

This is a repository copy of *Systematic identification of P. falciparum sporozoite membrane protein interactions reveals an essential role for the p24 complex in host infection.*

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/170140/>

Version: Accepted Version

---

**Article:**

Knöcke, Julia, Dundas, Kirsten, Yang, Annie S.P. et al. (7 more authors) (Accepted: 2020)  
Systematic identification of *P. falciparum* sporozoite membrane protein interactions reveals an essential role for the p24 complex in host infection. *Molecular and Cellular Proteomics*. ISSN 1535-9476 (In Press)

---

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

**Systematic identification of *P. falciparum* sporozoite membrane protein interactions reveals an essential role for the p24 complex in host infection**

Julia Knöckel<sup>1,2</sup>, Kirsten Dundas<sup>1,2,†</sup>, Annie S.P. Yang<sup>3,†</sup>, Francis Galaway<sup>1,2</sup>, Tom Metcalf<sup>2</sup>, Geert-Jan van Gemert<sup>3</sup>, Robert W. Sauerwein<sup>3</sup>, Julian C. Rayner<sup>2</sup>, Oliver Billker<sup>2,4</sup> and Gavin J. Wright<sup>1,2,5\*</sup>.

<sup>1</sup>Cell Surface Signalling Laboratory, and <sup>2</sup>Malaria Programme, Wellcome Sanger Institute, Cambridge CB10 1SA, United Kingdom.

<sup>3</sup>Department of Medical Microbiology, Radboud University Medical Center, Geert Grooteplein 28, Microbiology 268, 6500 HB Nijmegen, The Netherlands.

<sup>4</sup>The Laboratory for Molecular Infection Medicine Sweden (MIMS) and Department of Molecular Biology, Umeå University, Umeå, Sweden.

<sup>5</sup>Department of Biology, Hull York Medical School, York Biomedical Research Institute, University of York, Wentworth Way, York, YO10 5DD, United Kingdom.

† represents equal contribution.

\*Correspondence to: Gavin J. Wright, Cell Surface Signalling Laboratory, Wellcome Sanger Institute, Cambridge, CB10 1SA, United Kingdom. Tel +44 (0) 1223 496852, E-mail: [gw2@sanger.ac.uk](mailto:gw2@sanger.ac.uk)

Running title: *P. falciparum* sporozoite membrane protein interactions

Abbreviations: AVEXIS, avidity-based extracellular interaction screen; CSP, circumsporozoite protein; CyRPA, Cysteine-Rich Protective Antigen; RIPR, RH5 Interacting PRotein; GPI, glycosylphosphatidylinositol; HEK, human embryonic kidney; HBS, hepes-buffered saline; MSP, merozoite surface protein; PIESP15, parasite-infected erythrocyte surface protein 15; SPR, Surface Plasmon Resonance; TRAP, Thrombospondin-Related Anonymous Protein

## Abstract

Sporozoites are a motile form of malaria-causing *Plasmodium falciparum* parasites that migrate from the site of transmission in the dermis through the bloodstream to invade hepatocytes. Sporozoites interact with many cells within the host, but the molecular identity of these interactions and their role in the pathology of malaria is poorly understood. Parasite proteins that are secreted and embedded within membranes are known to be important for these interactions, but our understanding of how they interact with each other to form functional complexes is largely unknown. Here, we compile a library of recombinant proteins representing the repertoire of cell surface and secreted proteins from the *P. falciparum* sporozoite and use an assay designed to detect extracellular interactions to systematically identify complexes. We identify three protein complexes including an interaction between two components of the p24 complex that is involved in the trafficking of glycosylphosphatidylinositol (GPI)-anchored proteins through the secretory pathway. *Plasmodium* parasites lacking either gene are strongly inhibited in the establishment of liver stage infections. These findings reveal an important role for the p24 complex in malaria pathogenesis and show that the library of recombinant proteins represents a valuable resource to investigate *P. falciparum* sporozoite biology.

## Introduction

Malaria is a deadly mosquito-borne infectious disease caused by unicellular *Plasmodium* parasites, and despite intensive control efforts is still responsible for almost half a million deaths annually (1). Critical to mosquito-human transmission is the sporozoite: a motile stage that develops within oocysts located on the midgut of infected mosquitoes before invading the mosquito salivary glands and being transmitted during a subsequent blood meal. Once deposited in the dermis, sporozoites disperse in apparently random directions before entering the circulation and eventually enter the liver where they invade hepatocytes and continue their development (2). Targeting the sporozoite is an attractive choice for vaccine development because sporozoites are extracellular and therefore directly exposed to the host humoral immune system, this stage of the infection is asymptomatic, and sporozoites are few in number since only tens to hundreds are transmitted (3-5). Importantly, malaria vaccines based on attenuated whole sporozoites, as well as a subunit vaccine incorporating components of the major sporozoite surface protein CSP, have been shown to be effective in human trials, although protection was typically short-lived (6).

Proteins that are displayed on the surface of the parasite or released from intracellular secretory organelles such as the rhoptries and micronemes are considered good subunit vaccine targets because they are directly accessible to host antibodies. The main functions of these proteins are likely to be the subversion of the immune response or the interaction with host molecules to regulate different stages of infection; for example, to power parasite motility or bind specific receptors that control the tropism of host cell invasion. Because the blood stages of *P. falciparum* can be cultured *in vitro*, we have a comparatively good understanding of which parasite proteins are displayed on the surface of the merozoite and how they interact to form functional complexes. For example, the major merozoite surface protein MSP1 is found in a protein complex with MSP7 and MSP6 (7, 8), and the secreted P41 protein is tethered to

the surface of the merozoite by the GPI-anchored P12 protein (9). Parasite protein complexes that are known to be essential for *P. falciparum* erythrocyte invasion include the interactions between AMA1 and RON2 (10) and more recently those proteins that directly bind RH5 including CyRPA (11), RIPP (12), and P113 (13). Mechanistically, vaccine-elicited antibodies that target the components of these protein complexes on the parasite surface may neutralise blood stage parasites by inhibiting their function or formation (14). By contrast to the blood stages, however, relatively little is known about the biochemical composition and organisation of the sporozoite surface although genetic targeting has revealed that several parasite surface proteins are required for normal sporozoite development and infection. For example, circumsporozoite protein (CSP) is required for the development of sporozoites in the oocyst (15, 16), TRAP family proteins (TRAP and TLP) are important for sporozoite motility and infection (17, 18), and members of the 6-cys family (P36 and P52) are important for host receptor usage in hepatocyte invasion (19). While we have a good cellular description of how the sporozoite interacts with host cells (20), a molecular understanding of these behaviours is incomplete and yet could represent important new opportunities for pre-erythrocytic malaria vaccines.

Identifying interactions between membrane-embedded proteins is challenging because membrane proteins are both amphipathic, making them difficult to solubilise in solvents that retain their native conformation, and because interactions between them can be highly transient; together, these features make it difficult to confidently interpret data from commonly-used methods such as biochemical purifications using native material (21). To address these issues, we have developed an approach called AVEXIS (for AVidity-based EXtracellular Interaction Screening) (22), which, together with a library of functionally active recombinant blood-stage *Plasmodium* proteins (23, 24) has permitted the biochemical identification and characterisation of parasite protein complexes at the surface of the merozoite including the RH5 complex (13). Here, we use proteomic analyses of *Plasmodium* sporozoites to compile

a recombinant protein library representing the cell surface receptor repertoire and secretome of the *P. falciparum* sporozoite. We use this protein library and the AVEIXIS approach to identify interactions between sporozoite proteins and characterise their role in the pathogenesis of *Plasmodium* infections.

## Experimental Procedures

### *Ethics statement*

Experiments involving animals were performed under UK Home Office governmental regulations (project licence numbers PD3DA8D1F and P98FFE489) and European directive 2010/63/EU. Research was approved by the Sanger Institute Animal Welfare and Ethical Review Board.

### *Design of the Plasmodium sporozoite protein library.*

To obtain a catalogue of proteins that are expressed by sporozoites, published proteomic data using mass spectrometry-based profiling of midgut-and salivary gland sporozoites (25-27) were combined to create a list of unique proteins and protein expression constructs designed essentially as described (23); we additionally included several proteins expressed by ookinetes. Briefly, cell surface and secreted proteins were initially identified by the presence of a signal peptide and/or transmembrane domains using SignalP v4.0 (28) and TMHMM v2.0 (29) software. To define the ectodomain regions, signal sequences and transmembrane domains were removed and all potential N-linked glycosylation sites (N-X-S/T, where X is not proline) were systematically mutated by substituting the serine/threonine for alanine to prevent inappropriate glycosylation when expressed in mammalian cells. Sequences encoding the sporozoite protein ectodomains were codon optimised for expression in human cells and made by gene synthesis (GeneART). Four genes failed synthesis and three could not be subcloned into the mammalian expression plasmids, suggesting the insert was toxic in *E. coli* (Table S1).

All sequences were flanked with unique 5' NotI and 3' Ascl restriction enzyme sites to allow inframe cloning in the bait and prey plasmids (30). Both expression plasmids contain a highly efficient mouse variable  $\kappa$  light chain signal peptide and a rat Cd4d3+4 epitope tag followed by either a peptide sequence allowing biotinylation and 6-his tag (baits) or pentamerising COMP sequences followed by the beta-lactamase enzyme and a 6-his tag (preys) (31). The same process was followed for the *P. berghei* orthologues. All plasmids have been deposited in the Addgene plasmid repository ([www.addgene.org/express](http://www.addgene.org/express)).

#### *Production of the recombinant bait and prey proteins.*

All proteins were produced recombinantly as a monobiotinylated bait and a pentamerised prey by transient transfection of HEK293E (32) or HEK293-6E cells (33) essentially as described (30). All proteins were purified from the cell culture supernatant up to six days after transfection using Ni<sup>2+</sup>-NTA resin using either automated purification systems (GE Healthcare) or manually using Ni-NTA Agarose (Jena Biosciences) and dialysed extensively against HBS or PBS.

#### *Normalisation of bait and prey proteins.*

The expression levels of biotinylated bait and  $\beta$ -lactamase-tagged prey proteins were quantified essentially as described (30). Briefly, for bait proteins, serial dilutions of the purified, dialysed proteins were added to individual wells of a streptavidin-coated microtitre plate (NUNC). Proteins were captured and washed with HBS-T (HBS/0.1% Tween 20) followed by addition of 1  $\mu$ g/mL mouse anti-rat Cd4 OX68 monoclonal antibody in HBS/0.2% BSA and washed with HBS-T. Proteins were detected by addition of a goat-anti-mouse alkaline-phosphatase conjugated antibody (Sigma A4656), washing followed by addition of 1 mg/mL alkaline phosphatase substrate 104 dissolved in diethanolamine buffer and reading the absorbance at 405 nm in an automated plate reader (FluoStar Optima, BMG Labtech). For subsequent assays, the highest protein dilution that saturated the biotin binding capacity of the well was used. The activity of prey proteins was normalised by quantifying the activity of



the  $\beta$ -lactamase enzyme in the prey protein tag. Serial dilutions were prepared in a 96-well flat bottom microtitre plate and 20  $\mu$ L added to 60  $\mu$ L of a 242  $\mu$ M solution of the colourimetric  $\beta$ -lactamase substrate nitrocefin (Cayman Chemical Company). The kinetics of substrate hydrolysis were followed for 20 minutes and prey proteins normalised to an activity of  $\sim 2$  nmol  $\text{min}^{-1}$  for use in subsequent assays. To produce tetrameric preys to detect the CeITOS-PF3D7\_0721100 interaction, a dilution series of monomeric biotinylated protein (similar to bait normalisation) was incubated with 2-5  $\mu$ g/mL alkaline phosphatase coupled streptavidin (Sigma) for 45 min to allow the formation of tetramers and then used in an ELISA as described above. The lowest dilution which did not give a signal at 405 nm was used for AVEIXIS.

#### *Western blotting.*

Purified proteins were resolved by SDS-PAGE using Novex NuPAGE 4–12% Bis Tris precast gels (Life Technologies). Proteins were reduced by adding NuPAGE reducing agent to the sample and antioxidant (Life Technologies) to the gel running buffer according to the manufacturer's instructions. Proteins were blotted onto nitrocellulose membranes (Life Technologies) and protein binding sites blocked using a 2% BSA solution. Biotinylated proteins were detected by incubating membranes with High sensitivity Streptavidin-HRP (PIERCE), washing and addition of SuperSignal West Pico Chemiluminescent substrate (PIERCE) and developed on photographic film.

#### *Enzyme-linked immunosorbant assay (ELISA)*

ELISAs were performed by capturing biotinylated bait proteins into individual wells of streptavidin-coated 384-well plates (NUNC). Plates were washed for 30 mins with 50  $\mu$ L PBS-T (0.2% Tween) and blocked with PBS-2% BSA for a minimum of 3 hours. 20  $\mu$ L of a bait protein diluted in PBS-2% BSA at a concentration previously determined as the amount required to saturate the biotin binding capacity of the well were added in triplicate and incubated for at least 16 hours at 4°C. Where required, a similar set of bait protein dilutions

were heat inactivated for 30 min at 70 °C. Antisera from pooled Malawian adults (23) were centrifuged at 13,000 rpm for a minimum of one hour, diluted in PBS-2% BSA and incubated with rocking for at least 16 hours at 4°C before adding to the antigen-coated plates for one hour. Plates were washed 3x in PBS-T before addition of a 1:10,000 dilution of rabbit-anti-human polyvalent IgG/A/M (Sigma I8010) in PBS-2% BSA for one hour, and washed in PBS-T, followed by 1:10,000 dilution of goat anti-rabbit Poly-HRP IgG/A/M (PIERCE) in PBS-2% BSA for one hour. Plates were washed in PBS-T and the HRP substrate ABTS (KPL) was added and absorption at 405 nm determined using an automated plate reader (FluoStar Optima, BMG Labtech).

#### *AVIDITY-based EXtracellular protein Interaction Screening assay*

Systematic screening for protein interactions was performed essentially as described previously (22, 34). Briefly, normalised bait proteins were immobilised on streptavidin-coated 96-well microtitre plates (NUNC) and activity-normalised prey proteins added to the arrayed bait proteins and incubated at room temperature for 1 hour. The plates were washed twice in HBS-T (0.2% Tween 20) and once in HBS before 60 µL of 125 µg/mL nitrocefin (Cayman Chemical Company) was added and the absorbance measured at 485 nm in a plate reader. When tetramerised preys were used, 1 mg/mL phosphatase substrate 104 in diethanolamine buffer was added and the absorbance measured at 405 nm. The rat Cd200-Cd200R interaction was used as a positive control, and a biotinylated Ox68 antibody which recognises the Cd4-tag in the prey was used as a prey capture control. Some bait and prey proteins showed low levels of promiscuous binding in the screen (e.g. bait 56 and prey 27) but because of their lack of specificity and weak binding signals, were not investigated further.

#### *Experimental Design and Statistical Rationale*

The goal of the protein interaction screening was to identify directly interacting proteins within the library of recombinant *P. falciparum* sporozoite membrane proteins. Sixty-four bait proteins

were tested for interactions against 54 prey proteins; however, there was insufficient amounts of one of the baits (protein 53 - SIAP-1) and therefore this bait was tested against 35 prey proteins. In total, 3,437 binding tests were performed: 46 proteins were present in both prey and bait libraries and therefore tested in both bait-prey orientations; the other proteins were consequently tested in one orientation, making a total of 2,379 unique interactions tested. The initial screen was performed once but the identified interactions were repeated in triplicate and further characterised using surface plasmon resonance. Control baits and preys were present in all experiments; typically the rat Cd200-CD200R interaction is used as a positive control. Data are shown as mean absorbance readings with standard deviations to show reproducibility between replicates.

#### *Surface Plasmon Resonance*

A Biacore 8K instrument (GE Healthcare) was used for the surface plasmon resonance studies essentially as described (35). In brief, biotinylated bait proteins were immobilised on a streptavidin-coated sensor chip (GE Healthcare). Approximately 400 RUs of the control bait (biotinylated rat Cd4d3+4) were immobilised in the flow cell used as a reference and the approximate molar equivalents of the query protein immobilized in the other flow cells. Purified analyte proteins were resolved by size exclusion chromatography on a Superdex 200 Increase 10/300 column (GE Healthcare) in HBS-P (10 mM HEPES, 150 mM NaCl, 0.05% v/v P20 surfactant) immediately before use in SPR experiments thereby removing any aggregated protein that might influence kinetic measurements. The surface was regenerated with a pulse of 2 M NaCl at the end of each cycle, and all experiments were performed at 37°C. Duplicate injections of the same concentration in each experiment were superimposable demonstrating no loss of activity after regenerating the surface. Binding data were analysed in the manufacturer's Biacore 8K evaluation software version 1.1 (GE Healthcare).

#### *Cloning and expression of *P. berghei* orthologues*

The *P. berghei* orthologues of the six identified interacting *P. falciparum* proteins were identified using PlasmoDB (36). Sequences were modified in the same way described in *Design of the Plasmodium sporozoite protein library* and the genes synthesised by GeneART. The synthetic plasmids were digested with NotI and Ascl restriction enzymes (NEB) and were each subcloned into the bait and prey vectors. Recombinant proteins were produced by transient transfection of HEK293-E or 6E cells as described above.

#### *P. berghei* genetic targeting.

*P. berghei* targeting plasmids for *PIESP15* (*PBANKA\_0209200*, PbGEM-268595) and *PBANKA\_0800600*, PbGEM-331155 were obtained from the PlasmoGEM vector resource ([plasmogem.sanger.ac.uk](http://plasmogem.sanger.ac.uk)) (37). To make targeting plasmids that disrupt the *PBANKA\_0522500* and *PBANKA\_1241700* loci, oligonucleotides were designed to amplify 5'- and 3' homology regions of each gene. Each homology arm was approximately 1 kbp long and spans the 5' UTR into the open reading frame, and from inside the open reading frame into the 3'UTR of each gene respectively. Restriction sites were added to the primer sequence to enable cloning of the PCR products. PCRs were done using Kapa HiFi HotStart Ready Mix (Kapa Biosciences) and genomic DNA isolated from *P. berghei* GFP-LUC parasites (38) was used as a template. PCR conditions were 5 min at 95°C followed by 35 cycles of 95°C for 30 sec, 55°C for 30 sec and 68°C for 1-3.5 min and a final extension of 10 min at 68°C. PCR products were sequentially cloned into the KpnI and XmaI or the XhoI and SacII restriction sites of the pR6K attL1-3xHA-h*DHFR-yfcu*-attL2 plasmid backbone (39) and used to transfect *P. berghei* parasites.

#### *Transfection and cloning by limiting dilution*

Recombinant parasites were generated by transfecting a parental “wild-type” transgenic *P. berghei* strain expressing green fluorescent protein (GFP) and luciferase (38). Transfections were done essentially as described (40). Briefly, 6-12 week old Balb/C or Theiler's original

(TO) mice were used for routine propagation of the parasites. Prior to transfection, 10 µL DNA were digested with NotI restriction enzyme (NEB) overnight at 37 °C to release the transfection cassette from the plasmid backbone. DNA was purified by ethanol precipitation and resuspended in 10 µL dH<sub>2</sub>O. Parasites were cultured *in vitro* for 20-22 hours to obtain synchronised schizonts. Schizonts were purified using a 55% Nycodenz (Sigma)/PBS gradient and washed once in RPMI1640 medium. Pelleted schizonts were resuspended in P3 Primary cell Nucleofector (Lonza, 18 µL/transfection) and 25 µL of this mix was added to 5 µL of digested DNA. 20 µL of the mix was then added to an individual well of a 16-well Nucleocuvette™ Strip and pulsed using the AMAXA™ 4D-Nucleofector electroporator, programme FI-115. Parasites were immediately resuspended in 100 µL incomplete RPMI1640 medium and injected into the tail vein of Balb/c mice. The following day in the morning, pyrimethamine (0.07mg/ml, pH 4.5) was added to the drinking water for parasite selection. Blood smears were taken from day 7 after transfection to monitor parasitaemia and when parasitaemia reached 0.5-1%, blood was collected for stocks and genotyping. Subsequently, a drop of blood was taken from a mouse infected with the transfected parasites and parasites were cloned by limiting dilution. Clonality (absence of the WT locus) was confirmed by genotyping and cloned parasites were then used for experiments.

#### *Genotyping of P. berghei parasites for validation of gene disruption*

To confirm integration of the transfection cassette into the genome of the parasite, we designed specific oligonucleotides for each target gene, annealing to the chromosome just outside the homology arm of the plasmid (primer GT). This was paired with a universal primer which anneals to the *yfcu* cassette in the targeting plasmid (primer GW2a or b). To detect absence or presence of the WT locus, a gene specific primer located inside the gene sequence of the target gene (primer QCR1) was paired with a reverse primer in the homology arm (QCR2). Primers GW2 and QCR2 were then combined to detect the transfection plasmid, which could be either integrated or not integrated in the genome. Because the product size for

the primer combination *PbPIESP15*-GW2+GT was too large, two PCR reactions were done to amplify one half of the fragment each. Therefore, primer GW2b was combined with primer *PIESP15*-rev2 and primer GT was combined with primer *PbPIESP15*-fwd2. Sequences for primers QCR1, QCR2 and GT for *PbANKA\_0800600* and *PbPIESP15* were obtained from the PlasmoGEM web page. All primer sequences are listed in Table S2. PCR reactions were performed using Kapa HiFi HotStart Ready Mix (Kapa Biosciences) and cycling conditions were 5 min at 95°C followed by 35 cycles of 95°C for 30 sec, 55°C for 30 sec and 68°C for 1-3.5 min and a final extension of 10 min at 68°C. Samples were separated on a 0.8% Agarose gel using Hyperladder 1kb (Bioline) as a size standard.

#### *In vivo phenotyping of transgenic P. berghei parasites*

*Anopheles stephensi* mosquitoes were allowed to feed on anaesthetised *P. berghei* infected Balb/C mice for no more than 15 min and placed in an incubator at 19°C thereafter. Nine to eleven days after infection, mosquitoes were anaesthetised and kept on ice and midguts were dissected in PBS. Only mosquitoes with developed eggs were dissected to ensure mosquitoes had taken a blood meal at the time of feeding. Images were taken by fluorescence microscopy and used to quantify GFP-positive oocysts. Oocysts were counted manually using the cell counter plugin in ImageJ 1.45s software. To quantify the number of sporozoites in salivary glands, mosquitoes were anaesthetised 21 days after infection and salivary glands dissected in PBS. Salivary gland sporozoites from 15 to 20 mosquitoes were released by homogenising the glands with a micropestle and salivary gland debris were pelleted by centrifugation for 5 min at 100g. Sporozoites were subsequently counted in a haemocytometer. Groups of four mice were infected by intravenous injection of 5000 salivary gland sporozoites per mouse and the development of the parasites was followed in the whole animal every 24 hours for 72-96 hours by IVIS imaging. For imaging, mice were injected with 200 µL D-luciferin (30 mg/mL in DPBS, Source Biosciences) and imaged under isoflurane anaesthesia 10 min later using an IVIS Spectrum Imaging System (Perkin Elmer).

Once parasites were observed in the blood by IVIS (usually 72 hours post sporozoite injection), thin blood smears were taken daily for a minimum of five days and stained using Giemsa to monitor blood stage parasitaemia. To determine parasitaemia, at least 500 red blood cells were counted and parasitaemia calculated as a percentage of infected red blood cells per total number of red blood cells; percentages were then used for growth curves.

#### *Production of polyclonal antisera*

For the production of polyclonal antisera, proteins were expressed as non-biotinylated monomers by transient transfection of HEK293 cells, purified using the C-terminal his tag, and dialysed against PBS. For immunisation, 50 µg protein per animal was adjuvanted in Alhydrogel (Invivogen) and delivered intraperitoneally into six to eight-week-old female Balb/C mice. Two subsequent boosts were given with 20 µg per animal at two and four weeks after the initial injection. Sera were obtained one week after the final immunisation boost by cardiac bleed and centrifugation of the whole blood after clotting.

#### *P. falciparum cell invasion assays*

HC-04 cells were routinely cultured in Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM-F12; ThermoFisher Cat#31330038), supplemented with 10% heat inactivated fetal bovine serum and 1x Penicillin/Streptomycin. These cells are incubated at 37°C under 5% CO<sub>2</sub> and are split using trypsin when 80% confluency is reached. 12 hours before the assay, 50,000 cells per well were seeded in a collagen coated 96-well flat-bottom plate (Corning Cat# CLS3603-48EA). On the day of the experiment, salivary glands of *P. falciparum* NF-135 infected mosquitoes were dissected in 1ml DMEM-F12 media and crushed using a glass pestle to liberate the sporozoites. Sporozoite numbers were determined using a hemocytometer. Media from each of the hepatocyte-seeded wells was removed and replaced with media containing 50,000 sporozoites per well. The plate was centrifuged at 3000 rpm for 10 minutes and incubated at 37°C under 5% CO<sub>2</sub> for 3 hours. After three hours, supernatant

was removed and each well rinsed three times with PBS and trypsinized for 5 minutes to obtain single cell suspension. The trypsinized sample was neutralised in equal volumes of PBS supplemented with 10% FCS and transferred to a 96-well V-bottom plate (Costar Cat#3897). Samples were spun down at 1700 rpm for 4 minutes. Supernatants were removed and pellets were resuspended in fixation reagent as per manufacturer's instructions (eBioscience™ Foxp3/Transcription factor staining buffer set; ThermoFisher Cat# 00-5523-00) and incubated at 4°C for 30 minutes. Permeabilization buffer was added to each well after the incubation and the samples were pelleted at 1700 rpm for 4 minutes. The supernatant was removed, the pellet was resuspended in permeabilization buffer containing 3SP2 (antibody against PfCSP 1:400 dilution) conjugated to FITC and incubated at 4°C for 30 minutes. Permeabilization buffer was added to each well after the incubation and the samples were pelleted at 1700 rpm for 4 minutes. Supernatant was removed and the pellets were resuspended in 1% paraformaldehyde in PBS. Each sample was read by fluorescent activated cytometry sorting (FACS) to determine the percentage of FITC-positive (infected) cells.

#### *Immunostaining of P. berghei sporozoites*

To determine the amount and subcellular localisation of CSP, salivary gland sporozoites were dissected from infected *A. stephensi* mosquitoes 21-30 days after blood feeding on *P. berghei* infected mice. Sporozoites were isolated by homogenising the glands with a micropestle and centrifugation for 5 min at 100g to pellet salivary gland debris. Sporozoites were subsequently counted in a haemocytometer and 10,000 sporozoites each were placed onto round 13mm coverslips and air dried. Coverslips were then stored at -80°C with dessicant until used for immunostaining. For immunostaining, the coverslips were left to reach room temperature before being removed from the bag containing desiccant. Coverslips were then placed into individual wells of a 24-well plate (Nunc) and parasites were fixed with a 4% formalin solution in PBS for 10 mins on ice. After washing twice in PBS, the coverslips were incubated with 0.1% Triton in PBS on ice for 15 min to permeabilise the



parasites. All slides were washed twice in PBS, 1% BSA (Sigma) and incubated with anti-CSP monoclonal antibody, 1:1000 in PBS, 1% BSA for 1 hour at room temperature. After washing three times in PBS, 1% BSA parasites were incubated with Cy3 goat anti-mouse IgG (Life Technologies A10521), 1:500 in PBS, 1% BSA for 30 min at room temperature, protected from light. Coverslips were subsequently washed in PBS, 1% BSA and mounted on microscope slides with Vectashield HardSet containing DAPI (Vectorlabs).

## Results

### **A library of *P. falciparum* sporozoite surface and secreted proteins**

To create a recombinant protein library representing the cell surface receptor repertoire and secretome of the *P. falciparum* sporozoite, we analysed published proteome and transcriptome datasets from both *P. falciparum* and *P. yoelii* sporozoites (25, 26, 41). To identify proteins that are likely to be displayed on the surface of the sporozoite or secreted, we selected those that contain a predicted N-terminal signal peptide and/or transmembrane domain using a set of bioinformatic tools. The resulting list of 102 proteins included 57 proteins that contained one or more transmembrane-spanning regions, 11 that encoded a C-terminal hydrophobic sequence for the addition of a glycosylphosphatidylinositol (GPI) lipid anchor, and 34 that had no obvious way of being tethered to a membrane, and so were predicted to be secreted (Table S1). Of these 102 proteins, 25 are also known to be expressed on the blood stages (such as P38, RhopH3 and MSP1) and already represented within our library of previously published merozoite proteins (23, 24). For the remaining 77 proteins, we designed expression constructs in a similar manner by truncating the proteins just prior to the predicted transmembrane region or GPI-anchor, and using gene synthesis to codon optimise the coding regions for expression in human cells (23). Out of these, 69 genes could be successfully

synthesised and subcloned into the mammalian expression plasmid (Table S1), and were subsequently expressed as Cd4d3+4-tagged enzymatically monobiotinylated bait proteins by transient transfection of suspension-grown HEK293 cells. Using this approach, 58 out of the 69 proteins (84.1%) could be detected after his-tag purification from the spent cell culture medium by ELISA. As is typical for this approach, for those proteins that were detectably expressed, levels varied widely across several orders of magnitude, but could be as high as ~50 µg/mL for proteins such as TRAP and CSP. Forty-eight proteins could be produced in sufficient quantities to be used in the subsequent experiments and were resolved by SDS-PAGE and detected by anti-biotin Western blotting to determine their integrity (Fig. 1). With the exception of three proteins, PSOP7, B9 and PF3D7\_0624400, whose major species was smaller than expected, the proteins were expressed at the expected mass, although some proteins did show evidence of being processed. Those proteins that could not be detected or were expressed at very low levels included proteins that localise to the rhoptries (RALP1, ARNP and MAEBL), both the LCCL-domain containing secreted proteins LAP2 and LAP4, two members of the TRAP family (CTRP and TREP) and two members of the 6-cys family that are also expressed on gametes (P47 and P48/45). We have previously observed that proteins localised to the rhoptries of merozoites are also poorly expressed in this system suggesting that a rhoptry-specific chaperone may be required for expression using this approach (23).

### **The recombinant sporozoite proteins are immunoreactive and contain conformational epitopes**

To systematically evaluate which of the recombinant proteins within the sporozoite library contained conformational epitopes, we quantified the immunoreactivity to purified immunoglobulins pooled from adult donors living in a *P. falciparum* malaria-endemic region. Over half (35/60, 58%) of the proteins were recognised by the pooled immune serum (Fig. 2). We found that the overall immunoreactivity to sporozoite antigens was lower than for

merozoite antigens (23), as expected given that the immune system is exposed to many orders of magnitude fewer sporozoites than merozoites during an infection. The most immunoreactive proteins were those whose expression is not restricted to the sporozoite stage but which are also expressed during the blood stages. Sporozoite proteins that are not expressed by the blood stages generally exhibited a lower immunoreactivity with CSP, PIESP15, ETRAMP8, and three conserved proteins with unknown function: PF3D7\_0702900, PF3D7\_0713700 and PF3D7\_0505700, being the most immunodominant. Except for CSP, the immunoreactivity of the immune serum to heat-treated proteins was either completely lost or reduced compared to the untreated proteins, demonstrating that the immune serum was recognising conformational epitopes and suggesting that most proteins were folded (Fig. 2). The lack of heat-labile epitopes in recombinant CSP is consistent with the known immunodominance of the repetitive “NANP” repeats which are predicted to be natively unstructured (42). These data demonstrate that the majority of soluble recombinant *P. falciparum* sporozoite proteins expressed in mammalian cells contain conformational epitopes and represent a valuable resource for studying the molecular pathology of *P. falciparum* infections.

### **Identification of *P. falciparum* sporozoite surface protein complexes by systematic protein interaction screening**

To identify *P. falciparum* sporozoite extracellular protein complexes we used the AVEXIS assay which detects direct interactions between soluble recombinant ectodomains expressed as biotinylated bait and highly avid enzyme-tagged prey proteins (22). To produce the sporozoite library as prey proteins, the ectodomains were subcloned into a plasmid containing tags for pentamer formation and expressed in HEK293 cells. As expected, the levels of activity varied and 54 prey proteins were expressed at sufficient levels after their activities had been normalised to the threshold level required for the AVEXIS assay (Table S1). In total, 64 bait proteins were systematically tested for interactions against the 54 prey proteins (Fig. 3A). One

of the baits (SIAP-1 - protein 53) was expressed at low levels and therefore tested against just 35 prey proteins so that a total of 3,437 binding tests were performed. Forty-six of the proteins were present in both bait and prey libraries and consequently tested in both bait-prey orientations; the remaining proteins were therefore tested in a single orientation only, resulting in a total of 2,379 unique interactions tested. We identified three interactions: the first between a type I cell surface protein called PIESP15 and a secreted protein: PF3D7\_0702900; a second between two members of the p24 family of proteins (PF3D7\_0422100 and PF3D7\_0526900), and the third between CelTOS and a predicted secreted protein (PF3D7\_0721100). The PIESP15-PF3D7\_0702900 and p24 family interactions were identified in both bait-prey orientations in the initial screen; however, because the PF3D7\_0721100 protein did not express at sufficient levels in the prey format, the CelTOS-PF3D7\_0721100 interaction was initially identified in one orientation. For those five prey proteins that interacted with a specific bait, we repeated the binding assays against all the baits in triplicate and confirmed the interactions (Fig. 3B). The interaction between CelTOS and PF3D7\_0721100 was confirmed using a modified assay where monomeric biotinylated proteins were clustered around an alkaline phosphatase-streptavidin conjugate to form avid tetrameric preys, which then could be used to probe immobilised baits to test for interactions (43). Using this method, we could show binding of CelTOS and PF3D7\_0721100 in both orientations relative to controls (Fig. S1).

### **The interactions are conserved between orthologous proteins in *P. berghei***

To further characterise these interactions, we first used surface plasmon resonance (SPR) to establish whether the proteins could directly interact and quantify their biophysical binding parameters. For each interacting pair, one protein was purified and serial dilutions injected over its binding partner expressed as a biotinylated protein and immobilised on a streptavidin-coated sensor chip. Clear binding was observed for each interaction and the association and dissociation binding parameters were quantified by fitting the series of binding traces to a 1:1

binding model (Fig. 4A). As expected, the two interactions involving a protein predicted to be secreted exhibited slower dissociations than the membrane-tethered p24 complex proteins (21). Genome analysis revealed that the six proteins forming complexes had identifiable orthologues in the genomes of all *Plasmodium* species sequenced to date (36). To determine if the protein interactions were conserved across species, we expressed the orthologues of these proteins from the rodent parasite *P. berghei* as both baits and preys and tested if they could directly interact using the AVExis assay. We could detect robust interactions between *PbPIESP15* and *PbANKA\_0800600* and the two *P. berghei* p24 complex proteins in both bait-prey orientations (Fig. 4B, C). We were unable to determine if the interaction between CelTOS and PF3D7\_0721100 was conserved between the *P. berghei* orthologues of these proteins because repeated attempts to produce the *P. berghei* orthologue of PF3D7\_0721100 failed; therefore, this interaction was not investigated further. These data demonstrate evolutionary conservation of the interaction interfaces, which is striking since the amino acid sequence identity of the extracellular regions was relatively modest: 58 and 64% for the two p24 complex proteins; and 42 and 75% for PIESP15 and PF3D7\_0702900 respectively. We next sought to characterise the biochemical interaction between PIESP15 and PF3D7\_0702900 in more detail. Both proteins contain recognisable protein domains in their extracellular regions: PIESP15 has an L-type lectin domain, so named because it has homologies to lectins found in leguminous plants (44), and PF3D7\_0702900 contains tandem EF-hand domains. It seemed unlikely that the PIESP15- PF3D7\_0702900 interaction involved carbohydrate binding because *Plasmodium* parasites lack the glycosyltransferase enzymes necessary for complex carbohydrate modifications (45, 46), and because of the design of the recombinant proteins the only predicted glycosylation was present within the Cd4 protein tag common to all bait proteins. Both L-type lectin and EF hand domains are known to require calcium ion cofactors in their binding sites (47, 48), and consistent with this we observed that the PIESP15- PF3D7\_0702900 interaction was inhibited in a dose dependent manner in the presence of the divalent ion chelator, EDTA (Fig. 4D).

## Both genes encoding p24 complex proteins are required for *P. berghei* liver stage infection

The cross-species conservation of both the PIESP15-PF3D7\_0702900 and p24 complex protein interactions suggested that *P. berghei* was a suitable and experimentally tractable model, and enabled us to further validate these interactions by asking if parasites lacking genes encoding interacting proteins had similar infection phenotypes. To make gene-deficient parasites for all four genes, we used existing plasmids from the PlasmoGEM resource to target both *PbANKA\_0800600* and *PIESP15* (49). Gene targeting plasmids for the *P. berghei* orthologues of both p24 complex genes, *PBANKA\_1241700* and *PBANKA\_0522500*, were not available, so these were designed and constructed (Fig. S2). Gene-deficient parasites for all four genes were made by electroporating erythrocytes infected with a GFP-luciferase-transgenic “wild-type” *P. berghei* parental line and then cloned by limiting dilution. Parasites were genotyped to confirm recombination at the targeted locus (Fig. S2B, C, Fig. S3). Mosquitoes were infected with the cloned parasite lines and the infection was quantified throughout the life cycle in both the mosquito vector and mammalian host. We observed that both the  $\Delta P_{bPIESP15}$  and  $\Delta P_{bANKA_0800600}$  parasite lines had no overt infection phenotype in either the mosquito or in mice compared to the wild-type parental line (Fig. S4). This lack of infection phenotype was consistent with the inability of polyclonal antibodies to both PIESP15 and PF3D7\_0702900 to prevent invasion of *P. falciparum* sporozoites in human cells (Fig. S5). By contrast, while *P. berghei* parasites lacking either *PBANKA\_1241700* or *PBANKA\_0522500* infected mosquitoes indistinguishably from the wild-type line (Fig. 5A, B), intravenous infections of mice with salivary gland sporozoites resulted in a striking reduction in the number of parasites in the liver stage of infection for both genes (Fig. 5C, D). This was confirmed by a significant delay in the appearance of parasites in the blood of infected mice (Fig. 5E).

## Discussion

In this study we have compiled a library of soluble recombinant forms of the membrane-embedded and secreted proteins from the *P. falciparum* sporozoite. We have shown how this resource can be used to characterise the host humoral immune response in malaria patients and identify protein complexes that are important for sporozoite function. A key feature of this resource is that we have produced the proteins in a mammalian expression system to increase the chances that the proteins contain structurally-critical posttranslational modifications such as disulfide bonds and thereby adopt their native conformation, which is essential for retaining extracellular binding functions. A similar library representing the secreted and membrane-embedded proteins from the *P. falciparum* merozoite has proven very valuable to define the serology of the infected host (50), identify both parasite-parasite (9, 13) host-parasite protein interactions (51-54) as well as blood stage vaccine candidate screening (55). An important facet of this protein library is that many antigens can be assessed in parallel so that they can be directly compared in a systematic manner rather than the more common testing of proteins individually, which introduces technical variation that makes comparison very difficult. We therefore envision that this resource will be an important tool for the malaria community and have an impact in defining a better molecular understanding of the basic infection biology of the *P. falciparum* sporozoite; ultimately, these findings will contribute towards providing new opportunities for therapeutic interventions.

Of the three interactions we describe, the interaction between the p24 complex proteins PF3D7\_0526900 and PF3D7\_0422100 was of particular interest because *P. berghei* parasites lacking the orthologues of either of these genes exhibited strong sporozoite *in vivo* infection phenotypes. p24 proteins contain a GOLD (GOLgi Dynamics) domain and four can be readily identified from their primary protein sequence in *Plasmodium* spp. genomes. Most p24 family proteins are ubiquitously expressed, although some may show elevated expression in different

developmental contexts, and at least three are abundant in sporozoites and therefore represented in our library. We only observed interactions between two of these three proteins suggesting that there is binding specificity within the family, which is most likely mediated by the regions of predicted coiled coil in the luminal domain (56). p24 family proteins are mainly localised to the membranes of the early secretory pathway, and experiments in yeast have suggested that they play major roles as cargo receptors for GPI-linked proteins. Their luminal (ectodomain) regions function as a lectin by binding the oligo-mannose glycans in the GPI anchor (57) while their short cytoplasmic tails nucleate the formation of both COPI and COPII vesicle coats through their dilysine and phenylalanine-containing motifs respectively (58). One likely molecular interpretation for the phenotypes of sporozoites lacking the interacting p24 family proteins is that they are unable to appropriately traffic GPI-anchored proteins that are important for sporozoite invasion of host cells to the sporozoite surface. One obvious candidate is CSP, but we did not observe any obvious difference in the localisation or amount of CSP at the surface of sporozoites that lacked either *PBANKA\_1241700* or *PBANKA\_0522500* (Fig. S6). It is known that p24 family proteins are not essential for GPI-anchored protein trafficking since an engineered yeast strain lacking all known p24 family proteins only has a subtle phenotype that affects the rate and fidelity of protein trafficking (59). This suggests that most GPI-anchored proteins in yeast necessary for laboratory growth can be trafficked by “bulk flow” through the secretory pathway. Perhaps because CSP is so abundant at the sporozoite surface, it mainly relies on bulk flow for secretion and so is not affected by the loss of p24 family proteins. The infection phenotype we observed may therefore be due to mistrafficking of one or more of the several other GPI-anchored proteins that are expressed by sporozoites.

Of the two other interactions that we identified, we were able to show that *P. berghei* parasites genetically targeted for genes encoding the interacting proteins *PIESP15* and *PbANKA\_800600* had no overt infection phenotypes, although because sporozoites were



delivered by intravenous injection, there remains the formal possibility of a phenotype in sporozoite behaviour in the dermis. These findings are consistent with the lack of *in vivo* blood stage growth phenotypes for knockouts of these genes in *P. berghei* as part of a large scale genetic screen (49), as well as a previous smaller scale genetic screen focussed on identifying genes encoding cell surface and secreted proteins required by the ookinete to infect mosquitoes (60). There is proteomic evidence from several studies that these proteins are not restricted to sporozoites in *P. falciparum* but can be detected in other stages including gametocytes (61) and blood stages (62). Interestingly, insertional mutagenesis genetic screens have shown that both *PIESP15* and *PF3D7\_0702900* are very important for *in vitro* blood stage culture in *P. falciparum* providing evidence that these genes may play important roles in *P. falciparum* in other lifecycle stages (63). It is worth noting that the genes encoding *PF3D7\_0702900/PbANKA\_0800600* are currently annotated as a “putative centrin” in gene databases which appears to be a computational prediction based on very low (<30%) protein sequence identity. There is no functional evidence to support this claim and both *PF3D7\_0702900/PbANKA\_0800600* have confident signal peptide predictions strongly suggesting an extracellular localisation, which would not be compatible with the required cytoplasmic localisations of centrin and its known role in centrosome biology. We were not able to further investigate the functional role of the interaction between CelTOS, a protein of considerable interest due to its role in mosquito infection and cell traversal (64) and *PF3D7\_0721100* in *P. berghei* because initial attempts to genetically target the *P. berghei* orthologue *PBANKA\_0618600* did not result in any recombinants and we were unable to express the protein in a recombinant form.

In summary, we have produced a library of secreted and membrane-embedded proteins that are expressed by the *P. falciparum* sporozoite using a mammalian expression system. We believe that these proteins will be particularly useful to the malaria research community because of the challenges in obtaining large amounts of sporozoite material and solubilising

membrane-embedded proteins in their native conformation. Given that targeting the sporozoite stage either as attenuated whole sporozoites or as a sporozoite surface protein-based subunit vaccine has shown promise in malaria vaccine development (6), gaining a greater understanding of these proteins and how they function will advance not only a molecular understanding of sporozoite infection biology, but is also likely to provide important information to improve sporozoite-based malaria vaccine design.

## Acknowledgements

The authors would like to thank Ellen Bushell and Gareth Girling for plasmids, and Youri M. van Waardenburg for *P. falciparum* sporozoites. This work was supported by the Wellcome Trust grant 206194, Medical Research Council, UK (grant number MR/J002283/1), the Medical Research Council Doctoral Training (MR/J004111/1), and EU Horizon 2020 (Agreement 733273), and Dutch Research Council (NOW) talent programme veni (VI.Veni.192.171).

## Data availability

All the data are available within the manuscript and associated supplementary information. Expression plasmids encoding the proteins have been deposited and are available from the Addgene repository ([www.addgene.org](http://www.addgene.org)).

## Author contributions

JK, OB, GJW designed experiments; JK, TM, FG performed experiments except the *P. falciparum* sporozoite invasion assays which were performed by ASPY under the supervision of RWS and with help from G-JvG; KD, JK, GJW designed expression constructs; JK, OB,

GJW analysed data; OB, JCR and GJW acquired funding and supervised the research. JK, GJW wrote the manuscript.

## Competing interests

The authors declare that they have no competing interests.

## References

1. WHO (2019) World malaria report.
2. Dundas, K., Shears, M. J., Sinnis, P., and Wright, G. J. Important Extracellular Interactions between Plasmodium Sporozoites and Host Cells Required for Infection. (2019) *Trends Parasitol* 35, 129-139
3. Aliprandini, E., Tavares, J., Panatieri, R. H., Thiberge, S., Yamamoto, M. M., Silvie, O., Ishino, T., Yuda, M., Darteville, S., Traincard, F., Boscardin, S. B., and Amino, R. Cytotoxic anti-circumsporozoite antibodies target malaria sporozoites in the host skin. (2018) *Nat Microbiol* 3, 1224-1233
4. Daily, J. P. Shedding Light on the Role of the Skin in Vaccine-Induced Protection against the Malaria Sporozoite. (2018) *mBio* 9
5. Menard, R., Tavares, J., Cockburn, I., Markus, M., Zavala, F., and Amino, R. Looking under the skin: the first steps in malarial infection and immunity. (2013) *Nature reviews. Microbiology* 11, 701-712
6. Draper, S. J., Sack, B. K., King, C. R., Nielsen, C. M., Rayner, J. C., Higgins, M. K., Long, C. A., and Seder, R. A. Malaria Vaccines: Recent Advances and New Horizons. (2018) *Cell host & microbe* 24, 43-56
7. Pachebat, J. A., Ling, I. T., Grainger, M., Trucco, C., Howell, S., Fernandez-Reyes, D., Gunaratne, R., and Holder, A. A. The 22 kDa component of the protein complex on the surface

of Plasmodium falciparum merozoites is derived from a larger precursor, merozoite surface protein 7. (2001) *Molecular and biochemical parasitology* 117, 83-89

8. Kauth, C. W., Woehlbier, U., Kern, M., Mekonnen, Z., Lutz, R., Mucke, N., Langowski, J., and Bujard, H. Interactions between merozoite surface proteins 1, 6, and 7 of the malaria parasite Plasmodium falciparum. (2006) *The Journal of biological chemistry* 281, 31517-31527

9. Taechalerpaisarn, T., Crosnier, C., Bartholdson, S. J., Hodder, A. N., Thompson, J., Bustamante, L. Y., Wilson, D. W., Sanders, P. R., Wright, G. J., Rayner, J. C., Cowman, A. F., Gilson, P. R., and Crabb, B. S. Biochemical and functional analysis of two Plasmodium falciparum blood-stage 6-cys proteins: P12 and P41. (2012) *PloS one* 7, e41937

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

11. Reddy, K. S., Amlabu, E., Pandey, A. K., Mitra, P., Chauhan, V. S., and Gaur, D. Multiprotein complex between the GPI-anchored CyRPA with Pfrh5 and Pfripr is crucial for Plasmodium falciparum erythrocyte invasion. (2015) *Proceedings of the National Academy of Sciences of the United States of America* 112, 1179-1184

12. Chen, L., Lopaticki, S., Riglar, D. T., Dekiwadia, C., Uboldi, A. D., Tham, W. H., O'Neill, M. T., Richard, D., Baum, J., Ralph, S. A., and Cowman, A. F. An EGF-like protein forms a complex with Pfrh5 and is required for invasion of human erythrocytes by Plasmodium falciparum. (2011) *PLoS pathogens* 7, e1002199

13. Galaway, F., Drought, L. G., Fala, M., Cross, N., Kemp, A. C., Rayner, J. C., and Wright, G. J. P113 is a merozoite surface protein that binds the N terminus of Plasmodium falciparum RH5. (2017) *Nature communications* 8, 14333

14. Douglas, A. D., Baldeviano, G. C., Jin, J., Miura, K., Diouf, A., Zenonos, Z. A., Ventocilla, J. A., Silk, S. E., Marshall, J. M., Alanine, D. G. W., Wang, C., Edwards, N. J., Leiva, K. P., Gomez-Puerta, L. A., Lucas, C. M., Wright, G. J., Long, C. A., Royal, J. M., and Draper, S. J. A defined mechanistic correlate of protection against *Plasmodium falciparum* malaria in non-human primates. (2019) *Nature communications* 10, 1953
15. Coppi, A., Natarajan, R., Pradel, G., Bennett, B. L., James, E. R., Roggero, M. A., Corradin, G., Persson, C., Tewari, R., and Sinnis, P. The malaria circumsporozoite protein has two functional domains, each with distinct roles as sporozoites journey from mosquito to mammalian host. (2011) *The Journal of experimental medicine* 208, 341-356
16. Menard, R., Sultan, A. A., Cortes, C., Altszuler, R., van Dijk, M. R., Janse, C. J., Waters, A. P., Nussenzweig, R. S., and Nussenzweig, V. Circumsporozoite protein is required for development of malaria sporozoites in mosquitoes. (1997) *Nature* 385, 336-340
17. Sultan, A. A., Thathy, V., Frevert, U., Robson, K. J., Crisanti, A., Nussenzweig, V., Nussenzweig, R. S., and Menard, R. TRAP is necessary for gliding motility and infectivity of *Plasmodium* sporozoites. (1997) *Cell* 90, 511-522
18. Moreira, C. K., Templeton, T. J., Lavazec, C., Hayward, R. E., Hobbs, C. V., Kroeze, H., Janse, C. J., Waters, A. P., Sinnis, P., and Coppi, A. The *Plasmodium* TRAP/MIC2 family member, TRAP-Like Protein (TLP), is involved in tissue traversal by sporozoites. (2008) *Cellular microbiology* 10, 1505-1516
19. Manzoni, G., Marinach, C., Topcu, S., Briquet, S., Grand, M., Tolle, M., Gransagne, M., Lescar, J., Andolina, C., Franetich, J. F., Zeisel, M. B., Huby, T., Rubinstein, E., Snounou, G., Mazier, D., Nosten, F., Baumert, T. F., and Silvie, O. *Plasmodium* P36 determines host cell receptor usage during sporozoite invasion. (2017) *eLife* 6
20. Tavares, J., Formaglio, P., Thiberge, S., Mordelet, E., Van Rooijen, N., Medvinsky, A., Menard, R., and Amino, R. Role of host cell traversal by the malaria sporozoite during liver infection. (2013) *The Journal of experimental medicine* 210, 905-915

21. Wright, G. J. Signal initiation in biological systems: the properties and detection of transient extracellular protein interactions. (2009) *Molecular bioSystems* 5, 1405-1412
22. Bushell, K. M., Sollner, C., Schuster-Boeckler, B., Bateman, A., and Wright, G. J. Large-scale screening for novel low-affinity extracellular protein interactions. (2008) *Genome research* 18, 622-630
23. Crosnier, C., Wanaguru, M., McDade, B., Osier, F. H., Marsh, K., Rayner, J. C., and Wright, G. J. A library of functional recombinant cell-surface and secreted *P. falciparum* merozoite proteins. (2013) *Molecular & cellular proteomics : MCP* 12, 3976-3986
24. Zenonos, Z. A., Rayner, J. C., and Wright, G. J. Towards a comprehensive *Plasmodium falciparum* merozoite cell surface and secreted recombinant protein library. (2014) *Malaria journal* 13, 93
25. Lasonder, E., Janse, C. J., van Gemert, G. J., Mair, G. R., Vermunt, A. M., Douradinha, B. G., van Noort, V., Huynen, M. A., Luty, A. J., Kroeze, H., Khan, S. M., Sauerwein, R. W., Waters, A. P., Mann, M., and Stunnenberg, H. G. Proteomic profiling of *Plasmodium* sporozoite maturation identifies new proteins essential for parasite development and infectivity. (2008) *PLoS pathogens* 4, e1000195
26. Lindner, S. E., Swearingen, K. E., Harupa, A., Vaughan, A. M., Sinnis, P., Moritz, R. L., and Kappe, S. H. Total and putative surface proteomics of malaria parasite salivary gland sporozoites. (2013) *Molecular & cellular proteomics : MCP* 12, 1127-1143
27. Swearingen, K. E., Lindner, S. E., Shi, L., Shears, M. J., Harupa, A., Hopp, C. S., Vaughan, A. M., Springer, T. A., Moritz, R. L., Kappe, S. H., and Sinnis, P. Interrogating the *Plasmodium* Sporozoite Surface: Identification of Surface-Exposed Proteins and Demonstration of Glycosylation on CSP and TRAP by Mass Spectrometry-Based Proteomics. (2016) *PLoS pathogens* 12, e1005606
28. Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. (2011) *Nature methods* 8, 785-786

29. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. (2001) *Journal of molecular biology* 305, 567-580
30. Kerr, J. S., and Wright, G. J. Avidity-based extracellular interaction screening (AVEXIS) for the scalable detection of low-affinity extracellular receptor-ligand interactions. (2012) *Journal of visualized experiments : JoVE*, e3881
31. Sun, Y., Gallagher-Jones, M., Barker, C., and Wright, G. J. A benchmarked protein microarray-based platform for the identification of novel low-affinity extracellular protein interactions. (2012) *Analytical biochemistry* 424, 45-53
32. Durocher, Y., Perret, S., and Kamen, A. High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. (2002) *Nucleic acids research* 30, E9
33. Loignon, M., Perret, S., Kelly, J., Boulais, D., Cass, B., Bisson, L., Afkhamizarreh, F., and Durocher, Y. Stable high volumetric production of glycosylated human recombinant IFN $\alpha$ 2b in HEK293 cells. (2008) *BMC biotechnology* 8, 65
34. Bartholdson, S. J., Bustamante, L. Y., Crosnier, C., Johnson, S., Lea, S., Rayner, J. C., and Wright, G. J. Semaphorin-7A is an erythrocyte receptor for *P. falciparum* merozoite-specific TRAP homolog, MTRAP. (2012) *PLoS pathogens* 8, e1003031
35. Galaway, F., Yu, R., Constantinou, A., Prugnolle, F., and Wright, G. J. Resurrection of the ancestral RH5 invasion ligand provides a molecular explanation for the origin of *P. falciparum* malaria in humans. (2019) *PLoS Biol* 17, e3000490
36. Aurrecochea, C., Brestelli, J., Brunk, B. P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O. S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J. C., Kraemer, E., Li, W., Miller, J. A., Nayak, V., Pennington, C., Pinney, D. F., Roos, D. S., Ross, C., Stoeckert, C. J., Jr., Treatman, C., and Wang, H. PlasmoDB: a functional genomic database for malaria parasites. (2009) *Nucleic acids research* 37, D539-543

37. Schwach, F., Bushell, E., Gomes, A. R., Anar, B., Girling, G., Herd, C., Rayner, J. C., and Billker, O. PlasmoGEM, a database supporting a community resource for large-scale experimental genetics in malaria parasites. (2015) *Nucleic acids research* 43, D1176-1182
38. Franke-Fayard, B., Djokovic, D., Dooren, M. W., Ramesar, J., Waters, A. P., Falade, M. O., Kranendonk, M., Martinelli, A., Cravo, P., and Janse, C. J. Simple and sensitive antimalarial drug screening in vitro and in vivo using transgenic luciferase expressing *Plasmodium berghei* parasites. (2008) *International journal for parasitology* 38, 1651-1662
39. Pfander, C., Anar, B., Schwach, F., Otto, T. D., Brochet, M., Volkmann, K., Quail, M. A., Pain, A., Rosen, B., Skarnes, W., Rayner, J. C., and Billker, O. A scalable pipeline for highly effective genetic modification of a malaria parasite. (2011) *Nature methods* 8, 1078-1082
40. Janse, C. J., Ramesar, J., and Waters, A. P. High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite *Plasmodium berghei*. (2006) *Nature protocols* 1, 346-356
41. Mikolajczak, S. A., Silva-Rivera, H., Peng, X., Tarun, A. S., Camargo, N., Jacobs-Lorena, V., Daly, T. M., Bergman, L. W., de la Vega, P., Williams, J., Aly, A. S., and Kappe, S. H. Distinct malaria parasite sporozoites reveal transcriptional changes that cause differential tissue infection competence in the mosquito vector and mammalian host. (2008) *Molecular and cellular biology* 28, 6196-6207
42. Patra, A. P., Sharma, S., and Ainarapu, S. R. Force Spectroscopy of the *Plasmodium falciparum* Vaccine Candidate Circumsporozoite Protein Suggests a Mechanically Pliable Repeat Region. (2017) *The Journal of biological chemistry* 292, 2110-2119
43. Dundas, K., Shears, M. J., Sun, Y., Hopp, C. S., Crosnier, C., Metcalf, T., Girling, G., Sinnis, P., Billker, O., and Wright, G. J. Alpha-v-containing integrins are host receptors for the *Plasmodium falciparum* sporozoite surface protein, TRAP. (2018) *Proceedings of the National Academy of Sciences of the United States of America* 115, 4477-4482

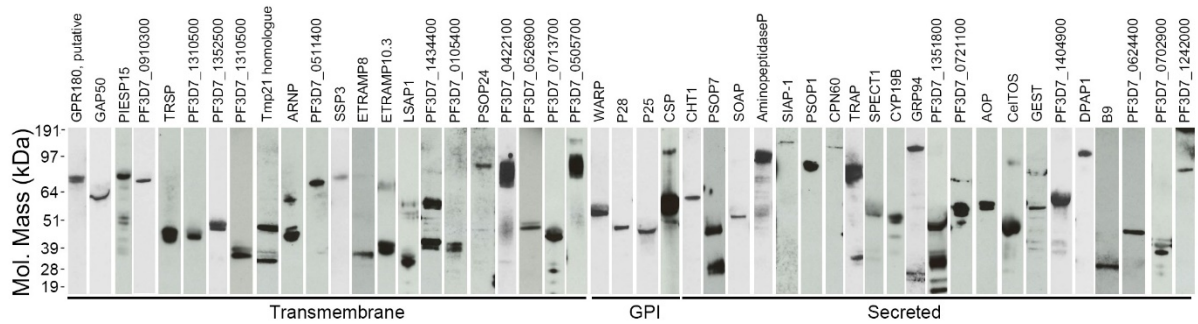


44. Fiedler, K., and Simons, K. A putative novel class of animal lectins in the secretory pathway homologous to leguminous lectins. (1994) *Cell* 77, 625-626
45. Bandini, G., Albuquerque-Wendt, A., Hegermann, J., Samuelson, J., and Routier, F. H. Protein O- and C-Glycosylation pathways in *Toxoplasma gondii* and *Plasmodium falciparum*. (2019) *Parasitology* 146, 1755-1766
46. Bushkin, G. G., Ratner, D. M., Cui, J., Banerjee, S., Duraisingh, M. T., Jennings, C. V., Dvorin, J. D., Gubbels, M. J., Robertson, S. D., Steffen, M., O'Keefe, B. R., Robbins, P. W., and Samuelson, J. Suggestive evidence for Darwinian Selection against asparagine-linked glycans of *Plasmodium falciparum* and *Toxoplasma gondii*. (2010) *Eukaryot Cell* 9, 228-241
47. Hamelryck, T. W., Loris, R., Bouckaert, J., Dao-Thi, M. H., Strecker, G., Imberty, A., Fernandez, E., Wyns, L., and Etzler, M. E. Carbohydrate binding, quaternary structure and a novel hydrophobic binding site in two legume lectin oligomers from *Dolichos biflorus*. (1999) *Journal of molecular biology* 286, 1161-1177
48. Chazin, W. J. Relating form and function of EF-hand calcium binding proteins. (2011) *Accounts of chemical research* 44, 171-179
49. Bushell, E., Gomes, A. R., Sanderson, T., Anar, B., Girling, G., Herd, C., Metcalf, T., Modrzynska, K., Schwach, F., Martin, R. E., Mather, M. W., McFadden, G. I., Parts, L., Rutledge, G. G., Vaidya, A. B., Wengelnik, K., Rayner, J. C., and Billker, O. Functional Profiling of a *Plasmodium* Genome Reveals an Abundance of Essential Genes. (2017) *Cell* 170, 260-272 e268
50. Osier, F. H., Mackinnon, M. J., Crosnier, C., Fegan, G., Kamuyu, G., Wanaguru, M., Ogada, E., McDade, B., Rayner, J. C., Wright, G. J., and Marsh, K. New antigens for a multicomponent blood-stage malaria vaccine. (2014) *Science translational medicine* 6, 247ra102
51. Crosnier, C., Bustamante, L. Y., Bartholdson, S. J., Bei, A. K., Theron, M., Uchikawa, M., Mboup, S., Ndir, O., Kwiatkowski, D. P., Duraisingh, M. T., Rayner, J. C., and Wright, G.

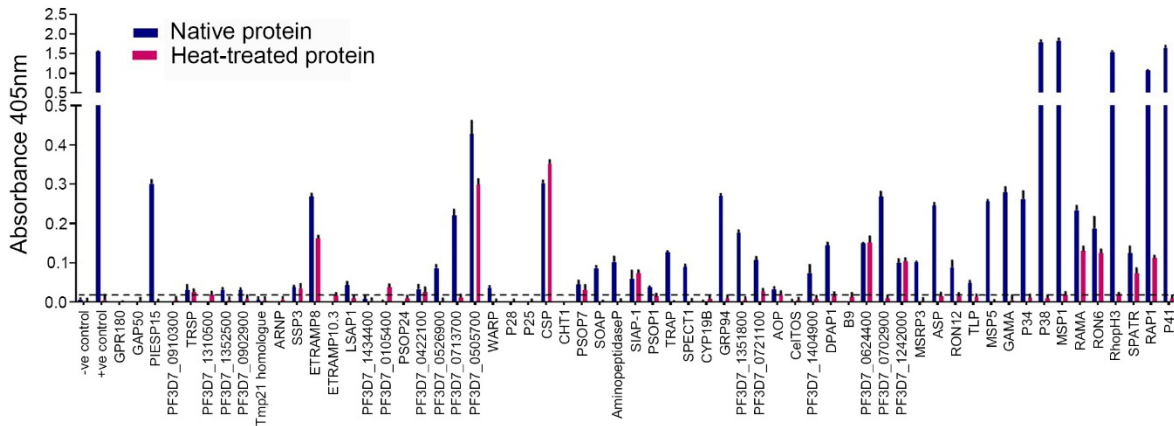
- J. Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*. (2011) *Nature* 480, 534-537
52. Wanaguru, M., Crosnier, C., Johnson, S., Rayner, J. C., and Wright, G. J. Biochemical analysis of the *Plasmodium falciparum* erythrocyte-binding antigen-175 (EBA175)-glycophorin-A interaction: implications for vaccine design. (2013) *The Journal of biological chemistry* 288, 32106-32117
53. Perrin, A. J., Bartholdson, S. J., and Wright, G. J. P-selectin is a host receptor for *Plasmodium* MSP7 ligands. (2015) *Malaria journal* 14, 238
54. Crosnier, C., Iqbal, Z., Knuepfer, E., Maciuca, S., Perrin, A. J., Kamuyu, G., Goulding, D., Bustamante, L. Y., Miles, A., Moore, S. C., Dougan, G., Holder, A. A., Kwiatkowski, D. P., Rayner, J. C., Pleass, R. J., and Wright, G. J. Binding of *Plasmodium falciparum* Merozoite Surface Proteins DBLMSP and DBLMSP2 to Human Immunoglobulin M Is Conserved among Broadly Diverged Sequence Variants. (2016) *The Journal of biological chemistry* 291, 14285-14299
55. Bustamante, L. Y., Powell, G. T., Lin, Y. C., Macklin, M. D., Cross, N., Kemp, A., Cawkill, P., Sanderson, T., Crosnier, C., Muller-Sienerth, N., Doumbo, O. K., Traore, B., Crompton, P. D., Cicuta, P., Tran, T. M., Wright, G. J., and Rayner, J. C. Synergistic malaria vaccine combinations identified by systematic antigen screening. (2017) *Proceedings of the National Academy of Sciences of the United States of America* 114, 12045-12050
56. Ciuffo, L. F., and Boyd, A. Identification of a luminal sequence specifying the assembly of Emp24p into p24 complexes in the yeast secretory pathway. (2000) *The Journal of biological chemistry* 275, 8382-8388
57. Manzano-Lopez, J., Perez-Linero, A. M., Aguilera-Romero, A., Martin, M. E., Okano, T., Silva, D. V., Seeberger, P. H., Riezman, H., Funato, K., Goder, V., Wellinger, R. E., and Muniz, M. COPII coat composition is actively regulated by luminal cargo maturation. (2015) *Current biology : CB* 25, 152-162

58. Pastor-Cantizano, N., Montesinos, J. C., Bernat-Silvestre, C., Marcote, M. J., and Aniento, F. p24 family proteins: key players in the regulation of trafficking along the secretory pathway. (2016) *Protoplasma* 253, 967-985
59. Springer, S., Chen, E., Duden, R., Marzioch, M., Rowley, A., Hamamoto, S., Merchant, S., and Schekman, R. The p24 proteins are not essential for vesicular transport in *Saccharomyces cerevisiae*. (2000) *Proceedings of the National Academy of Sciences of the United States of America* 97, 4034-4039
60. Ecker, A., Bushell, E. S., Tewari, R., and Sinden, R. E. Reverse genetics screen identifies six proteins important for malaria development in the mosquito. (2008) *Molecular microbiology* 70, 209-220
61. Silvestrini, F., Lasonder, E., Olivieri, A., Camarda, G., van Schaijk, B., Sanchez, M., Younis Younis, S., Sauerwein, R., and Alano, P. Protein export marks the early phase of gametocytogenesis of the human malaria parasite *Plasmodium falciparum*. (2010) *Molecular & cellular proteomics : MCP* 9, 1437-1448
62. Treeck, M., Sanders, J. L., Elias, J. E., and Boothroyd, J. C. The phosphoproteomes of *Plasmodium falciparum* and *Toxoplasma gondii* reveal unusual adaptations within and beyond the parasites' boundaries. (2011) *Cell host & microbe* 10, 410-419
63. Zhang, M., Wang, C., Otto, T. D., Oberstaller, J., Liao, X., Adapa, S. R., Udenze, K., Bronner, I. F., Casandra, D., Mayho, M., Brown, J., Li, S., Swanson, J., Rayner, J. C., Jiang, R. H. Y., and Adams, J. H. Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. (2018) *Science* 360
64. Kariu, T., Ishino, T., Yano, K., Chinzei, Y., and Yuda, M. CelTOS, a novel malarial protein that mediates transmission to mosquito and vertebrate hosts. (2006) *Molecular microbiology* 59, 1369-1379

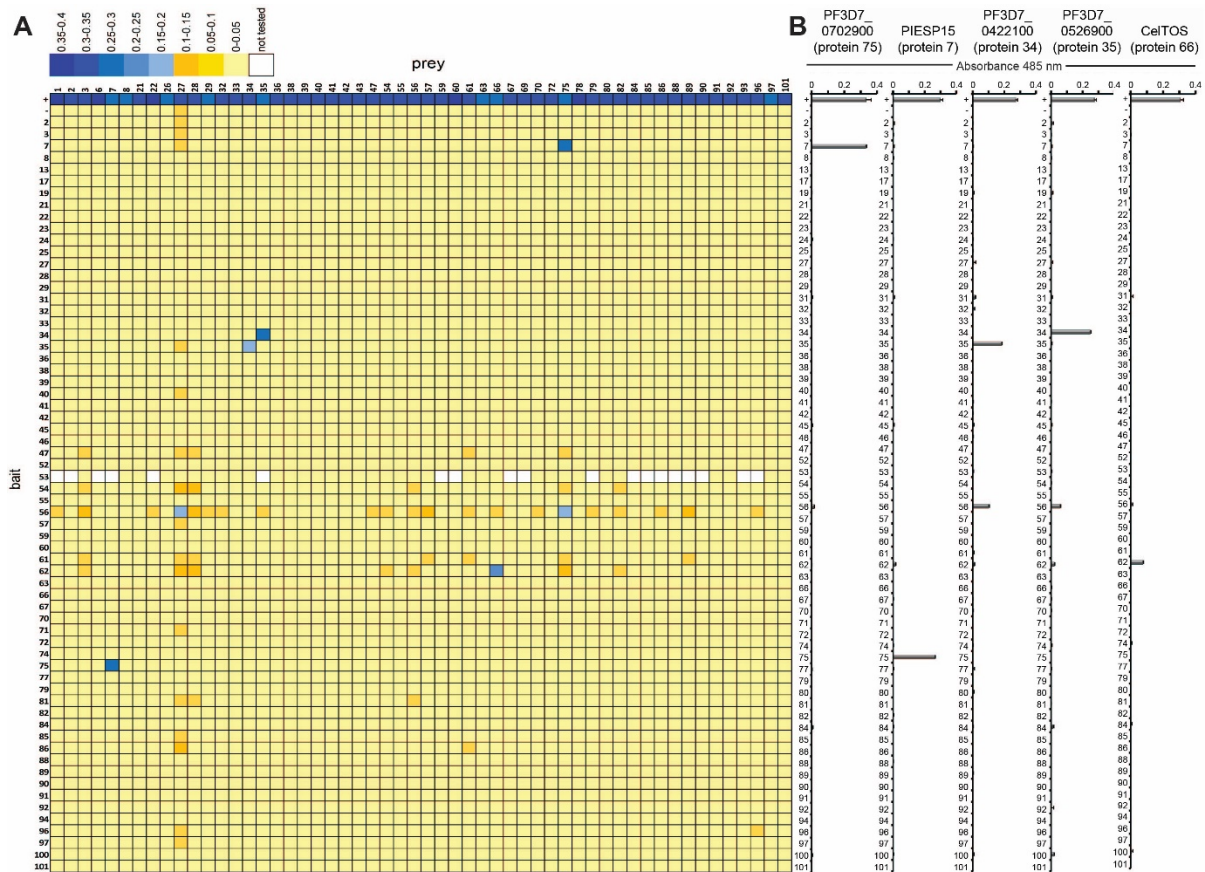
## Figures and figure legends



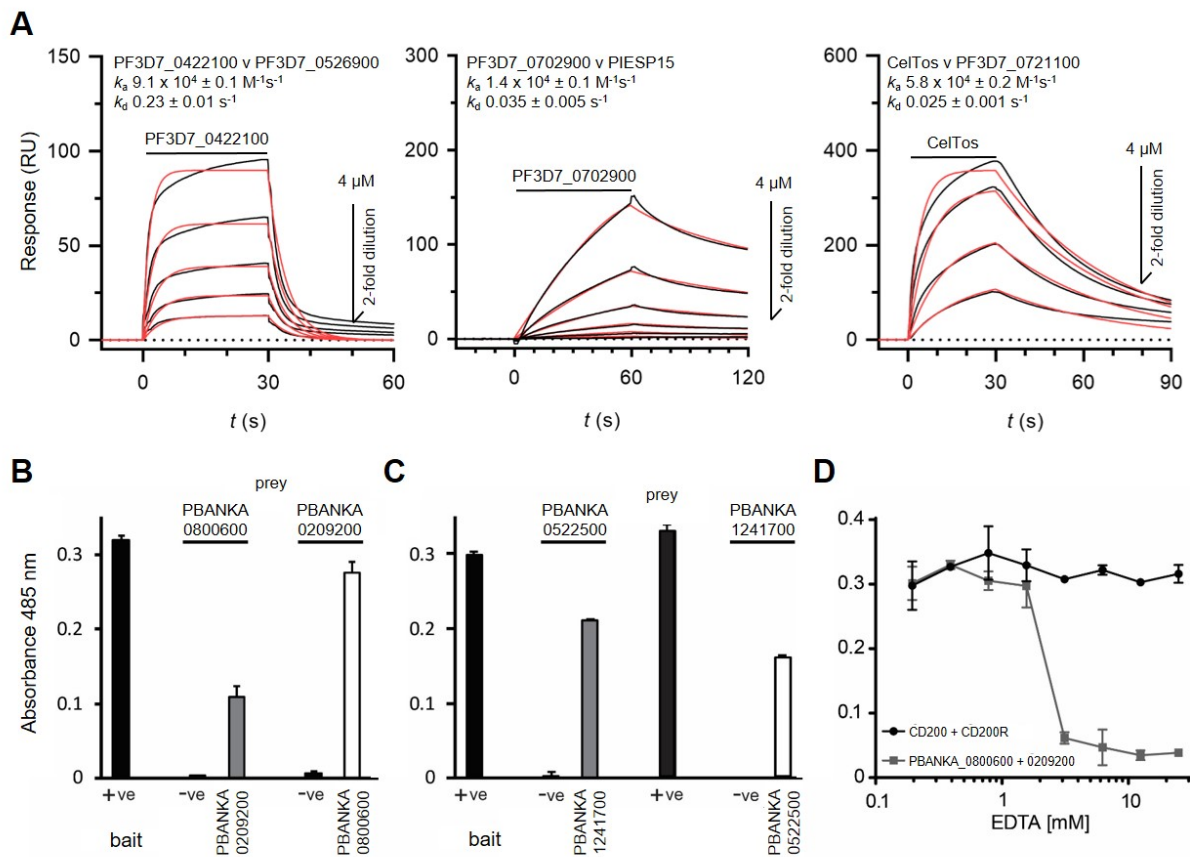
**Figure 1. A library of recombinant *P. falciparum* sporozoite cell surface and secreted proteins.** The library of sporozoite receptor ectodomains was expressed as soluble recombinant enzymatically biotinylated proteins in HEK293 cells and the purified proteins resolved by SDS-PAGE under reducing conditions. Proteins were blotted and detected using streptavidin-HRP.



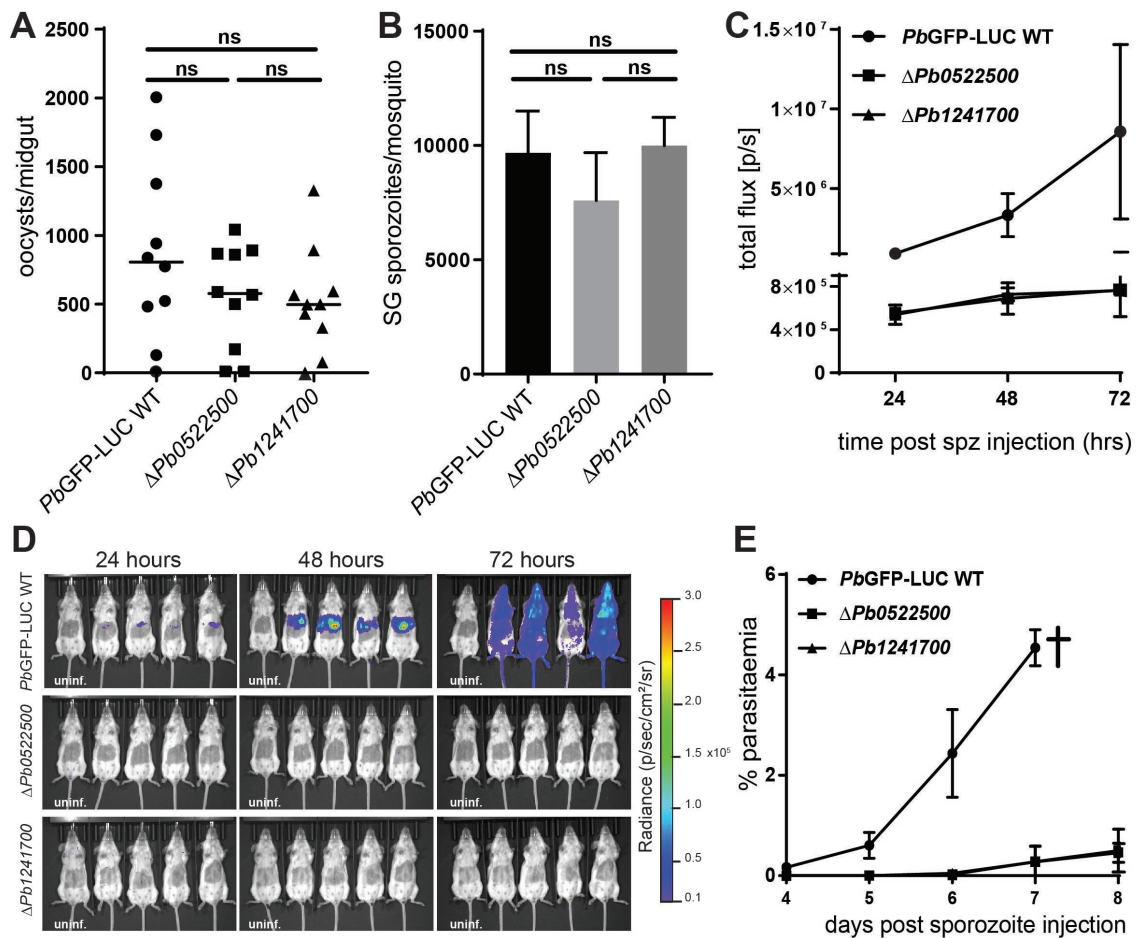
**Figure 2. The recombinant *P. falciparum* sporozoite proteins contain heat-labile conformational epitopes.** The immunoreactivity of the recombinant sporozoite proteins to immune serum was quantified by ELISA. Biotinylated proteins were normalised and captured on a streptavidin-coated microtitre plate and probed with pooled immune sera from Malawian adults (blue). Reduced response of immune serum to heat-treated (70°C for 30 min) proteins (pink) demonstrates the presence of heat-labile conformational epitopes. Note that the eleven proteins between MSP5 and P41 that are grouped on the right of the graph are additionally expressed in blood-stage parasites. Bar charts show means  $\pm$  SEM;  $n = 3$ . Positive immunoreactivity was defined as mean responses  $> 3$  SEM above the averaged control (dashed line).



**Figure 3. Identification of three *P. falciparum* sporozoite complexes by systematic extracellular protein interaction screening.** (A) The library of cell surface and secreted sporozoite proteins was expressed as biotinylated bait and enzyme-tagged pentameric prey proteins and systematically screened for interactions using the AVEXIS assay. The results are presented as a quantitative binding grid. The biotinylated baits (arranged vertically according to the numbers in Table S1) were immobilised in individual wells of streptavidin-coated microtitre plates and tested for direct binary interactions with the same proteins expressed as pentameric  $\beta$ -lactamase-tagged preys (arranged horizontally). Binding by specific baits is quantified by detecting the capture of prey proteins and hydrolysis of the colorimetric  $\beta$ -lactamase substrate nitrocefin. (B) The five named prey proteins which interacted with a bait were re-tested in triplicate against the arrayed baits and the interactions confirmed. Controls were: positive (+) Cd200 bait -Cd200R prey; negative (-) Cd200 bait with corresponding sporozoite protein prey. Bars represent mean  $\pm$  SD;  $n = 3$ .



**Figure 4. Biophysical characterisation of the three sporozoite complex interactions and conservation of interactions with orthologous proteins in *P. berghei*.** (A) Biophysical characterisation of the three *P. falciparum* sporozoite interactions by SPR. For each interacting pair, one protein was immobilised on the sensor chip and 2-fold serial dilutions of the indicated corresponding binding partner was injected to quantify binding parameters. The raw sensorgram data (black lines) were fitted to a 1:1 binding model (red lines) to derive the association ( $k_a$ ) and dissociation ( $k_d$ ) rates constants. (B) The entire ectodomains of the *P. berghei* orthologues of PIESP15 (PBANKA\_0209200) and PF3D7\_0702900 (PBANKA\_0800600), as well as PF3D7\_0422100 (PBANKA\_0522500) and PF3D7\_0526900 (PBANKA\_1241700) (C), were expressed as both bait and prey proteins and shown to directly interact using the AVEXIS assay in both bait-prey orientations. The rat Cd200-Cd200R interaction was used as a positive control (+ve); negative control bait (-ve) was Cd200. (D) The PIESP15-PBANKA\_0800600 interaction but not the control Cd200-Cd200R interaction was inhibited in a dose-dependent manner by addition of EDTA as detected by AVEXIS. Bars and data points indicate means  $\pm$  SD;  $n = 3$ .



**Figure 5. Both genes encoding the interacting p24 complex proteins are required for *P. berghei* mammalian host liver stage infection.** (A) The number of oocysts per midgut in *A. stephensi* mosquitoes infected with *P. berghei*  $\Delta$ PBANKA\_1241700 or  $\Delta$ PBANKA\_0522500 were counted and found to be not significantly different from the GFP-luciferase transgenic wild-type parental strain. Data points are individual midguts; median is indicated and statistical comparisons were performed using the Mann-Whitney Test; P-values > 0.05 were considered non-significant (ns). (B) Average number of salivary gland (SG) sporozoites per mosquito. 15-20 mosquitoes were dissected and sporozoites isolated from the salivary glands were counted. The columns represent the average number of salivary gland sporozoites across four independent experiments. Columns represent the mean with SD, statistical analysis was performed using an unpaired t-test. (C) *P. berghei*  $\Delta$ PBANKA\_1241700 or  $\Delta$ PBANKA\_0522500 parasites do not establish robust liver stage infections in mice compared to wild-type control. Groups of four mice were infected by intravenous administration of 5000 isolated salivary gland sporozoites and the resulting parasitaemia quantified in the whole animal using luciferase-based bioluminescence using an *in vivo* imaging system. Data points are mean  $\pm$  SD;  $n = 4$ . (D) Normalised bioluminescence images collected on each day after



infection; the left-most animal in each cohort is an uninfected control. **(E)** The pre-patency period before detection of blood stage parasitaemia is significantly delayed in  $\Delta PBANKA\_1241700$  and  $\Delta PBANKA\_0522500$  parasites relative to the wild-type control. WT animals were culled on day 7 because they showed signs of disease (indicated with a cross). Asexual blood stage parasitaemia was quantified by microscopic analysis of blood smears taken between four and eight days after infection. Shown are the results from one of two independent experiments with similar results. Data points are mean  $\pm$  SD;  $n = 4$ . The prepatency period was less than four days for the parental line and  $7.0 \pm 0.7$  (mean  $\pm$  S.D.  $n = 4$ ) for  $\Delta PBANKA\_0522500$  and  $7 \pm 1$  for  $\Delta PBANKA\_1241700$ .