4 *P. VIVAX* RECOMBINANT PROTEIN LIBRARY

Publication note: portions of the introduction, results (4.2.1, 4.2.2, 4.2.4.2, 4.2.5), and discussion were slightly modified from a published manuscript (Hostetler et al., 2015). I drafted the text in these sections, which was edited by co-authors prior to publication. I am solely responsible for the work described in this chapter under the supervision of my PhD supervisors, Rick Fairhurst and Julian Rayner, except where noted in the text.

4.1 Introduction

A vaccine targeting *P. vivax* will be an essential component of any comprehensive malaria elimination program, but relatively few *P. vivax* antigens have been investigated as vaccine candidates [e.g., (Perera et al., 1998, Malkin et al., 2005, Valderrama-Aguirre et al., 2005, Castellanos et al., 2007, Devi et al., 2007, Wu et al., 2008, Herrera et al., 2011, Mizutani et al., 2014, Yadava et al., 2014)]. This is at least partly due to the difficulty in studying *P. vivax*, which lacks a reliable *in vitro* culture system, and consequently there is limited knowledge about the molecular mechanisms of *P. vivax* merozoite invasion and antimalarial immunity. Reticulocyte invasion represents an essential, completely extracellular part of the asexual blood stage and would make a sensible target for prioritizing new vaccine candidates. The understanding of the interactions required for erythrocyte invasion in *P. vivax* lags significantly behind that of *P. falciparum*. Currently, 5 ligand-receptor interactions are known in *P. falciparum*

(Figure 1.5) including Glycophorin A-EBA175, Glycophorin B-EBL-1, Glycophorin C-EBA140, Complement receptor 1-RH4, and Basigin-RH5 (Camus and Hadley, 1985, Sim, 1995, Mayer et al., 2009, Maier et al., 2002, Crosnier et al., 2011). This is in stark contrast to only 1 known ligand-receptor interaction in *P. vivax*, that between *P. vivax* Duffy Binding Protein (DBP) and the Duffy Antigen Receptor for Chemokines (DARC, also referred to as the Duffy antigen or Fy).

Due to its central role in *P. vivax* invasion, DBP remains a primary target for vaccine development and the only *P. vivax* candidate currently in clinical trials (Table 1.1). DARC binding is mediated by a cysteine-rich region II domain (DBP-II or DBP-RII), which is polymorphic and under immune pressure (VanBuskirk et al., 2004a, Cole-Tobian et al., 2002, Batchelor et al., 2014, Batchelor et al., 2011), leading to strain-specific antibody responses (Ceravolo et al., 2009). Thus, while DBP remains a high-priority target, these complicating factors and previous experience with monovalent *P. falciparum* blood-stage vaccines, which have not thus far provided protection, suggest that a multivalent vaccine that generates cumulative immune responses should be explored. For such a multi-target *P. vivax* vaccine to be generated, a comprehensive and systematic approach is needed to better understand *P. vivax* invasion and identify additional vaccine candidates.

To address this knowledge gap, we initiated a large-scale study of *P. vivax* merozoite proteins that are potentially involved in reticulocyte binding and invasion. Several publications facilitated the assembly of a *P. vivax* merozoite protein library. The primary resource was the 2008 publication of the *P. vivax* Sal 1 reference genome (Carlton et al., 2008), which provided gene model annotation for over 5000 genes and the identification of over 3700 *P. vivax* and *P. falciparum* ortholog pairs. This publication identified 30 proteins with predicted GPI-anchors (Figure 4.1), several of which are known to localize to the merozoite surface and/or invasive secretory organelles in *P. falciparum* orthologs. A subset of proteins with predicted GPI-anchors are part of the 6-cys gene family, members of which are leading vaccine candidates in other stages of the life cycle, for instance, P48/45 and P230 (transmission-blocking) and P36 and P36p (pre-erythrocytic) (Ishino et al., 2005, van Dijk et al., 2005, van Dijk et al., 2010). At least two 6-cys proteins are expressed in each life stage, including P12, P12p, P41, and P38 during the asexual blood stages. *P. vivax* apparently has an additional 6-cys member not found in *P. falciparum* (PVX_001015) (Figure 4.1).

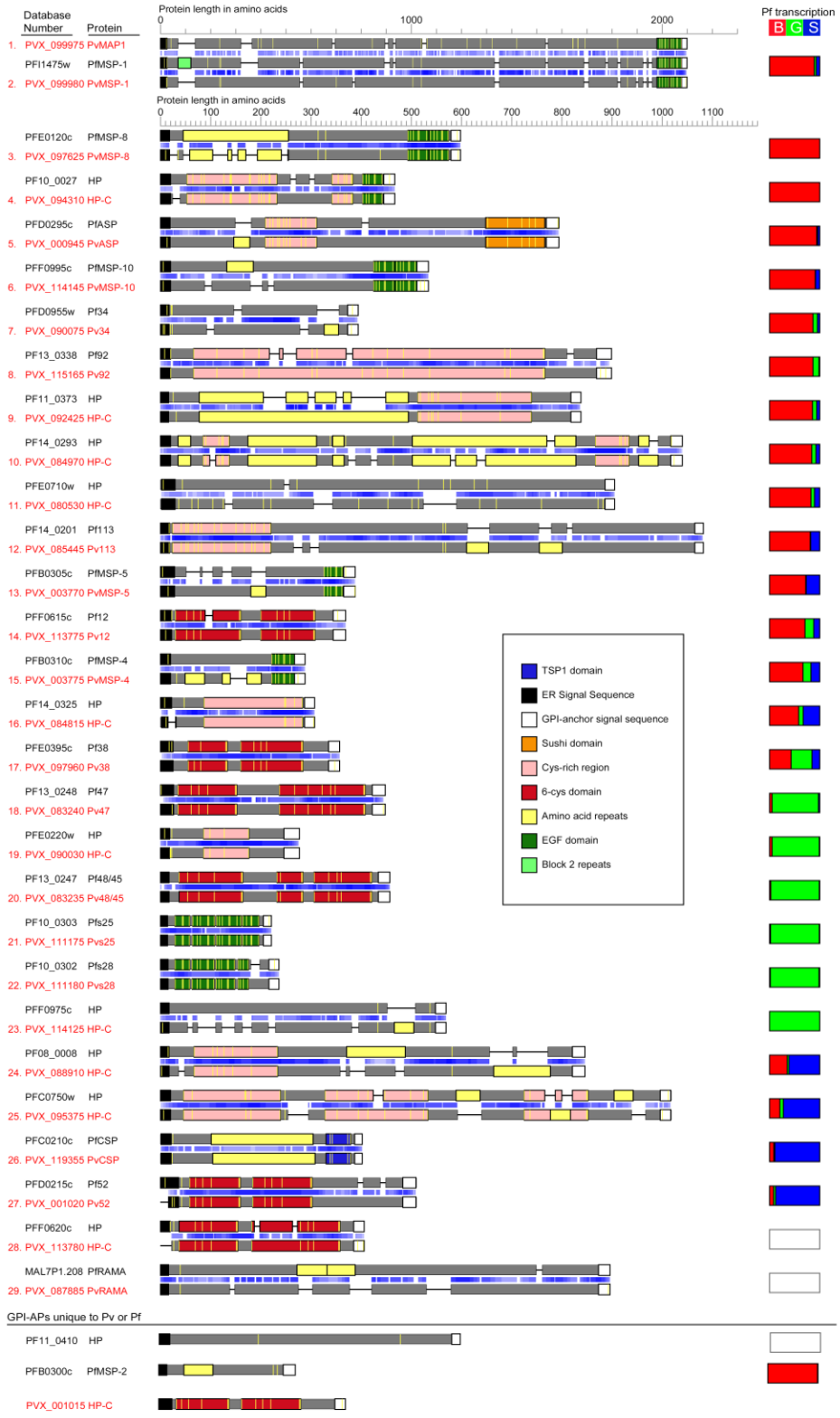

Figure 4.1: GPI-anchored *P. vivax* proteins

Alignments between *P. falciparum* (black gene IDs) and *P. vivax* (red gene IDs) orthologs with predicted GPI anchors, with amino acid similarity proportional to the intensity of blue shading. Depictions include predicted domains (colored boxes), cysteine residues (yellow bars), and microarray expression data for *P. falciparum* (rectangles to the right). HP: hypothetical; HP-C: hypothetical conserved. Reprinted by permission from Macmillan Publishers Ltd: Nature (Carlton et al., 2008) © 2008.

In addition to the reference genome, there are 2 published microarray datasets for *P. vivax* (Bozdech et al., 2008, Westenberger et al., 2010). The Westenberger *et al*. dataset contained average expression data for the asexual blood stages, while the Bozdech *et al*. study spanned 9 time points over the intraerythrocytic development cycle (IDC) of *P. vivax*. These studies, combined with the expression data we generated for 4 schizont-enriched Cambodian clinical isolates (presented in Chapter 3), provide the expression data needed to identify and prioritize candidates with potential involvement in merozoite invasion (i.e., with peak expression in the late blood-stages).

Multiple systems have now been used to express panels of *Plasmodium* proteins (Mehlin et al., 2006, Vedadi et al., 2007, Doolan et al., 2008, Tsuboi et al., 2008, Crompton et al., 2010, Trieu et al., 2011, Chen et al., 2010, Crosnier et al., 2013), and all have their strengths and weaknesses. The wheat germ extract and *E. coli in vitro* protein expression systems are scalable, but lack the context of a complete eukaryotic secretory system including some post-translational modifications, though the lack of *N*-linked glycosylation is helpful in the case of *Plasmodium*, where it is thought to occur rarely or not at all (contrary to most eukaryotes) (Gowda and Davidson, 1999). The mammalian human embryonic kidney, HEK293E, system is more medium-throughput, but is well suited to large proteins, may better represent native post-translational modifications, and has already demonstrated its utility in *Plasmodium* immunoepidemiological and functional studies (Crosnier et al., 2011, Osier et al., 2014, Taechalertpaisarn et al., 2012). This system was applied to uncover the *P. falciparum* RH5 interaction with the erythrocyte receptor basigin, an essential interaction for invasion in all parasite isolates tested (Crosnier et al., 2011), and a library of over 40 *P. falciparum* proteins has now been successfully expressed (Crosnier et al., 2013). Based on these initial successful studies in *P. falciparum* protein expression, we selected the mammalian HEK293E system for large-scale *P. vivax* recombinant protein expression.

The Cell Surface Signalling Laboratory led by Gavin Wright at WTSI was actively developing the methods for expressing *Plasmodium* proteins in the HEK293E expression system with several publications describing protocols for expression and high-throughput screening (Bushell et al., 2008, Crosnier et al., 2013, Kerr and Wright, 2012). Support and training by members of the Wright laboratory were essential to developing the *P. vivax* recombinant protein library. Full-length protein ectodomains, often synthesized by an external company (e.g., GeneArtAG, Germany) were inserted into derivatives of the

pTT3 expression vector (Durocher et al., 2002) including biotinylated 'bait' constructs and/or pentamerized 'prey' constructs (Figure 4.2). The biotinylation amino acid recognition sequence in bait protein constructs allows the proteins to be enzymatically biotinylated when transiently co-transfected with a plasmid containing the BirA enzyme in expression media containing free biotin. This enables bait proteins to be bound to streptavidin-coated surfaces (i.e., plates or beads) for high-throughput screens without the need for initial purification, which renders large interaction screens more feasible. Both baits and prey constructs contain C-terminal rat Cd4d3+d4 domains, which serve as tags to check for protein expression, as polyclonal antibodies against most *Plasmodium* proteins are not available for such purposes, whereas commercial antibodies to this domain are readily available. The C-terminal cartilage oligomeric matrix protein (COMP) in the β-lactamase enzyme-tagged prey constructs enable ectodomains to be pentamerized, which has been shown to increase the avidity with potential interacting partners in downstream screens (Bushell et al., 2008).

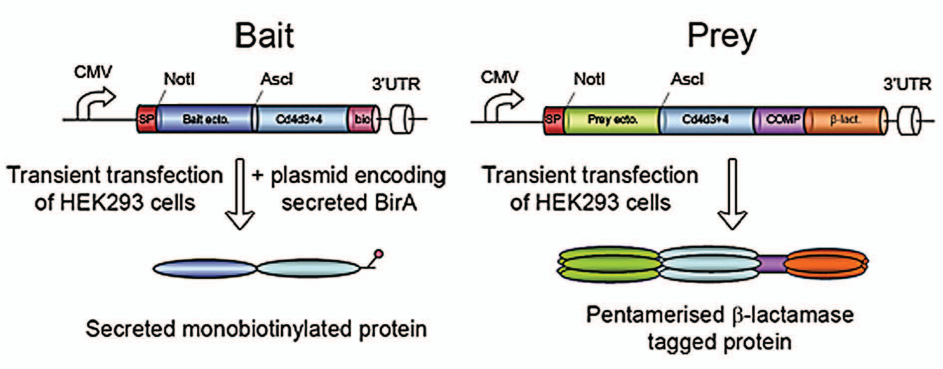

Figure 4.2: Bait and prey protein constructs

Schematic representations of constructs for expression in the mammalian HEK293E cell expression system. Both construct ectodomain (ecto.) regions contain flanking Not1 (5') and Asc1 (3') cut sites and a C-terminal rat Cd4d3+d4 terminal tag (~25 kDa). Bait proteins contain a C-terminal biotinylation recognition sequence that leads to enzymatic mono-biotinylation when transiently co-transfected with free biotin and a BirA enzyme-expressing plasmid. Prey proteins contain a C-terminal pentamerization sequence from the cartilage oligomeric matrix protein (COMP) and the β-lactamase enzyme leading to the pentamerization of prey ectodomains. Figure modified from (Kerr and Wright, 2012).

Once expressed, the *P. vivax* recombinant protein library can be utilized to uncover potential invasion ligands through erythrocyte binding experiments. Such experiments can be performed using several methods. One approach is to employ a cell-based binding assay. Most commonly, this involves transfecting COS7 cells to express membrane-bound protein fragments of interest, followed by incubation with erythrocytes. Binding is detected through counting the clusters ("rosettes") of erythrocytes surrounding the COS7 cells. This method was used to confirm that region II of *P. vivax* DBP and *P. falciparum* EBA175 was sufficient for binding to erythrocytes (Chitnis and Miller, 1994, Sim et al., 1994). However, this method involves individual experiments per protein and manual counting, equating to significant labor and time to evaluate several dozen proteins in a larger scale library. A flow cytometry-based assay published by Tran et al. in 2005 detected the DBP-RII interaction with erythrocytes using a His-tagged recombinant DBP-RII and a fluorescently-labeled anti-His antibody (Tran et al., 2005). This would enable high-throughput screening in plate formats, but requires His-purified proteins at high concentrations (~50 µg/ml), also a significant limitation for screening a protein library.

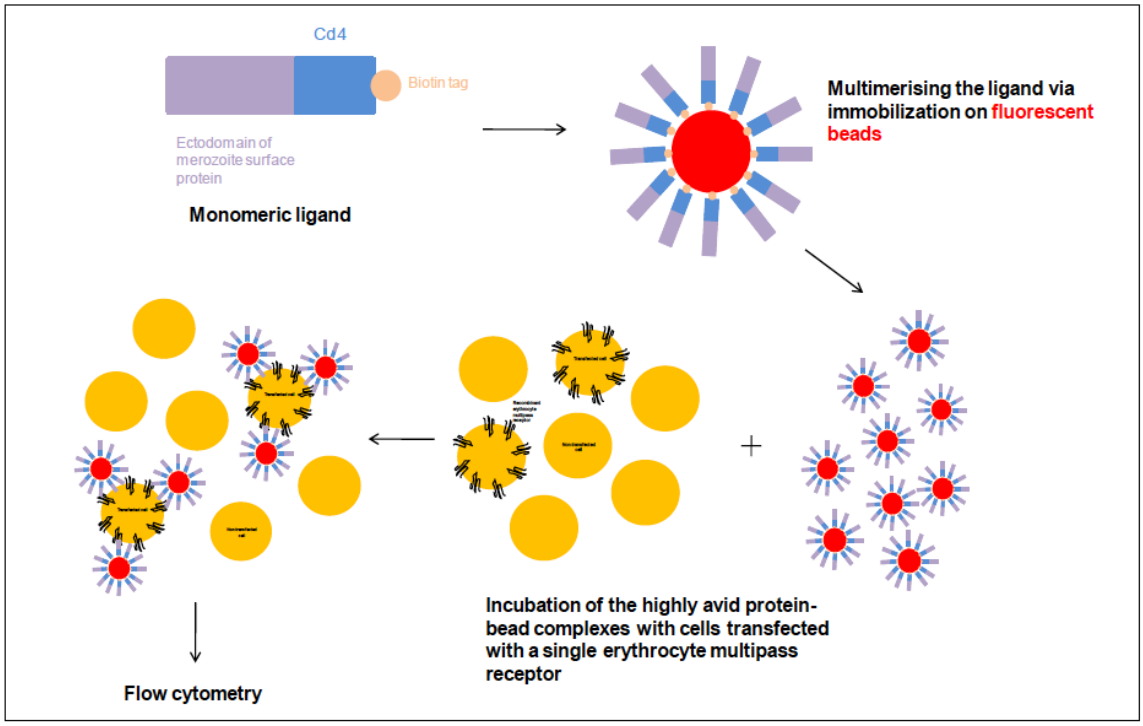

Figure 4.3: High-throughput bead-based interaction screening method
Plate-based interaction screening method developed in Gavin Wright's laboratory. Mono-biotinylated recombinant proteins are bound to streptavidin-coated beads and incubated with cells. Binding events are detected via a fluorescence shift by flow cytometry. Figure provided by Madushi Wanaguru.

Madushi Wanaguru, in Gavin Wright's laboratory, optimized a flow-cytometry based method for detecting binding based on previous methods in the laboratory (Wright et al., 2000). This method involved binding biotinylated recombinant proteins to streptavidin-coated fluorescent beads and incubating them with cells expressing multi-pass proteins in order to identify specific receptor-ligand interactions (Figure 4.3). The recombinant proteins were multimerized on the beads, potentially increasing the avidity of interactions with erythrocyte receptors, which would prolong the usually transient cell-surface interactions and facilitate their detection. Early testing showed significant fluorescence shifts when incubating erythrocytes with recombinant *P. falciparum* EBA175 and EBA140, which was ultimately published (Crosnier et al., 2013). This high-throughput approach for testing erythrocyte binding would be useful to evaluate a *P. vivax* recombinant protein library in a 96-well format.

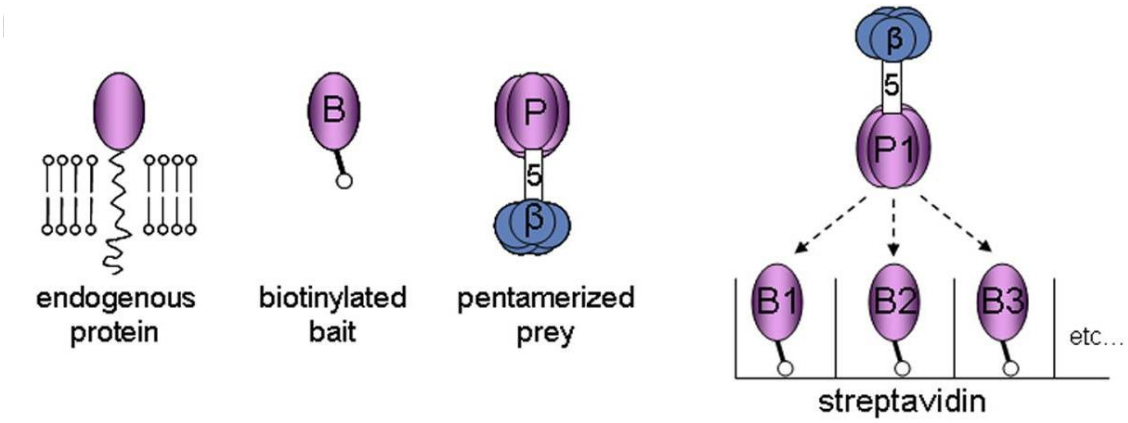

Figure 4.4: Overview of AVEXIS

Depiction of AVEXIS showing biotinylated bait proteins bound to streptavidin-coated plates. Baits are probed by pentamerized prey proteins tagged with ß-lactamase. Interactions are detected by nitrocefin turnover resulting in a color change from yellow to red. This figure was originally published in (Martin et al., 2010), © the American Society for Biochemistry and Molecular Biology.

The Wright and Rayner laboratories had also developed a library of over 40 proteins abundant on the surface of mature erythrocytes, which led to the detection of the essential *P. falciparum* RH5 interaction with the erythrocyte receptor basigin (Crosnier et al., 2011). This library was comprised of single-pass erythrocyte surface proteins, expressed in soluble form as full-length ectodomain fragments (Crosnier et al., 2011). The largest surface-exposed ectodomain of several multipass membrane proteins, were also included. We hypothesized that several uncharacterized merozoite-erythrocyte protein interactions

are involved in *P. vivax* recognition and invasion of reticulocytes, and planned to perform large-scale interaction screens to test for specific binding partners to the recombinant *P. vivax* merozoite protein library. Gavin Wright's laboratory had developed a large-scale screening method for detecting specific receptor-ligand interactions called AVidity-based EXtracellular Interaction Screen or AVEXIS (Figure 4.4). The assay reliably detects highly-transient interactions with half-lives ≤0.1 seconds (Bushell et al., 2008). This transiency is a common feature of interactions between cell surface proteins (van der Merwe et al., 1994, Dustin and Springer, 1991), and it is plausible that initial *P. vivax* merozoite-reticulocyte interactions will be similarly transient. We plan to use this assay to study *P. vivax* erythrocyte invasion by screening the library of *P. vivax* merozoite surface and invasion antigens against the existing erythrocyte surface protein library. This method can additionally be used to screen for intra-library parasite protein-protein interactions, which may be essential to merozoite function and invasion.

4.1.1 Benefits of these studies

Research into *P. vivax* invasion of reticulocytes lags significantly behind *P. falciparum*, partly due to the lack of an *in vitro* culture system and the difficulty in expressing recombinant *Plasmodium* proteins. Our studies aim to address this need and benefit the scientific community in several ways. First, we aim to assemble a library of *P. vivax* antigens known or predicted to localize to the merozoite surface and/or be involved in erythrocyte invasion. Second, we plan to screen this library in large protein-protein interaction screens and erythrocyte-binding assays to determine protein function and potentially prioritize proteins for further development as vaccine candidates. Lastly, we plan to make this protein library publicly available to facilitate future functional and immunoepidemiological studies.

4.1.2 Objectives

i. To generate a library of *P. vivax* merozoite proteins and express them in a mammalian expression system

ii. To investigate protein library function in high-throughput binding and interaction screens

4.2 Results

4.2.1 *P. vivax* merozoite library candidate selection

I incorporated the RNA-Seq data from Chapter 3, existing *P. vivax* microarray data (Westenberger et al., 2010, Bozdech et al., 2008), and homology comparisons to *P. falciparum* to identify candidate merozoite proteins to evaluate in downstream protein interaction screens. This search concentrated on proteins known or predicted to be exported to the merozoite surface, micronemes, or rhoptries (often based on homology with *P. falciparum* proteins) and/or whose transcripts are most abundant during schizogony. A library of 39 *P. vivax* merozoite proteins was identified, and constructs were designed and synthesized for expression in HEK293E cells (GeneArt AG, Germany, see section 2.3.1 for greater detail). These 39 proteins were subdivided into 4 groups Table 4.1.

Table 4.1: *P. vivax* recombinant merozoite proteins

Group	Accession #	Name	Product	Size (kDa)	Pf3D7 homologs /ortholog group (OG)	Subcell. location	GPI anchor (y/n)	Expr level
MSPs[1]	PVX_ 099980	MSP1	merozoite surface protein 1	215	PF3D7_ 0930300	Merozoite surface[8]	y	low
	PVX_ 097670	MSP3.1 MSP3A[2] MSP3γ	merozoite surface protein 3	114	None[6]		n	high
	PVX_ 097680	MSP3.3 MSP3C[2] MSP3β	merozoite surface protein 3	133	None[6]	Merozoite surface[9]	n	high
	PVX_ 097685	MSP3.4 MSP3D1	merozoite surface protein 3	140	OG5_ 126854		n	high
	PVX_ 097720	MSP3.10 MSP3H[2] MSP3α	merozoite surface protein 3	113	None[6]	Merozoite surface[9]	n	med
	PVX_ 003775	MSP4[2]	merozoite surface protein 4, putative	45	PF3D7_ 0207000	Merozoite surface[10]	y	med
	PVX_ 003770	MSP5[2]	merozoite surface protein 5	62	PF3D7_ 0206900	Microneme and or/Apical[10]	y	high

	PVX_082700	MSP7.1[3]	merozoite surface protein 7 (MSP7)	70	None		n	low
	PVX_082675	MSP7.6[3]	merozoite surface protein 7 (MSP7)	72	PF3D7_1334300		n	med
	PVX_082655	MSP7.9[3]	merozoite surface protein 7 (MSP7), putative	65	OG5_151950	Merozoite surface[7, 11]	n	low
	PVX_114145	MSP10[2]	merozoite surface protein 10, putative	72	PF3D7_0620400	Merozoite surface[12, 13]	y	high
6-cys proteins	PVX_113775	P12	6-cysteine protein	61	PF3D7_0612700	Merozoite surface[13], Rhoptry[14]	y	high
	PVX_113780	P12p[2]	6-cysteine protein	68	PF3D7_0612800		n	med
	PVX_097960	P38[2]	6-cysteine protein	60	PF3D7_0508000	Merozoite surface and/or Apical[7, 15]	y	low
	PVX_000995	P41	6-cysteine protein	67	PF3D7_0404900	Merozoite surface[16]	n	med
	PVX_115165	P92	6-cysteine protein	118	PF3D7_1364100	Merozoite surface[7, 15]	y	low
Mero. proteins not in other families	PVX_001725	RON12	rhoptry neck protein 12, putative	54	PF3D7_1017100	Rhoptry[7, 17]	n	med
	PVX_088910	GAMA	GPI-anchored micronemal antigen, putative	103	PF3D7_0828800	Merozoite surface[7, 18]	y	high
	PVX_090075	Pv34[2]	apical merozoite protein[2]	61	PF3D7_0419700	Rhoptry[7, 19]	y	high
	PVX_090210	ARP	asparagine-rich protein	54	PF3D7_0423400	Rhoptry[7, 20]	n	high
	PVX_090240	CyRPA	cysteine-rich protective antigen, putative	65	PF3D7_0423800	Apical[7, 21]	y[25]	med
	PVX_095055	RIPR	Rh5 interacting protein,	144	PF3D7_0323400	Microneme and merozoite	n	very low

			putative			surface[7, 22]		
	PVX_098712	RhopH3	high molecular weight rhoptry protein 3, putative	125	PF3D7_0905400	Rhoptry	n	low
	PVX_110810	DBP	Duffy receptor precursor	135	OG5_148188	Microneme [10, 23]	n	low
		DBP-RII	Duffy binding protein region II	64				low
	PVX_111290	MTRAP	merozoite TRAP-like protein, putative	58	PF3D7_1028700	Microneme[7, 24]	n	high
	PVX_121885	CLAG[4]	cytoadherence linked asexual protein, CLAG, putative	159	OG5_138272	Rhoptry[7]	n	very low
	PVX_081550		StAR-related lipid transfer protein, putative	80	PF3D7_0104200		n	low
	PVX_084815		conserved *Plasmodium* protein, unknown function[2]	53	PF3D7_1434400		y	very low
	PVX_084970		conserved *Plasmodium* protein, unknown function[2]	121	PF3D7_1431400		y	low
	PVX_116775		conserved *Plasmodium* protein, unknown function[2]	60	PF3D7_1321900		n	low
No known *Pf* 3D7 homolog	PVX_101590	RBP2-like[4]	reticulocyte-binding protein 2 (RBP2), like	97	None		n	med
	PVX_001015		6-cysteine protein, putative[5]	62	None		y	low
	PVX_110945		hypothetical protein	61	None		n	low
	PVX_		conserved	66	None		n	low

	110950	*Plasmodium* protein, unknown function[2]				
	PVX_ 110960	hypothetical protein	92	None	n	low
	PVX_ 110965	conserved *Plasmodium* protein, unknown function[2]	69	None	n	low

Proteins and their Plasmodb.org accession numbers are classified into 4 groups, based on whether they are members of the MSP family, the 6-cysteine family, neither family and have a known *P. falciparum* homolog, or neither family and have no known *P. falciparum* homolog. Common names and products are listed based on existing *P. vivax* annotation. The expected size includes a ~25-kDa C-terminal rat Cd4d3+d4 tag. *P. falciparum* accession numbers are listed when only 1 homolog exists and OrthoMCL clusters are listed when multiple homologs exist. Subcellular locations are based on existing *P. vivax* literature or predictions based on existing *P. falciparum* literature, where noted. The presence of a glycosylphosphatidylinositol (GPI) anchor sequence is indicated (Carlton et al., 2008). Expression levels are based on an estimated 0.3-0.5 μg/ml required to saturate biotin binding sites on plates (Osier et al., 2014, Kerr and Wright, 2012). Levels are listed as a guide, as significant batch-to-batch variability was observed. Groups include "high" (>5 μg/ml), "med" for medium (0.5-5 μg/ml), and "low" (<0.5 μg/ml) expressors, as defined in section 2.3.3. Accession links to plasmodb.org, amino acid lengths and boundaries of full-length ectodomains are listed in Supplementary Table D.

[1]MSP9 and MSP1P failed to be sub-cloned or expressed, respectively

[2]Based on genedb.org annotation

[3]Based on MSP7 family naming in (Kadekoppala and Holder, 2010)

[4]Based on *P. vivax* product description

[5]6-cysteine protein upstream of P52 and P36 in *P. vivax* but not *P. falciparum*

[6]*P. vivax* and *P. falciparum* MSP3s have the same numeric designation, but do not appear to be orthologs (Rice et al., 2014) or fall into the same OrthoMCL clusters

[7]Based on localization in *P. falciparum*

[8](del Portillo et al., 1988, del Portillo et al., 1991)

[9](Jiang et al., 2013)

[10](Black et al., 2002)

[11](Kadekoppala and Holder, 2010, Kadekoppala et al., 2010)

[12](Perez-Leal et al., 2005, Giraldo et al., 2009)

[13](Moreno-Perez et al., 2013)

[14](Li et al., 2012)

[15](Sanders et al., 2005)

[16]Cheng et al., 2013

[17](Knuepfer et al., 2014)

[18](Arumugam et al., 2011)
[19](Proellocks et al., 2007)
[20](Wickramarachchi et al., 2008)
[21](Dreyer et al., 2012)
[22](Chen et al., 2011)
[23](Adams et al., 1990, Black et al., 2002)
[24](Baum et al., 2006)
[25](Reddy et al., 2015)

Merozoite surface proteins. Merozoite surface proteins (MSPs) are primarily exposed on the plasma membrane of the invasive merozoite, and several *P. falciparum* MSPs have been tested as vaccine candidates. We selected a total of 13 MSPs, 8 of which have known subcellular localizations in *P. vivax* (del Portillo et al., 1988, Jiang et al., 2013, Black et al., 2002, Perez-Leal et al., 2005, Moreno-Perez et al., 2013, del Portillo et al., 1991, Giraldo et al., 2009, Oliveira-Ferreira et al., 2004, Cheng et al., 2014) that are similar to those of their *P. falciparum* homologs – except MSP5, which has a micronemal and/or apical distribution in *P. vivax* but a merozoite surface distribution in *P. falciparum* (Black et al., 2002). GPI-anchor sequences were predicted for MSP5 (Carlton et al., 2008) and both predicted and supported by merozoite surface localization experiments in *P. vivax* for MSP1, MSP4, and MSP10 (del Portillo et al., 1988, del Portillo et al., 1991, Black et al., 2002, Perez-Leal et al., 2005, Moreno-Perez et al., 2013, Giraldo et al., 2009, Carlton et al., 2008). GPI anchors were experimentally confirmed in *P. falciparum* homologs for MSP1, MSP4, and MSP5 (Sanders et al., 2005). *P. vivax* has 12 members of the MSP3 family, which are not clear homologues of the *P. falciparum* MSP3 family but may serve similar functions (Rice et al., 2014); 4 representative MSP3s were selected for expression in this library. Similarly 3 MSP7s were selected to represent the 11 MSP7 family members, which Kadekoppala and Holder (Kadekoppala and Holder, 2010) have named sequentially in the order they occur on chromosome 12, beginning with PVX_082700 as MSP7.1 and ending with PVX_082650 as MSP7.11. Despite repeated attempts, we were unable to express the MSP1 paralog, MSP1P, or to sub-clone MSP9 into the expression vector.

6-cysteine proteins. 6-cysteine proteins are defined by a characteristic arrangement of cysteine residues and are expressed at multiple stages of the *P. falciparum* life cycle; several members are considered stage-specific *P. falciparum* vaccine candidates. We selected 5 of 13 *P. vivax* 6-cysteine proteins, 4 of which are most homologous to the *P. falciparum* proteins known to be involved in blood-stage parasite development (Sanders

et al., 2005). Transcriptional data suggest that P12p may also be involved in the blood-stage development of *P. berghei* parasites (van Dijk et al., 2010). GPI anchors were predicted for 3 proteins (P12, P38, and P92) (Carlton et al., 2008), all of which were experimentally confirmed in *P. falciparum* orthologs (Sanders et al., 2005). The predicted GPI anchor in P12 is supported by a merozoite surface and/or rhoptry localization in *P. vivax* (Moreno-Perez et al., 2013, Li et al., 2012).

Merozoite proteins not in other families. Twenty other proteins, which do not belong to the aforementioned *P. vivax* MSP or 6-cysteine protein families, were selected because they are known or suspected to localize to the merozoite surface, micronemes, or rhoptries. Localization to these regions is predicted based on *P. falciparum* homologs (Knuepfer et al., 2014, Arumugam et al., 2011, Proellocks et al., 2007, Wickramarachchi et al., 2008, Dreyer et al., 2012, Chen et al., 2011, Baum et al., 2006) and several have predicted GPI anchors (GAMA, Pv34, PVX_084815, PVX_084970, and PVX_001015) (Carlton et al., 2008). *P. falciparum* CyRPA was also recently determined to have a GPI anchor (Reddy et al., 2015). These 20 proteins fall into 2 groups, with 14 (including PvDBP) having a *P. falciparum* homolog or ortholog, and 6 having neither. The latter 6 proteins include a reticulocyte binding protein (RBP2-like, PVX_101590) that is much smaller than other RBPs, which are typically > 250 kDa and therefore unlikely to express well in the HEK293E cell system; a hypothetical 6-cysteine protein (PVX_001015) that is located upstream from the 6-cysteine proteins P52 and P36 at a microsyntenic breakpoint on *P. vivax* chromosome 3, and has a predicted GPI-anchor (Carlton et al., 2008); and 4 hypothetical proteins (Frech and Chen, 2011) that are syntenic to an MSP-encoding region in *P. falciparum* but have no *P. falciparum* orthologs. Little is known about the hypothetical proteins in our library, except that their transcripts are most abundant in early-to-late schizonts [(Bozdech et al., 2008), Chapter 3 RNA-Seq data] and 1 of them (PVX_110960) contains a merozoite SPAM domain (Pfam accession number, PF07133).

4.2.2 *P. vivax* merozoite protein library expression in HEK293E cells

The expression of 95% (37/39) of the biotinylated full-length protein ectodomains was confirmed by an ELISA that detected the rat Cd4d3+d4 tag (Figure 4.5); 34 of these proteins were also detected by western blot analysis (Figure 4.6). The sizes of expressed proteins ranged from 45 to 215 kDa (which includes the ~25-kDa rat Cd4d3+d4 tag), and generally conform to predicted sizes (Table 4.1).

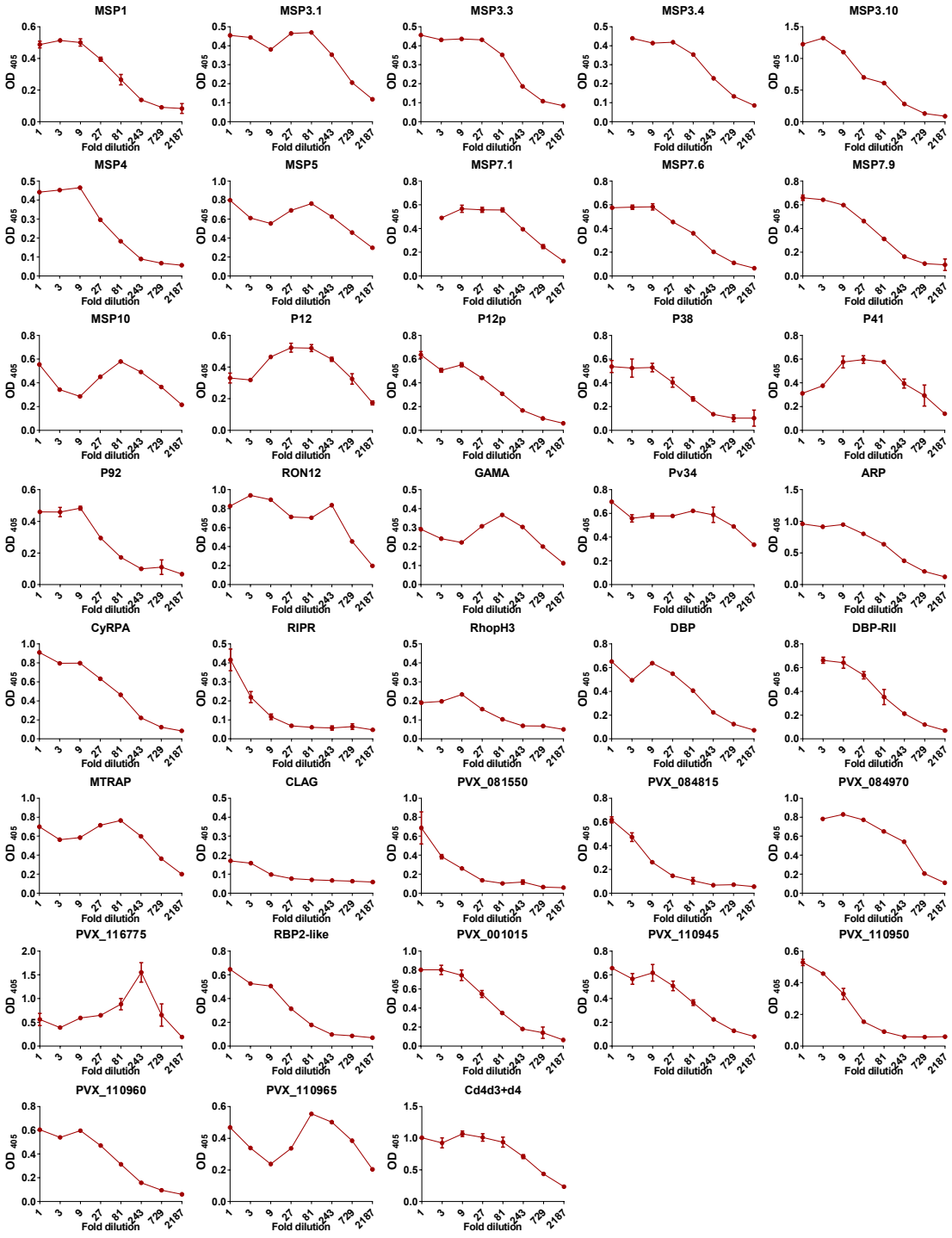

Figure 4.5: *P. vivax* recombinant protein expression detected by ELISA

Three-fold serial dilutions of 37 mono-biotinylated *P. vivax* recombinant proteins were bound to streptavidin-coated plates, and C-terminal Cd4d3+d4 tags were detected using OX68 antibody (primary) and alkaline phosphatase-conjugated anti-mouse antibody (secondary), with alkaline phosphatase activity measured as an increase in optical density (OD) at 405 nm. Mean ± standard deviation; n=3.

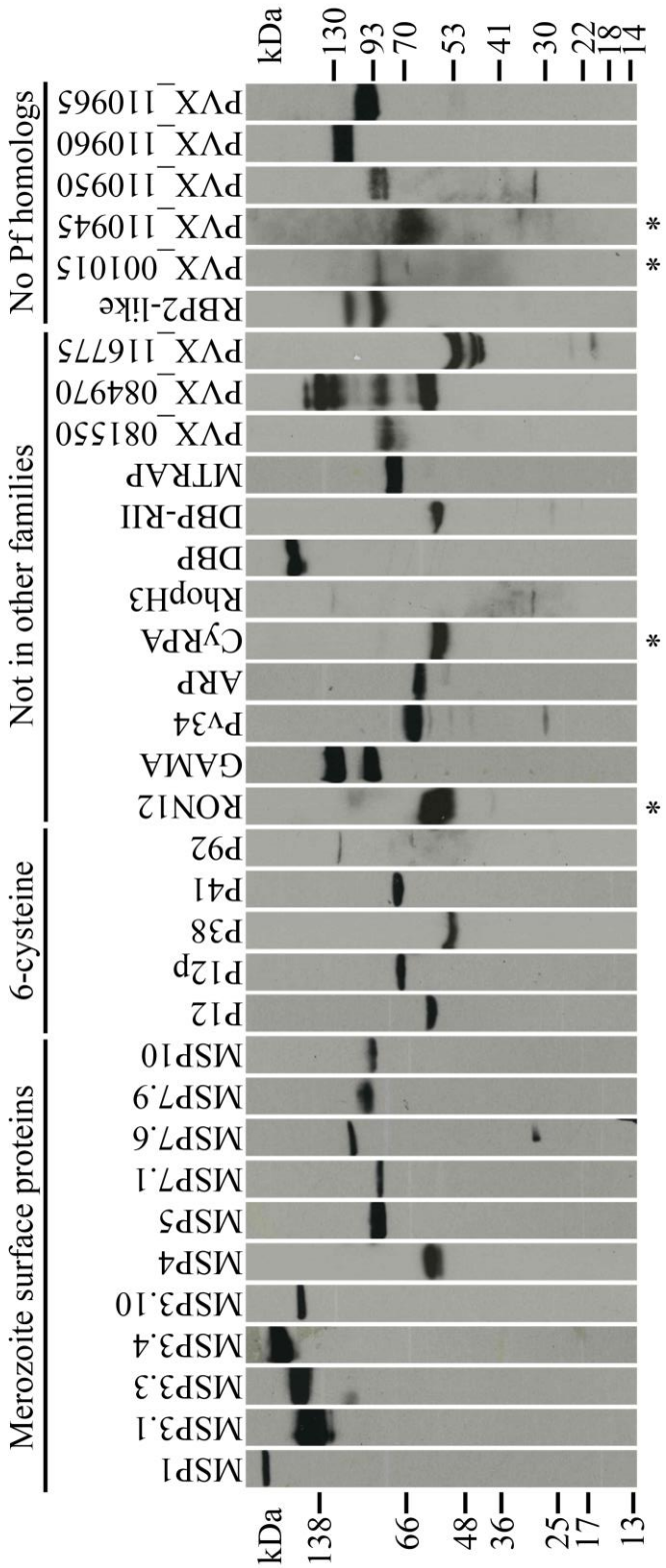

Figure 4.6: Western blot analysis confirms expression of 34/37 *P. vivax* recombinant proteins.

Biotinylated proteins were resolved by SDS-PAGE under reducing conditions, blotted, and probed using streptavidin-HRP. All proteins contain a ~25-kDa rat Cd4d3+4 tag. (*) indicates proteins that were run with the right ladder; all others were run with the left or both ladders. Reprinted from (Hostetler, 2015) under the Creative Commons Attribution (CC BY) license.

Several proteins (e.g., GAMA, RBP2-like) showed multiple bands, suggesting they were proteolytically processed, as was seen for some recombinant proteins in an analogous *P. falciparum* library (Crosnier et al., 2013). Protein expression varied widely: 44% (17/39) of the *P. vivax* bait proteins expressed well at over 0.5 μg/ml (see "high" and "medium" in section 2.3.3), and 51% (20/39) of proteins had low expression (<0.5 μg/ml). Most could be effectively concentrated and used in downstream screens. Proteins CLAG, RIPR, and PVX_084815 showed very low expression and were detected by ELISA only.

4.2.3 Erythrocyte binding experiments by flow cytometry

I tested the *P. vivax* recombinant protein library in 96-well format for binding to erythrocytes using a method optimised by Madushi Wanaguru (Wright Laboratory), who provided significant training. The method successfully detected binding of recombinant EBA175 and EBA140 to erythrocytes (Crosnier et al., 2013). The method involved binding recombinant biotinylated *Plasmodium* proteins to streptavidin-coated fluorescent beads (Nile red) and subsequently incubating them with erythrocytes, with binding indicated by a fluorescence shift by flow cytometry. I expressed the complete library as mono-biotinylated proteins in HEK293E cells as described in 2.3.3, and 34/37 proteins showed some expression by dilution series ELISA detection of the Cd4d3+d4 tag (Figure 4.7). Proteins RIPR, RhopH3, and CLAG had no detectable expression while 9 additional proteins had an immediately linear slope in the dilution series, indicating the proteins did not saturate the streptavidin binding sites on plates and were considered to have low protein levels (<0.5 μg/ml after concentrating). The proteins with low expression included MSP1, RON12, PVX_081550, PVX_084815, PVX_084970, PVX_116775, PVX_001015, PVX_110950, and PVX_110960. 34 proteins were multimerized by direct attachment to streptavidin-coated Nile red beads based on the initial ELISA results (excluding CLAG, RIPR, RhopH3, PVX_081550, PVX_110950). I determined the minimum amount of protein needed for saturating the beads by pre-incubating dilutions of protein with uniform quantities of Nile red beads, and assessing the remaining unbound protein by ELISA (Figure 4.7). This was important as any free protein (not bound to beads) in the assay would potentially inhibit binding of the protein-coated beads and reduce the ability to detect erythrocyte binding. Similarly, too little initial protein would result in beads which were not saturated, thereby decreasing the avidity of binding to erythrocytes and again reducing the ability to detect binding.

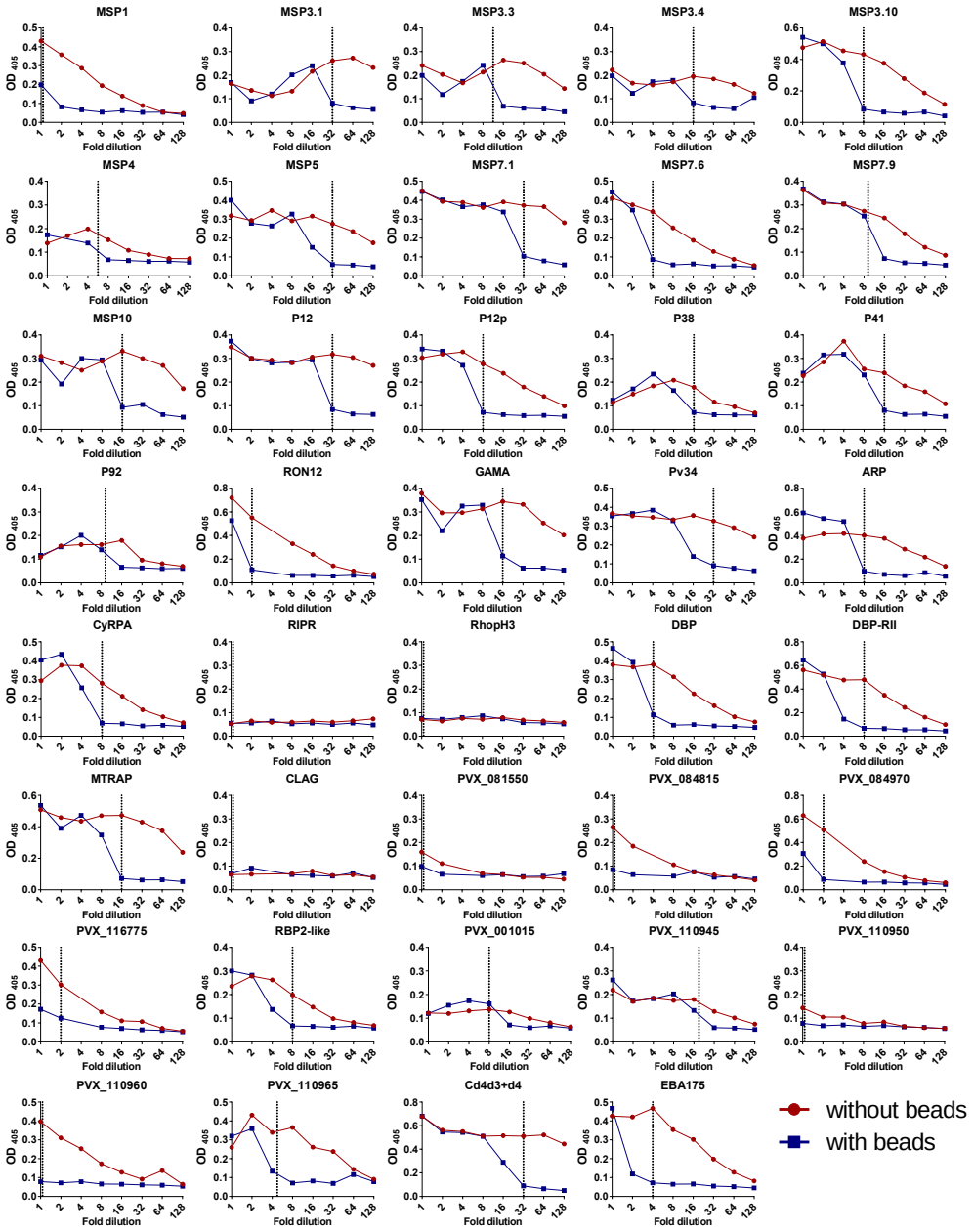

Figure 4.7: *P. vivax* recombinant library expression and bead saturation assay by ELISA

Two-fold serial dilutions of 37 mono-biotinylated *P. vivax* recombinant proteins with (blue line) or without (red line) pre-incubation with streptavidin-coated Nile red beads were bound to streptavidin-coated plates, and C-terminal Cd4d3+d4 tags were detected using OX68 antibody (primary) and alkaline phosphatase-conjugated anti-mouse antibody (secondary), with alkaline phosphatase activity measured as an increase in optical density (OD) at 405 nm. Dotted lines represent the dilution needed for complete saturation of beads with minimal unbound protein. Any unbound protein after pre-incubation with beads, was subsequently bound to the streptavidin-coated plates and resulted in higher OD values, while dilutions where all protein was bound to beads resulted in low OD values (blue lines).

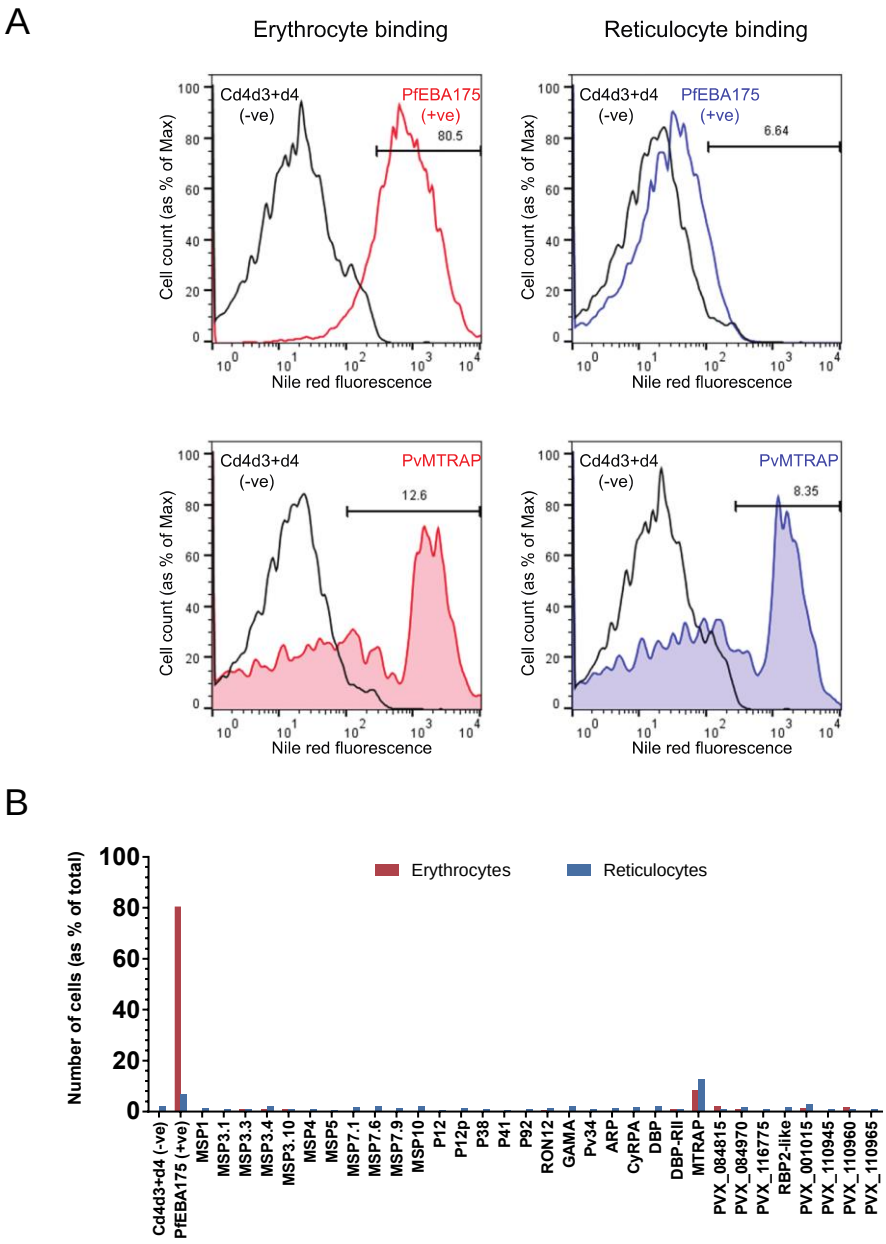

Figure 4.8: Erythrocyte and reticulocyte binding to *P. vivax* recombinant proteins

Recombinant *P. vivax* proteins (34) were multimerized on fluorescent beads (streptavidin-coated, Nile red) and incubated with either erythrocytes (left panel) or hematopoietic stem cell-derived reticulocytes (right panel). (A) Histograms showing fluorescence intensities (at the Nile red emission wavelength) and the number of erythrocytes (as a percentage of total) associated with beads with the gate for binding set based on the negative control, Cd4d3+d4 (-ve) (black). Positive control, *P. falciparum* EBA175 showed 80.5% of erythrocytes (red) and 6.6% of reticulocytes (blue) associated with beads. *P. vivax* MTRAP (lower panel, shaded) displayed some binding with 12.6% of erythrocytes (red) and 8.35% of reticulocytes (blue) associated with beads (B) number of erythrocytes (red) or reticulocytes (blue) (as a percentage of total) associated with beads for 34 *P. vivax* recombinant proteins.

Protein-saturated Nile red beads were then incubated with either erythrocytes or hematopoietic stem cell-derived reticulocytes (provided by NHS Blood and Transplant), and binding was measured using flow cytometry (Figure 4.8). A fluorescence intensity threshold was set based on Cd4d3+d4-coated beads (negative control) in order to calculate the number of erythrocytes (percentage of the total) bound to each protein. *P. falciparum* EBA175-coated beads served as a positive control and 80.5% of erythrocytes and 6.6% of reticulocytes associated with beads. The percentage of erythrocytes or reticulocytes associating with beads was very low (under 2%) for all *P. vivax* recombinant proteins tested except for MTRAP which showed 12.6% of erythrocytes and 8.35% of reticulocytes associated with beads (Figure 4.8B). *P. falciparum* MTRAP has been found to bind with erythrocyte receptor, Semaphorin-7a, so binding may be predicted for the orthologous *P. vivax* MTRAP (Bartholdson et al., 2012). The fluorescence profile however was atypical (i.e. compared to PfEBA175) with a bimodal distribution rather than a unimodal distribution. A repeat of the binding experiment for *P. vivax* MTRAP including re-expression and binding in triplicate failed in the final steps due to plate handling issues and was not able to be repeated within the timeline of this project. Future experiments will be needed to replicate the MTRAP binding results and rule out non-specific binding or bead aggregation. Overall, the assay appeared to distinguish binding in the case of positive control *P. falciparum* EBA175, which binds to glycophorin A, a highly abundant erythrocyte surface protein (10^6 copies/cell), but failed to detect any binding with receptors with low abundance (i.e., no binding was detected in the case of DBP, which bind to lower abundance surface receptor, DARC (10^4 copies/cell) (Anstee, 1990).

4.2.4 High-throughput protein interaction screens

4.2.4.1 AVEXIS between *P. vivax* recombinant proteins and an erythrocyte receptor library

The sensitivity of the bead-based erythrocyte-binding assay was limited in detecting interactions with lower abundance erythrocyte surface receptors. Thus, it was possible that binding between members of the recombinant protein library and erythrocytes would be below the limit of detection of the assay. The Wright and Rayner laboratories had recently developed a library of recombinant erythrocyte prey proteins to facilitate the screening for specific receptor-ligand interactions. This led to the discovery of the

essential *P. falciparum* RH5 interaction with basigin using AVEXIS technology (Crosnier et al., 2011). I performed a similar experiment screening the *P. vivax* recombinant protein library as baits and the erythrocyte recombinant protein library as preys using AVEXIS (Figure 4.4).

AVEXIS consists of testing for interactions between "bait" proteins, which are biotinylated and captured by the streptavidin-coated wells of a 96-well plate, and "prey" proteins, which are enzymatically tagged and contain a pentamerization domain to increase interaction avidity. Thus, I first expressed the recombinant *P. vivax* protein library as mono-biotinylated bait proteins (as described in 2.3.3). Individual protein expression was evaluated by a dilution series ELISA detecting the Cd4d3+d4 tag prior to performing AVEXIS as described in 2.4.1 (Figure 4.5). Nearly 95% (37/39) of the library showed some expression; the remaining 2 proteins either consistently showed no expression (MSP1P) or could not be subcloned in the bait expression construct (MSP9). Seven proteins showed an immediately linear slope in the dilution series, indicating the proteins did not saturate the streptavidin binding sites on plates and were considered to have low protein levels (<0.5 μg/ml after concentrating) (noted with asterisks on Figure 4.10).

Recombinant erythrocyte receptors (37) were expressed (by both Sumana Sharma and myself) as preys similarly to bait proteins as described in section 2.3.3. Prey protein activity was evaluated through the use of a β-lactamase normalization assay described in section 2.4.2. Over 80% (30/37) of preys showed optimal activity, defined as complete turnover of 60 μl at 125 μg/ml nitrocefin by β-lactamase within 20 minutes (Kerr and Wright, 2012, Bushell et al., 2008). The remaining 7 prey proteins (noted with asterisks on Figure 4.10) failed to completely turnover nitrocefin within 20 minutes and were considered to have activity below the threshold of the assay. All 37 bait and 39 prey proteins were included in the AVEXIS screen.

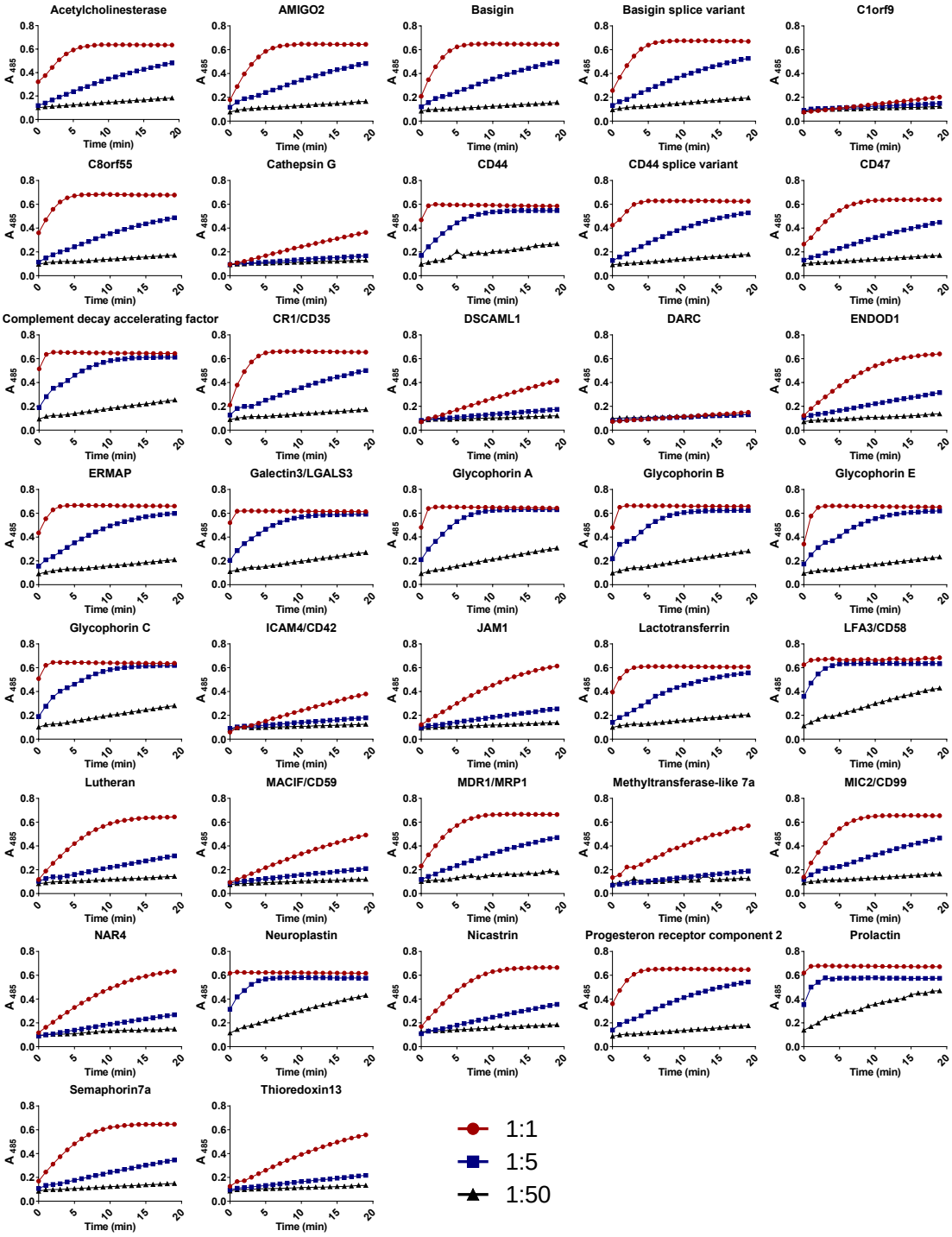

Figure 4.9: Erythrocyte prey normalization assay

Recombinant erythrocyte prey activity was assessed by incubating dilutions of prey proteins (1:1, 1:5, 1:50) with nitrocefin, with β-lactamase turnover of nitrocefin resulting in a color change from yellow to red and increasing absorbance (A) at 485 nm. Optimal activity defined as saturation of signal within 20 minutes.

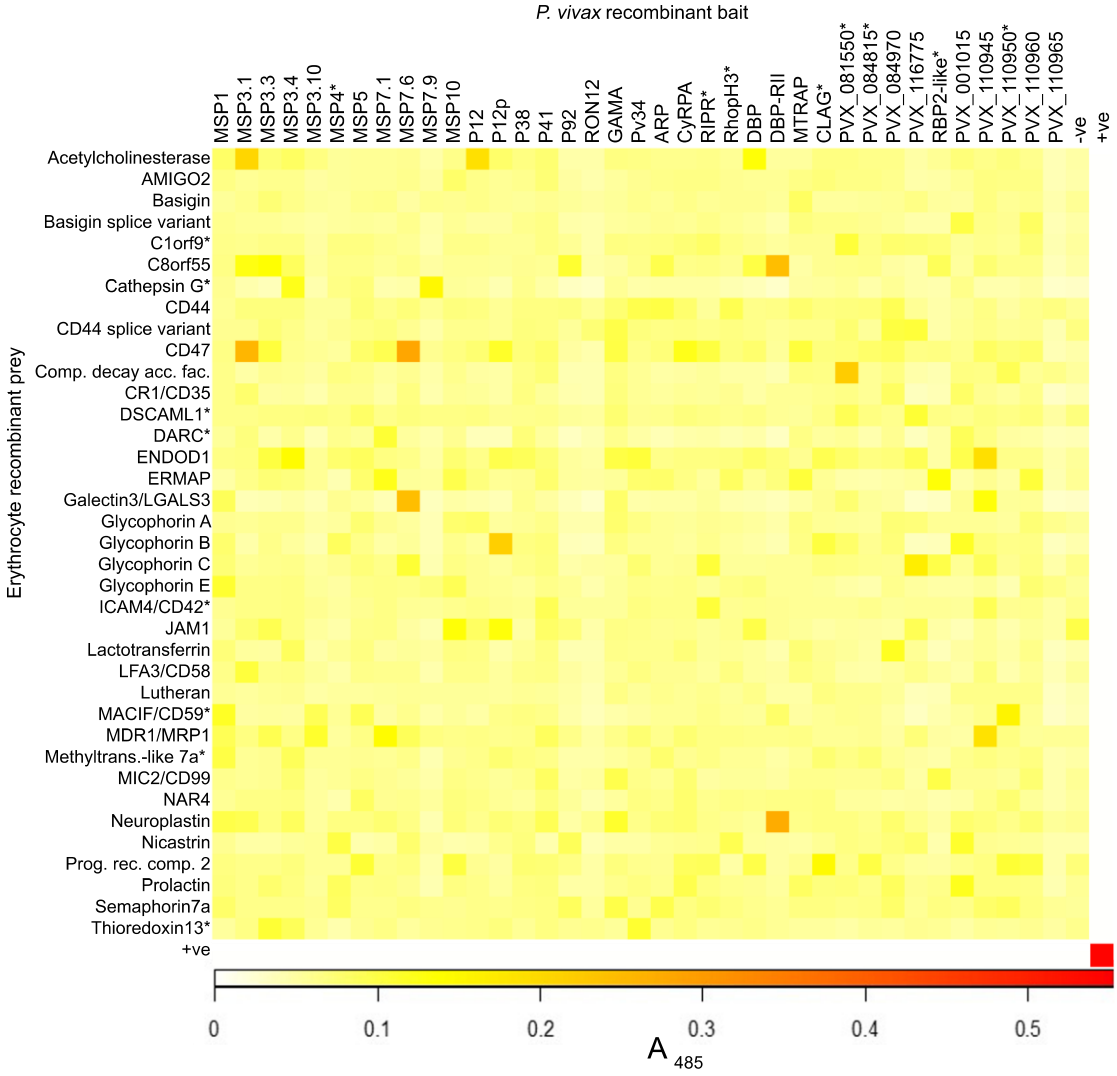

Figure 4.10: AVEXIS between *P. vivax* recombinant proteins and erythrocyte receptor library

Heat map of AVEXIS between *P. vivax* merozoite proteins as baits and erythrocyte receptor proteins as preys, with the intensity of absorbance (A) values at 485 nm after 90 min (values for Neuroplastin-DBP-RII and MSP7.6-CD47 after 18 h). No signal saturations indicating no clear interactions were detected. (*) indicates baits with low protein levels (<0.5 µg/ml after concentrating) and preys with activity below the threshold required by the assay. Positive control (+ve) is the *P. falciparum* P12-P41 interaction. Negative controls (–ve) are rat Cd4d3+d4 tag.

The AVEXIS between the *P. vivax* recombinant protein library and the erythrocyte receptor library detected no clear interactions (Figure 4.10). There were 7 wells with absorbance between 0.22 and 0.29, which was above the general background level, with mean absorbance for the assay 0.07, but well below the absorbance of the positive control (0.57), which is generally observed for validated interactions. These elevated readings

often occurred for the same prey or bait (i.e., CD47 prey and MSP7.6, DBP-RII baits), likely indicating an elevated level of non-specific interactions in these recombinant proteins. A second screen was performed using a complete set of re-expressed proteins, which also failed to detect any interactions (no signals saturated). In order to further characterize the function of proteins in the library in a high-throughput manner, I moved forward to investigating protein-protein interactions between proteins within the *P. vivax* merozoite library.

4.2.4.2 AVEXIS detects predicted and novel interactions between *P. vivax* recombinant proteins

Reflecting the general lack of knowledge about *P. vivax* biology, most of the recombinant proteins in our library have no known function, making it impossible to establish whether they recapitulate the function of native proteins. In some cases, however, protein-protein interactions between members of the library were either known or predicted based on homology to *P. falciparum*. I therefore performed AVEXIS in order to identify predicted and novel *P. vivax* protein-protein interactions. Of the 37 constructs that expressed successfully in the bait vector (Figure 4.5), 34 were successfully sub-cloned and expressed in the prey vector (Figure 4.11), though several had low expression or activity (indicated by asterisks in Figure 4.12). AVEXIS was then performed using all 37 bait and 34 prey *P. vivax* proteins.

This intra-library AVEXIS (Figure 4.12) identified 3 *P. vivax* protein-protein interactions in the bait-prey orientation: P12-P41, P12-PVX_110945, and MSP3.10-MSP7.1. The P12-P41 interaction was predicted, as the *P. falciparum* homologs of these proteins are known to form a heterodimer, an interaction that has been validated by both AVEXIS and immunoprecipitation from parasite interactions (Taechalertpaisarn et al., 2012). The P12-P41 interaction was also detected in the reciprocal prey-bait orientation (Figure 4.12). While interaction between *P. vivax* P12 and P41 is predicted based on the function of the *P. falciparum* homologues, this is the first time it has been confirmed, implying that the ability of these proteins to interact is essential for their functional activity and predates the evolutionary divergence of *P. falciparum* and *P. vivax*.

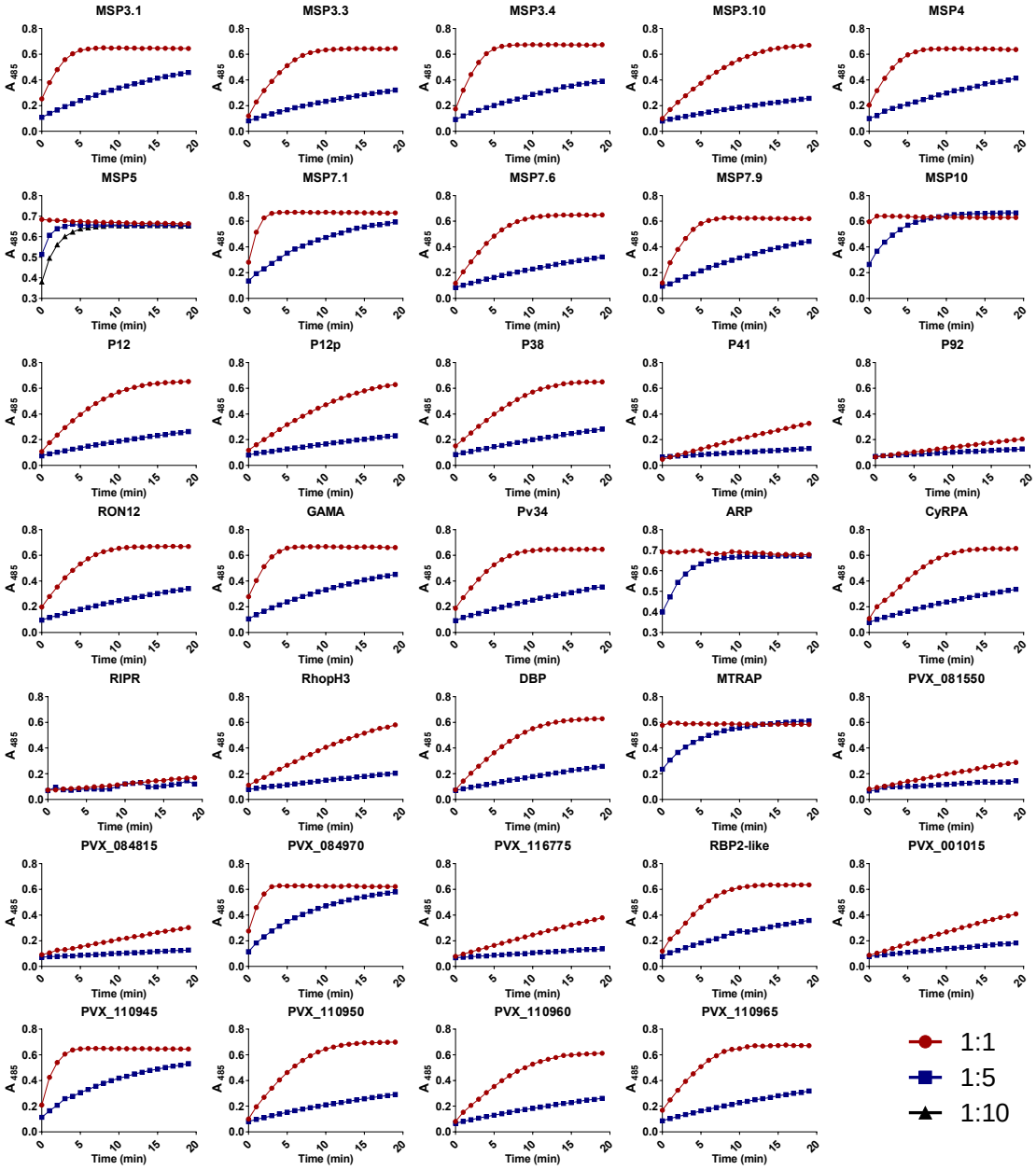

Figure 4.11: *P. vivax* recombinant prey normalization

Recombinant *P. vivax* prey activity was assessed by incubating dilutions of prey proteins (1:1, 1:5, 1:10) with nitrocefin, with β-lactamase turnover of nitrocefin resulting in a color change from yellow to red and increasing absorbance (A) at 485 nm. Optimal activity defined as saturation of signal within 20 min.

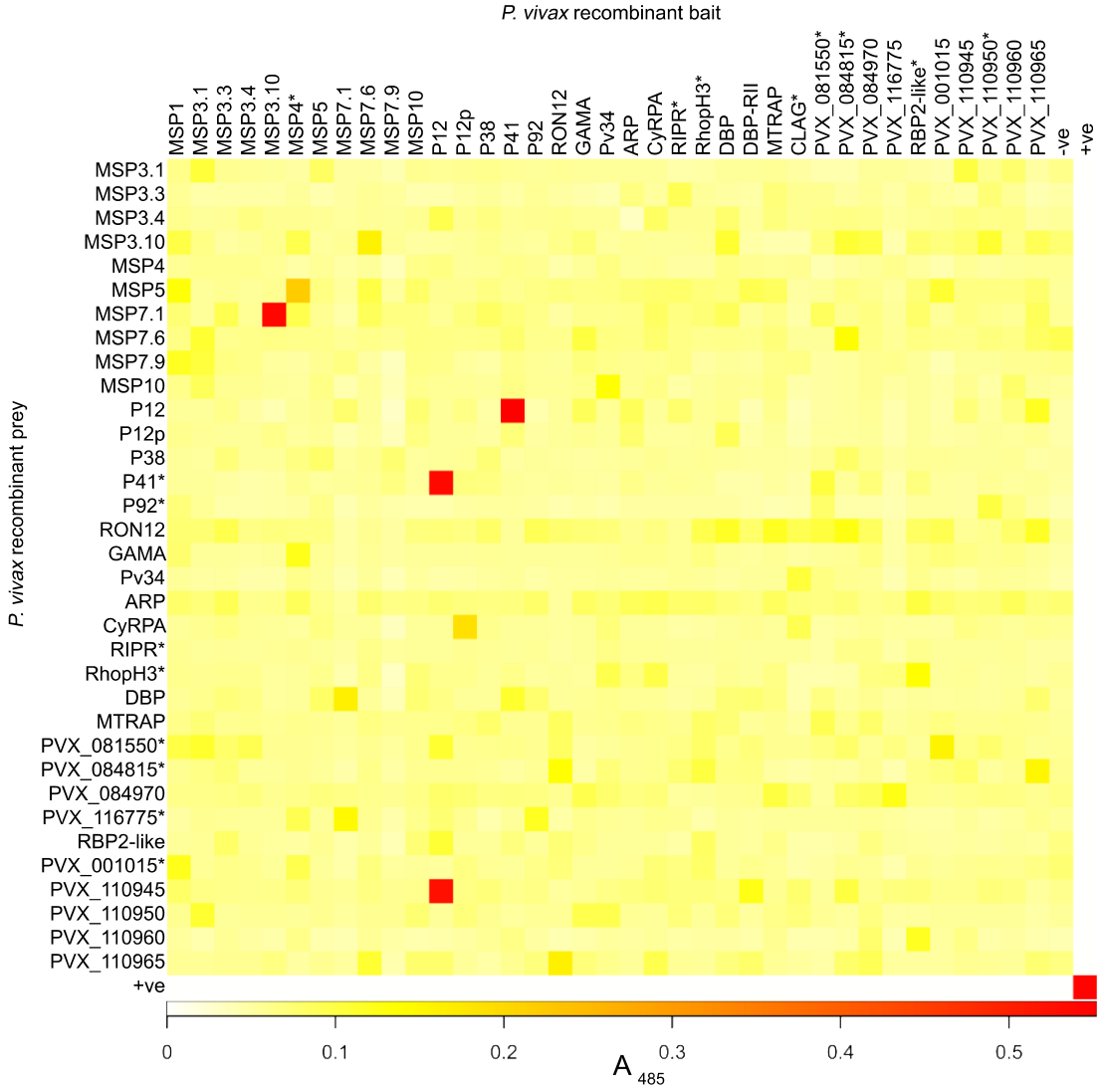

Figure 4.12: AVEXIS reveals novel interactions involving *P. vivax* recombinant proteins

Heat map of the initial *P. vivax* intra-library AVEXIS, with the intensity of absorbance (A) values at 485 nm and positive putative interactions in red: P12-P41 (bait-prey and prey-bait orientations), P12-PVX_110945 (bait-prey orientation), and MSP3.10-MSP7.1 (bait-prey orientation). (*) indicates baits with low protein levels (<0.5 µg/ml after concentrating) and preys with activity below the threshold required by the assay. Positive control (+ve) is the *P. falciparum* P12-P41 interaction. Negative controls (–ve) are rat Cd4d3+d4 tag. Figure reprinted from (Hostetler et al., 2015) under the Creative Commons Attribution (CC BY) license.

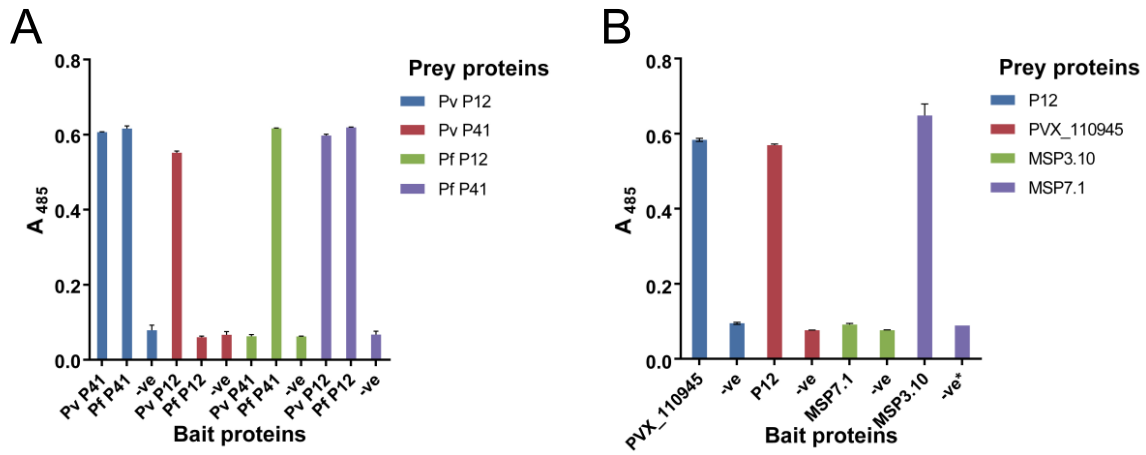

Figure 4.13: Replicated intra-library interactions with AVEXIS

Bar chart shows mean absorbance (A) at 485 nm with range; n=2 (*) indicates n=1. Replicated *P. vivax* intra-library AVEXIS using re-expressed bait proteins. (A) P12-P41 interaction within and between *P. vivax* (Pv) and *P. falciparum* (Pf) proteins by AVEXIS, which detected an interaction between Pv P12-Pv P41 and Pv P12-Pf P41 in both bait and prey orientations. (B) P12-PVX_110945 interaction confirmed in both orientations, and the MSP3.10-MSP7.1 interaction confirmed in only the bait-prey orientation. Positive control (+ve) is the *P. falciparum* P12-P41 interaction. Negative control (–ve) is the rat Cd4d3+d4 tag. Figure reprinted from (Hostetler et al., 2015) under the Creative Commons Attribution (CC BY) license.

To test whether specific amino acid components of the interaction have been conserved over the long evolutionary time frame since the *P. falciparum* and *P. vivax* divergence, I tested the ability of *P. vivax* and *P. falciparum* P12 and P41 to interact with each other. *P. vivax* P12 and *P. falciparum* P41 interacted, suggesting that conserved amino acid contacts do exist (Figure 4.13A), but *P. falciparum* P12 and *P. vivax* P41 did not interact. Investigating the sequences and structures of these proteins in more detail may improve our understanding of their interactions.

While the P12-P41 interaction was predicted, 2 novel interactions, P12-PVX_110945 and MSP3.10-MSP7.1, were also identified during the intra-library screen. Re-expression of PVX_110945 and MSP7.1 as baits instead of preys confirmed the PVX_110945-P12 interaction, but the MSP3.10-MSP7.1 interaction was not detected in the reciprocal prey-bait orientation (Figure 4.13B). The inability to recapitulate the MSP3.10-MSP7.1 interaction does not necessarily indicate that this interaction is biologically insignificant. Protein-protein interactions that are orientation-dependent have been reproducibly detected in other studies (Bushell et al., 2008, Martin et al., 2010, Sollner and Wright,

2009), and in this case may indicate a loss of activity when MSP3.10 is pentamerized in the prey vector.

4.2.5 Biophysical analysis with SPR

4.2.5.1 Size-exclusion chromatographic purification of P12, P41, and MSP7.1

In order to confirm these *P. vivax* protein-protein interactions using surface plasmon resonance (SPR), I subcloned the entire ectodomains of P12, P41, and MSP7.1 into a modified plasmid containing a 6-His tag (Bushell et al., 2008), expressed them (described in 2.3.3), and purified the proteins by immobilized metal-ion affinity chromatography (described in 2.6.1). Purified proteins were subjected to size-exclusion chromatography (SEC) immediately before use, as described in (Taechalertpaisarn et al., 2012) and section 2.6.1. In SEC, recombinant purified *P. vivax* P12 and P41 appeared to elute at apparent molecular masses of 112 and 111 kDa, which are greater than their expected sizes of 60 kDa and 66 kDa, respectively (Figure 4.14A,B). Since the apparent masses are about 2 times the expected sizes, this finding may indicate that these proteins form homodimers in solution, which is supported by gel data obtained under native conditions (Figure 4.15). *P. falciparum* P12 and P41 also eluted at higher-than-expected masses of 90 kDa and 105 kDa, respectively (Taechalertpaisarn et al., 2012). No P12-P12 or P41-P41 self-binding was observed by SPR (Figure 4.17) or AVEXIS (Figure 4.12, Figure 4.13), indicating that the proteins, whether monomers or dimers, are unable to self-associate into higher-order structures.

The recombinant purified *P. vivax* MSP7.1 eluted as a main peak with evidence of higher molecular mass forms by SEC. The main peak eluted at a much larger-than-expected molecular mass of 69 kDa, suggesting oligomerization or aggregation of the protein in solution (Figure 4.14C), which is additionally supported by gel data obtained under native conditions (Figure 4.15). Despite little sequence conservation between MSP7 family members for *P. vivax* and *P. falciparum* (Kadekoppala and Holder, 2010), this is also seen in *P. falciparum* MSP7 purifications, though with additional smaller forms present (Perrin et al., 2015). Other *Plasmodium* surface proteins, such as *P. falciparum* MSP2 and MSP3, are also known to form higher-order structures (Adda et al., 2009, Zhang et al., 2012, Gondeau et al., 2009, Imam et al., 2014).

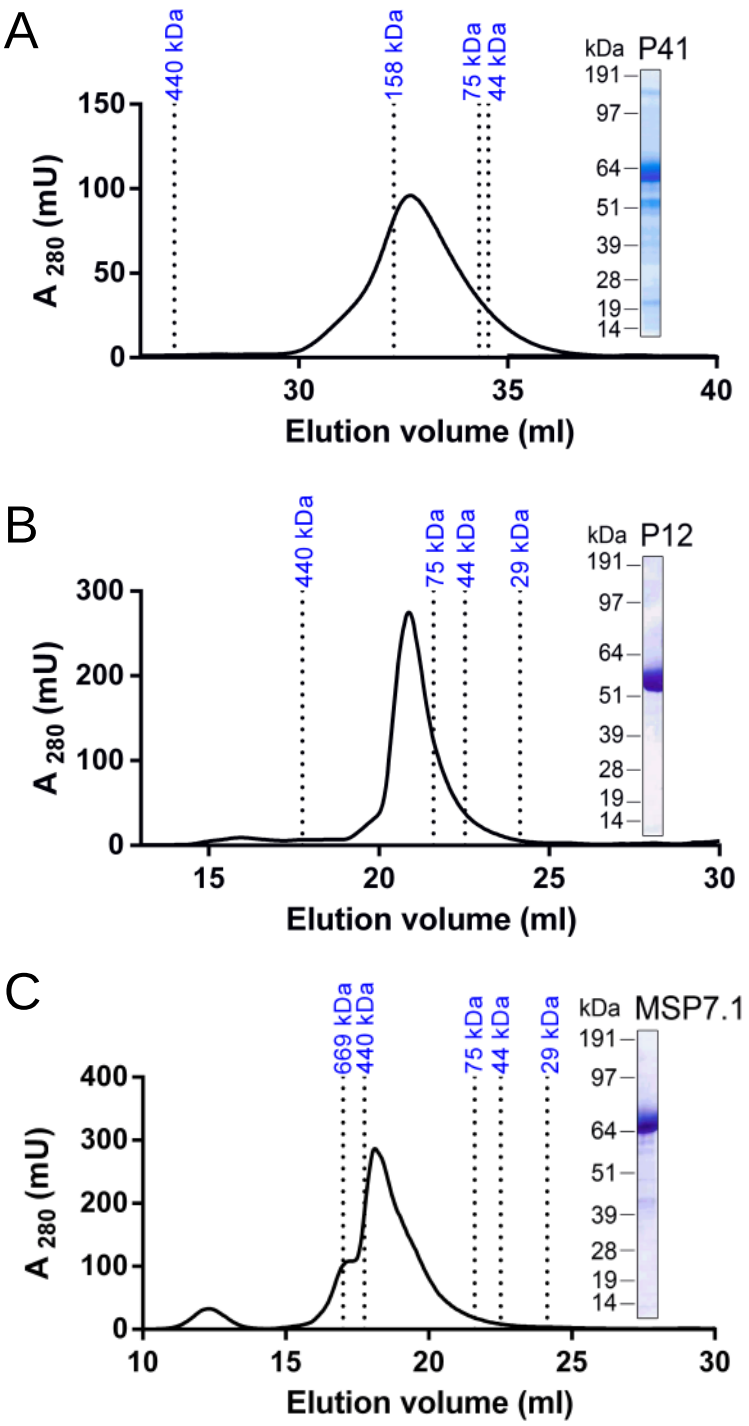

Figure 4.14: SEC for *P. vivax* P12 and P41 and MSP7.1

Recombinant 6-His-tagged *P. vivax* P12 (A) and P41 (B), each eluted as a monodisperse peak after SEC, which was resolved as a single band of the expected size by SDS-PAGE (insets). Recombinant 6-His-tagged *P. vivax* MSP7.1 eluted as a main peak with a small shoulder at higher-than-expected masses after SEC, likely due to oligomerization, and with a main band of the expected size by SDS-PAGE (inset). Figure adapted from (Hostetler et al., 2015) under the Creative Commons Attribution (CC BY) license.

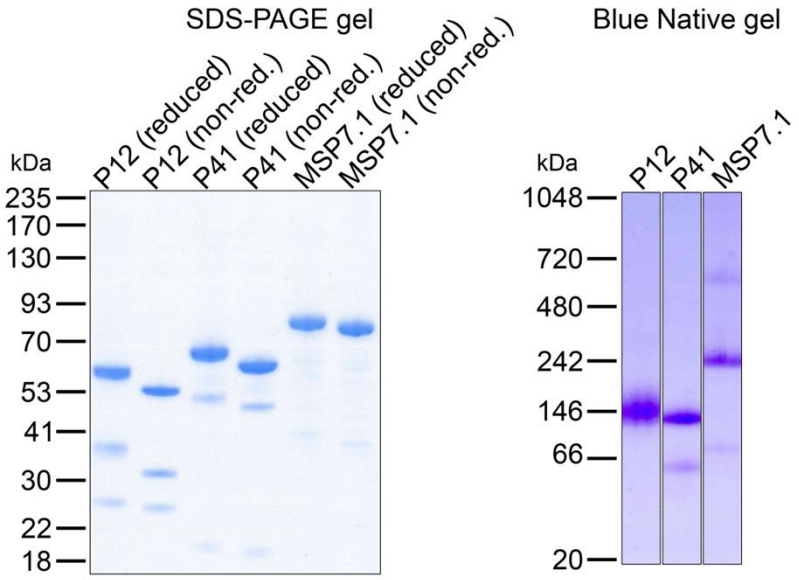

Figure 4.15: P12, P41, and MSP7.1 may exist as homodimers or oligomers

Recombinant 6-His-tagged *P. vivax* P12, P41, and MSP7.1 were fractionated by SDS-PAGE under reducing conditions and non-reducing conditions, and stained. An increase in molecular mass upon reduction suggested that the proteins contain disulphide bonds, as expected. No homodimers or oligomers are observed by SDS-PAGE. Samples run under native conditions support the presence of homodimers for P12 and P41 with bands near 146 kDa. Monomers for these 2 proteins correspond to 60 and 66 kDa, respectively. MSP7.1 shows a faint band near the expected monomer size of 69 kDa and a dominant band near 242 kDa, suggesting that this protein potentially forms an oligomer of 3 or 4 monomers. Figure reprinted from (Hostetler et al., 2015) under the Creative Commons Attribution (CC BY) license.

4.2.5.2 SPR confirms *P. vivax* P12-P41 interaction

SPR was used to validate all interactions and determine the biophysical binding parameters for the P12-P41 interaction. Equilibrium binding experiments between *P. vivax* P12-P41 showed clear evidence of saturation, thus demonstrating the interaction's specificity (Figure 4.16). However, due to limited amounts of protein, several of the lower concentrations of analyte did not achieve equilibrium in both bait-prey orientations with the consequence that the calculated equilibrium dissociation constants (K_D) are likely to be slightly overestimated; that is, the interaction has a higher affinity. When *P. vivax* P12 was used as the purified 6-His-tagged analyte and *P. vivax* P41 as the immobilized biotinylated ligand, the K_D was 120 ± 10 nM, and when the proteins were used in the reverse orientation the K_D was 77 ± 6 nM. The *P. vivax* P12-P41 interaction

therefore has an affinity that is at least 3 times higher than the *P. falciparum* P12-P41 interaction, with the *P. vivax* P12-P41 K_D of <100 nM being much lower than the *P. falciparum* P12-P41 K_D of 310 nM (Taechalertpaisarn et al., 2012). Binding between *P. falciparum* and *P. vivax* homologues that had been detected using AVEXIS was confirmed by SPR using *P. vivax* P12 as the purified 6-His-tagged analyte and *P. falciparum* P41 as the immobilized biotinylated ligand (Figure 4.16C). The K_D of 31 ± 10 nM for this interaction also suggests that it has a much higher affinity than the *P. falciparum* P12-P41 interaction.

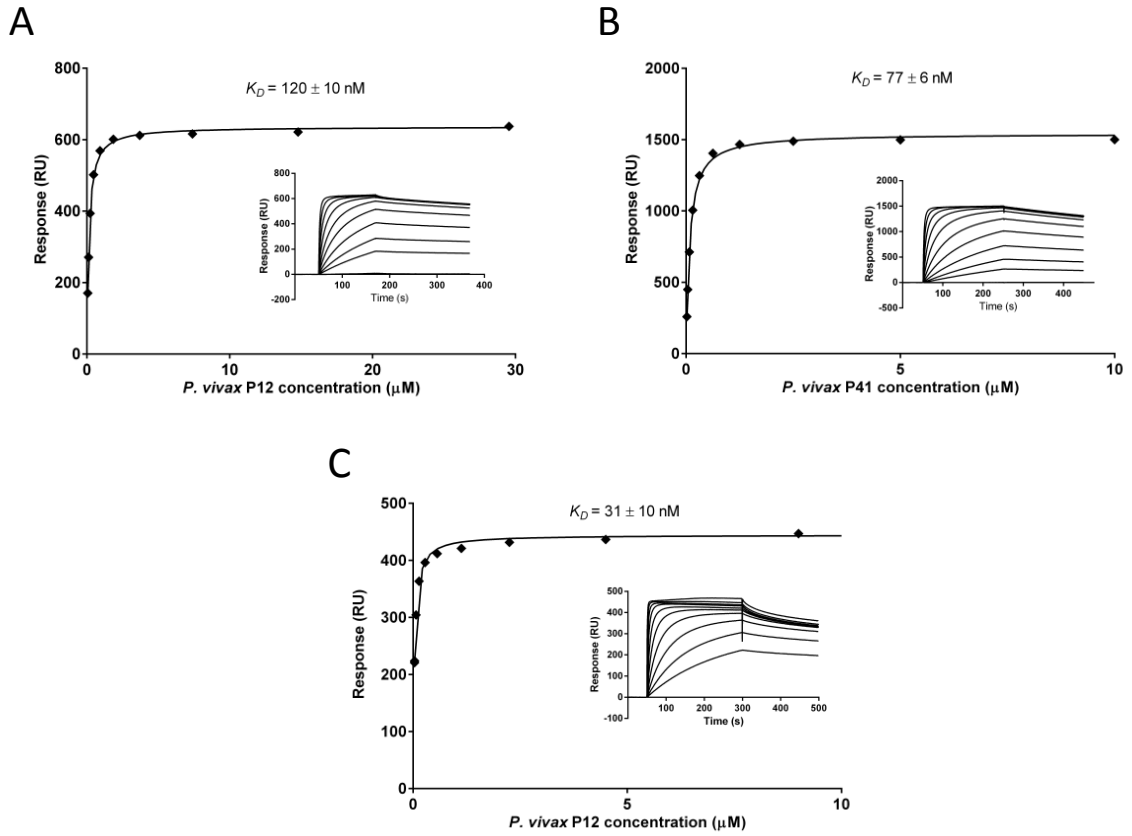

Figure 4.16: Quantification of the *P. vivax* P12-P41 interaction affinity by surface plasmon resonance

Increasing concentrations of analyte protein were injected over immobilized biotinylated ligand protein. Reference-subtracted binding data were plotted as a binding curve and the equilibrium dissociation constant was calculated using $R_{eq}=CR_{max}/(C+K_D)$. Experiments included analyte-ligand combinations *P. vivax* P12-P41 (A), *P. vivax* P41-P12 (B), and *P. vivax* P12-*P. falciparum* P41 (C). Lower concentrations failed to reach equilibrium, which resulted in an overestimated K_D. SEC, size-exclusion chromatography. Figure adapted from (Hostetler et al., 2015) under the Creative Commons Attribution (CC BY) license.

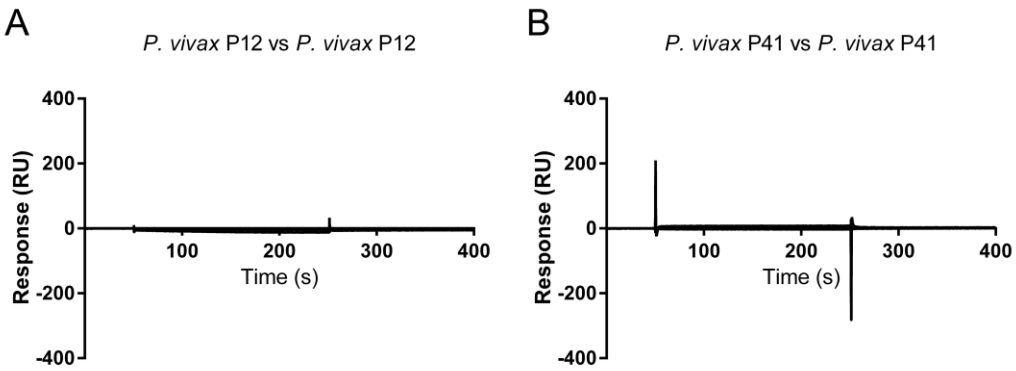

Figure 4.17: *P. vivax* P12 and *P. vivax* P41 show no self-binding by surface plasmon resonance

Increasing concentrations of *P. vivax* P12 (A) or P41 (B) were injected over immobilized biotinylated *P. vivax* P12 (A) or P41 (B), with no interactions observed. Figure reprinted from (Hostetler et al., 2015) under the Creative Commons Attribution (CC BY) license.

4.2.5.3 SPR for novel interactions

SPR was also used to study the *P. vivax* MSP3.10-MSP7.1 interaction. Equilibrium binding experiments using *P. vivax* MSP3.10 as ligand and *P. vivax* MSP7.1 as analyte showed a relatively high binding affinity (Figure 4.18), but a K_D could not be calculated since the binding did not reach equilibrium at any of the concentrations tested. In addition, the binding did not fit a 1:1 model (Figure 4.18, shown in red), most likely due to oligomerization of the purified *P. vivax* MSP7.1. Such complex binding behavior was also observed between P-selectin and *P. falciparum* MSP7, which also oligomerize (Perrin et al., 2015).

SPR was also used to explore the P12-PVX_110945 interaction identified by AVEXIS. This confirmed an extremely weak interaction between recombinant purified *P. vivax* P12 as analyte and biotinylated PVX_110945 as ligand (Figure 4.19). None of the concentrations used reached equilibrium, which prevented the calculation of an equilibrium dissociation constant (K_D). Further functional studies will be needed to determine whether the interaction is biologically relevant.

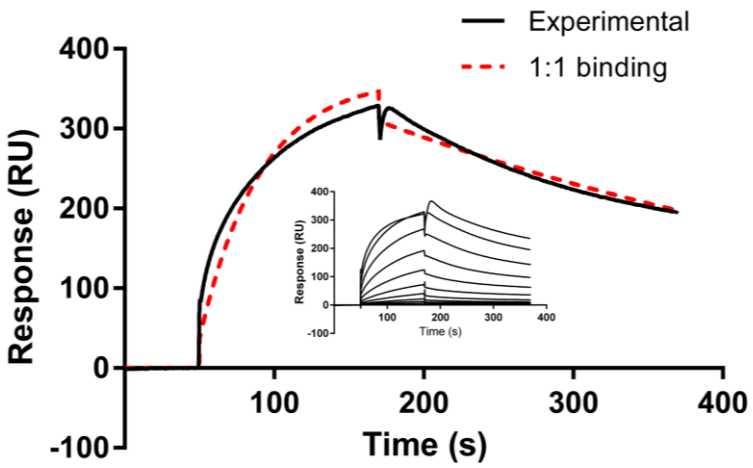

Figure 4.18: Surface plasmon resonance confirms the *P. vivax* MSP3.10-MSP7.1 interaction

Increasing concentrations of *P. vivax* MSP7.1 were injected over immobilized biotinylated *P. vivax* MSP3.10. Relatively high-affinity binding was observed, although none of the concentrations used reached equilibrium (inset). The binding did not fit a 1:1 model (red dashed line). The increase in response units at the start of the dissociation phase at the higher analyte concentrations of *P. vivax* MSP7.1 is likely an artefactual buffer effect. SEC, size-exclusion chromatography. Figure reprinted from (Hostetler et al., 2015) under the Creative Commons Attribution (CC BY) license.

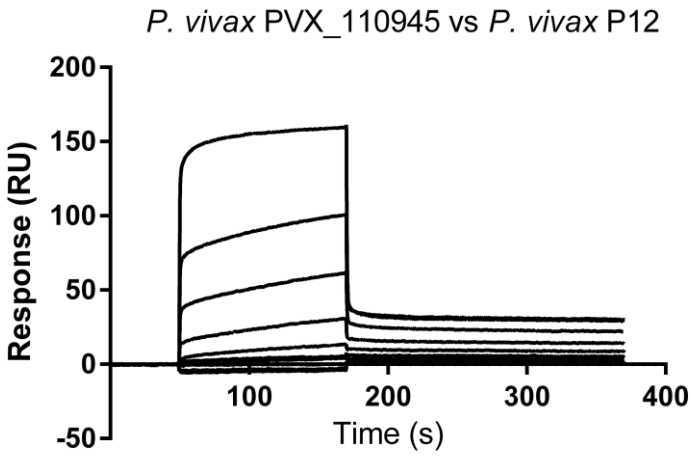

Figure 4.19: Surface plasmon resonance supports a weak interaction between *P. vivax* P12 and *P. vivax* PVX_110945

Increasing concentrations of *P. vivax* P12 were injected over immobilized biotinylated *P. vivax* PVX_110945 with weak binding observed. None of the concentrations used reached equilibrium, which prevented the calculation of an equilibrium dissociation constant (K_D). Figure reprinted from (Hostetler et al., 2015) under the Creative Commons Attribution (CC BY) license.

4.3 Discussion

In this Chapter, I presented the production and characterization of a recombinant *P. vivax* merozoite protein library. Research into *P. vivax* biology generally lags behind *P. falciparum,* in large part due to the lack of a robust *in vitro* culture system for *P. vivax.* This complicates investigation of *P. vivax* erythrocyte invasion and thus identification of new *P. vivax* blood-stage vaccine candidates. Multiple studies have investigated single or few *P. vivax* merozoite proteins, but a comprehensive library of full-length ectodomains of merozoite surface, microneme, and rhoptry proteins for functional studies has not yet been assembled. To build such a library, we selected 39 proteins that are known or predicted to localize to the surface or apical organelles of merozoites based on published *P. vivax* and *P. falciparum* studies, and published microarray data showing gene upregulation during *P. vivax* schizogony. I was able to express 37 of these proteins as recombinant ectodomains, including several members of the MSP and 6-cysteine protein families, additional invasion-related and/or GPI-anchored proteins, as well as several proteins with no *P. falciparum* homologs. This success in expressing 95% of *P. vivax* merozoite proteins at levels useable for biochemical studies is comparable to similar efforts at expressing *P. falciparum* merozoite proteins (Crosnier et al., 2013), and demonstrates the broad utility of the human HEK293E cell expression system in producing high-yield, high-quality proteins for *Plasmodium* research.

Protein expression levels were evaluated for patterns that predicted success. From previous experience using the HEK293E system, proteins larger than 250 kDa are frequently not expressed or expressed at low levels. However, size alone was not a useful predictor of expression levels in this library, as I sometimes observed high and low expression levels for large and small proteins, respectively [e.g., high expression for MSP3.4 (140 kDa]) and low expression for PVX_084815 (53 kDa)]. The effects of amino acid composition and predicted protein folding on expression are not yet clear, but useful predictors of expression levels may become apparent as we expand our *Plasmodium* plasmid library.

In order to maximize our chance for successful expression, I used codon-optimized, protein ectodomains with mutated *N*-linked glycosylation sites. In this work I did not seek to systematically study the factors that aid *Plasmodium* expression, seeking instead a “one size fits all” approach to maximize the number of proteins expressed, but minimize the

time and resources invested in optimizing expression. However, prior work in the Wright laboratory suggests that codon optimization is necessary for successful expression of some but not all *Plasmodium* proteins in mammalian expression systems. Previous studies have often used native sequences [reviewed in (Birkholtz et al., 2008)], even though other studies have noted that optimization of codon usage can aid expression (Kocken et al., 2002, Russell et al., 2005). As another example, the Wright and Rayner laboratories previously found that codon optimization alone had little effect on the expression of PfRH5, while the inclusion of an exogenous signal peptide and the mutation of *N*-linked glycosylation sites significantly increased its expression (Crosnier et al., 2013). In addition, the AT content of the *P. vivax* genome (50-60%) is much closer to that of humans (59%) than to that of *P. falciparum* (80%). This significantly impacts codon usage between the 2 *Plasmodium* species [as noted in (Chen and Cheng, 1999)], and potentially means that *P. vivax* proteins would be more easily expressed without codon optimization than *P. falciparum* proteins. If expression constructs are to be generated synthetically, as they were in this study, then codon optimization appears to be a useful step, but not a panacea, in the expression of *P. vivax* proteins.

A flow cytometry-based erythrocyte-binding assay was used to investigate *P. vivax* protein library binding to both erythrocytes and hematopoietic stem cell-derived reticulocytes. The assay allowed for the complete library to be screened in 96-well format using small quantities of protein culture supernatant (i.e., 100 µl per binding replicate), which is a significant improvement over assays requiring His-purification or manual counting of rosettes. The results demonstrated significant shifts in fluorescence indicative of binding for *P. falciparum* EBA175. *P. vivax* MTRAP also appeared to show binding with less dramatic shifts in fluorescence. However, the atypical distribution may indicate bead aggregation and follow-up experiments are needed to fully explore this potential interaction. Overall, the assay appeared to distinguish binding in the case of positive control *P. falciparum* EBA175, which binds to glycophorin A, a highly abundant erythrocyte surface protein (10^6 copies/cell), but failed to detect any novel interactions for proteins in the *P. vivax* protein library. This could reflect biological reality (i.e., none of these proteins bind erythrocytes), but it could also suggest methodological issues, specifically that the binding detection method has relatively low sensitivity. Supporting this theory is the fact that no binding was detected in the case of DBP, which binds to a relatively low abundance surface receptor, DARC (10^4 copies/cell). More sensitive

binding assays need to be developed to systematically interrogate this and other *Plasmodium* merozoite protein libraries.

The AVEXIS interaction screen between the *P. vivax* recombinant protein library and specific erythrocyte receptors also failed to detect any clear interactions, including the expected interaction between DBP and the N-terminal sequence of DARC. Multipass membrane proteins, such as DARC, are difficult to express as properly-folded amphiphilic fragments, and the recombinant prey protein included in the library represented the largest surface-exposed ectodomain, which included the expected binding region with DBP-RII. However, this tail alone may not be enough to provide binding in the context of an AVEXIS assay, which has not been applied to this interaction before. Again, more sensitive methods may need to be developed to use the library to identify new interactions.

The intra-library AVEXIS assay detected 3 interactions: P12-P41, MSP3.10-MSP7.1, and P12-PVX_110945. This result confirmed for the first time that the known P12-P41 interaction in *P. falciparum* (Taechalertpaisarn et al., 2012) is conserved in the evolutionarily distant but related parasite, *P. vivax*. Biophysical measurements using SPR showed that the *P. vivax* P12-P41 interaction appears to be at least 3 times stronger than the *P. falciparum* P12-P41 interaction. Gene knockout experiments have shown that *P. falciparum* P12 is not essential for parasite invasion or growth *in vitro*, and antibodies against P12 do not significantly block erythrocyte invasion (Taechalertpaisarn et al., 2012). The difference in protein-protein interaction affinity between the species indicates that there could be functional differences between them. Interestingly, an interaction between *P. vivax* P12 and *P. falciparum* P41 was also detected, suggesting the existence of conserved binding sites, which may be attractive targets for vaccines against both *P. falciparum* and *P. vivax*. The reverse interaction between *P. falciparum* P12 and *P. vivax* P41 (in both bait-prey and prey-bait orientations) was not detected, which may assist in mapping the P12-P41 binding site within *Plasmodium* species.

Two novel interactions were also detected with AVEXIS and investigated by SPR; both putative interactions were validated by SPR and/or reciprocation of prey-bait and bait-prey interactions in AVEXIS. The failure of AVEXIS to detect the *P. vivax* MSP3.10-MSP7.1 interaction in the prey-bait orientation may indicate that artificially pentamerized MSP3.10 prey interferes with the formation of an oligomeric structure necessary for

binding. While *P. falciparum* MSP3.1 contains a conserved C-terminus important for oligomerization (Burgess et al., 2005), *P. vivax* MSP3s are not clear homologs and lack this feature (Rice et al., 2014, Jiang et al., 2013). Their potential for forming higher order structures is unknown. SPR showed that the *P. vivax* MSP3.10-MSP7.1 interaction has a relatively high avidity, potentially increased due to oligomerization of MSP7.1. Members of the *P. vivax* MSP3 family are known to peripherally associate with the merozoite surface (Jiang et al., 2013) and *P. vivax* MSP7s are predicted to have a similar location, based on their *P. falciparum* orthologs (Kadekoppala and Holder, 2010, Kadekoppala et al., 2010). MSP3 has no known binding partners, and *P. falciparum* MSP7 is known to form a complex with MSP1 (Kadekoppala and Holder, 2010). An interaction between MSP7 and MSP1, while perhaps predicted based on *P. falciparum* data, has not yet been established for *P. vivax*. In my study, the AVEXIS screen also did not detect such an interaction although only 3 of the 11 *P. vivax* MSP7 family members were included in our library, meaning that 1 of the members not expressed to date could interact with *P. vivax* MSP1. SPR data suggest that the *P. vivax* P12-PVX_110945 interaction is extremely weak. PVX_110945 is a hypothetical protein with no known function, but its transcription and genomic location give it a possible function in merozoite development, invasion of erythrocytes, or both. Additional experiments to co-localize these 2 interacting pairs in parasite isolates may help validate and shed light on the possible biological relevance of these interactions.

4.3.1 Limitations and future work

The *P. vivax* recombinant library AVEXIS screen against erythrocyte receptors failed to detect any interactions, including the known PvDBP-DARC interaction. For several reasons, however, these results do not conclusively indicate that none of these proteins interact. AVEXIS is optimized to have a very low false-positive rate with 2/8 expected interactions detected and 0/37 false positives in validation experiments described in (Bushell et al., 2008). This focus on eliminating false positives may also reduce the number of real positives. In addition, several baits and preys used in the screens were expressed at levels below the optimal concentrations for detecting interactions. Future studies exploring additional erythrocyte receptors and/or focusing on individual proteins in the library will be useful to fully evaluate the proteins' potential for erythrocyte binding.

The *P. vivax* P12-P41 interaction appears to have a much higher affinity than the analogous *P. falciparum* P12-P41 interaction. The difference in protein-protein interaction affinity between species indicates that there could be functional differences between them; therefore, we plan to investigate the potential function of the *P. vivax* P12-P41 interaction in merozoite invasion of reticulocytes and the possible invasion-blocking effects of immune IgG specific for P12 or P41 in Cambodian *P. vivax* isolates *ex vivo*.

We believe that this *P. vivax* recombinant library and expression approach will significantly improve our ability to study the biology of *P. vivax* erythrocyte invasion and the natural development of *P. vivax* immunity. This approach has already been successfully applied to both fronts in *P. falciparum* research (Crosnier et al., 2013, Crosnier et al., 2011, Bartholdson et al., 2012, Osier et al., 2014). Not only will this *P. vivax* library enable new screens for parasite-host interactions, it can also contribute to protein structure and immunoepidemiological studies. Structural studies can greatly increase our understanding of the function of these proteins. Such studies can also enhance the possibility of defining protective versus decoy conformational epitopes, which may have a tremendous impact on their ultimate potential as vaccine candidates (Chen et al., 2013). Future immunoepidemiological studies that use a panel of proteins should enable more systematic comparisons between proteins, in contrast to previous studies that have used only one or several proteins. The *P. vivax* library of expression plasmids is now freely available through the non-profit plasmid repository, Addgene (www.addgene.org), facilitating community efforts to identify, validate, and develop promising vaccines for *P. vivax* malaria. Finally, systematic generation of antibodies against library proteins will generate reagents that can be used in *ex vivo* invasion assays to test for invasion-blocking potential. Identification and prioritisation of new *P. vivax* vaccine antigens is therefore significantly enhanced by the work described here, even if specific receptors were not identified in this work.

4.4 Conclusion

We selected 39 *P. vivax* proteins that are predicted to localize to the merozoite surface or invasive secretory organelles, some of which show homology to *P. falciparum* vaccine candidates. Of these, I was able to express 37 full-length protein ectodomains in a mammalian expression system, which has been previously used to express *P. falciparum*

invasion ligands such as PfRH5. Using a method specifically designed to detect low-affinity, extracellular protein-protein interactions, I confirmed a predicted interaction between *P. vivax* 6-cysteine proteins P12 and P41, suggesting that the proteins are natively-folded and functional. This screen also identified 2 novel protein-protein interactions, between P12 and PVX_110945 and between MSP3.10 and MSP7.1, the latter of which was confirmed by surface plasmon resonance. As well as identifying new interactions for further biological studies, this library will be useful in identifying *P. vivax* proteins with vaccine potential, and studying *P. vivax* malaria pathogenesis and immunoepidemiology.