







# The structure of a *Plasmodium vivax* Tryptophan Rich Antigen domain suggests a lipid binding function for a pan-*Plasmodium* multi-gene family

Received: 24 October 2022

Accepted: 10 August 2023

Published online: 14 September 2023

 Check for updates


Prasun Kundu<sup>1,6</sup>, Deboki Naskar<sup>1,6</sup>, Shannon J. McKie<sup>1</sup>, Sheena Dass<sup>2</sup>,  
Usheer Kanjee<sup>2</sup>, Viola Introini<sup>1,3</sup>, Marcelo U. Ferreira<sup>4,5</sup>, Pietro Cicuta<sup>3</sup>,  
Manoj Duraisingh<sup>1,7</sup>  , Janet E. Deane<sup>1,7</sup>   & Julian C. Rayner<sup>1,7</sup>  

Tryptophan Rich Antigens (TRAg)s are encoded by a multi-gene family found in all *Plasmodium* species, but are significantly expanded in *P. vivax* and closely related parasites. We show that multiple *P. vivax* TRAg)s are expressed on the merozoite surface and that one, PVP01\_0000100 binds red blood cells with a strong preference for reticulocytes. Using X-ray crystallography, we solved the structure of the PVP01\_0000100 C-terminal tryptophan rich domain, which defines the TRAg family, revealing a three-helical bundle that is conserved across *Plasmodium* and has structural homology with lipid-binding BAR domains involved in membrane remodelling. Biochemical assays confirm that the PVP01\_0000100 C-terminal domain has lipid binding activity with preference for sulfatide, a glycosphingolipid present in the outer leaflet of plasma membranes. Deletion of the putative orthologue in *P. knowlesi*, *PKNH\_1300500*, impacts invasion in reticulocytes, suggesting a role during this essential process. Together, this work defines an emerging molecular function for the *Plasmodium* TRAg family.

*Plasmodium vivax* is the malaria parasite with the broadest geographic distribution, with nearly 3 billion people at risk of infection, primarily in Latin America and Asia<sup>1</sup>. It differs from *P. falciparum* in multiple aspects of its biology, most notably by the presence of the hypnozoite, a dormant liver stage that can lead to recurrent relapses even in the absence of transmission. In addition, the sexual stage gametocytes form much more rapidly, leading to more frequent transmission that is harder to control. As a result, *P. vivax*-specific intervention strategies are a defined goal for the WHO Malaria Control Plan<sup>2</sup>. However,

progress in understanding *P. vivax* biology has been slowed in part by its absolute requirement for growth in immature red blood cells called reticulocytes, which has hampered development of in vitro culture systems.

The process by which *Plasmodium* parasites identify and invade human red blood cells has long been of interest for vaccine development, as it is essential for parasite survival and therefore malaria pathology. In *P. falciparum*, the essential invasion protein Pfrh5, which binds the erythrocyte receptor Basigin, is a high priority vaccine

<sup>1</sup>Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK. <sup>2</sup>Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA. <sup>3</sup>Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge, UK. <sup>4</sup>Department of Parasitology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil. <sup>5</sup>Global Health and Tropical Medicine, Associate Laboratory in Translation and Innovation Towards Global Health, LA-REAL, Institute of Hygiene and Tropical Medicine, NOVA University of Lisbon, Lisbon, Portugal. <sup>6</sup>These authors contributed equally: Prasun Kundu, Deboki Naskar. <sup>7</sup>These authors jointly supervised this work: Julian C. Rayner, Janet E. Deane, Manoj Duraisingh.  e-mail: [mduraisi@hsph.harvard.edu](mailto:mduraisi@hsph.harvard.edu); [jed55@cam.ac.uk](mailto:jed55@cam.ac.uk); [jcr1003@cam.ac.uk](mailto:jcr1003@cam.ac.uk)

target<sup>3,4</sup> and Phase IIa trials show that even early-stage Pfrh5 vaccines can reduce parasite growth rates in vivo<sup>5</sup>. The phylogenetic distance between *P. falciparum* and *P. vivax* means that there is no clear identifiable orthologue of Pfrh5 in *P. vivax*, and blood-stage vaccine attention has focussed primarily on Duffy Binding Protein, PvDBP, with early stage vaccine trials under way<sup>6</sup>. However, sequence and copy number variation in PvDBP between *P. vivax* isolates is well-known<sup>7</sup> raising the potential for vaccine escape and emphasising that other blood-stage targets require investigation. Of particular interest are *P. vivax* proteins that specify reticulocytes for invasion. The recognition of reticulocytes depends, at least in part, on the interaction between *P. vivax* Reticulocyte Binding Protein 2b (PvRBP2b) and CD71/Transferrin Receptor<sup>8</sup>, but the process of red blood cell invasion is a highly complex, multi-stage process<sup>9</sup>, which likely involves additional proteins.

Potential candidate proteins include the Tryptophan rich antigen multi-gene family (TRAGs), which all share a Tryptophan-Threonine-rich domain of unknown function (PFAM ID PF12319, OrthoMCL ID OG6\_145873), usually at their C-terminus. First identified in the *Plasmodium yoelii* (a species of parasite that causes malaria in rodents), TRAGs have been shown to generate highly protective antibodies in mice upon immunisation<sup>10,11</sup>. All *Plasmodium* species possess some TRAGs, but they are particularly numerous in *P. vivax* and closely related species, where they are frequently found in clusters along chromosomes, indicative of expansion through gene duplication and diversification. The precise function of TRAGs is not known in any *Plasmodium* species, but several members of the *P. vivax* TRAG family have been reported to have red blood cell binding properties<sup>12</sup> and in some cases interaction partners have been reported<sup>13,14</sup>. TRAGs also differ in expression between *P. vivax* infections in *Samiri* and *Aotus* monkeys, suggestive of a variable role, potentially as alternative invasion ligands to PvDBP<sup>15</sup>. Members of the *P. vivax* TRAG family are highly immunoreactive, with a mean seropositivity rate of around 60%, even in samples from low endemic regions<sup>16</sup>.

In this work we explored the function of *P. vivax* TRAGs in a systematic manner using a eukaryotic recombinant protein expression system that we have previously employed to identify the interaction between Pfrh5 and Basigin<sup>3</sup> and to screen for *P. vivax* vaccine candidates<sup>17,18</sup>. Antibodies raised against multiple *P. vivax* TRAGs suggested a merozoite surface localisation and revealed some limitations in current genome annotation. Focusing on one family member, PVPO1\_0000100, we establish that it binds to both erythrocytes and reticulocytes, with a preference for the latter, mediated via the C-terminal domain. Solving the crystal structure of this domain reveals an extended and curved three-helical bundle, with structural homology to lipid binding BAR domains. AlphaFold2 predictions suggest this core fold is widely conserved across *P. vivax* TRAGs and also TRAGs in other *Plasmodium* species. Lipid and liposome binding assays confirm that the PVPO1\_0000100 TRAG domain binds specifically to sulfatide and that other PvTRAGs also have lipid binding activity. Deletion of the

homologous gene in *P. knowlesi* results in a decreased preference for reticulocytes in invasion assays. This combination of biochemistry, structural biology and experimental genetics reveals a compelling molecular function for PVPO1\_0000100 and raises interesting hypotheses for the broader *Plasmodium* TRAG gene family.

## Results

### Tryptophan Rich Antigens (TRAGs) are expanded in some *Plasmodium* species

To facilitate a systematic exploration of the TRAG family we used three different methods to identify TRAGs: searching PFAM and OrthoMCL databases with the relevant domain ID for their conserved Tryptophan/Threonine-rich domain (PF12319 and OG6\_145873 respectively) and BLAST searching PlasmoDB using the C-terminal domain from PVPO1\_0000100, eliminating low confidence hits. The numbers of TRAGs identified differed depending on the method used (Table 1), as the databases all use different reference genomes and different domain definitions. However, it is clear that the TRAG family, while present in all *Plasmodium* species, has expanded significantly in the primate *Plasmodium* clade that includes four of the five major human *Plasmodium* species—*P. vivax*, *P. knowlesi*, *P. ovale* and *P. malariae*. Multiple different nomenclatures have been used for the TRAG family, which complicates analysis and comparison<sup>16,19</sup>. To facilitate understanding we use the current *P. vivax* reference genome gene IDs throughout this work and provide a systematic summary of all previous nomenclature systems in Supplementary Table 1. A neighbour joining tree based on full-length TRAG sequences from *P. vivax*, *P. knowlesi* and *P. cynomolgi* (Supplementary Fig. 1) reveals that many genes have 1:1 orthologues in all three species, suggesting these TRAGs originated in a common ancestor before these species diverged.

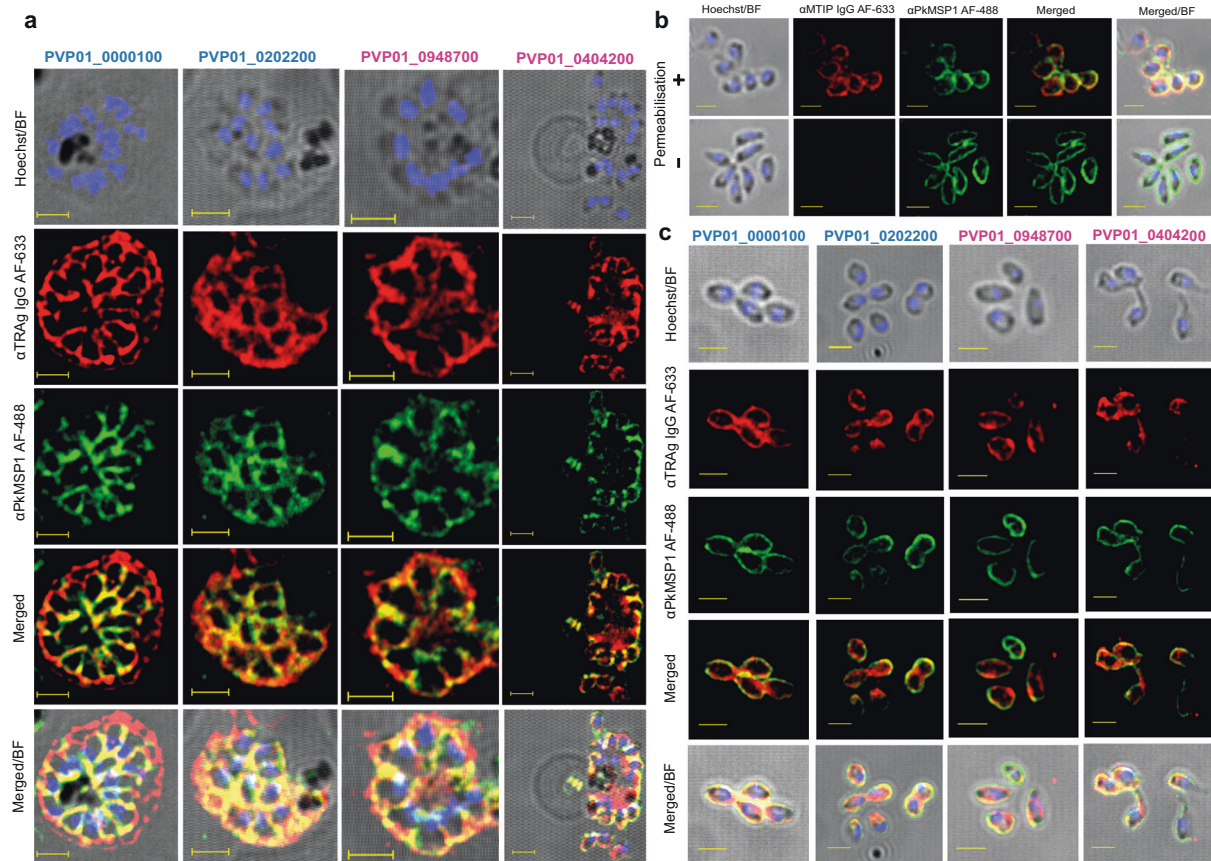
There were however, examples of closely related genes in *P. vivax* and *P. cynomolgi* with no homologue present in *P. knowlesi*, suggesting they originated after the divergence of a common ancestor of *P. vivax* and *P. cynomolgi* from *P. knowlesi*. There are also *P. vivax* TRAGs with no clear 1:1 orthologue in either species, suggesting an origin after the split of *P. vivax* and *P. cynomolgi*. This pattern suggests the process that led to the expansion of the TRAG family in the *Plasmodium* subgenus is still occurring, with new genes evolving as *Plasmodium* species diverge.

### Certain *P. vivax* TRAGs are localised on the merozoite surface irrespective of signal peptide prediction

In keeping with previous analyses<sup>16</sup> some of the *P. vivax* TRAGs we identified are predicted to have Signal Peptides (SPs) in PlasmoDB, while some are not. However, all predictions in PlasmoDB are made using one SP prediction server, SignalP<sup>20</sup>. Analysis of *P. vivax* TRAG sequences using alternative servers, Phobius<sup>21</sup> and PredSci<sup>22</sup>, identified potential SPs in several TRAGs that are not identified by SignalP (Supplementary Table 2). Notably, all *P. vivax* TRAGs lacking a predicted SP instead contained a domain annotated as 'N-terminal transmembrane', which we hypothesised may function as an SP. To test this, we generated polyclonal antibodies in rabbits against four PvTRAGs, both with and without predicted SPs. Antibodies were raised against recombinant proteins spanning the predicted ectodomain (ECD) (ie. from the end of the N-terminal hydrophobic region to the C-terminus of the protein, predicted using servers TMHMM 2.0 and Phobius; amino acid boundaries and sequences of all constructs are included in Supplementary Table 1 and Supplementary Table 3). These recombinant proteins were expressed in the HEK293E cell system, which has been used extensively for *Plasmodium* protein expression<sup>23</sup>, and purified using affinity chromatography. To test for cross-reactivity, we used the antisera to probe Western blots containing a panel of 12 purified recombinant *P. vivax* TRAGs. Since all the constructs have a Biolinker peptide and hexa-His tag in common, some level of

**Table 1 | Expansion of the TRAGs in specific *Plasmodium* species**

| Organism                | PFAM PF12319 | OrthoMCL OG6_145873 | PlasmoDB BLAST |
|-------------------------|--------------|---------------------|----------------|
| <i>P. vivax</i>         | 36           | 35                  | 38             |
| <i>P. knowlesi</i>      | 29           | 24                  | 29             |
| <i>P. cynomolgi</i>     | 31           | 33                  | 35             |
| <i>P. ovale curtesi</i> | 43           | 32                  | 40             |
| <i>P. malariae</i>      | 27           | 25                  | 35             |
| <i>P. falciparum</i>    | 3            | 3                   | 3              |
| <i>P. berghei</i>       | 7            | 5                   | 5              |
| <i>P. yoelii</i>        | 5            | 5                   | 5              |



**Fig. 1 | Multiple PvTRAGs localise to the merozoite surface, regardless of signal peptide prediction.** **a** Rabbit polyclonal IgGs raised against four PvTRAGs (colour coded as Blue for -SP and magenta +SP) were used to study the localisation of the proteins in *P. vivax* isolates. A rat antibody against PkMSP<sub>1-19</sub> was used as a merozoite surface marker for co-localisation studies. The first column is an overlay of Hoechst with brightfield images, while the last column is an overlay of Alexa Fluor 633, Alexa Fluor 488, Hoechst and brightfield (BF). The highly colocalised fluorescence intensity of the green and red channel indicates the presence of the TRAGs on the surface of the individual Pv merozoites. The colocalised fluorescence intensities are quantified in Supplementary Fig. 4. **b** Triton X100 treated purified *Plasmodium knowlesi* merozoites were stained with anti-*Plasmodium falciparum* MyoA Tail Interacting Protein (MTIP) antibodies, which localise to the Inner Membrane Complex.  $\alpha$ MTIP antibody yielded strong peripheral staining following

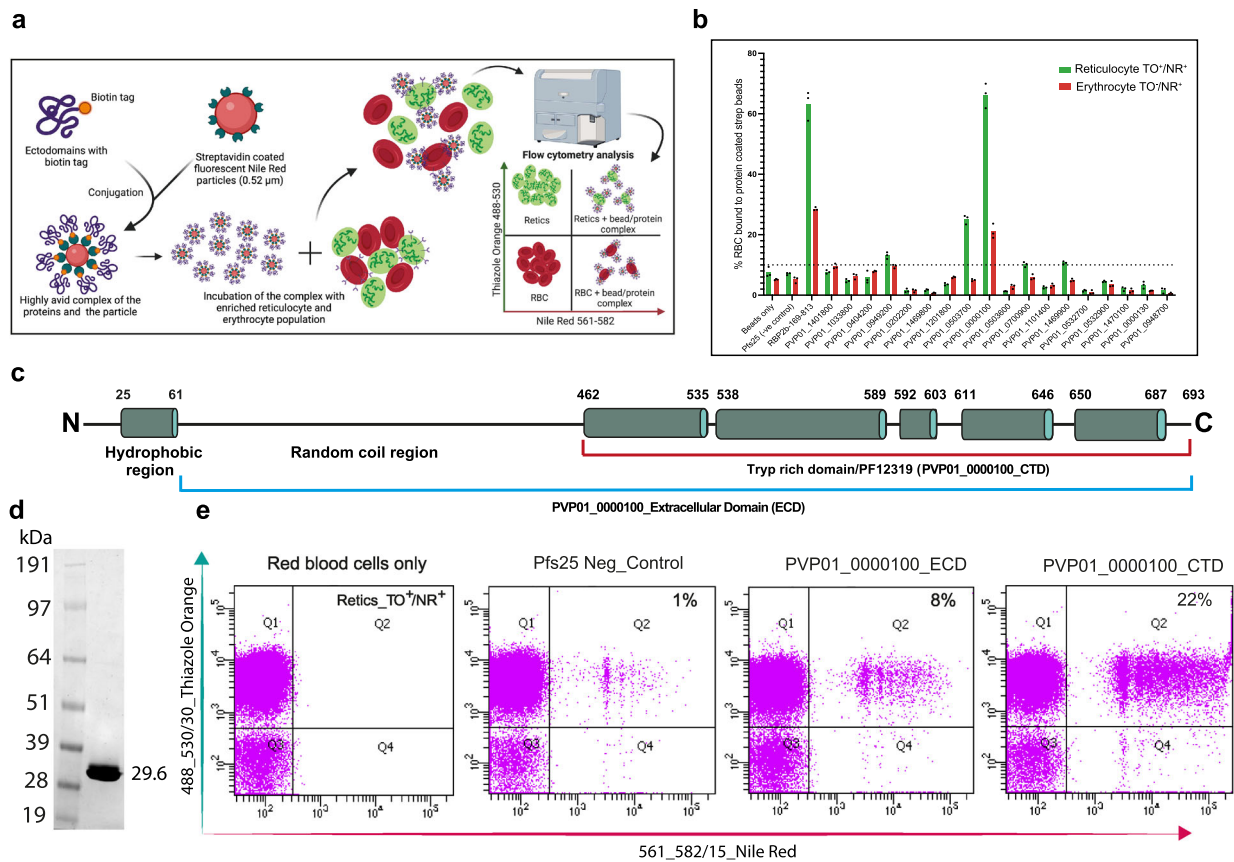
membrane permeabilisation (upper panel), but not in the absence of permeabilisation (lower panel) when the antibodies are unable to penetrate the non-permeabilised membrane. Anti PkMSP<sub>1-19</sub> was used as a merozoite surface marker and the signal remain constant regardless of permeabilization as expected. **c** Purified non-permeabilised *P. knowlesi* merozoites were stained with rabbit antibodies generated against four PvTRAGs ( $\pm$  predicted SP); the panel layout is the same as mentioned in Fig. 1a. Each  $\alpha$ TRAG rabbit antibody binds to its target on the merozoite outer surface and is colocalised with the merozoite surface marker  $\alpha$ PkMSP<sub>1-19</sub> even in the absence of permeabilisation. The Scale bars are two microns and all the images were taken in 100X with AiryScan settings. Alexa Fluor 633 goat-anti rabbit (red) and Alexa Fluor 488 goat-anti rat (green) were used as secondary antibodies for staining. Hoechst 33342 (blue) was used to stain the nucleus. All the micrographs presented are representative of  $n = 3$  independent experiments.

background was expected and Pfs25, which includes a Cd4d3 + 4 tag at the C-terminus along with the Biolinker peptide and hexa-His tag was used as a negative control. Only a very low-level of cross-reactivity was found for the anti-PVP01\_0000100, PVP01\_0532700 and PVP01\_0202200 antisera, while anti-PVP01\_0404200 had some level of cross-reactivity with one other TRAG, PVP01\_0202200 (Supplementary Fig. 2).

Immunofluorescence assays (IFAs) using ex vivo *P. vivax* parasites enriched for schizonts showed merozoite surface localisation for all antisera, whether they were raised against PvTRAGs without (PVP01\_0000100 and PVP01\_0202200) or with (PVP01\_0948700 and PVP01\_0404200) a predicted SP. Merozoite surface localisation was defined based on co-localisation with merozoite surface stain marker PkMSP<sub>1-19</sub> (Fig. 1a), with the Pearson Correlation Coefficient ( $r$ ) or colocalisation averaging 0.8 (Supplementary Fig. 4a). All four anti-PvTRAG antibodies also yielded predominantly merozoite surface localisation in *P. knowlesi*, again irrespective of whether the TRAGs were predicted to contain an SP or not (Supplementary Fig. 3). The  $r$  value for colocalisation with PkMSP1 was lower in *P. knowlesi*,

presumably because of lower cross-reactivity of anti-PvTRAG antibodies with their *P. knowlesi* orthologues (Supplementary Fig. 4b).

To confirm that staining represented the merozoite surface and not internal organelles such as the inner membrane complex (IMC, a cytoskeleton like structure that sits immediately under the plasma membrane in mature merozoites), IFAs were repeated with purified *P. knowlesi* merozoites using both permeabilising and non-permeabilising conditions. To validate the permeabilisation process the staining was optimised with a control IMC antibody raised against *Plasmodium falciparum* MyoA Tail Interacting Protein (MTIP)<sup>24,25</sup>. As expected the anti-MTIP signal was absent in the non permeabilised merozoites (Fig. 1b, Lower panel), establishing that in the absence of detergent the antibodies were unable to gain access through the merozoite surface membrane to internal organelles, whereas a strong signal for anti-MTIP was observed for the permeabilised merozoites (Fig. 1b, Upper panel). By contrast, non-permeabilised merozoites all stained strongly with antibodies against all four PvTRAGs, as did antibodies against the merozoite surface marker PkMSP<sub>1-19</sub>, confirming that they are all



**Fig. 2 | Screening PvTRAGs for erythrocyte and reticulocyte binding activity.** **a** Affinity purified biotinylated recombinant proteins were multimerised on streptavidin coated Nile Red beads and incubated with a mixture of erythrocytes and reticulocytes. Reticulocytes were labelled with Thiazole Orange (staining rRNA) before binding was assessed using flow cytometry. Erythrocytes are devoid of any genetic material, so are not labelled with TO and are in the lower half of the plots. Any binding with protein-labelled beads will cause a shift in the population of cells towards upper right (reticulocyte, green) or lower right (erythrocyte, red). The figure is created with Biorender.com. **b** Bar graph showing the binding of the multimerised PvTRAGs to reticulocytes and erythrocytes. PVRBP2b<sub>169-813</sub> previously reported to have binding activity against reticulocytes was used as positive control, while Pfs25, a sexual stage antigen, was used as a negative control for this assay. A cut-off value for positive binding was set as 10% of reticulocytes bound, based on the maximum binding value observed for the bead-only control sample. PVP01\_0000100 and PVP01\_0503700 are both substantially above this cut-off for both reticulocyte and erythrocyte binding and are termed 'strong binders', while PVP01\_0949200, PVP01\_0700900 and PVP01\_1469900 are above the cut-off for

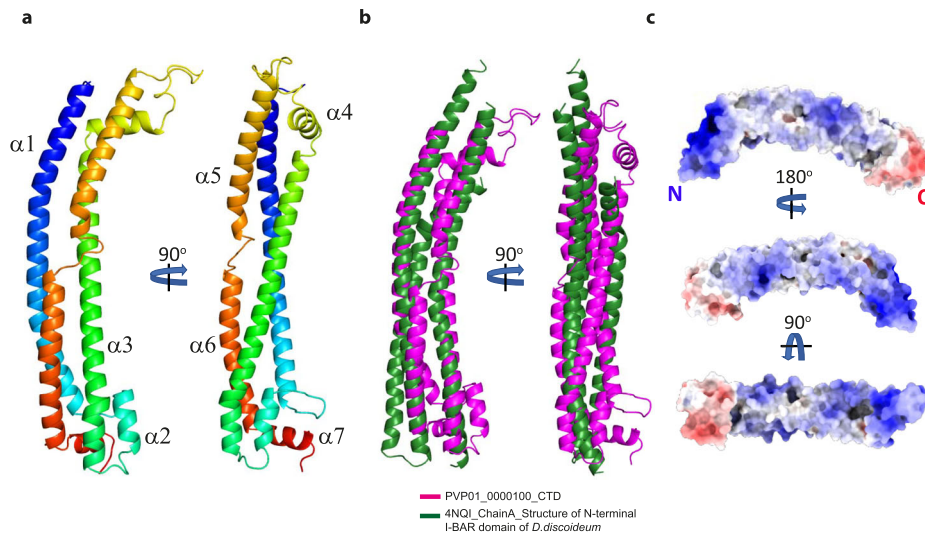
reticulocytes only, and are considered much weaker. The data represent the average of  $n = 3$  technical replicates. The gating strategy and source data are provided in the Supplementary Fig. 7 and Source data file. **c** Schematic diagram of PVP01\_0000100 secondary structure prediction using NetSurfP 3.0 and Phobius identifies a long N-terminal random coil region and a C-terminal trp-rich region which is predicted to possess an extensive helical secondary structure. The alpha helices are deep green and the random coil regions are black lines. The ectodomain (ECD) was used in binding assays and antibody generation and the C-terminal domain was used for experimental structure determination and binding assays. The boundaries are indicated for the respective regions. **d** SDS PAGE of purified PVP01\_0000100 CTD (amino acids 459-693) following a two-step purification (affinity and size exclusion); protein was stained with Coomassie Blue. The SDS page image is a representative of  $n = 3$  independent experiment. **e** The reticulocyte binding activity of the PVP01\_0000100 CTD and ECD was assessed using flow cytometry as above. All the Retic\_TO<sup>+</sup>/NR<sup>-</sup> (Q2) populations were calculated after eliminating background population (Q2) from the red blood cells only. This experiment has been repeated  $n = 2$  times independently.

localised to the outer surface of the merozoites. Together, these data support the hypothesis that the lack of a predicted SP in at least two PvTRAGs does not preclude them from entering the secretory pathway and becoming localised to the merozoite surface. Given differences in the prediction of SPs in PvTRAGs between servers, and that these servers are often trained primarily on model eukaryotes where amino acid usage can be very different to *Plasmodium* parasites, we hypothesise that the N-terminal hydrophobic region found in all PvTRAGs may actually function as an SP and support their secretion. Whether this hypothesis holds true for all TRAGs requires further validation.

### PVP01\_0000100 binds preferentially to human reticulocytes via its C-terminal domain

Given the localisation of multiple TRAGs to the merozoite surface (Fig. 1b and refs. <sup>14,26</sup>) we hypothesised that they may bind directly to

the surface of erythrocytes or reticulocytes and play a role in parasite invasion. To test this, the ECDs (as defined above) of all 33 PvTRAGs were trialled for expression using the HEK293E system. The expressed proteins are secreted into the media with a C-terminal hexa-His tag and a 31-amino acid biotinylation peptide, which can be used for affinity purification and to detect expression. Expression of 27 of the 33 PvTRAGs was confirmed by Western blot using HRP conjugated anti-His antibody (Supplementary Fig. 5). Sufficient yields for binding studies were obtained for 15 PvTRAGs. To increase avidity, affinity purified PvTRAG proteins were multimerised on Nile Red (NR) coated streptavidin beads (Fig. 2a). Reticulocytes enriched using Percoll cushions followed by anti-CD71 magnetic beads (Supplementary Fig. 6) were mixed 1:1 with mature erythrocytes and then incubated with the PvTRAG-NR beads. After incubation and washing, the mixture was stained with Thiazole Orange (TO), which binds the RNA present in reticulocytes but not mature erythrocytes and allows



**Fig. 3 | Experimentally determined structure of the PVP01\_0000100 CTD and its structural homology with BAR domain proteins.** **a** Ribbon diagram of the experimentally determined structure of the CTD of PVP01\_0000100 coloured from blue (N-terminus) to red (C-terminus). Two orientations are displayed (rotated by 90°) demonstrating the extended, slightly bent, three-helical bundle structure. **b** Superposition of the CTD of PVP01\_0000100 (magenta) with the BAR domain

from *Dictyostelium discoideum* (PDB ID 4NQI, green) showing the similarity between the fold of these two domains. **c** Electrostatic potential analysis of the surface of PVP01\_0000100 identifies positively charged patches at the N-terminal end and along the surface including the concave side of the protein. The small domain at the C-terminal end of the TRAg domain possesses a negatively charged surface.

detection of binding to reticulocytes and/or erythrocytes by flow cytometry, double gating with NR and TO. Pfs25, a sexual stage protein from *P. falciparum*<sup>27–29</sup> with no known red cell binding activity was used as a negative control, while the reticulocyte-specific CD71-interacting domain of PvRBP2b<sub>169–813</sub><sup>8</sup> acted as a positive control. The cut-off for positive binding was established at 10% of reticulocytes bound, which corresponds to the highest value observed for the bead-only control sample and is nearly equivalent to the binding level of the negative control protein (Pfs25) plus six times the standard deviation of Pfs25 binding; proteins displaying binding above this cut-off were considered significant. Of the 15 PvTRAGs tested in binding assays, PVP01\_0000100 (>60%) and PVP01\_0503700 (>20%) were the most significant, and both showed a clear preference for reticulocytes over erythrocytes (Fig. 2b), similar to the preference of PvRBP2b. (Fig. 2b; gating strategy shown in the Supplementary Fig. 7). PVP01\_0949200, PVP01\_0700900 and PVP01\_1469900 also showed binding above the cut-off but were considered to be much weaker binders.

PVP01\_0000100 and PVP01\_0503700 are close to each other in the TRAg phylogenetic tree (Supplementary Fig. 1) and share 34% sequence identity across the full-length proteins, and 48% identity in their C-terminal Tryptophan/Threonine-rich domains. Secondary structure prediction for PVP01\_0000100 using NetSurfP 3.0 suggests that the Tryptophan/Threonine-rich domain possesses a high  $\alpha$ -helical content (Fig. 2c). To test which domain was responsible for reticulocyte binding, the C-terminal region (amino acids D459–693L) was expressed alone (Fig. 2d) and assessed alongside the ECD (amino acids K62–693L). The C-terminal Tryptophan/Threonine-rich domain (CTD) showed a higher level of binding towards reticulocytes than the ECD (Fig. 2e), strongly suggesting that reticulocyte binding activity is mediated by this domain.

To control for potential interactions between the TRAg-coated Nile Red beads and the anti-CD71 beads used to enrich for reticulocytes, the assay was also performed with reticulocytes enriched using Percoll alone. Both PVP01\_0000100 and the positive control PvRBP2b bound to Percoll-enriched reticulocytes, while Pfs25 did not, establishing that binding was not impacted by the method of reticulocyte purification (Supplementary Fig. 8).

### The PVP01\_0000100 C-terminal Tryptophan/Threonine-rich domain adopts a helical fold with structural homology to Bin/Amphiphysin/Rvs (BAR) domains

To establish how the C-terminal Tryptophan/Threonine-rich domain of PVP01\_0000100 mediates reticulocyte binding we used X-ray crystallography to experimentally determine the structure of this domain. The structure was refined to 1.45 Å resolution and had one molecule in the asymmetric unit (Fig. 3a and Supplementary Table 4). Consistent with the NetSurfP predictions, the structure is composed primarily of  $\alpha$ -helices that are arranged into an extended three-helical bundle with small helical domains at each end. The conserved tryptophan residues that define the domain are distributed along the length of the helical bundle and are all buried within the fold, contributing to the arrangement of the long helices. Since experimental determination of this structure, new methods for prediction of 3D protein structures have been accelerated through use of deep-learning strategies as exemplified by AlphaFold2 (AF2)<sup>30</sup>. AF2 predictions of the PVP01\_0000100 Tryptophan/Threonine-rich domain structure closely resemble the structure we experimentally determined, with an RMSD of 1.56 Å over 213 C $\alpha$  atoms (Supplementary Fig. 9a).

The accuracy of the predicted AF2 structure therefore allows for comparison of our experimental structure for PVP01\_0000100 with AF2 predictions for the Tryptophan/Threonine-rich domain of TRAGs from other *Plasmodium* species. The quality of the predicted models was assessed by analysis of the pLDDT plots which showed high confidence scores for the helical regions of these predicted models (Supplementary Fig. 9c). Superposition of PVP01\_0000100 Tryptophan/Threonine-rich domain with its closest orthologue in *P. knowlesi* (PKNH\_1300500, RMSD 1.83 Å over 211 C $\alpha$  atoms, sequence identity 77.4%) and a much more distant orthologue in *P. falciparum* (PF3D7\_0102700, RMSD 2.87 Å over 179 C $\alpha$  atoms, sequence identity 30.2%) suggests strongly that the three-helical fold we experimentally determined is highly conserved across *Plasmodium* species (Supplementary Fig. 9b). To compare the conservation of the fold we also generated AF2 models for multiple paralogues in *P. vivax*. The C-terminal domains of four PvTRAGs were selected across the phylogenetic tree which range in sequence identity relative to PVP01\_0000100 CTD from 28 to 42%. Superposing the C $\alpha$  atoms of

AF2 models onto our experimentally confirmed PVPO1\_0000100 CTD structure indicate a high level of conservation, with RMSD values ranging from 1.2 to 2 Å (Supplementary Fig. 10a). In addition, we compared the TRAG domain of PVPO1\_0000100 with AF2 predictions of the TRAG domains from seven different species of *Plasmodium* (*P. falciparum*, *P. malariae*, *P. berghei*, *P. yoelii*, *P. ovale*, *P. cynomolgi* and *P. knowlesi*; in each case we selected the TRAG within the species that had the lowest primary sequence identity to PVPO1\_0000100). The RMSD values strongly support that the overall fold of TRAG domains are highly conserved across the entire *Plasmodium* genus, with the main differences being within short helices and loops at the distal ends of the extended helical core (Supplementary Fig. 10b).

The structure of the Tryptophan/Threonine-rich domain of PVPO1\_0000100 was compared to other known protein folds using DALI<sup>31</sup>. This identified structural homology to several membrane-binding proteins containing BAR (Bin/Amphiphysin/Rvs) domains (PDB ID 4NQI with a DALI Z-score of 7.2) (Fig. 3b). BAR domains possess extended helical bundles with a bent shape that can bind lipids and in some cases reshape membranes. Typically, BAR domains interact with negatively charged lipids on the plasma membrane via positively charged patches on the protein surface<sup>32–35</sup>. Analysis of the surface charge of the PVPO1\_0000100 revealed extended positively charged patches, mainly on the N-terminal and concave side of the protein surface (Fig. 3c), similar to the shape and surface charge of BAR domains, suggesting a potential lipid binding function.

### The PVPO1\_0000100 Tryptophan/Threonine-rich domain binds to sulfatide

In order to carry out comparative in vitro lipid-binding assays, the ECD and CTD of PVPO1\_0000100, and a negative control protein, Pfs25, were all purified as above for RBC binding assays (Fig. 4a). As PVPO1\_0000100 is expressed on the merozoite surface, the lipid species it is likely to interact with will be those present in the outer leaflet of the host plasma membrane. We therefore performed lipid-binding assays with binding strips containing a range of sphingolipids, which are present in the outer leaflet of cell surface membranes. PVPO1\_0000100 CTD showed specific and strong binding to sulfatide (Fig. 4b), an anionic glycosphingolipid found on the outer leaflet of multiple cell types, including red blood cells<sup>36,37</sup>. Commercial sphingostrips are made by dissolving lipids in organic solvent and spotting them onto a nitrocellulose membrane, resulting in lipid head groups being randomly arrayed rather than presented in a specific orientation as they would be in a membrane. These strips therefore provide a useful rapid screening tool to test binding to a wide range of lipids, but do not necessarily recapitulate physiological lipid binding. To test binding in a more relevant context, the interaction with sulfatide was confirmed using giant multilamellar liposomes composed of 48% phosphatidylcholine (PC), 2% Rhodamine-labelled phosphatidylethanolamine (Rhod-PE) and 50% sulfatide. These liposomes were incubated with 1 µM purified protein, pelleted via centrifugation, washed, resolved using SDS-PAGE and stained with Coomassie. Both PVPO1\_0000100 ECD and CTD bound more strongly to liposomes containing 50% sulfatide than they did to those without (Fig. 4c). Pfs25, the negative control, did not appreciably bind either liposome composition.

These two independent assays confirm the lipid-binding properties of the PVPO1\_0000100 Tryptophan/Threonine-rich domain that were suggested by the structural homology to BAR domains. To explore the specificity of lipid binding and specifically the influence of charge, liposome assays were repeated with liposomes consisting of phosphatidylserine (PS), another negatively charged lipid but with a very different headgroup. A titration experiment was performed with liposomes composed of increasing concentration (0–80%) of sulfatide or PS and binding quantitated using densitometry (Fig. 4d). PVPO1\_0000100 binding to sulfatide followed a sigmoidal binding

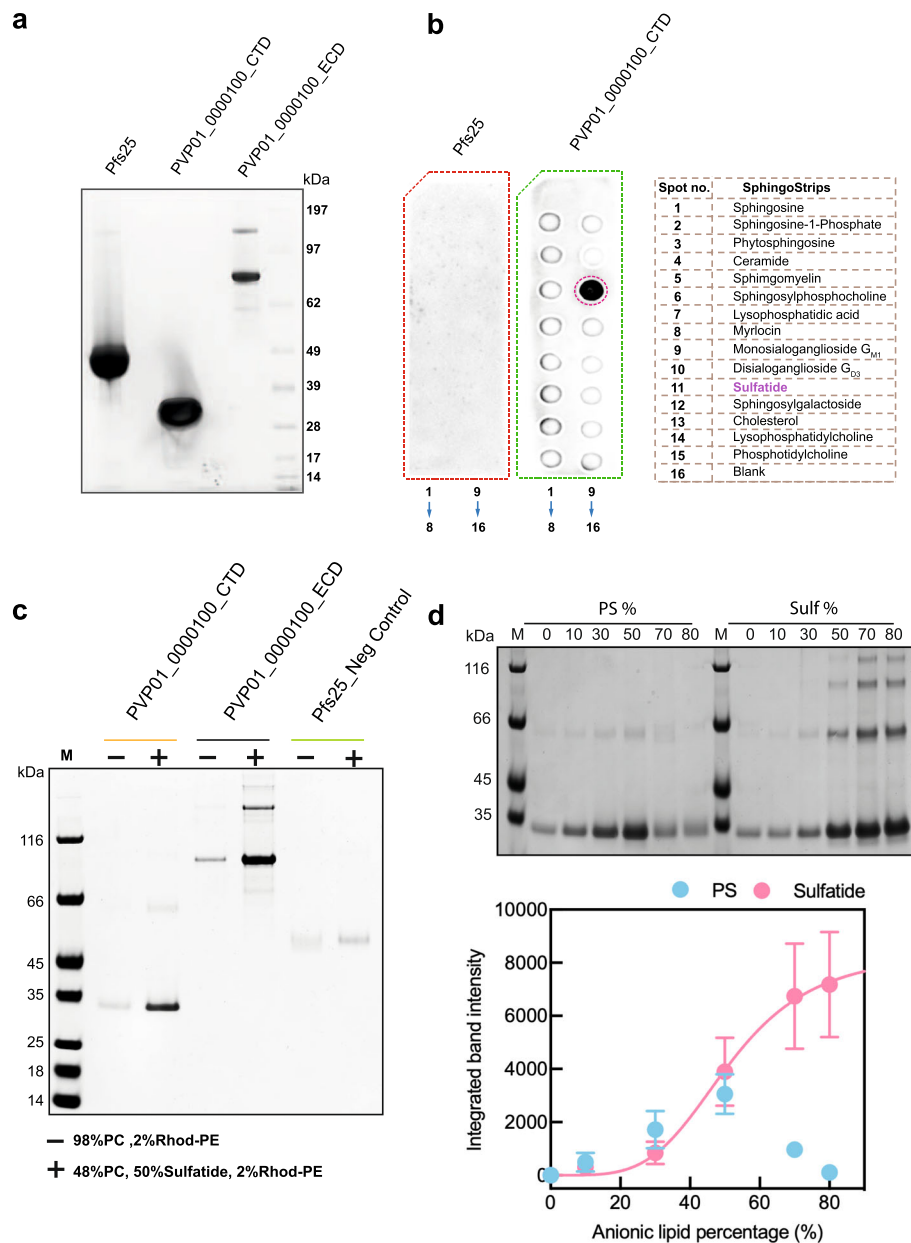
curve, suggesting a saturable and hence specific interaction. The sigmoidal binding relationship for the interaction between sulfatide and PVPO1\_0000100 also suggests there are multiple cooperative sulfatide binding sites. Binding to PS-containing liposomes did not follow this pattern, instead the binding appeared linear up to 50% PS and then decreased at higher percentages, suggesting the interaction is less specific. Therefore, lipid headgroup charge alone does not appear to be the sole mediator of PVPO1\_0000100 binding to sulfatide.

Given the predicted structural conservation of the core *Plasmodium* Tryptophan/Threonine-rich domain fold (Fig. 3b, Supplementary Fig. 9), we hypothesised that lipid-binding may be a conserved feature of this domain in multiple *Plasmodium* TRAGs. To test whether other PvTRAGs also have lipid-binding properties, five were chosen based on their expression level in HEK293E cells and to represent a range of primary amino acid identity to PVPO1\_0000100 (PVPO1\_0202200-31%, PVPO1\_0404200-29%, PVPO1\_0948700-19%, PVPO1\_1401800-30%). The ECDs of this small subset of PvTRAGs were purified using affinity and size exclusion chromatography (Supplementary Fig. 11a) and assayed for lipid binding using liposome binding assays. All four PvTRAGs had lipid-binding activity, but with different preferences (Supplementary Fig. 11b). Similarly to the PVPO1\_0000100 ECD, the PVPO1\_0948700 ECD showed a strong preference specifically towards sulfatide-containing liposomes compared to those containing only PC or both PC and PS. By contrast, PVPO1\_0202200, PVPO1\_0404200 and PVPO1\_1401800 all showed some level of binding to all three liposome compositions but with an apparent weaker affinity. This result supports the hypothesis that multiple PvTRAGs bind to membranes and that the lipid headgroup preference differs. Neurofascin 155 (NF155), a known sulfatide binding protein<sup>38</sup>, was used as the positive control for this assay (Supplementary Fig. 11b).

### The *P. knowlesi* homologue of PVPO1\_0000100 plays a role in reticulocyte invasion

*P. vivax* parasites cannot be cultured in vitro, which means the function of PVPO1\_0000100 cannot be directly tested in this species. To test the implications of sulfatide binding we therefore generated a knockout line of the *PVPO1\_0000100* orthologue (*PKNH\_1300500*) in *P. knowlesi*, a close phylogenetic relative of *P. vivax* which can be cultured in human erythrocytes<sup>39</sup>. Purified *PKNH\_1300500* bound both sulfatide and reticulocytes, like PVPO1\_0000100 (Supplementary Fig. 12). Deletion of *PKNH\_1300500* was carried out using a CRISPR-Cas9 two-plasmid approach, resulting in deletion of the entire gene and replacement with a GFP expression cassette (Fig. 5a; see Methods for details). After transfection and selection, genotyping and GFP fluorescence revealed a mixed population of wildtype and edited parasites, from which knockout lines were isolated by dilution cloning (Fig. 5b). We have previously shown that anti-PVPO1\_0000100 antiserum detects merozoite surface staining in wild type *P. knowlesi* schizonts (Supplementary Fig. 3). The fluorescence intensity of this staining was reduced by >80% in the *PKNH\_1300500* knockout line (Fig. 5c), confirming gene deletion. The weak residual staining suggests that there might be some low-level cross-reaction of anti-PVPO1\_0000100 antisera with PkTRAGs other than its closest homologue *PKNH\_1300500*, although previous tests showed only limited cross-reaction (Supplementary Fig. 2).

To assess the impact of *PKNH\_1300500* deletion, growth assays were performed in duplicate for 14 days (c. 14 cycles, as *P. knowlesi* has a 27-h cycle). A slight reduction in cumulative growth of the *PKNH\_1300500* KO line compared to wildtype *P. knowlesi* was observed (Fig. 5d). As both PVPO1\_0000100 and *PKNH\_1300500* bind preferentially to reticulocytes, we carried out invasion assays comparing the ability of the KO line to invade both reticulocytes and mature erythrocytes. Late stage parasites were incubated with a mixture of erythrocytes and reticulocytes and invasion was quantitated using flow cytometry, with anti-CD71 staining distinguishing between

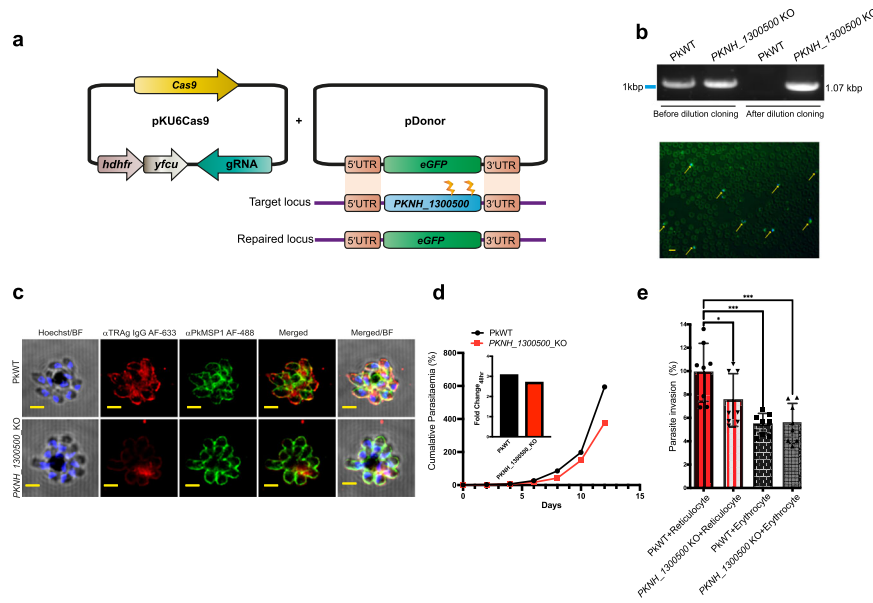


**Fig. 4 | Lipid binding properties of PVP01\_0000100.** **a** SDS-PAGE of purified PVP01\_0000100 ECD, PVP01\_0000100 CTD and Pfs25 following affinity and size exclusion chromatography. The SDS page image is a representative of  $n = 3$  independent experiment. **b** PVP01\_0000100 CTD and Pfs25 (1  $\mu$ g) were incubated with sphingostrips and binding detected using HRP conjugated anti-His antibody (1:5000) and chemiluminescent substrate. The spot numbers and orientation are shown under the blot, and show strong binding of PVP01\_0000100 CTD to sulfatide (spot no.11, magenta circle). The lipid dot blot is representative of  $n = 3$  independent experiments. **c** Liposomes composed of PC:Rhod-PE only (-, 98:2) or PC:sulfatide:Rhod-PE (+, 48:50:2) were incubated with 1  $\mu$ M PVP01\_0000100 ECD, CTD or Pfs25 and then run on SDS-PAGE and stained with Coomassie. The band intensity is increased in the presence of sulfatide for PVP01\_0000100 CTD and ECD,

while the negative control, Pfs25, did not significantly bind liposomes. The SDS page image is representative of  $n = 3$  independent experiments **d** Liposomes containing increasing percentages of sulfatide and another anionic lipid phosphatidylserine (PS) (0-80%) were incubated with 1  $\mu$ M PVP01\_0000100 CTD before being run on SDS-PAGE and stained with Coomassie. The additional bands above the CTD are higher order oligomers produced as a by-product of sample preparation prior to SDS-PAGE. Assay performed in **(d)** was done in triplicate and quantified using ImageJ and GelAnalyzer. The integrated band intensity, proportional to the amount of protein bound, was plotted against lipid percentage (average  $\pm$  S.E.M.,  $n = 3$  independent experiments) and revealed a sigmoidal binding relationship for sulfatide and a linear binding relationship for PS up to 50%, followed by a dramatic decrease in binding at higher percentages.

erythrocytes and reticulocytes, and invasion events quantitated using SYBR Green to label parasite DNA (Supplementary Fig. 13). Using a CD71 positive cutoff to define reticulocytes carries the risk of false negatives as older reticulocytes may have very low CD71 levels<sup>40</sup>, but it reduces the risk of false positives from low levels of contaminating erythrocytes that would also be CD71 negative. Wildtype *P. knowlesi*, while not restricted to reticulocytes like *P. vivax*, does have a clear preference for these immature red cells<sup>41</sup>. The preference was

observed as expected in wildtype *P. knowlesi*, whereas in *PKNH\_1300500* knockout parasites the preference was significantly reduced (Fig. 5e). The impact of *PKNH\_1300500* deletion specifically on reticulocyte invasion likely explains the lack of difference in cumulative growth rate between KO and wildtype lines, as under standard growth conditions, parasites are cultured only in the presence of mature erythrocytes, with few if any reticulocytes present (as these mature rapidly in culture).



**Fig. 5 | Gene deletion in *P. knowlesi* reveals a function for *PKNH\_1300500* in reticulocyte invasion.** **a** The homologous repair template used for gene editing of the *PVPO1\_0000100* orthologue in *P. knowlesi* (*PKNH\_1300500*) contained an eGFP expression cassette flanked by 5' and 3' UTR 800 bp homology regions. This construct was co-transfected with a plasmid expressing both Cas9 and a guide RNA. Two guide RNAs were used that are positioned near the 3' end of the gene. **b** After transfection and drug (100 nM pyrimethamine) selection, genotyping PCR revealed a mixed population of PkWT and *PKNH\_1300500* KO parasites, as indicated by PCR specific for the wildtype (WT) and knockout (KO) locus. A pure population of the edited line was obtained by dilution cloning, with no WT genotyping visible (lane 3), (uncropped, unprocessed image is provided in the source data file). Successful editing was also assessed using GFP expression (yellow arrows showing the positively edited cells) using fluorescence microscopy,  $n = 3$  independent experiment. Scale bar = 10  $\mu\text{m}$ . **c** Rabbit polyclonal IgG raised against PVPO1\_0000100 was used to detect its orthologue *PKNH\_1300500* in *P. knowlesi*. Top panel is the *P. knowlesi* wild type population and bottom panel is *PKNH\_1300500* gene edited population.

*PkMSP*<sub>19</sub> antibody (raised in rat) was used as merozoite surface marker and Hoechst was used to stain the parasite nucleus. The images were taken in 100X with AiryScan settings and are representative of  $n = 2$  independent experiments. Scale bars are two microns. **d** Growth rate of the PkWT and *PKNH\_1300500* KO line was assessed for 14 days, with parasites split every 2 days and reset to 1% parasitemia. Cumulative parasitaemia was calculated by multiplying parasitemia values detected with flow cytometry with the cumulative split values. Data points represent two independent biological replicate. **e** Invasion was quantitated using flow cytometry, with parasites detected using SYBR green staining (Q1), and erythrocytes and reticulocytes distinguished using anti-CD71. The percentage of reticulocytes (Q2) or erythrocytes (Q1) that were invaded is shown (mean  $\pm$  S.E.M) and the data combines three biological replicates, one containing technical duplicate wells, and the other two technical triplicate wells.  $P^* = 0.0207$ ,  $P^{***} = 0.0003$  (One-way ANOVA with Sidak's multiple comparisons test was used to calculate  $p$  values). The gating strategy is shown in Supplementary Fig. 13 and source data is provided in the source data file.

To explore the mechanism by which PVPO1\_0000100 might play a role in invasion, we incubated purified PVPO1\_0000100 CTD with reticulocytes and assessed the impact on reticulocyte biophysical properties using flickering spectroscopy<sup>42</sup>. The PVPO1\_0000100 CTD led to a significant decrease in membrane tension, which in *P. falciparum* has previously been shown to be directly linked with invasion efficiency, with lower tension erythrocytes being more accessible for invasion than higher tension erythrocytes<sup>43</sup>. Purified Pfs25 used at the same concentration had no effect (Fig. 6a). Tension is thought to affect invasion through its impact on the membrane wrapping energy, essentially the amount of energy required for the parasite to wrap the cell membrane around itself during the invasion process<sup>44</sup>. Addition of PVPO1\_0000100 reduced the wrapping energy compared to untreated reticulocytes (Fig. 6b), suggesting a model by which PVPO1\_0000100 binding to sulfatide modulates invasion by decreasing the energy required for merozoites to enter their host.

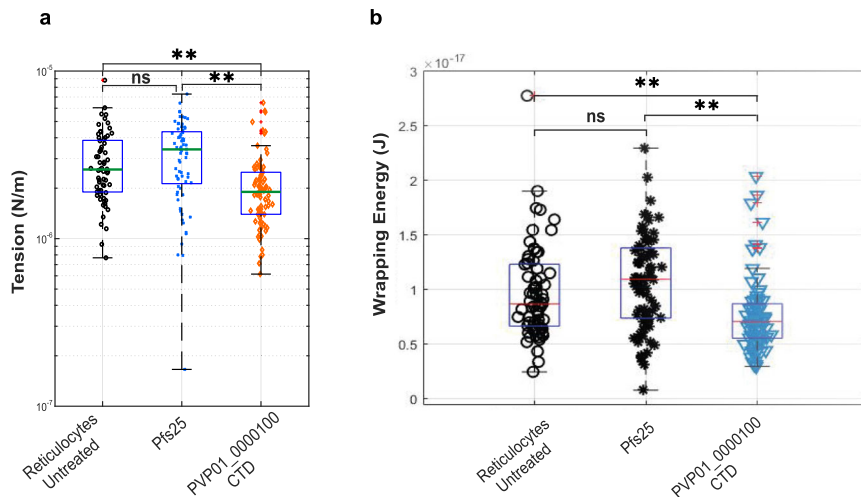
### Discussion

TRAGs have been studied in multiple *Plasmodium* species and are potential vaccine targets. However, a clearly defined molecular function for these antigens has not yet been established and some of the published literature is contradictory, suggesting a thorough systematic analysis is required. A comprehensive search and annotation of TRAGs from *P. vivax*, *P. cynomolgi* and *P. knowlesi* revealed that half of the genes are predicted to contain signal peptides, but the other half are not. From a cell biology and evolutionary perspective, it would be unusual for the same protein

family to contain members both with and without signal peptides, as secreted and cytosolic proteins generally have distinct functions. A more prosaic, technical explanation therefore seems likely. The genomes of *Plasmodium* species are all biased towards a high AT content and have patterns of codon usage that are unusual compared to other eukaryotes<sup>45,46</sup>. This can affect the efficiency of automated genome annotation for protein features such as signal peptides, which use domain prediction programmes that are usually trained on genomes from humans or model eukaryotes. We observed several TRAGs had SPs annotated by some prediction programmes but not others, and all contained a hydrophobic sequence at their extreme N-terminus, regardless of SP annotation. Immunofluorescence assays using antibodies raised against four *P. vivax* TRAGs, two with predicted SPs and two without, revealed that all four co-localised with a merozoite surface marker, indicating that all enter the secretory pathway. Based on this, we hypothesise that all PvTRAGs are potentially secreted, regardless of whether the hydrophobic region at their N-terminus is formally annotated as a signal peptide; formal proof of this clearly requires testing of further PvTRAGs. This also emphasises the limitations that automated genome annotation can sometimes have in inferring gene function, and the importance of researchers validating annotation and exploring the use of multiple annotation servers when analysing *Plasmodium* genes, particularly those of unknown function.

The presence of multiple PvTRAGs on the merozoite surface, confirmed by this and previous studies<sup>26</sup>, suggests a role for some





**Fig. 6 | Impact of PVP01\_0000100 binding on reticulocyte membrane properties.** **a** Membrane tension was measured using flickering spectroscopy for CD71+ reticulocytes in the absence of additional protein (untreated), and incubated with 100  $\mu\text{g/ml}$  proteins (Pfs25 and PVP01\_0000100 CTD). Number of cells per sample: untreated  $n = 65$ , Pfs25  $n = 73$ , and PVP01\_0000100  $n = 79$ . Pairwise comparisons between groups used Tukey's honest significant difference (HSD) test and Bonferroni correction.  $P^{**} = 0.0015$  (Reticulocytes Untreated vs PVP01\_0000100 CTD),  $P^{**} = 1.7 \times 10^{-6}$  (Pfs25 vs PVP01\_0000100 CTD). The fluctuations with contour detection of untreated red blood cell and in presence of proteins are shown in Supplementray Movies 1,2 and 3. **b** Wrapping energy was calculated using MATLAB.

Data points for untreated are 57, for Pfs25 are 73 and for PVP01\_0000100 sample are 77. Statistical comparison across groups was performed by one-way ANOVA with Bonferroni correction for multiple comparisons, while pairwise comparisons between groups used the Tukey HSD test.  $P^{**} = 5.2 \times 10^{-4}$  (Reticulocytes Untreated vs PVP01\_0000100 CTD),  $P^{**} = 1.3551 \times 10^{-5}$  (Pfs25 vs PVP01\_0000100 CTD). Each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' marker symbol. Source data is provided in the source data file.

family members in erythrocyte and reticulocyte binding. Using a new bead binding assay we observed strong binding for PVP01\_0000100 and PVP01\_0503700 with a preference for reticulocytes in both cases, which has not been reported previously in studies of other PvTRAGs<sup>12,13</sup>. PKNH\_1300500, the *P. knowlesi* homologue of PVP01\_0000100, also showed reticulocyte binding activity, although this was weaker than PVP01\_0000100, potentially due to lower protein expression affecting the bead conjugation efficiency. Experimental determination of the crystal structure of the Tryptophan/Threonine-rich domain of PVP01\_0000100 suggested a hypothesis for the molecular mechanism of this binding activity. The three helical bundle structure of this domain has homology with the BAR domain superfamily, which is characterised by  $\alpha$ -helical coiled coils that form a crescent shape with positively charged patches on the concave side, which interact with phospholipid membranes<sup>32,47</sup>. BAR domains are often involved in detecting or modulating membrane curvature, for example by interacting with negatively charged membranes and contributing to membrane bending during endocytosis<sup>48</sup>, and can also serve as membrane-binding scaffolds for other proteins, thereby participating in signalling pathways<sup>49</sup>. The PVP01\_0000100 Tryptophan/Threonine-rich domain has positively charged patches on the concave side (driven by high lysine content, 18%), and lipid dot blot and liposome binding assays confirmed that both it and its orthologue in *P. knowlesi* (PKNH\_1300500) specifically binds to the anionic lipid sulfatide. Comparison of PVP01\_0000100 TRAg binding to liposomes that maintain a similar charge but different headgroup (PS versus sulfatide) suggest the membrane binding of PVP01\_0000100 possesses specificity for sulfatide. Furthermore, the sigmoidal relationship between sulfatide percentage and TRAg binding suggests the TRAg domain may possess multiple co-operative lipid binding sites. Although several well-characterised BAR domain-containing proteins primarily interact with membranes via relatively non-specific electrostatic interactions this is not the case across the diverse BAR-domain containing protein family. Several BAR domains have been shown to possess specific

differences in their lipid binding properties<sup>50,51</sup> consistent with the binding to sulfatide seen here.

Sulfatides (also known as 3-O-sulfogalactosylceramides, sulfated galactocerebrosides, or SM4) are sphingolipids that participate in a wide range of cellular processes including protein trafficking, cell adhesion and aggregation, axon-myelin interactions, neural plasticity, and immune responses, among others<sup>52</sup>. They are particularly enriched in the myelin sheath surrounding nerve cells but are also found in the extracellular leaflet of the plasma membrane of most eukaryotic cells including red blood cells<sup>53</sup>. Sulfatides have previously been reported to play a role in malaria infection, specifically during the invasion of hepatocytes where circumsporozoite protein has been reported to bind to sulfatide<sup>54,55</sup>. Sulfatide binding has not previously been implicated in erythrocyte invasion in any *Plasmodium* species, but deletion of *PKNH\_1300500*, which binds sulfatide like its *P. vivax* orthologue, affected parasite invasion of reticulocytes, suggesting a direct link between sulfatide binding and invasion. The structural homology between the PVP01\_0000100 Tryptophan/Threonine-rich domain fold and the BAR domain suggests a clear hypothesis for how TRAGs may promote invasion. TRAg binding to lipids on the surface of erythrocytes could lead to membrane remodelling, potentially affecting biophysical properties and aiding the invasion process, which requires significant reshaping of the erythrocyte membrane in order for the merozoite to enter. In keeping with this hypothesis, addition of purified PVP01\_0000100 to reticulocytes reduced their surface tension (Fig. 6a). Lower surface tension has recently been associated with increased invasion efficiency in *P. falciparum*<sup>43</sup>, supporting the hypothesis that TRAg-lipid interactions could directly promote invasion of reticulocytes. This interaction is not absolutely essential for invasion, as PKNH\_1300500 knockout parasites can still invade both erythrocytes and reticulocytes, albeit at a reduced rate in reticulocytes. This is likely due to the large number of genes within the TRAg family, with other PKTRAGs potentially binding to either sulfatide or other lipids, providing partially overlapping functions. Whether the sulfatide

binding preference of PVP01\_0000100 also explains its reticulocyte binding preference is not clear, as we were unable to find any data quantifying the abundance of sulfatide on reticulocytes compared to erythrocytes. It is worth noting though that CD71<sup>+</sup> reticulocytes have a c. 30% larger surface area compared to mature erythrocytes, and membrane lipids vesiculate out during the process of RBC maturation, suggesting that many lipids, including sulfatide, are likely to be different between reticulocytes and mature erythrocytes<sup>40,56</sup>.

We have used the localisation and structure of a specific merozoite surface localised PvTRAg, PVP01\_0000100, to generate and test a specific hypothesis that this protein binds lipids and thereby promotes the invasion of red blood cells by *P. vivax* merozoites. An important question is whether this hypothesis holds true across all TRAGs within *P. vivax*, or indeed for all TRAGs across the *Plasmodium* genus. AlphaFold2 predicts that the structure we experimentally determined for the PVP01\_0000100 Tryptophan/Threonine-rich domain is likely to be conserved across *Plasmodium* TRAGs, despite primary sequence conservation levels <50%. This strongly suggests that all will possess a similar extended, helical fold with homology to BAR domains. This conserved fold supports the hypothesis that all TRAGs may possess lipid-binding properties. Using five different PvTRAGs we demonstrate that all can bind liposomes but with some differences in lipid-binding specificities. Given the low primary sequence homology and different surface charge patterns seen between different *Plasmodium* Tryptophan/Threonine-rich domain models, it is very possible that individual PvTRAGs may bind to different lipids, and initial tests of other PvTRAGs support that they have differences in binding activity to sulfatide. However, future studies are clearly required to explore the extent and specificity of lipid-binding across the C-terminal domain of TRAGs from different *Plasmodium* species. Other non-lipid binding functions for individual TRAGs have also been proposed, including binding to the most abundant RBC surface protein, Band 3<sup>13</sup>. While the evidence for this interaction currently lacks structural or quantitative confirmation, it is not necessarily inconsistent with the data presented here. The focus of this work is on the structurally conserved C-terminal domain, and it is quite possible that the more variable N-terminal domains of PvTRAGs could encode other functions such as specific receptor binding, including to Band 3.

It is also worth noting that it has been suggested that other *P. vivax* TRAGs may be exported into the infected erythrocyte cytosol, where they associate with Maurer's Clefts or other membrane structures<sup>57</sup>. Similarly, a *P. berghei* TRAG has recently been shown to be found inside infected erythrocytes and proposed to be involved in membrane remodelling<sup>58</sup>. These observations could still be consistent with the hypothesis that many TRAGs may have lipid binding properties, but with potentially different lipid specificities and in different intracellular locations. It is also possible that TRAG function could be different in *P. berghei* and *P. falciparum*, which only possess a small number of TRAGs, likely representing the ancestral state, compared to *Plasmodium* subgenus species such as *P. vivax* where the family has expanded 5–10-fold, and which from phylogenetic analysis appears to be still expanding within species. Importantly, this work has established a defined molecular function for the TRAG domain and provides clear hypotheses to test across species in the further exploration of this unique *Plasmodium* protein family.

## Methods

### Phylogenetic tree generation

The amino acid sequences of TRAGs from *P. vivax* (38), *P. cynomolgi* (33), *P. knowlesi* (29) were obtained from PlasmoDB<sup>20,59</sup>. The full length (including the predicted signal peptide /N-terminal domain) amino acid sequences of these TRAGs were aligned using Clustal Omega<sup>60</sup> and trees generated using PhyML<sup>61</sup>. Phylogenetic trees were constructed

using maximum likelihood method<sup>62</sup>. The evolutionary distances were computed using Jones-Taylor-Thornton Matrix model with a bootstrap test of 100 replicates<sup>63</sup> and remove gap from alignment option.

### Gene synthesis and protein expression

The region corresponding to the entire extracellular domain of 33 *P. vivax* Tryptophan rich antigens were determined using signal peptide<sup>64</sup> and transmembrane<sup>65</sup> prediction servers. Five of the 38 total *P. vivax* TRAGs were omitted because they were >200 kDa in size; proteins above this size are usually not expressed well in HEK293E cells<sup>23</sup>. To prevent the aberrant glycosylation of the proteins in mammalian cells all potential N-linked glycosylation sites (N-X-S/T/P), which are non-functional in *Plasmodium* parasites, were mutated by substituting alanine for serine/threonine/proline. The designed constructs were synthesised (TWIST biosciences, UK) with codons optimised based on the human genome. The gene region encoding the entire extracellular domain was flanked by was flanked by NotI and Ascl sites and subcloned using these enzymes into a mammalian expression plasmid that contained a 5' mouse variable  $\kappa$  light chain signal peptide and an enzymatic biotinylation sequence/hexa-His tag<sup>66</sup> at the C-terminus, just as we have previously done for multiple proteins in both *P. falciparum* and *P. vivax*<sup>67,17,23</sup>. Pfs25, the negative control used in this study was a gift from Gavin Wright (Addgene plasmid # 110991; <http://n2t.net/addgene:110991>; RRID:Addgene\_110991). The synthesised plasmids were amplified, purified, and diluted to 1 mg/ml then mixed in a 1:2.5 ratio with Polyethylenimine (PEI 40 kDa, Polysciences, Inc, #24765-1) used as a transfectant (for transfection in 50 ml, 25  $\mu$ l plasmid (1 mg/ml) in 2.5 ml and 62.5  $\mu$ l PEI in 2.5 ml freestyle media were mixed). To biotinylate the recombinant proteins, 1.25  $\mu$ l *E. coli* BirA ligase plasmid was co-transfected with the plasmid of interest. The DNA/PEI mixture was then incubated for 8 min and added to HEK 293E cells. The cells were cultured in Freestyle 293 media (Invitrogen, #12338-018) supplemented with Geneticin (#10131-027) and 1% foetal bovine serum (Sigma, #F2442) and were maintained in a shaker flask at 125 rpm, 37°C, 5% CO<sub>2</sub> and 70% relative humidity. The recombinant protein boundaries of the constructs and the amino acid sequences are listed in Supplementary Table 1 and Supplementary Table 3.

### Western blotting

The constructs cloned in the mammalian cell expression plasmids have an intrinsic signal peptide causing the recombinant proteins to get secreted into the culture supernatant. To confirm the expression of the recombinant proteins, 10  $\mu$ l of culture supernatant was collected from each transfection and mixed with 3.5  $\mu$ l of NuPAGE 4X LDS sample buffer and 1.5  $\mu$ l of sample reducing agent (Invitrogen, #NP0009). The samples were heated in 80 °C for 10 min and resolved using SDS PAGE (4–12% Polyacrylamide gel, Invitrogen, #NP0323) followed by Western blotting to Nitrocellulose membrane (Amersham, #10600003), with transfer carried out at 100 V for 60 min at 4 °C. Membranes were blocked in 3% BSA in TBST (Tris buffer saline with 0.05% Tween20) for 2 h at room temperature and probed using HRP conjugated anti-His antibody (Proteintech, Ptglab, #HRP66005, 1:20,000) overnight at 4 °C. After 3 × 10 min washes with TBST buffer, expressed proteins were detected using ECL chemiluminescent substrate (GE Healthcare, USA) in Biorad ChemiDoc MP<sup>TM</sup> Imaging System. The volume of culture required was optimised for each construct based on the expression level in the blot.

To assess cross reactivity of PVP01\_0000100 with other TRAGs, an equal amount (5  $\mu$ g) of 14 purified PvTRAGs were run on a single gel, and probed with polyclonal antibodies raised against PVP01\_0000100. Pfs25 was used as control, as it shares the biolinker hexa-His tag that is common to all the constructs. The protocol for the blotting was largely

as above, except blocking was carried out in 5% milk TBST at room temperature (RT) for 2 h followed by overnight incubation at 4 °C with Rabbit IgGs against the respective TRAgS (1:5000). Blots were washed for 3 × 10 min with TBST, then incubated with HRP conjugated anti rabbit secondary antibody (1:20,000) for 45 min. Following three further washes, antibodies were detected using ECL chemiluminescent substrate (GE Healthcare, USA) in a Biorad ChemiDoc MP™ Imaging System.

### Protein purification

To purify proteins from supernatant, HEK 293E cells were removed by pelleting at 3000 g for 20 min followed by filtering (0.22 µm, Corning). Culture supernatant was adjusted to 25 mM Imidazole and 300 mM NaCl, then loaded on to a 1 ml HiTrap Ni column (Cytiva) pre-equilibrated with binding buffer (40 mM Imidazole, 20 mM Sodium Phosphate, 500 mM NaCl, pH 7.4) at 1 ml/min flow rate. After completion of sample loading, the column was washed with 5 column volume of binding buffer to release any non-specifically bound protein. Elution was performed using step gradients with elution buffer containing 400 mM Imidazole, 20 mM Sodium Phosphate, 500 mM NaCl, pH 7.4. The eluted fractions containing recombinant protein were collected and pooled together for buffer exchange in PBS and concentration in 10 kDa Vivaspin Ultracel concentrators. The final concentrations were measured by Nanodrop OneC UV-Vis Spectrophotometer (Thermo Fisher) and kept in 4 °C until further use.

### Immunisation and IgG purification

Polyclonal IgGs were developed in New Zealand White Rabbits (Eurogentec, Belgium) as per the company's protocol. In brief, 0.8 mg of purified recombinantly *P. vivax* TRAg protein was injected in two rabbits (0.4 mg each) with 100 µg/immunisation/rabbit. Freund's complete adjuvant was used for first immunisation followed by Freund's incomplete adjuvant in booster doses given on days 14, 28 and 58. The final bleed was collected on day 87.

Total IgG was purified from the final bleed rabbit sera using Protein G Gravitrapp (GE Healthcare, Cytiva). The sera were diluted as 1:5 in binding buffer (20 mM Sod. Phosphate, pH 7.4). The purification was done using Ab purification kit and the protocol was followed as per the kit's manual. In brief, the diluted sera were loaded in the pre-equilibrated Gravitrapp column and the eluant reloaded three times to enrich the binding and the beads washed with 10 ml of binding buffer. The IgG was finally eluted with 5 ml of elution buffer (0.1 M glycine-HCl, pH 2.7) poured into the column and collected directly in a 10 kDa centrifugal filter unit (Millipore) containing 0.225 ml of neutralisation buffer (1 M Tris-HCl, pH 9.0). The IgGs were buffer exchanged three times in RPMI 1640 media (Gibco) for any follow up assay. The final concentration of the purified IgG was determined using a Nanodrop one UV-Vis Spectrophotometer (Thermo Fisher).

### Reticulocyte isolation

**Enrichment of CD71<sup>+</sup> reticulocytes.** Fresh peripheral whole blood (withdrawn within 48 h) provided by NHS Blood and Transplant (ethical approval University of Cambridge REC HBREC.2019.40 and NHS REC 20/EE/0100). Whole blood was centrifuged for 7 min at 1500 g and the plasma and the white cell layer aspirated, then resuspended with equal volume of phosphate-buffered saline buffer, PBS (Gibco, UK). The solution was passed through a Plasmodipur filter to remove leucocytes (EuroProxima, #8011 Filter 10U) and washed twice with PBS at 1000 g for 10 min. The leucocyte-depleted red blood cell pellet was then resuspended in 10 ml PBS and 5 ml were carefully layered on 6 ml of 70% (v/v) isotonic Percoll cushion (GE Healthcare, #17089101). After centrifugation at 1200 × g for 15 min, brakes off, a thin band(s) of reticulocytes was collected at the Percoll interface and spun down at 300 g for 10 min. All the following steps were carried out at 4 °C to preserve reticulocytes. Microscopic counting of

reticulocytes stained with new methylene blue (NMB, Sigma-Aldrich, #R4132), and the Percoll separation achieved purity of 7–20% reticulocytemia, depending on the blood sample, as previously reported<sup>68</sup>. To purify immature CD71<sup>+</sup> at purities up to 90%, we used CD71 magnetic MicroBeads (Miltenyi Biotec, # 130-046-201) and followed the company protocol. We confirmed the high purity through sub vital staining (NMB) and microscopy. Reticulocytes were stored at 4 °C and used for binding experiments within 24 h. A binding assay was also performed only with percoll enriched retics without CD71<sup>+</sup> Ab enrichment to rule out any non specific binding with Anti CD71<sup>+</sup> coupled beads.

### Reticulocyte binding assay

Purified biotinylated proteins were multimerised on streptavidin coated Nile Red fluorescent beads (Spherotech Inc, 0.52 µm, 0.1% w/v, binding capacity 1.65 nM Biotin-FITC to 1 mg beads). The beads were sonicated to disperse them evenly using a bath sonicator for 10 min with a pulse duration 30 s and 10 s interval between each pulse, all carried out at 4 °C. The sonicated beads were then blocked using 0.2% BSA buffer (in PBS). In a 96 U bottom plate (Nunc 165306, Thermo), 1 µl bead was incubated with 1 µg protein per well for 45 min at 4 °C. Unbound protein was then removed by washing the beads by centrifugation at 3000 g for 10 min at 4 °C. Purified/CD71 enriched reticulocytes were mixed with equal number of erythrocytes (from the same blood sample, using material pelleted under the Percoll cushion) and prestained with Thiazole Orange (10 mg/ml stock solution, 1 µl from the stock added in 10 ml of 0.2% BSA buffer). From this mix 0.5 million cells were added to the washed protein coated beads per well of 100 µl reaction volume and incubated for 45 min at 4 °C. The plate was centrifuged at 250 g for 10 min and the supernatant containing unbound protein was removed carefully. The cell pellet was then resuspended gently with 0.2% BSA buffer and was centrifuged for 10 min at 250 g, this washing step was repeated three times. The cells were finally resuspended in 200 µl 0.2% BSA and analysed using flow cytometry (BD Fortessa). Data were acquired using the FACS Diva software and post processing was performed using FlowJo (ver 10.2). PvRBP2b<sub>169-813</sub> and Pfs25<sub>23-193</sub> proteins were used as positive and negative controls respectively. Nile red and Thiazole Orange emissions were detected using 561 and 488 nm lasers respectively. The voltage settings for the respective lasers were FSC-367, SSC-293, 640/670-500, 488-433 and 561/582-480. The gating strategy has been shown in Supplementary Fig. 6. The NR positive binding populations for both retics and erythrocytes has been calculated in terms of percentage with respect to the total retics (TO positive population) and total erythrocytes respectively.

### Structural methods

**Protein cloning, expression and purification.** All protein structural work was conducted prior to the availability of AlphaFold2. To identify constructs best suited for structural analysis, secondary structure composition for PVP01\_0000100 was predicted using NetSurf 3.0<sup>69</sup> identifying potential domain boundaries. This identified a C-terminal domain (CTD, D459-693L) of PVP01\_0000100 possessing high helical content that was selected for expression and crystallisation studies. The *PVP01\_0000100* CTD was PCR amplified from the plasmid containing the *PVP01\_0000100* ectodomain (ECD, K62-693L) construct and cloned into the pHLSec vector with AgeI/KpnI restriction sites. The construct was expressed in 200 ml of HEK293F cells (Thermo #R79007) and purified with Ni-NTA (HiTrap, Cytiva) followed by buffer exchange in size exclusion buffer (20 mM Tris, 150 mM NaCl, pH 7.4) for three times in a concentrator, MWCO 10 kDa (Millipore) prior to size exclusion (Superdex 75 16/600, analytical Hi load column, Cytiva). The protein was eluted and collected fractions were pooled together and concentrated using a centrifugal concentrator, 10 kDa MWCO (Millipore, Fischer scientific) to 10 mg/ml for crystallisation trials.

**Crystallisation and structure determination.** Initial crystallisation trials were performed in 96-well plates using sitting drops composed of 200 nL protein plus 200 nL reservoir solution. Initial crystals were identified in broad matrix screens (PACT PREMIER, Molecular Dimensions) and diffraction-quality crystals grew in 0.1M SPG, pH 4.0, 25% (w/v) PEG1500. The crystals were cryo-protected with 20% (v/v) glycerol added as supplement to the reservoir solution prior to flash-freezing in liquid nitrogen. Data were collected at the Diamond Light Source I04 beamline. Data processing and scaling was performed using Xia2 DIALS<sup>70</sup>. Phaser<sup>71</sup> was used for molecular replacement using a model generated from RoseTTAFold<sup>72</sup>. ArpWarp<sup>73</sup> was used to build an initial model of PVP01\_0000100 CTD, followed by manual building in Coot<sup>74</sup>. Structure refinement was carried out with Coot<sup>74</sup>, ISOLDE<sup>75</sup> and phenix.refine<sup>76</sup>. Data collection and refinement statistics are shown in Supplementary Table 4. Illustrations were generated using The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. The surface charge analysis for the structure of PVP01\_0000100 CTD was carried out with Adaptive Poisson-Boltzmann Solver (APBS) programme (server.poissonboltzmann.org)<sup>77</sup>. The results were analysed in UCSF chimera<sup>78</sup>. The atomic coordinates and structure factors have been deposited in the Protein Data Bank, [www.pdb.org](http://www.pdb.org) under accession code 8ARL.

**Lipid dot blot.** The membrane lipid strips (Echelon P-6002) were blocked with 3% fatty acid free BSA (Sigma, A4612) overnight at 4 °C. The blocked membrane was incubated with 1 µg of PVP01\_0000100 CTD, PVP01\_0000100 ECD and Pfs25 (negative control) for 2 h at RT. The proteins were incubated overnight followed by three washes with PBST. The strips were then incubated with anti-His-rabbit HRP 1:5000 (Proteintech, Ptglab, USA). Detection was performed with an ECL reagent (GE Healthcare, USA).

**Liposome binding assay.** Phosphatidylcholine (PC) was purchased from Merck (#840051 C). Sulfatide (#131305 P), 1,2-dimyristoyl-sn-glycero-3-phosphoethanolamine-N-(lissamine rhodamine B sulfonyl) (Rhod-PE) (#810157 P) and phosphatidylserine (PS) (#840029 P) were purchased from Avanti®. Lipids were dissolved/suspended in chloroform and the final molar mixtures added to a glass vial before the chloroform was evaporated under a nitrogen stream. To all liposomes prepared, 2% Rhod-PE was incorporated to aid with liposome visualisation and PC was the carrier lipid, with percentages of sulfatide or PS varying from 0 to 80%. The resultant lipid film was hydrated using 50 mM HEPES (pH 7.4) and 150 mM NaCl to form multilamellar liposomes. A final liposome concentration of 0.8 mM in 50 µL was incubated with 1 µM protein for 30 min at room temperature with rotation. The liposomes were then pelleted via centrifugation at 21,000 g and the supernatant removed. The pellet was washed twice before resuspension in 10 µL of 50 mM HEPES (pH 7.4) and 150 mM NaCl and 10 µL of 2X loading dye and then ran on a NuPAGE™ 4–12% Bis-Tris Gel (Invitrogen, #NPO335) and stained with InstantBlue® (Abcam, #HSB1L). Densitometric quantitation of the PS and sulfatide titration assay was performed using ImageJ and GelAnalyzer and then graphed in Prism 9.

**In vitro culture of *P. knowlesi*.** Human O<sup>+</sup> erythrocytes were used for the parasite culture and were purchased from NHSBT. The use of human cells for this work was approved by the NHS Cambridge South Research Ethics Committee (REC reference<sup>20</sup>/EE/0100) and by the University of Cambridge Human Biology Research Ethics Committee (HBREC.2019.40). *Plasmodium knowlesi* AI-H.1 were grown in RPMI 1640 medium supplemented with Albumax (Gibco, # 041-96698 A), L-Glutamine (Gibco, # 25030-024), 10% Horse serum (Gibco, # 26050-70), and Gentamicin (Gibco, # 15750-045). The cultures were kept at 37 °C with 2% haematocrit and kept in an incubator containing gas

mixture of 3% CO<sub>2</sub>, 1% O<sub>2</sub> and 96% Nitrogen. The cultures were monitored thrice a week by counting parasitaemia using light microscopy with media change or splitting as appropriate<sup>39</sup>

### Genetic modification of *P. knowlesi* parasites

**PVP01\_0000100 knockout construct design and assembly.** The constructs, guide RNAs and primers were designed in Benchling. The 800 bp fragments immediately 5' and 3' upstream of *PVP01\_0000100* were amplified using *P. knowlesi* genomic DNA obtained from Pk infected erythrocytes using a DNA blood kit (Qiagen) according to the manufacturer's protocol. The eGFP gene was amplified from a plasmid which is used in transfection as positive control. The 5' and 3' homology regions (88 and 85 ng/µl), eGFP gene (40 ng/µl) and pUC19 EcoRI/HindIII digested vector (70 ng/µl) were then added to the Gibson assembly mix at a 1:1:1 ratio and incubated for 1 h at 50 °C.

**Cas9 vector and gRNA assembly.** The cloning vector pKU6-Cas9*ccdB* (1 µg) was digested and dephosphorylated using Fast Digest *SapI* (*Lgul*) (Thermo, FD1934) and Alkaline phosphatase (Thermo, EF0654) and ran on a 1% agarose gel followed by purification using agarose gel purification kit (Macherey and Nagel). The Guide RNAs were synthesised from Sigma and reconstituted by mixing 10 µM of the forward and reverse strands for each guide with 1 µl of 10X ligation buffer (NEB), 0.5 µl T4 polynucleotide kinase (NEB) and 6.5 µl nuclease free PCR water. Annealing was carried out by incubating at 37 °C for 30 min, then increasing to 95 °C for 5 min before cooling at 25 °C at a ramp speed of 0.1 °C/sec. Annealed primers were then diluted to 1 µl in 200 µl and ligated to the digested and dephosphorylated pKU6Cas9*ccdB* (50 ng/µl). Two gRNAs were constructed for knocking out *PVP01\_0000100* orthologue *PKNH\_1300500*. The details of the reaction mixtures and primers for repair template and gRNA vector assemblies are provided in the Supplementary Table 5.

The assembled pUC19 Repair template and gRNA/Cas9 vector were transformed in *E. coli* Ultracompetent cells (XL-10 gold, Agilent) and plated onto Amp<sup>+</sup> LB plates. Colonies positive for the insertion were sequence verified (Genewiz) before preparation using a midprep purification kit (Macherey and Nagel) for transfection.

### *P. knowlesi* synchronisation, schizont enrichment and transfection.

Synchronization was performed by enriching late-stage parasites using Histodenz (Sigma Aldrich, #D2158). The parasites were centrifuged at 1500 g, and the pellet was resuspended in 5 ml complete media and then layered on top of 5 ml of 55% Histodenz (prepared by adding 2.25 ml of complete medium + 2.75 ml of 100% Histodenz) in a 15 ml tube (Greiner). The mixture was then centrifuged for 10 min at room temperature, 1500 g, acceleration three and brake one, resulting in late-stage parasites becoming enriched at the interface<sup>79</sup>. For immunofluorescence assays and transfections, this was repeated over three consecutive cell cycles to create tightly synchronised parasites, with schizont samples from a fourth cycle of Histodenz purification used for transfection. For each transfection, 20–25 µl of pure late schizonts were mixed with 90 µl of P3 solution in a cuvette (Lonza) containing 20 µg of repair template and 10 µg each of two guide vectors (total 40 µg). Transfections were carried out using programme FP158 (Amaya Nucleofector, Lonza), and the contents were immediately transferred into a 1.5 ml sterile Eppendorf tube containing 500 µl of prewarmed complete culture media and 150 µl of uninfected erythrocytes (50% Hct). The transfection mix was incubated at 37 °C while shaking at 800 rpm in a thermomixer for 30 min, before being transferred into a six well plate at 37 °C for one parasite life cycle. After 24 h, the parasites were subjected to 100 nM pyrimethamine (Sigma, #P4200000) selection media. The parasites were kept under selection for 7 days with drug media changed every 2 days. On day 7, the drug media was replaced with complete media. The smears of the parasites

were checked for their recovery and parasitaemia around 1%. The parasite cultures were collected for gDNA extraction using a DNA Blood kit (Qiagen) for subsequent genotyping PCR to confirm editing. To clone out the WT population, limiting dilution and plaque cloning in flat-bottomed 96-well plates were performed. Wells containing single plaques were identified using an EVOS microscope (4× objective, transmitted light), expanded, and DNA isolated and genotyped as described above.

**Growth rate assays for genetically modified lines.** Wild type and genetically modified *P. knowlesi* strains were synchronised as described above. Tightly synchronised ring parasites for both the strains were maintained in six-well plates in normal complete RPMI media. Every alternate day, a 100 µL aliquot was collected, centrifuged and thickly smeared on the slide, before being fixed with methanol, air dried and stained with Giemsa for no longer than 10 min. Cells were counted manually under the microscope with 100× oil immersion objective. A total of ~5000 cells were counted for each slide for each line. The parasitaemia were determined and remaining culture was cut down to 1% parasitaemia and grown again. Data were collected this way for 14 days with 7 time points collected. The assay was done twice with technical replicates and analysed by plotting cumulative parasitaemia over time.

**Reticulocyte invasion assays.** A 20 ml culture of *P. knowlesi* were synchronised as described above and schizonts were purified and washed in complete media. A 0.2 µl aliquot of schizonts from PkWT and *PKNH\_1300500* KO (~90% pure) were added separately to two different tubes containing 4 µl (100% Hct) of enriched reticulocytes purified using CD71<sup>+</sup> coated magnetic beads as described above. The final parasitaemia was ~4.5% and the haematocrit was adjusted to 2% by adding 196 µl complete media (total volume 200 µl). A 50 µl volume of culture was added to each well and a set of three wells were set up for the PkWT and *PKNH\_1300500* KO line. The plate was incubated at 37 °C for 12 h to ensure at least one round of invasion. The next day, the cells were fixed in 100 µl of 2% PFA, 0.008% GA in PBS for 30 min at 4 °C followed by two times wash with PBS at 450 g, 3 min each. The cells were then permeabilised for 10 min at RT with 0.3% Triton X-100 in PBS followed by three washes with PBS. The cells were treated with 10 µg/ml Ribonuclease A in PBS per 100 µl (per well) for 1 h at 37 °C followed by three washes with PBS. Finally, the cells were stained with SyBr green (1:5000) in PBS at 37 °C for 30 min followed by one wash with PBS at 450 g for 3 min and finally resuspend in 100 µl PBS for analysis using flow cytometry (BD Fortessa).

***P. vivax* ex vivo culture and immunofluorescence slide preparation.** A cryopreserved Brazilian *P. vivax* isolate was thawed and the samples were then enriched on 1.080 g/ml KCl high Percoll gradients<sup>80</sup>. Briefly, 2 ml of parasitised red blood cells resuspended using Iscove's modified Dulbecco's medium (IMDM) were layered on 3 ml 1.080 g/mL KCl high Percoll gradient and spun for 15 min at 1200 g with slow acceleration and no break. The interface post Percoll-gradient spin was removed, washed in incomplete IMDM, and subjected to in vitro culture conditions with 1% haematocrit in IMDM (Gibco) containing 10% AB<sup>+</sup> heat-inactivated sera and 50 µg/ml gentamicin at 37 °C in 5% CO<sub>2</sub>, 1% O<sub>2</sub>, and 94% N<sub>2</sub><sup>81</sup>. For immunofluorescence assays, slide cytopspins were made after 42–44 h of maturation, fixed in ice-cold methanol for 10 min, air-dried and stored at -20 °C.

**Immunofluorescence assays.** To generate *P. knowlesi* slides for immunofluorescence, parasites were synchronised as described above and schizonts purified by density gradient centrifugation using a Histodenz (Sigma) cushion. They were then incubated with the Protein Kinase G inhibitor Compound 2 (1 µM) to prevent egress<sup>82</sup>, resulting in the majority of parasites arresting as

extreme late stage mature schizonts. After an hour of this treatment, the media was replaced with normal media for half an hour and finally washed with PBS and fixed for 30 min in 4% paraformaldehyde/0.0075% glutaraldehyde in PBS at room temperature. Fixing buffer was removed by PBS washing, and a schizont smear prepared and fixed on a glass slide. The smear was surrounded by hydrophobic ink from a PAP pen (Daido Sangyo, Japan) to allow low volume staining and washing. For permeabilization, the schizonts were incubated in permeabilization buffer (0.1% Triton X100 in PBS) for 30 min at room temperature. To prevent nonspecific antibody interaction, the cells were blocked overnight at 4 °C using a blocking buffer (3% BSA in PBS or 10% Goat Sera in PBS). For colocalisation studies (both for Pk and Pv), purified rabbit IgGs (described in the IgG purification section) against the selected TRAGs were used (in 1:750 dilution) along with rat PkMSP<sub>1-19</sub> (in 1:2000 dilution)<sup>18</sup>. With these primary antibodies, goat anti rabbit Alexa Fluor 633 (Thermo scientific, #A21070) and goat anti rat Alexa Fluor 488 (Thermo Scientific, #A11006) antibodies were used as secondary antibodies respectively in 1:500 dilutions. Primary antibodies were incubated with the samples sequentially for 1 h each at room temperature. All the secondary antibodies and counterstain Hoechst 33342 (Invitrogen, #H3570) together were incubated with the sample for 45 min at room temperature in the dark. In between each incubation, washing was performed using PBS buffer with 0.05% Tween20 (3 times, 10 min each). All primary and secondary antibodies were prepared in blocking buffer for incubation. After a final wash, the excess washing buffer was removed completely from the smear and the cells were left to mount with a mounting medium (Prolong gold antifade agent, Thermo Fisher) under a coverslip to cure overnight in the dark. The slides were either stored or visualised using a confocal microscope (LSM880 with airyscan setting). The captured images were processed using Zen Blue (2.0) software. The colocalization of PkMSP<sub>1-19</sub> and the TRAGs antibodies were measured by determining the Pearson correlation coefficient (*r*) using additional JACoP in ImageJ.

**Purification of *Plasmodium knowlesi* merozoites and Immunofluorescence assay.** *P. knowlesi* merozoites were purified as described previously<sup>83</sup>. In brief, *Plasmodium knowlesi* AIH1 culture was grown up to a parasitaemia around 9%. The culture was synchronised three times with 55% Histodenz as described above. After the final synchronisation of the day 3 synchronised schizonts (24 h life cycle stage of the parasite on the final day), the top layer of the schizonts was collected and washed with 10 ml RPMI complete media. The schizont pellet was resuspended in RPMI complete media containing Compound 2 (1 µM) for 2 h to block schizont rupture and allow merozoite maturation<sup>84</sup>. The cells were then washed to remove Compound 2 and finally resuspended in the same for 30 min at 37 °C in 3% CO<sub>2</sub>, 1% O<sub>2</sub>, and 96% N<sub>2</sub> to revive the parasites. The cells were first passed through 3 µm (Nucleopore etched membrane, Whatman 110612) and then immediately through 2 µm filters (Nucleopore etched membrane, Whatman 110611). The eluant, containing free merozoites (schizonts can not fit through the filters) was directly collected into the fixing buffer. The fixed cells were centrifuged at 8801g for 5 min, resuspended in PBS and stored at 4 °C. The free *P. knowlesi* merozoites (2 µl) were smeared on to glass slides. One set was permeabilised with permeabilisation buffer (0.1% Triton X100 in PBS) for 30 min at room temperature and the other set was remained non-permeabilised. The cells were washed with PBS to remove traces of Triton X-100; this was used as washing buffer all throughout the experiment. The cells were blocked overnight at 4 °C using a blocking buffer (3% BSA in PBS). The colocalization studies were performed as described for the *P. vivax* schizont staining with same

dilutions of the primary and secondary antibodies. The slides were visualised using a confocal microscope (LSM880 with Airyscan setting). The captured images were processed using Zen Blue (2.0) software.

**Flickering spectrometry of reticulocytes coated with proteins.** 50  $\mu$ l of CD71+ reticulocytes were used as control and incubated at 4 °C for 45 min in PBS + 0.2% BSA with 100  $\mu$ g/ml Pfs25 (negative control), and 100  $\mu$ g/ml PVP01\_0000100, respectively. Samples were washed once and diluted in PBS supplemented with 0.2% BSA at 0.01% Hct and transferred in a Secure Seal Hybridisation Chamber (Sigma-Aldrich, UK) attached to a glass cover slip; the same was kept at 37 °C temperature through a heated collar. Cell membrane thermal fluctuations were recorded at a high frame rate (514 frames per s) and short exposure time (0.8 ms) for 20 s with a 60X Plan Apo VC 1.40 NA oil objective using a Nikon Eclipse Ti-E inverted microscope (Nikon, Japan). Videos were acquired in bright-field with a red filter using a CMOS camera (model GS3-U3-23S6M-C, Point Grey Research/FLIR Integrated Imaging Solutions (Machine Vision, Ri Inc., Canada) (Supplementary Movies 1–3). The equatorial contour for each frame was tracked by an in-house Matlab programme as previously described<sup>44,85</sup>.

The radial component of each contour is then decomposed into its Fourier modes, and the average power spectrum for each cell is:

$$\langle |\tilde{u}(q_x)|^2 \rangle = \frac{k_B T}{L} \frac{1}{2\sigma} \left( \frac{1}{q_x} - \frac{1}{\sqrt{\frac{\sigma}{\kappa} + q_x^2}} \right) \quad (1)$$

$u$  is the amplitude of fluctuations,  $q_x$  is the mode, brackets denote averaging across all contours for a cell,  $k_B$  is the Boltzmann constant,  $T$  is the temperature, and  $L$  is the mean circumference of the cell contour.  $\sigma$  and  $\kappa$  are the tension and bending modulus, respectively, for which the same programme fits the average power spectrum for mode numbers between 5 and 20, as done in ref.<sup>86</sup>. Average mean-square amplitude spectra for each condition are shown in Supplementary Fig. 14. Higher modes are dominated by noise, whereas lower modes are dominated by the shape of the cell. Recent work has reported the effects of out of plane fluctuations for vesicles<sup>87</sup> that has not been implemented yet on red blood cells.

**Cell membrane wrapping energy during invasion.** The wrapping energy was calculated by considering changes to the free energy of the erythrocyte membrane due to a localised deformation and depends on its tension and bending modulus. We consider the deformation to be localised to a disc of radius  $R$  which is deformed by a merozoite to a half-sphere of the same radius. Assuming a very simple hemi-spherical geometry, the sum of these two contributions is given by

$$\Delta F = \pi R^2 \sigma + 4\pi \kappa \quad (2)$$

where  $R = 1 \mu\text{m}$  is the radius of the merozoite<sup>88,89</sup>,  $\sigma$  is the tension, and  $\kappa$  is the bending modulus. The above equation is derived in ref.<sup>44</sup>.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The authors declare that the data supporting the findings of this study are available within Main Text, Supplementary Figures, Tables and videos. Additional data generated during the peer review process are provided in the peer review file. The structure factors and atomic coordinates for PVP01\_0000100 C terminal domain are deposited in

the PDB under the accession code [8ARL](https://doi.org/10.1038/s41467-023-40885-8). Source data are provided with this paper.

### References

- Battle, K. E. et al. Mapping the global endemicity and clinical burden of *Plasmodium vivax*, 2000–17: a spatial and temporal modelling study. *Lancet* **394**, 332–343 (2019).
- Global technical strategy for malaria 2016–2030 (2015).
- Crosnier, C. et al. Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*. *Nature* **480**, 534–537 (2011).
- Douglas, A. D. et al. A PfRH5-based vaccine is efficacious against heterologous strain blood-stage *Plasmodium falciparum* infection in aotus monkeys. *Cell Host Microbe* **17**, 130–139 (2015).
- Minassian, A. M. et al. Reduced blood-stage malaria growth and immune correlates in humans following RH5 vaccination. *Medicine* **2**, 701–719.e719 (2021).
- Singh, K. et al. Malaria vaccine candidate based on duffy-binding protein elicits strain transcending functional antibodies in a phase I trial. *NPJ Vaccines* **3**, 48 (2018).
- Lo, E. et al. Frequent expansion of *Plasmodium vivax* duffy binding protein in ethiopia and its epidemiological significance. *PLoS Negl. Trop. Dis.* **13**, e0007222 (2019).
- Gruszczyk, J. et al. Transferrin receptor 1 is a reticulocyte-specific receptor for *Plasmodium vivax*. *Science* **359**, 48–55 (2018).
- Cowman, A. F. & Crabb, B. S. Invasion of red blood cells by malaria parasites. *Cell* **124**, 755–766 (2006).
- Burns, J. M., Adeeku, E. K., Belk, C. C. & Dunn, P. D. An unusual tryptophan-rich domain characterizes two secreted antigens of *Plasmodium yoelii*-infected erythrocytes. *Mol. Biochem Parasitol.* **110**, 11–21 (2000).
- Oguariri, R. M., Borrmann, S., Klinkert, M. Q., Kremsner, P. G. & Kun, J. F. High prevalence of human antibodies to recombinant duffy binding-like alpha domains of the *Plasmodium falciparum*-infected erythrocyte membrane protein 1 in semi-immune adults compared to that in nonimmune children. *Infect. Immun.* **69**, 7603–7609 (2001).
- Zeeshan, M., Tyagi, R. K., Tyagi, K., Alam, M. S. & Sharma, Y. D. Host-parasite interaction: selective Pv-fam-a family proteins of *Plasmodium vivax* bind to a restricted number of human erythrocyte receptors. *J. Infect. Dis.* **211**, 1111–1120 (2015).
- Alam, M. S. et al. Interaction of *Plasmodium vivax* tryptophan-rich antigen PvTRAg38 with band 3 on human erythrocyte surface facilitates parasite growth. *J. Biol. Chem.* **290**, 20257–20272 (2015).
- Tyagi, K. et al. *Plasmodium vivax* tryptophan rich antigen PvTRAg36.6 interacts with PvETRAMP and PvTRAg56.6 interacts with PvMSP7 during erythrocytic stages of the parasite. *PLoS One* **11**, e0151065 (2016).
- Gunalan, K. et al. Transcriptome profiling of *Plasmodium vivax* in Saimiri monkeys identifies potential ligands for invasion. *Proc. Natl Acad. Sci. USA* **116**, 7053–7061 (2019).
- Wang, B. et al. Immunoprofiling of the tryptophan-rich antigen family in *Plasmodium vivax*. *Infect. Immun.* **83**, 3083–3095 (2015).
- Hostetler, J. B. et al. A library of *Plasmodium vivax* recombinant merozoite proteins reveals new vaccine candidates and protein-protein interactions. *PLoS Negl. Trop. Dis.* **9**, e0004264 (2015).
- Ndegwa, D. N. et al. Using *Plasmodium knowlesi* as a model for screening *Plasmodium vivax* blood-stage malaria vaccine targets reveals new candidates. *PLoS Pathog.* **17**, e1008864 (2021).
- Ntumngia, F. B., Bahamontes-Rosa, N. & Kun, J. F. Genes coding for tryptophan-rich proteins are transcribed throughout the asexual cycle of *Plasmodium falciparum*. *Parasitol. Res.* **96**, 347–353 (2005).

20. Bahl, A. et al. PlasmoDB: the Plasmodium genome resource. A database integrating experimental and computational data. *Nucleic Acids Res.* **31**, 212–215 (2003).
21. Kall, L., Krogh, A. & Sonnhammer, E. L. Advantages of combined transmembrane topology and signal peptide prediction: the phobius web server. *Nucleic Acids Res.* **35**, W429–W432 (2007).
22. Hiller, K., Grote, A., Scheer, M., Munch, R. & Jahn, D. PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res.* **32**, W375–W379 (2004).
23. Crosnier, C. et al. A library of functional recombinant cell-surface and secreted *P. falciparum* merozoite proteins. *Mol. Cell Proteom.* **12**, 3976–3986 (2013).
24. Jones, M. L., Kitson, E. L. & Rayner, J. C. Plasmodium falciparum erythrocyte invasion: a conserved myosin associated complex. *Mol. Biochem. Parasitol.* **147**, 74–84 (2006).
25. Bergman, L. W. et al. Myosin A tail domain interacting protein (MTIP) localizes to the inner membrane complex of Plasmodium sporozoites. *J. Cell Sci.* **116**, 39–49 (2003).
26. Ntumngia, F. B. et al. Characterisation of a tryptophan-rich Plasmodium falciparum antigen associated with merozoites. *Mol. Biochem. Parasitol.* **137**, 349–353 (2004).
27. Barr, P. J. et al. Recombinant Pfs25 protein of Plasmodium falciparum elicits malaria transmission-blocking immunity in experimental animals. *J. Exp. Med.* **174**, 1203–1208 (1991).
28. Kaslow, D. C., Syin, C., McCutchan, T. F. & Miller, L. H. Comparison of the primary structure of the 25 kDa ookinete surface antigens of Plasmodium falciparum and Plasmodium gallinaceum reveal six conserved regions. *Mol. Biochem. Parasitol.* **33**, 283–287 (1989).
29. Vermeulen, A. N. et al. Sequential expression of antigens on sexual stages of Plasmodium falciparum accessible to transmission-blocking antibodies in the mosquito. *J. Exp. Med.* **162**, 1460–1476 (1985).
30. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
31. Holm, L. Using dali for protein structure comparison. *Methods Mol. Biol.* **2112**, 29–42 (2020).
32. Peter, B. J. et al. BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. *Science* **303**, 495–499 (2004).
33. Gallop, J. L. et al. Mechanism of endophilin N-BAR domain-mediated membrane curvature. *EMBO J.* **25**, 2898–2910 (2006).
34. Masuda, M. et al. Endophilin BAR domain drives membrane curvature by two newly identified structure-based mechanisms. *EMBO J.* **25**, 2889–2897 (2006).
35. Zimmerberg, J. & McLaughlin, S. Membrane curvature: how BAR domains bend bilayers. *Curr. Biol.* **14**, R250–R252 (2004).
36. Roberts, D. D. et al. Laminin binds specifically to sulfated glycolipids. *Proc. Natl Acad. Sci. USA* **82**, 1306–1310 (1985).
37. Zhou, Z., Thiagarajan, P., Udden, M., Lopez, J. A. & Guchhait, P. Erythrocyte membrane sulfatide plays a crucial role in the adhesion of sickle erythrocytes to endothelium. *Thromb. Haemost.* **105**, 1046–1052 (2011).
38. McKie, S. J. et al. Altered plasma membrane abundance of the sulfatide-binding protein NF155 links glycosphingolipid imbalances to demyelination. *Proc. Natl Acad. Sci. USA* **120**, e2218823120 (2023).
39. Moon, R. W. et al. Adaptation of the genetically tractable malaria pathogen Plasmodium knowlesi to continuous culture in human erythrocytes. *Proc. Natl Acad. Sci. USA* **110**, 531–536 (2013).
40. Malleret, B. et al. Significant biochemical, biophysical and metabolic diversity in circulating human cord blood reticulocytes. *PLoS One* **8**, e76062 (2013).
41. Amir, A. et al. Invasion characteristics of a Plasmodium knowlesi line newly isolated from a human. *Sci. Rep.* **6**, 24623 (2016).
42. Yoon, Y. Z. et al. Flickering analysis of erythrocyte mechanical properties: dependence on oxygenation level, cell shape, and hydration level. *Biophys. J.* **97**, 1606–1615 (2009).
43. Kariuki, S. N. et al. Red blood cell tension protects against severe malaria in the Dantu blood group. *Nature* **585**, 579–583 (2020).
44. Introini, V. et al. The erythrocyte membrane properties of beta thalassaemia heterozygotes and their consequences for Plasmodium falciparum invasion. *Sci. Rep.* **12**, 8934 (2022).
45. Gardner, M. J. et al. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419**, 498–511 (2002).
46. Auburn, S. et al. A new Plasmodium vivax reference sequence with improved assembly of the subtelomeres reveals an abundance of pir genes. *Wellcome Open Res.* **1**, 4 (2016).
47. Frost, A. et al. Structural basis of membrane invagination by F-BAR domains. *Cell* **132**, 807–817 (2008).
48. Dawson, J. C., Legg, J. A. & Machesky, L. M. Bar domain proteins: a role in tubulation, scission and actin assembly in clathrin-mediated endocytosis. *Trends Cell Biol.* **16**, 493–498 (2006).
49. Lyon, A. M., Dutta, S., Boguth, C. A., Skiniotis, G. & Tesmer, J. J. Full-length Galpha(q)-phospholipase C-beta3 structure reveals interfaces of the C-terminal coiled-coil domain. *Nat. Struct. Mol. Biol.* **20**, 355–362 (2013).
50. Salzer, U., Kostan, J. & Djinic-Carugo, K. Deciphering the BAR code of membrane modulators. *Cell Mol. Life Sci.* **74**, 2413–2438 (2017).
51. Zhao, H. et al. Membrane-sculpting BAR domains generate stable lipid microdomains. *Cell Rep.* **4**, 1213–1223 (2013).
52. Takahashi, T. & Suzuki, T. Role of sulfatide in normal and pathological cells and tissues. *J. Lipid Res.* **53**, 1437–1450 (2012).
53. Kyogashima, M. The role of sulfatide in thrombogenesis and haemostasis. *Arch. Biochem. Biophys.* **426**, 157–162 (2004).
54. Cerami, C., Kwakye-Berko, F. & Nussenzweig, V. Binding of malarial circumsporozoite protein to sulfatides [Gal(3-SO4)beta 1-Cer] and cholesterol-3-sulfate and its dependence on disulfide bond formation between cysteines in region II. *Mol. Biochem. Parasitol.* **54**, 1–12 (1992).
55. Pancake, S. J., Holt, G. D., Mellouk, S. & Hoffman, S. L. Malaria sporozoites and circumsporozoite proteins bind specifically to sulfated glycoconjugates. *J. Cell Biol.* **117**, 1351–1357 (1992).
56. Namvar, A. et al. Surface area-to-volume ratio, not cellular viscoelasticity, is the major determinant of red blood cell traversal through small channels. *Cell Microbiol.* **23**, e13270 (2021).
57. Burns, J. M. Jr., Adeeku, E. K. & Dunn, P. D. Protective immunization with a novel membrane protein of Plasmodium yoelii-infected erythrocytes. *Infect. Immun.* **67**, 675–680 (1999).
58. Gabelich, J. A. et al. A member of the tryptophan-rich protein family is required for efficient sequestration of Plasmodium berghei schizonts. *PLoS Pathog* **18**, e1010846 (2022).
59. Bahl, A. et al. PlasmoDB: the Plasmodium genome resource. An integrated database providing tools for accessing, analyzing and mapping expression and sequence data (both finished and unfinished). *Nucleic Acids Res.* **30**, 87–90 (2002).
60. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* **7**, 539 (2011).
61. Dereeper, A. et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465–W469 (2008).
62. Felsenstein, J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376 (1981).
63. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791 (1985).
64. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).

65. Sonnhammer, E. L., von Heijne, G. & Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **6**, 175–182 (1998).
66. Sun, Y., Gallagher-Jones, M., Barker, C. & Wright, G. J. A bench-marked protein microarray-based platform for the identification of novel low-affinity extracellular protein interactions. *Anal. Biochem.* **424**, 45–53 (2012).
67. Crosnier, C., Staudt, N. & Wright, G. J. A rapid and scalable method for selecting recombinant mouse monoclonal antibodies. *BMC Biol.* **8**, 76 (2010).
68. Sorette, M. P., Shiffer, K. & Clark, M. R. Improved isolation of normal human reticulocytes via exploitation of chloride-dependent potassium transport. *Blood* **80**, 249–254 (1992).
69. Hoie, M. H. et al. NetSurfP-3.0: accurate and fast prediction of protein structural features by protein language models and deep learning. *Nucleic Acids Res.* **50**, W510–W515 (2022).
70. Kabsch, W. Xds. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).
71. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
72. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
73. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171–1179 (2008).
74. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
75. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D. Struct. Biol.* **74**, 519–530 (2018).
76. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 352–367 (2012).
77. Jurrus, E. et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112–128 (2018).
78. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
79. Moon, R. W. et al. Normocyte-binding protein required for human erythrocyte invasion by the zoonotic malaria parasite *Plasmodium knowlesi*. *Proc. Natl Acad. Sci. USA* **113**, 7231–7236 (2016).
80. de Oliveira, T. C. et al. Genome-wide diversity and differentiation in new world populations of the human malaria parasite *Plasmodium vivax*. *PLoS Negl. Trop. Dis.* **11**, e0005824 (2017).
81. Rangel, G. W. et al. Enhanced ex vivo *Plasmodium vivax* intraerythrocytic enrichment and maturation for rapid and sensitive parasite growth assays. *Antimicrob. Agents Chemother.* **62**, e02519–17 (2018).
82. Collins, C. R. et al. Malaria parasite cGMP-dependent protein kinase regulates blood stage merozoite secretory organelle discharge and egress. *PLoS Pathog.* **9**, e1003344 (2013).
83. Lyth, O. et al. Cellular dissection of malaria parasite invasion of human erythrocytes using viable *Plasmodium knowlesi* merozoites. *Sci. Rep.* **8**, 10165 (2018).
84. Mohring, F. et al. Rapid and iterative genome editing in the malaria parasite *Plasmodium knowlesi* provides new tools for *P. vivax* research. *Elife* **8**, e45829 (2019).
85. Introini, V., Govendir, M. A., Rayner, J. C., Cicuta, P. & Bernabeu, M. Biophysical tools and concepts enable understanding of asexual blood stage malaria. *Front. Cell. Infect. Microbiol.* **12**, 908241 (2022).
86. Koch, M. et al. *Plasmodium falciparum* erythrocyte-binding antigen 175 triggers a biophysical change in the red blood cell that facilitates invasion. *Proc. Natl Acad. Sci. USA* **114**, 4225–4230 (2017).
87. Faizi, H. A., Reeves, C. J., Georgiev, V. N., Vlahovska, P. M. & Dimova, R. Fluctuation spectroscopy of giant unilamellar vesicles using confocal and phase contrast microscopy. *Soft Matter* **16**, 8996–9001 (2020).
88. Dasgupta, S. et al. Membrane-wrapping contributions to malaria parasite invasion of the human erythrocyte. *Biophys. J.* **107**, 43–54 (2014).
89. Dasanna, A. K., Hillringhaus, S., Gompper, G. & Fedosov, D. A. Effect of malaria parasite shape on its alignment at erythrocyte membrane. *Elife* **10**, e68818 (2021).

## Acknowledgements

We would like to thank Prof. Gavin J. Wright for his help during the construct designing for expression vectors and for providing the HEK 293E cell line for protein expression. Dr. Robert W. Moon for providing the CRISPR Cas9 plasmid and *Plasmodium knowlesi* A1.H1 line. Dr. Reiner Schulte, Gabriela Grondys-Kotarba and Chiara Cossetti of the CIMR Flow Cytometry facility for providing required training and assistance during the flow cytometry experiments. We also like to convey our thanks to Matthew Gratian and Mark Bowen for providing training and data acquisition in LSM880 confocal microscope with Airyscan setting. We are grateful to Alison Kemp for her guidance to use the CRISPR Cas9 system. Ellen Knuepfer for the kind donation of Rat anti-PkMSP1-19 serum. We thank Stephen Graham for help with crystallographic model building. We acknowledge Diamond Light Source for time on beamline IO4 under proposal MX21426. This work was funded by the National Institutes of Health (R01AI137154) and the Wellcome Trust (220266/Z/20/Z). J.E.D. and S.J.M. are supported by a Wellcome Trust Senior Research Fellowship (219447/Z/19/Z) awarded to J.E.D. As this research was funded in part by the Wellcome Trust, for the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

## Author contributions

P.K., J.C.R., J.E.D. and M.D. conceived and designed the research; P.K., D.N., S.M., V.I., performed the experiments, P.K., D.N., S.M., V.I., P.C., M.D., J.E.D. and J.C.R. analysed the data; P.K. and J.E.D. performed structural refinement and analysis, M.U.F., M.D., S.D. and U.K. provided the Pv isolates for the IFA studies. P.K. J.C.R., D.N., J.E.D., S.M. wrote the first draft of the manuscript and all authors provided input.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-40885-8>.

**Correspondence** and requests for materials should be addressed to Manoj Duraisingh, Janet E. Deane or Julian C. Rayner.

**Peer review information** *Nature Communications* thanks Hammad Faizi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023