both ligands can bind individually to a head structure, with the head structure:EPCR complex consistent with the crystal structures solved in the previous chapter. Together, these experiment demonstrate that within the context of a larger *Pf*EMP1 construct, ligand-binding domains act as individuals and are able to simultaneously bind ligands. These findings will have consequences on both pathology studies and therapy design.

5.2 Declaration

This chapter contains unpublished data. Work described in this chapter was devised and carried out in collaboration with Frank Lennartz, a postdoctoral researcher in the Higgins group, who studies the interaction between *Pf*EMP1 and ICAM-1. P1 virus for the three-domain head structures and a sample of HB3var03 head structure was supplied by the Lavstsen group, University of Copenhagen.

5.3 Production of *Pf*EMP1 head structures

5.3.1 Construct design

We decided to use larger *Pf*EMP1 constructs from natural strains rather than merely fusing two recombinant *Pf*EMP1 domains together, in order to study dual-binding interactions that may have relevance *in vivo*. All discovered EPCR-binding CIDRα1 domains are positioned as the second domain from the N-terminus, while the identified ICAM-1 binding DBLβ domains are usually positioned as the third domain from the N-terminus. We observed that EPCR and ICAM-1 binding domain are present in some natural *Pf*EMP1 variants, and so decided to produce constructs consisting of the two adjacent ligand-binding domains, plus the N-terminal DBL domain, henceforth known as "head structures".

The Lavstsen group has previously expressed longer *Pf*EMP1 constructs in the *Sf*9/-baculovirus system [140]. We hence decided to use this system to express three-domain head structures for testing. P1 virus stocks for the PF11_0521 and IT4var07 three domain head structures were supplied by the Lavstsen group. These viruses were generated using the BaculoGold transfection kit to cotransfect *Sf*9 cells with linearised viral DNA plus the head structure constructs, which included a histidine tag for purification.

5.3.2 Expression of *Pf*EMP1 head structures in *Sf*9/baculovirus

Each P1 virus was amplified to P3 through subsequent rounds of *Sf*9 infection. Three litres of Hi5 cells were grown up to infect with each P3 virus, harvesting the supernatant containing the head structure three days later. Cells were spun down, and the supernatant was filtered before purification.

Purification of the head structures from the supernatant required some optimisation. I initially attempted to purify protein from the supernatant by buffer exchanging the protein into 20 mM Tris pH 8; 300 mM NaCl. An SDS-PAGE analysis of the nickel-NTA affinity purification showed contaminating bands in the Ni-NTA elution fractions, as well as a band at the expected size (Figure 5.1, left). I hence optimised the imidazole concentration, binding and eluting small amounts of the supernatant in the presence of 0 mM, 25 mM, 50 mM, 75 mM and 100 mM imidazole (Figure 5.1, right). A concentration of 25 mM imidazole successfully prevented non-specific interactions of the contaminating proteins, leading to elution of pure protein from the Ni-NTA beads.

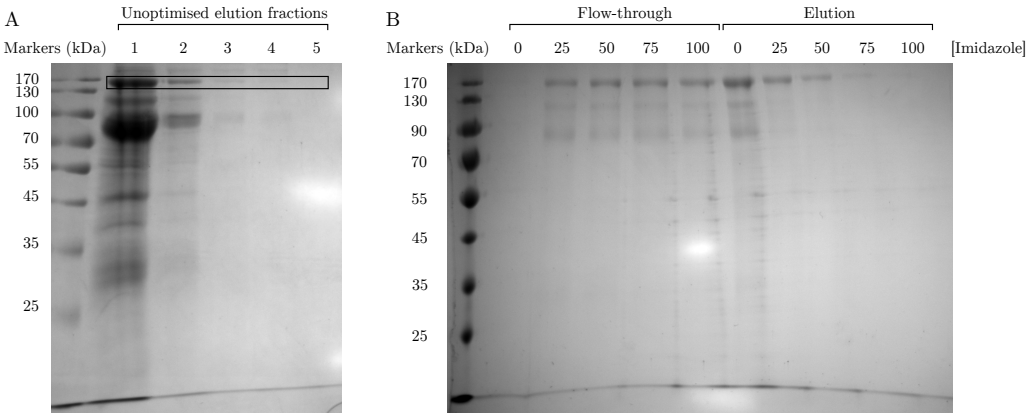


Figure 5.1: SDS-PAGE analysis of the initial purification showed that the head structure (black box) purified with contaminants (A). SDS-PAGE analysis of the optimisation experiments shows 25 mM imidazole as the optimal concentration to separate the head structure, which elutes from the beads, from the contaminants in the flow-through (B). The head structure band runs at 170 kDa.

An initial expression attempt suggested that much less protein binds to the nickel-NTA beads than suggested by the manufacturer's protocol. In this attempt 3 ml of Ni-NTA beads (enough to bind up to 150 mg of protein according to the manufacturer) were

incubated with the TFF eluant for 15 h. A subsequent Western blot revealed a band in the Ni-NTA flow through at the correct size and similar intensity for the head structure band seen in both the initial sample and the Ni-purified sample. As a result, 10 ml of nickel beads were added to the subsequent TFF eluants, incubating at 4°C for 15 h. Protein was then purified from these nickel beads, eluting using 200 mM imidazole. This protein was subsequently gel filtered on an S200 column. The gel filtration traces and subsequent SDS-PAGE analyses show pure protein in the monodisperse peak (Figure 5.3). IT4var07 was expressed and purified giving a yield at 50 µg/L culture; PF11_0521 gave a yield of 50 mg/L culture.

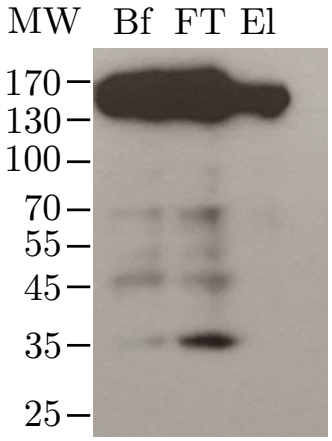


Figure 5.2: A Western blot of the initial head structure Ni-NTA purification. Comparing the three lanes (before, Bf; flow-through, Ft; elution, El) shows that there is still a considerable amount of protein left in the flow-through.

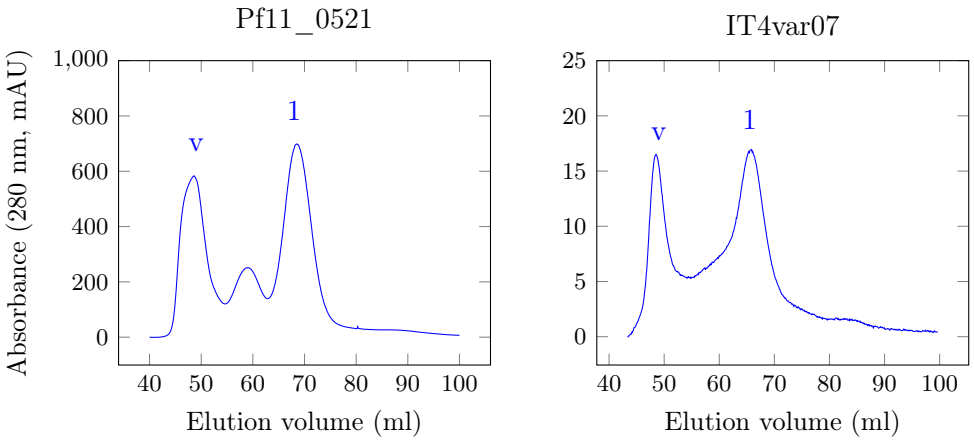


Figure 5.3: Preparative gel filtration for the PF11_0521 and IT4var07 head structures. The peaks at the correct size for the head structures are marked with a 1.

5.4 Dual binding of *Pf*EMP1s to ICAM-1 and EPCR

To test the dual binding to these *Pf*EMP1 constructs, we used a range of biophysical techniques to assess the interaction between head structures and ICAM-1, produced in HEK293F cells by Frank Lennartz, and EPCR (produced in Section 3.3.4). In addition to the head structures that I expressed, we tested a DD2var32 NTSDBL-CIDR-DBL construct, expressed by Frank Lennartz, and a small sample of HB3var03, provided by the Lavstsen group.

5.4.1 Gel filtration of complex

We intiailly set up a series of gel filtration experiments with DD2var32 in complex with EPCR and an ICAM-1 D1D2 construct (consisting of the first two Ig domains) to assess dual binding. Head structures were incubated with EPCR and ICAM-1 individually and with both ligands together for thirty minutes to give time for complexes to form. After incubation, these samples were run on a Superdex 200 column to assess complex formation. Compared to gel filtration of head structures alone, the peak containing DD2var32 shifts to a higher molecular weight when mixed with each combination of ligands. The fact that DD2var32:ICAM-1:EPCR elutes before DD2var32:ICAM-1, which in turn elutes before DD2var32:EPCR is consistent with the molecular weight added (EPCR is 21 kDa; the ICAM-1 construct, fully glycosylated, is 45 kDa) In each case the shifted peak was analysed by SDS-PAGE, showing that these peaks contained all components, suggesting that the complex has formed and is stable over the course of a gel filtration experiment.

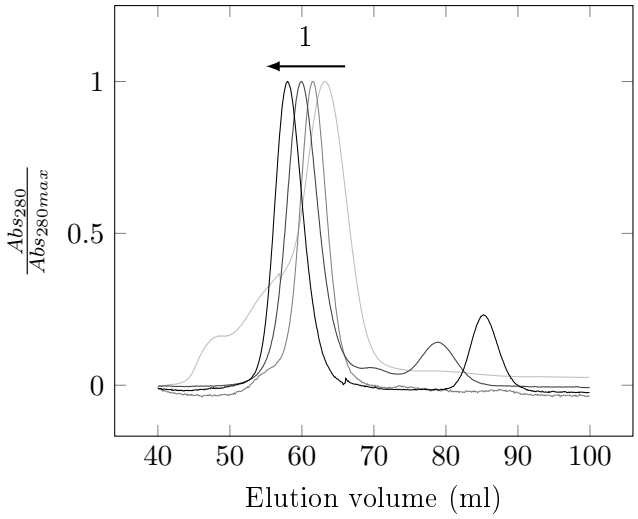


Figure 5.4: Gel filtration experiments, carried out on an S200 column, show a shift in peak (shown as the arrow labelled 1) of the DD2var32 head structure (light grey) upon addition of EPCR (grey); ICAM-1 (dark grey); or EPCR and ICAM-1 (black).

5.4.2 Surface plasmon resonance

We used surface plasmon resonance to quantify the interaction between the *Pf*EMP1 and both EPCR and ICAM-1. In these experiments we exploited the slow dissociation rates of both the CIDR:EPCR and DBL:ICAM-1 interactions; these slow dissociation rates allowed us to form the triple complex on the SPR chip.

One ligand was first bound to an SPR chip, using the CAP system to couple biotinylated EPCR, or a Protein A/Protein G chip to couple an Fc-tagged ICAM-1 construct. A concentration series of each *Pf*EMP1 head structure was then flown over the chip surface to collect data for the head structures binding to each ligand individually. To assess dual binding, we flowed a concentration series of EPCR over the ICAM-1:head structure, immobilised via the ICAM-1 Fc tag to the Protein A/Protein G chip.

Full datasets were collected for the DD2var32 and Pf11_0521 head structures, with a preliminary dataset collected for HB3var03, due to limited sample. Data for the IT4var07 head structure could not be collected, due to non-specific binding of the head structure to both chip surfaces. This could be due to the particularly high isoelectric point for this head structure ($pI = 9$), causing non-specific interactions between the head structure and

the negatively charged chip surfaces at the pH of the running buffer.

The datasets collected are shown in Figures 5.5 and 5.6 for DD2var32 and Pf11_0521 respectively. The DD2var32 and Pf11_0521 datasets were fit to a kinetic 1:1 binding model.

The DD2var32 (Figure 5.5) and Pf11_0521 (Figure 5.6) were flown over a Protein A chip, with datasets fit to a kinetic 1:1 binding model. These head structures bound to both EPCR and ICAM-1 individually with low nanomolar K_D values and slow dissociation rates (Table 5.1). Dual binding was observed for both head structures in the triple complex experiments. In these experiments, the head structures were first flown over ICAM-1, immobilised on the chip surface via its Fc tag. This was followed by a concentration range of EPCR, which fit to a 1:1 kinetic model. The calculated K_D value for EPCR binding to the head structure:ICAM-1 complexes is similar to the K_D of the head structure binding to EPCR alone, showing that the EPCR interaction is not affected by the binding of ICAM-1 to the DBLβ domain within the head structure (Table 5.1).

HB3var03 head structure bound to EPCR with low nanomolar affinities (Figure 5.7, HB3var03:EPCR, Table 5.1). However the HB3var03 head structure bound non-specifically to the Protein A chip, preventing accurate data collection for the ICAM-1 interaction and for dual binding. A preliminary dataset was collected, instead using a Protein G chip to immobilise the ICAM-1 for these two experiments (Figure 5.7). These sets of binding curves clearly show HB3var03 binds to ICAM-1 and that EPCR can bind to the HB3var03:ICAM-1 complex, but the data could not be fit to kinetic models, preventing the calculation of kinetic constants.

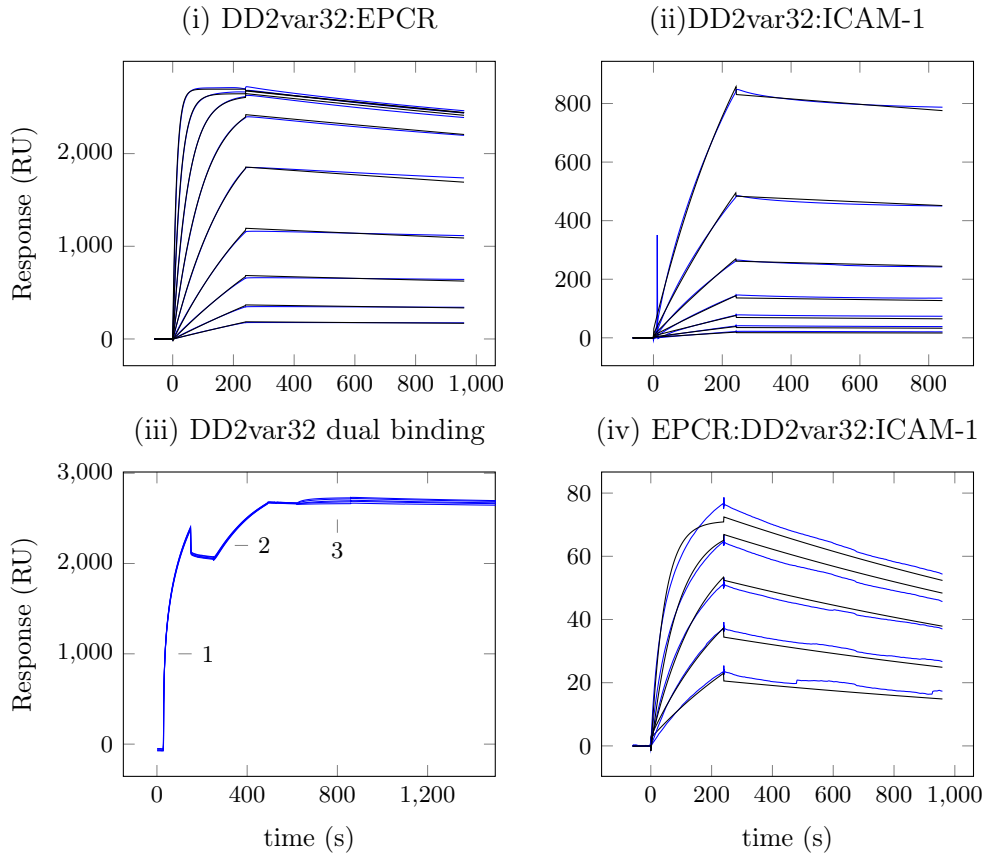


Figure 5.5: SPR curves show the interaction of the DD2var32 head structure with EPCR (i) and ICAM-1 (ii) individually. (iii) shows the sequential binding of ICAM-1 (1), DD2var32 (2) then a concentration series of EPCR (3). This EPCR concentration series is enlarged in (iv). Blank-subtracted data is shown in blue, with kinetic fits for quantitative data are shown in black.

Bound	Ligand	Range	K_D (nM)	k_{assoc} (M^{-1} s^{-1})	k_{dissoc} (s^{-1})
EPCR	DD2var32	1.9 nM - 500 nM	0.8	4.99×10^4	4.70×10^{-5}
ICAM-1	DD2var32	7.8 nM - 500 nM	20.6	4.99×10^4	4.70×10^{-5}
DD2var32:ICAM-1	EPCR	16.125 nM - 250 nM	5.1	8.00×10^4	3.09×10^{-4}
EPCR	Pf11_0521	1.9 nM - 500 nM	1.2	4.99×10^4	4.70×10^{-5}
ICAM-1	Pf11_0521	7.8 nM - 500 nM	2.3	4.99×10^4	4.70×10^{-5}
Pf11_0521:ICAM-1	EPCR	16.125 nM - 250 nM	11.4	8.00×10^4	3.09×10^{-4}
EPCR	HB3var03	7.8 nM - 250 nM	4.6	8.17×10^3	3.76×10^{-5}

Table 5.1: SPR summary table for quantitative measurements of head structure interactions.

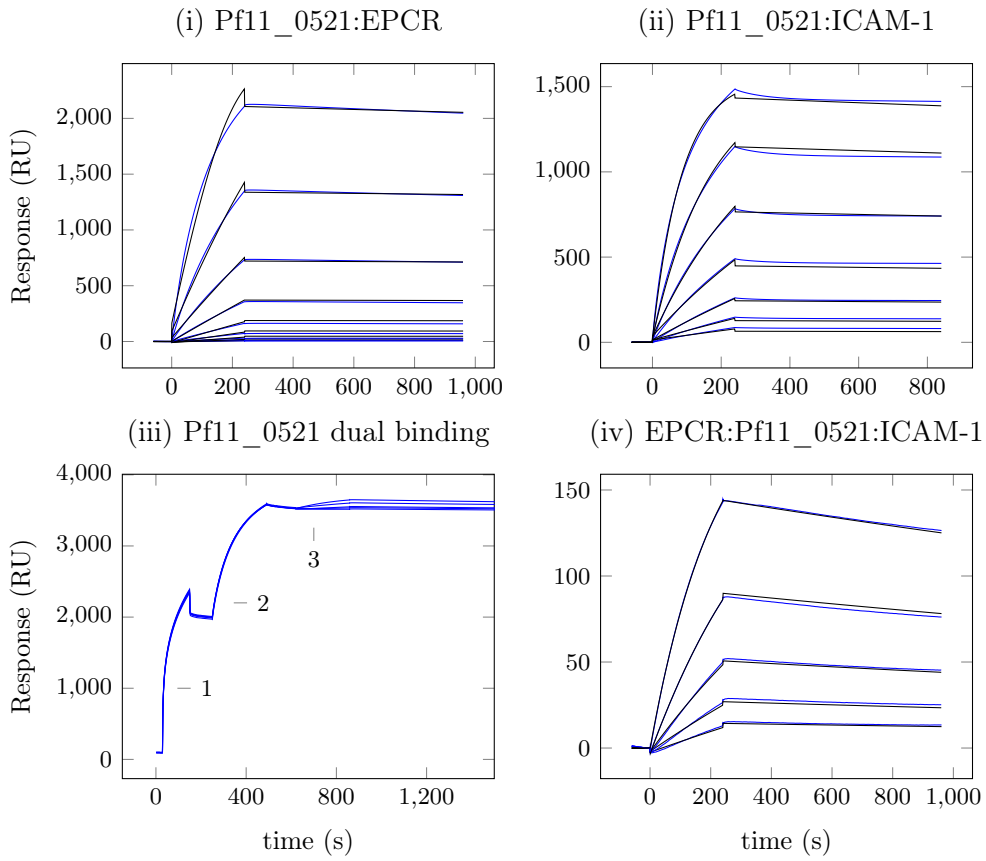


Figure 5.6: SPR curves show the interaction of the Pf11_0521 head structure with EPCR (i) and ICAM-1 (ii) individually. (iii) shows the sequential binding of ICAM-1 (1), Pf11_0521 (2) then a concentration series of EPCR (3). This EPCR concentration series is enlarged in (iv). Blank-subtracted data is shown in blue, with kinetic fits for quantitative data are shown in black.

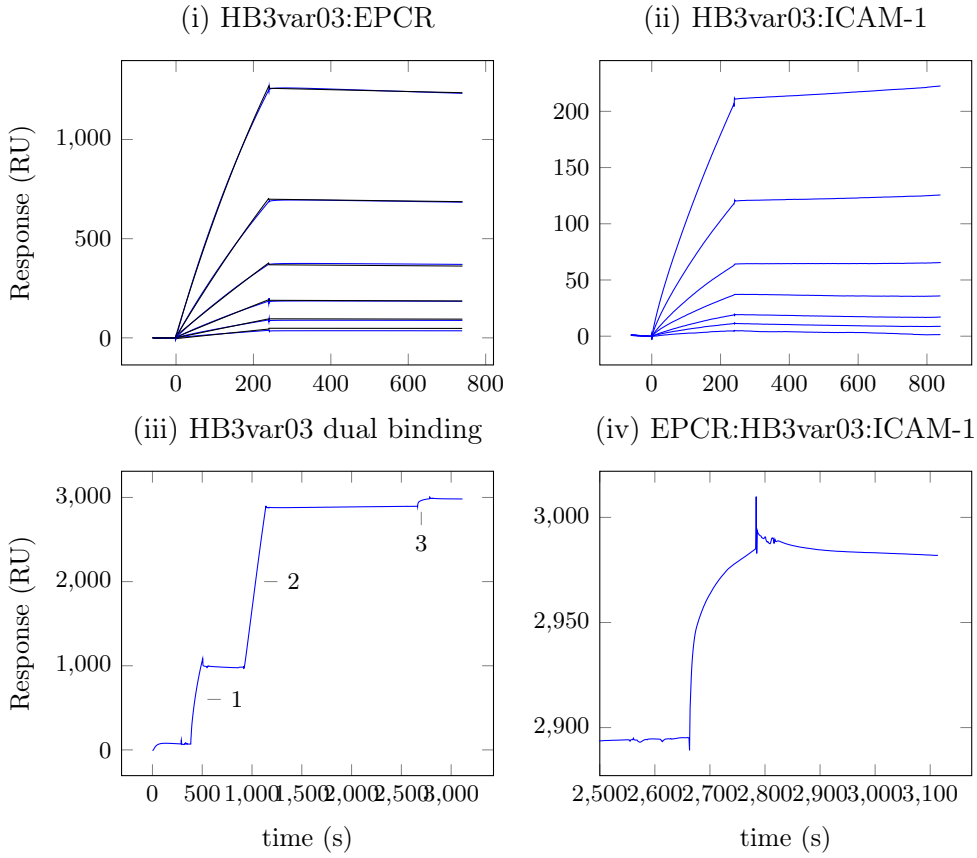


Figure 5.7: SPR curves show the interaction of the HB3var03 head structure with EPCR (i) and ICAM-1 (ii) individually. (iii) shows the sequential binding of ICAM-1 (1), HB3var03 (2) then a concentration series of EPCR (3). This EPCR concentration series is enlarged in (iv). Blank-subtracted data is shown in blue, with kinetic fits for quantitative data are shown in black.

5.4.3 Small angle X-ray scattering

Small angle X-ray measurements were taken using the best expressing head structures, DD2var32 and PF11_0521, to gain low resolution structural data about how both ligands bind to a longer construct. Data were collected for the head structures individually, and in combination with EPCR, ICAM-1 (D1D2) and for the triple complex.

SAXS measurements for DD2var32 complexes were taken on the P12 beamline at DESY (Hamburg, Germany). Data were collected for a series of concentrations for each sample, and examined for signs of aggregation. The reciprocal R_g values were consistent for the DD2var32 ($\bar{x}$ = 5.08 nm, s = 0.18 nm) and DD2var32:EPCR ($\bar{x}$ = 5.21 nm, s = 0.28 nm) series, so scattering curves with linear Guinier plots were both selected for further processing. These curves were fully processed to give real space distance distributions and real space radius of gyration, R_g (Figure 5.8). Comparing these two distance distributions shows that the DD2var32:EPCR has a similar R_g and D_{max} to DD2var32 alone (Table 5.2). This is not surprising, as EPCR accounts for 13% of the complex molecular weight and is expected to bind to the middle domain, the CIDRα domain, in the head structure. A comparison of the DD2var32 and the DD2var32:EPCR envelopes reveal an additional density in the complex. Previously solved DBL crystal structures and the HB3var03 CIDR structure fit into the both envelopes, with EPCR fitting in the correct orientation to bind the CIDR domain in the complex envelope (Figure 5.10).

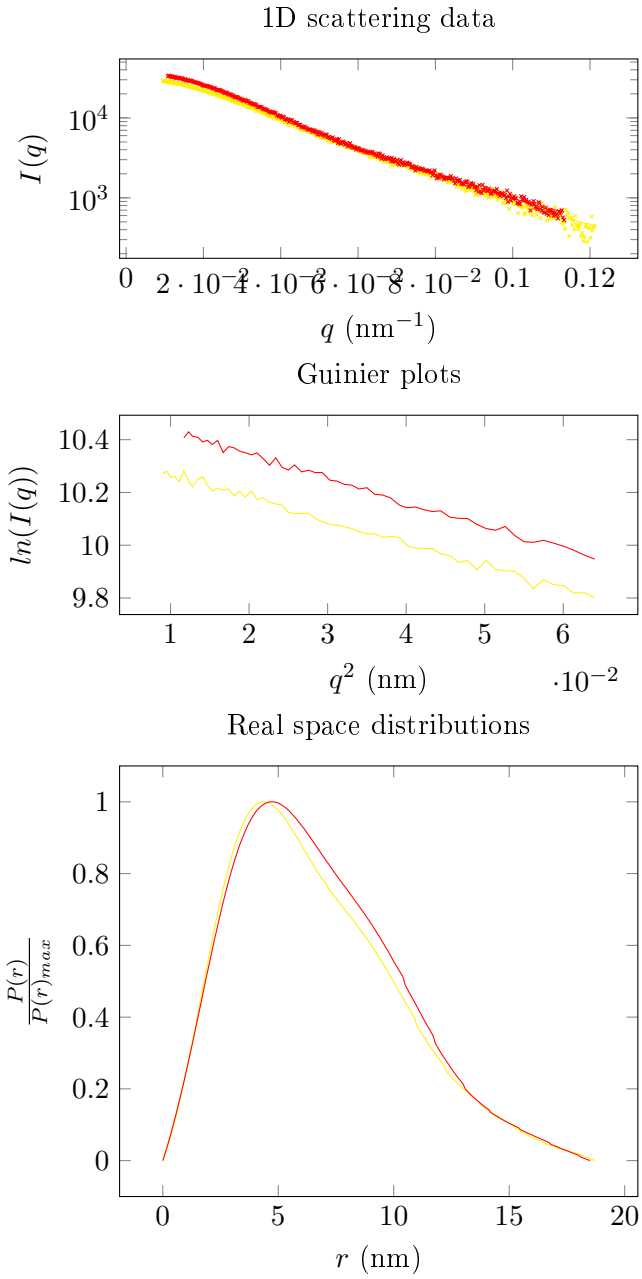


Figure 5.8: SAXS measurements for DD2var32 head structure complexes. In each plot the DD2var32 head structure alone is in yellow; the head structure:EPCR in red. 1D scattering data (top) shows the original data (x), and the fit of the generated envelope (Figure 5.10) to these data (line).

The R_g values for DD2var32:ICAM-1 ($\bar{x}$ = 7.67 nm, s = 1.01 nm) and for the DD2var32 triple complex ($\bar{x}$ = 8.82 nm, s = 0.96 nm) were less consistent, suggesting aggregation in the sample. In a second set of experiments, performed at the ESRF we ran these two complexes down a gel filtration column to separate any unbound components, isolating the complex peak for SAXS measurements (Section 5.4.1). A concentration series of these two complexes were run similarly to previous measurements. DD2var32:ICAM-1 R_g values were more consistent in this set of experiments ($\bar{x}$ = 5.581 nm, s = 0.23 nm), allowing us to process the dataset, similarly to above (Figure 5.9). The real space R_g for the DD2var32:ICAM-1 distance distribution is slightly larger than that of DD2var32 sample (5.51 compared to 5.3), while the D_{max} is the same (Table 5.2). This suggests that the addition of ICAM-1 does not significantly change the overall shape of the molecule. The calculated envelope for this dataset is shown in Figure 5.10, with the previously solved ICAM-1 structure fit into the envelope next to the C-terminal DBL domain. Unfortunately the triple complex concentration still gave an inconsistent R_g ($\bar{x}$ = 7.02 nm, s = 0.55 nm), so these data were not processed.

Species	Real space R_g (nm)	Volume (nm^3)	MW_{app} (kDa)	D_{max} (nm)	$\sqrt{\chi}$
DD2var32					
Head structure	5.30	263.89	155.23	18.7	1.034
+ EPCR	5.38	309	181.76	18	0.979
+ ICAM-1	5.51	371	218.2	18	0.965

Table 5.2: Summary table of the head structure SAXS.

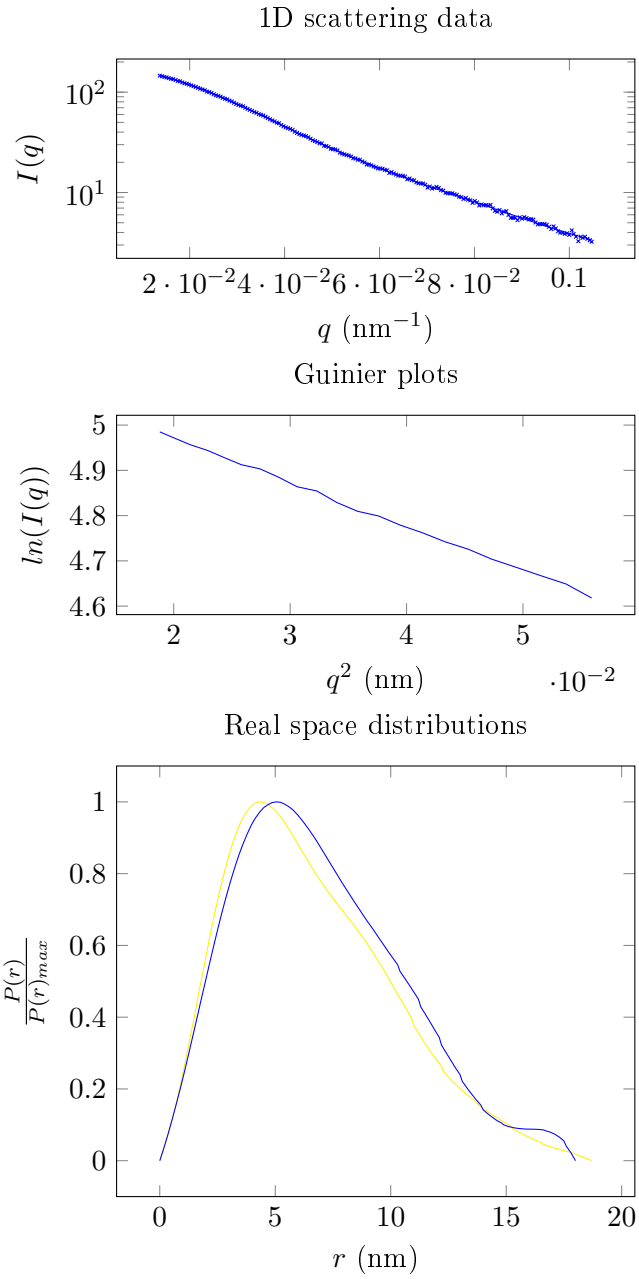


Figure 5.9: SAXS measurements for DD2var32 head structure complexes. In each plot the head structure:ICAM-1 is shown in blue, with the DD2var32 real space distribution (yellow) shown for reference. 1D scattering data (top) shows the original data (x), and the fit of the generated envelope (Figure 5.10) to these data (line).

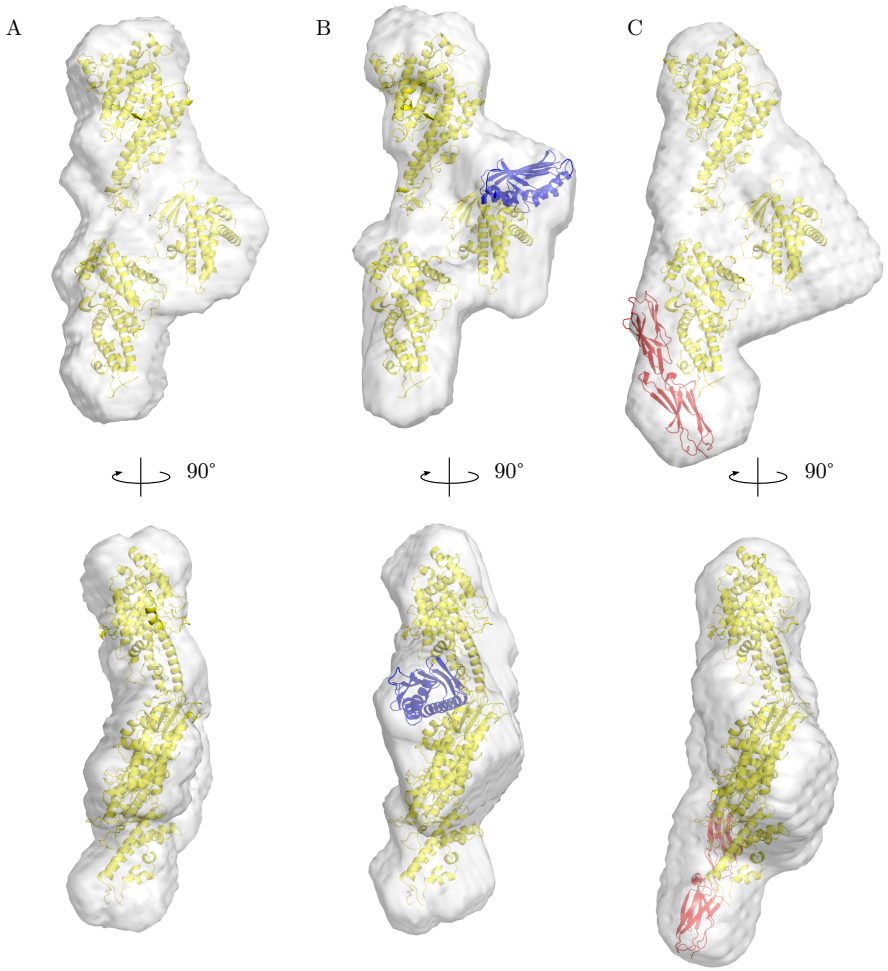


Figure 5.10: SAXS envelopes generated for the DD2var32 head structure alone (A), in complex with EPCR (B) or in complex with ICAM-1 (C). The crystal structures of var0 NTSDBL, HB3var03 CIDR and DBL3x are fitted into the envelopes as a model of the head structure (all yellow). The EPCR is modelled in blue and ICAM-1 is modelled in red.

5.5 Discussion

Fresh parasite isolates from patients have been shown to bind multiple ligands [313], suggesting *Pf*EMP1s to contain multiple ligand binding domains. This is advantageous to the parasite, as there will be a greater number of attachment sites to which the infected erythrocytes can cytoadhere. In addition to this, a single parasitised red blood cell line, from the FCR3 lab strain, has been shown to mediate interactions between cytoadhesion and rosetting ligands [314]. This could allow the concentration of uninfected red blood cells at sites of sequestration, facilitating faster invasion by the next generation of merozoites. While these studies suggest a dual functionality for some *Pf*EMP1 proteins, simultaneous binding of two different ligands has not previously been demonstrated. Given the relationship between two *Pf*EMP1 ligands EPCR and ICAM-1, we investigated the possibility of simultaneous binding of these two ligands to a singular *Pf*EMP1. The results of this chapter show that one *Pf*EMP1 can simultaneously mediate interactions with two ligands.

*Pf*EMP1 interactions with EPCR and ICAM-1 were chosen for this study as the expression of ICAM-1 is linked to thrombin-activated PAR-1 [233], which may occur when Protein C signalling is preventing by *Pf*EMP1 binding to EPCR [140]. The *Pf*EMP1 head structures tested in this thesis come from lab strains, containing adjacent CIDRα1.4 and DBLβ3 domains, which bind EPCR and ICAM-1 respectively. I expressed and purified two NTSDBL-CIDR-DBL head structures that contained both EPCR and ICAM-1 binding domains using the baculovirus/*Sf*9 system. Optimisation of imidazole concentration allowed protein to be purified with few contaminants, while adding an excess of Ni-NTA increased the yield of protein.

The gel filtration experiments and the SPR data clearly show that different *Pf*EMP1 head structures can bind both EPCR and ICAM-1 at the same time. Quantitative kinetic data gained from the SPR experiments, using the DD2var32 and Pf11_0521 head structures, also show that the affinity of these proteins to EPCR is not affected by the binding of ICAM-1 to the *Pf*EMP1 (Table 5.1). Unfortunately we could not conclusively prove ICAM-1 binding to be independent of EPCR, though we expect this to be the case. ICAM-1 could not bind to the DD2var32 head structure when bound to EPCR, immo-

bilised on the CAP chip in our assays. This could be as EPCR is a small, globular protein compared to the longer ICAM-1. This means that the head structure is held closer to the chip surface in the CAP chip assay, perhaps partially occluding the ICAM-1 binding site.

As well as allowing us to verify simultaneous interactions, the SPR data allowed me to study the CIDR:EPCR interaction in the context of a larger *Pf*EMP1 construct. It has been shown previously that EPCR binds the intact *Pf*EMP1 ectodomain with a similar affinity to the CIDR domain alone [140]. Comparing these kinetic parameters to values generated in Chapter 4 suggests that the affinity, K_D, is similar to the individual domains alone, in line with the previous data. However, the kinetic components of K_D differ up to a factor of 100, with a lower $k_{association}$ and higher $k_{dissocation}$ values for both head structures than for the corresponding CIDR domains alone (Table 5.3). Whether these differences are trends or artefacts of the SPR system is unclear. However, given my finding that the dissociation rate is particularly well conserved, the fact that larger *Pf*EMP1 constructs have a slower dissociation rate to a surface-bound ligand could be a physiological advantage to the parasite, driving towards a stable interaction between the infected erythrocyte and the endothelial membrane.

Bound	Ligand	K_D (nM)	k_{assoc} (M^{-1} s^{-1})	k_{dissoc} (s^{-1})
EPCR	DD2var32 head structure	0.8	4.99×10^4	4.70×10^{-5}
EPCR	DD2var32 CIDR	0.3	1.06×10^6	3.43×10^{-4}
EPCR	HB3var03 head structure	4.6	8.17×10^3	3.76×10^{-5}
EPCR	HB3var03 CIDR	0.37	5.32×10^5	1.97×10^{-4}

Table 5.3: SPR data summary table for EPCR interacting with different size *Pf*EMP1 constructs from DD2var32 and HB3var03.

There is a limited range of structural information available about the multidomain architecture *Pf*EMP1. SAXS measurements by Alan Brown showed that IT4var13 is long and relatively straight [315], whilst SAXS measurements made using the var2CSA ectodomain show that the extracellular domain of var2CSA adopts a much more compact structure [316]. In addition to these two SAXS reconstructions, the crystal structure of the NTSDBL-CIDR from var0 suggests that the NTSDBL domain folds back onto the CIDR. This observation is not consistent with the SAXS reconstruction of the complete IT4var13 *Pf*EMP1,

where the NTSDBL domain is extended. We attempted to use SAXS as a technique to probe how EPCR and ICAM-1 interact with a longer *Pf*EMP1 construct individually and in combination.

Unfortunately, problems with aggregation prevented us from gaining a structural insight into the triple complex between the *Pf*EMP1 head structure, EPCR and ICAM-1. However, the processed datasets allowed us to look at how EPCR and ICAM-1 individually interact with multidomain *Pf*EMP1 constructs. Comparing the DD2var32 and DD2var32:-EPCR envelopes reveals a clear density for EPCR. Fitting the solved complex structure into this envelopes with two additional DBL domains places EPCR in an antiparallel orientation with respect to the *Pf*EMP1. This is consistent with the two proteins coming together from opposite membranes. Without a crystal structure of the DBLβ:ICAM-1 complex, it is difficult to place the DBLβ:ICAM-1 interaction in context of the head structure. Placing the three *Pf*EMP1 domains into the DD2var32:ICAM-1 envelope as they are in the DD2var32 envelope reveals additional density next to the C-terminal DBL domain, which can fit ICAM-1 (Figure 5.10, C). However, this is different to ICAM-1 binding site in the SAXS-derived architecture of the IT4var13:ICAM-1 complex [315]: in this envelope ICAM-1 appears to stick out roughly perpendicular from the head structure. The differences between this envelope and the one shown here are difficult to reconcile with such low resolution data. It is possible that the DBLβdomain in the head structure is constrained in a certain conformation by the rest of the *Pf*EMP1 ectodomain in the IT4var13 envelope. More high resolution structural data of multidomain *Pf*EMP1 constructs is needed to resolve these differences and to assess the conformation of an intact *Pf*EMP1 domain. This data could be collected using X-ray crystallography or cryo-electron microscopy.

Despite failing to gain structural information about a triple complex, the experiments in this chapter show that *Pf*EMP1 proteins can simultaneously bind to two ligands. It remains to be seen whether the same is true for infected erythrocytes that have these *Pf*EMP1 proteins on the surface. The capability *Pf*EMP1 proteins to interact simultaneously with two ligands could be advantageous for the parasite. These multiple interactions would allow for the parasited red blood cell to anchor tighter to the endothelium surface by interacting with twice the number of ligands, perhaps allowing the infected erythrocyte

to sequester under a higher shear force. Parasited red blood cell flow assays, assessing the shear force required to remove an infected blood cell from a surface coated with EPCR and ICAM-1 individually or together would explore this possibility.

This demonstration that the EPCR binding and ICAM-1 binding sites act independently on a *Pf*EMP1 variant may also pose problem for vaccine developers who hope to use *Pf*EMP1 as a vaccine target. Would antibodies raised against the EPCR binding site be enough to prevent the cytoadhesion of a *Pf*EMP1 that could also bind ICAM-1? Since the ligand-binding sites are shown here to act independently, this is unlikely, though antibody binding could sterically hinder the other binding site.

Work is ongoing to establish whether the presence of both ICAM-1-binding and EPCR-binding domains is linked to malaria severity. It will be interesting to see whether this DBLβ domain is also correlated to some extent with domain cassette 13, and whether other domain cassettes, containing EPCR- and ICAM-1-binding domains, can be identified. If the proposed model [140] holds true, an antibody against the *Pf*EMP1:EPCR could lead to a decrease in ICAM-1 expression. It is hence clear that the link between these two *Pf*EMP1 ligands should be explored.

Chapter 6

Rational immunogen design

6.1 Outline

Inhibition of the CIDR:EPCR interaction may be an important requirement for a vaccine against severe malaria. Having identified the conserved determinants of the CIDR:EPCR interaction, I designed a range of immunogens that could focus the antibody response onto the EPCR-binding surface of the CIDR. A glycosylated CIDR construct was designed and expressed in S2 cells. This immunogen bound to EPCR with similar affinity to the wild type CIDR, and so was used as an immunogen to raise an antibody response in rats. We immunised also rats with a stapled helical peptide, designed to replicate the structure of the helix including the central phenylalanine, F656. Analysis of immunised rat sera revealed that these two immunogens had no effect on the proportion of inhibitory antibody raised in rats.

I also grafted the binding epitope to a helical bundle scaffold *in silico*, attempting to design a soluble, highly expressed immunogen. This was possible as the seven of the nine binding residues were a continuous stretch of peptide, and mutating the other two residues did not abolish the CIDR:EPCR interaction. Epitope grafting was accomplished by using the Rosetta software suite: the three helical bundle scaffold was first folded around the binding epitope, then the sequence was optimised for stability of this fold. I

expressed the 8 most stable helical bundles and characterised them for stability and EPCR binding. Two of these bundles were stable to high temperatures, as confirmed by circular dichroism, though no bundle from this first round of design bound to EPCR. A second round of design introduced three cysteine pairs into the most stable bundle, designed to form disulfide bonds to stabilise the target epitope structure. These cysteine pairs were mutated individually and in combination, with the resulting bundles again tested for fold stability and EPCR binding. One bundle showed EPCR binding, with a characteristically slow off rate, suggesting this mutant has a binding site reminiscent of the parent CIDR. This immunogen may be useful alone, or as part of a multi-component vaccine, to elicit inhibitory antibodies.

6.2 Declaration

Rat immunisation and preparation of the sera samples were carried out by the Lavstsen lab, University of Copenhagen, for the assessment of the immunogen design strategies. Gareth Watson, an undergraduate project student, set up the fluorescence inhibition assay with my guidance as project supervisor. Helical bundle expression and testing was done with the help of Gareth Watson and Jack Davey (a summer student), again with my guidance, expressing some of the bundles, and carrying out some of the biophysical charcaterisation.

6.3 Assessment of immunisation strategies

6.3.1 Fluorescent labelling of CIDRs

The inhibition assays required a labelled CIDR, in order to quickly assess the success of the assay without the requirement of an extra detection step. The CIDRs were fluorescently labelled with Alexafluor 488 nm, coupling the fluorophore to lysine residues on the CIDR. After 15 h the reaction was quenched using diethanolamine and the free fluorophore was removed. Attempts to remove the free fluorophore from the reaction mixture using a supplied size exclusion resin resulted in around 90% loss of protein. Subsequent purifications

used dialysis to remove the free fluorophore from the reaction mixture.

The fluoroscently labelled CIDR domains (fCIDRs) were checked for binding by SPR (Figure 6.1). These experiments showed that the fluorophores did not alter the binding of the fCIDRs to EPCR, producing similar kinetic and thermodynamic constants to that of the unlabelled CIDR (Table 6.1).

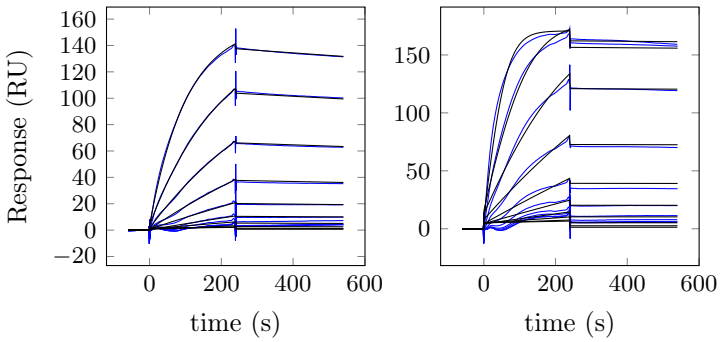


Figure 6.1: Surface plasmon resonance traces showing fCIDRs binding to EPCR. Each dataset consists of binding data for 2-fold dilution series from 500 nM, with blank-subtracted data shown in blue and kinetic fits shown in black.

fluorescent CIDR	Range	K_D (nM)	k_{assoc} (M^{-1} s^{-1})	k_{dissoc} (s^{-1})
HB3var03 fCIDR	500 nM - 0.9 nM	0.2	7.51×10^4	1.56×10^{-5}
IT4var20 fCIDR	500 nM - 0.9 nM	7.6	1.90×10^4	1.44×10^{-4}

Table 6.1: SPR data summary table for fluorescent CIDR binding, tested for EPCR binding.

6.3.2 Assay design

I designed two assays to assess the sera, raised using the immunisation strategies. The first was an ELISA to assess the reactivity of each sera sample against two sequence diverse CIDRα domains from IT4var20 and HB3var03, which have 44% sequence identity. The CIDR domains in question were immobilised on a Nunc Maxisorp® ELISA plate, the plate blocked, then the sera incubated on the plate for 1 h. Bound antibody was detected using a α-rat IgG-HRP antibody, using TMB substrate to quantify the amount of bound antibody. Sera, raised by each immunogen, was tested for binding to both CIDR domains, to test their reactivity against the homologous CIDR, HB3var03, upon which

all the immunogens were based, and a heterologous CIDR, IT4var20, to see if the sera contained any cross-reactive antibodies.

The second assessment method was a fluorescence inhibition assay. In this assay, fCIDR was incubated with the sera, raised against each immunogen. This mixture was then incubated on an EPCR-coated plate for one hour, before extensively washing of the wells. If the immunogen design strategy was successful in raising inhibitory antibodies, no fCIDR would remain bound to EPCR on the plate and no fluorescent signal would be observed. Each batch of fluorescent CIDR was calibrated for signal in the assay, by incubated the EPCR-coated plate with different quantities of the fCIDRs (Figure 6.2). The concentration of fCIDR that saturated the plate without leaving excess fCIDR in the solution was selected, so that any inhibitory antibody binding would cause a decrease in the observed fluorescence.

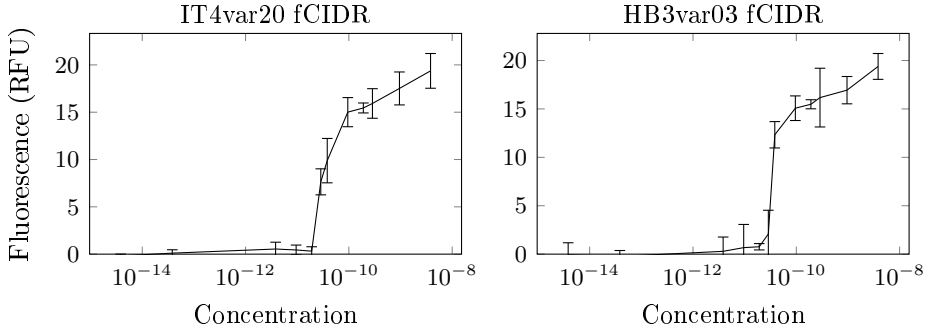


Figure 6.2: An example set of calibration curves for fluorescent CIDR binding to immobilised EPCR, with 95% confidence limits. The IT4var20 fCIDR curve is shown on the left; the HB3var03 fCIDR on the right. From these curves a concentration of 8×10^{-11} was chosen for both fCIDRs. Data and figure by Gareth Watson.

This assay protocol was tested by incubating the fCIDR with different concentrations of soluble EPCR before adding these mixtures to the plate (Figure 6.3). These mixtures were incubated with the EPCR-coated plate for one hour before washing. The fluorescence of the plate, measured after washing, decreases as the soluble EPCR concentration increases. This shows that soluble EPCR binding to the fCIDR prevents fCIDR from binding to the immobilised EPCR. These test experiments show that the assay is very sensitive to even small amounts of sEPCR. This assay was used to assess the success of the immunisation strategies.

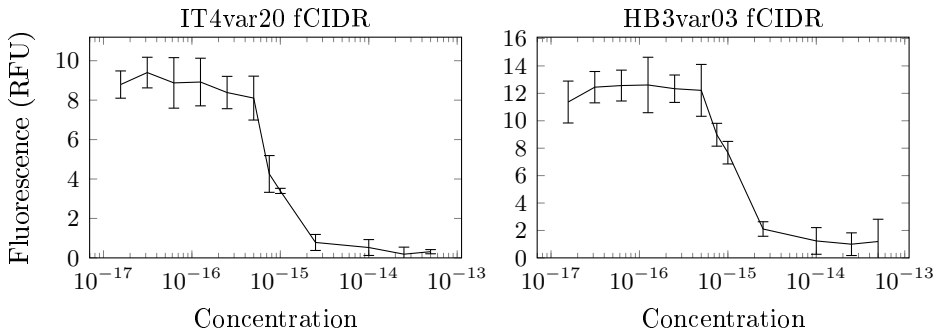


Figure 6.3: Proof-of-principle experiment showing soluble EPCR can prevent both IT4-var20 and HB3var03 fCIDR from binding to immobilised EPCR with 95% confidence limits. Data and figure by Gareth Watson.

6.4 Whole CIDR domains as immunogens

As a reference for the following immunogen strategies, rats were immunised with the HB3var03 CIDR domain. The rats were immunised subcutaneously three times at 3 week intervals with 20 µg/10 µg/10 µg of protein in Freund's incomplete adjuvant. Sera were harvested from the rats two weeks after the final immunisation and tested for reactivity against CIDRα domains from HB3var03 and IT4var20, and for the inhibition of two CIDR:EPCR interactions, using the two assays described above. The reactivity ELISA shows that there are antibodies in the polyclonal sera that react with both CIDR variants, whilst not reacting significantly to the His-tagged control, EPCR. Preliminary inhibition assays (n=1) show that these sera samples inhibit the interaction of the HB3var03 CIDR to EPCR in a concentration-dependent manner, but not the interaction between IT4var20 CIDR and EPCR, suggesting the cross-reactive antibodies are not cross-inhibitory.

6.5 Glycan-masking of CIDR domains

6.5.1 Construct design

The glycosylated construct (gCIDR) was based on the CIDRα domain from HB3var03. We used the crystal structure of this domain to identify sites to introduce the N-linked glycosylation motif, N X T/S, where X is a small uncharged residue (not proline) [317].

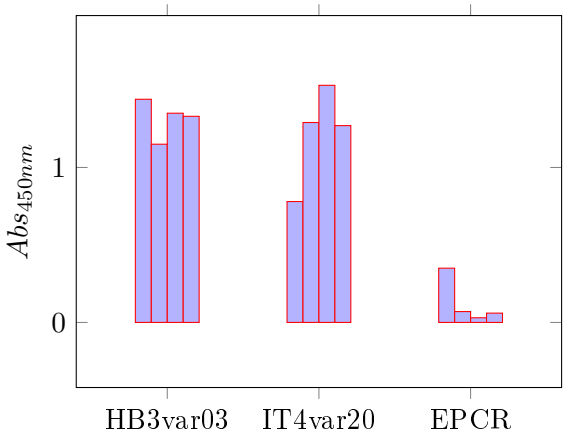


Figure 6.4: ELISA showing reactivity of ∝-HB3var03CIDR sera to HB3var03 and IT4var20 CIDR domains, and EPCR as a His-tag control. Sera from four mice per immunogen were tested in each case.

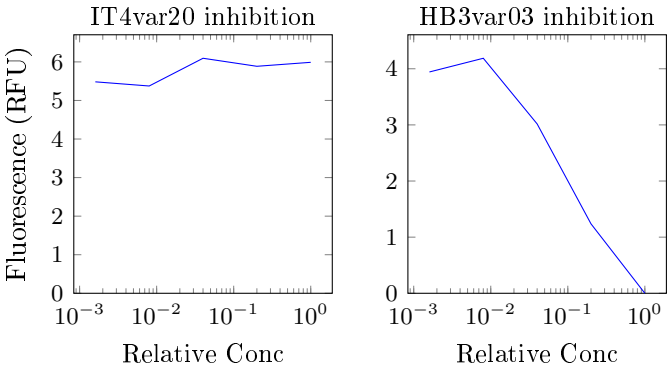


Figure 6.5: Preliminary inhibition assay comparing the inhibitory activity of sera raised against the HB3var03 CIDR produced in bacteria. Sera were tested for inhibition of the IT4var20 CIDR(left) and the HB3var03 CIDR (right). Relative concentration refers to the sera, incubated with fluorescent CIDR, with a value of 1 (log value of 0) denoting undiluted sera

Seven potential glycosylation sites were introduced into the construct, with the majority of sites in loops between secondary structure features. One glycosylation site was introduced in the β-sheet, and one at the C-terminus of the protein (Figure 6.6). This construct was synthesised by Geneart, optimising the codon usage for expression in S2 cells, since a eukaryotic system is required to glycosylate the construct on expression. This construct was subcloned into the pExpres2-1 backbone, previously converted to include a histidine tag and TEV cleavage site.

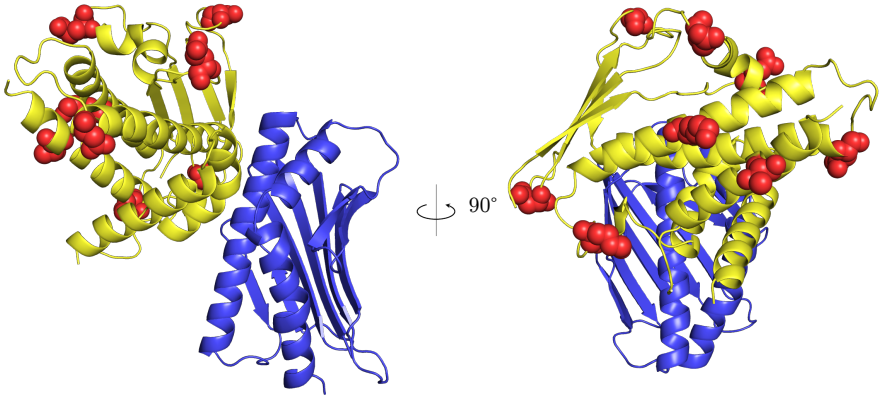


Figure 6.6: Model of the HB3var03 CIDR (yellow) in complex with EPCR (blue). Engineered glycosylation sites are marked by side chains in sphere representation (red).

6.5.2 Expression in S2 cells

This construct was transfected into S2 cells to create a stable S2 line that expresses the glycosylated HB3var03 CIDR domain. After initial transfection, the cell line was selected for three weeks with Zeocin, then checked for expression by Western blot. Subsequently 1 L of S2 cells was grown, with the supernatant harvested four days after the final split. This supernatant was loaded onto a TFF, exchange the media for 20 mM Tris pH 8; 500 mM NaCl, then incubated with 3 ml Ni-NTA for 15 h at 4°C. Protein was then washed on the nickel-NTA beads, then eluted in 20 mM Tris pH 8; 500 mM NaCl, 200 mM imidazole. A comparison of these elution fractions with the HB3var03 CIDR domain, expressed by bacteria, show that the gCIDR is glycosylated (Figure 6.7). Multiple bands in the elution fractions suggest multiple glycosylation states. The eluant was concentrated and gel filtered on a Superdex 75, with the majority of protein eluting as a monodisperse peak (Figure

6.8).

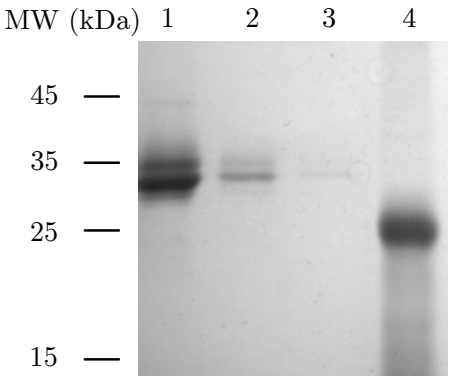


Figure 6.7: An SDS-PAGE gel comparing the gCIDR Ni-NTA elution fractions (1, 2, 3) with the bacterially produced HB3var03 CIDR domain (4).

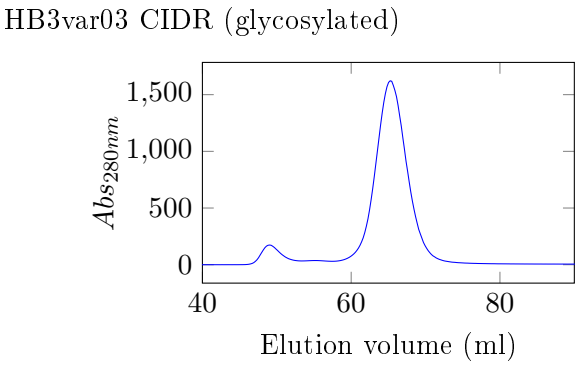


Figure 6.8: Preparative gel filtration shows the gCIDR to elute as a monodisperse peak (at 65 ml).

To check whether the protein was indeed glycosylated, a mixture of endoHf and endoF3 was added to the CIDR, incubating for 15 h. These two deglycosidase enzymes would cleave any glycans on the protein, with endoHf cleaving non-fucosylated high mannose glycans and endoF3 cleaving fucosylated high mannose glycans after the proximal N-acetyl-D-glucosamine residue. SDS-PAGE analysis before and after deglycosylation shows that the CIDR is glycosylated, although it does not reveal how many sites have glycans attached (Figure 6.9).

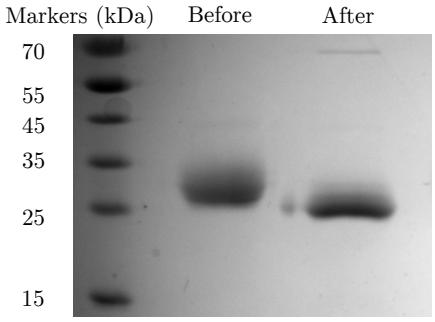


Figure 6.9: Deglycosylation of the gCIDR was verified by SDS-PAGE analysis. Comparison of bands before and after deglycosylation shows a band shift, signifying the removal of glycans.

6.5.3 Confirmation of glycosylated CIDR interaction with EPCR

I confirmed the binding activity of gCIDR using the SPR assay designed in Chapter 3. The glycosylated construct was flown over an EPCR-coated chip at different concentrations (Figure 6.10). The observed K_D, $k_{association}$ and $k_{dissociation}$ show the gCIDR binds to EPCR with a similar affinity to the wild-type CIDR (Table 6.2). This confirms that the addition of glycans to the CIDR domain does not alter EPCR-binding activity of the CIDR domain.

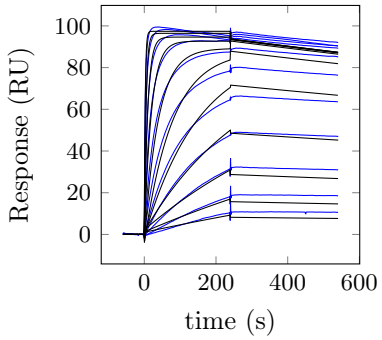


Figure 6.10: Surface plasmon resonance traces showing the gCIDR binding to EPCR. Each dataset consists of binding data for 2-fold dilution series from 1 µM, with blank-subtracted data shown in blue and kinetic fits shown in black.

6.5.4 Assessment of glycosylated CIDR as an immunogen

We immunised 4 rats with the glycosylated CIDR (similar regimen to the whole CIDR, Section 6.4), harvesting sera two weeks after the final immunisation. I used an ELISA

CIDR species	Range	K_D (nM)	k_{assoc} (M^{-1} s^{-1})	k_{dissoc} (s^{-1})
wild type	1 µM - 0.9 nM	0.37	5.32×10^5	1.97×10^{-4}
gCIDR	1 µM - 0.9 nM	0.58	4.03×10^5	2.35×10^{-4}

Table 6.2: SPR data summary table for the glycosylated CIDR domain, tested for EPCR binding. Wild-type parameters are included for reference.

to assess the reactivity of sera raised using the glycosylated CIDR to both the HB3var03 CIDR and the IT4var20 CIDR (Figure 6.11). This assay shows that the glycosylated CIDR raises antibodies that are reactive to HB3var03 but no longer reactive against IT4var20, compared to the antibodies raised against the wild type CIDR, HB3var03. No sera had significant reactivity against EPCR, a negative control protein with a His-tag. Preliminary inhibition assays showed that the sera inhibited the HB3var03 CIDR:EPCR interaction in a similar concentration-dependent manner to the sera raised using the whole CIDR. However this sera did not inhibit the IT4var20 CIDR:EPCR interaction. This is consistent with the observation that antibodies raised against the whole CIDR are cross-reactive, but not cross-inhibitory.

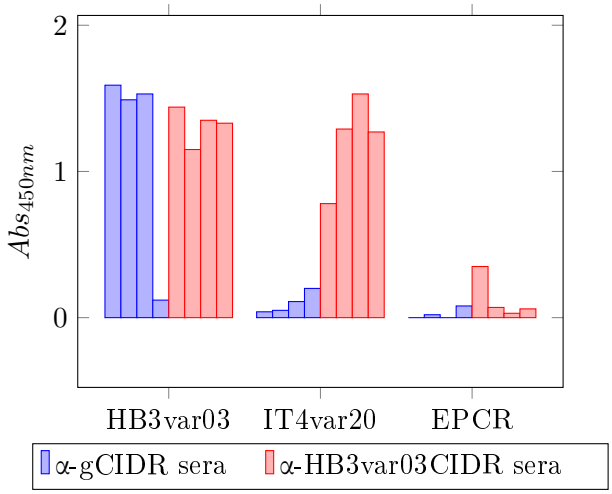


Figure 6.11: ELISA showing reactivity of α-gCIDR and α-HB3var03CIDR to HB3var03 and IT4var20 CIDR domains, and EPCR as a His-tag control. Sera from four mice per immunogen were tested in each case.

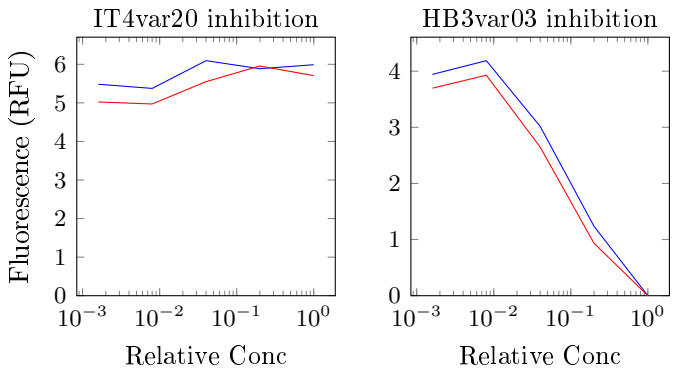


Figure 6.12: Preliminary inhibition assay comparing the inhibitory activity of sera raised against the glycosylated CIDR (red) to sera raised against the CIDR produced in bacteria. Relative concentration refers to the sera, incubated with fluorescent CIDR, with a value of 1 denoting undiluted sera

6.6 Stapled helical peptide

We designed our stapled helical peptide based on the central phenylalanine and the succeeding peptide sequence (Figure 6.13). This was based on the rationale that whilst a helical peptide has previously been stapled successfully, replicating the kinked angle of the helix comprising of the majority of the binding residues using a molecular staple is not currently possible. As a result of this, we chose to synthesise the peptide sequence C terminal from the pair of phenyalanines (F655 to E664) Our peptide sequence, FFQVIYKFNE was synthesised by Biosynthesis, substituting the V658 and F662 to the peptide staples, (S)-2-(4-pentenyl)alanine. We also mutated the sequence in two positions, N663E and E664K, to improve solubility of the peptide. These residues at the C-terminal end of the peptide, away from the binding interface.

6.6.1 Biophysical characterisation of the stapled peptide

To assess the conformation of the stapled peptide, we collected circular dichroism spectra for the peptide between 200 nm and 260 nm. This spectra was analysed for α-helix and β-sheet content using the K2D3 webserver [261], which suggested the protein to have 57% α-helix and 4% β-sheet. While this helical content is inconsistent with the entire stapled peptide adopting a helical conformation, the data are consistent with a helical stretch of

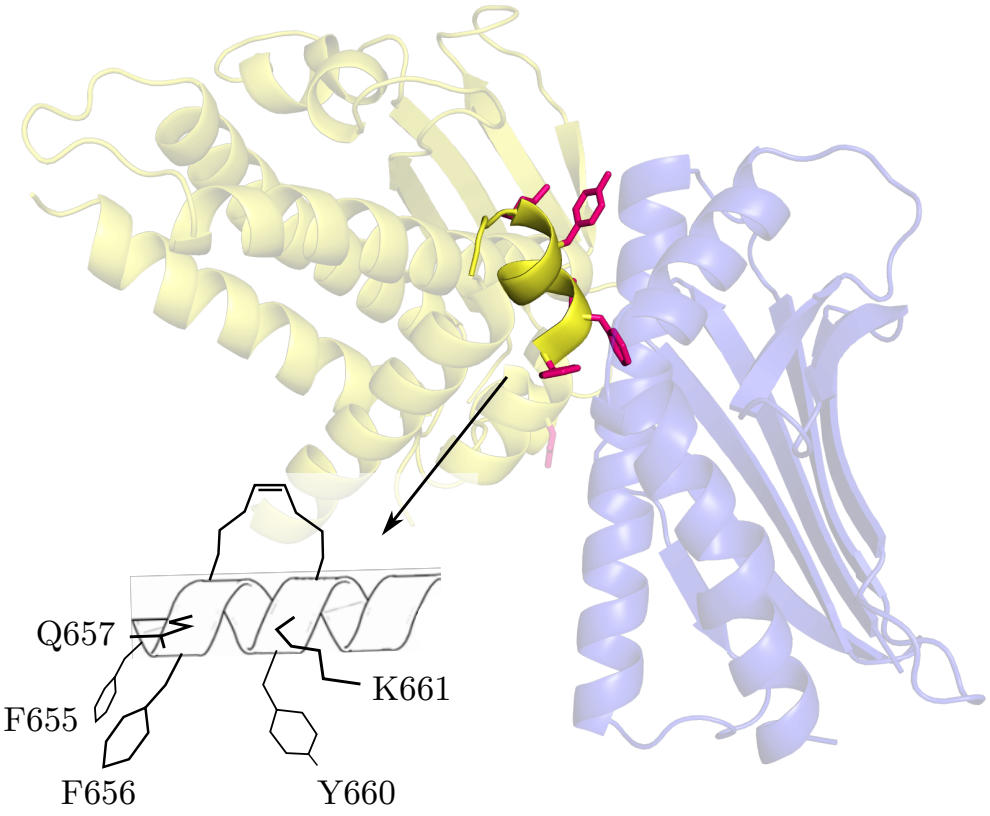


Figure 6.13: The helical motif selected for the stapled helical peptide (yellow, solid) is shown in the context of the HB3var03 CIDR(yellow):EPCR(blue) interaction. Interacting CIDR sidechains are shown in stick representation (pink). The stapled helical peptide design is shown bottom left, with the interacting side chains labelled.

amino acids including the two stapled linkers in a helical confirmation.

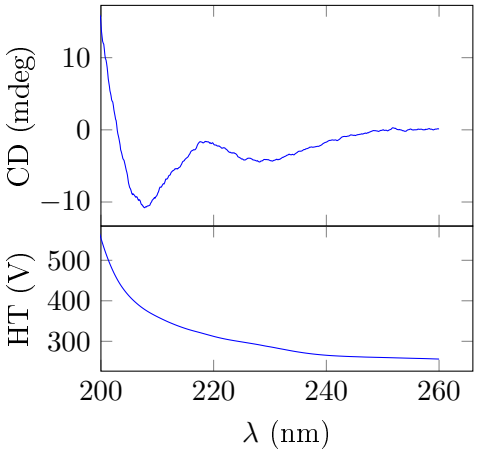


Figure 6.14: Circular Dichroism trace of the stapled helical peptide. Data were disregarded if high tension (HT) voltage was over 700nm.

Due to the small size of the peptide, I collected ITC data for the stapled helical peptide with EPCR to investigate binding activity of the peptide to EPCR. In this experiment EPCR at 500 µM was injected into a cell containing the peptide at 50 µM. I observed little heat change upon addition of EPCR into the peptide solution (Figure 6.15). This dataset shows the peptide does not interact with EPCR at these concentrations.

6.6.2 Assessement of stapled peptide as an immunogen

We immunised 4 rats with the stapled helical peptide (similar regimen to the whole CIDR, Section 6.4), harvesting sera two weeks after the final immunisation. This sera was tested for inhibition of the HB3var03 CIDR:EPCR interaction using the fluorescence assay. This assay showed that the sera, raised against the peptide immunogen, did not inhibit the CIDR:EPCR interaction. (Figure 6.16). Due to this result, and the fact that the peptide showed no binding activity to the CIDR, this immunogen strategy was not developed further.

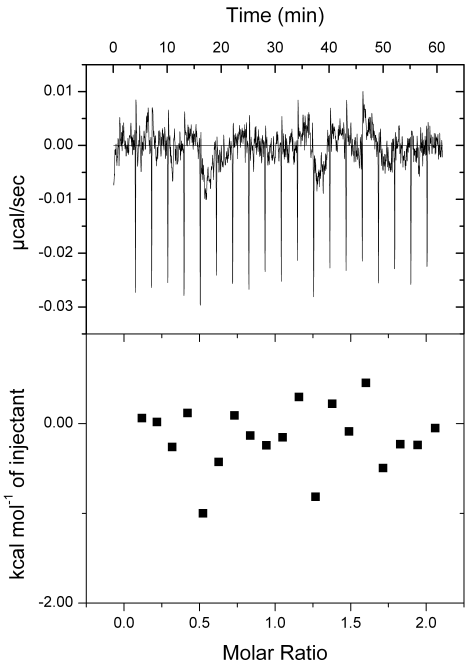


Figure 6.15: Isothermal Calorimetry, injecting the EPCR into peptide suggests no binding at the expected stoichiometry (1:1)

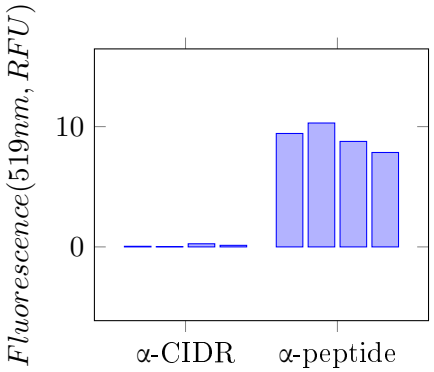


Figure 6.16: Comparison of α-peptide sera and α-CIDR sera, testing the ability to inhibit the HB3var03 CIDR:EPCR interaction. Sera from four mice per immunogen were tested in each case.

6.7 Binding site epitope grafting

I investigated methods to take the binding epitope from the CIDR and graft it onto a scaffold. Since seven of the nine residues, identified as responsible for CIDR:EPCR binding were on the binding arm, consisting of the kinked helix-turn helix motif, I targeted this epitope for grafting studies.

6.7.1 Justification of binding motif

To justify this choice of binding motif, the two interacting residues not present in the chosen motif, D576 and K642 in the HB3var03 CIDRα1 sequence, were mutated to alanine to ensure these two residues were not critical to the interaction (Figure 6.17). These two mutations were introduced separately and in combination by site-directed mutagenesis, expressing each mutant construct in B834 cells, purifying under similar conditions to the wild-type CIDR domain.

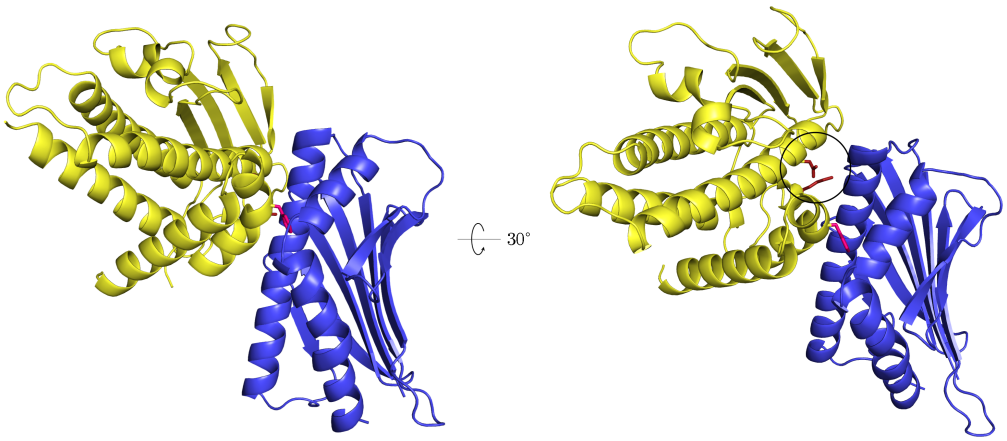


Figure 6.17: Crystal structure of the HB3var03 CIDR:EPCR, highlighting the two binding residues (D576 and K642, pink, circled) not on the kinked helix.

These mutants were tested using surface plasmon resonance to assess binding to EPCR, using the CAP chip as described in chapter 3. Two-fold dilution concentration series were measured from 250 nM to 0.9 nM for each mutant, with binding observed in each case. Subsequent kinetic analyses reveal that the two single site mutants bound to EPCR with a similar kinetic constants and thermodynamic affinity to the wild type, showing that

mutating the residues individually does not affect the binding to EPCR. Kinetic analysis of the double mutant revealed that the two single site mutants combine synergistically in their interaction with EPCR. The analysis showed the double mutant to bind EPCR with a K_D of 220 nM, with a lower $k_{association}$ and a higher $k_{dissociation}$ rate constant than the wild type. I decided to continue on with the epitope grafting despite this drop in affinity, as this affinity value for this CIDR mutant was no lower we observed some of the tested natural variants.

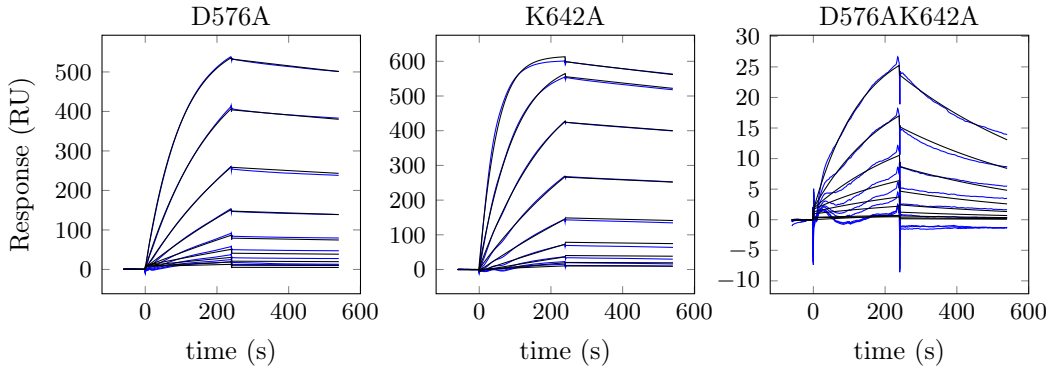


Figure 6.18: Surface plasmon resonance traces looking at D576A, K642A and D576A-K642A double mutant binding to EPCR. The D576A and K642A datasets consist of binding data for 2-fold dilution series from 1 µM, while the double mutant dataset consists of binding data for 2-fold dilution series from 250 nM. Blank-subtracted data is shown in blue, with kinetic fits shown in black.

CIDR mutant	Range	K_D (nM)	k_{assoc} (M^{-1} s^{-1})	k_{dissoc} (s^{-1})
wild type	1 µM - 0.9 nM	0.37	5.32×10^5	1.97×10^{-4}
D576A	1 µM - 0.9 nM	5.25	3.87×10^4	2.02×10^{-4}
K642A	1 µM - 0.9 nM	2.17	1.05×10^5	2.27×10^{-4}
D576AK642A	250 nM - 0.9 nM	220	9.10×10^3	2.00×10^{-3}

Table 6.3: SPR data summary table for D576A and K642A mutants, tested for EPCR binding. The wild type summary is included for reference.

6.7.2 Scaffold search

I used the programs PDBeFold and DALI to search for proteins with structurally similar segments to assess the possibility of epitope grafting by sidechain mutation or backbone replacement. These searches showed whilst the angle of the kink was adopted in several

structures, none were presented on the surface in a conformation compatible with the presentation of the interacting residues. As a suitable scaffold could not be found, I turned to the Rosetta-based design.

6.7.3 Computational design of three helical bundles immunogens

I based my workflow on the Fold-From-Loops protocol [252]. I attempted to graft the binding helix (from N648 to K678) to a loop in three-helix bundle, which included seven interacting residues, and the three inwards-facing residues that stabilise the kinked conformation (Figure 6.19, A). The bundle scaffold (PDB 3LHP, chain S) was selected based on previous papers, which show that this bundle is produced in high quantities in bacteria in the soluble fraction (Figure 6.19, B). The precise epitope boundaries were chosen by alignment of the helical termini to the three-helical bundle. This alignment revealed that the overlapping helical portions overlay with an RMSD of 2.5 Å. The two termini of the epitope were selected based on the proximity of the backbone polypetide to the equivalent polypeptide backbone in the bundle. These alignments led to the hybrid bundle-epitope input sequence, with the epitope sequence inserted in place of the bundle sequence.

The first step was to use Rosetta's main program to fold the bundle sequence around the epitope, creating an initial model that could be stabilised. I modified Rosetta's functionality using a script, available with the Rosetta software (Appendix B), [248]. In this protocol the binding epitope is held, locked in position, while the remaining sequence from the helical bundle is generated and folded around this locked epitope. This initial step is supplemented with C_α-C_α constraints to drive the model to the target conformation, as the input sequence may not be stable. These C_α-C_α constraints were generated from the three-helix bundle PDB file, excluding the substituted segment, using a simple Python program (Appendix B). Ten thousand structures were generated to provide an initial model for sequence design. After each complete model is generated, the epitope is conformationally unlocked, and a relaxation step is performed, allowing the model to adopt a conformation in the local energy nadir.

The results of this first step are shown in Figure 6.20. The ten thousand output structures

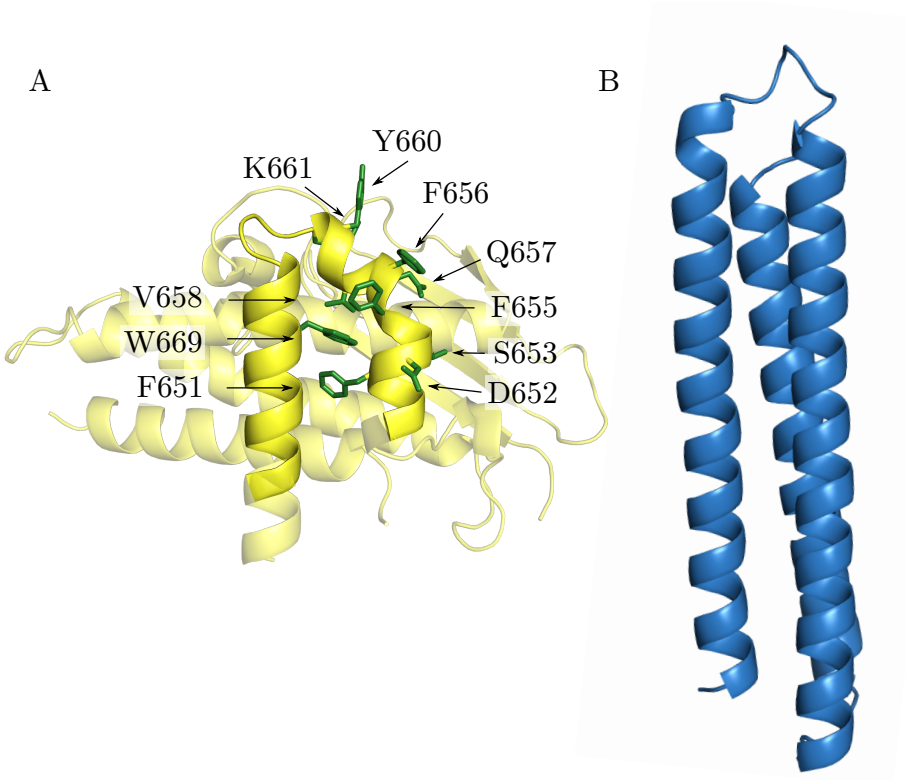
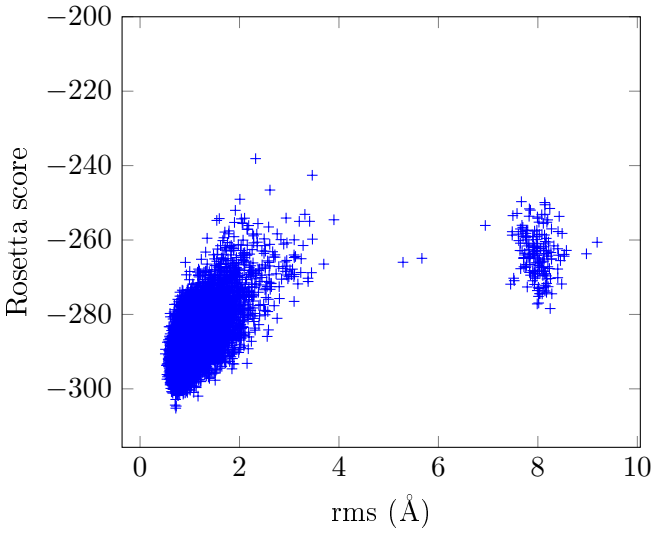


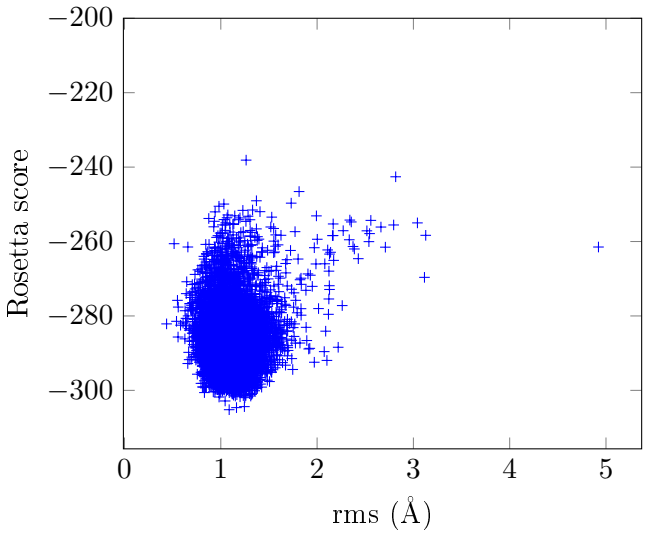
Figure 6.19: Models of the binding epitope (A) in the context of the CIDR and the target scaffold (B). Residues responsible for the interaction, or for conformation of the motif are shown in stick representation (green), with the residue identities shown.

were plotted comparing the Rosetta score, a measure of the intramolecular interaction strength, with the root mean standard deviation (rmsd) compared to the target scaffold in one plot, and the rmsd compared to the epitope in another. The first plot, comparing Rosetta score to rmsd (scaffold) shows the models generated broadly cluster into one of two groups. The first group with rmsd 0-4 Ångströms were similar to the scaffold conformation, whilst the second group, with rmsd 7-9 Ångströms had the third helix flipped round to the front of the other two. Within the first group, models differed in the precise orientation of backbone, but were largely similar. The second plot, comparing Rosetta score to rmsd of the binding epitope, showed the majority of the epitopes to be similar to the desired conformation, with an rmsd of 0.5-3 Ångströms (Figure 6.20b).

The next step in the Fold-from-loops protocol is sequence design, to stabilise the conformation of the immunogen. To maximise the chances that the target epitope conformation would recapitulated, I selected the model with the lowest rmsd compared to the target epitope as a starting point for this process. This sequence design used a custom Rosetta script (Appendix B) using the RosettaScripts program [292]. I allowed every residue to mutate apart from interacting residues and the three inwards-facing residues that stabilise the kinked helix. These mutations were semi-random, assigning each residue with one of three categories, core, boundary and surface residues based on their locations and allowing each residue to mutate to any other residue appropriate for that position. Four cycles of mutagenesis were completed using the flxbb mover; each cycle consisting of a round of mutation then a relaxation step, to allow the model to adopt a low energy, local conformation. This sequence design included C_α-C_α constraints for the epitope, to bias the sequence design to maintain the epitope conformation. Ten thousand models were generated by this step with different primary sequences. These models were filtered firstly according to the Rosetta score, selecting only the models that had a lower Rosetta score than the input model, then filtered for models that had a packstat score greater than 0.5. The packstat filter assesses the stability of the core of the protein by looking at how well the core of the protein is packed [293]. As a final design step, these constraints were removed, then the model relaxed, to allow the model to adopt a low energy confirmation without these constraints. These models were filtered once more with the packstat filter, selecting models



(a) Comparing Rosetta score against rmsd from the scaffold.



(b) Comparing Rosetta score against rmsd from the target epitope.

Figure 6.20: The ten thousand output models were scored using the Rosetta software, and compared against the structures of the scaffold and the target epitope, calculating the rmsd from each structure.

with a score greater than 0.55 as the final model set.

Initial test runs of this protocol (outputting ten structures) resulted in models with relatively large rmsd values (>3 Å), compared to the binding epitope (Figure 6.21, left). Examination of the output pdb files showed that loop in the epitope, between the kinked helix and the following helix does not pack properly, affecting the conformation of the kinked helix side chains. As a result I added an N-terminal extension to the scaffold, which was designed to provide some support at the back of the epitope (Figure 6.21, right). Using this new sequence with the extended N-terminus, ten thousand instances of sequence design were run, producing 771 sequences that passed the filters.

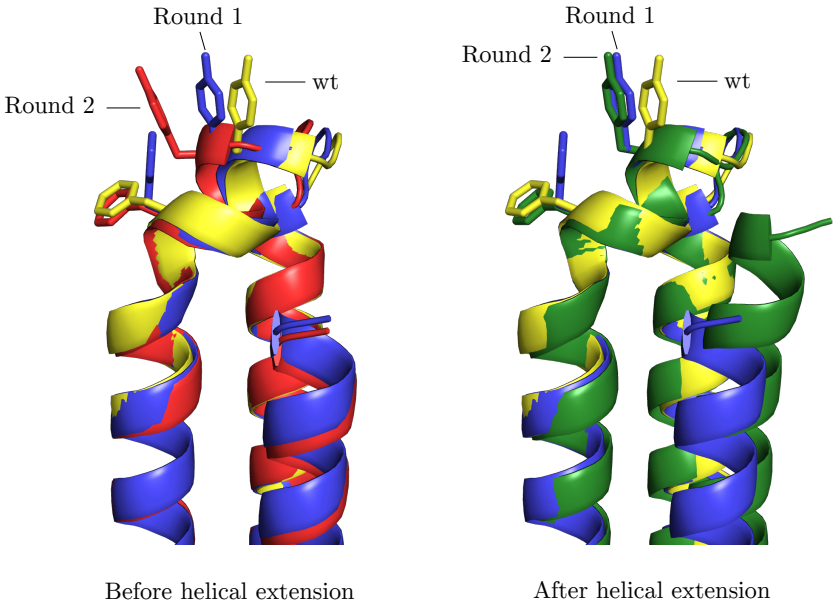


Figure 6.21: Initial trials of the sequence design step (Round 2, red) resulted in large rmsd values compared to the wild type (wt, yellow), due to differences in the top of the target epitope (left). After the addition of an N-terminal extension, the sequence design step (Round 2, green) outputs structures closer to the wild type (right).

These output models were then assessed to choose a selection of bundle immunogens to express and test. A plot of Rosetta score versus the rmsd compared to the epitope visualised the stability of the dataset as well as the similarity to the epitope confirmation. The models were fairly closely clustered, with Rosetta scores between -320 and -355, and rmsd values between 0.6 and 1.7.

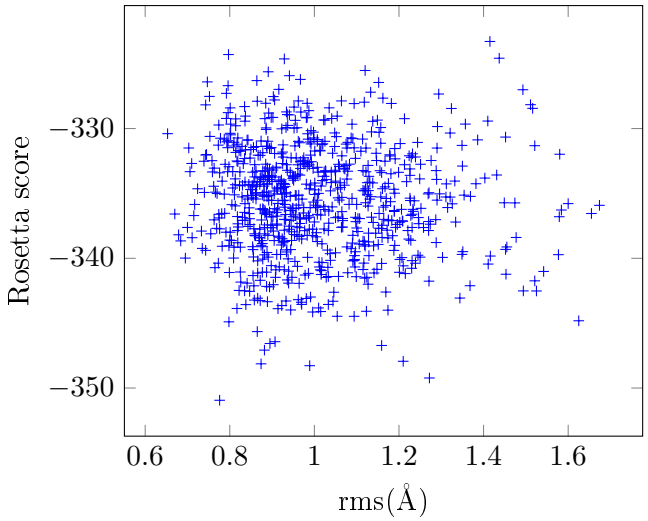


Figure 6.22: Output models from the sequence design stage that pass the filters are plotted as Rosetta score versus the rmsd of each output compared to the epitope.

Since the output models have as little as 35% sequence identity, I chose to cluster the output models by sequence before choosing the best immunogens to test. I divided the top 100 immunogens, based on rmsd to the binding epitope, into eight groups based on sequence similarity (Figure 6.23), then selected the best immunogen in each group based on a combination of the lowest rmsd value, compared to the epitope, and the highest packstat filter score. These 8 immunogens were named as they were for the sequence design, 18, 24, 321, 398, 469, 555, 631, 866, with sequence identity ranging from 38% to 60% (Figure 6.24).

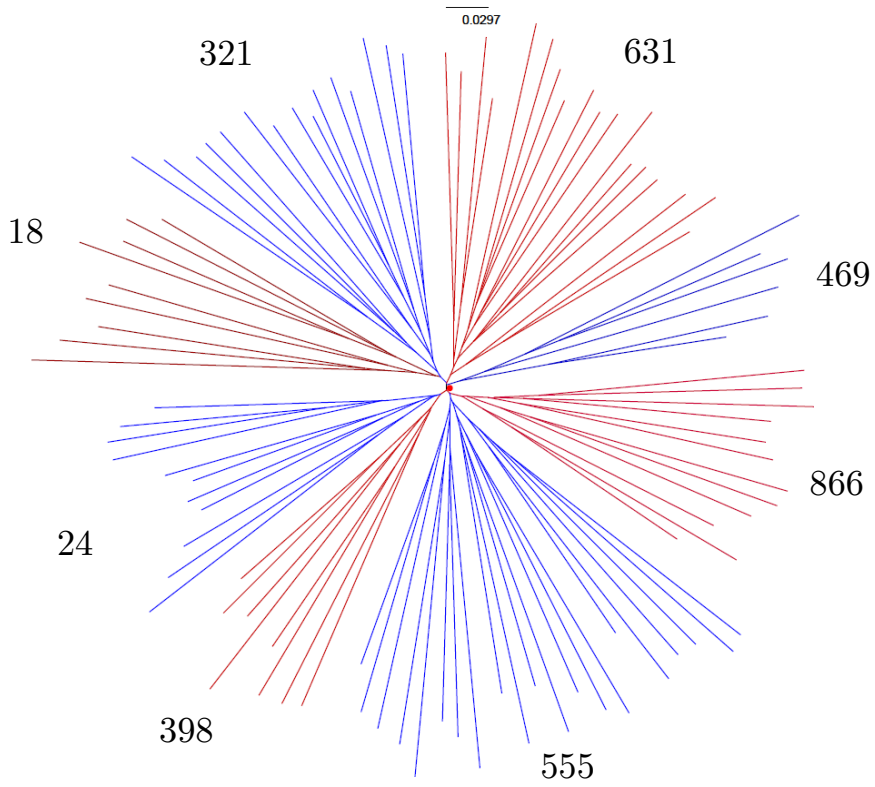


Figure 6.23: Cladogram showing the sequence diversity of the top 100 sequences (generated by MUSCLE). The selected bundle immunogen for each clade is shown outside the respective clade.

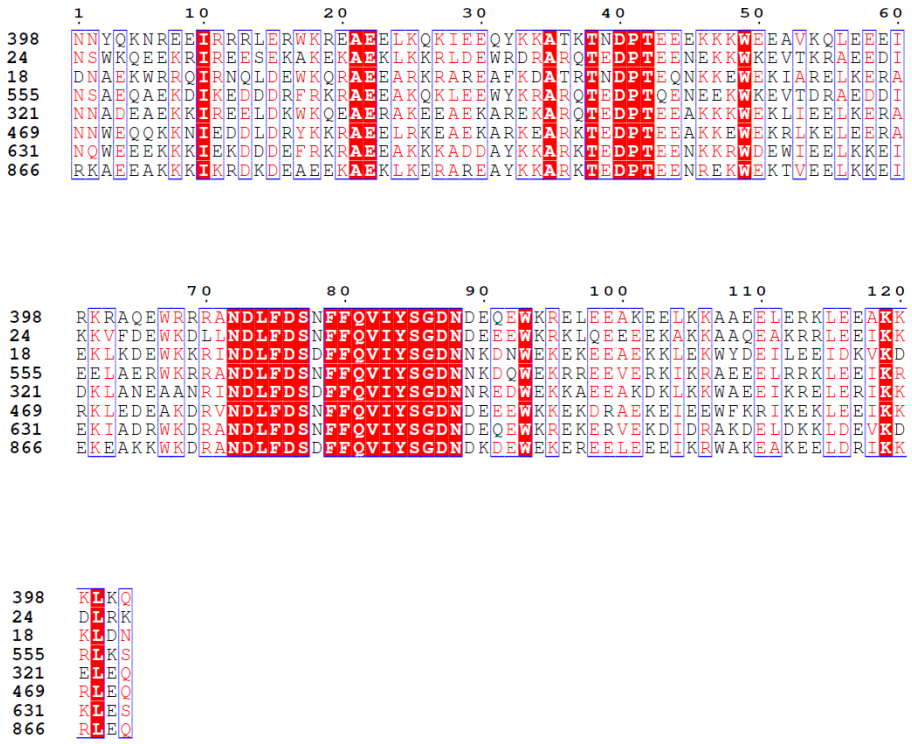


Figure 6.24: Sequence alignment of the eight selected immunogen, showing the conserved epitope residues (between residues 72 and 93) amongst the diversity generated by the sequence design.

6.7.4 Expression of bundle immunogens

The 8 immunogens were synthesised by GeneArt. I optimised the codon usage for *Escherichia coli*, and subcloned the genes into the pEt15b(converted) backbone for expression. These immunogens were tested in *E. coli* cell strain, BL-21. Expression trials showed that all constructs apart from bundle immunogen 18 were expressed (Table 6.4). The majority of these bundles expressed a large amount of protein in the insoluble fraction, with little or no protein in the soluble fraction.

Large-scale expression was initially attempted for bundles 24 and 555. 8 L of each constructs were grown, with protein purified and prepared from both the soluble and insoluble fractions. Protein was purified from the soluble fractions via their histidine tag, through affinity purification using Ni-NTA. Protein was refolded from the insoluble fraction using on-column refolding in a decreasing gradient of guanidine hydrochloride. Eluted fractions were analysed by SDS-PAGE to assess protein content. In contrast to our expression trials,

Bundle	Chosen expression conditions			
	E. coli strain	Temperature (°C)	Time (h)	solubility
18	BL21	-	-	-
24	BL21	25	3	s/i
321	BL21	37	15	s/i
398	BL21	25	15	s/i
469	BL21	25	15	s/i
555	BL21	25	3/15	s/i
631	BL21	25	15	s/i
866	BL21	37	15	s/i

Table 6.4: Summary table of optimal expression conditions for the bundle immunogens. Bundle 18 did not express.

a large amount of protein was expressed in the soluble fraction, as well as in the insoluble fraction. We subsequently expressed 8 L of each bundle in turn except bundle 18, purifying protein from the soluble fraction. These preparations were gel filtered to ensure purity (Figure 6.25). The gel filtrations in each case show the bundle to elute as a single peak. In the case of 398 and 469, the peak is symmetrical, suggesting a monodisperse, homogeneous protein. However the gel filtration traces of other bundle immunogens show these bundles elute in asymmetric peaks, with the peak eluting at an earlier volume. SDS-PAGE gels show that the molecular weight over the peak was consistent, suggesting that the asymmetric peak may represent a heterogeneous single species.

6.7.5 Characterisation of bundle immunogens

As these bundle immunogens were designed *in silico*, we characterised these bundles extensively to see assess their fold as well as their function. We assessed the initial fold and the fold stability using circular dichroism, collecting spectra at room temperature, then a spectrum every 0.2°C up to 90°C (Figure 6.26). All of the bundle immunogens are largely α-helical at 20°C. The CD analysis software, K2D3, estimates these bundles to have over 90% α-helicity, based on their spectra.

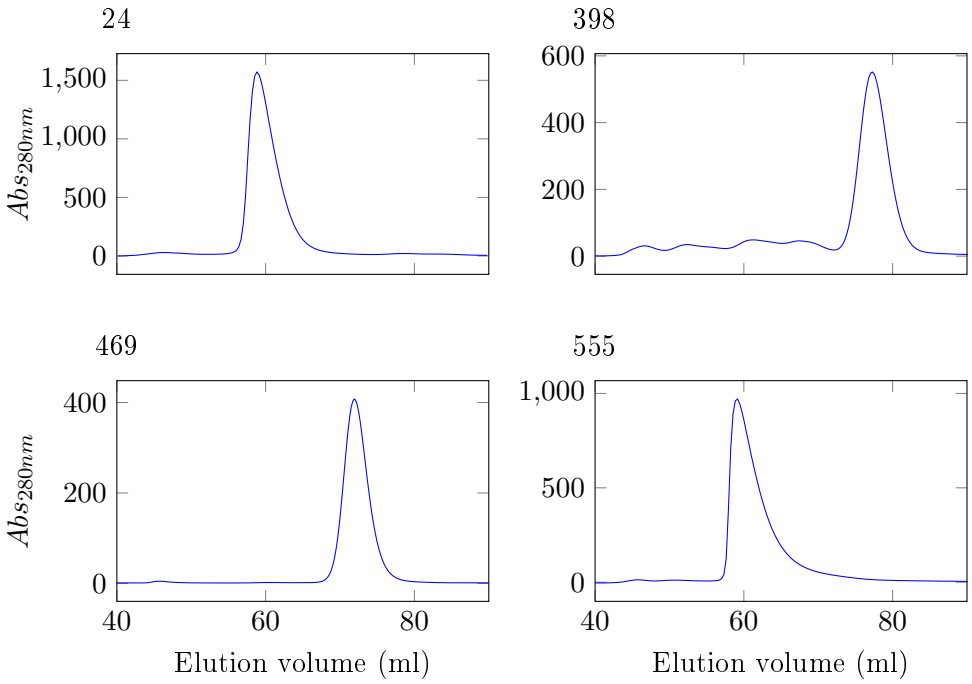


Figure 6.25: Preparative gel filtration shows each bundle immunogen to elute as a singular peak (321, 631 and 866, expressed by Gareth Watson, not shown). Bundles 398 and 469 elute from the column as symmetrical peaks, whereas bundles 24 and 555 elute as asymmetric peaks, suggesting inhomogeneity.

Studying the variance of circular dichroism with temperature allows us to infer the loss of structure with temperature. As the temperature rises, the CD spectrum loses its shape as the intramolecular bonds holding the tertiary fold and the α-helices break (Figure 6.26). This loss of α-helical structure is typically measured at 222 nm (Figure 6.27), with the increase in the CD at this wavelength representing loss of structure. The thermal melts reveal differences between the bundles in terms of stability of structure. This is not surprising since their sequence identity is between 38% and 60%. Looking at the variance in circular dichroism with temperature, 398 and 469 are stable, with 398 retaining 90% secondary structure (as estimated by K2D3) even at 90°C. The CD at 222 nm for these two bundles is consistent with stable bundles - 398 maintains a low CD to 90°C; 469 starts to lose secondary structure at around 70°C. Bundles 555, 631, 866 and 24 all lose some of their helical structure in a cooperative fashion before 60°C, signified by a sigmoidal increase in CD at 222 nm in each case. The α-helical content decreases to around 60%, as estimated by K2D3. 321 has an intermediate stability, not completely unfolded at 90°C (70% alpha helix).

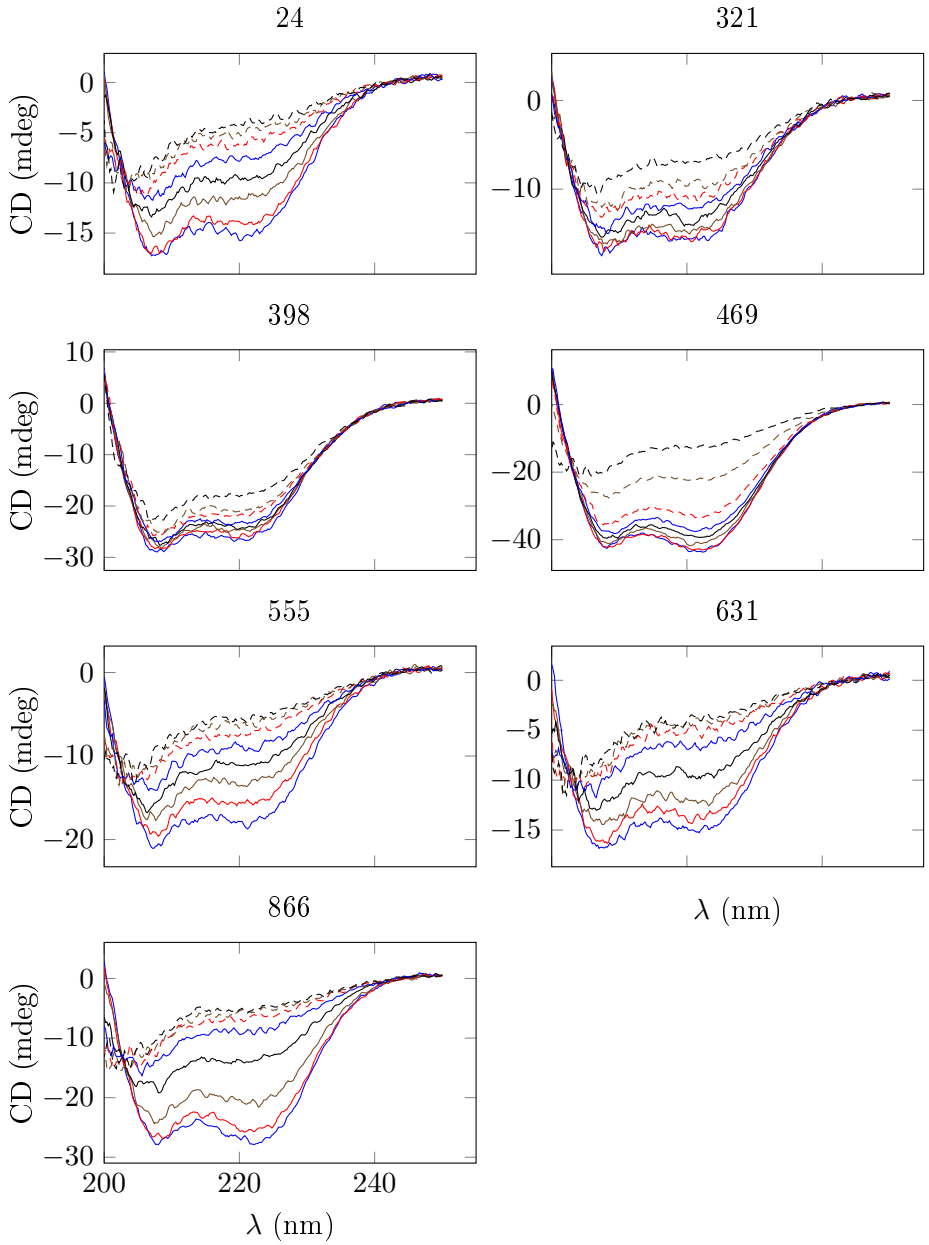


Figure 6.26: Measuring the CD spectra of the bundle immunogens as the temperature is increased reveals loss of α-helicity at higher temperatures. The CD spectra between 200 and 250 nm is plotted every 10°C from 20°C to 90°C.

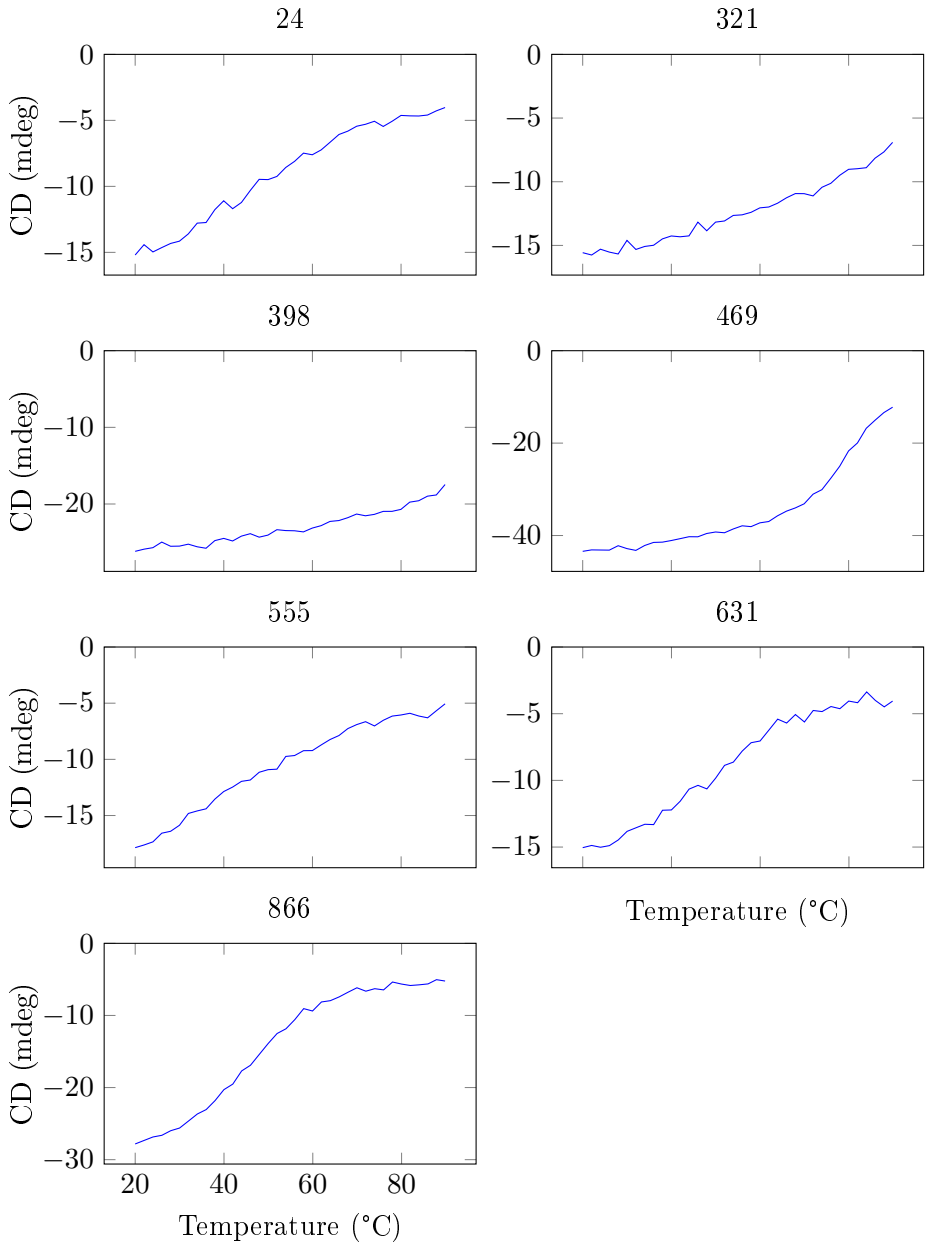


Figure 6.27: Following the circular dichroism at 222 nm between 20°C and 90°C shows the loss of α-helicity of the bundle immunogens.

Since the unstable bundles identified by circular dichroism also gave asymmetric peak in gel filtration, I chose one example of a stable and an unstable bundle to probe the molar mass of these single peaks (Figure 6.28). The stable bundle, 469, gave a symmetric peak in the SEC elution trace, with a consistent molar mass given by MALLS. This molar mass, 15500 Da, matches well with the calculated molecular weight of 15620 Da, suggesting 469 elutes as a single species. In contrast the unstable bundle, 555, elutes from SEC as an asymmetric peak, with the zenith at an earlier elution volume than 469. The molar mass of the species is 21500 Da at the peak zenith, decreasing steadily through the tail, towards 15500 Da. The molecular weight is larger than the predicted molecular weight of 15400 Da for the majority of the peak. This is consistent with bundle 555 in a monomer:dimer equilibrium, with the monomer:dimer ratio of roughly 3:2.

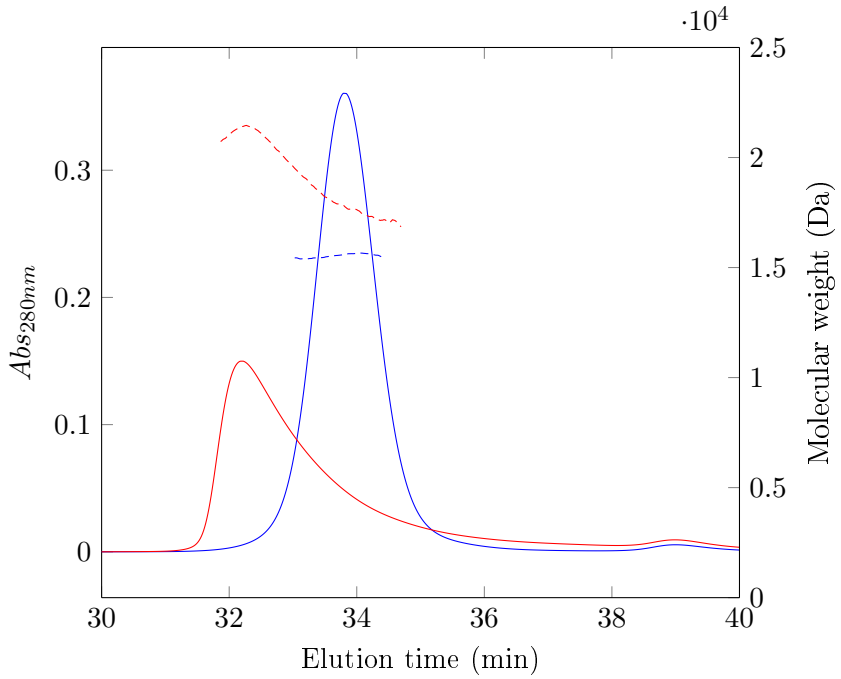


Figure 6.28: SECMALLS of the bundle immunogens, 469 (blue) and 555 (red). Molecular weights are shown as dashed lines, while the absorbance at 280 nm is shown as solid lines.

6.7.6 Characterisation of bundle immunogen binding to EPCR

As well as assessing the stability and fold of the immunogens themselves we tested binding to EPCR, using the SPR set-up described earlier in this thesis. Each immunogen had three concentrations (50 µM, 5 µM & 500 nM) flown over the EPCR-coated chip (Figure 6.29). These concentration curves show that the majority of immunogens have little to no binding to the chip, at concentrations two orders of magnitude above the K_D of the HB3var03 D576AK642A CIDR:EPCR interaction. Bundle 555 showed the most binding with a response of 60 RU. However, an extended data series, flowing 555 concentrations between 3.4 µM and 220 µover EPCR failed to reproduce preliminary experiments.

I also assessed the ability of three of these bundles to pull out inhibitory antibodies from CIDR-immunised rat sera. I chose to test 398 and 469, which were both structurally stable, and 555, which was unstable, yet showed very limited EPCR interaction. Sera from CIDR-immunised rats was reactive against both 398 and 469, with limited reactivity with 555 (Figure 6.30). Due to this, 0.5 mg of HB3var03 CIDR and 469 was covalently bound to cyanogen-bromide activated sepharose (Sigma), using these beads to purify antibodies from the HB3var03 CIDR-immunised rat sera. Samples of the purified antibody (pAb) and unbound antibodies (FT) were tested for inhibition of the CIDR:EPCR interaction using antibody, purified by the HB3var03 CIDR domain as a control (Figure 6.31). These assays showed while the control antibody, purified by the HB3var03 CIDR domain, inhibited the interaction between the fluorescent CIDR and EPCR, antibody purified using 469 did not inhibit the interaction. Rather, the flow through sample from these preparations inhibited the interaction. This suggests that inhibitory antibodies will not be raised if these bundles were used as immunogens.

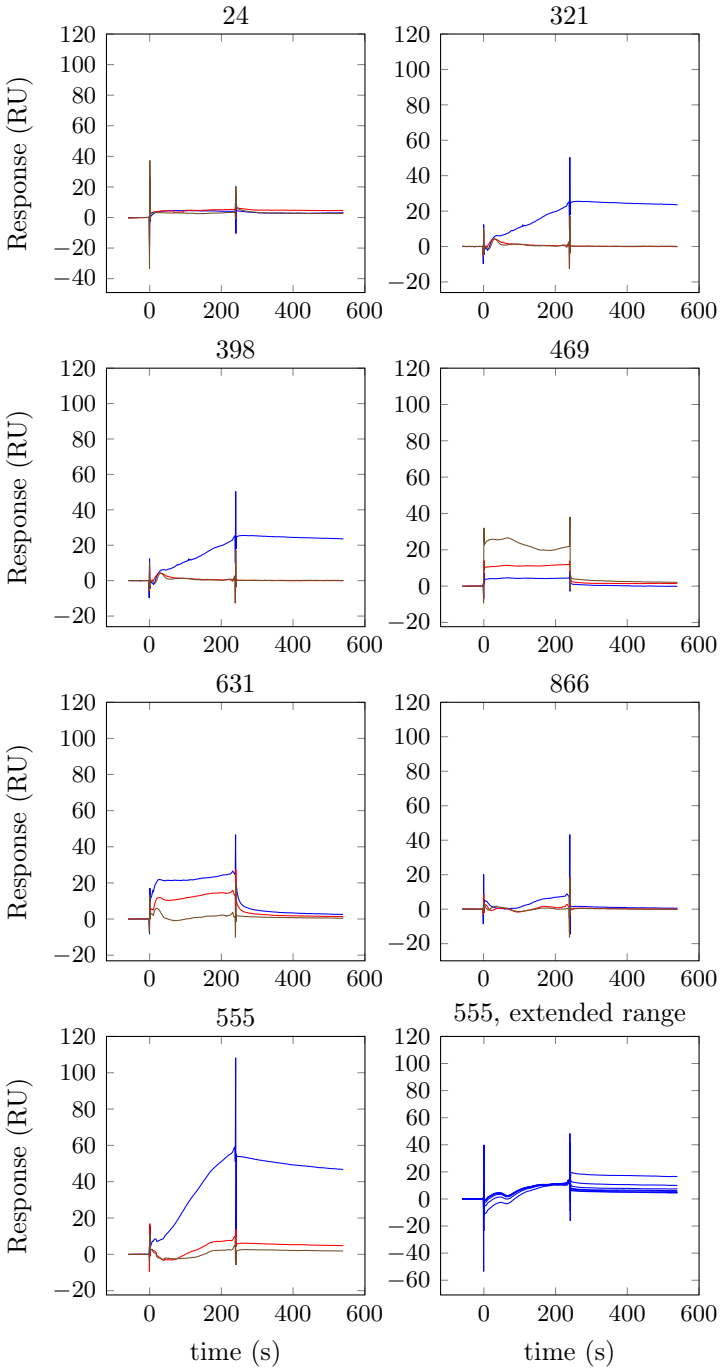


Figure 6.29: Surface plasmon resonance traces, assessing the binding of the bundle immunogens to EPCR. Each bundle was flown over the chip at 50 µM (green), 5 µM (red) and 500 nM (blue). 555, the bundle that gave the largest response, was flown over EPCR at a wider concentration range (220 µM - 3.4 µM in a series of two-fold dilutions) shows no binding of 555 to EPCR (blank-subtracted data in blue).

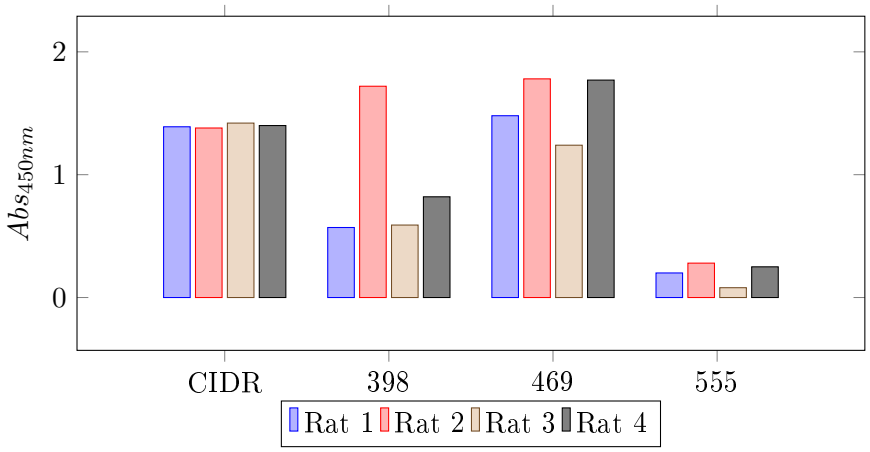


Figure 6.30: Reactivity ELISA of α-HB3var03 CIDR sera against bundle immunogens 398, 469 and 555. Sera from four rats, immunised with the CIDR, was tested against each of the proteins.

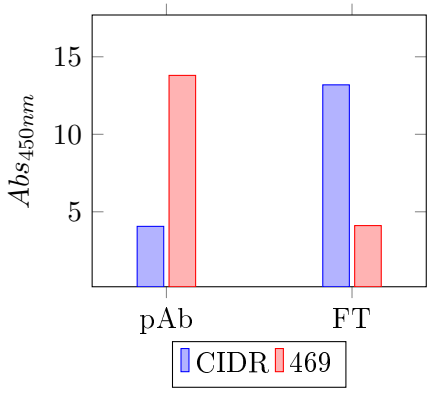


Figure 6.31: Purified antibody from the α-CIDR sera (pAb) and flowthrough (FT) from the purification procedures, using CIDR (blue) or 469 (red), was tested for inhibition of the CIDR:EPCR interaction.

6.7.7 Second round of immunogen design

After identifying that the first round of immunogens neither bound to EPCR nor reacted with inhibitory antibodies from CIDR-immunised rat sera, I looked back at the *in silico* models to investigate whether additional steps were needed to stabilise the desired structure.

I used the AbinitioRelax program, included in Rosetta, to predict the conformation of the immunogen with the highest helicity, 469, generating ten thousand folding trajectories to gain clues as to why these bundles did not bind to EPCR (Figure 6.32). This dataset, when plotting the Rosetta score against the rmsd comparing the model to the output pdb from the sequence design showed that no trajectory in this protocol folded into the desired confirmation, as outputted by the sequence design protocol. The majority of models had an rmsd value between 5 and 13 Å(the model with the lowest rmsd is shown in Figure 6.32b), suggesting that bundle 469 had not folded into the conformation output by the Fold-from-loops protocol.

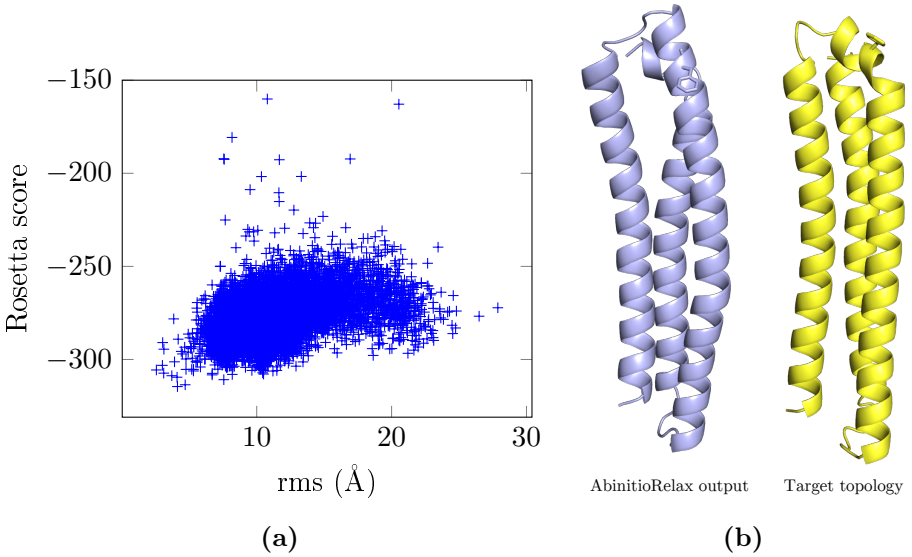


Figure 6.32: *in silico* folding of 469 using the AbinitioRelax program suggests the 469 does not fold into the target topology. The results of the ten thousand trajectories is shown in 6.32a, with no model closer than 3 Å rmsd compared to the target topology. The closest model (6.32b, blue), shows significant differences to the target topology (yellow), especially in the epitope region.

To try and stabilise the desired conformation, I opted to introduce disulfide bonds based

on the orientation of three pairs of residues in bundle 469, the bundle immunogen that was stable up to high temperatures, with the highest helical content. Each pair of residues had a $C_\beta - C_\beta$ distance between 4 and 5 Å, within the ideal distance for a disulfide bond [318]. These three disulfide bonds, named Cys1, Cys2 and Cys 3, were introduced individually using site-directed mutagenesis: mutating I10 and L74 for Cys1; L53 and L115 for Cys2; V71 and A100 for Cys3 (Figure 6.33). In addition, combinations of Cys1/Cys2 and Cys2/Cys3 were created. Combinations Cys1/3 and Cys1/2/3 were not considered, as the close proximity of the disulfide bonds Cys1 and Cys3 may cause disruption of the immunogen fold.

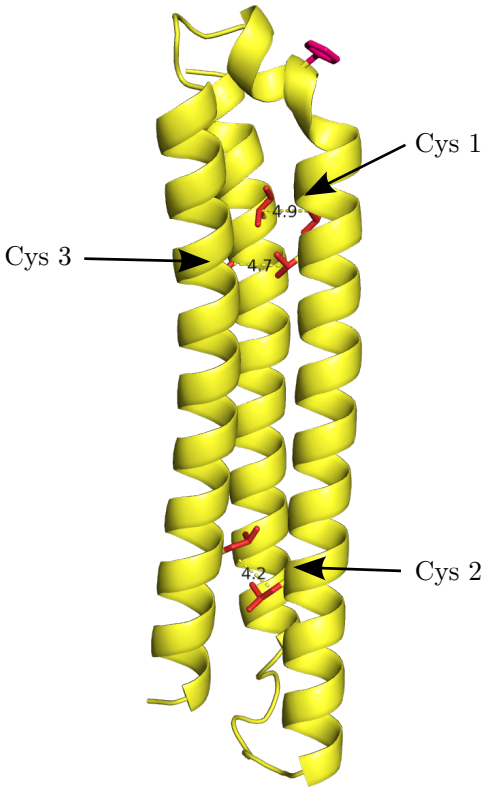


Figure 6.33: Model of bundle 469 showing the mutation sites and $C_\beta - C_\beta$ distances for each disulfide bond in Ångströms. Residues highlighted in red are mutated to cysteine in different combinations to make the mutants. The central phenylalanine is highlighted in pink.

6.7.8 Cysteine mutant expression

Each cysteine mutant was expressed in BL-21 cells, using expression trials to identify optimum conditions, and to see if each mutant construct expressed in the insoluble or

soluble fraction. All constructs except Cys 2 expressed well, with all expressed constructs expressing in the insoluble fraction, and Cys1, Cys1/2 and Cys2/3 expressing in the soluble fraction. Cys2 expressed in a small quantity in the insoluble fraction.

8 L of each Cys mutant were grown in BL-21 cells, according to the optimum expression conditions identified in the expression trials. To maximise yield, protein was purified from both the soluble and insoluble fractions. Protein from the soluble fraction was purified using the histidine tag on Ni-NTA, before unfolding the bound protein using 6 M guanidine. Unfolded protein from the insoluble fraction was then added to the Ni-NTA, refolding in a decreasing gradient of guanidine. Protein was gel filtered, with each protein eluting from the column as a singular monodisperse peak. In each case samples from the peak were checked using SDS-PAGE, showing a species at the correct molecular weight.

6.7.9 Cysteine mutant characterisation

CD spectra were taken for the set of mutant bundles, comparing their stability to the parent bundle, 469. CD spectra were again taken every 0.2°C up to 90°C to assess the thermal stability of the bundles (Figure 6.34). As expected, at 20°C all of the cysteine mutants had very similar spectra to the parent bundle, 469. However, the change in spectra at increasing temperature shows differences in bundle stability. The circular dichroism at 222 nm increases less in Cys1 and Cys3 than in the parent bundle, especially at higher temperatures, suggesting that these two bundles are more stable than 469. Conversely, the circular dichroism of Cys2 and the double cysteine mutants increases faster than the parent bundle at higher temperatures, suggesting these bundles to be less stable than 469.

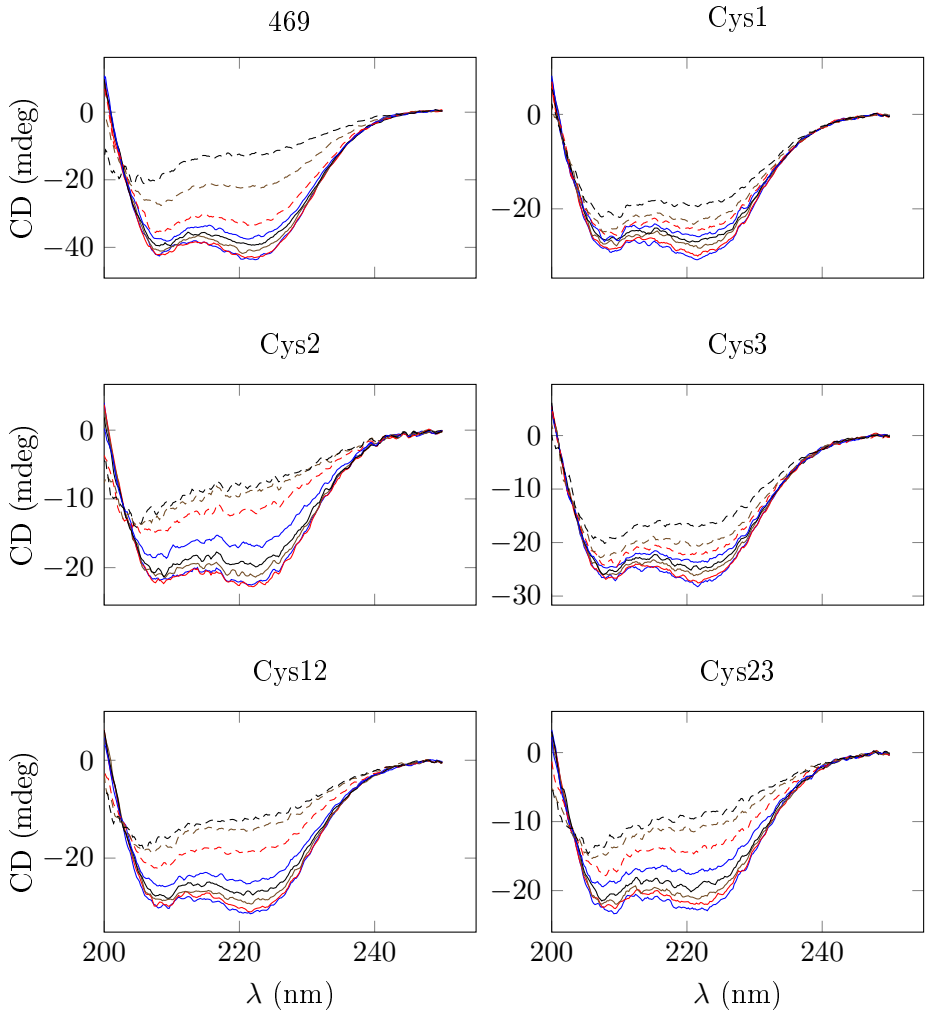


Figure 6.34: Measuring the CD spectra of the second generation bundle immunogens as the temperature is increased shows that Cys1 and 3 are more stable than 469. The CD spectra between 200 and 250 nm is plotted every 10°C from 20°C to 90°C.

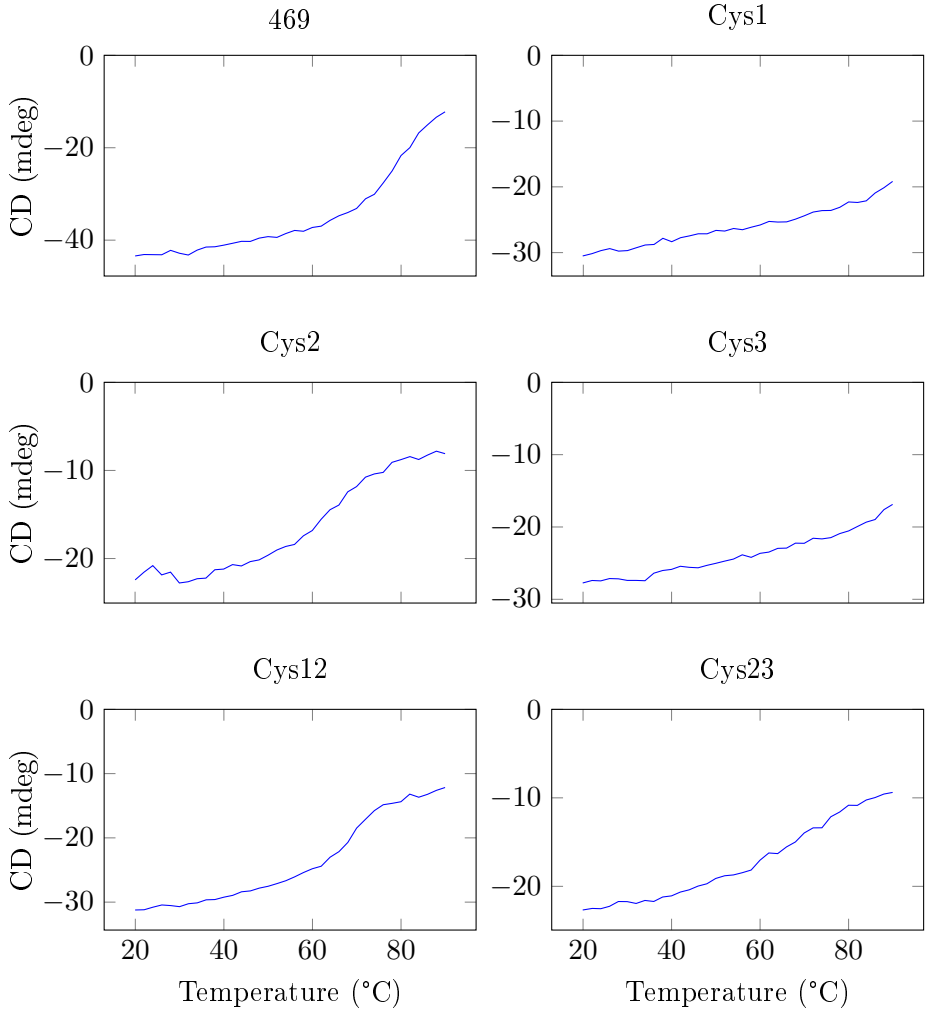


Figure 6.35: Following the circular dichroism at 222 nm between 20°C and 90°C shows that the single disulfide mutants are more stable than the parent bundle, 469, shown here for reference.

6.7.10 Characterisation of cysteine mutant binding to EPCR

These mutants were screened for binding to EPCR using SPR, similar to the first round of mutations. Concentration series from 12 µM to 47 nM were flown over immobilised EPCR. All of the single disulfide Cys mutants showed binding activity to EPCR. While Cys1 and Cys3 bound with extremely fast on and off rates, Cys 2 bound with a slow off rate, reminiscent of the CIDR:EPCR interaction. The data for Cys2 fit well ($\chi^2 = 6.58$), though the results should be repeated due to a baseline problem. However these results suggest that Cys 2 binds EPCR with a comparable affinity with the HB3var03 mutant excluded both residues outside of the binding epitope (D576AK642A). This immunogen should hence be taken forward for immunisation.

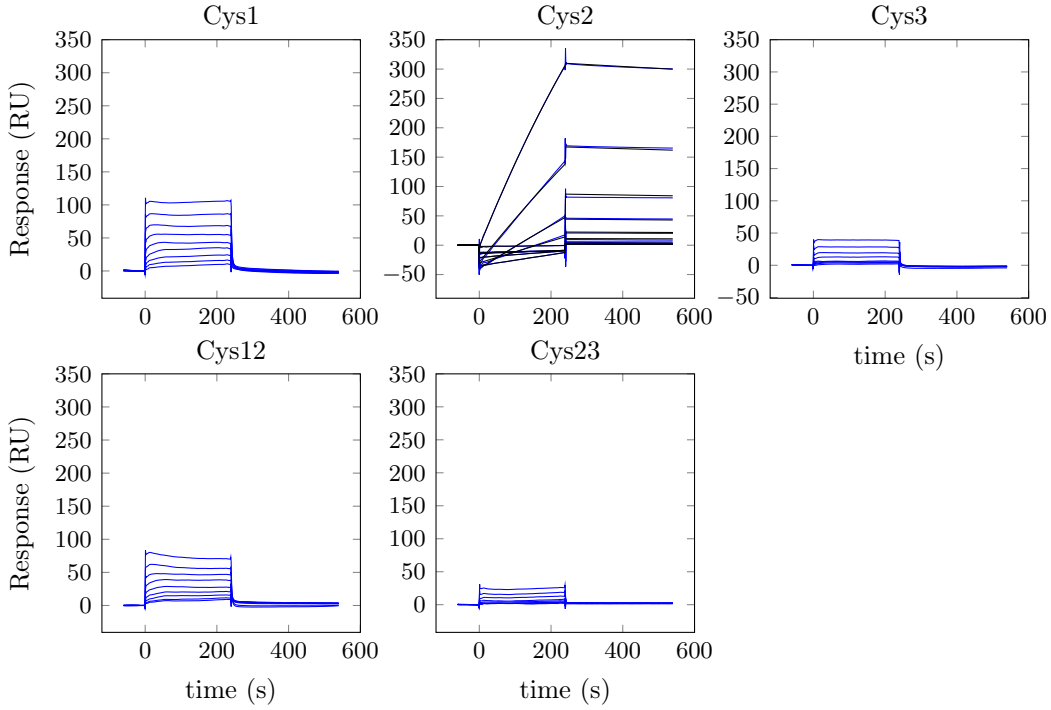


Figure 6.36: Surface plasmon resonance traces showing concentration series for each cysteine mutant bundle for binding to EPCR. Blank-subtracted data is shown in blue, with the kinetic fits shown in black. Data are summarised in Table 6.5.

Species	Range	K_D (nM)	k_{assoc} (M^{-1} s^{-1})	k_{dissoc} (s^{-1})
D576AK642A	250 nM - 0.9 nM	220	9.10×10^3	2.00×10^{-3}
469 Cys2	12 µM - 47 nM	988	1.12×10^2	1.11×10^{-4}

Table 6.5: SPR data summary table for D576A and K642A mutants, tested for EPCR binding. The HB3var03 D576AK642A CIDR mutant results are included for reference.

6.8 Discussion

Once I solved the CIDR:EPCR structure, the Lavstsen group used peptides encompassing the CIDR binding site residues to purify inhibiting antibodies [319]. In these experiments they purified antibodies using a peptide, consisting of the loop between the second and third conserved helix including the binding helices, from the HB3var03 CIDR or the IT4var20 CIDR to purify antibodies from patient sera. These purified antibodies prevented sequence-similar CIDR domains from binding EPCR, but could not inhibit the interaction between more diverse CIDR variants and EPCR. Antibodies purified using the HB3var03 peptide could not inhibit the interaction between IT4var20CIDR and EPCR, and vice versa.

In line with these data, antibodies raised against the whole HB3var03 CIDR did inhibit the homotypic interaction, whilst not inhibiting the interaction between IT4var20 CIDR and EPCR (Figure 6.5). Due to this, I explored strategies to elicit cross-inhibitory antibodies, searching for immunogens that could elicit a greater proportion of inhibitory antibodies compared to immunising with the CIDR alone.

We had two strategies to attempt to increase the inhibitory response - glycan masking, to focus the immune response towards the binding site; and taking the binding epitope out of the context of the CIDR using stapled peptides or a scaffold to recreate the conformation of the binding site.

Glycan masking has been attempted previously in viruses and in *Plasmodium vivax* [244]. This strategy attempts to focus the immune response, either away from immunodominant epitopes, or towards ligand binding sites, to inhibit an interaction. To emulate this, I

designed a glycosylated CIDR construct, incorporating seven glycosylation sites. These were located in loops between secondary structure elements, or on β-sheet surfaces, avoiding α-helices. These sites are where glycosylations are commonly found in solved structures of proteins in the PDB [320].

Though these glycosylated CIDR proteins expressed well in S2 cells, sera raised against this construct contained the same percentage of inhibitory antibodies than sera raised against HB3var03 CIDR, and reduced cross-reactivity against IT4var20 CIDR compared to using the unmodified CIDR. This result is consistent with past attempts at glycan masking; though some studies do have success in dampening a target epitope, attempts to use this technique to select an epitope to present to the immune system have only been partially successful [243, 244]. In one study antibodies raised against two hyperglycosylated hemagglutinin proteins from a H5N1 virus cross-inhibited hemagglutannin from two other viruses, but neither immunogen raised cross-inhibitory antibodies that inhibited hemagglutanin from both viruses [321]. Perhaps in this present study, this strategy has eliminated one or two cross-reactive, non-inhibitory sites on the surface without effectively focussing the response to the binding site.

The second method of focussing the immune response to the binding site took the binding motif out of the CIDR and expressing it, either by itself or attached to a scaffold.

Stapled helical peptides are beginning to attract attention as vaccine targets in HIV-1 vaccine design [246], though this strategy is yet to have been used to raise antibodies. I based our stapled helical peptide on the binding site in the HB3var03 structure, selecting a straight peptide to synthesise, including the two interacting phenylalanines at the N-terminus of the peptide. This stapled helical peptide was not a successful immunogen not eliciting reactive antibodies. It is not clear whether this is due to the size or conformation of the immunogen. Peptides are poor immunogens [247], so it is possible that, despite the aid of the adjuvants used, antibodies were not raised against the peptide at all. The CD traces here, coupled with the lack of binding activity to EPCR, suggest that the central phenylalanine is not presented in the correct orientation for EPCR binding. This is not surprising, since the two phenylalanines are at the end of the helical peptide. A limiting

factor of this approach is that stapled helical peptides are currently limited to straight α-helices. In addition, this peptide only included five of the nine binding residues, with two of the others present on the other helical segment forming the kinked helix. Stabilisation of a kinked helix would likely require the use of more complex staples, with three of four attachment points to stabilise the angle of the kink as well as the N and C helical portions. This sort of staple would require considerable design and optimisation, but is conceptually possible for future staple peptide immunogens.

The kinked helix-turn-helix motif between the second and third conserved helices contains seven of the nine residues responsible for EPCR interaction, so may be suitable for epitope grafting. To check this, we mutated the two outlying interacting residues, D576 and K642, individually and in combination. SPR data of show that these two mutations act synergistically to increase the K_D from 0.37 nM in the wild type, to 220 nM in the double mutant. The double mutant, D576AK642A, has a k_{assoc} that was 50-fold lower and a k_{dissoc} that was 10-fold higher than the wild-type. The synergy of these two mutations could be linked to their location on the binding site (Figure 6.17). Both of these interactions are hydrophilic in nature, possibly involved with the initial reorientation of the CIDR and the EPCR before the hydrophobic phenylalanine locks into the groove of EPCR. Though these two residues are significant in the interaction, alanine mutation does not abolish the interaction. Indeed the observed kinetic and thermodynamic parameters of the double mutant are within the observed range of the natural CIDR variants. This justifies their exclusion in the stapled peptide and helical bundle immunogens.

I based my epitope grafting approach on the Fold-From-Loops [252], which was not available as an executable at the time. Our attempt, though similar, was more complex due to the presence of the kinked helix in our target epitope. Despite this, the termini of the helices matched well with the scaffold, so I proceeded with the workflow used in the previous study [252]. The first round of immunogen design, using the Rosetta suite, predicted stable helical bundles in the desired conformation. Of the eight sequences brought forward to expression, seven expressed protein with a high α-helix content, with two constructs stable to high temperatures. Given the D576AK642A CIDR mutant bound EPCR with a K_D of 220 nM, I tested these bundles at 50 µM, 5 µM and 500 nM for EPCR binding

activity by SPR. None of these bundles bound to EPCR. The reason for the lack of binding is perhaps hinted to by the AbinitioRelax folding prediction run. Though the conformational sampling is low (20,000 runs is suggested as a minimum for finding the lowest energy conformation), the fact that this program did not predict the bundle 469 to fold into the desired conformation is consistent with lack of EPCR binding observed. An encouraging observation from this experiment was that the output model with the lowest Rosetta score had a score of -314, compared to the desired conformation score of -340. This is consistent with incomplete conformational sampling [322], as the desired conformation is predicted to be more stable than the output of the AbinitioRelax experiment.

The second round of design produced five stable helical bundles, three containing one putative disulfide bond (Cys1, 2 and 3), and two containing two putative disulfide bonds (Cys12 and Cys23). It is interesting that bundles including Cys2 were less stable at high temperatures than 469, while Cys 1 and Cys 3 mutations alone are more stable than 469 (Figure 6.34). Since Cys1 and Cys3 are both stabilising disulfide bonds, this suggests that Cys2 destabilises the bundle slightly, most likely in the epitope-distal end of the helical bundle. However since the T_M of Cys2 is around 60°C, Cys2 can still be considered a stable helical bundle, suitable for use as an immunogen. SPR traces revealed Cys2 as the one helical bundle that shows binding activity to EPCR comparable to the HB3var03 D576AK642A double mutant. This successful bundle contained the disulfide bond furthest from the binding epitope. This link, between the second and third helix, perhaps was enough to allow the binding epitope to fold into the correct conformation. The other two disulfide bonds by comparison were much closer to the binding epitope, perhaps disrupting the pitch of the linked helices upon bond formation. This could prevent the epitope from folding correctly to bind EPCR. This successful bundle, 469Cys2, could be used as part of a prime/boost strategy. Future work will structurally characterise this immunogen, to assess the conformation of the binding site, and immunise rats with this immunogen to assess its ability to raise inhibitory antibodies.

Chapter 7

Conclusions and Perspective

7.1 Summary of the findings from this DPhil

The aim of this thesis was to characterise the *Pf*EMP1:EPCR interaction, providing the first structural insight into how a *Pf*EMP1 protein binds its ligand. I also wanted to gain insight into how this interaction was conserved between sequence-diverse *Pf*EMP1 proteins, and attempt to target this conservation to design an immunogen to raise inhibitory antibodies.

7.1.1 The crystal structures of the CIDR:EPCR complex

After optimising expression for of EPCR-binding CIDR domains from *Pf*EMP1 proteins and for the host protein EPCR, I solved the structure for two CIDR:EPCR complexes. These crystal structures represent the first molecular view of a *Pf*EMP1:ligand interaction, revealing the nature of the interaction between the CIDRα1 domain and EPCR, with a hydrophobic phenylalanine from the CIDR protruding into the hydrophobic groove of EPCR. Comparison with the previous EPCR:protein C structure show that the CIDR domains bind to the same surface as the natural EPCR ligand. This explains the previous observation that CIDR binding to EPCR can prevent protein C from binding, which may have important immunological consequences for the host.

7.1.2 Conservation of the interaction between *Pf*EMP1 and EPCR

The combination of these structures with CIDRα1 sequence data sheds light on the conservation of this interaction. Amongst EPCR-binding CIDRs, inwards-facing residues are well conserved, forming the core fold of the CIDR domain. The interaction site does not show sequence conservation, instead conserved by the hydrophobic or hydrophilic character of the interacting residues. These data have identified the structurally and chemically conserved features of the EPCR-binding site that would need to be mimicked by a vaccine immunogen.

7.1.3 Simultaneous dual-binding by *Pf*EMP1 proteins

This thesis also reports the first direct observation of simultaneous ligand binding by a *Pf*EMP1 construct. EPCR and ICAM-1 are both shown to bind to a three domain head construct, with neither interaction disrupting the other. Such simultaneous interactions may need to be considered in the design of a vaccine.

7.1.4 Immunogen design

Three immunogen design strategies were attempted, targeting the surface of the CIDRα domains that bind to EPCR, as identified in the crystal structures. Glycan masking of the CIDR domain decreased the cross-reactivity of the antibody response, while not significantly changing the proportion of inhibitory antibodies. The stapled helical peptide did not yield any inhibitory antibodies. In the third strategy attempted, the EPCR-binding site on the CIDR was grafted to a small, soluble scaffold using Rosetta software suite *in silico*. After the addition of a disulfide bond, one immunogen was expressed that bound EPCR with an affinity in line with natural variants. This immunogen should now be tested to assess its ability to raise cross-inhibitory antibodies.

7.2 On the conservation of *Pf*EMP1 proteins and their interactions

Characterisation of the vast diversity in *Pf*EMP1 sequence is key to understanding how *Pf*EMP1 proteins interact with their ligands and enables the identification of conserved elements that could be targeted in a therapy. The results of this present study contribute to this characterisation, both suggesting conserved structural features of the CIDR fold, and by showing how a combination of structural, biophysical and sequence data can allow the identification of conserved elements that facilitate an interaction between a diverse parasite protein family and its ligands.

7.2.1 Conservation of the CIDR fold

Whilst the determination of ten DBL domain structures has led to the identification of a conserved DBL fold, the precise nature of CIDR fold was previously unclear, due to the conflicting structural data available. The crystal structures described in this thesis represent the third and fourth CIDR structures in the protein databank, after MC179 [137] and the var0 CIDR [189].

The four structures share three conserved helices (Figure 7.1), with a variable portion between the second and third conserved helices. This matches well with homology blocks predicted to make up the CIDR domain [170], with HB8, HB1 and HB10 matching to conserved helices 1, 2 and 3 respectively. The variable portion between the second and third conserved helices maps to HB121 in the CIDRα1 subclass, containing seven of the nine EPCR-binding residues, while mapping to other homology blocks in other CIDR classes. Given that EPCR binding site is mostly contained within this variable portion, it is plausible that the interaction site for CD36 also rests in the equivalent part of the CIDR.

In addition to this, the two EPCR-binding CIDR domains include an N-terminal β-sheet segment, that broadly resembles the β-sheet segment found in the var0 CIDR. This suggests that perhaps this β-sheet is common to all CIDRs. Indeed sequence alignments of the

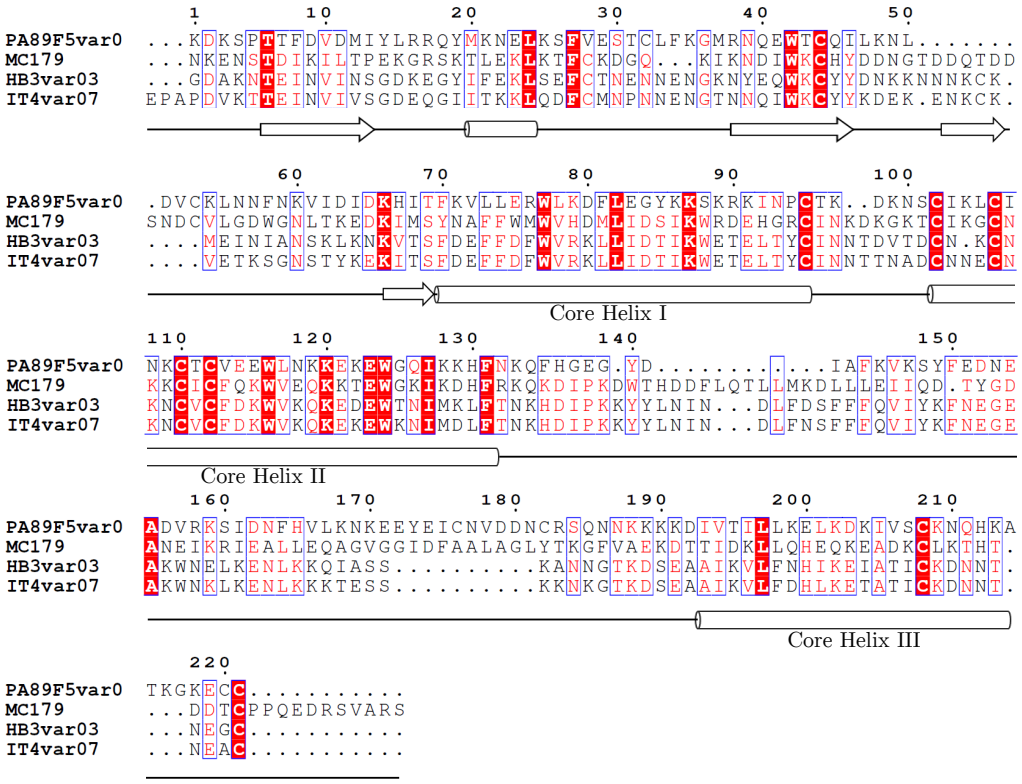


Figure 7.1: Sequence alignments of the four crystallised CIDRs, including the upsteam sequence to the MC179 construct. This alignments shows conservation, suggesting the MC179 upstream sequence adopts a β-sheet.

MC179 sequence equivalent to this β-sheet segment show similarities in key structural residues to the other three crystallised CIDRs (Figure 7.1). Moreover, the placement of residues on the conserved helical bundle suggest that the β-sheet segment in MC179 may more closely resemble that of the CIDRα1 class, with the MC179 residues that contact the β-sheet more similar to the HB3var03 CIDR than to the var0 CIDR (Figure 7.2).

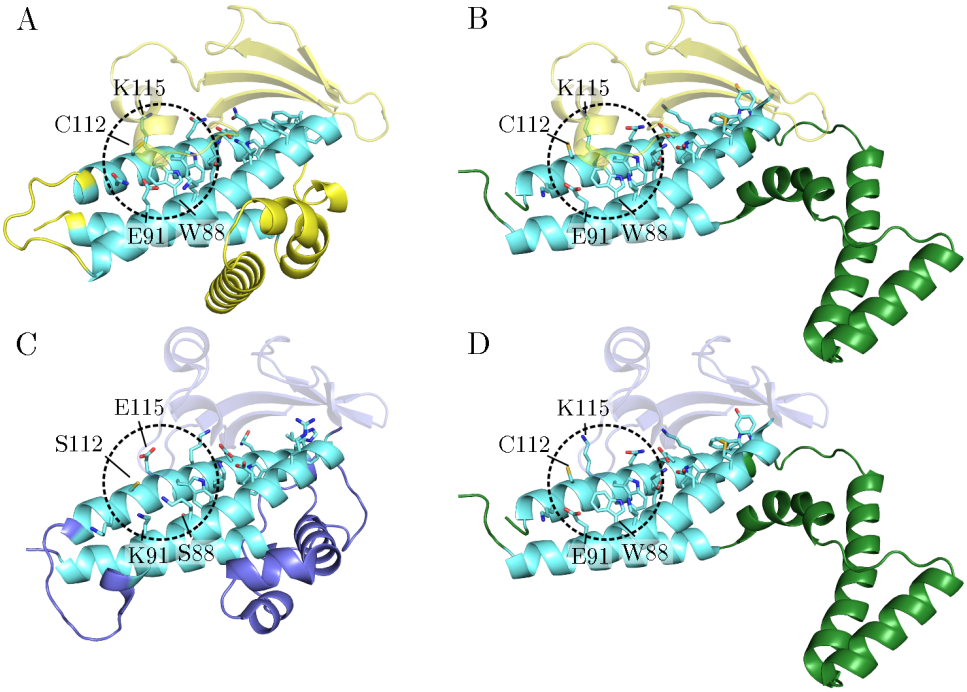


Figure 7.2: A comparison of the residues which contact the β-sheet subdomain in the crystallised CIDR domains from HB3var03 (yellow, A), MC179 (green, B,D) and var0(blue, C), aligned by the conserved helices (light blue). The β-sheet subdomains are overlayed from HB3var03 in A and B; from var0 in C and D. The contacting residues, numbered as in Figure 7.1, are more similar between HB3var03 and the MC179 CIDR.

Whilst the overall fold in the two EPCR-binding CIDRs strongly resembles the var0 CIDR, only the three core helices of MC179 align to these three structures, with the rest of the MC179 domain adopting an open conformation. The conformational differences between the MC179 construct and the other three solved CIDR structures are surprising, given higher sequence identity between the EPCR-binding CIDRα1 domains and MC179, a CIDRα2 domain (34% identity), than between these domains and var0, which is a CIDRγ domain (23% identity). Indeed, there are differences the structural context of a patch of sequence conservation [137] on the face formed by the first and third conserved helices in the CIDRα subclass. Whilst this patch sits between the two helical sections in the MC179

structure, in the two EPCR-binding CIDRs these surfaces are buried, with well conserved residues L583 and L699 responsible for the packing of the binding are onto the conserved bundle (Figure 7.3). Overall, this conservation means that the difference between the structure of MC179 and the two structures in this thesis is hard to rationalise. Future structural data, especially of the CIDR domain in complex with CD36 will reveal whether this open conformation is important to the interaction between the MC179 CIDR domain and CD36.

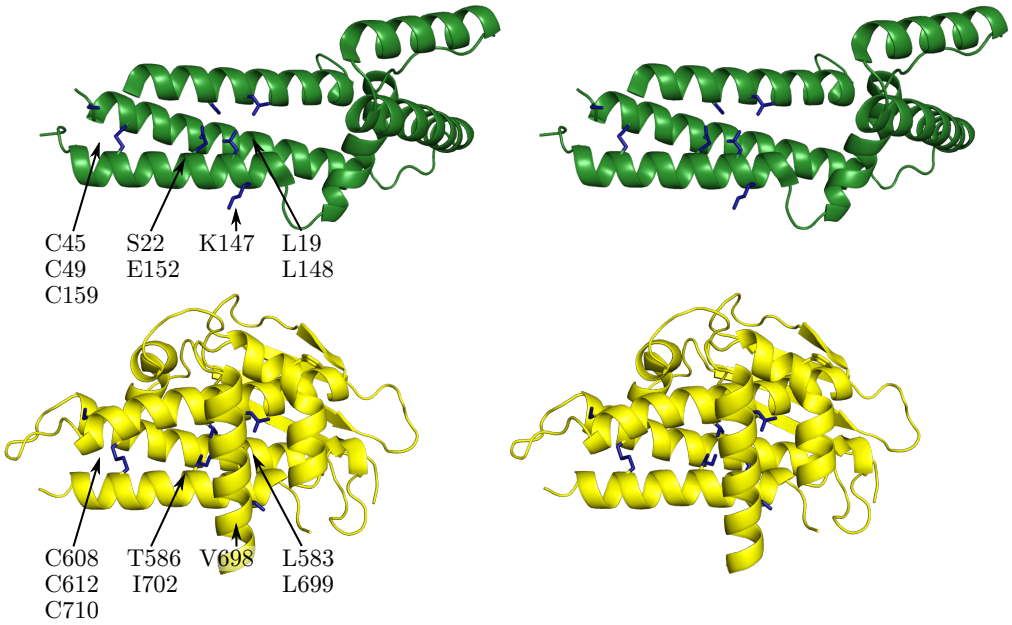


Figure 7.3: The conserved patch proposed in [137] is shown in stereo view for the MC179 (green) and HB3var03 (yellow) structures. Residues previously identified as conserved are coloured deep blue and represented as sticks.

7.2.2 Conservation of *Pf*EMP1:ligand interactions

Though the crystal structures alone represent a step forward in the molecular characterisation of the *Pf*EMP1:EPCR interaction, without the formidable sequence database of EPCR-binding CIDRs, provided by the Lavstsen group, it may not have been possible to generalise the information gained from the structures to the sequence-diverse CIDRα1 subclass. The plotting of the sequence conservation onto the structures allowed the quantification of diversity in the binding site and the identification of the molecular determinants that allow the CIDR to interact with EPCR. Given the large increase in the amount of sequence data available for *Pf*EMP1 proteins, it may be possible to apply this sequence-structure approach to other *Pf*EMP1:ligand interactions, providing sufficient structural insight is first gained through means such as X-ray crystallography.

The biophysical studies hint at how *Pf*EMP1s, with such variable affinity for a ligand, could maintain their cytoadhering nature *in vivo*. The slow dissociation rate seems to be conserved in the natural variants, whilst the association rate is variable. Naturally-occuring, low affinity ligand binding domains can be explained by avidity effects taking place in the context of an infected erythrocyte, where multiple *Pf*EMP1 are presented on a knob surface, compensating for the low affinity of the individual protein. Studies using parasites expressing high and low affinity *Pf*EMP1 variants in flow assays could assess whether the low affinity variant reduces the shear force required to remove an adhered infected erythrocyte. It remains to be seen whether the conservation of a slow dissociation rate is crucial in other *Pf*EMP1:ligand interactions.

7.3 The link between EPCR, ICAM-1 and severe malaria

The findings reported in this thesis add to the suggestion that binding to EPCR could have important immunological consequences in the host. The crystal structures described here validate the previous observation that CIDRs and protein C compete for similar binding surfaces on EPCR. The model linking EPCR and severe malaria, suggesting by Turner *et al* [140], remains unproven, but includes ICAM-1 expression as a downstream effect.

In this thesis I observed simultaneous dual-binding of EPCR and ICAM-1 to multiple *Pf*EMP1 variants. Linking EPCR and ICAM-1 with severe malaria is very appealing, as it could partially explain the aberrant inflammation often seen in patients with severe malaria. Further study of this dual binding phenotype in *Pf*EMP1 should be carried out to ascertain how common these two ligands can be bound by the same variant, and whether this dual-binding phenotype is predictive of severe malaria.

7.4 Towards a vaccine against malaria

In this thesis I identified the conserved molecular determinants of the CIDR:EPCR interaction. Despite the interaction being conserved, the use of one CIDR domain to elicit a broadly neutralising antibody response against the CIDR:EPCR interaction seems an unlikely prospect. Purification of antibodies from human sera using binding site peptides from HB3var03 or IT4var20 shows that these antibodies exhibit only limited cross-inhibition of the interaction [319]. Studies in this thesis confirm this, as sera from rats, immunised with the HB3var03 CIDRα1 domain, do not inhibit the IT4var20:EPCR interaction despite being cross-reactive.

Different immunogen design strategies were used in an effort to focus the immune response towards the EPCR-binding site of the CIDRαdomains. Although the glycan masking was not successful in increasing the proportion of inhibitory antibodies in the serum, lessons can be learnt from this process. The sera raised against the glycosylated HB3var03 CIDR showed no cross-reactivity to the IT4var20 CIDR, in contrast to the sera raised against the unglycosylated control. This suggests there are patches of conservation on the CIDR that are away from the interaction site. This is consistent with the findings of our sequence-structure studies, which suggested a patch of conservation on the opposite face of the CIDR to the EPCR-binding site (Figure 4.10). This should be taken into account, if a non-redundant set of CIDRs, representing the binding site diversity, is used in a vaccine strategy.

Another option to focus the immune response is by removing the binding epitope from

the CIDR and presenting it to the immune system on a scaffold. This would reduce the potential for non-inhibitory antibodies, especially if used in a prime/boost strategy in which the immune response is first raised against a whole CIDR, then focussed using this immunogen. The *in silico* design protocol used is still in its infancy, but has already been used successfully to induce RSV-neutralising antibodies in macaques [252]. This design protocol, with additional modifications, produced one synthetic immunogen that recapitulated the CIDR:EPCR interaction. This immunogen will be carried forwards to assess its ability to produce inhibitory antibodies. This shows the "Fold from loops" protocol to be a very powerful tool to produce immunogens, though much manual guidance is currently required in order to properly recreate the epitope structure.

An immunogen that raises inhibitory antibodies targeting the CIDR:EPCR is unlikely to work alone, due to the number of different *Pf*EMP1 interactions, as well as the complexity of the malarial parasite life cycle, a multicomponent vaccine could hold more promise to vaccinate against malaria. This strategy has been attempted with petussis and serogroup B meningococcus vaccines, with multicomponent vaccines more effective than single component vaccines [323, 324]. If a suitable immunogen and strategy is found, a vaccine that prevents the CIDR:EPCR interaction could form a key component of a vaccine against severe malaria.

Bibliography

[1] "WHO | World Malaria Report 2014." [Online]. Available: http://www.who.int/malaria/publications/world_malaria_report_2014/report/en/

[2] "Global, regional, and national agesex specific all-cause and cause-specific mortality for 240 causes of death, 19902013: a systematic analysis for the Global Burden of Disease Study 2013," *The Lancet*, vol. 385, no. 9963, pp. 117–171, Dec. 2014. [Online]. Available: http://www.thelancet.com/article/S0140673614616822/fulltext

[3] J. Sachs and P. Malaney, "The economic and social burden of malaria." *Nature*, vol. 415, no. 6872, pp. 680–5, Feb. 2002. [Online]. Available: http://dx.doi.org/10.1038/415680a

[4] R. Hecht, P. Wilson, and A. Palriwala, "Improving health r&d financing for developing countries: a menu of innovative policy options." *Health affairs (Project Hope)*, vol. 28, no. 4, pp. 974–85, Jan. 2009. [Online]. Available: http://content.healthaffairs.org/content/28/4/974

[5] C. J. L. Murray, L. C. Rosenfeld, S. S. Lim, K. G. Andrews, K. J. Foreman, D. Haring, N. Fullman, M. Naghavi, R. Lozano, and A. D. Lopez, "Global malaria mortality between 1980 and 2010: a systematic analysis." *Lancet (London, England)*, vol. 379, no. 9814, pp. 413–31, Feb. 2012. [Online]. Available: http://www.lancet.com/article/S0140673612600348/fulltext

[6] N. Dhingra, P. Jha, V. P. Sharma, A. A. Cohen, R. M. Jotkar, P. S. Rodriguez, D. G. Bassani, W. Suraweera, R. Laxminarayan, and R. Peto, "Adult and child malaria mortality in India: a nationally representative mortality survey." *Lancet (London, England)*, vol. 376, no. 9754, pp. 1768–74, Nov. 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140673610608318

[7] P. C. Bull, B. S. Lowe, M. Kortok, C. S. Molyneux, C. I. Newbold, and K. Marsh, "Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria." *Nature medicine*, vol. 4, no. 3, pp. 358–60, Mar. 1998. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3836255&tool=pmcentrez&rendertype=abstract

[8] S. J. Rogerson, L. Hviid, P. E. Duffy, R. F. G. Leke, and D. W. Taylor, "Malaria in pregnancy: pathogenesis and immunity." *The Lancet. Infectious diseases*, vol. 7, no. 2, pp. 105–17, Feb. 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1473309907700221

[9] B. Lüthi and P. Schlagenhauf, "Risk factors associated with malaria deaths in travellers: a literature review." *Travel medicine and infectious disease*, vol. 13, no. 1, pp. 48–60, Jan. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1477893914001161

[10] G. M. Sternberg, "THE MALARIAL GERM OF LAVERAN." *Science (New York, N.Y.)*, vol. 7, no. 165, pp. 297–9, Apr. 1886. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/17829595

[11] W. Osler, "An Address on the Haematozoa of Malaria." *British medical journal*, vol. 1, no. 1367, pp. 556–62, Mar. 1887. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2534096&tool=pmcentrez&rendertype=abstract

[12] J. C. Geiger and F. L. Kelly, "PLASMODIUM MALARIAE (QUARTAN)-A TYPE NEW TO CALIFORNIA: Report of two cases." *California state journal of medicine*, vol. 14, no. 5, p. 198, May 1916. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1642226&tool=pmcentrez&rendertype=abstract

[13] C. C. Bass and F. M. Johns, "THE CULTIVATION OF MALARIAL PLASMODIA (PLASMODIUM VIVAX AND PLASMODIUM FALCIPARUM) IN VITRO." *The Journal of experimental medicine*, vol. 16, no. 4, pp. 567–79, Oct. 1912. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2124976&tool=pmcentrez&rendertype=abstract

[14] D. Wilson and M. E. Wilson, "Infections with Plasmodium ovale Stephens, in Tanganyika territory," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 28, no. 5, pp. 469–474, Mar. 1935. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0035920335900025

[15] S. Jongwutiwes, C. Putaporntip, T. Iwasaki, T. Sata, and H. Kanbara, "Naturally Acquired Plasmodium knowlesi Malaria in Human, Thailand," *Emerging Infectious Diseases*, vol. 10, no. 12, pp. 2211–2213, Dec. 2004. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3323387&tool=pmcentrez&rendertype=abstract

[16] I. Vythilingam, C. H. Tan, M. Asmad, S. T. Chan, K. S. Lee, and B. Singh, "Natural transmission of Plasmodium knowlesi to humans by Anopheles latens in Sarawak, Malaysia." *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 100, no. 11, pp. 1087–8, Nov. 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16725166

[17] J. Cox-Singh and B. Singh, "Knowlesi malaria: newly emergent and of public health importance?" *Trends in parasitology*, vol. 24, no. 9, pp. 406–10, Sep. 2008. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2843823&tool=pmcentrez&rendertype=abstract

[18] R. Amino, S. Thiberge, B. Martin, S. Celli, S. Shorte, F. Frischknecht, and R. Ménard, "Quantitative imaging of Plasmodium transmission from mosquito to mammal." *Nature medicine*, vol. 12, no. 2, pp. 220–4, Feb. 2006. [Online]. Available: http://dx.doi.org/10.1038/nm1350

[19] D. L. Medica and P. Sinnis, "Quantitative dynamics of Plasmodium yoelii sporozoite transmission by infected anopheline mosquitoes." *Infection and*

immunity, vol. 73, no. 7, pp. 4363–9, Jul. 2005. [Online]. Available: http://iai.asm.org/content/73/7/4363.abstract

[20] A. U. Krettli, L. A. B. Dantas;, and J. A. Vaughan, "Which Routes Do Plasmodium Sporozoites Use for Successful Infections of Vertebrates?" *Infection and Immunity*, vol. 68, no. 5, pp. 3064–3065, May 2000. [Online]. Available: http://iai.asm.org/content/68/5/3064

[21] S. C. Shin, J. P. Vanderberg, and J. A. Terzakis, "Direct infection of hepatocytes by sporozoites of Plasmodium berghei." *The Journal of protozoology*, vol. 29, no. 3, pp. 448–54, Aug. 1982. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/6752394

[22] P. Sinnis, "Remnant lipoproteins inhibit malaria sporozoite invasion of hepatocytes," *Journal of Experimental Medicine*, vol. 184, no. 3, pp. 945–954, Sep. 1996. [Online]. Available: http://jem.rupress.org/content/184/3/945.abstract

[23] J. Tavares, P. Formaglio, S. Thiberge, E. Mordelet, N. Van Rooijen, A. Medvinsky, R. Ménard, and R. Amino, "Role of host cell traversal by the malaria sporozoite during liver infection." *The Journal of experimental medicine*, vol. 210, no. 5, pp. 905–15, May 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3646492&tool=pmcentrez&rendertype=abstract

[24] T. Kariu, T. Ishino, K. Yano, Y. Chinzei, and M. Yuda, "CelTOS, a novel malarial protein that mediates transmission to mosquito and vertebrate hosts." *Molecular microbiology*, vol. 59, no. 5, pp. 1369–79, Mar. 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16468982

[25] O. Silvie, E. Rubinstein, J.-F. Franetich, M. Prenant, E. Belnoue, L. Rénia, L. Hannoun, W. Eling, S. Levy, C. Boucheix, and D. Mazier, "Hepatocyte CD81 is required for Plasmodium falciparum and Plasmodium yoelii sporozoite infectivity." *Nature medicine*, vol. 9, no. 1, pp. 93–6, Jan. 2003. [Online]. Available: http://dx.doi.org/10.1038/nm808

[26] A. Sturm, R. Amino, C. van de Sand, T. Regen, S. Retzlaff, A. Rennenberg, A. Krueger, J.-M. Pollok, R. Menard, and V. T. Heussler, "Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids." *Science (New York, N.Y.)*, vol. 313, no. 5791, pp. 1287–90, Sep. 2006. [Online]. Available: http://www.sciencemag.org/content/313/5791/1287.full

[27] A. M. Vaughan, S. A. Mikolajczak, E. M. Wilson, M. Grompe, A. Kaushansky, N. Camargo, J. Bial, A. Ploss, and S. H. I. Kappe, "Complete Plasmodium falciparum liver-stage development in liver-chimeric mice." *The Journal of clinical investigation*, vol. 122, no. 10, pp. 3618–28, Oct. 2012. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3461911&tool=pmcentrez&rendertype=abstract

[28] W. A. Krotoski, D. M. Krotoski, P. C. Garnham, R. S. Bray, R. Killick-Kendrick, C. C. Draper, G. A. Targett, and M. W. Guy, "Relapses in primate malaria: discovery of two populations of exoerythrocytic stages. Preliminary note." *BMJ*, vol. 280, no. 6208, pp. 153–154, Jan. 1980. [Online]. Available: http://www.bmj.com/content/280/6208/153.2.abstract

[29] W. A. Krotoski, P. C. Garnham, R. S. Bray, D. M. Krotoski, R. Killick-Kendrick, C. C. Draper, G. A. Targett, and M. W. Guy, "Observations on early and late post-sporozoite tissue stages in primate malaria. I. Discovery of a new latent form of Plasmodium cynomolgi (the hypnozoite), and failure to detect hepatic forms within the first 24 hours after infection." *The American journal of tropical medicine and hygiene*, vol. 31, no. 1, pp. 24–35, Jan. 1982. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/7058977

[30] K. Baer, C. Klotz, S. H. I. Kappe, T. Schnieder, and U. Frevert, "Release of hepatic Plasmodium yoelii merozoites into the pulmonary microvasculature." *PLoS pathogens*, vol. 3, no. 11, p. e171, Nov. 2007. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2065874&tool=pmcentrez&rendertype=abstract

[31] V. K. Goel, X. Li, H. Chen, S.-C. Liu, A. H. Chishti, and S. S. Oh, "Band 3 is a host receptor binding merozoite surface protein 1 during the Plasmodium falciparum invasion of erythrocytes." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 9, pp. 5164–9, Apr. 2003. [Online]. Available: http://www.pnas.org/content/100/9/5164.short

[32] A. F. Cowman and B. S. Crabb, "Invasion of red blood cells by malaria parasites." *Cell*, vol. 124, no. 4, pp. 755–66, Feb. 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0092867406001814

[33] H. Sakamoto, S. Takeo, A. G. Maier, J. Sattabongkot, A. F. Cowman, and T. Tsuboi, "Antibodies against a Plasmodium falciparum antigen PfMSPDBL1 inhibit merozoite invasion into human erythrocytes." *Vaccine*, vol. 30, no. 11, pp. 1972–80, Mar. 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0264410X12000138

[34] S. Lopaticki, A. G. Maier, J. Thompson, D. W. Wilson, W.-H. Tham, T. Triglia, A. Gout, T. P. Speed, J. G. Beeson, J. Healer, and A. F. Cowman, "Reticulocyte and erythrocyte binding-like proteins function cooperatively in invasion of human erythrocytes by malaria parasites." *Infection and immunity*, vol. 79, no. 3, pp. 1107–17, Mar. 2011. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3067488&tool=pmcentrez&rendertype=abstract

[35] D. T. Riglar, D. Richard, D. W. Wilson, M. J. Boyle, C. Dekiwadia, L. Turnbull, F. Angrisano, D. S. Marapana, K. L. Rogers, C. B. Whitchurch, J. G. Beeson, A. F. Cowman, S. A. Ralph, and J. Baum, "Super-resolution dissection of coordinated events during malaria parasite invasion of the human erythrocyte." *Cell host & microbe*, vol. 9, no. 1, pp. 9–20, Jan. 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1931312810004130

[36] G. E. Ward, L. H. Miller, and J. A. Dvorak, "The origin of parasitophorous vacuole membrane lipids in malaria-infected erythrocytes." *Journal of cell science*, vol. 106 (Pt 1, pp. 237–48, Sep. 1993. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/8270628

[37] M. Aikawa, L. H. Miller, J. Johnson, and J. Rabbege, "Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite." *The Journal of cell biology*, vol. 77, no. 1, pp. 72–82, Apr. 1978. [Online].

Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2110026&tool=pmcentrez&rendertype=abstract

[38] M. J. Boyle, C. Langer, J.-A. Chan, A. N. Hodder, R. L. Coppel, R. F. Anders, and J. G. Beeson, "Sequential processing of merozoite surface proteins during and after erythrocyte invasion by Plasmodium falciparum." *Infection and immunity*, vol. 82, no. 3, pp. 924–36, Mar. 2014. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3958018&tool=pmcentrez&rendertype=abstract

[39] S. A. Lauer, "A Membrane Network for Nutrient Import in Red Cells Infected with the Malaria Parasite," *Science*, vol. 276, no. 5315, pp. 1122–1125, May 1997. [Online]. Available: http://www.sciencemag.org/content/276/5315/1122.full

[40] M. Marti, R. T. Good, M. Rug, E. Knuepfer, and A. F. Cowman, "Targeting malaria virulence and remodeling proteins to the host erythrocyte." *Science (New York, N.Y.)*, vol. 306, no. 5703, pp. 1930–3, Dec. 2004. [Online]. Available: http://www.sciencemag.org/content/306/5703/1930.full

[41] D. I. Baruch, B. L. Pasloske, H. B. Singh, X. Bi, X. C. Ma, M. Feldman, T. F. Taraschi, and R. J. Howard, "Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes." *Cell*, vol. 82, no. 1, pp. 77–87, 1995. [Online]. Available: http://dx.doi.org/10.1016/0092-8674(95)90054-3

[42] L. Dudgeon and C. Clarke, "A CONTRIBUTION TO THE MICROSCOPICAL HISTOLOGY OF MALARIA," *The Lancet*, vol. 190, no. 4901, pp. 153–156, Aug. 1917. [Online]. Available: http://dx.doi.org/10.1016/S0140-6736(01)50818-1

[43] J. W. Barnwell, P. Ingravallo, M. R. Galinski, Y. Matsumoto, and M. Aikawa, "Plasmodium vivax: Malarial proteins associated with the membrane-bound caveola-vesicle complexes and cytoplasmic cleft structures of infected erythrocytes," *Experimental Parasitology*, vol. 70, no. 1, pp. 85–99, Jan. 1990. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-0025017735&partnerID=tZOtx3y1

[44] R. Suwanarusk, B. M. Cooke, A. M. Dondorp, K. Silamut, J. Sattabongkot, N. J. White, and R. Udomsangpetch, "The deformability of red blood cells parasitized by Plasmodium falciparum and P. vivax." *The Journal of infectious diseases*, vol. 189, no. 2, pp. 190–4, Jan. 2004. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-0742324945&partnerID=tZOtx3y1

[45] M. J. Blackman and V. B. Carruthers, "Recent insights into apicomplexan parasite egress provide new views to a kill." *Current opinion in microbiology*, vol. 16, no. 4, pp. 459–64, Aug. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1369527413000659

[46] M. Smalley, J. Brown, and N. M. Bassett, "The rate of production of Plasmodium falciparum gametocytes during natural infections," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 75, no. 2, pp. 318–319, Jan. 1981. [Online]. Available: http://trstmh.oxfordjournals.org/content/75/2/318.abstract

[47] M. C. Bruce, P. Alano, S. Duthie, and R. Carter, "Commitment of the malaria parasite Plasmodium falciparum to sexual and asexual development." *Parasitology*, vol. 100 Pt 2, pp. 191–200, Apr. 1990. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/2189114

[48] A. M. Talman, O. Domarle, F. E. McKenzie, F. Ariey, and V. Robert, "Gametocytogenesis: the puberty of Plasmodium falciparum." *Malaria journal*, vol. 3, no. 1, p. 24, Jul. 2004. [Online]. Available: http://www.malariajournal.com/content/3/1/24

[49] A. S. I. Aly, A. M. Vaughan, and S. H. I. Kappe, "Malaria parasite development in the mosquito and infection of the mammalian host." *Annual review of microbiology*, vol. 63, pp. 195–221, Jan. 2009. [Online]. Available: http://www.annualreviews.org/doi/full/10.1146/annurev.micro.091208.073403

[50] A. Kan, Y.-H. Tan, F. Angrisano, E. Hanssen, K. L. Rogers, L. Whitehead, V. P. Mollard, A. Cozijnsen, M. J. Delves, S. Crawford, R. E. Sinden, G. I. McFadden, C. Leckie, J. Bailey, and J. Baum, "Quantitative analysis of P lasmodium ookinete motion in three dimensions suggests a critical role for cell shape in the biomechanics of malaria parasite gliding motility," *Cellular Microbiology*, vol. 16, no. 5, pp. 734–750, May 2014. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4286792&tool=pmcentrez&rendertype=abstract

[51] A. S. I. Aly and K. Matuschewski, "A malarial cysteine protease is necessary for Plasmodium sporozoite egress from oocysts." *The Journal of experimental medicine*, vol. 202, no. 2, pp. 225–30, Jul. 2005. [Online]. Available: http://jem.rupress.org/content/202/2/225.full

[52] J. M. Crutcher and S. L. Hoffman, "Malaria," 1996. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK8584/

[53] A. Trampuz, M. Jereb, I. Muzlovic, and R. M. Prabhu, "Clinical review: Severe malaria." *Critical care (London, England)*, vol. 7, no. 4, pp. 315–23, Aug. 2003. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=270697&tool=pmcentrez&rendertype=abstract

[54] N. M. Anstey, B. Russell, T. W. Yeo, and R. N. Price, "The pathophysiology of vivax malaria." *Trends in parasitology*, vol. 25, no. 5, pp. 220–7, May 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1471492209000646

[55] G. Pasvol, "The treatment of complicated and severe malaria." *British medical bulletin*, vol. 75-76, no. 1, pp. 29–47, Jan. 2005. [Online]. Available: http://bmb.oxfordjournals.org/content/75-76/1/29.full

[56] E. Marchiafava and A. Bignami, *On Summer-Autumn Malarial Fevers*. Lond., 1894. [Online]. Available: http://www.amazon.co.uk/Summer-Autumn-Malarial-Fevers-Ettore-Marchiafava/dp/5518539320

[57] G. E. Grau, P. F. Piguet, P. Vassalli, and P. H. Lambert, "Tumor-necrosis factor and other cytokines in cerebral malaria: experimental and clinical data." *Immunological reviews*, vol. 112, pp. 49–70, Dec. 1989. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/2575074

[58] G. N. JAKEMAN, A. SAUL, W. L. HOGARTH, and W. E. COLLINS, "Anaemia of acute malaria infections in non-immune patients primarily results from destruction of uninfected erythrocytes," *Parasitology*, vol. 119, no. 02, pp. 127–133, Aug. 1999. [Online]. Available: http://journals.cambridge.org/abstract_S0031182099004564

[59] R. Price, J. Simpson, F. Nosten, C. Luxemburger, L. Hkirjaroen, F. ter Kuile, T. Chongsuphajaisiddhi, and N. White, "Factors contributing to anemia after uncomplicated falciparum malaria," *Am J Trop Med Hyg*, vol. 65, no. 5, pp. 614–622, Nov. 2001. [Online]. Available: http://www.ajtmh.org/content/65/5/614

[60] W. E. COLLINS, G. M. JEFFERY, and J. M. ROBERTS, "A RETROSPECTIVE EXAMINATION OF ANEMIA DURING INFECTION OF HUMANS WITH PLASMODIUM VIVAX," *Am J Trop Med Hyg*, vol. 68, no. 4, pp. 410–412, Apr. 2003. [Online]. Available: http://www.ajtmh.org/content/68/4/410.full

[61] E. Pongponratn, M. Riganti, B. Punpoowong, and M. Aikawa, "Microvascular sequestration of parasitized erythrocytes in human falciparum malaria: A pathological study," *American Journal of Tropical Medicine and Hygiene*, vol. 44, no. 2, pp. 168–175, 1991. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-0025764099&partnerID=tZOtx3y1

[62] S. Lewallen, V. A. White, R. O. Whitten, J. Gardiner, B. Hoar, J. Lindley, J. Lochhead, A. McCormick, K. Wade, M. Tembo, J. Mwenechanyana, M. E. Molyneux, and T. E. Taylor, "Clinical-Histopathological Correlation of the Abnormal Retinal Vessels in Cerebral Malaria," *Archives of Ophthalmology*, vol. 118, no. 7, pp. 924–928, Jul. 2000. [Online]. Available: http://archopht.jamanetwork.com/article.aspx?articleid=413361

[63] T. E. Taylor, W. J. Fu, R. A. Carr, R. O. Whitten, J. S. Mueller, N. G. Fosiko, S. Lewallen, N. G. Liomba, M. E. Molyneux, and J. G. Mueller, "Differentiating the pathologies of cerebral malaria by postmortem parasite counts." *Nature medicine*, vol. 10, no. 2, pp. 143–5, Feb. 2004. [Online]. Available: http://dx.doi.org/10.1038/nm986

[64] A. J. Cunnington, E. M. Riley, and M. Walther, "Stuck in a rut? Reconsidering the role of parasite sequestration in severe malaria syndromes." *Trends in parasitology*, vol. 29, no. 12, pp. 585–92, Dec. 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3880783&tool=pmcentrez&rendertype=abstract

[65] A. M. Dondorp, F. Nosten, P. Yi, D. Das, A. P. Phyo, J. Tarning, K. M. Lwin, F. Ariey, W. Hanpithakpong, S. J. Lee, P. Ringwald, K. Silamut, M. Imwong, K. Chotivanich, P. Lim, T. Herdman, S. S. An, S. Yeung, P. Singhasivanon, N. P. J. Day, N. Lindegardh, D. Socheat, and N. J. White, "Artemisinin resistance in Plasmodium falciparum malaria." *The New England journal of medicine*, vol. 361, no. 5, pp. 455–67, Jul. 2009. [Online]. Available: http://www.nejm.org/doi/full/10.1056/NEJMoa0808859#t=article

[66] A. J. Cunnington, M. T. Bretscher, S. I. Nogaro, E. M. Riley, and M. Walther, "Comparison of parasite sequestration in uncomplicated and severe childhood Plasmodium falciparum malaria." *The Journal of infection*, vol. 67, no. 3, pp. 220–30, Sep. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0163445313001023

[67] S. Bhatt, D. J. Weiss, E. Cameron, D. Bisanzio, B. Mappin, U. Dalrymple, K. E. Battle, C. L. Moyes, A. Henry, P. A. Eckhoff, E. A. Wenger, O. Briët, M. A. Penny, T. A. Smith, A. Bennett, J. Yukich, T. P. Eisele, J. T. Griffin, C. A. Fergus,

M. Lynch, F. Lindgren, J. M. Cohen, C. L. J. Murray, D. L. Smith, S. I. Hay, R. E. Cibulskis, and P. W. Gething, "The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015." *Nature*, vol. advance on, Sep. 2015. [Online]. Available: http://dx.doi.org/10.1038/nature15535

[68] G. Chapin and R. Wasserstrom, "Agricultural production and malaria resurgence in Central America and India," *Nature*, vol. 293, no. 5829, pp. 181–185, Sep. 1981. [Online]. Available: http://dx.doi.org/10.1038/293181a0

[69] K. Mendis, A. Rietveld, M. Warsame, A. Bosman, B. Greenwood, and W. H. Wernsdorfer, "From malaria control to eradication: The WHO perspective." *Tropical medicine & international health : TM & IH*, vol. 14, no. 7, pp. 802–9, Jul. 2009. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/19497083

[70] C. F. Curtis, J. E. Miller, M. H. Hodjati, J. H. Kolaczinski, and I. Kasumba, "Can anything be done to maintain the effectiveness of pyrethroid-impregnated bednets against malaria vectors?" *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 353, no. 1376, pp. 1769–75, Oct. 1998. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1692389&tool=pmcentrez&rendertype=abstract

[71] C. Lengeler, "Insecticide-treated bed nets and curtains for preventing malaria." 2004. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-3843070027&partnerID=tZOtx3y1

[72] M. L. H. Mabaso, B. Sharp, and C. Lengeler, "Historical review of malarial control in southern African with emphasis on the use of indoor residual house-spraying." *Tropical medicine & international health : TM & IH*, vol. 9, no. 8, pp. 846–56, Aug. 2004. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/15303988

[73] T. Narahashi, "Modification of nerve membrane sodium channels by the insecticide pyrethroids," *Comparative Biochemistry and Physiology Part C: Comparative Pharmacology*, vol. 72, no. 2, pp. 411–414, Jan. 1982. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0306449282901125

[74] Y. Du, Y. Nomura, G. Satar, Z. Hu, R. Nauen, S. Y. He, B. S. Zhorov, and K. Dong, "Molecular evidence for dual pyrethroid-receptor sites on a mosquito sodium channel," *Proceedings of the National Academy of Sciences*, vol. 110, no. 29, pp. 11 785–11 790, Jul. 2013. [Online]. Available: http://www.pnas.org/content/110/29/11785.full

[75] I. Petersen, R. Eastman, and M. Lanzer, "Drug-resistant malaria: molecular mechanisms and implications for public health." *FEBS letters*, vol. 585, no. 11, pp. 1551–62, Jun. 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0014579311003012

[76] F. P. Mockenhaupt, K. Reither, P. Zanger, F. Roepcke, I. Danquah, E. Saad, P. Ziniel, S. Y. Dzisi, M. Frempong, P. Agana-Nsiire, F. Amoo-Sakyi, R. Otchwemah, J. P. Cramer, S. D. Anemana, E. Dietz, and U. Bienzle, "Intermittent Preventive Treatment in Infants as a Means of Malaria Control: a Randomized, Double-Blind, Placebo-Controlled Trial in Northern Ghana," *Antimicrobial Agents and Chemotherapy*, vol. 51, no. 9, pp. 3273–3281, Jul. 2007.

[Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2043181&tool=pmcentrez&rendertype=abstract

[77] J. J. Aponte, D. Schellenberg, A. Egan, A. Breckenridge, I. Carneiro, J. Critchley, I. Danquah, A. Dodoo, R. Kobbe, B. Lell, J. May, Z. Premji, S. Sanz, E. Sevene, R. Soulaymani-Becheikh, P. Winstanley, S. Adjei, S. Anemana, D. Chandramohan, S. Issifou, F. Mockenhaupt, S. Owusu-Agyei, B. Greenwood, M. P. Grobusch, P. G. Kremsner, E. Macete, H. Mshinda, R. D. Newman, L. Slutsker, M. Tanner, P. Alonso, and C. Menendez, "Efficacy and safety of intermittent preventive treatment with sulfadoxine-pyrimethamine for malaria in African infants: a pooled analysis of six randomised, placebo-controlled trials." *Lancet (London, England)*, vol. 374, no. 9700, pp. 1533–42, Oct. 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140673609612587

[78] T. E. Wellems and C. V. Plowe, "Chloroquine-resistant malaria." *The Journal of infectious diseases*, vol. 184, no. 6, pp. 770–6, Sep. 2001. [Online]. Available: http://jid.oxfordjournals.org/content/184/6/770.long

[79] F. O. ter Kuile, A. M. van Eijk, and S. J. Filler, "Effect of sulfadoxine-pyrimethamine resistance on the efficacy of intermittent preventive therapy for malaria control during pregnancy: a systematic review." *JAMA*, vol. 297, no. 23, pp. 2603–16, Jun. 2007. [Online]. Available: http://jama.jamanetwork.com/article.aspx?articleid=207561

[80] D. L. Klayman, "Qinghaosu (Artemisinin): An antimalarial drug from China," pp. 1049–1055, 1985. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-0021948029&partnerID=tZOtx3y1

[81] R. T. Eastman and D. A. Fidock, "Artemisinin-based combination therapies: a vital tool in efforts to eliminate malaria," *Nature Reviews Microbiology*, vol. 7, no. 12, pp. 864–874, Dec. 2009. [Online]. Available: http://dx.doi.org/10.1038/nrmicro2239

[82] U. Eckstein-Ludwig, R. J. Webb, I. D. A. Van Goethem, J. M. East, A. G. Lee, M. Kimura, P. M. O'Neill, P. G. Bray, S. A. Ward, and S. Krishna, "Artemisinins target the SERCA of Plasmodium falciparum." *Nature*, vol. 424, no. 6951, pp. 957–61, Aug. 2003. [Online]. Available: http://dx.doi.org/10.1038/nature01813

[83] A. V. Pandey, B. L. Tekwani, R. L. Singh, and V. S. Chauhan, "Artemisinin, an Endoperoxide Antimalarial, Disrupts the Hemoglobin Catabolism and Heme Detoxification Systems in Malarial Parasite," *Journal of Biological Chemistry*, vol. 274, no. 27, pp. 19 383–19 388, Jul. 1999. [Online]. Available: http://www.jbc.org/content/274/27/19383.full

[84] F. Ariey, B. Witkowski, C. Amaratunga, J. Beghain, A.-C. Langlois, N. Khim, S. Kim, V. Duru, C. Bouchier, L. Ma, P. Lim, R. Leang, S. Duong, S. Sreng, S. Suon, C. M. Chuor, D. M. Bout, S. Ménard, W. O. Rogers, B. Genton, T. Fandeur, O. Miotto, P. Ringwald, J. Le Bras, A. Berry, J.-C. Barale, R. M. Fairhurst, F. Benoit-Vical, O. Mercereau-Puijalon, and D. Ménard, "A molecular marker of artemisinin-resistant Plasmodium falciparum malaria." *Nature*, vol. 505, no. 7481, pp. 50–5, Jan. 2014. [Online]. Available: http://dx.doi.org/10.1038/nature12876

[85] P. L. Alonso and M. Tanner, "Public health challenges and prospects for malaria control and elimination." *Nature medicine*, vol. 19, no. 2, pp. 150–5, Feb. 2013. [Online]. Available: http://dx.doi.org/10.1038/nm.3077

[86] W. Peters, "Drug resistance in malariaA perspective," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 63, no. 1, pp. 25–40, Jan. 1969. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0035920369900637

[87] A. Färnert, J. Lindberg, P. Gil, G. Swedberg, Y. Berqvist, M. M. Thapar, N. Lindegå rdh, S. Berezcky, and A. Björkman, "Evidence of Plasmodium falciparum malaria resistant to atovaquone and proguanil hydrochloride: case reports." *BMJ (Clinical research ed.)*, vol. 326, no. 7390, pp. 628–9, Mar. 2003. [Online]. Available: http://www.bmj.com/content/326/7390/628.abstract

[88] S. Kuhn, M. J. Gill, and K. C. Kain, "Emergence of atovaquone-proguanil resistance during treatment of Plasmodium falciparum malaria acquired by a non-immune north American traveller to west Africa." *The American journal of tropical medicine and hygiene*, vol. 72, no. 4, pp. 407–9, Apr. 2005. [Online]. Available: http://www.ajtmh.org/content/72/4/407.abstract

[89] A. K. Boggild, R. Lau, D. Reynaud, K. C. Kain, and M. Gerson, "Failure of atovaquone-proguanil malaria chemoprophylaxis in a traveler to Ghana." *Travel medicine and infectious disease*, vol. 13, no. 1, pp. 89–93, Jan. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1477893914002658

[90] A. P. Phyo, S. Nkhoma, K. Stepniewska, E. a. Ashley, S. Nair, R. McGready, C. ler Moo, S. Al-Saai, A. M. Dondorp, K. M. Lwin, P. Singhasivanon, N. P. Day, N. J. White, T. J. Anderson, and F. Nosten, "Emergence of artemisinin-resistant malaria on the western border of Thailand: a longitudinal study," *The Lancet*, vol. 6736, no. 12, pp. 1–7, Apr. 2012. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S014067361260484X

[91] O. Miotto, J. Almagro-Garcia, M. Manske, B. Macinnis, S. Campino, K. A. Rockett, C. Amaratunga, P. Lim, S. Suon, S. Sreng, J. M. Anderson, S. Duong, C. Nguon, C. M. Chuor, D. Saunders, Y. Se, C. Lon, M. M. Fukuda, L. Amenga-Etego, A. V. O. Hodgson, V. Asoala, M. Imwong, S. Takala-Harrison, F. Nosten, X.-Z. Su, P. Ringwald, F. Ariey, C. Dolecek, T. T. Hien, M. F. Boni, C. Q. Thai, A. Amambua-Ngwa, D. J. Conway, A. A. Djimdé, O. K. Doumbo, I. Zongo, J.-B. Ouedraogo, D. Alcock, E. Drury, S. Auburn, O. Koch, M. Sanders, C. Hubbart, G. Maslen, V. Ruano-Rubio, D. Jyothi, A. Miles, J. O'Brien, C. Gamble, S. O. Oyola, J. C. Rayner, C. I. Newbold, M. Berriman, C. C. A. Spencer, G. McVean, N. P. Day, N. J. White, D. Bethell, A. M. Dondorp, C. V. Plowe, R. M. Fairhurst, and D. P. Kwiatkowski, "Multiple populations of artemisinin-resistant Plasmodium falciparum in Cambodia." *Nature genetics*, vol. 45, no. 6, pp. 648–55, Jun. 2013. [Online]. Available: http://dx.doi.org/10.1038/ng.2624

[92] N. Elissa, J. Mouchet, F. Riviere, J. Y. Meunier, and K. Yao, "Resistance of Anopheles gambiae s.s. to pyrethroids in Côte d'Ivoire." *Annales de la Societe belge de medecine tropicale*, vol. 73, no. 4, pp. 291–4, Dec. 1993. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/8129474

[93] J. M. VULULE, R. F. BEACH, F. K. ATIELI, J. M. ROBERTS, D. L. MOUNT, and R. W. MWANGI, "Reduced susceptibility of Anopheles gambiae to permethrin associated with the use of permethrin-impregnated bednets and curtains in Kenya,"

Medical and Veterinary Entomology, vol. 8, no. 1, pp. 71–75, Jan. 1994. [Online]. Available: http://doi.wiley.com/10.1111/j.1365-2915.1994.tb00389.x

[94] K. Hargreaves, L. L. Koekemoer, B. D. Brooke, R. H. Hunt, J. Mthembu, and M. Coetzee, "Anopheles funestus resistant to pyrethroid insecticides in South Africa," *Medical and Veterinary Entomology*, vol. 14, no. 2, pp. 181–189, Jun. 2000. [Online]. Available: http://doi.wiley.com/10.1046/j.1365-2915.2000.00234.x

[95] J. M. Riveron, M. Chiumia, B. D. Menze, K. G. Barnes, H. Irving, S. S. Ibrahim, G. D. Weedall, T. Mzilahowa, and C. S. Wondji, "Rise of multiple insecticide resistance in Anopheles funestus in Malawi: a major concern for malaria vector control." *Malaria journal*, vol. 14, no. 1, p. 344, Jan. 2015. [Online]. Available: http://www.malariajournal.com/content/14/1/344

[96] S. A. Charman, S. Arbe-Barnes, I. C. Bathurst, R. Brun, M. Campbell, W. N. Charman, F. C. K. Chiu, J. Chollet, J. C. Craft, D. J. Creek, Y. Dong, H. Matile, M. Maurer, J. Morizzi, T. Nguyen, P. Papastogiannidis, C. Scheurer, D. M. Shackleford, K. Sriraghavan, L. Stingelin, Y. Tang, H. Urwyler, X. Wang, K. L. White, S. Wittlin, L. Zhou, and J. L. Vennerstrom, "Synthetic ozonide drug candidate OZ439 offers new hope for a single-dose cure of uncomplicated malaria." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 11, pp. 4400–5, Mar. 2011. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3060245&tool=pmcentrez&rendertype=abstract

[97] J. Marfurt, F. Chalfein, P. Prayoga, F. Wabiser, G. Wirjanata, B. Sebayang, K. A. Piera, S. Wittlin, R. K. Haynes, J. J. Möhrle, N. M. Anstey, E. Kenangalem, and R. N. Price, "Comparative ex vivo activity of novel endoperoxides in multidrug-resistant plasmodium falciparum and P. vivax." *Antimicrobial agents and chemotherapy*, vol. 56, no. 10, pp. 5258–63, Oct. 2012. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3457353&tool=pmcentrez&rendertype=abstract

[98] B. Baragaña, I. Hallyburton, M. C. S. Lee, N. R. Norcross, R. Grimaldi, T. D. Otto, W. R. Proto, A. M. Blagborough, S. Meister, G. Wirjanata, A. Ruecker, L. M. Upton, T. S. Abraham, M. J. Almeida, A. Pradhan, A. Porzelle, M. S. Martínez, J. M. Bolscher, A. Woodland, S. Norval, F. Zuccotto, J. Thomas, F. Simeons, L. Stojanovski, M. Osuna-Cabello, P. M. Brock, T. S. Churcher, K. A. Sala, S. E. Zakutansky, M. B. Jiménez-Díaz, L. M. Sanz, J. Riley, R. Basak, M. Campbell, V. M. Avery, R. W. Sauerwein, K. J. Dechering, R. Noviyanti, B. Campo, J. A. Frearson, I. n. Angulo-Barturen, S. Ferrer-Bazaga, F. J. Gamo, P. G. Wyatt, D. Leroy, P. Siegl, M. J. Delves, D. E. Kyle, S. Wittlin, J. Marfurt, R. N. Price, R. E. Sinden, E. A. Winzeler, S. A. Charman, L. Bebrevska, D. W. Gray, S. Campbell, A. H. Fairlamb, P. A. Willis, J. C. Rayner, D. A. Fidock, K. D. Read, and I. H. Gilbert, "A novel multiple-stage antimalarial agent that inhibits protein synthesis," *Nature*, vol. 522, no. 7556, pp. 315–320, Jun. 2015. [Online]. Available: http://dx.doi.org/10.1038/nature14451

[99] P. D. Crompton, S. K. Pierce, and L. H. Miller, "Advances and challenges in malaria vaccine development." *The Journal of clinical investigation*, vol. 120, no. 12, pp. 4168–78, Dec. 2010. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2994342&tool=pmcentrez&rendertype=abstract

[100] E. S. Bergmann-Leitner, Q. Li, D. Caridha, M. T. ONeil, C. F. Ockenhouse, M. Hickman, and E. Angov, "Protective immune mechanisms against pre-erythrocytic forms of Plasmodium berghei depend on the target antigen," *Trials in Vaccinology*, vol. 3, pp. 6–10, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1879437813000156

[101] A. Dolo, D. Modiano, O. Doumbo, A. Bosman, T. Sidibé, M. M. Keita, S. Naitza, K. J. Robson, and A. Crisanti, "Thrombospondin related adhesive protein (TRAP), a potential malaria vaccine candidate." *Parassitologia*, vol. 41, no. 1-3, pp. 425–8, Sep. 1999. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/10697897

[102] M. J. Blackman, H. G. Heidrich, S. Donachie, J. S. McBride, and A. A. Holder, "A single fragment of a malaria merozoite surface protein remains on the parasite during red cell invasion and is the target of invasion-inhibiting antibodies." *The Journal of experimental medicine*, vol. 172, no. 1, pp. 379–82, Jul. 1990. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2188181&tool=pmcentrez&rendertype=abstract

[103] A. W. Stowers, M. C. Kennedy, B. P. Keegan, A. Saul, C. A. Long, and L. H. Miller, "Vaccination of monkeys with recombinant Plasmodium falciparum apical membrane antigen 1 confers protection against blood-stage malaria." *Infection and immunity*, vol. 70, no. 12, pp. 6961–7, Dec. 2002. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=133036&tool=pmcentrez&rendertype=abstract

[104] J. Baum, L. Chen, J. Healer, S. Lopaticki, M. Boyle, T. Triglia, F. Ehlgen, S. A. Ralph, J. G. Beeson, and A. F. Cowman, "Reticulocyte-binding protein homologue 5 - an essential adhesin involved in invasion of human erythrocytes by Plasmodium falciparum." *International journal for parasitology*, vol. 39, no. 3, pp. 371–80, Feb. 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0020751908003883

[105] S. Jones, L. Grignard, I. Nebie, J. Chilongola, D. Dodoo, R. Sauerwein, M. Theisen, W. Roeffen, S. K. Singh, R. K. Singh, S. Singh, E. Kyei-Baafour, K. Tetteh, C. Drakeley, and T. Bousema, "Naturally acquired antibody responses to recombinant Pfs230 and Pfs48/45 transmission blocking vaccine candidates." *The Journal of infection*, vol. 71, no. 1, pp. 117–27, Jul. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0163445315000912

[106] "Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial," *The Lancet*, vol. 386, no. 9988, pp. 31–45, Apr. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140673615607218

[107] O. C. Finney, G. J. Keitany, H. Smithers, A. Kaushansky, S. Kappe, and R. Wang, "Immunization with genetically attenuated P. falciparum parasites induces long-lived antibodies that efficiently block hepatocyte invasion by sporozoites." *Vaccine*, vol. 32, no. 19, pp. 2135–8, Apr. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0264410X14002400

[108] R. A. Seder, L.-J. Chang, M. E. Enama, K. L. Zephir, U. N. Sarwar, I. J. Gordon, L. A. Holman, E. R. James, P. F. Billingsley, A. Gunasekera, A. Richman,

S. Chakravarty, A. Manoj, S. Velmurugan, M. Li, A. J. Ruben, T. Li, A. G. Eappen, R. E. Stafford, S. H. Plummer, C. S. Hendel, L. Novik, P. J. M. Costner, F. H. Mendoza, J. G. Saunders, M. C. Nason, J. H. Richardson, J. Murphy, S. A. Davidson, T. L. Richie, M. Sedegah, A. Sutamihardja, G. A. Fahle, K. E. Lyke, M. B. Laurens, M. Roederer, K. Tewari, J. E. Epstein, B. K. L. Sim, J. E. Ledgerwood, B. S. Graham, and S. L. Hoffman, "Protection against malaria by intravenous immunization with a nonreplicating sporozoite vaccine." *Science (New York, N.Y.)*, vol. 341, no. 6152, pp. 1359–65, Sep. 2013. [Online]. Available: http://www.sciencemag.org/content/341/6152/1359

[109] M. J. Blackman, "Antibodies inhibit the protease-mediated processing of a malaria merozoite surface protein," *Journal of Experimental Medicine*, vol. 180, no. 1, pp. 389–393, Jul. 1994. [Online]. Available: http://jem.rupress.org/content/180/1/389

[110] B. R. Ogutu, O. J. Apollo, D. McKinney, W. Okoth, J. Siangla, F. Dubovsky, K. Tucker, J. N. Waitumbi, C. Diggs, J. Wittes, E. Malkin, A. Leach, L. A. Soisson, J. B. Milman, L. Otieno, C. A. Holland, M. Polhemus, S. A. Remich, C. F. Ockenhouse, J. Cohen, W. R. Ballou, S. K. Martin, E. Angov, V. A. Stewart, J. A. Lyon, D. G. Heppner, and M. R. Withers, "Blood stage malaria vaccine eliciting high antigen-specific antibody concentrations confers no protection to young children in Western Kenya." *PloS one*, vol. 4, no. 3, p. e4708, Jan. 2009. [Online]. Available: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0004708

[111] I. Sagara, A. Dicko, R. D. Ellis, M. P. Fay, S. I. Diawara, M. H. Assadou, M. S. Sissoko, M. Kone, A. I. Diallo, R. Saye, M. A. Guindo, O. Kante, M. B. Niambele, K. Miura, G. E. D. Mullen, M. Pierce, L. B. Martin, A. Dolo, D. A. Diallo, O. K. Doumbo, L. H. Miller, and A. Saul, "A randomized controlled phase 2 trial of the blood stage AMA1-C1/Alhydrogel malaria vaccine in children in Mali." *Vaccine*, vol. 27, no. 23, pp. 3090–8, May 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0264410X09004022

[112] A. D. Douglas, A. R. Williams, J. J. Illingworth, G. Kamuyu, S. Biswas, A. L. Goodman, D. H. Wyllie, C. Crosnier, K. Miura, G. J. Wright, C. A. Long, F. H. Osier, K. Marsh, A. V. Turner, A. V. S. Hill, and S. J. Draper, "The blood-stage malaria antigen PfRH5 is susceptible to vaccine-inducible cross-strain neutralizing antibody." *Nature communications*, vol. 2, p. 601, Jan. 2011. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3504505&tool=pmcentrez&rendertype=abstract

[113] K. E. Wright, K. A. Hjerrild, J. Bartlett, A. D. Douglas, J. Jin, R. E. Brown, J. J. Illingworth, R. Ashfield, S. B. Clemmensen, W. A. de Jongh, S. J. Draper, and M. K. Higgins, "Structure of malaria invasion protein RH5 with erythrocyte basigin and blocking antibodies." *Nature*, vol. 515, no. 7527, pp. 427–30, Nov. 2014. [Online]. Available: http://dx.doi.org/10.1038/nature13715

[114] A. D. Douglas, G. C. Baldeviano, C. M. Lucas, L. A. Lugo-Roman, C. Crosnier, S. J. Bartholdson, A. Diouf, K. Miura, L. E. Lambert, J. A. Ventocilla, K. P. Leiva, K. H. Milne, J. J. Illingworth, A. J. Spencer, K. A. Hjerrild, D. G. W. Alanine, A. V. Turner, J. T. Moorhead, K. A. Edgel, Y. Wu, C. A. Long, G. J. Wright, A. G. Lescano, and S. J. Draper, "A PfRH5-based vaccine is efficacious against Heterologous strain blood-stage Plasmodium falciparum infection in aotus monkeys." *Cell host & microbe*, vol. 17, no. 1, pp. 130–9, Jan. 2015.

[Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4297294&tool=pmcentrez&rendertype=abstract

[115] J. K. Baird, T. R. Jones, E. W. Danudirgo, B. A. Annis, M. J. Bangs, H. Basri, Purnomo, and S. Masbar, "Age-dependent acquired protection against Plasmodium falciparum in people having two years exposure to hyperendemic malaria," *American Journal of Tropical Medicine and Hygiene*, vol. 45, no. 1, pp. 65–76, 1991. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-0025777095&partnerID=tZOtx3y1

[116] G. K. K. Cham, L. Turner, J. Lusingu, L. Vestergaard, B. P. Mmbando, J. D. Kurtis, A. T. R. Jensen, A. Salanti, T. Lavstsen, and T. G. Theander, "Sequential, ordered acquisition of antibodies to Plasmodium falciparum erythrocyte membrane protein 1 domains." *Journal of immunology (Baltimore, Md. : 1950)*, vol. 183, no. 5, pp. 3356–63, Sep. 2009. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/19675168

[117] R. Cholera, N. J. Brittain, M. R. Gillrie, T. M. Lopera-Mesa, S. a. S. Diakité, T. Arie, M. a. Krause, A. Guindo, A. Tubman, H. Fujioka, D. a. Diallo, O. K. Doumbo, M. Ho, T. E. Wellems, and R. M. Fairhurst, "Impaired cytoadherence of Plasmodium falciparum-infected erythrocytes containing sickle hemoglobin." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 3, pp. 991–6, Jan. 2008. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2242681&tool=pmcentrez&rendertype=abstract

[118] J. S. McBride, D. Walliker, and G. Morgan, "Antigenic diversity in the human malaria parasite Plasmodium falciparum." *Science (New York, N.Y.)*, vol. 217, no. 4556, pp. 254–257, Jul. 1982. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/6178159

[119] J. H. Leech, J. W. Barnwell, L. H. Miller, and R. J. Howard, "Identification of a strain-specific malarial antigen exposed on the surface of Plasmodium falciparum-infected erythrocytes." *The Journal of experimental medicine*, vol. 159, no. 6, pp. 1567–75, Jun. 1984. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2187322&tool=pmcentrez&rendertype=abstract

[120] S. A. Kyes, Z. Christodoulou, A. Raza, P. Horrocks, R. Pinches, J. A. Rowe, and C. I. Newbold, "A well-conserved Plasmodium falciparum var gene shows an unusual stage-specific transcript pattern," *Molecular Microbiology*, vol. 48, no. 5, pp. 1339–1348, Apr. 2003. [Online]. Available: http://doi.wiley.com/10.1046/j.1365-2958.2003.03505.x

[121] C. Magowan, W. Wollish, L. Anderson, and J. Leech, "Cytoadherence by Plasmodium falciparum-infected erythrocytes is correlated with the expression of a family of variable proteins on infected erythrocytes." *The Journal of experimental medicine*, vol. 168, no. 4, pp. 1307–1320, Oct. 1988. [Online]. Available: http://jem.rupress.org/content/168/4/1307

[122] R. J. Howard, J. W. Barnwell, E. P. Rock, J. Neequaye, D. Ofori-Adjei, W. L. Maloy, J. A. Lyon, and A. Saul, "Two approximately 300 kilodalton Plasmodium falciparum proteins at the surface membrane of infected erythrocytes." *Molecular and biochemical parasitology*, vol. 27, no. 2-3, pp. 207–23, Jan. 1988. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/3278227

[123] S. G. Langreth and R. T. Reese, "Antigenicity of the infected-erythrocyte and merozoite surfaces in Falciparum malaria." *The Journal of experimental medicine*, vol. 150, no. 5, pp. 1241–1254, Nov. 1979. [Online]. Available: http://jem.rupress.org/content/150/5/1241.abstract

[124] S. a. Luse and L. H. Miller, "Plasmodium falciparum malaria. Ultrastructure of parasitized erythrocytes in cardiac vessels." *American Journal of Tropical Medicine and Hygiene*, vol. 20, no. 5, pp. 655–660, Sep. 1971. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/4999241

[125] B. S. Crabb, B. M. Cooke, J. C. Reeder, R. F. Waller, S. R. Caruana, K. M. Davern, M. E. Wickham, G. V. Brown, R. L. Coppel, and A. F. Cowman, "Targeted Gene Disruption Shows That Knobs Enable Malaria-Infected Red Cells to Cytoadhere under Physiological Shear Stress," *Cell*, vol. 89, no. 2, pp. 287–296, Apr. 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S009286740080207X

[126] S. Moghimi, "Mechanisms of splenic clearance of blood cells and particles: towards development of new splenotropic agents," *Advanced Drug Delivery Reviews*, vol. 17, no. 1, pp. 103–115, Oct. 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0169409X95000437

[127] R. Udomsangpetch, "Plasmodium falciparum-infected erythrocytes form spontaneous erythrocyte rosettes," *Journal of Experimental Medicine*, vol. 169, no. 5, pp. 1835–1840, May 1989. [Online]. Available: http://jem.rupress.org/content/169/5/1835.abstract

[128] J. A. Rowe, J. M. Moulds, C. I. Newbold, and L. H. Miller, "P. falciparum rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1," vol. 388, no. 6639, pp. 292–295, Jul. 1997.

[129] Q. Chen, A. Barragan, V. Fernandez, A. Sundstrom, M. Schlichtherle, A. Sahlen, J. Carlson, S. Datta, and M. Wahlgren, "Identification of Plasmodium falciparum Erythrocyte Membrane Protein 1 (PfEMP1) as the Rosetting Ligand of the Malaria Parasite P. falciparum," *Journal of Experimental Medicine*, vol. 187, no. 1, pp. 15–23, Jan. 1998. [Online]. Available: http://jem.rupress.org/content/187/1/15

[130] A. M. Dondorp, E. Pongponratn, and N. J. White, "Reduced microcirculatory flow in severe falciparum malaria: pathophysiology and electron-microscopic pathology," *Acta Tropica*, vol. 89, no. 3, pp. 309–317, Feb. 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0001706X03002742

[131] D. K. Kaul, E. F. Roth, R. L. Nagel, R. J. Howard, and S. M. Handunnetti, "Rosetting of Plasmodium falciparum-infected red blood cells with uninfected red blood cells enhances microvascular obstruction under flow conditions." *Blood*, vol. 78, no. 3, pp. 812–9, Aug. 1991. [Online]. Available: http://www.bloodjournal.org/content/78/3/812.abstract

[132] A. Pain, D. J. Ferguson, O. Kai, B. C. Urban, B. Lowe, K. Marsh, and D. J. Roberts, "Platelet-mediated clumping of Plasmodium falciparum-infected erythrocytes is a common adhesive phenotype and is associated with severe malaria." *Proceedings of the National Academy of Sciences of the United States*

of America, vol. 98, no. 4, pp. 1805–10, Feb. 2001. [Online]. Available: http://www.pnas.org/cgi/content/abstract/98/4/1805

[133] X.-z. Su, V. M. Heatwole, S. P. Wertheimer, F. Guinet, J. A. Herrfeldt, D. S. Peterson, J. A. Ravetch, and T. E. Wellems, "The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of plasmodium falciparum-infected erythrocytes," *Cell*, vol. 82, no. 1, pp. 89–100, Jul. 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/009286749590055 1http://dx.doi.org/10.1016/0092-8674(95)90055-1

[134] J. D. Smith, C. E. Chitnis, A. G. Craig, D. J. Roberts, D. E. Hudson-Taylor, D. S. Peterson, R. Pinches, C. I. Newbold, and L. H. Miller, "Switches in expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes." *Cell*, vol. 82, no. 1, pp. 101–10, Jul. 1995. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3730239&tool=pmcentrez&rendertype=abstract

[135] D. Camus and T. J. Hadley, "A Plasmodium falciparum antigen that binds to host erythrocytes and merozoites," *Science*, vol. 230, no. 4725, pp. 553–556, 1985. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-0022373071&partnerID=tZOtx3y1

[136] J. H. Adams, B. K. Sim, S. A. Dolan, X. Fang, D. C. Kaslow, and L. H. Miller, "A family of erythrocyte binding proteins of malaria parasites." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 15, pp. 7085–9, Aug. 1992. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=49650&tool=pmcentrez&rendertype=abstract

[137] M. M. Klein, A. G. Gittis, H.-P. Su, M. O. Makobongo, J. M. Moore, S. Singh, L. H. Miller, and D. N. Garboczi, "The cysteine-rich interdomain region from the highly variable plasmodium falciparum erythrocyte membrane protein-1 exhibits a conserved structure." *PLoS pathogens*, vol. 4, no. 9, p. e1000147, Jan. 2008. [Online]. Available: http://www.plospathogens.org/article/info:doi/10.1371/journal.ppat.1000147#s3

[138] J. D. Smith, A. G. Craig, N. Kriek, D. Hudson-Taylor, S. Kyes, T. Fagan, T. Fagen, R. Pinches, D. I. Baruch, C. I. Newbold, and L. H. Miller, "Identification of a Plasmodium falciparum intercellular adhesion molecule-1 binding domain: a parasite adhesion trait implicated in cerebral malaria." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 4, pp. 1766–71, Feb. 2000. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-12944270593&partnerID=tZOtx3y1

[139] J. Smith, S. Kyes, A. Craig, T. Fagan, D. Hudson-Taylor, L. Miller, D. Baruch, and C. Newbold, "Analysis of adhesive domains from the A4VAR Plasmodium falciparum erythrocyte membrane protein-1 identifies a CD36 binding domain," *Molecular and Biochemical Parasitology*, vol. 97, no. 1-2, pp. 133–148, Nov. 1998. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0166685198001455

[140] L. Turner, T. Lavstsen, S. S. Berger, C. W. Wang, J. E. V. Petersen, M. Avril, A. J. Brazier, J. Freeth, J. S. Jespersen, M. A. Nielsen, P. Magistrado, J. Lusingu, J. D. Smith, M. K. Higgins, and T. G. Theander, "Severe

malaria is associated with parasite binding to endothelial protein C receptor." *Nature*, vol. 498, no. 7455, pp. 502–5, Jun. 2013. [Online]. Available: http://www.nature.com/nature/journal/v498/n7455/nature12216/metrics/googleplus

[141] M. Dahlbäck, L. M. Jø rgensen, M. A. Nielsen, T. M. Clausen, S. B. Ditlev, M. Resende, V. V. Pinto, D. E. Arnot, T. G. Theander, and A. Salanti, "The chondroitin sulfate A-binding site of the VAR2CSA protein involves multiple N-terminal domains." *The Journal of biological chemistry*, vol. 286, no. 18, pp. 15 908–17, May 2011. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3091200&tool=pmcentrez&rendertype=abstract

[142] A. Srivastava, S. Gangnard, S. Dechavanne, F. Amirat, A. Lewit Bentley, G. A. Bentley, and B. Gamain, "Var2CSA minimal CSA binding region is located within the N-terminal region." *PloS one*, vol. 6, no. 5, p. e20270, Jan. 2011. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3098292&tool=pmcentrez&rendertype=abstract

[143] S. M. Kraemer and J. D. Smith, "A family affair: var genes, PfEMP1 binding, and malaria disease." *Current opinion in microbiology*, vol. 9, no. 4, pp. 374–80, Aug. 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16814594

[144] Y. Adams, P. Kuhnrae, M. K. Higgins, A. Ghumra, and J. A. Rowe, "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." *Infection and immunity*, vol. 82, no. 3, pp. 949–59, Mar. 2014. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3958005&tool=pmcentrez&rendertype=abstract

[145] K. L. Waller, "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)," *Journal of Biological Chemistry*, vol. 274, no. 34, pp. 23 808–23 813, Aug. 1999. [Online]. Available: http://www.jbc.org/cgi/content/abstract/274/34/23808

[146] A. K. Ganguly, P. Ranjan, A. Kumar, and N. S. Bhavesh, "Dynamic association of PfEMP1 and KAHRP in knobs mediates cytoadherence during Plasmodium invasion," *Scientific Reports*, vol. 5, p. 8617, Mar. 2015. [Online]. Available: http://www.nature.com/srep/2015/150302/srep08617/full/srep08617.html

[147] M. J. Gardner, N. Hall, E. Fung, O. White, M. Berriman, R. W. Hyman, J. M. Carlton, A. Pain, K. E. Nelson, S. Bowman, I. T. Paulsen, K. James, J. A. Eisen, K. Rutherford, S. L. Salzberg, A. Craig, S. Kyes, M.-S. Chan, V. Nene, S. J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A. B. Vaidya, D. M. A. Martin, A. H. Fairlamb, M. J. Fraunholz, D. S. Roos, S. A. Ralph, G. I. McFadden, L. M. Cummings, G. M. Subramanian, C. Mungall, J. C. Venter, D. J. Carucci, S. L. Hoffman, C. Newbold, R. W. Davis, C. M. Fraser, and B. Barrell, "Genome sequence of the human malaria parasite Plasmodium falciparum." *Nature*, vol. 419, no. 6906, pp. 498–511, Oct. 2002. [Online]. Available: http://dx.doi.org/10.1038/nature01097

[148] J. P. Rubio, J. K. Thompson, and A. F. Cowman, "The var genes of Plasmodium falciparum are located in the subtelomeric region of most

chromosomes." *The EMBO journal*, vol. 15, no. 15, pp. 4069–77, Aug. 1996. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=452127&tool=pmcentrez&rendertype=abstract

[149] K. Fischer, P. Horrocks, M. Preuss, J. Wiesner, S. Wünsch, A. A. Camargo, and M. Lanzer, "Expression of var genes located within polymorphic subtelomeric domains of Plasmodium falciparum chromosomes." *Molecular and cellular biology*, vol. 17, no. 7, pp. 3679–86, Jul. 1997. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=232219&tool=pmcentrez&rendertype=abstract

[150] R. Hernandez-Rivas, D. Mattei, Y. Sterkers, D. S. Peterson, T. E. Wellems, and A. Scherf, "Expressed var genes are found in Plasmodium falciparum subtelomeric regions." *Molecular and cellular biology*, vol. 17, no. 2, pp. 604–11, Feb. 1997. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=231785&tool=pmcentrez&rendertype=abstract

[151] M. Hommel, P. H. David, and L. D. Oligino, "Surface alterations of erythrocytes in Plasmodium falciparum malaria. Antigenic variation, antigenic diversity, and the role of the spleen." *The Journal of experimental medicine*, vol. 157, no. 4, pp. 1137–1148, Apr. 1983. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2186973&tool=pmcentrez&rendertype=abstract

[152] D. J. Roberts, A. G. Craig, A. R. Berendt, R. Pinches, G. Nash, K. Marsh, and C. I. Newbold, "Rapid switching to multiple antigenic and adhesive phenotypes in malaria." *Nature*, vol. 357, no. 6380, pp. 689–92, Jun. 1992. [Online]. Available: http://dx.doi.org/10.1038/357689a0

[153] A. Scherf, R. Hernandez-Rivas, P. Buffet, E. Bottius, C. Benatar, B. Pouvelle, J. Gysin, and M. Lanzer, "Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum." *The EMBO journal*, vol. 17, no. 18, pp. 5418–26, Sep. 1998. [Online]. Available: http://dx.doi.org/10.1093/emboj/17.18.5418

[154] A. Bachmann, S. Predehl, J. May, S. Harder, G. D. Burchard, T.-W. Gilberger, E. Tannich, and I. Bruchhaus, "Highly co-ordinated var gene expression and switching in clinical Plasmodium falciparum isolates from non-immune malaria patients." *Cellular microbiology*, vol. 13, no. 9, pp. 1397–409, Sep. 2011. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/21740496

[155] J. Guizetti and A. Scherf, "Silence, activate, poise and switch! Mechanisms of antigenic variation in Plasmodium falciparum." *Cellular microbiology*, vol. 15, no. 5, pp. 718–26, May 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3654561&tool=pmcentrez&rendertype=abstract

[156] M. T. Duraisingh, T. S. Voss, A. J. Marty, M. F. Duffy, R. T. Good, J. K. Thompson, L. H. Freitas-Junior, A. Scherf, B. S. Crabb, and A. F. Cowman, "Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum." *Cell*, vol. 121, no. 1, pp. 13–24, Apr. 2005. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/15820675

[157] K. Perez-Toledo, A. P. Rojas-Meza, L. Mancio-Silva, N. A. Hernandez-Cuevas, D. M. Delgadillo, M. Vargas, S. Martinez-Calvillo, A. Scherf, and R. Hernandez-Rivas, "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 Research*, vol. 37, no. 8, pp. 2596–2606, Mar. 2009. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2677873&tool=pmcentrez&rendertype=abstract

[158] J. Guizetti, R. M. Martins, S. Guadagnini, A. Claes, and A. Scherf, "Nuclear pores and perinuclear expression sites of var and ribosomal DNA genes correspond to physically distinct regions in Plasmodium falciparum." *Eukaryotic cell*, vol. 12, no. 5, pp. 697–702, May 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3647773&tool=pmcentrez&rendertype=abstract

[159] R. Dzikowski, M. Frank, and K. Deitsch, "Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production." *PLoS pathogens*, vol. 2, no. 3, p. e22, Mar. 2006. [Online]. Available: http://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.0020022

[160] T. S. Voss, J. Healer, A. J. Marty, M. F. Duffy, J. K. Thompson, J. G. Beeson, J. C. Reeder, B. S. Crabb, and A. F. Cowman, "A var gene promoter controls allelic exclusion of virulence genes in Plasmodium falciparum malaria." *Nature*, vol. 439, no. 7079, pp. 1004–8, Mar. 2006. [Online]. Available: http://dx.doi.org/10.1038/nature04407

[161] J. J. Lopez-Rubio, A. M. Gontijo, M. C. Nunes, N. Issar, R. Hernandez Rivas, and A. Scherf, "5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites." *Molecular microbiology*, vol. 66, no. 6, pp. 1296–305, Dec. 2007. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2228885&tool=pmcentrez&rendertype=abstract

[162] P. C. Bull, M. Kortok, O. Kai, F. Ndungu, A. Ross, B. S. Lowe, C. I. Newbold, and K. Marsh, "Plasmodium falciparum-infected erythrocytes: agglutination by diverse Kenyan plasma is associated with severe disease and young host age." *The Journal of infectious diseases*, vol. 182, no. 1, pp. 252–9, Jul. 2000. [Online]. Available: http://jid.oxfordjournals.org/content/182/1/252.long

[163] M. A. Nielsen, T. Staalsoe, J. r. A. L. Kurtzhals, B. Q. Goka, D. Dodoo, M. Alifrangis, T. G. Theander, B. D. Akanmori, and L. Hviid, "Plasmodium falciparum variant surface antigen expression varies between isolates causing severe and nonsevere malaria and is modified by acquired immunity." *Journal of immunology (Baltimore, Md. : 1950)*, vol. 168, no. 7, pp. 3444–50, Apr. 2002. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/11907103

[164] M. R. Hilleman, "Strategies and mechanisms for host and pathogen survival in acute and persistent viral infections." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101 Suppl, no. suppl_2, pp. 14 560–6, Oct. 2004. [Online]. Available: http://www.pnas.org/content/101/suppl_2/14560.abstract

[165] M. Nacher, P. Singhasivanon, U. Silachamroon, S. Treeprasertsuk, T. Tosukhowong, S. Vannaphan, F. Gay, D. Mazier, and S. Looareesuwan, "Decreased hemoglobin

concentrations, hyperparasitemia, and severe malaria are associated with increased Plasmodium falciparum gametocyte carriage," *Journal of Parasitology*, vol. 88, no. 1, pp. 97–101, 2002. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-0036200466&partnerID=tZOtx3y1

[166] T. Lavstsen, A. Salanti, A. T. R. Jensen, D. E. Arnot, and T. G. Theander, "Sub-grouping of Plasmodium falciparum 3D7 var genes based on sequence analysis of coding and non-coding regions." *Malaria journal*, vol. 2, no. 1, p. 27, Sep. 2003. [Online]. Available: http://www.malariajournal.com/content/2/1/27

[167] A. T. R. Jensen, P. Magistrado, S. Sharp, L. Joergensen, T. Lavstsen, A. Chiucchiuini, A. Salanti, L. S. Vestergaard, J. P. Lusingu, R. Hermsen, R. Sauerwein, J. Christensen, M. A. Nielsen, L. Hviid, C. Sutherland, T. Staalsoe, and T. G. Theander, "Plasmodium falciparum associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A var genes." *The Journal of experimental medicine*, vol. 199, no. 9, pp. 1179–90, May 2004. [Online]. Available: http://jem.rupress.org/content/199/9/1179.long

[168] A. Bengtsson, L. Joergensen, T. S. Rask, R. W. Olsen, M. A. Andersen, L. Turner, T. G. Theander, L. Hviid, M. K. Higgins, A. Craig, A. Brown, and A. T. R. Jensen, "A novel domain cassette identifies Plasmodium falciparum PfEMP1 proteins binding ICAM-1 and is a target of cross-reactive, adhesion-inhibitory antibodies." *Journal of immunology (Baltimore, Md. : 1950)*, vol. 190, no. 1, pp. 240–9, Jan. 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3539686&tool=pmcentrez&rendertype=abstract

[169] J. D. Smith, G. Subramanian, B. Gamain, D. I. Baruch, and L. H. Miller, "Classification of adhesive domains in the Plasmodium falciparum Erythrocyte Membrane Protein 1 family," *Molecular and Biochemical Parasitology*, vol. 110, no. 2, pp. 293–310, Oct. 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0166685100002796

[170] T. S. Rask, D. A. Hansen, T. G. Theander, A. Gorm Pedersen, and T. Lavstsen, "Plasmodium falciparum erythrocyte membrane protein 1 diversity in seven genomes–divide and conquer." *PLoS computational biology*, vol. 6, no. 9, Jan. 2010. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2940729&tool=pmcentrez&rendertype=abstract

[171] D. P.-G. Howell, E. A. Levin, A. L. Springer, S. M. Kraemer, D. J. Phippard, W. R. Schief, and J. D. Smith, "Mapping a common interaction site used by Plasmodium falciparum Duffy binding-like domains to bind diverse host receptors." *Molecular microbiology*, vol. 67, no. 1, pp. 78–87, Jan. 2008. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/18047571

[172] B. A. Robinson, T. L. Welch, and J. D. Smith, "Widespread functional specialization of Plasmodium falciparum erythrocyte membrane protein 1 family members to bind CD36 analysed across a parasite genome." *Molecular microbiology*, vol. 47, no. 5, pp. 1265–78, Mar. 2003. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/12603733

[173] S. M. Kraemer, S. A. Kyes, G. Aggarwal, A. L. Springer, S. O. Nelson, Z. Christodoulou, L. M. Smith, W. Wang, E. Levin, C. I. Newbold, P. J. Myler,

and J. D. Smith, "Patterns of gene recombination shape var gene repertoires in Plasmodium falciparum: comparisons of geographically diverse isolates." *BMC genomics*, vol. 8, p. 45, Jan. 2007. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1805758&tool=pmcentrez&rendertype=abstract

[174] G. Winter, Q. Chen, K. Flick, P. Kremsner, V. Fernandez, and M. Wahlgren, "The 3D7var5.2 (varCOMMON) type var gene family is commonly expressed in non-placental Plasmodium falciparum malaria," *Molecular and Biochemical Parasitology*, vol. 127, no. 2, pp. 179–191, Apr. 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0166685103000045

[175] A. Salanti, T. Staalsoe, T. Lavstsen, A. T. R. Jensen, M. P. K. Sowa, D. E. Arnot, L. Hviid, and T. G. Theander, "Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria," *Molecular Microbiology*, vol. 49, no. 1, pp. 179–191, May 2003. [Online]. Available: http://doi.wiley.com/10.1046/j.1365-2958.2003.03570.x

[176] C. W. Wang, T. Lavstsen, D. C. Bengtsson, P. A. Magistrado, S. S. Berger, A. M. Marquard, M. Alifrangis, J. P. Lusingu, T. G. Theander, and L. Turner, "Evidence for in vitro and in vivo expression of the conserved VAR3 (type 3) plasmodium falciparum erythrocyte membrane protein 1." *Malaria journal*, vol. 11, p. 129, Jan. 2012. [Online]. Available: /pmc/articles/PMC3407477/?report=abstract

[177] T. Lavstsen, L. Turner, F. Saguti, P. Magistrado, T. S. Rask, J. S. Jespersen, C. W. Wang, S. S. Berger, V. Baraka, A. M. Marquard, A. Seguin-Orlando, E. Willerslev, M. T. P. Gilbert, J. Lusingu, and T. G. Theander, "Plasmodium falciparum erythrocyte membrane protein 1 domain cassettes 8 and 13 are associated with severe malaria in children." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 26, pp. E1791–800, Jun. 2012. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22619319

[178] S. S. Berger, L. Turner, C. W. Wang, J. E. V. Petersen, M. Kraft, J. P. A. Lusingu, B. Mmbando, A. M. Marquard, D. B. A. C. Bengtsson, L. Hviid, M. A. Nielsen, T. G. Theander, and T. Lavstsen, "Plasmodium falciparum expressing domain cassette 5 type PfEMP1 (DC5-PfEMP1) bind PECAM1." *PloS one*, vol. 8, no. 7, p. e69117, Jan. 2013. [Online]. Available: /pmc/articles/PMC3706608/?report=abstract

[179] G. I. Bertin, T. Lavstsen, F. Guillonneau, J. Doritchamou, C. W. Wang, J. S. Jespersen, S. Ezimegnon, N. Fievet, M. J. Alao, F. Lalya, A. Massougbodji, N. T. Ndam, T. G. Theander, and P. Deloron, "Expression of the domain cassette 8 Plasmodium falciparum erythrocyte membrane protein 1 is associated with cerebral malaria in Benin." *PloS one*, vol. 8, no. 7, p. e68368, Jan. 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3726661&tool=pmcentrez&rendertype=abstract

[180] M. Avril, A. K. Tripathi, A. J. Brazier, C. Andisi, J. H. Janes, V. L. Soma, D. J. Sullivan, P. C. Bull, M. F. Stins, and J. D. Smith, "A restricted subset of var genes mediates adherence of Plasmodium falciparum-infected erythrocytes to brain endothelial cells." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 26, pp. E1782–90, Jun. 2012. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22619321

[181] A. Claessens, Y. Adams, A. Ghumra, G. Lindergard, C. C. Buchan, C. Andisi, Z. Bozdech, and J. A. Rowe, "A subset of group A-like var genes encodes the malaria parasite ligands for binding to human brain endothelial cells," vol. 109, no. 26, 2012.

[182] M. K. Higgins and M. Carrington, "Sequence variation and structural conservation allows development of novel function and immune evasion in parasite surface protein families." *Protein science : a publication of the Protein Society*, vol. 23, no. 4, pp. 354–65, Apr. 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/24442723

[183] N. H. Tolia, E. J. Enemark, B. K. L. Sim, and L. Joshua-Tor, "Structural basis for the EBA-175 erythrocyte invasion pathway of the malaria parasite Plasmodium falciparum." *Cell*, vol. 122, no. 2, pp. 183–93, Jul. 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0092867405005428

[184] S. K. Singh, R. Hora, H. Belrhali, C. E. Chitnis, and A. Sharma, "Structural basis for Duffy recognition by the malaria parasite Duffy-binding-like domain." *Nature*, vol. 439, no. 7077, pp. 741–4, Feb. 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16372020

[185] M. K. Higgins, "The structure of a chondroitin sulfate-binding domain important in placental malaria." *The Journal of biological chemistry*, vol. 283, no. 32, pp. 21 842–6, Aug. 2008. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2494935&tool=pmcentrez&rendertype=abstract

[186] P. J. B. D. H. M. Khunrae, P., "Structural Comparison of Two Cspg-Binding Dbl Domains from the Var2Csa Protein Important in Malaria During Pregnancy." *J.Mol.Biol.*, vol. 393, p. 202, 2009. [Online]. Available: http://www.rcsb.org/pdb/explore/explore.do?structureId=2WAU

[187] A. N. Hodder, P. E. Czabotar, A. D. Uboldi, O. B. Clarke, C. S. Lin, J. Healer, B. J. Smith, and A. F. Cowman, "Insights into Duffy binding-like domains through the crystal structure and function of the merozoite surface protein MSPDBL2 from Plasmodium falciparum." *The Journal of biological chemistry*, vol. 287, no. 39, pp. 32 922–39, Sep. 2012. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3463323&tool=pmcentrez&rendertype=abstract

[188] D. H. Lin, B. M. Malpede, J. D. Batchelor, and N. H. Tolia, "Crystal and solution structures of Plasmodium falciparum erythrocyte-binding antigen 140 reveal determinants of receptor specificity during erythrocyte invasion." *The Journal of biological chemistry*, vol. 287, no. 44, pp. 36 830–6, Oct. 2012. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3481285&tool=pmcentrez&rendertype=abstract

[189] I. Vigan-Womas, M. Guillotte, A. Juillerat, A. Hessel, B. Raynal, P. England, J. H. Cohen, O. Bertrand, T. Peyrard, G. A. Bentley, A. Lewit-Bentley, and O. Mercereau-Puijalon, "Structural basis for the ABO blood-group dependence of Plasmodium falciparum rosetting." *PLoS pathogens*, vol. 8, no. 7, p. e1002781, Jan. 2012. [Online]. Available: http://dx.plos.org/10.1371/journal.ppat.1002781

[190] S. Gangnard, C. Badaut, S. Ramboarina, B. Baron, T. Ramdani, B. Gamain, P. Deloron, A. Lewit-Bentley, and G. A. Bentley, "Structural and Immunological Correlations between the Variable Blocks of the Var2Csa Domain Dbl6Epsilon

from Two Plasmodium Falciparum Parasite Lines." *J.Mol.Biol.*, vol. 425, p. 1697, 2013. [Online]. Available: http://www.rcsb.org/pdb/explore/explore.do?structureId=2Y8D

[191] B. M. Malpede, D. H. Lin, and N. H. Tolia, "Molecular basis for sialic acid-dependent receptor recognition by the Plasmodium falciparum invasion protein erythrocyte-binding antigen-140/BAEBL." *The Journal of biological chemistry*, vol. 288, no. 17, pp. 12 406–15, Apr. 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3636924&tool=pmcentrez&rendertype=abstract

[192] E. Chen, M. M. Paing, N. Salinas, B. K. L. Sim, and N. H. Tolia, "Structural and functional basis for inhibition of erythrocyte invasion by antibodies that target Plasmodium falciparum EBA-175." *PLoS pathogens*, vol. 9, no. 5, p. e1003390, Jan. 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3662668&tool=pmcentrez&rendertype=abstract

[193] J. D. Batchelor, B. M. Malpede, N. S. Omattage, G. T. DeKoster, K. A. Henzler-Wildman, and N. H. Tolia, "Red Blood Cell Invasion by Plasmodium vivax: Structural Basis for DBP Engagement of DARC," *PLoS Pathogens*, vol. 10, no. 1, p. e1003869, Jan. 2014. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3887093&tool=pmcentrez&rendertype=abstract

[194] D. I. Baruch, J. A. Gormely, C. Ma, R. J. Howard, and B. L. Pasloske, "Plasmodium falciparum erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 8, pp. 3497–502, Apr. 1996. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=39638&tool=pmcentrez&rendertype=abstract

[195] J. C. Reeder, A. F. Cowman, K. M. Davern, J. G. Beeson, J. K. Thompson, S. J. Rogerson, and G. V. Brown, "The adhesion of Plasmodium falciparum-infected erythrocytes to chondroitin sulfate A is mediated by P. falciparum erythrocyte membrane protein 1." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 9, pp. 5198–202, Apr. 1999. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=21841&tool=pmcentrez&rendertype=abstract

[196] A. K. Biswas, A. Hafiz, B. Banerjee, K. S. Kim, K. Datta, and C. E. Chitnis, "Plasmodium falciparum uses gC1qR/HABP1/p32 as a receptor to bind to vascular endothelium and for platelet-mediated clumping." *PLoS pathogens*, vol. 3, no. 9, pp. 1271–80, Sep. 2007. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2323294&tool=pmcentrez&rendertype=abstract

[197] A. Ghumra, J. P. Semblat, R. S. McIntosh, A. Raza, I. B. Rasmussen, R. Braathen, F. E. Johansen, I. Sandlie, P. K. Mongini, J. A. Rowe, and R. J. Pleass, "Identification of residues in the Cμ4 domain of polymeric IgM essential for interaction with plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1)," *Journal of Immunology*, vol. 181, no. 3, pp. 1988–2000, 2008. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-49649106080&partnerID=tZOtx3y1

[198] A. M. Senczuk, "Plasmodium falciparum erythrocyte membrane protein 1 functions as a ligand for P-selectin," *Blood*, vol. 98, no. 10, pp. 3132–3135, Nov. 2001. [Online]. Available: http://www.bloodjournal.org/content/98/10/3132.abstract

[199] T. Hatabu, S.-I. Kawazu, M. Aikawa, and S. Kano, "Binding of Plasmodium falciparum-infected erythrocytes to the membrane-bound form of Fractalkine/CX3CL1." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 26, pp. 15 942–6, Dec. 2003. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=307672&tool=pmcentrez&rendertype=abstract

[200] C. F. Ockenhouse, "Human vascular endothelial cell adhesion receptors for Plasmodium falciparum-infected erythrocytes: roles for endothelial leukocyte adhesion molecule 1 and vascular cell adhesion molecule 1," *Journal of Experimental Medicine*, vol. 176, no. 4, pp. 1183–1189, Oct. 1992. [Online]. Available: http://jem.rupress.org/content/176/4/1183

[201] J. H. Janes, C. P. Wang, E. Levin-Edens, I. Vigan-Womas, M. Guillotte, M. Melcher, O. Mercereau-Puijalon, and J. D. Smith, "Investigating the host binding signature on the Plasmodium falciparum PfEMP1 protein family." *PLoS pathogens*, vol. 7, no. 5, p. e1002032, May 2011. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3088720&tool=pmcentrez&rendertype=abstract

[202] C. Ockenhouse, M. Ho, and N. Tandon, "Molecular Basis of Sequestration in Severe and Uncomplicated Plasmodium falciparum Malaria: Differential Adhesion of Infected Erythrocytes to CD36 and ICAM-l," *Journal of Infectious Diseases*, vol. 164, no. 1, pp. 163–169, 1991. [Online]. Available: http://jid.oxfordjournals.org/content/164/1/163.short

[203] J. W. Barnwell, C. F. Ockenhouse, and D. M. Knowles, "Monoclonal antibody OKM5 inhibits the in vitro binding of Plasmodium falciparum-infected erythrocytes to monocytes, endothelial, and C32 melanoma cells." *Journal of immunology (Baltimore, Md. : 1950)*, vol. 135, no. 5, pp. 3494–7, Nov. 1985. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/3900210

[204] J. W. Barnwell, A. S. Asch, R. L. Nachman, M. Yamaya, M. Aikawa, and P. Ingravallo, "A human 88-kD membrane glycoprotein (CD36) functions in vitro as a receptor for a cytoadherence ligand on Plasmodium falciparum-infected erythrocytes." *The Journal of clinical investigation*, vol. 84, no. 3, pp. 765–72, Sep. 1989. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=329717&tool=pmcentrez&rendertype=abstract

[205] M. Febbraio, D. P. Hajjar, and R. L. Silverstein, "CD36: a class B scavenger receptor involved in angiogenesis, atherosclerosis, inflammation, and lipid metabolism." *The Journal of clinical investigation*, vol. 108, no. 6, pp. 785–91, Sep. 2001. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=200943&tool=pmcentrez&rendertype=abstract

[206] I. N. Baranova, R. Kurlander, A. V. Bocharov, T. G. Vishnyakova, Z. Chen, A. T. Remaley, G. Csako, A. P. Patterson, and T. L. Eggerman, "Role of Human CD36 in Bacterial Recognition, Phagocytosis, and Pathogen-Induced JNK-Mediated Signaling," *The Journal of Immunology*, vol. 181, no. 10, pp. 7147–7156, Nov. 2008. [Online]. Available: http://www.jimmunol.org/content/181/10/7147.full

[207] I. D. McGilvray, L. Serghides, A. Kapus, O. D. Rotstein, and K. C. Kain, "Nonopsonic monocyte/macrophage phagocytosis of Plasmodium falciparum-parasitized erythrocytes: a role for CD36 in malarial clearance." *Blood*, vol. 96, no. 9, pp. 3231–40, Nov. 2000. [Online]. Available: http://www.bloodjournal.org/content/96/9/3231.abstract

[208] S. N. Patel, Z. Lu, K. Ayi, L. Serghides, D. C. Gowda, and K. C. Kain, "Disruption of CD36 Impairs Cytokine Response to Plasmodium falciparum Glycosylphosphatidylinositol and Confers Susceptibility to Severe and Fatal Malaria In Vivo," *The Journal of Immunology*, vol. 178, no. 6, pp. 3954–3961, Mar. 2007. [Online]. Available: http://www.jimmunol.org/content/178/6/3954.full

[209] B. C. Urban, D. J. Ferguson, A. Pain, N. Willcox, M. Plebanski, J. M. Austyn, and D. J. Roberts, "Plasmodium falciparum-infected erythrocytes modulate the maturation of dendritic cells." *Nature*, vol. 400, no. 6739, pp. 73–7, Jul. 1999. [Online]. Available: http://dx.doi.org/10.1038/21900

[210] C. Newbold, P. Warn, G. Black, A. Berendt, A. Craig, B. O. B. Snow, M. Msobo, and N. Peshu, "Receptor-specific adhesion and clinical disease in Plasmodium falciparum," *Tropical Medicine*, vol. 57, no. 4, pp. 389–398, 1997.

[211] M. L. Dustin, R. Rothlein, A. K. Bhan, C. A. Dinarello, and T. A. Springer, "Induction by IL 1 and interferon-gamma: tissue distribution, biochemistry, and function of a natural adherence molecule (ICAM-1)." *Journal of immunology (Baltimore, Md. : 1950)*, vol. 137, no. 1, pp. 245–54, Jul. 1986. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/3086451

[212] G. B. Nash, C. D. Buckley, and G. Ed Rainger, "The local physicochemical environment conditions the proinflammatory response of endothelial cells and thus modulates leukocyte recruitment." *FEBS letters*, vol. 569, no. 1-3, pp. 13–7, Jul. 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0014579304006817

[213] G. D. Turner, H. Morrison, M. Jones, T. M. Davis, S. Looareesuwan, I. D. Buley, K. C. Gatter, C. I. Newbold, S. Pukritayakamee, and B. Nagachinta, "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." *The American journal of pathology*, vol. 145, no. 5, pp. 1057–69, Nov. 1994. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1887431&tool=pmcentrez&rendertype=abstract

[214] K. Silamut, N. H. Phu, C. Whitty, G. D. Turner, K. Louwrier, N. T. Mai, J. A. Simpson, T. T. Hien, and N. J. White, "A quantitative analysis of the microvascular sequestration of malaria parasites in the human brain." *The American journal of pathology*, vol. 155, no. 2, pp. 395–410, Aug. 1999. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S000294401065136X

[215] F. Lennartz, A. Bengtsson, R. W. Olsen, L. Joergensen, A. Brown, L. Remy, P. Man, E. Forest, L. K. Barfod, Y. Adams, M. K. Higgins, and A. T. R. Jensen, "Mapping the Binding Site of a Cross-Reactive Plasmodium falciparum PfEMP1 Monoclonal Antibody Inhibitory of ICAM-1 Binding." *Journal of immunology (Baltimore, Md. : 1950)*, Aug. 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/26320251

[216] C. Bertonati and A. Tramontano, "A model of the complex between the PfEMP1 malaria protein and the human ICAM-1 receptor." *Proteins*, vol. 69, no. 2, pp. 215–22, Nov. 2007. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/17640071

[217] A. R. Berendt, A. McDowall, A. G. Craig, P. A. Bates, M. J. Sternberg, K. Marsh, C. I. Newbold, and N. Hogg, "The binding site on ICAM-1 for plasmodium falciparum-infected erythrocytes overlaps, but is distinct from, the LFA-1-binding site," *Cell*, vol. 68, no. 1, pp. 71–81, Jan. 1992. [Online]. Available: http://www.sciencedirect.com/science/article/pii/009286749290207S

[218] M. Shimaoka, T. Xiao, J.-H. Liu, Y. Yang, Y. Dong, C.-D. Jun, A. McCormack, R. Zhang, A. Joachimiak, J. Takagi, J.-H. Wang, and T. A. Springer, "Structures of the αL I Domain and Its Complex with ICAM-1 Reveal a Shape-Shifting Pathway for Integrin Regulation," *Cell*, vol. 112, no. 1, pp. 99–111, Jan. 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0092867402012576

[219] M. Fried and P. E. Duffy, "Adherence of Plasmodium falciparum to chondroitin sulfate A in the human placenta." *Science (New York, N.Y.)*, vol. 272, no. 5267, pp. 1502–4, Jun. 1996. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/8633247

[220] A. Salanti, M. Dahlbäck, L. Turner, M. A. Nielsen, L. Barfod, P. Magistrado, A. T. R. Jensen, T. Lavstsen, M. F. Ofori, K. Marsh, L. Hviid, and T. G. Theander, "Evidence for the involvement of VAR2CSA in pregnancy-associated malaria." *The Journal of experimental medicine*, vol. 200, no. 9, pp. 1197–203, Nov. 2004. [Online]. Available: http://jem.rupress.org/cgi/content/abstract/200/9/1197

[221] N. G. Tuikue Ndam, A. Salanti, G. Bertin, M. Dahlbäck, N. Fievet, L. Turner, A. Gaye, T. Theander, and P. Deloron, "High level of var2csa transcription by Plasmodium falciparum isolated from the placenta." *The Journal of infectious diseases*, vol. 192, no. 2, pp. 331–5, Jul. 2005. [Online]. Available: http://jid.oxfordjournals.org/content/192/2/331.long

[222] B. Gamain, A. R. Trimnell, C. Scheidig, A. Scherf, L. H. Miller, and J. D. Smith, "Identification of multiple chondroitin sulfate A (CSA)-binding domains in the var2CSA gene transcribed in CSA-binding parasites." *The Journal of infectious diseases*, vol. 191, no. 6, pp. 1010–3, Mar. 2005. [Online]. Available: http://jid.oxfordjournals.org/content/191/6/1010.long#ref-10

[223] I. Naka, J. Patarapotikul, H. Hananantachai, H. Imai, and J. Ohashi, "Association of the endothelial protein C receptor (PROCR) rs867186-G allele with protection from severe malaria." *Malaria journal*, vol. 13, no. 1, p. 105, Jan. 2014. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84899062985&partnerID=tZOtx3y1

[224] C. A. Moxon, S. C. Wassmer, D. A. Milner, N. V. Chisala, T. E. Taylor, K. B. Seydel, M. E. Molyneux, B. Faragher, C. T. Esmon, C. Downey, C.-H. Toh, A. G. Craig, and R. S. Heyderman, "Loss of endothelial protein C receptors links coagulation and inflammation to parasite sequestration in cerebral malaria in African children." *Blood*, vol. 122, no. 5, pp. 842–51, Aug. 2013. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84885941107&partnerID=tZOtx3y1

[225] A. Moussiliou, M. J. Alao, L. Denoeud-Ndam, R. Tahar, S. Ezimegnon, G. Sagbo, A. Amoussou, A. J. F. Luty, P. Deloron, and N. Tuikue Ndam, "High plasma levels of soluble endothelial protein C receptor are associated with increased mortality among children with cerebral malaria in Benin." *The Journal of infectious diseases*, vol. 211, no. 9, pp. 1484–8, May 2015. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84928911761&partnerID=tZOtx3y1

[226] Z. Laszik, A. Mitro, F. B. Taylor, G. Ferrell, and C. T. Esmon, "Human protein C receptor is present primarily on endothelium of large blood vessels: implications for the control of the protein C pathway." *Circulation*, vol. 96, no. 10, pp. 3633–40, Nov. 1997. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/9396465

[227] V. Oganesyan, N. Oganesyan, S. Terzyan, D. Qu, Z. Dauter, N. L. Esmon, and C. T. Esmon, "The crystal structure of the endothelial protein C receptor and a bound phospholipid." *The Journal of biological chemistry*, vol. 277, no. 28, pp. 24 851–4, Jul. 2002. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/12034704

[228] K. Fukudome, S. Kurosawa, D. J. Stearns-Kurosawa, X. He, A. R. Rezaie, and C. T. Esmon, "The Endothelial Cell Protein C Receptor: CELL SURFACE EXPRESSION AND DIRECT LIGAND BINDING BY THE SOLUBLE RECEPTOR," *Journal of Biological Chemistry*, vol. 271, no. 29, pp. 17 491–17 498, Jul. 1996. [Online]. Available: http://www.jbc.org/content/271/29/17491.long

[229] C. T. Esmon, "The Protein C Pathway," *CHEST Journal*, vol. 124, no. 3_suppl, p. 26S, Sep. 2003. [Online]. Available: http://journal.publications.chestnet.org/article.aspx?articleid=1081893

[230] M. Riewald, R. J. Petrovan, A. Donner, B. M. Mueller, and W. Ruf, "Activation of endothelial cell protease activated receptor 1 by the protein C pathway." *Science (New York, N.Y.)*, vol. 296, no. 5574, pp. 1880–2, Jun. 2002. [Online]. Available: http://www.sciencemag.org/content/296/5574/1880.long

[231] J.-S. Bae and A. R. Rezaie, "Protease activated receptor 1 (PAR-1) activation by thrombin is protective in human pulmonary artery endothelial cells if endothelial protein C receptor is occupied by its natural ligand." *Thrombosis and haemostasis*, vol. 100, no. 1, pp. 101–9, Jul. 2008. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2562749&tool=pmcentrez&rendertype=abstract

[232] C. T. Esmon, K. Fukudome, T. Mather, W. Bode, L. M. Regan, D. J. Stearns-Kurosawa, and S. Kurosawa, "Inflammation, sepsis, and coagulation." *Haematologica*, vol. 84, no. 3, pp. 254–9, Mar. 1999. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/10189392

[233] A. Rahman, "Galphaq and Gbetagamma Regulate PAR-1 Signaling of Thrombin-Induced NF-kappaB Activation and ICAM-1 Transcription in Endothelial Cells," *Circulation Research*, vol. 91, no. 5, pp. 398–405, Aug. 2002. [Online]. Available: http://circres.ahajournals.org/content/91/5/398

[234] J. Doritchamou, P. Bigey, M. A. Nielsen, S. Gnidehou, S. Ezinmegnon, A. Burgain, A. Massougbodji, P. Deloron, A. Salanti, and N. T. Ndam, "Differential adhesion-inhibitory patterns of antibodies raised against two major variants of the NTS-DBL2X region of VAR2CSA." *Vaccine*, vol. 31, no. 41, pp. 4516–22,

Sep. 2013. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84883600510&partnerID=tZOtx3y1

[235] M. A. Nielsen, M. Resende, W. A. de Jongh, S. B. Ditlev, B. Mordmüller, S. Houard, N. T. Ndam, M. O. Agerbæ k, M. Hamborg, A. Massougbodji, S. Issifou, A. Strø bæk, L. Poulsen, O. Leroy, P. G. Kremsner, J.-P. Chippaux, A. J. F. Luty, P. Deloron, T. G. Theander, C. Dyring, and A. Salanti, "The Influence of Sub-Unit Composition and Expression System on the Functional Antibody Response in the Development of a VAR2CSA Based Plasmodium falciparum Placental Malaria Vaccine." *PloS one*, vol. 10, no. 9, p. e0135406, Jan. 2015. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4556615&tool=pmcentrez&rendertype=abstract

[236] A. Babakhanyan, R. G. F. Leke, A. Salanti, N. Bobbili, P. Gwanmesia, R. J. I. Leke, I. A. Quakyi, J. J. Chen, and D. W. Taylor, "The Antibody Response of Pregnant Cameroonian Women to VAR2CSA ID1-ID2a, a Small Recombinant Protein Containing the CSA-Binding Site," *PLoS ONE*, vol. 9, no. 2, p. e88173, Feb. 2014. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3913775&tool=pmcentrez&rendertype=abstract

[237] A. T. Heiny, O. Miotto, K. N. Srinivasan, A. M. Khan, G. L. Zhang, V. Brusic, T. W. Tan, and J. T. August, "Evolutionarily conserved protein sequences of influenza a viruses, avian and human, as vaccine targets." *PloS one*, vol. 2, no. 11, p. e1190, Jan. 2007. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-43149084860&partnerID=tZOtx3y1

[238] R. Pantophlet and D. R. Burton, "GP120: target for neutralizing HIV-1 antibodies." *Annual review of immunology*, vol. 24, pp. 739–69, Jan. 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16551265

[239] R. R. Garrity, G. Rimmelzwaan, A. Minassian, W. P. Tsai, G. Lin, J. J. de Jong, J. Goudsmit, and P. L. Nara, "Refocusing neutralizing antibody response by targeted dampening of an immunodominant epitope." *Journal of immunology (Baltimore, Md. : 1950)*, vol. 159, no. 1, pp. 279–89, Jul. 1997. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/9200464

[240] R. Pantophlet, I. A. Wilson, and D. R. Burton, "Hyperglycosylated mutants of human immunodeficiency virus (HIV) type 1 monomeric gp120 as novel antigens for HIV vaccine design." *Journal of virology*, vol. 77, no. 10, pp. 5889–901, May 2003. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=154011&tool=pmcentrez&rendertype=abstract

[241] R. D. Astronomo, H.-K. Lee, C. N. Scanlan, R. Pantophlet, C.-Y. Huang, I. A. Wilson, O. Blixt, R. A. Dwek, C.-H. Wong, and D. R. Burton, "A glycoconjugate antigen based on the recognition motif of a broadly neutralizing human immunodeficiency virus antibody, 2G12, is immunogenic but elicits antibodies unable to bind to the self glycans of gp120." *Journal of virology*, vol. 82, no. 13, pp. 6359–68, Jul. 2008. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2447108&tool=pmcentrez&rendertype=abstract

[242] R. Pantophlet, I. A. Wilson, and D. R. Burton, "Improved design of an antigen with enhanced specificity for the broadly HIV-neutralizing antibody b12." *Protein*

engineering, design & selection : PEDS, vol. 17, no. 10, pp. 749–58, Oct. 2004. [Online]. Available: http://peds.oxfordjournals.org/content/17/10/749.long

[243] S.-C. Lin, Y.-F. Lin, P. Chong, and S.-C. Wu, "Broader neutralizing antibodies against H5N1 viruses using prime-boost immunization of hyperglycosylated hemagglutinin DNA and virus-like particles." *PloS one*, vol. 7, no. 6, p. e39075, Jan. 2012. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3374787&tool=pmcentrez&rendertype=abstract

[244] S. Sampath, C. Carrico, J. Janes, S. Gurumoorthy, C. Gibson, M. Melcher, C. E. Chitnis, R. Wang, W. R. Schief, and J. D. Smith, "Glycan masking of Plasmodium vivax Duffy Binding Protein for probing protein binding function and vaccine development." *PLoS pathogens*, vol. 9, no. 6, p. e1003420, Jan. 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3681752&tool=pmcentrez&rendertype=abstract

[245] C. E. Schafmeister, J. Po, and G. L. Verdine, "An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides," *Journal of the American Chemical Society*, vol. 122, no. 24, pp. 5891–5892, Jun. 2000. [Online]. Available: http://dx.doi.org/10.1021/ja000563a

[246] G. H. Bird, A. Irimia, G. Ofek, P. D. Kwong, I. A. Wilson, and L. D. Walensky, "Stapled HIV-1 peptides recapitulate antigenic structures and engage broadly neutralizing antibodies." *Nature structural & molecular biology*, vol. 21, no. 12, pp. 1058–67, Dec. 2014. [Online]. Available: http://dx.doi.org/10.1038/nsmb.2922

[247] A. W. Purcell, J. McCluskey, and J. Rossjohn, "More than one reason to rethink the use of peptides in vaccine design." *Nature reviews. Drug discovery*, vol. 6, no. 5, pp. 404–14, May 2007. [Online]. Available: http://dx.doi.org/10.1038/nrd2224

[248] A. Leaver-Fay, M. Tyka, S. M. Lewis, O. F. Lange, J. Thompson, R. Jacak, K. Kaufman, P. D. Renfrew, C. A. Smith, W. Sheffler, I. W. Davis, S. Cooper, A. Treuille, D. J. Mandell, F. Richter, Y.-E. A. Ban, S. J. Fleishman, J. E. Corn, D. E. Kim, S. Lyskov, M. Berrondo, S. Mentzer, Z. Popović, J. J. Havranek, J. Karanicolas, R. Das, J. Meiler, T. Kortemme, J. J. Gray, B. Kuhlman, D. Baker, and P. Bradley, "ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules." *Methods in enzymology*, vol. 487, pp. 545–74, Jan. 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780123812704000196

[249] J. S. McLellan, B. E. Correia, M. Chen, Y. Yang, B. S. Graham, W. R. Schief, and P. D. Kwong, "Design and characterization of epitope-scaffold immunogens that present the motavizumab epitope from respiratory syncytial virus." *Journal of molecular biology*, vol. 409, no. 5, pp. 853–66, Jun. 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022283611004736

[250] M. L. Azoitei, B. E. Correia, Y.-E. A. Ban, C. Carrico, O. Kalyuzhniy, L. Chen, A. Schroeter, P.-S. Huang, J. S. McLellan, P. D. Kwong, D. Baker, R. K. Strong, and W. R. Schief, "Computation-Guided Backbone Grafting of a Discontinuous Motif onto a Protein Scaffold," *Science*, vol. 334, no. 6054, pp. 373–376, Oct. 2011. [Online]. Available: http://www.sciencemag.org/content/334/6054/373.full

[251] M. L. Azoitei, Y.-E. A. Ban, J.-P. Julien, S. Bryson, A. Schroeter, O. Kalyuzhniy, J. R. Porter, Y. Adachi, D. Baker, E. F. Pai, and W. R. Schief, "Computational design of high-affinity epitope scaffolds by backbone grafting of a linear epitope." *Journal of molecular biology*, vol. 415, no. 1, pp. 175–92, Jan. 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022283611011272

[252] B. E. Correia, J. T. Bates, R. J. Loomis, G. Baneyx, C. Carrico, J. G. Jardine, P. Rupert, C. Correnti, O. Kalyuzhniy, V. Vittal, M. J. Connell, E. Stevens, A. Schroeter, M. Chen, S. Macpherson, A. M. Serra, Y. Adachi, M. A. Holmes, Y. Li, R. E. Klevit, B. S. Graham, R. T. Wyatt, D. Baker, R. K. Strong, J. E. Crowe, P. R. Johnson, and W. R. Schief, "Proof of principle for epitope-focused vaccine design." *Nature*, vol. 507, no. 7491, pp. 201–6, Mar. 2014. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4260937&tool=pmcentrez&rendertype=abstract

[253] J. Shi, T. L. Blundell, and K. Mizuguchi, "FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties." *Journal of molecular biology*, vol. 310, no. 1, pp. 243–57, Jun. 2001. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/11419950

[254] F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, and D. G. Higgins, "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega." *Molecular systems biology*, vol. 7, no. 1, p. 539, Jan. 2011. [Online]. Available: http://msb.embopress.org/content/7/1/539.abstract

[255] X. Robert and P. Gouet, "Deciphering key features in protein structures with the new ENDscript server." *Nucleic acids research*, vol. 42, no. Web Server issue, pp. W320–4, Jul. 2014. [Online]. Available: http://nar.oxfordjournals.org/content/42/W1/W320.full

[256] Schrödinger, LLC, "The PyMOL molecular graphics system, version 1.3r1," August 2010.

[257] F. Studier and B. A. Moffatt, "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes," *Journal of Molecular Biology*, vol. 189, no. 1, pp. 113–130, May 1986. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-0023042283&partnerID=tZOtx3y1

[258] G. E. Smith, M. D. Summers, and M. J. Fraser, "Production of human beta interferon in insect cells infected with a baculovirus expression vector." *Mol. Cell. Biol.*, vol. 3, no. 12, pp. 2156–2165, Dec. 1983. [Online]. Available: http://mcb.asm.org/content/3/12/2156.abstract?ijkey=fd0dd608b911413a91837e761894f65878ed1318&keytype2=tf_ipsecsha

[259] M. M. van Oers, G. P. Pijlman, and J. M. Vlak, "Thirty years of baculovirus-insect cell protein expression: from dark horse to mainstream technology." *The Journal of general virology*, vol. 96, no. Pt 1, pp. 6–23, Jan. 2015. [Online]. Available: http://vir.sgmjournals.org/content/96/Pt_1/6.long

[260] S. M. Kelly, T. J. Jess, and N. C. Price, "How to study proteins by circular dichroism." *Biochimica et biophysica acta*, vol. 1751, no. 2, pp. 119–39, Aug. 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1570963905001792

[261] C. Louis-Jeune, M. A. Andrade-Navarro, and C. Perez-Iratxeta, "Prediction of protein secondary structure from circular dichroism using theoretically derived spectra." *Proteins*, vol. 80, no. 2, pp. 374–81, Feb. 2012. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22095872

[262] J. L. Cole, J. W. Lary, T. P Moody, and T. M. Laue, "Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium." *Methods in cell biology*, vol. 84, pp. 143–79, Jan. 2008. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2711687&tool=pmcentrez&rendertype=abstract

[263] P. Schuck, "Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling." *Biophysical journal*, vol. 78, no. 3, pp. 1606–19, Mar. 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0006349500767130

[264] J. Vistica, J. Dam, A. Balbo, E. Yikilmaz, R. A. Mariuzza, T. A. Rouault, and P. Schuck, "Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition." *Analytical biochemistry*, vol. 326, no. 2, pp. 234–56, Mar. 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0003269704000363

[265] M. M. Pierce, C. S. Raman, and B. T. Nall, "Isothermal titration calorimetry of protein-protein interactions." *Methods (San Diego, Calif.)*, vol. 19, no. 2, pp. 213–21, Oct. 1999. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/10527727

[266] R. S. Stein and G. Hadziioannou, "Generalization of the Zimm equation for scattering from concentrated solutions," *Macromolecules*, vol. 17, no. 5, pp. 1059–1062, Sep. 1984. [Online]. Available: http://dx.doi.org/10.1021/ma00135a014

[267] P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, and D. I. Svergun, "PRIMUS : a Windows PC-based system for small-angle scattering data analysis," *Journal of Applied Crystallography*, vol. 36, no. 5, pp. 1277–1282, Sep. 2003. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?ks5004

[268] D. I. Svergun, "Determination of the regularization parameter in indirect-transform methods using perceptual criteria," *Journal of Applied Crystallography*, vol. 25, no. 4, pp. 495–503, Aug. 1992. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?wi0087

[269] D. Franke and D. I. Svergun, "DAMMIF , a program for rapid ab-initio shape determination in small-angle scattering," *Journal of Applied Crystallography*, vol. 42, no. 2, pp. 342–346, Jan. 2009. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?aj5117

[270] V. V. Volkov and D. I. Svergun, "Uniqueness of ab initio shape determination in small-angle scattering," *Journal of Applied Crystallography*, vol. 36, no. 3, pp. 860–864, Apr. 2003. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?ks7125

[271] D. I. Svergun, "Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing." *Biophysical journal*, vol. 76, no. 6, pp. 2879–86, Jun. 1999. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1300260&tool=pmcentrez&rendertype=abstract

[272] W. Wriggers, "Using Situs for the integration of multi-resolution structures." *Biophysical reviews*, vol. 2, no. 1, pp. 21–27, Feb. 2010. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2821521&tool=pmcentrez&rendertype=abstract

[273] S. Birmanns, M. Rusu, and W. Wriggers, "Using Sculptor and Situs for simultaneous assembly of atomic components into low-resolution shapes." *Journal of structural biology*, vol. 173, no. 3, pp. 428–35, Mar. 2011. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3352678&tool=pmcentrez&rendertype=abstract

[274] J. C. KENDREW, G. BODO, H. M. DINTZIS, R. G. PARRISH, H. WYCKOFF, and D. C. PHILLIPS, "A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis," *Nature*, vol. 181, no. 4610, pp. 662–666, Mar. 1958. [Online]. Available: http://dx.doi.org/10.1038/181662a0

[275] T. Bergfors, "Seeds to crystals." *Journal of structural biology*, vol. 142, no. 1, pp. 66–76, Apr. 2003. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/12718920

[276] G. Winter and K. E. McAuley, "Automated data collection for macromolecular crystallography." *Methods (San Diego, Calif.)*, vol. 55, no. 1, pp. 81–93, Sep. 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1046202311001265

[277] G. Winter, "xia2 : an expert system for macromolecular crystallography data reduction," *Journal of Applied Crystallography*, vol. 43, no. 1, pp. 186–190, Dec. 2009. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?S0021889809045701

[278] M.-F. Incardona, G. P. Bourenkov, K. Levik, R. A. Pieritz, A. N. Popov, and O. Svensson, "EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis." *Journal of synchrotron radiation*, vol. 16, no. Pt 6, pp. 872–9, Nov. 2009. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?S0909049509036681

[279] A. G. W. Leslie and H. R. Powell, "Processing diffraction data with mosflm," *Evolving Methods for Macromolecular Crystallography*, p. 41, 2007.

[280] P. R. Evans, "An introduction to data reduction: space-group determination, scaling and intensity statistics." *Acta crystallographica. Section D, Biological crystallography*, vol. 67, no. Pt 4, pp. 282–92, Apr. 2011. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?ba5158

[281] P. Evans, "Scaling and assessment of data quality." *Acta crystallographica. Section D, Biological crystallography*, vol. 62, no. Pt 1, pp. 72–82, Jan. 2006. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?S0907444905036693

[282] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read, "Phaser crystallographic software." *Journal of applied crystallography*, vol. 40, no. Pt 4, pp. 658–674, Aug. 2007. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2483472&tool=pmcentrez&rendertype=abstract

[283] J. Navaza, "AMoRe : an automated package for molecular replacement," *Acta Crystallographica Section A Foundations of Crystallography*, vol. 50, no. 2, pp. 157–163, Mar. 1994. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?li0145

[284] A. Vagin and A. Teplyakov, "MOLREP : an Automated Program for Molecular Replacement," *Journal of Applied Crystallography*, vol. 30, no. 6, pp. 1022–1025, Dec. 1997. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?li5010

[285] A. J. McCoy, "Liking likelihood." *Acta crystallographica. Section D, Biological crystallography*, vol. 60, no. Pt 12 Pt 1, pp. 2169–83, Dec. 2004. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?S0907444904016038

[286] P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, "Features and development of Coot." *Acta crystallographica. Section D, Biological crystallography*, vol. 66, no. Pt 4, pp. 486–501, Apr. 2010. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?ba5144

[287] K. Cowtan, "The Buccaneer software for automated model building. 1. Tracing protein chains." *Acta crystallographica. Section D, Biological crystallography*, vol. 62, no. Pt 9, pp. 1002–11, Sep. 2006. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?S0907444906022116

[288] G. N. Murshudov, A. A. Vagin, and E. J. Dodson, "Refinement of macromolecular structures by the maximum-likelihood method." *Acta crystallographica. Section D, Biological crystallography*, vol. 53, no. Pt 3, pp. 240–55, May 1997. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?li0235

[289] G. Bricogne, E. Blanc, M. Brandl, C. Flensburg, P. Keller, W. Paciorek, P. Roversi, A. Sharff, O. S. Smart, C. Vonrhein, and T. O. Womack, "Buster version 2.10.0," *Global Phasing Ltd.*, 2011.

[290] G. J. Kleywegt and T. A. Jones, "Homo Crystallographicus - Quo Vadis?" *Structure*, vol. 10, no. 4, pp. 465–472, Apr. 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0969212602007438

[291] K. T. Simons, C. Kooperberg, E. Huang, and D. Baker, "Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions." *Journal of molecular biology*, vol. 268, no. 1, pp. 209–25, Apr. 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022283697909591

[292] S. J. Fleishman, A. Leaver-Fay, J. E. Corn, E.-M. Strauch, S. D. Khare, N. Koga, J. Ashworth, P. Murphy, F. Richter, G. Lemmon, J. Meiler, and D. Baker, "RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite." *PloS one*, vol. 6, no. 6, p. e20161, Jan. 2011. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3123292&tool=pmcentrez&rendertype=abstract

[293] W. Sheffler and D. Baker, "RosettaHoles: rapid assessment of protein core packing for structure prediction, refinement, design, and validation." *Protein science : a publication of the Protein Society*, vol. 18, no. 1, pp. 229–39, Jan. 2009. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2708028&tool=pmcentrez&rendertype=abstract

[294] V. K. Bhasin and W. Trager, "Gametocyte-forming and non-gametocyte-forming clones of Plasmodium falciparum." *The American journal of tropical medicine and hygiene*, vol. 33, no. 4, pp. 534–7, Jul. 1984. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/6383092

[295] J. Mu, P. Awadalla, J. Duan, K. M. McGee, D. A. Joy, G. A. T. McVean, and X.-z. Su, "Recombination hotspots and population structure in Plasmodium falciparum." *PLoS biology*, vol. 3, no. 10, p. e335, Oct. 2005. [Online]. Available: http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0030335

[296] D. E. Neafsey, S. F. Schaffner, S. K. Volkman, D. Park, P. Montgomery, D. A. Milner, A. Lukens, D. Rosen, R. Daniels, N. Houde, J. F. Cortese, E. Tyndall, C. Gates, N. Stange-Thomann, O. Sarr, D. Ndiaye, O. Ndir, S. Mboup, M. U. Ferreira, S. d. L. Moraes, A. P. Dash, C. E. Chitnis, R. C. Wiegand, D. L. Hartl, B. W. Birren, E. S. Lander, P. C. Sabeti, and D. F. Wirth, "Genome-wide SNP genotyping highlights the role of natural selection in Plasmodium falciparum population divergence." *Genome biology*, vol. 9, no. 12, p. R171, Jan. 2008. [Online]. Available: http://genomebiology.com/2008/9/12/R171

[297] L. Ellgaard and A. Helenius, "Quality control in the endoplasmic reticulum." *Nature reviews. Molecular cell biology*, vol. 4, no. 3, pp. 181–91, Mar. 2003. [Online]. Available: http://dx.doi.org/10.1038/nrm1052

[298] Y. Soejima, J. Lee, Y. Nagata, H. Mon, K. Iiyama, H. Kitano, M. Matsuyama, and T. Kusakabe, "Comparison of signal peptides for efficient protein secretion in the baculovirus-silkworm system," *Open Life Sciences*, vol. 8, no. 1, pp. 1–7, Jan. 2013. [Online]. Available: http://www.degruyter.com/view/j/biol.2013.8.issue-1/s11535-012-0112-6/s11535-012-0112-6.xml

[299] Y. K. Kim, H. S. Shin, N. Tomiya, Y. C. Lee, M. J. Betenbaugh, and H. J. Cha, "Production and N-glycan analysis of secreted human erythropoietin glycoprotein in stably transfected Drosophila S2 cells." *Biotechnology and bioengineering*, vol. 92, no. 4, pp. 452–61, Nov. 2005. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16025538

[300] H. P. Erickson, "Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy." *Biological procedures online*, vol. 11, no. 1, pp. 32–51, Jan. 2009. [Online]. Available: http://www.biologicalproceduresonline.com/content/11/1/32

[301] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton, "PROCHECK: a program to check the stereochemical quality of protein structures," *Journal of Applied Crystallography*, vol. 26, no. 2, pp. 283–291, Apr. 1993. [Online]. Available: http://scripts.iucr.org/cgi-bin/paper?gl0276

[302] P. Stumptner-Cuvelette and P. Benaroch, "Multiple roles of the invariant chain in MHC class II function," *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 1542, no. 1-3, pp. 1–13, Jan. 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167488901001665

[303] A. Claessens, W. L. Hamilton, M. Kekre, T. D. Otto, A. Faizullabhoy, J. C. Rayner, and D. Kwiatkowski, "Generation of antigenic diversity in Plasmodium falciparum by structured rearrangement of Var genes during mitosis."

PLoS genetics, vol. 10, no. 12, p. e1004812, Dec. 2014. [Online]. Available: http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004812

[304] T. A. Larsen, A. J. Olson, and D. S. Goodsell, "Morphology of proteinprotein interfaces," *Structure*, vol. 6, no. 4, pp. 421–427, Apr. 1998. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0969212698000446

[305] J. Chen, N. Sawyer, and L. Regan, "Protein-protein interactions: general trends in the relationship between binding affinity and interfacial buried surface area." *Protein science : a publication of the Protein Society*, vol. 22, no. 4, pp. 510–5, Apr. 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3610057&tool=pmcentrez&rendertype=abstract

[306] D. P. Doane and L. E. Seward, "Measuring Skewness: A Forgotten Statistic?." *Journal of Statistics Education*, vol. 19, no. 2, Nov. 2010. [Online]. Available: http://eric.ed.gov/?id=EJ945954

[307] G. E. Crooks, G. Hon, J.-M. Chandonia, and S. E. Brenner, "WebLogo: a sequence logo generator." *Genome research*, vol. 14, no. 6, pp. 1188–90, Jun. 2004. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=419797&tool=pmcentrez&rendertype=abstract

[308] L. A. Mirny and E. I. Shakhnovich, "Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function." *Journal of molecular biology*, vol. 291, no. 1, pp. 177–96, Aug. 1999.

[309] R. M. MacCallum, A. C. Martin, and J. M. Thornton, "Antibody-antigen interactions: contact analysis and binding site topography." *Journal of molecular biology*, vol. 262, no. 5, pp. 732–45, Oct. 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022283696905483

[310] R. J. S. Preston, A. Villegas-Mendez, Y.-H. Sun, J. Hermida, P. Simioni, H. Philippou, B. Dahlbäck, and D. A. Lane, "Selective modulation of protein C affinity for EPCR and phospholipids by Gla domain mutation." *The FEBS journal*, vol. 272, no. 1, pp. 97–108, Jan. 2005. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/15634335

[311] X. V. Yang, Y. Banerjee, J. A. Fernández, H. Deguchi, X. Xu, L. O. Mosnier, R. T. Urbanus, P. G. de Groot, T. C. White-Adams, O. J. T. McCarty, and J. H. Griffin, "Activated protein C ligation of ApoER2 (LRP8) causes Dab1-dependent signaling in U937 cells." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 1, pp. 274–9, Jan. 2009. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2629184&tool=pmcentrez&rendertype=abstract

[312] A. R. Fersht, J.-P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M. M. Y. Waye, and G. Winter, "Hydrogen bonding and biological specificity analysed by protein engineering," *Nature*, vol. 314, no. 6008, pp. 235–238, Mar. 1985. [Online]. Available: http://dx.doi.org/10.1038/314235a0

[313] A. Heddini, F. Pettersson, O. Kai, J. Shafi, J. Obiero, Q. Chen, A. Barragan, M. Wahlgren, and K. Marsh, "Fresh isolates from children with severe Plasmodium falciparum malaria bind to multiple receptors." *Infection and immunity*, vol. 69,

no. 9, pp. 5849–56, Sep. 2001. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=98703&tool=pmcentrez&rendertype=abstract

[314] V. Fernandez, C. J. Treutiger, G. B. Nash, and M. Wahlgren, "Multiple adhesive phenotypes linked to rosetting binding of erythrocytes in Plasmodium falciparum malaria." *Infection and immunity*, vol. 66, no. 6, pp. 2969–75, Jun. 1998. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=108296&tool=pmcentrez&rendertype=abstract

[315] A. Brown, L. Turner, S. Christoffersen, K. A. Andrews, T. Szestak, Y. Zhao, S. Larsen, A. G. Craig, and M. K. Higgins, "Molecular architecture of a complex between an adhesion protein from the malaria parasite and intracellular adhesion molecule 1." *The Journal of biological chemistry*, vol. 288, no. 8, pp. 5992–6003, Feb. 2013. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3581401&tool=pmcentrez&rendertype=abstract

[316] T. M. Clausen, S. Christoffersen, M. Dahlbäck, A. E. Langkilde, K. E. Jensen, M. Resende, M. O. Agerbæ k, D. Andersen, B. Berisha, S. B. Ditlev, V. V. Pinto, M. a. Nielsen, T. G. Theander, S. Larsen, and A. Salanti, "Structural and Functional Insight into How the Plasmodium falciparum VAR2CSA Protein Mediates Binding to Chondroitin Sulfate A in Placental Malaria." *The Journal of biological chemistry*, vol. 287, no. 28, pp. 23 332–45, Jul. 2012. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22570492

[317] G.-Y. Chuang, J. C. Boyington, M. G. Joyce, J. Zhu, G. J. Nabel, P. D. Kwong, and I. Georgiev, "Computational prediction of N-linked glycosylation incorporating structural properties and patterns." *Bioinformatics (Oxford, England)*, vol. 28, no. 17, pp. 2249–55, Sep. 2012. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3426846&tool=pmcentrez&rendertype=abstract

[318] R. Sowdhamini, N. Srinivasan, B. Shoichet, D. V. Santi, C. Ramakrishnan, and P. Balaram, "Stereochemical modeling of disulfide bridges. Criteria for introduction into proteins by site-directed mutagenesis." *Protein engineering*, vol. 3, no. 2, pp. 95–103, Nov. 1989. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/2594728

[319] C. K. Y. Lau, L. Turner, J. S. Jespersen, E. D. Lowe, B. Petersen, C. W. Wang, J. E. V. Petersen, J. Lusingu, T. G. Theander, T. Lavstsen, and M. K. Higgins, "Structural Conservation Despite Huge Sequence Diversity Allows EPCR Binding by the PfEMP1 Family Implicated in Severe Childhood Malaria." *Cell host & microbe*, vol. 17, no. 1, pp. 118–129, Dec. 2014. [Online]. Available: http://www.cell.com/article/S1931312814004235/fulltext

[320] A.-J. Petrescu, A.-L. Milac, S. M. Petrescu, R. A. Dwek, and M. R. Wormald, "Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding." *Glycobiology*, vol. 14, no. 2, pp. 103–14, Feb. 2004. [Online]. Available: http://glycob.oxfordjournals.org/content/14/2/103.long

[321] S.-C. Lin, Y.-F. Lin, P. Chong, and S.-C. Wu, "Broader neutralizing antibodies against H5N1 viruses using prime-boost immunization of hyperglycosylated hemagglutinin DNA and virus-like particles." *PloS one*, vol. 7, no. 6, p. e39075, Jan. 2012. [Online]. Available: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0039075#pone.0039075-Garrity1

[322] D. E. Kim, B. Blum, P. Bradley, and D. Baker, "Sampling bottlenecks in de novo protein structure prediction." *Journal of molecular biology*, vol. 393, no. 1, pp. 249–60, Oct. 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022283609009309

[323] P. Olin, F. Rasmussen, L. Gustafsson, H. O. Hallander, and H. Heijbel, "Randomised controlled trial of two-component, three-component, and five-component acellular pertussis vaccines compared with whole-cell pertussis vaccine," *The Lancet*, vol. 350, no. 9091, pp. 1569–1577, Nov. 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140673697065082

[324] M. M. Giuliani, J. Adu-Bobie, M. Comanducci, B. Aricò, S. Savino, L. Santini, B. Brunelli, S. Bambini, A. Biolchi, B. Capecchi, E. Cartocci, L. Ciucchi, F. Di Marcello, F. Ferlicca, B. Galli, E. Luzzi, V. Masignani, D. Serruto, D. Veggi, M. Contorni, M. Morandi, A. Bartalesi, V. Cinotti, D. Mannucci, F. Titta, E. Ovidi, J. A. Welsch, D. Granoff, R. Rappuoli, and M. Pizza, "A universal vaccine for serogroup B meningococcus." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 29, pp. 10 834–9, Jul. 2006. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2047628&tool=pmcentrez&rendertype=abstract

Appendix A

Primer sequences

CIDRα1 variant	Vector	Direction	Sequence
HB3var03	pEt15bcon	f	GCGGATCCGGTGATGCAAAAAATACTGAAATTAATG
		r	GCGCTAGCTTAACATCCTTCGTTTGTATTATTATCTTTGC
IT4var06	pEt15bcon	f	GCGGATCCGAAATTTTACAAAGTGAGGAAACTAC
		r	GCGCTAGCTTAGCAATTTTTATCTTCTTTTAATGAATTATTG
IT4var07	pEt15bcon	f	GCGGATCCGAACCTGCACCGGATGTGAAGACTAC
		r	GCCATATGTTAACATGCTTCGTTTGTATTATTGTCTTTAC
IT4var19	pEt15bcon	f	GCGGATCCAAAATTCTAAAATATGAGACGCCTAC
		r	GCGCTAGCTTAGCAATTTTCAACTGCTAATGAATTATTG
IT4var20	pEt15bcon	f	GCGGATCCCAAAGTGAGGAACCTACTGAAATTC
		r	GCGCTAGCTTAACATGATTCGTTTGAATTATTGTCTC
PFD0020c	pEt15bcon	f	GCGGATCCAAAATTCTAGAAAGTGAGACGCC
		r	GCGCTAGCTTAACATGGTTTGATTGAATTATTGTC

Table A.1: Primer pairs for cloning CIDR domains from genomic DNA.

EPCR construct	Vector	Direction	Sequence
EPCRL	pEt/pAc	f	CCGGATCCAGCCAAGACGCCTCAGATG
		r	GGGCTAGCCTACGAAGTGTAGGAGCGGC
EPCRS	pEt/pAc	f	AAGGATCCTCAGATGGCCTCCAAAGAC
		r	GGGCTAGCCTATTCCGCGGAAATATGTTTCTG

Table A.2: Primer pairs for cloning CIDRα1 variants from genomic DNA.

Mutant	Sequence
F656A	GATAGTTTTTTTGCCCAAGTTATATA
	TATATAACTTGGGCAAAAAAACTATC
F656R	GATAGTTTTTTTCGCCAAGTTATATA
	TATATAACTTGGCGAAAAAAACTATC
F656Y	GATAGTTTTTTTTACCAAGTTATATA
	TATATAACTTGGTAAAAAAAACTATC
F656V	GATAGTTTTTTTGTCCAAGTTATATA
	TATATAACTTGGACAAAAAAACTATC
F655L	CTTTTTGATAGTTTTCTGTTCCAAGTTATATAT
	ATATATAACTTGGAACAGAAAACTATCAAAAAG
F655Y	TTTGATAGTTTTTATTTCCAAGTTATA
	TATAACTTGGAAATAAAAACTATCAAA
Q657K	AGTTTTTTTTTCAAAGTTATATATA
	TATATATAACTTTGAAAAAAAAACT
Q657R	AGTTTTTTTTTCCGTGTTATATATAAG
	CTTATATATAACACGGAAAAAAAAACT
F651A	AATATTAATGATCTTGCCGATAGTTTTTTTTTC
	GAAAAAAAAACTATCGGCAAGATCATTAATATT
V658A	TTTTTTTTCCAAGCCATATATAAGTTT
	TATATATAACTTTGAAAAAAAAACT
W669S	GAAGGAGAAGCAAAAAGCAATGAACTTAAAGAA
	TTCTTTAAGTTCATTGCTTTTTGCTTCTCCTTC
S653F	TGATCTTTTTGATTTCTTTTTTTTCCAAG
	CTTGGAAAAAAAAGAAATCAAAAAGATCA
S653FQ657R	TGATCTTTTTGATTTCTTTTTTTTCCGTG
	CACGGAAAAAAAAGAAATCAAAAAGATCA
D576A	GATGAATTTTTTGCGTTTTGGGTTAGA
	TCTAACCCAAAACGCAAAAAATTCATC
K642A	CACGATATACCGAAAGCGTATTATCTTAATATT
	AATATTAAGATAATACGCTTTCGGTATATCGTG

Table A.3: Primer pairs for HB3var03 CIDR mutants.

Cysteine bond	Mutation	**Sequence**
Cys1	I10C	CAGCAGAAAAAAAACTGCGAGGATGATCTGGAC
		GTCCAGATCATCCTCGCAGTTTTTTTTCTGCTG
	L74C	GATCGTGTGAACGACTGCTTTGATAGCAACTTT
		AAAGTTGCTATCAAAGCAGTCGTTCACACGATC
Cys2	L53C	GAATGGGAAAAACGTTGCAAAGAACTGGAAGAAC
		GTTCTTCCAGTTCTTTGCAACGTTTTTCCCATTC
	L115C	CGCATCAAAGAAAAATGCGAAGAGATCAAAAAAC
		GTTTTTTGATCTCTTCGCATTTTTCTTTGATGCG
Cys3	V71C	GAAGCAAAAGATCGTTGCAACGACCTGTTTGATAG
		CTATCAAACAGGTCGTTGCAACGATCTTTTGCTTC
	A100C	GAAAAAGACCGTTGCGAGAAAGAAATCGAAGAATG
		CATTCTTCGATTTCTTTCTCGCAACGGTCTTTTTC

Table A.4: Primer pairs for the bundle immunogen 469 cysteine mutants.

Appendix B

Custom python scripts

B.1 Rosetta arguments file to fold a scaffold around an epitope

```
-run:protocol broker
-run:shuffle
-in:file:fullatom
-abinitio:close_loops
## Loop building options
-short_frag_cycles 1
-scored_frag_cycles 1
-non_ideal_loop_closing
-alternative_closure_protocol
-fast_loops:window_accept_ratio .01
-fast_loops:nr_scored_sampling_passes 4
-fast_loops:min_breakout_good_loops 5
-fast_loops:min_breakout_fast_loops 80
-fast_loops:min_fast_loops 3
-fast_loops:min_good_loops 0
-fast_loops:nr_scored_fragments 20
-fast_loops:vdw_delta 0.5
-fast_loops:give_up 1000
-random_grow_loops_by  4
-idealize_after_loop_close
-out:file:silent_struct_type binary
-bGDT
-increase_cycles 1
-jumps:ramp_chainbreaks
-sep_switch_accelerate 0.8
-skip_convergence_check
-overlap_chainbreak
-out::shuffle_nstruct 10000
-mute core.chemical
-mute core.conformation
-fail_on_bad_hbond false
-run:write_failures
-cm:min_loop_size 2
-relax_failures
```

```
-abinitio:relax
-relax::fast
-relax::default_repeats 12
-final_clean_fastrelax
```

B.2 Python script for generating C_α-C_α constraints from a PDB file

```
from numpy import array
#input for file has to be pdb.
input_file = input("Please enter input pdb file, Clinton: ")
stand_dev = input("Please enter required standard deviation (usual is
    2): ")
output_file = input("What shall I call the output?: ")
#CAs will go in here.
CAs = {}
#this reads in the pdb and splits it into its components
for line in open(input_file):
    list = line.split()
    id = list[0]
    if id == 'ATOM':
        atom_no = int(list[1])
        #atom_type is CA, N, CO etc.
        atom_type = list[2]
        resi_type = list[3]
        chain_id = list[4]
        resi_no = int(list[5])
        coord_x = float(list[6])
        coord_y = float(list[7])
        coord_z = float(list[8])
        #occupancy = float(list[9])
        #B_fact = float(list[10])
        if atom_type == 'CA':
            CAs[resi_no] = array((coord_x, coord_y, coord_z))
f = open((str(output_file)+".cst"), "w")
for key, value in CAs.items():
    for key2,value2 in CAs.items():
        if key2>key:
            subtract_array = value2-value
            distance = ((subtract_array[0])**2 + (subtract_array[1])**2
                + (subtract_array[2])**2)**0.5
            f.write("AtomPair " + "CA " + str(key)+ " " + "CA " + str(
                key2)
                    + " " + "GAUSSIANFUNC " + str(distance) + " " +
                        stand_dev + "\n")
f.close()
```

B.3 RosettaScripts custom script for sequence design

```
<dock_design>
  <SCOREFXNS>
    <SFXN1 weights="score12"/>
  </SCOREFXNS>
```

```
<TASKOPERATIONS>
  Mutates residues to residues found in the appropriate places core,
      surface, etc.
  <LayerDesign name="layer" layer="core_boundary_surface"/>
  Restricts the residues to repacking only Chose outwards facing arm
      residues and kinks
  <RestrictResiduesToRepacking name="RRTR" residues
      =72,73,75,76,77,79,80,81,82,84,85,93 />
</TASKOPERATIONS>
<FILTERS>
  Defining filter for rmsd to native structure CHECK NUMBERS
  <Rmsd name="rmsd" chains="A" threshold=0.7 superimpose=0>
    <span begin_res_num="72" end_res_num="85"/>
    <span begin_res_num="89" end_res_num="105"/>
  </Rmsd>
  Filter for backbone RMSD to two phenylalanines
  <Rmsd name="rmsd2" chains="A" threshold=0.4 superimpose=0>
    <span begin_res_num="79" end_res_num="80"/>
  </Rmsd>
  Filter for backbone RMSD to native struct strict
  <Rmsd name="rmsd3" chains="A" threshold=0.6 superimpose=0>
    <span begin_res_num="72" end_res_num="85"/>
    <span begin_res_num="89" end_res_num="105"/>
  </Rmsd>
  Defining score filter
  <ScoreType name="score" scorefxn="score12" score_type="total_score"
      threshold=-280.0/>
  Packstat from RosettaHoles
  <PackStat name="pstat" threshold=0.5 chain=0 repeats=1/>
  <PackStat name="pstat2" threshold=0.55 chain=0 repeats=1/>
</FILTERS>
<MOVERS>
  That flexible backbone design thing
  <FlxbbDesign name="flxbb" ncycles=4 sfxn_design="SFXN1" sfxn_relax
      ="SFXN1" clear_all_residues=0 task_operations="layer,RRTR"/>
  Superimposition mover for RMSD calculation vs binding arm
  <Superimpose name="super" ref_start=72 ref_end=105 target_start=72
      target_end=105 ref_pose="newnative.pdb"/>
  To clear constraints
  <ClearConstraintsMover name="rmconstraints" />
  Add constraints
  <ConstraintSetMover name="addconstraints" cst_file="constraints.cst
      "/>
  Fast relax
  <FastRelax name="relax" task_operations="RRTR"/>
</MOVERS>
<PROTOCOLS>
  <Add mover = addconstraints/>
  <Add mover = flxbb/>
  <Add filter = pstat/>
  <Add filter = score/>
  <Add mover = rmconstraints/>
  <Add mover = relax/>
  <Add filter = pstat2/>
</PROTOCOLS>
<OUTPUT />
```

```
</dock_design>
```

Appendix C

Publications arising from this thesis

Lau, C.K.Y., Turner, L., Jespersen, J.S., Lowe, E.D., Petersen, B., Wang, C.W., Petersen, J.E.V., Lusingu, J., Theander, T.G., Lavstsen, T., Higgins, M.K.. Structural conservation despite huge sequence diversity allows EPCR binding by the *Pf*EMP1 family implicated in severe childhood malaria . *Cell Host & Microbe*. 2015 Jan 14;17(1):118-29. doi: 10.1016/j.chom.2014.11.007. Epub 2014 Dec 4.