CHAPTER 5

Systematic screening for extracellular receptor-ligand interactions involved in erythrocyte recognition and invasion by *P. falciparum*

5.1 Summary and Aims

Invasion of erythrocytes by *Plasmodium falciparum* merozoites, is an essential step in parasite's lifecycle and has therefore long been thought of as a desirable target for the development of therapeutics, particularly a vaccine. However, the development of a highly efficacious blood stage malaria vaccine has been made difficult by the highly polymorphic nature of merozoite ligands as well as by functional redundancy of the receptor-ligand pairs involved in the invasion process. A more detailed molecular understanding of invasion could lead to identification of cell surface receptor-ligand interactions, which are essential for erythrocyte invasion and therefore, promising intervention targets can be revealed. However, such studies have been hampered by the difficulties in recombinant expression of *Plasmodium* spp. proteins and by the technical challenges in systematically identifying interacting partners for cell surface proteins (Birkholtz *et al.*, 2008; Bartholdson *et al.*, 2013).

In previous work from our laboratory, 53 *P. falciparum* merozoite cell surface and secreted proteins were expressed in a biochemically active recombinant form, by using the HEK293E protein expression system (Crosnier et al, unpublished). By using AVEXIS (Bushell *et al.*, 2008), a system designed to detect low affinity protein-protein interactions that is discussed in more depth below, the *P. falciparum* recombinant merozoite protein array was systematically screened against a protein library consisting of erythrocyte cell surface receptors which was also expressed recombinantly in our laboratory (Crosnier and Bartholdson, unpublished data). The screen identified BSG and Semaphorin7A as the erythrocyte receptors for merozoite RH5 and MTRAP respectively (Crosnier *et al.*, 2011; Bartholdson *et al.*, 2012). Importantly, the interaction between RH5 and BSG is the only one known so far that is essential and universally required for erythrocyte invasion (Crosnier *et al.*, 2011).

In this Chapter I, aimed to expand the previous work from our laboratory, and identify novel receptor-ligand pairs involved in erythrocyte invasion. For this purpose, I expanded the already existing *P. falciparum* merozoite recombinant protein library by 26 proteins, using transcription microarray data and information available in the published literature. Of 26 new *P. falciparum* merozoite proteins targeted, 21 were recombinantly expressed at usable amounts. Biochemical analysis provided evidence that the recombinant merozoite ligands were biochemically active and correctly folded. AVEXIS was then used to systematically screen these new

members of the *P. falciparum* recombinant protein library, against an equivalent library consisting of erythrocyte receptors (Bushell *et al.*, 2008). This screen identified one putative interaction (PF13_0125 – P4HB). Further characterisation of the identified interaction provided inconclusive results, and more experiments are required to confirm its validity. The recombinant *P. falciparum* merozoite proteins reported in this project should prove a useful tool for the deeper understanding of erythrocyte invasion, and *P. falciparum* biology in general.

5.2 Introduction

5.2.1 Systems for recombinant expression of *Plasmodium* spp. proteins

Expressing *Plasmodium* proteins in heterologous systems has been a technically challenging problem (Birkholtz *et al.*, 2008; Fernández-Robledo and Vasta, 2010). Although the causal reasons why *Plasmodium* proteins are difficult to express in recombinant expression systems are not known, several protein characteristics such as high molecular mass (> 56 kDa) and the presence of export motifs and atypical signal peptide sequences were found to make soluble expression of recombinant *Plasmodium* proteins, in commonly used heterologous expression systems, challenging (Birkholtz *et al.*, 2008). Moreover, the problematic recombinant expression of *Plasmodium* proteins has been attributed to the unusually high (~80%) A+T content of parasite genes which can result in the prevalence of repetitive amino acid stretches (Tsuboi *et al.*, 2008), and the use of codons that are not frequently used by organisms selected for heterologous expression affecting translation efficiency.

Despite, or because of, these challenges, recombinant expression of *Plasmodium* proteins has been tested in a number of expression systems. The *E. coli* bacterial expression system is the most popular (Fernández-Robledo and Vasta, 2010). While *E. coli* expression is inexpensive system and can produce high yields of recombinant proteins, an important drawback is that recombinant proteins lack post-translational modifications and therefore, do not represent eukaryotic proteins as they found *in* vivo (Birkholtz *et al.*, 2008; Fernández-Robledo and Vasta, 2010). Moreover, recombinant proteins are expressed in an insoluble form and, often, sequestered in cytoplasmic inclusion bodies (Birkholtz *et al.*, 2008). Such recombinant protein aggregates require solubilisation followed by refolding

(Rodriguez *et al.*, 2008; Baum *et al.*, 2009; Chen *et al.*, 2011), and the overall success rate for obtaining soluble, immunogenic recombinant products remains low and variable depending on the individual protein being expressed (Fernández-Robledo and Vasta, 2010).

In other implementations of *E. coli* expression systems, the recombinant protein is expressed in a soluble form and is secreted into the periplasmic space. The oxidative environment of the periplasmic space facilitates the formation of disulfide bonds and therefore, increases the chances of obtaining a correctly folded, recombinant protein (De Marco, 2009). Nevertheless, recombinant expression of eukarytotic proteins (including *Plasmodium* proteins) in this way, often requires the co-expression of the protein of interest with other proteins which facilitate secretion to the periplasmic space, or mediate correct protein folding by controlling the disulfide bond formation (Outchkourov *et al.*, 2008; de Marco, 2009). A good example of a recombinant *Plasmodium* protein purified from the periplasmic space is *Pf*s48/45 which is a major candidate for the development of a malaria transmission blocking vaccine (section1.13.1 ;Outchkourov and Roeffen, 2008)

Yeast-based systems have also been used for recombinant expression of *Plasmodium* proteins (Ballou *et al.*, 2004). *Saccharomyces cerevisiae* and *Pichia pastoris* are the most commonly used yeast species for heterologous protein expression of *Plasmodium* proteins (Birkholtz *et al.*, 2008). The advantage of yeast over bacterial expression systems is that the recombinant proteins are expressed in their native conformation and can be secreted in a soluble form in the culture medium, simplifying their subsequent purification (Birkholtz *et al.*, 2008). Similar advantages are shared by baculovirus-insect cell and HEK293 expression systems, both eukaryotic expression systems that have been used for the heterologous expression of *Plasmodium* proteins (Li *et al.*, 2002; Ballou *et al.*, 2004; Fernández-Robledo and Vasta, 2010; Crosnier *et al.*, 2011; Douglas *et al.*, 2011; Williams *et al.*, 2012; Bustamante *et al.*, 2013). One common problem that arises with expression of *Plasmodium* proteins in eukaryotic expression systems is aberrant N-linked and O-linked glycosylation. *Plasmodium* parasites lack the enzymatic machinery necessary for glycosylation, but their proteins do contain canonical N-linked glycosylation sites which can become occupied during expression in heterologous eukaryotic cells (Dieckmann-Schuppert *et al.*, 1992). This problem can easily be circumvented by

mutating the N-linked glycosylation sites found in the transgene coding sequence, or by inclusion of N-glycosylation inhibitors in the culture medium (Birkholtz *et al.*, 2008; Fernández-Robledo and Vasta, 2010)

The protozoa *Dictyostelium discoideum* and *Tetrahymena thermophile* are another example of protein expression systems, employed for recombinant expression of *Plasmodium* proteins. They originally became an attractive system for heterologous expression of *Plasmodium* proteins because of the unusual A+T bias of their gene coding sequence, a characteristic they share with *P. falciparum* genome (Birkholtz *et al.*, 2008; Fernández-Robledo and Vasta, 2010). However, while multiple expression systems have been trialled, it is fair to say that until systematic application of HEK293E expression, no single approach has been consistently successful.

5.2.2 Studying extracellular protein-protein interactions between merozoite and erythrocytes- the AVEXIS assay.

Cell surface protein-protein interactions between *Plasmodium* merozoite and erythrocytes have been traditionally studied by using naturally occurring erythrocyte variants which either lack, or express a mutant form of a specific cell surface receptor (Bei and Duraisingh, 2012; Bartholdson *et al.*, 2013). Such erythrocytes variants have been used to derive evidence about the binding specificity of different parasite ligands or to determine the impact of specific mutations on parasite invasion efficiency (Bei and Duraisingh, 2012; Bartholdson *et al.*, 2013). For example, the binding specificity of EBA140 for GYPC was established by using erythrocytes lacking expression of GYPC, or lacking specific exons in the *GYPC* gene (Lobo *et al.*, 2003; Maier *et al.*, 2003; Bartholdson *et al.*, 2013). Likewise, GYPB-deficient red blood cells were utilised to show that GYPB is the receptor for EBL1 (Mayer *et al.*, 2009; Bartholdson *et al.*, 2013). The availability, however, of erythrocytes expressing particular receptor variants can be limited. This fact, combined with the relatively short shelf life of cellular biopsies and the possibility that the absence of some receptors might be incompatible with normal erythrocyte function, make this approach generally impractical for systematically identifying new ligand-receptor pairs (Bartholdson *et al.*, 2013).

Information about the binding characteristics of parasite ligands to erythrocytes can also be obtained by testing the sensitivity of the binding of parasite proteins to enzyme-treated erythrocytes (Bartholdson *et al.*, 2013). For example, initial studies with EBA-175 demonstrated that EBA-175 binds to a neuraminidase and trypsin sensitive receptor (Camus and Hadley, 1985), which was later identified to be GYPA (Sim *et al.*, 1994; Bartholdson *et al.*, 2013). Nevertheless, using binding sensitivity to erythrocytes treated with broad substrate specificity enzymes (e.g. trypsin and chymotrypsin), is only rarely helpful because a large number of receptors are sensitive to this treatment.

Cell-surface interactions between erythrocytes and merozoites can also be examined by using recombinant proteins. Besides the difficulties in recombinant expression of *Plasmodium* proteins (section 5.2.1), another major challenge of this approach has been the limited availability of high-throughput assays for the systematic identification of cell-surface protein-protein interaction which are normally highly transient with half-lives of just a few seconds (Wright, 2009; Bartholdson *et al.*, 2013). High throughput assays for the identification of protein-protein interactions, like the yeast-two-hybrid (Young, 1998) and TAP-tagging methods followed by mass spectrometry (Rigaut *et al.*, 1999; Puig *et al.*, 2001), are generally regarded as unsuitable to detect transient interactions between extracellular proteins (Bushell *et al.*, 2008; Wright, 2009): structurally-important post-translational modifications such as disulfide bonds and glycans are not generally added to proteins within the yeast nucleus and the stringent washing steps of biochemical purifications do not allow the detection of transient interactions (Wright, 2009).

To address the problems associated with identifying novel low affinity extracellular protein interactions (including those involved in erythrocyte invasion), an assay termed AVEXIS (for AVidity-based EXtracellular Interaction Screen) was developed (Bushell *et al.*, 2008) (Fig. 5.1). AVEXIS was designed specifically to detect highly transient extracellular protein-protein interactions. The entire ectodomains of cell surface or secreted proteins are recombinantly expressed as soluble fragments in HEK293 cells. Each protein is expressed as both monomeric biotinylated bait and β-lactamase tagged pentameric prey. The pentamerization is required to increase the overall interaction avidity, allowing the assay to detect very transient interactions (with halflives<0.1s) that are a common feature of this class of

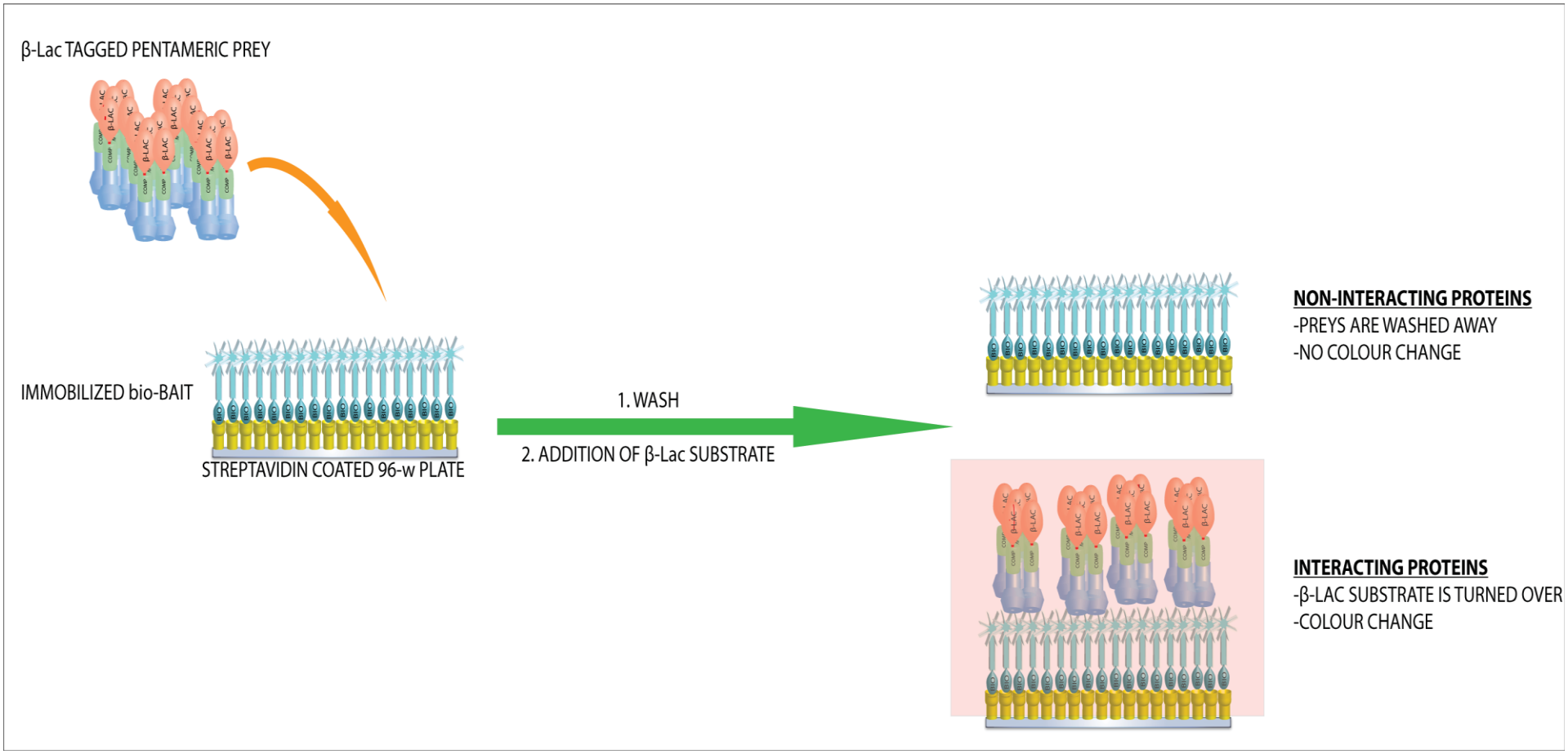

Figure 5.1 The AVEXIS assay. Recombinant monobiotinylated bait ectodomains are captured in an oriented manner on a streptavidin-coated micro well plate and systematically screened against β-lactamase tagged pentameric prey ectodomains. Interactions between baits and preys are detected by colourimetric enzymatic turnover of nitrocefin (β-lactamase substrate).

protein interactions. Baits are arrayed by orientated capture on streptavidin-coated plates, and probed for direct interactions with the enzyme-tagged prey proteins (Bushell *et al.*, 2008). By using this assay, BSG and Semaphorin7A were identified as the erythrocyte receptors for merozoite RH5 and MTRAP, respectively (Crosnier *et al.*, 2011; Bartholdson *et al.*, 2012). Notably, the RH5-BSG interaction is the only one known so far that is essential and universally required for erythrocyte invasion (Crosnier *et al.*, 2011). Given the complexity of erythrocyte invasion, it is clear that many additional interactions remain to be identified.

5.3 Results

To identify novel cell surface protein-protein interactions implicated in erythrocyte invasion, I sought to expand the *P. falciparum* merozoite recombinant protein library (section 5.1) which is available in our laboratory, and employ AVEXIS to *in vitro* screen the new library members against an erythrocyte cell surface receptor repertoire. To use the AVEXIs assay for the screen, it was necessary to recombinantly express both the merozoite and human erythrocyte protein libraries. The expression constructs encoding the erythrocyte cell surface receptors were available in our laboratory from previous experiments (Crosnier *et al.*, 2011).

5.3.1 Compilation of a list of candidate *P. falciparum* merozoite cell surface or secreted proteins

To expand the existing *P. falciparum* merozoite recombinant protein library, I compiled a list of merozoite cell surface or secreted proteins with possible roles in erythrocyte invasion. To identify candidate merozoite cell surface or secreted proteins, I took advantage of publicly available transcription microarray data of *P.falciparum* intra-erythrocytic stages (Bozdech *et al.*, 2003). Bioinformatic analysis of these data, demonstrated that several known merozoite ligands (e.g RH5, AMA1, EBA140, EBA175; Cowman *et al.*, 2012; Bartholdson *et al.*, 2013) follow a similar transcriptional expression pattern which passes through a minimum 20±6 hours post invasion and peak at about 42±6 hours after invasion. Therefore, the latter time-windows were set as criteria for short listing 405 candidate merozoite cell surface proteins (Bozdech *et al.*, 2003). Of 405 shortlisted proteins, 217 carried a signal peptide and/or a single transmembrane domain, as predicted by using the SignalP (http://www.cbs.dtu.dk/services/SignalP/) and TMHMM prediction servers

(http://www.cbs.dtu.dk/services/TMHMM/) respectively, suggesting they are likely cell surface or secreted proteins and therefore considered as candidate proteins. Multipass membrane proteins were excluded because such proteins are unlikely to be expressed in a soluble, secreted form by using the HEK293 expression system.

The 217 candidate proteins included 40 out of the 53 members of the already existing *P. falciparum* merozoite recombinant protein library (Crosnier *et al.*, 2013). This is translated to 75% coverage, indicative of the validity of my methodology. The 40 merozoite proteins which were already available in our laboratory were excluded for further analysis, reducing the number of candidate proteins to 177.

Because I was focused on identifying merozoite cell surface and secreted proteins, all candidate proteins were subject to gene ontology analysis by using the relevant tool in PlasmoDB (www.plasmoDB.org), and to protein domain mapping analysis by using Pfam (www.pfam.sanger.ac.uk). Out of 177 proteins, 121 were removed, because their predicted gene ontology (e.g. involvement in lipid metabolism or nuclear localization) or predicted protein domains (e.g. RNA or DNA binding motif) suggested that they are unlikely to participate in cell surface protein interactions. Another 42 proteins were excluded due to their large size (>1400 amino acids) which are not only prohibitively expensive to codon optimise by gene synthesis, but often expressed at levels below that required for performing additional experiments. After this analysis, 14 candidate merozoite cell surface proteins remained on our shortlist. Seven putative merozoite cell surface proteins were also added onto the candidate list, based on previously published reports in the literature (Haase *et al.*, 2008; Hu *et al.*, 2010) (Fig. 5.2).

The availability of the complete set of proteins belonging to a specific *Plasmodium* paralogous protein family, would allow direct comparison between individual members of the family in the case that one protein-member was found to interact with an erythrocyte cell surface receptor. Several members of the MSP3 (Burgess *et al.*, 2005; Singh *et al.*, 2009) and MSP-7 like (Kadekoppala *et al.*, 2008, 2010; Kadekoppala and Holder, 2010) *P. falciparum* protein families were already available in the laboratory from the previously developed *P. falciparum* merozoite recombinant protein library (Crosnier *et al.*, 2013). However, not all the members of these two families were included in this initial library. The members of

No	Name	Accession Code	Type	Region to be expressed
1	rhoptry-associated membrane antigen (RAMA)	MAL7P1.208	GPI	Y17–S840
2	prohibitin, putative	PF08_0006	secreted	L20–F272
3	conserved Plasmodium protein, unknown function	PF10_0166	secreted	Y25–E310
4	GLURP	PF10_0344	secreted	K24–I1233
5	MSP3.5	PF10_0350	secreted	A20–F710
6	MSP3.6	PF10_0351	secreted	N22–P566
7	MSRP5	PF13_0191	secreted	N22–I459
8	MSRP4	MAL13P1.173	secreted	D22–Q309
9	MSP8	PFE0120c	GPI	E26–S576
10	conserved Plasmodium protein, unknown function	PF13_0125	secreted	N20–S292
11	conserved Plasmodium protein, unknown function	PF14_0044	secreted	Q21–K290
12	merozoite-associated tryptophan-rich antigen, putative	PFA0135w	secreted	I25–K276
13	LCCL domain-containing protein	PFA0445w	secreted	K22–I1029
14	SERA1	PFB0360c	secreted	M1–V997
15	SERA2	PFB0355c	secreted	E23–V1105
16	SERA3	PFB0350c	secreted	T23–I930
17	SERA4	PFB0345c	secreted	S26–V962
18	SERA5	PFB0340c	secreted	T23–V997
19	SERA6	PFB0335c	secreted	N25–V1031
20	SERA7	PFB0330c	secreted	Q23–V946
21	SERA9	PFI0135c	secreted	E23–V932
22	conserved Plasmodium protein, unknown function	PFA0210c	secreted	Y24–D466
23	conserved Plasmodium protein, unknown function	PFB0475c	secreted	L23–D446
24	conserved Plasmodium protein, unknown function	PFD1130w	secreted	D29–E362
25	EBA-165 (corrected ORF)	PFD1155w	I	K22–S1340
26	RIPR	PFC1045c	secreted	I20-N1086

Figure 5.2 The merozoite cell surface proteins that were chosen for recombinant expression. Protein sequences from 3D7 reference *P.falciparum* strain were truncated to remove endogenous signal peptides and sequences corresponding to cytoplasmic and transmembrane domains. The region of each truncated protein that was targeted for recombinant expression is showed on the right column. Abbreviations: GPI, GPI-anchored; I, type I; sec., secreted. Proteins shortlisted from bioinformatic analysis of transcription microarray data are highlighted with grey colour. Proteins chosen from Hu *et al.*, 2010 or Haase *et al.*, 2008, are indicated in green and cyan respectively. The proteins chosen to enable having a complete set of proteins belonging to the MSP3, MSP7 like and SERA families of proteins are highlighted in yellow. EBA-165 was enlisted because it could potentially be useful for experiments that may help to the understanding of host tropism evolution of *P. falciparum.*

these protein families that were missing (MSP3.5, MSP3.6, MSRP4) were also added to our candidate protein list (Fig. 5.2). Similarly, SERA9 was selected because it would complete the set of proteins belonging to the SERA family of proteins (Fig. 5.2); the rest of the SERA family members were short listed from the bioinfomatic analysis of transcription microarray data or from the previously published studies (see above).

*Pf*EBA-165, which is encoded by a likely pseudogene in *Plasmodium falciparum* was also included in my list. Six *P. falciparum* strains investigated in earlier studies, all contained at least one frameshift mutation in the *Pf*eba165 gene that leads to premature termination of protein translation (Triglia *et al.*, 2001; Rayner *et al.*, 2004). The reference *P. falciparum* strain 3D7 contains two such frameshifts, and no *Pf*EBA-165 protein was detected in blood stage parasites (Triglia *et al.*, 2001). It is not currently known whether any naturally occurring *P. falciparum* parasites encode a full-length *Pf*EBA165. Nevertheless, *eba-165* appears to be functional in *Plasmodium reichenowi* in which the frameshifts are missing (Rayner *et al.*, 2004). *P. reichenowi* infects chimpanzee, and is the closest *Plasmodium* species to *Plasmodium falciparum.* Considering the close evolutionary relationship between the two parasite species and also between their primate hosts, EBA-165 could potentially be useful for experiments that may help to understand the evolution of host tropism in *P. falciparum* (see Chapter 6). Therefore, we felt that it was reasonable to add EBA-165 on our protein candidate list, even though it is not currently known whether any *P. falciparum* parasites express it. The two frameshifts in *Pf*eba-165 gene (from 3D7 genome) were corrected before gene synthesis. To correct the 5' frameshift, an adenine was inserted at position 219 of the *Pf*eba-165 coding sequence (CDS). To restore the reading frame towards the 3' end, the adenine at position 1252 of *Pf*eba-165 CDS was deleted (Rayner *et al.*, 2004). The final list consisted of 26 merozoite cell surface proteins (Fig. 5.2).

5.3.2 Recombinant expression of an expanded *P.falciparum* merozoite protein library

The 26 shortlisted merozoite candidate proteins were expressed recombinantly as described in section 2.1. Briefly, the sequences encoding for full-length secreted proteins or for ectodomains of cell surface proteins were codon optimised for expression in the human cell line HEK293E, and the potential N-linked glycosylation

sites were systematically removed by substituting alanine for serine/threonine at these sites. Codon optimised expression constructs were synthesised by gene synthesis and provided already sub-cloned into a pTT3-based vector which enables recombinant expression of C-terminally Cd4-hexa-histidine tagged proteins (Fig. 2.1). For recombinant protein expression, the plasmid vectors carrying the synthesised merozoite protein coding sequences, were transfected into HEK293E cells and tissue culture supernatant was harvested six days post transfection. Merozoite proteins were purified from tissue culture supernatant by affinity chromatography on a 96-well nickel column plate.

Purified proteins were quantified relative to Cd4 by ELISA (Fig. 5.3). Most of the proteins were expressed at the levels required for the AVEXIS assay (sufficient amount of protein for complete saturation of the available binding surface Bushell *et al.*, 2008). However, five proteins (GLURP, PF14_0044, PFA0445w, SERA1, SERA6) were almost undetectable by ELISA after five different transfection batches and therefore, excluded from further analysis (indicated in red in Fig. 5.3). The expression level for another three proteins which are indicated in cyan in Fig. 5.3 (SERA7, PFA0210c, RIPR) was below the threshold required for the AVEXIS screen assay, but because some protein was detected, it was considered worth testing them in the experiments described below. These data demonstrate that of 26 recombinant merozoite proteins, 21 (>80%) were successfully expressed at usable amounts.

To confirm expression at the correct size, biotinylated merozoite baits were analysed by western blot (Fig. 5.4). No band was detected for PF14_0044, PFA0445w and SERA1 (data not shown). This observation is consistent with the results obtained by ELISA, and suggests that these proteins failed to express. For SERA 7 and PF08_0006 only a band much lower than the expected size (probably representing the protein tags; section 5.4.1) was detected. In all the remaining proteins, a band representing the full length ectodomain was obtained, but a number of smaller bands were also evident probably due to protein degradation or likely because of proteolytic processing (Fig. 5.4) (section 5.4.1). These observations establish the successful expression of the majority of recombinant merozoite proteins. Overall, 19 of 26 proteins (>73%) were expressed at usable amounts, and a band corresponding to full length recombinant protein was detected by immunoblot.

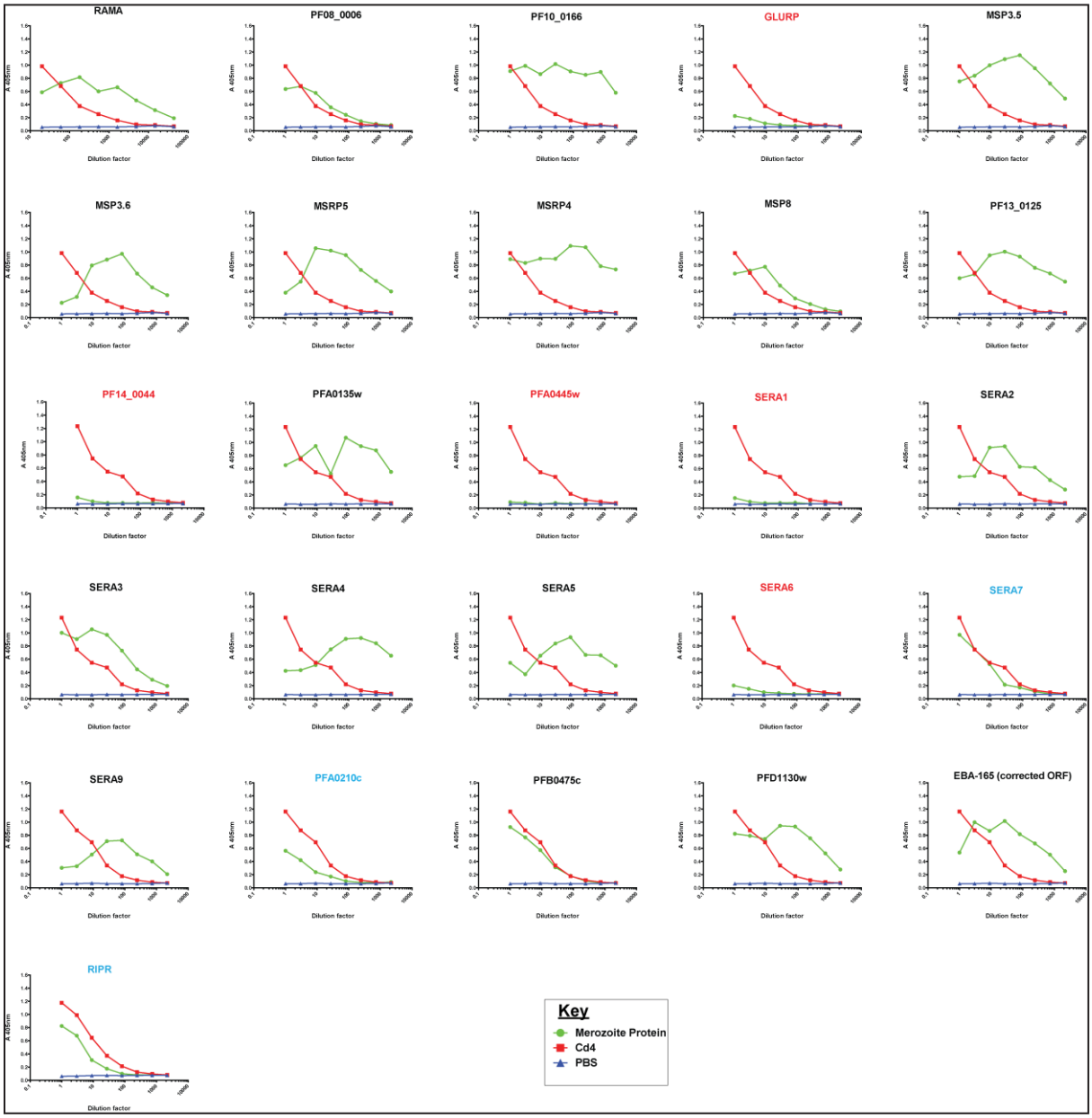

Figure 5.3 Quantitation of merozoite cell surface protein library by ELISA. Purified, recombinant merozoite proteins were serially diluted and immobilised on a streptavidin-coated microtiter plate. The anti-Cd4 mouse monoclonal OX68, was used as the primary antibody and an alkaline phosphatase-conjugated anti-mouse IgG antibody as the secondary antibody. Biotinylated Cd4 was used as reference. The proteins where the expression was almost undetectable by ELISA after five different transfection batches are indicated in red. The proteins which did not reach the threshold activity required for the AVEXIS assay are indicated in cyan.

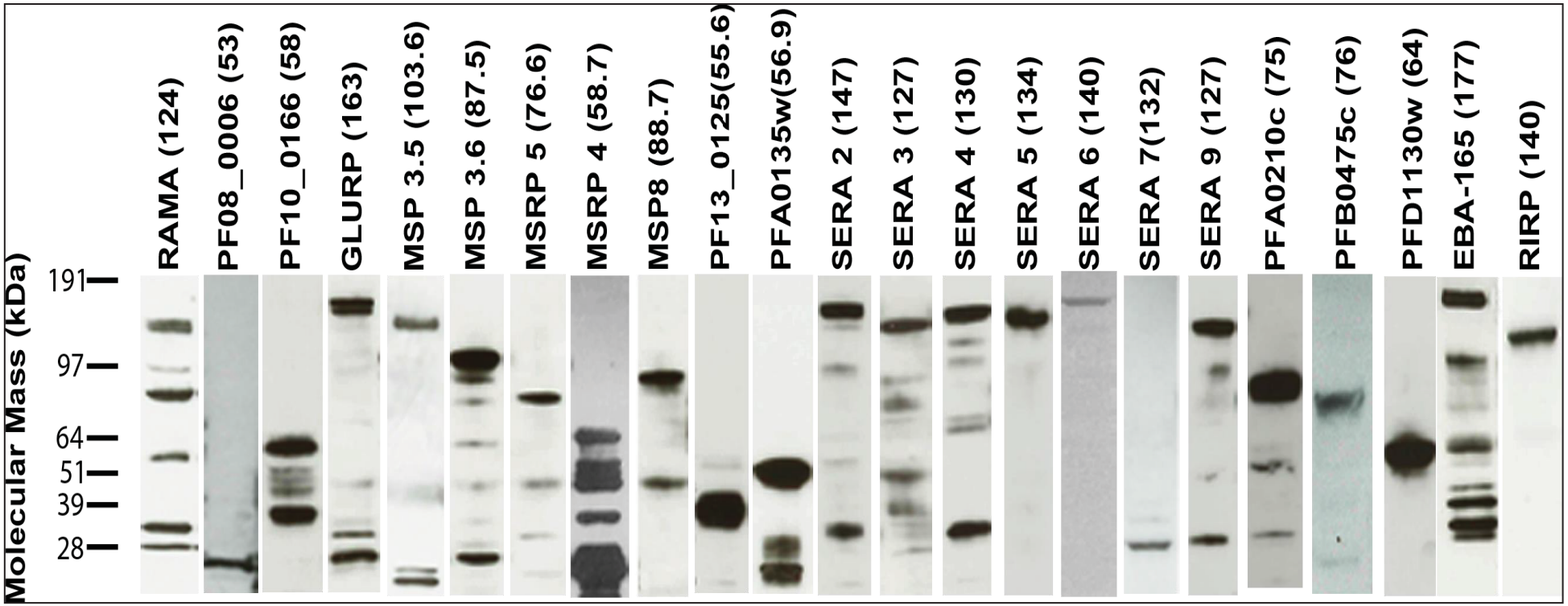

Figure 5.4 The majority of recombinant merozoite extracellular proteins were expressed at the expected size. Normalised amounts of purified biotinylated merozoite proteins were resolved under reducing conditions by SDS-PAGE, blotted, and probed using streptavidin-HRP. The expected molecular weight for the each recombinant protein is indicated in brackets above each lane; this includes the Cd4-BioLinker-6xHis tags (28kDa).

5.3.3 Biochemical characterisation of the recombinant merozoite protein library

Specific interactions can only detected between proteins that are correctly folded (Bartholdson *et al.*, 2013). Thus, it was considered essential to biochemically characterise the members of the merozoite bait library and confirm that the proteins are active and correctly folded, prior to performing the AVEXIS screen. For this purpose a series of three experiments were designed (described below) to derive evidence for the correct conformation of the recombinant merozoite proteins.

5.3.3.1 Members of the merozoite protein library contain heat-labile epitopes

In nature, proteins are recognised by antibodies in their native conformation. The latter implies that antibodies recognise and bind to conformational epitopes on target proteins. Therefore, denaturation of these epitopes by heat treatment normally reduces immunoreactivity. This logic was used to derive evidence whether recombinant merozoite proteins were correctly folded. Recombinant proteins at the correct conformation were expected to be recognised by naturally occurring antibodies, and immunoreactivity was anticipated to decrease after heat treatment of the same proteins.

For this purpose, recombinant merozoite cell surface proteins were assayed by ELISA, for their immunoreactivity against pooled hyperimmune sera obtained from previously malaria exposed Malawian adults (provided by Dr Faith Osier; Taylor *et al.*, 1992). Each biotinylated merozoite protein was immobilised on a 96-well streptavidin-coated plate (to levels sufficient for complete saturation of the available binding surface) and incubated with protein G purified total IgG from hyperimmune or non-immune sera. In this experiment AMA-1 was used as positive control. AMA-1 is a merozoite cell surface protein which has been previously shown to elicit high antibody titers (Thera *et al.*, 2010; Ellis *et al.*, 2012). The expression construct for AMA-1 was available in our laboratory for previous experiments and it was used for recombinant expression of AMA-1 (Crosnier *et al.*, 2013). The OX68 anti-Cd4 monoclonal antibody was used to confirm bait immobilisation.

Most of the merozoite recombinant proteins gave a signal well above the baseline against the hyperimmune sera and they were negative against the non-immune control sera (Fig. 5.5). For the vast majority of proteins the immunoreactivity

against the immune sera, decreased or completely disappeared when the proteins were heat-inactivated. These data show that antibodies contained in hyperimmune sera bind to heat-labile epitopes on recombinant merozoite proteins. When these proteins are linearized by heat-treatment, they are no longer as immunoreactive suggesting that merozoite recombinant proteins adopt their native conformation.

5.3.3.2 *In vitro* interaction of RIPR with RH5

RIPR is a merozoite cell surface protein which has been recently proposed to be an RH5 interaction partner (Chen *et al.*, 2011). Therefore, the recapitulation of this interaction by AVEXIS, using recombinant proteins, would provide evidence that RIPR has the correct conformation. Thus, monomeric biotinylated RIPR bait was recombinantly expressed, purified by affinity chromatography on nickel loaded Sepharose column and quantitated by ELISA (Fig. 5.6 A). Pentameric, β-lactamase tagged RH5 prey was also expressed and quantified by using β-lactamase activity as a proxy (Fig. 5.6A). For the AVEXIS assay, RIPR was immobilised on a streptavidin-coated plate and probed with pentameric RH5. Even though RIPR expression level was below the threshold required for the AVEXIS (sufficient amount of protein for complete saturation of the available binding surface) the interaction between RIPR and RH5 was still detected (Fig. 5.6B). This result suggests that RIPR is active and correctly folded.

5.3.3.3 EBA165 demonstrates differential glycan binding in comparison to other Erythrocyte Binding Ligands (EBLs)

The AVEXIS assay was primarily designed for the detection of low affinity protein-protein interactions, but can also be used to detect a range of extracellular interactions (Bushell *et al.*, 2008; Bartholdson *et al.*, 2013). By using AVEXIS, EBA-175 and EBA-140 preys were previously shown to bind a number of biotinylated glycan baits (Dr Madushi Wanaguru, unpublished data). These glycans are members of a larger glycan panel available in our laboratory which is enriched for glycans present on human erythrocytes (www.glycotech.org). It was interesting to test whether EBA-165, which belongs to the same protein family as EBA-175 and EBA-140 (Triglia *et al.*, 2001; Rayner *et al.*, 2004), was also capable of binding to certain

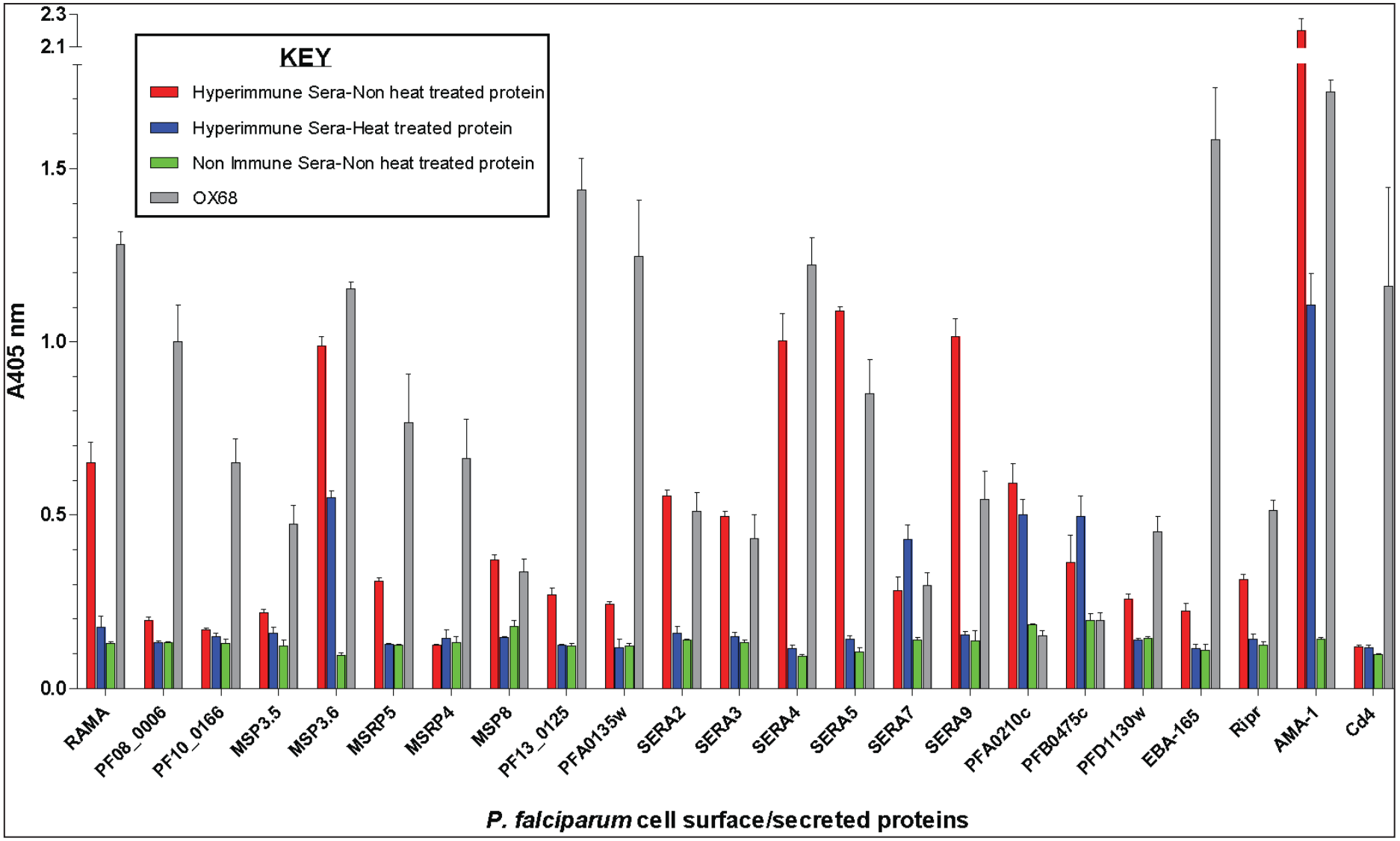
KEY
Hyperimmune Sera-Non heat treated protein
Hyperimmune Sera-Heat treated protein
Non Immune Sera-Non heat treated protein
OX68
A405 nm
2.3
2.1
1.5
1.0
0.5
0.0
RAMA
PF08_0006
PF10_0166
MSP3.5
MSP3.6
MSRP5
MSRP4
MSP8
PF13_0125
PFA0135w
SERA2
SERA3
SERA4
SERA5
SERA7
SERA9
PFA0210c
PFB0475c
PFD1130w
EBA-165
Ripr
AMA-1
Cd4
P. falciparum cell surface/secreted proteins

Figure 5.5 Members of the merozoite cell surface protein library are reactive against hyperimmune sera obtained from previously malaria exposed Malawian adults.

Biotinylated, purified merozoite cell surface proteins (native or heat treated at 80^{o}C for 10mins) were immobilised on a streptavidin-coated plate and incubated for 2 hours in the presence of immune or non-immune sera or OX68 (anti-Cd4) followed by 1 hour incubation with an anti-hIgG or anti-mIgG alkaline phosphatase conjugated secondary antibody. Biotinylated AMA-1 and Cd4 were used as positive and negative control, respectively. Data points are shown as mean ± s.e.m; *n*=3.

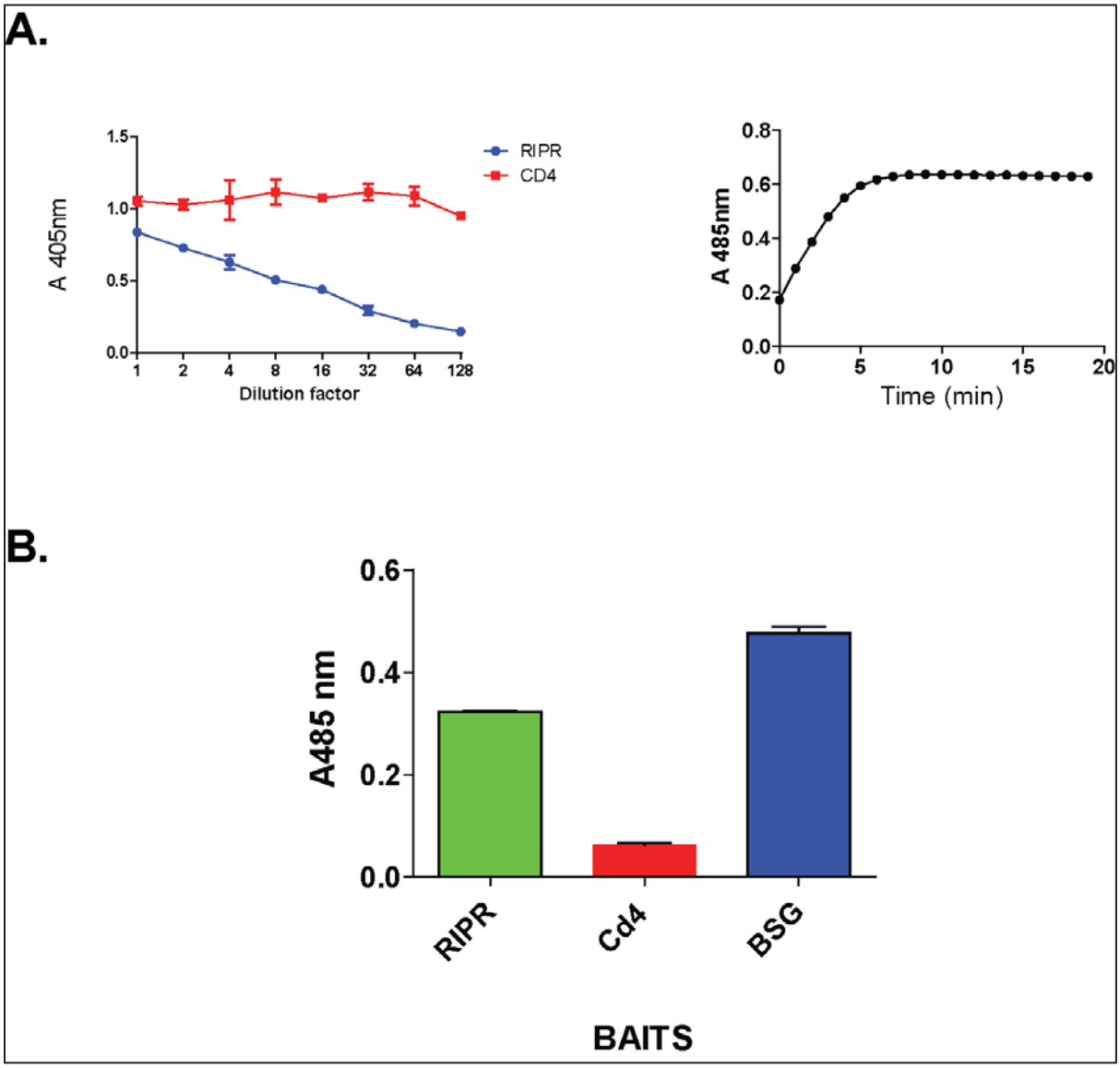

Figure 5.6 RIPR binds RH5 in AVEXIS.

A. Purified biotinylated RIPR bait (left) and RH5 pentameric prey (right), were quantitated by ELISA and time course monitoring of nitrocefin turnover, respectively. While RH5 activity reached the level required for AVEXIS (section 2.5), the expression level of RIPR was below the threshold required for AVEXIS. Biotinylated Cd4 was used as reference in ELISA.

B. RIPR was clearly bound by RH5 in an AVEXIS assay. Biotinylated RIPR was immobilised on a streptavidin-coated plate and probed with RH5-prey, in an AVEXIS assay. Biotinylated Cd4 and BSG were used as negative and positive control, respectively. Data is shown as mean ± s.e.m; *n*=3.

glycans. Such a case would provide evidence that recombinant EBA-165 has adopted the correct conformation.

First, EBA-140, EBA-175 and EBA-165 preys were recombinantly expressed and their activity in tissue culture supernatant was assessed by monitoring nitrocefin turnover over time (Fig 5.7A). For the AVEXIS screen, each glycan was immobilised on a streptavidin-coated plate to levels sufficient for complete saturation of the available binding surface, following by incubation with each of the EBL preys. The EBA-165 prey clearly bound multiple glycans, suggesting that folded domains do exist on the recombinant protein (Fig 5.7B). Interestingly, EBA-165 was able to bind only to glycans containing N-glycolylneuraminic acid (Neu5Gc), whereas EBA-140 and EBA-175 were capable of binding to both N-aceytlneuraminic acid (Neu5Ac) and Neu5Gc containing glycans. Because humans lack the enzyme that converts Ne5Ac to Neu5Gc, Neu5Gc are the dominant sialic-acids in apes but not in humans. (Chou *et al.*, 1998; Varki, 2001). Given that EBA-165 appears to be a pseudogene in the human malaria parasite *P.falciparum,* but is functional *P. reichenowi* (chimpanzee restricted parasite), this observation may have implications for the evolution of *P. falciparum* host specificity, which are discussed Chapter 6.

5.3.4 Recombinant expression of a human erythrocyte cell surface protein library

Crosnier and colleagues compiled a list of cell surface and secreted proteins expressed by human erythrocytes (Crosnier *et al.*, 2011), based on published proteomics data (Pasini *et al.*, 2006). Proteins for which the entire ectodomain was expected to be expressed as a soluble recombinant protein were selected and recombinantly expressed in HEK293E cells (Crosnier *et al.*, 2011). The expression constructs encoding for the prey forms of 31 of those erythrocyte cell surface proteins, were employed for the purposes of my experiments (Fig. 5.8).

For recombinant protein expression the above constructs were transfected in HEK293E cells and tissue culture supernatants were analysed for the presence of recombinant β-lactamase-tagged protein preys, by monitoring nitrocefin turnover six days post transfection (Fig. 5.9). To achieve the prey activity which was required for the AVEXIS assay (sufficient protein in 20µl prey tissue culture supernatants to turnover 60 µl of 250 µg/mL nitrocefin within 10mins) (Bushell *et al.*, 2008), repeated

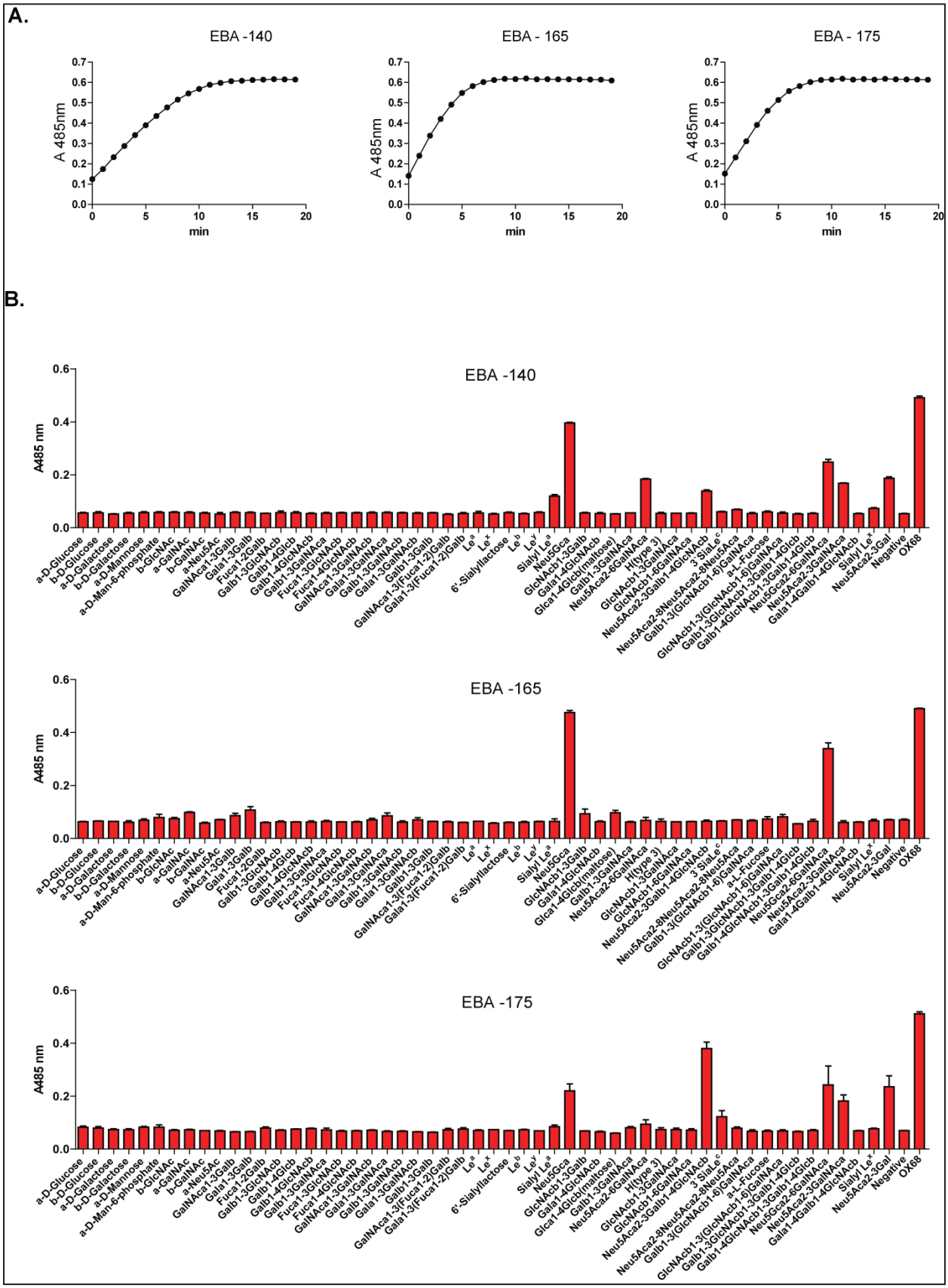
A.
EBA -140
EBA - 165
EBA - 175
A 485nm
min
B.
EBA -140
EBA -165
EBA -175
A485 nm
a-D-Glucose
b-D-Glucose
a-D-Galactose
b-D-Galactose
a-D-Mannose
a-D-Man-6-phosphate
b-GlcNAc
a-GalNAc
b-GalNAc
a-Neu5Ac
GalNAca1-3Galb
Gala1-3Galb
Fuca1-2Galb
Galb1-3GlcNAcb
Galb1-4Glcb
Galb1-4GlcNAcb
Galb1-3GalNAca
Fuca1-3GlcNAcb
Fuca1-4GlcNAcb
GalNAca1-3GalNAcb
Gala1-3GalNAca
Galb1-3GalNAcb
Galb1-3Galb
GalNAca1-3(Fuca1-2)Galb
Gala1-3(Fuca1-2)Galb
Le^a
Le^x
6'-Sialyllactose
Le^b
Le^y
Sialyl Le^a
Neu5Gca
GlcNAcb1-3Galb
Gala1-4GlcNAcb
Glca1-4Glcb(maltose)
Galb1-3GalNAca
Neu5Aca2-6GalNAca
H(type 3)
GlcNAcb1-3GalNAca
GlcNAcb1-6GalNAca
Galb1-3GlcNAcb1-4GlcNAcb
3'SiaLe^c
Neu5Aca2-8Neu5Aca
Neu5Aca2-3Galb1-4GlcNAcb
a-L-Fucose
GlcNAcb1-3(GlcNAcb1-6)GalNAca
Galb1-3(GlcNAcb1-6)GalNAca
GlcNAcb1-3Galb1-3Galb1-4Glcb
Galb1-3GlcNAcb1-3Galb1-4Glcb
Galb1-4GlcNAcb1-3Galb1-4Glcb
Neu5Aca2-3GalNAca
Neu5Gca2-6GalNAca
Gala1-4Galb1-4GlcNAcb
Sialyl Le^x
Neu5Aca2-3Gal
Negative
OX68

Figure 5.7 EBA-165 demonstrates differential binding to a glycan panel, in comparison to other members of the EBL family of proteins.

A. EBA-165, EBA-140, EBA-175 were expressed as β lactamase-tagged pentameric preys and quantitated by monitoring their enzymatic activity.

B. EBA-165, EBA-140, EBA-175 were screened against a glycan panel, by using AVEXIS. Normalised amounts of biotinylated glycans were immobilised on a streptavidin-coated plate and probed with pentameric preys. The biotinylated monoclonal anti-Cd4 antibody OX68 was used to confirm prey activity. The biotinylated bait used as the negative control contained the linker region carried by all the synthetic glycans. Data are shown as mean ± standard error; *n*=3

No	Protein Name	Trunc.	Type	Accession code	Blood group
1	GLYCOPHORIN A	P90	I	P02724	MNS
2	GLYCOPHORIN B	P60	I	P06028	MNS
3	GLYCOPHORIN E	N51	I	P15421	
4	GLYCOPHORIN C	D58	III	P04921	Gerbich
5	MIC2/CD99	P124	I	P14209	Xg
6	CD44_ISOFORM 1	P648	I	P16070	Indian
7	COMPLEMENT RECEPTOR 1_ISOFORM S	L2423	I	NP_000642	Knops
8	BASIGIN_ISOFORM 1	L322	I	P35613_1	Ok
9	BASIGIN_ISOFORM 2	L206	I	P35613_2	Ok
10	BCAM	G548	I	P50895	Lutheran
11	AMIGO2	T398	I	Q86SJ2	
12	NICASTRIN	L670	I	Q92542	
13	ERMAP	S154	I	Q96PL5	Scianna/Radin
14	C1ORF9	R1010	I	Q9UBS9	
15	JAM1	V238	I	Q9Y624	
16	NEUROPLASTIN	P223	I	Q9Y639	
17	PROGESTERON RECEPTOR COMPONENT 2	A46	I	O15173	
18	LFA3/CD58	L218	I	P19256	
19	ENDOD1	P338	I	O94919	
20	SEMAPHORIN7A	A644	GPI	O75326	John Milton Hagen
21	COMPLEMENT DECAY ACCELERATING FACTOR	S353	I/GPI	P08174	Cromer
22	MACIF/CD59	N102	GPI	P13987	
23	ACETYLCHOLINESTERASE_ISOFORM H	T581	GPI	P22303-2	Yt
24	MULTIDRUG RESISTANCE PROTEIN 1/MRP1	T36	17TM	P33527-2	
25	DUFFY_ISOFORM 1	P63	7TM	Q16570-2	
26	PROLACTIN	n.a.	sec.	P01236	
27	LACTOTRANSFERRIN	n.a.	sec.	P02788	
28	C8ORF55	n.a.	sec.	Q8WUY1	
29	METHYLTRANSFERASE-LIKE PROTEIN 7A	n.a.	sec.	Q9H8H3	
30	GALECTIN-3/LGALS3	n.a.	sec.	P17931	
31	PROTEIN DISULFIDE-ISOMERASE/P4HB	n.a.	sec.	P07237	

Figure 5.8 The erythrocyte cell surface receptor library. Abbreviations: Trunc.,the ectodomain truncation residue; I, type I; III, type III; GPI, GPI-anchored; #TM, multipass transmembrane protein with number of predicted transmembrane domains; n.a, not applicable; sec., secreted. Figure adapted from Crosnier *et al.*, 2011.

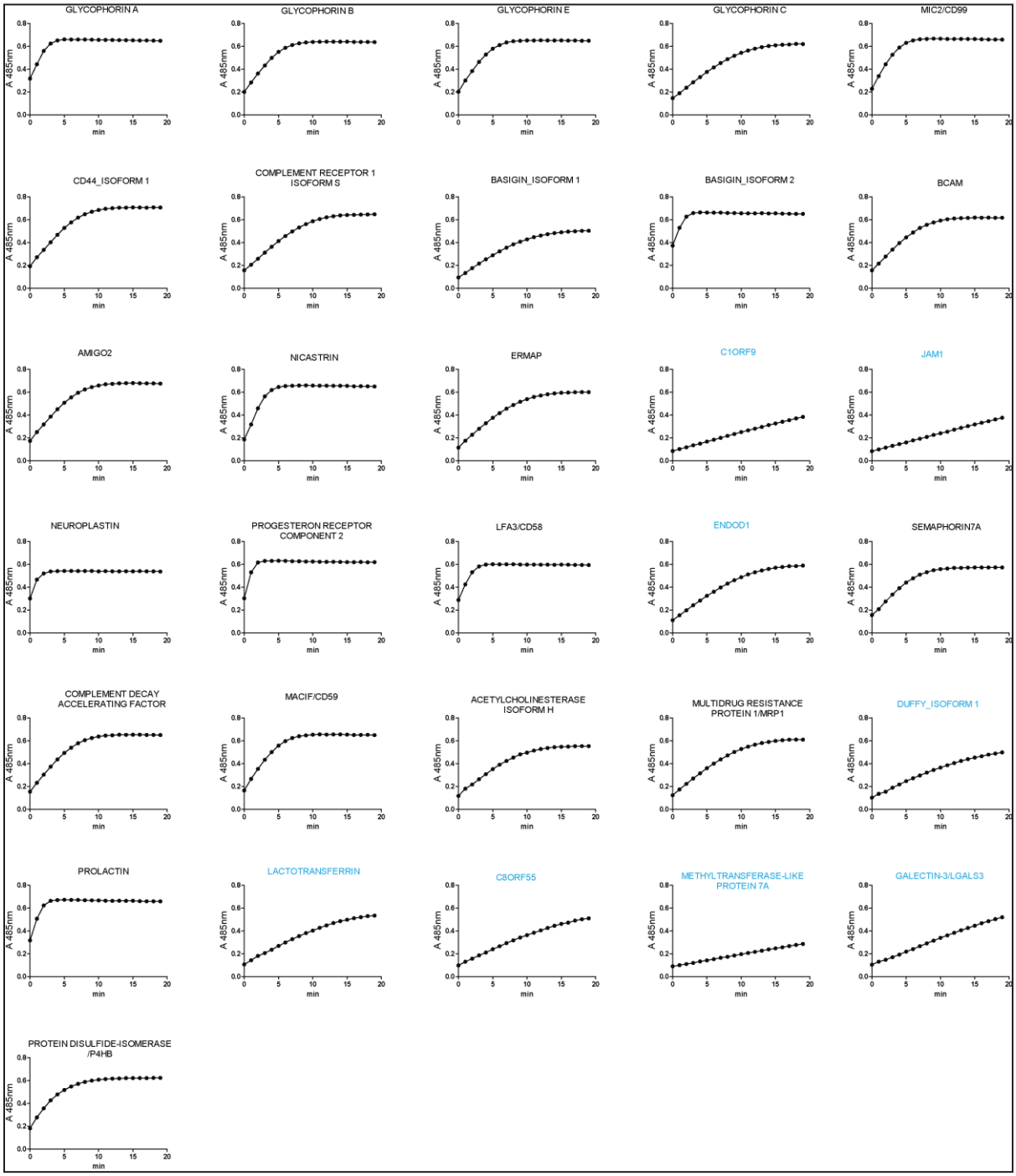

Figure 5.9 Expression and activity assessment of erythrocyte receptor preys. The activity of erythrocyte β-lactamase tagged pentameric preys was assessed by quantifying the turnover of the colourimetric β-lactamase substrate as a function of time. The proteins which did not reach the threshold activity required for the AVEXIS assay are indicated in cyan. The tissue culture supernatants containing the latter proteins were concentrated five times before assessment of β-lactamase activity.

rounds of transfection-analysis were performed. Proteins which did not reach the required expression level after five transfection batches, were concentrated five times, before used for probing on AVEXIS (indicated in cyan in Fig. 5.9)

5.3.5 The AVEXIS screen identified a putative interaction

After biochemically characterising the recombinant merozoite cell surface protein library, and recombinantly expressing a panel of erythrocyte cell surface receptors, I proceeded to the AVEXIS screen. For the screen the merozoite bait array was immobilised on a streptavidin-coated 96 well plate followed by incubation with each of the erythrocyte pentameric preys. The interaction between RH5 and BSG was used as a positive control (Crosnier *et al.*, 2011). Because the activity of several baits and preys was below the threshold required for the AVEXIS assay, to cover a range of bait and prey activities, a dilution series of RH5 bait was probed against BSG prey and vice versa (Fig. 5.10A). Cd4 was used as negative control, and biotinylated OX68 was used to capture the preys on the plate and thereby to confirm prey activity. No robust interaction with z-scores > 2 was detected (Fig 5.10B). Therefore, the three interactions (PF13_0125 – Prolactin, PF13_0125 – P4HB, PFA0135w – P4HB) which had z-scores > 1.5 (Fig 5.11) were investigated further.

Interactions detected by AVEXIS are only considered positive if they are independent of the bait–prey orientation (Bushell *et al.*, 2008). Therefore, I attempted to reciprocate the orientation of the proteins implicated in the three putative interactions identified in my screen. For this purpose, pentamerized β-lactamase tagged PF13_0125 and PFA0135w were recombinantly expressed and quantitated by monitoring β-lactamase activity over time (Fig. 5.12A). Monomeric biotinylated P4HB and Prolactin were also recombinantly expressed and quantitated by ELISA (Fig. 5.13C). All recombinant proteins were used for another AVEXIS assay. While the interactions PF13_0125 - P4HB, and PFA0135w – P4HB were detected again, the interaction between PF13_0125 and Prolactin was not (Fig. 5.12B). These reciprocation data suggest that while the interactions PF13_0125 – P4HB, and PFA0135w – P4HB were likely to be valid, the interaction between PF13_0125 and Prolactin was probably a false positive hit of the AVEXIS screen.

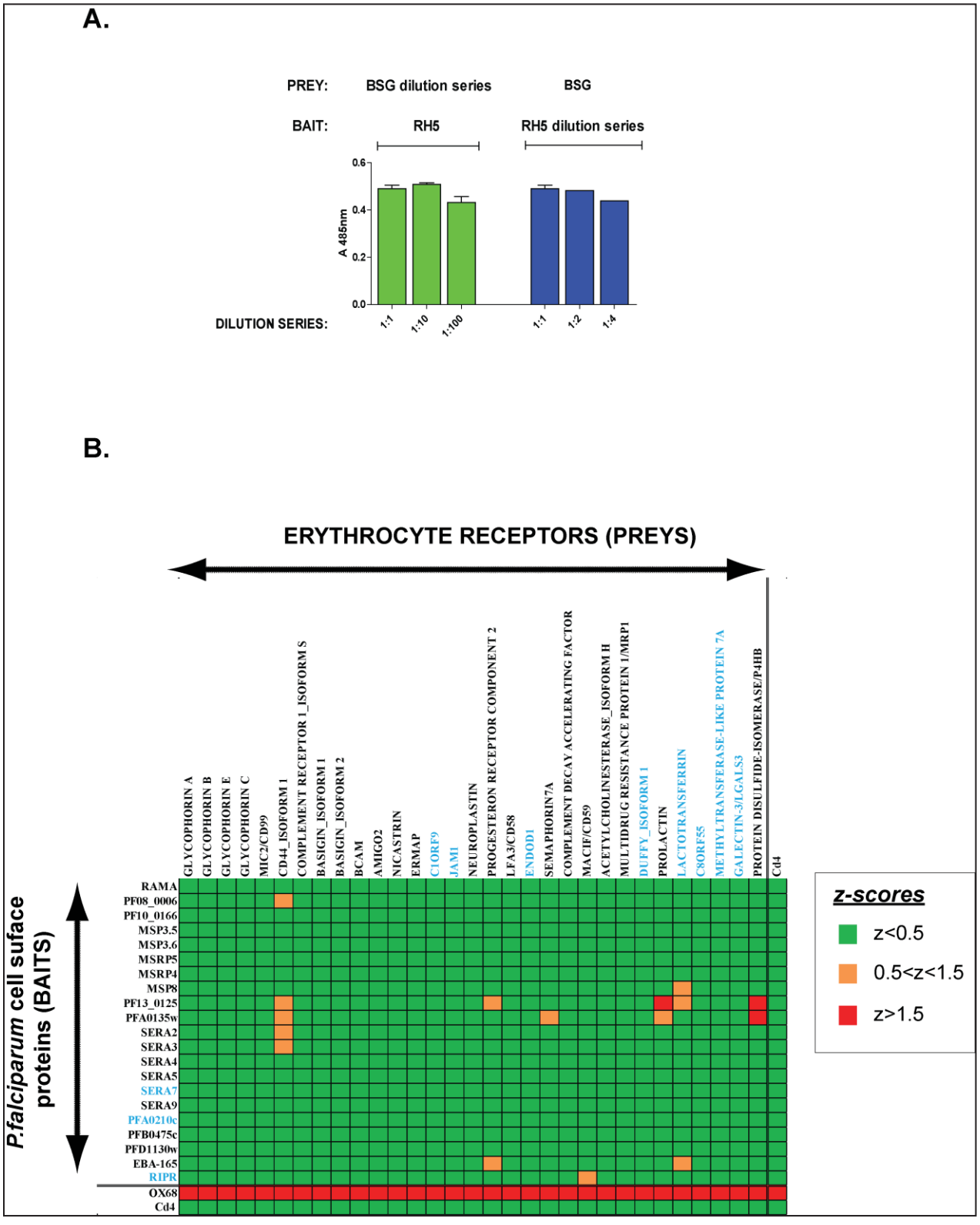
A.
PREY:
BSG dilution series
BSG
BAIT:
RH5
RH5 dilution series
A 485nm
0.6
0.4
0.2
0.0
DILUTION SERIES:
1:1
1:10
1:100
1:1
1:2
1:4
B.
ERYTHROCYTE RECEPTORS (PREYS)
GLYCOPHORIN A
GLYCOPHORIN B
GLYCOPHORIN E
GLYCOPHORIN C
MIC2/CD99
CD44_ISOFORM 1
COMPLEMENT RECEPTOR 1_ISOFORMS
BASIGIN_ISOFORM 1
BASIGIN_ISOFORM 2
BCAM
AMIGO2
NICASTRIN
ERMAP
C1ORF9
JAM1
NEUROPLASTIN
PROGESTERON RECEPTOR COMPONENT 2
LFA3/CD58
ENDOD1
SEMAPHORIN7A
COMPLEMENT DECAY ACCELERATING FACTOR
MACIF/CD59
ACETYLCHOLINESTERASE_ISOFORM H
MULTIDRUG RESISTANCE PROTEIN 1/MRP1
DUFFY_ISOFORM 1
PROLACTIN
LACTOTRANSFERRIN
C8ORF55
METHYLTRANSFERASE-LIKE PROTEIN 7A
GALECTIN-3/LGALS3
PROTEIN DISULFIDE-ISOMERASE/P4HB
Cd4
P.falciparum cell suface proteins (BAITS)
RAMA
PF08_0006
PF10_0166
MSP3.5
MSP3.6
MSRP5
MSRP4
MSP8
PF13_0125
PFA0135w
SERA2
SERA3
SERA4
SERA5
SERA7
SERA9
PFA0210c
PFB0475c
PFD1130w
EBA-165
RIPR
OX68
Cd4
z-scores
z<0.5
0.5<z<1.5
z>1.5

Figure 5.10 The systematic screen between the merozoite and erythrocyte protein libraries resulted in three interactions with z-score>1.5.

A. A bar chart demonstrating the interaction between Basigin-prey and RH5-bait (both from tissue culture supernatant) which was used as positive control in the screen. To cover a range of bait and prey activities a dilution series of RH5 was probed against Basigin and vice versa. Data are shown as mean ± s.e.m; *n*=3.

B. A z-score heat-map summarizing the results of the screen**.** Three interactions (PF13_0125 – Prolactin, PF13_0125 – P4HB, PFA0135w – P4HB) had z-score>1.5. Biotinylated merozoite protein baits were immobilised on a streptavidin-coated plate and probed with erythrocyte pentameric preys. Where applicable preys or baits activities were normalised against each other. The biotinylated monoclonal anti-Cd4 antibody OX68 was used to confirm prey activity. The proteins which did not reach the expression levels required for the AVEXIS assay are indicated in cyan.

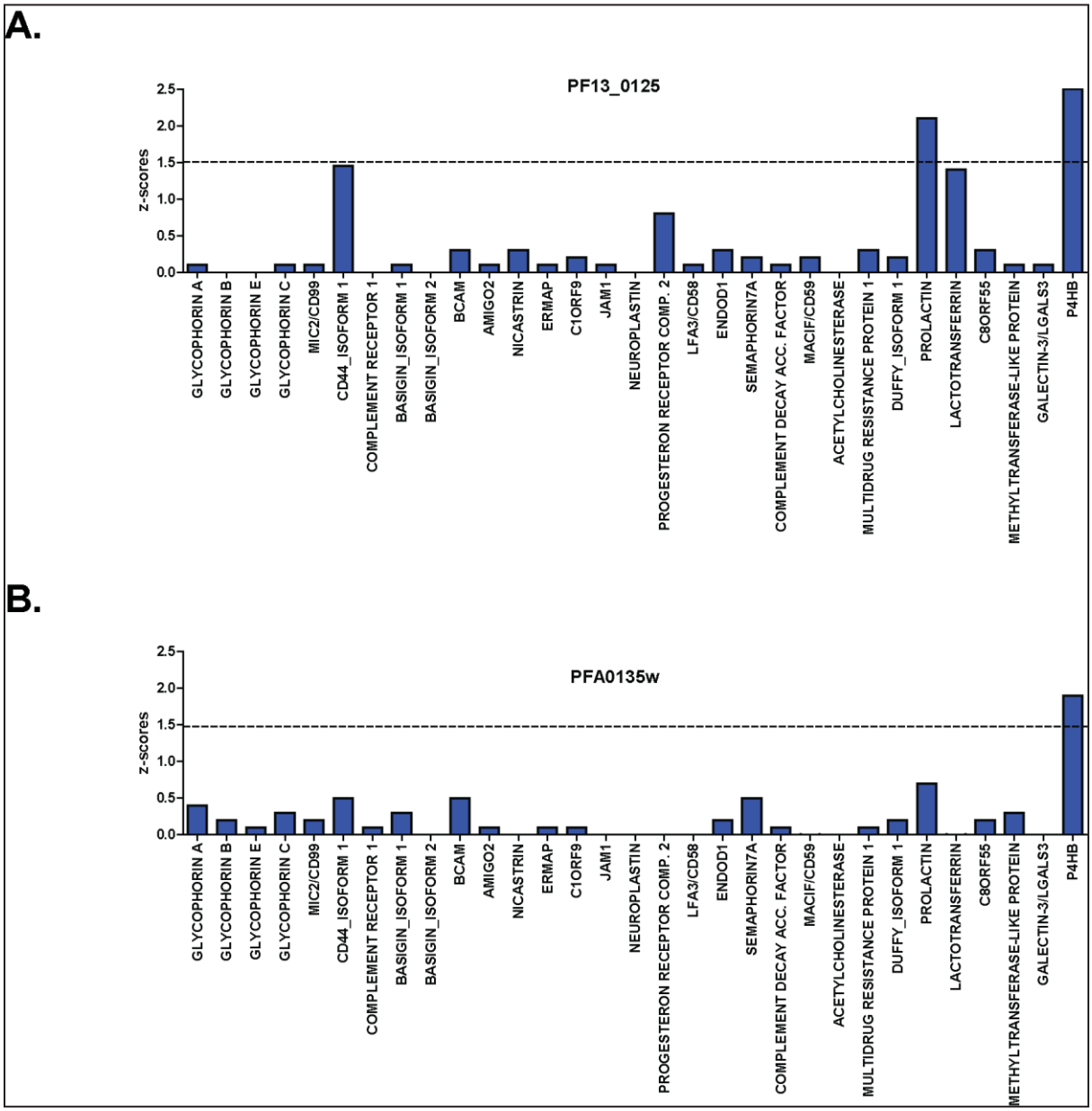

Figure 5.11 The interactions between PF13_0125, PFA0135w and P4HB, and between PF13_0125 and Prolactin had z scores>1.5. The bar charts show the normalised binding (z scores) of PF13_0125 (***A.***) and PFA0135w (***B***.) to the panel of erythrocyte receptors. Dashed line indicates z-score=1.5 which was set as threshold.

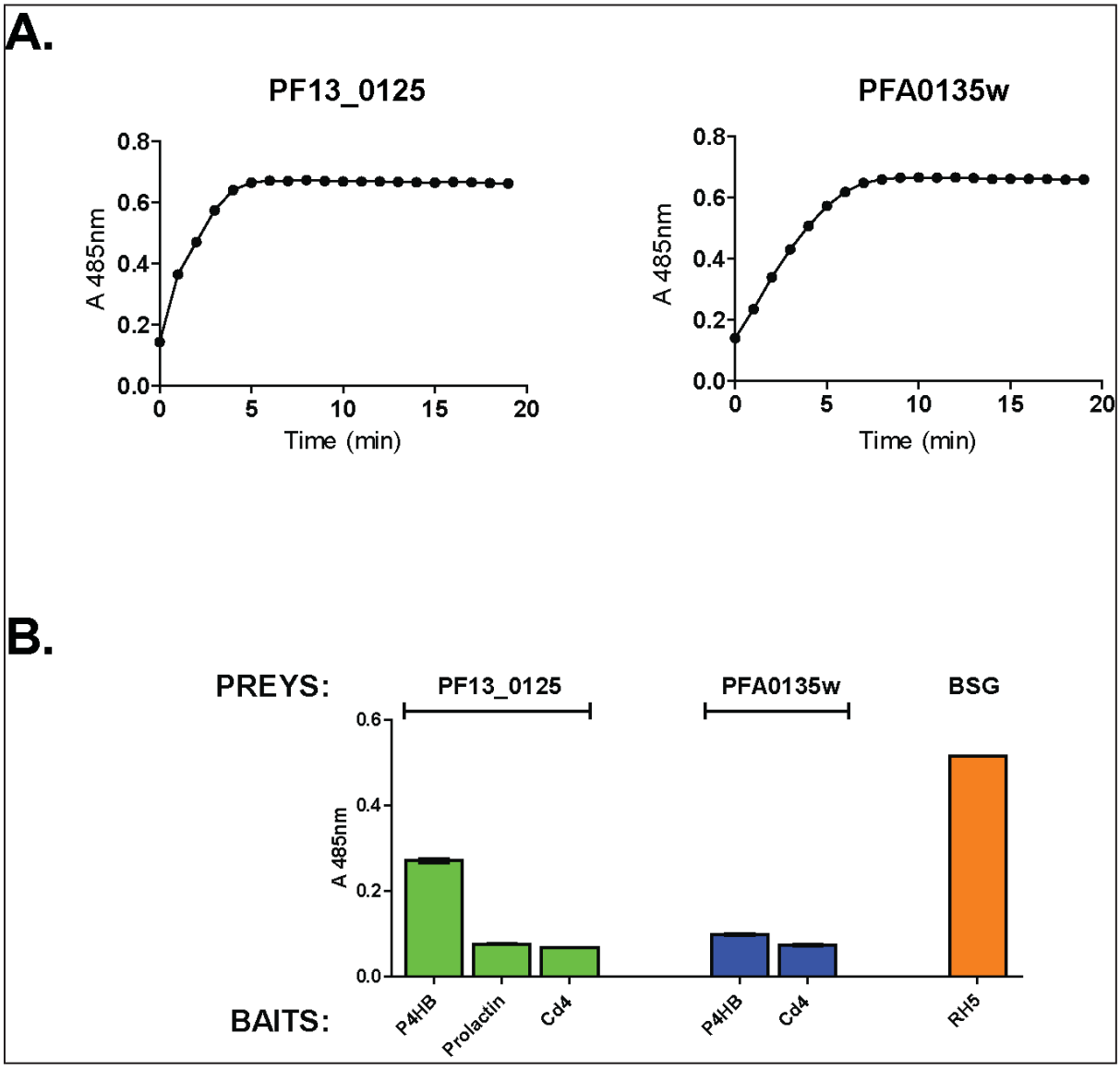

Figure 5.12 The interactions between PF13_0125, PFA0135w and P4HB are independent of bait-prey orientation.

A. PF13_0125 and PFA0135w were expressed in the pentameric β-lactamse prey form and their activity was normalised by monitoring nitrocefin turnover

B. Normalised amounts of biotinylated P4HB and Prolactin baits from tissue culture supernatant, were immobilised on a streptavidin-coated plate and probed with affinity purified PF13_0125 or PFA0135w preys. The interactions PF13_0125 - P4HB, and PFA0135w – P4HB were detectable, whereas the interaction between PF13_0125 and Prolactin was not. The interaction between RH5 and BSG was used as positive control. Data are shown as mean ± s.e.m; *n*=3. Cd4-bio was used as negative control.

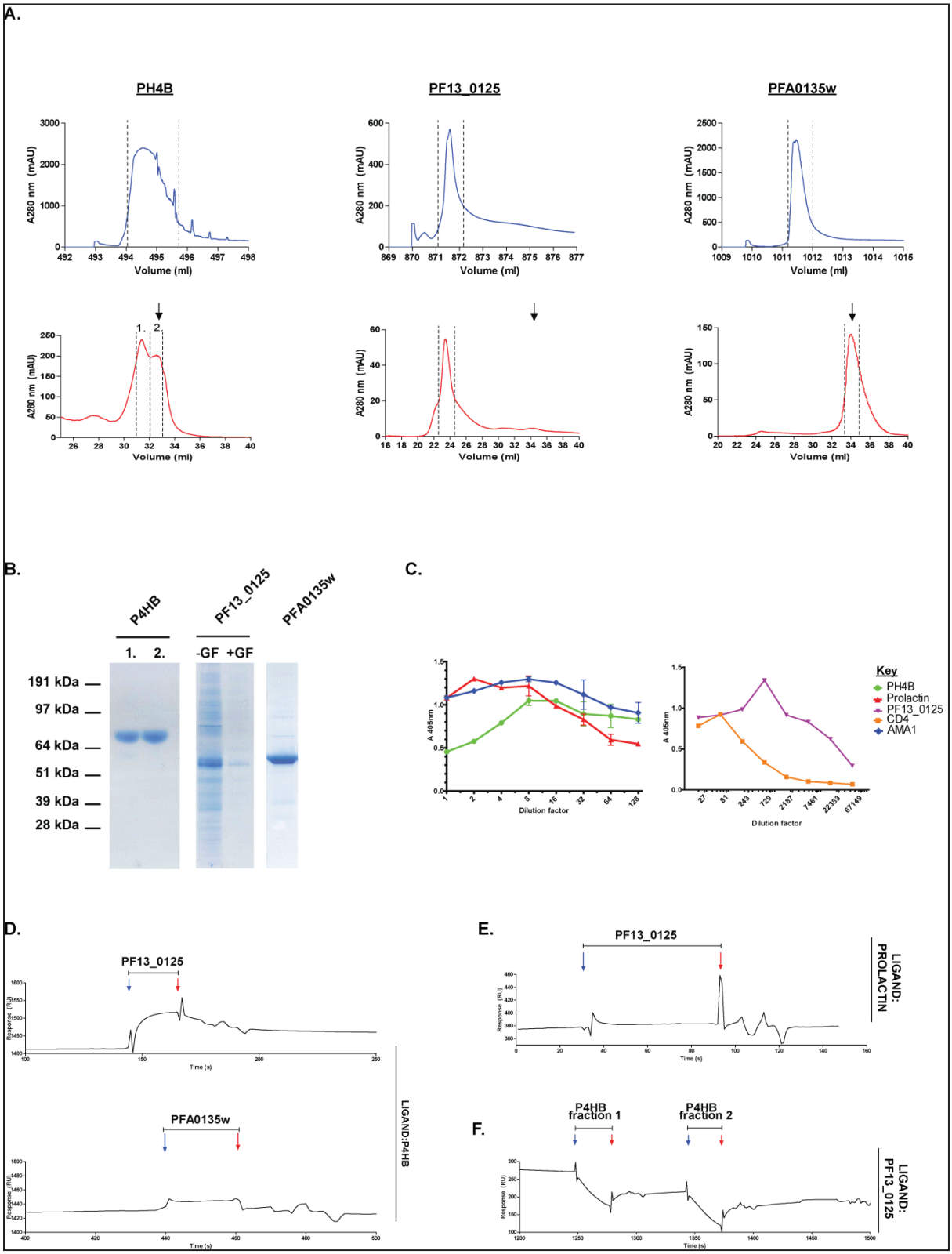
A.
PH4B
PF13_0125
PFA0135w
A280 nm (mAU)
Volume (ml)
1. 2.
B.
P4HB
PF13_0125
PFA0135w
1. 2.
-GF +GF
191 kDa
97 kDa
64 kDa
51 kDa
39 kDa
28 kDa
C.
A 405nm
Dilution factor
Key
PH4B
Prolactin
PF13_0125
CD4
AMA1
D.
PF13_0125
PFA0135w
Response (RU)
Time (s)
LIGAND:P4HB
E.
PF13_0125
LIGAND: PROLACTIN
F.
P4HB fraction 1
P4HB fraction 2
LIGAND: PF13_0125

Figure 5.13 Biophysical analysis of the putative interactions between PF13_0125, PFA0135w and P4HB, and between PF13_0125 and Prolactin, by Surface Plasmon Resonance.

A. Histidine tagged P4HB, PF13_0125 and PFA0135w were affinity purified from tissue culture supernatant on a nickel column (top), followed by gel filtration (bottom). The eluate in each experiment, was monitored at 280 nm in real-time and the peak fractions containing protein (between dashed lines) were pooled. Two gel filtration peak fractions (1 and 2) were collected for P4HB. The expected gel filtration elution volumes, which are indicated by arrows, were 32.8ml, 34.6ml, and 34.5ml for P4HB, PF13_0125 and PFA0135w respectively.

B. The gel filtrated proteins were analysed by denaturing SDS-PAGE and visualised using Coomassie brilliant blue. Because of the non-expected elution volume of PF13_0125, a purified non gel-filtrated fraction of this protein was analysed as well. The expected band sizes are 83kDa, 55.3kDa and 56.6kDa for P4HB, PF13_0125 and PFA0135w respectively. Abbreviations: -GF, without gel filtration; +GF, gel filtrated.

C. P4HB, Prolactin and PF13_0125 were expressed as monomeric biotinylated forms and quantified by ELISA**.** A dilution series of each protein was immobilised on a streptavidin-coated plate and probed with the anti-Cd4 mouse monoclonal OX68. An alkaline phosphatase-conjugated anti-mouse IgG antibody as the secondary antibody. In this experiment P4HB and Prolactin were used unpurified from tissue culture supernatant whereas PF13_0125 was affinity purified on nickel column. Biotinylated AMA1 and Cd4 were used as references. Data is shown as mean ± s.e.m; *n*=3.

D. Reference subtracted sensorgrams from the pulse injection of gel filtrated PF13_0125 (top) or PFA0135w (bottom), each at 3 μM, over biotinylated P4HB immobilised on a streptavidin-coated sensor chip. PF13_0125 and PFA0135w were injected at a flow rate of 20 μl/min. The biotinylated ligands were immobilised at: Cd4 (reference)-800 RU, P4HB-2600 RU.

E. Reference subtracted sensorgram from the injection of gel filtrated PF13_0125, at 3 μM, over biotinylated Prolactin immobilised on a streptavidin-coated sensor chip. PF13_0125 was injected at a flow rate of 20 μl/min. The biotinylated ligands were immobilised at: Cd4 (reference)-800 RU, Prolactin-1600RU

F. Reference subtracted sensorgram from the injection of gel filtrated P4HB, at 3 μM, over biotinylated PF13_0125. P4HB was injected at a flow rate of 20 μl/min. The biotinylated baits were immobilised at: Cd4 (reference)-500RU, PF13_0125-1100RU.

Blue and red arrows indicate the start and end of analyte injection and bars above the sensograms represent the duration of injection

5.3.6 SPR analysis of the putative novel interactions identified by using AVEXIS

Specific interactions are saturable, and this is normally determined by biophysical analysis. To test whether the putative interactions identified by AVEXIS, were saturable they were further examined by SPR. For this purpose, PFA0135w and PF13_0125 analytes were purified from tissue culture supernatant by affinity chromatography on a nickel column followed by gel filtration (Fig 5.13A). PFA0135w was eluted from the gel filtration column as a monodispersed peak at the expected elution volume (34.5ml) corresponding to an estimated size of 56.6kDa which was confirmed by SDS-PAGE (Fig 5.13B). Despite the expected elution volume for PF13_0125 being 32.8ml, a single elution peak at about 24ml (corresponding to an approximate size of 195kDa) was consistently obtained, suggesting that PF13_0125 could have been aggregated or associated with other proteins (Fig 5.13A). Analysis of PF13_0125 prior to and after gel filtration demonstrated that a band at the expected size (55.3kDa) was present, confirming the presence of protein in the elutant. However, a number of other bands with higher and lower molecular weights were also evident, suggesting protein aggregation or association with other proteins.

To test whether saturable binding can be observed between the putative interacting partners PF13_0125 – P4HB, PFA0135w – P4HB and PF13_0125 – Prolactin, biotinylated P4HB and Prolactin contained in tissue culture supernatant (Fig. 5.13C) were immobilised on the surface of a sensor chip. Gel filtrated PF13_0125 and PFA0135w were injected over the immobilised proteins and reference subtracted sensorgrams were obtained (Fig. 5.13D, E). No binding of PFA0135w to P4HB and of PF13_0125 to Prolactin was observed, suggesting that these interactions were most likely false positive hits of the AVEXIS screen. In contrast, PF13_0125 consistently bound to P4HB and the binding was saturable (Fig. 5.13D). Nevertheless, PF13_0125 did not dissociate, even after injecting 5M NaCl to facilitate dissociation (data not shown). Complexes from specific interaction between two proteins are expected to dissociate after some time and therefore, the latter observation raised questions for the specificity of this interaction.

To further investigate this result, I tested whether the observation could be replicated when P4HB is used as analyte and PF13_0125 as ligand. P4HB analyte was recombinantly expressed and purified from tissue culture on a nickel column

followed by gel filtration (Fig 5.13A). Two peaks were obtained from gel filtration, and both appeared to contain P4HB as confirmed after analysis by SDS-PAGE (Fig 5.13B). To repeat the SPR experiment in the reverse orientation, biotinylated PF13_0125 (Fig 3.13C) was captured on the surface of a sensor chip, followed by injection of P4HB analyte. Reference subtracted sensorgrams indicated that P4HB contained in both elution peak fractions (Fig 5.13A) demonstrated higher level of binding to Cd4, which was used as reference, than to PF13_0125 (Fig. 513F). These results are not conclusive whether PF13_0125 interacts with P4HB and more experiments are required to validate this interaction (Chapter 6).

5.4 Discussion

5.4.1 Recombinant expression of recombinant merozoite proteins

Based on microarray data and observations reported in the literature a total of 26 merozoite cell surface proteins were chosen for recombinant expression. Of 26 recombinant merozoite proteins, 21 (>80%) were successfully expressed at levels detectable by ELISA. Five proteins (GLURP, SERA6, SERA1, PF14_0044, PFA0445w) were not readily detectable by ELISA (Fig. 5.3). However, a band corresponding to full length ectodomain of GLURP and SERA6 was detected by western, indicating that these proteins were expressed at low levels. The low expression level for SERA6 is in agreement with previous attempts for recombinant expression of a full length form of this protein (Ruecker *et al.*, 2012). Full length recombinant GLURP was previously expressed in *E.coli* but the authors did not report the expression levels (Theisen *et al.*, 1995).

As mentioned in the introduction of this Chapter, recombinant expression of *Plasmodium* proteins is generally problematic (Birkholtz *et al.*, 2008). A possible explanation for the low or no recombinant expression observed for five merozoite proteins might simply be their incompatibility with the HEK293 protein expression system or with heterologous expression in general. An experiment which could potentially improve recombinant expression is discussed in Chapter 6.

For twenty one out of 23 merozoite recombinant proteins that were detectable by western (>88%), a band corresponding the full length ectodomain was observed (Fig. 5.4). Some processing was also evident for several proteins. The lower than the expected size bands likely origin from proteolytic activity of foetal bovine serum

proteases, contained in complete tissue culture medium (Shimomura *et al.*, 1992) or from other proteases released from dead cells in culture. A common band at about 28kDa was detected in most of the processed proteins. This band was observed in previous experiments in our laboratory and represents the protein tags (biotinylatable peptide-Cd4-6xHis). The latter protein fragment appears to be the result of a cleavage event within a presumed proteolytically prone area, nearby the C-terminus of the protein ectodomain sequence. It is possible that full length ectodomain was still recombinantly expressed, but because all tags are located C-terminally, purification on nickel column and detection with streptavidin based probes cannot be achieved.

Naturally occurring proteolytic processing for RAMA (Topolska *et al.*, 2004), GLURP (Borre *et al.*, 1991) and MSP8 (Black *et al.*, 2001; Drew *et al.*, 2005) has been reported in earlier studies. Nonetheless, it is difficult to compare the processing pattern I observed with those reported previously, due to differences in the detection probes that were used in each case. Of note is that only the unprocessed, full length SERA5 and SERA6 and a slightly processed form of SERA4 were obtained, consistent with previous findings that processing of those proteins requires the presence of *Pf*SUB1(Yeoh *et al.*, 2007; Blackman, 2008; Ruecker *et al.*, 2012).

5.4.2 Recombinant merozoite proteins are biochemically active

Evidence that recombinant merozoite proteins were active and correctly folded was provided by a series of experiments. Twenty out of 21 (~95%) merozoite cell surface proteins tested, were immunoreactive against hyperimmune sera obtained from previously malaria experienced individuals (Fig. 5.5). In 18 out of the 20 immunoreactive proteins (~86% of the total proteins tested) the immunoreactivity decreased, or was completely lost, when these proteins were heat-treated prior to probing with hyperimmune sera. These results suggest that antibodies contained in hyperimmune sera recognised conformational epitopes on the merozoite proteins, suggesting correct protein folding.

Another experiment that provided evidence for the correct conformation and activity of merozoite proteins was that RIPR was capable of interacting with RH5 by AVEXIS. Because only active and correctly folded proteins are expected to specifically interact, the observation that RIPR interacts with RH5 *in vitro* suggests

that RIPR has the correct conformation. Indications that this interaction happens *in vivo* was first provided by Chen and colleagues (Chen *et al.*, 2011), but there are a few important questions that are yet to be answered. It would be interesting to derive the kinetic and affinity parameters of this interaction and to test how the interaction between RH5 and RIPR (Chen *et al.*, 2011) influences the interaction between RH5 and BSG (Crosnier *et al.*, 2011). Experiments to further characterise the interaction between RH5-RIPR were hampered by the low expression yields of RIPR. Attempts to recombinantly express the two naturally occurring RIPR halves (Chen *et al.*, 2011) did not improve expression levels.

Finally, evidence for the correct folding of the recombinant merozoite proteins were provided by the observation that *Pf*EBA-165 was able to interact with a number of glycans on AVEXIS, suggesting that folded domains do exist on *Pf*EBA-165.

5.4.3 PF13_0125, PFA0135w, Prolactin and P4HB participated in the three interactions with the highest z-scores

Three interactions (PF13_0125 – Prolactin, PF13_0125 – P4HB, PFA0135w – P4HB) with z-scores higher than 1.5 were observed in the AVEXIS screen where the recombinant merozoite protein library was systematically screened against an equivalent protein library of erythrocyte receptors. While no information was found in the literature about PF13_0125, apart from that is localised apically in the mature merozoite (Hu *et al.*, 2010), PFA0135w (also known as MaTrA, for merozoite associated tryptophan-rich antigen) has been shown to co-localize with MSP-1 in segmented schizonts and free merozoites (Ntumngia *et al.*, 2004). It contains a tryptophan-rich domain which is also found in the *P. falciparum* paralogues TrpA-3 (PF10_0026), TryThrA (PF08_0003) and LysTrpA (MAL13P1.269) (Ntumngia *et al.*, 2005). MaTrA has orthologues in *P vivax, P. yoelii and P.knowlesi* (Burns *et al.*, 1999, 2000; Uhlemann *et al.*, 2001; Jalah *et al.*, 2005; Ntumngia *et al.*, 2005; Bora *et al.*, 2013). Synthetic peptides spanning the PFA0135w orthologue TryThrA sequence, have been shown to be highly active in binding human erythrocytes and were able to inhibit erythrocyte invasion (Curtidor *et al.*, 2006).

Interestingly, recombinant PFA0135w tryptophan rich domain, expressed in *E. coli,* was specifically recognised by IgG serum antibodies from previously *P. falciparum* exposed individuals living in Lambarene, Gabon, an area of high malaria

transmission (Ntumngia *et al.*, 2004). However, out of 135 serum samples tested, only 30 (23%) contained antibodies against the recombinant protein. These results are in agreement with my observation that full length recombinant PFA0135w exhibits relatively low immunoreactivity against pooled hyperimmune sera obtained from Malawians (Fig. 5.5), suggesting that PFA0135w is not under significant immune pressure.

On the other hand prolactin is an all α-helix protein hormone, secreted by the anterior pituitary gland (Teilum *et al.*, 2005). More than 300 biological functions have been assigned to prolactin, all mediated by prolactin receptor (Bole-Feysot *et al.*, 1998). Apart from the binding to prolactin receptor, an extracellular interaction between prolactin and cyclophilin B has also been described (Rycyzyn *et al.*, 2000).

Finally P4HB (also known as Prolyl 4-hydroxylase subunit beta or p55) is a protein disulfde isomerase and is found both in the endoplasmic reticulum (ER) and at the cell surface. Its function on the cell surface is not well understood (Wilkinson and Gilbert, 2004). Of note, is that P4HB has been linked with the events that mediate HIV entry into human leukocytes. It has been shown to that P4HB is clustered together with CD4 on human $CD4^+$ lymphocytes during HIV entry (Fenouillet *et al.*, 2001). Although P4HB is not necessary for the initial adherence of HIV (Barbouche *et al.*, 2003), it is indispensable from the fusion of the viral envelope with the host cell membrane (Fenouillet *et al.*, 2001; Gallina *et al.*, 2002; Barbouche *et al.*, 2003, 2005; Markovic *et al.*, 2004).

Structurally, P4HB harbours two thioredoxin-like domains and it interacts with a number of proteins where it catalyzes the reduction, formation and rearrangement of disulfide bonds (Wilkinson and Gilbert, 2004). P4HB has been shown to exhibit both chaperone (prevents protein aggregation) and anti-chaperone (facilitates protein aggregation) activity, depending on its concentration (Mezghrani *et al.*, 2000; Wilkinson and Gilbert, 2004). Because there is no apparent consensus sequence that P4HB recognises (Wilkinson and Gilbert, 2004) it can theoretically interact with a range of different proteins which carry disulfide bonds. This might be a reason why two putative interactions in my screen involved P4HB. Other independent protein screens in our laboratory (Dr Yi Sun, unpublished), also found P4HB to interact with numerous proteins. It is likely that a number of those interactions were specific, without though implying their existence in nature.

The interaction between P4HB and PF13_0125 was reversible in AVEXIS (Fig. 5.12). This interaction was saturable by SPR, only when P4HB was used as ligand and PF13_0125 as analyte. PF13_0125 remained bound to P4HB and did not dissociate, even after injecting a regeneration solution to facilitate dissociation (Fig. 5.13). The latter observation, together with the fact that PF13_0125 was consistently eluted from the gel filtration column much earlier than expected (Fig. 5.13), suggests that PF13_0125 analyte could have been aggregated. However, it is not yet clear why PF13_0125 demonstrated saturable binding to P4HB. Moreover, the reason why P4HB, when used as analyte, bound more to the Cd4 reference than PF13_0125 still remains an enigma.

In contrast to the interaction between PF13_0125 and P4HB, the interaction between PF13_0125 and Prolactin could not be repeated in the reciprocate bait-prey orientation (Fig. 5.12). Similarly, the interaction between PFA0135w and P4HB could only barely be detected when PFA0135w was used as prey and P4HB as bait. None of the latter two interactions was reproducible by SPR, suggesting that they may be false positive hits of AVEXIS. AVEXIS has a general low false-positive rate, and a stringent prey activity is thought to be appropriate for the testing of most interaction pairs by AVEXIS (Bushell *et al.*, 2008). However, because the stringency range varies significantly from protein pair to protein pair under investigation, inevitably some false positives and negatives do exist (Fig. 5.14). Therefore it is not unlikely that some of the interactions I observed are simply false positives.

5.5 Conlcusions

In this chapter I described the recombinant expression of 23 putative *P. falciparum* merozoite cell surface and secreted proteins. Of 23 *P. falciparum* merozoite recombinant proteins, 19 were expressed at usable amounts, and a band corresponding to full length recombinant protein was detected by immunoblot. Biochemical characterisation of the recombinant proteins provided evidence that the the majority of merozoite ligands were biochemically active and correctly folded.

The *P. falciparum* recombinant protein library was then systematically screened against an equivalent library consisting of erythrocyte receptors, by using AVEXIS

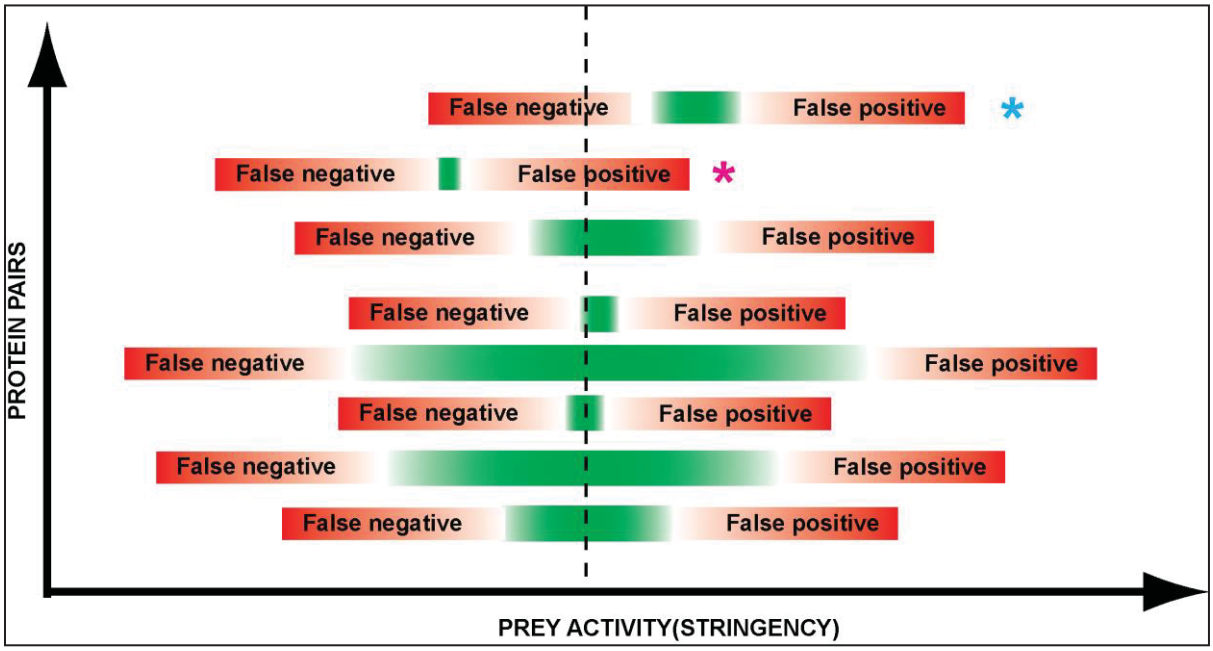

Figure 5.14 The false positive and negative rate in AVEXIS is dependent on prey activity. A diagram showing the prey activity in relation to the stringency of each protein pair tested for interaction in AVEXIS. Prey activity (dashed line) is chosen to fall within the stringency range (green) of most protein pairs. Nevertheless, because the stringency range varies significantly from protein pair to protein pair, inevitably some false positives (magenta asterisk) and negatives (cyan asterisk) arise.

(Bushell *et al.*, 2008). The screen identified one putative interaction (PF13_0125 – P4HB). Further characterisation of the identified interaction provided inconclusive results, and more experiments are required to confirm its validity. However, the expanded library of recombinant *P. falciparum* merozoite proteins reported in this project, should prove to be a useful tool for the deeper understanding of erythrocyte invasion, and *P. falciparum* merozoite biology.