Pfcerli2, a duplicated gene in the malaria parasite *Plasmodium falciparum* essential for invasion of erythrocytes as revealed by phylogenetic and cell biological analysis.

Benjamin Liffner[1¶], Juan Miguel Balbin[1¶], Gerald J. Shami[2], Jan Strauss[3,4,5‡], Sonja Frölich[1], Gary K. Heinemann[6], Arne Alder[3,4], Jan Stephan Wichers[3,4], Leann Tilley[3], Matthew W. A. Dixon[3], Tim-Wolf Gilberger[3,4,5] Danny W. Wilson[1,7*]

[1]Research Centre for Infectious Diseases, School of Biological Sciences, University of Adelaide, Adelaide 5005, Australia.

[2]Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Melbourne 3010, Victoria, Australia.

[3]Centre for Structural Systems Biology, 22607 Hamburg, Germany.

[4]Bernhard Nocht Institute for Tropical Medicine, 20359 Hamburg, Germany.

[5]Biology Department, University of Hamburg, 20146, Hamburg, Germany.

[6]Experimental Therapeutics Laboratory, School of Pharmacy & Medical Sciences, University of South Australia Cancer Research Institute, Adelaide 5005, Australia.

[7]Burnet Institute, 85 Commercial Road, Melbourne 3004, Victoria, Australia.

* Corresponding author

E-mail: Danny.wilson@adelaide.edu.au

[¶]These authors contributed equally to this work.

[‡]Current address: GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany.

ABSTRACT

Merozoite invasion of host red blood cells (RBCs) is essential for survival of the human malaria parasite *Plasmodium falciparum*. Proteins involved with RBC binding and invasion are secreted from dual-club shaped organelles at the apical tip of the merozoite called the rhoptries. Here we characterise *P. falciparum* Cytosolically Exposed Rhoptry Leaflet Interacting protein 2 (PfCERLI2), as a rhoptry bulb protein that is essential for merozoite invasion. Phylogenetic analyses show that *cerli2* arose through an ancestral gene duplication of *cerli1*, a related cytosolically exposed rhoptry bulb protein. We show that PfCERLI2 is essential for blood-stage growth and localises to the cytosolic face of the rhoptry bulb. Inducible knockdown of PfCERLI2 led to an inhibition of merozoite invasion after tight junction formation. PfCERLI2 knockdown was associated with inhibition of rhoptry antigen processing and a significant elongation of the rhoptries, suggesting that the inability of merozoites to invade is caused by aberrant rhoptry function due to PfCERLI2 deficiency. These findings identify PfCERLI2 as a protein that has key roles in rhoptry biology during merozoite invasion.

INTRODUCTION

Plasmodium falciparum, a human malaria parasite, is the cause of ~400,000 deaths each year; predominantly in children under the age of five[1]. *P. falciparum* is transmitted from *Anopheles* mosquitoes, its definitive host, to humans through the blood meal of an infected mosquito where invasive sporozoites are injected subcutaneously and migrate to the liver. Following invasion of liver hepatocytes, the parasite develops into thousands of daughter merozoites inside a hepatic schizont, which ruptures and releases the daughter merozoites into the blood stream where they invade red blood cells (RBCs)[2]. Over the following ~48 hours, the parasite develops inside the RBC until it forms 16-32 new daughter merozoites inside of a schizont that ruptures, releasing the merozoites that go on to infect new RBCs[2]. Replication of *P. falciparum* in this asexual blood-stage of the lifecycle is the cause of all the clinical symptoms of malaria.

RBC invasion by merozoites occurs over a period of ~30 seconds and involves the attachment of the merozoite to the RBC, which then reorients so its apical tip contacts the RBC membrane[3]. Specialised invasion organelles, known as the rhoptries and micronemes, then secrete their contents to form an irreversible tight junction before the merozoite pulls the RBC plasma membrane around itself; forming a parasitophorous vacuole (PV) and completing invasion[4]. Rhoptries, the largest of the invasion organelles, are dual club-shaped and divided into a bulb and neck, with the neck positioned at the apical tip of the merozoite. Rhoptry neck proteins are released early in the invasion process and mediate initial attachment to the RBC and formation of the tight junction. Rhoptry bulb contents are secreted following junction formation and are typically involved in establishing the PV.

Most rhoptry luminal proteins exclusively localise to the rhoptry bulb or neck, with minimal overlap between the two[5]. During the process of RBC invasion, the neck of the rhoptries fuse to the parasite plasma membrane (PPM) to allow secretion of rhoptry contents.

Additionally, the dual club-shaped rhoptries fuse from neck to bulb, leaving a singular rhoptry that retains its neck and bulb structure[6]. What controls both PPM and rhoptry fusion has not been elucidated.

It is likely that rhoptry and PPM fusion are controlled by proteins that localise to the cytosolic face of the rhoptry membrane, as they can contact both membranes to facilitate their fusion. To date, however, only three proteins in *P. falciparum* have been shown to localise to the cytosolic face of the rhoptry membrane: Armadillo repeats only (ARO)[7], ARO interacting protein (AIP)[8], and Cytosolically exposed rhoptry leaflet interacting protein 1 (CERLI1; also known as Rhoptry apical surface protein (RASP) 2)[9,10]. Here we describe PfCERLI2 (Pf3D7_0405200), a paralogue of *Pfcerli1*, which localises to the cytosolic face of the rhoptry bulb membrane and is essential for merozoite invasion of RBCs.

RESULTS

PfCERLI2 is conserved among Apicomplexa and arose from an ancestral gene duplication

Plasmodium falciparum Cytosolically Exposed Rhoptry Leaflet Interacting protein 2 (PfCERLI2) (Pf3D7_0405200) is a protein of 579 amino acids in *P. falciparum* isolate 3D7. PfCERLI2 contains a predicted lipid-calcium binding C2 domain towards its N-terminus and possesses a decapeptide tandem repeat (consensus sequence QTEIkNDhi; upper case = fixed amino acid, lower case = variable amino acid) at its C-terminus, with the number of repeats ranging from 10 to 20 between different *P. falciparum* isolates (Figure 1a; Supplementary Table 1). PfCERLI2 shares homology (~20% amino acid identity) (Supplementary Figure 1) with the recently characterised protein PfCERLI1[9,10], with PfCERLI2 being most similar protein to PfCERLI1 in the *P. falciparum* 3D7 proteome on PlasmoDB[11]. Notably, PfCERLI1 and PfCERLI2 both possess a C2 domain towards their N-terminus as well as a short but highly conserved motif we have termed PHIS, which has the consensus sequence PHIS[-]xxP ([-] = negatively charged) in *P. falciparum*. PfCERLI2 is highly conserved amongst *Laverania* (>90% amino acid identity with *Laverania* homologues) and shares approximately 65% amino acid identity amongst more distantly related *Plasmodium* spp. (Figure 1b) (Supplementary Figure 2). Notably, PfCERLI2 shares high similarity with its homologues over the first ~420 amino acids of its sequence, but this similarity decreases significantly at the *Laverania*-specific C-terminal repeat region.

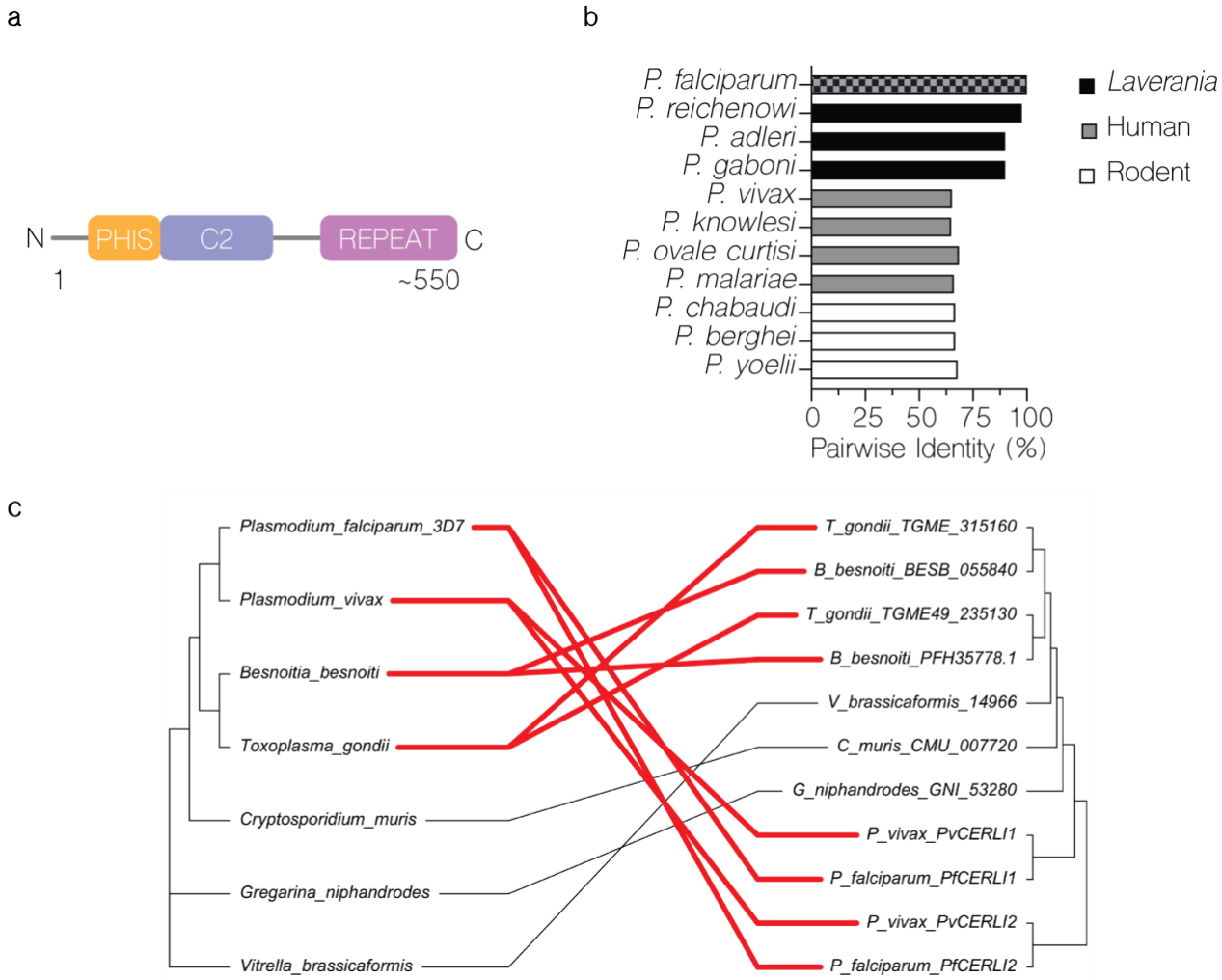


Figure 1: PfCERLI2 is conserved among Apicomplexa and may have evolved from an ancestral gene duplication.

(a) *PfCERLI2 is a protein of 579 amino acids in P. falciparum 3D7. Towards its N-terminus PfCERLI2 contains a motif with the consensus sequence PHIS[-]xxP we have termed PHIS, a C2 domain, and a decapeptide tandem repeat with the consensus sequence QTEIkNDhi at its C-terminus. Repeat number (10-20), and therefore PfCERLI2 amino acid length (559-659), is highly variable between P. falciparum isolates.* ***(b)*** *Amino acid sequence identity for Plasmodium spp. PfCERLI2 orthologues in Laverania, human-infecting, and rodent-infecting parasites was compared using multiple pairwise alignments.* ***(c)*** *Tanglegram comparing general evolutionary relationships between selected Apicomplexa and Chromerids, as described in[12](left), with phylogenetic tree constructed with PfCERLI1, PfCERLI2 and homologus sequences retrieved form EuPathDB using the unweighted pair group method with arithmetic mean (UPGMA) method (right). Branch length of UPGAME tree corresponds to amino acid substitutions per site. Taxa containing CERLI1 and CERLI2 homologues are joined by red edges, while taxa with a single CERLI are joined by black edges to visualise timing of ancestral gene duplication giving rise to CERLI2.*

All organisms that diverged from *Plasmodium* more recently than *Cryptosporidium* contain homologues of both *Pfcerli1* and *Pfcerli2*, while earlier diverging organisms, such as *Gregarina niphandrodes* and the Chomerid *Vitrella brassicaformis*, contain only a single identifiable homologue (Figure 1c) (Supplementary Figure 3); suggesting that *Pfcerli1* and *Pfcerli2* are paralogues that arose from an ancestral gene duplication event. As genomic DNA (14.6 %) and amino acid (22.9 %) sequence identities werelow across Apicomplexa and Chromera, analysis of sequence conservation, outside of *Plasmodium*, was not able to identify the direct orthologues of either *Pfcerli1* or *Pfcerli2* outside the *Plasmodium* genus (Supplementary Figure 4a). Therefore, we compared the gene structure of *Pfcerli1* and *Pfcerli2*, along with homologous sequences in Apicomplexa and Chromerids (Supplementary Table 2) to identify the putative *cerli1* and *cerli2* orthologues across distantly related species. The comparison of *cerli1* and *cerli2* sequences across *Plasmodium spp.* indicated that *cerli1* orthologues have fewer introns and decreased exon length than the orthologues of *cerli2* (Supplementary Table 3). This gene structure pattern was also observed for homologues outside the *Plasmodium* genus, allowing for a classification of more divergent homologues as the direct orthologue of either *Pfcerli1* or *Pfcerli2*. Using this approach, we determined that the direct orthologue of *Pfcerli2* in *Toxoplasma gondii* is TGME49_315160 (TgRASP2), which was previously reported to be the direct orthologue of *Pfcerli1*[10]. Additionally, our classification suggests that the direct orthologue of *Pfcerli1* is TGME49_235130 (TgRASP1). As Apicomplexa and Chromerids that are distantly related to *P. falciparum* only possess a single homologue of *Pfcerli1* and *Pfcerli2*, we compared their gene structure to determine which gene is more similar to the conserved ancestral state and which one represents a more diverged duplicated copy. All *Cryptosporidium* spp. homologues have an intron number and exon length similar to the *cerli1* lineage, suggesting that *cerli1* represents the ancestral locus and *cerli2* the duplicated locus. There is no consistent trend in gene structure in more distantly

related organisms (Gregarines or Chromerids), suggesting the gene structure may not have been conserved in more divergent organisms.

With our analysis supporting that *Pfcerli2* arose from a gene duplication event, we next assessed the selection pressure acting on the *Pfcerli1* and *Pfcerli2* sequences. To investigate this, we determined Ka/Ks ratios, which calculate the number of non-synonymous amino acid substitutions (Ka) relative to the number of synonymous substitutions (Ks). *Pfcerli1* and *Pfcerli2* across *Plasmodium* spp. had Ka/Ks of 0.091 and 0.29 respectively, indicating strong negative selection (favouring low sequence diversity), which is consistent with their essentiality for *P. falciparum*[9,10,13] and *P. berghei*[14] blood-stage survival. This negative selection in both lineages may be indicative of subfunctionalisation in Apicomplexa, where both paralogues have maintained a subset of the function of the ancestral gene.

Interestingly, homologues of PfCERLI2 within the *Plasmodium* genus no longer contain a Pleckstrin homology (PH) domain that is predicted for homologous sequences outside the *Plasmodium* genus (Supplementary Figure 5) and has been shown to be involved in rhoptry localisation and secretion for TgRASP2[10]. Analysis of predicted protein structures for PfCERLI1 and PfCERLI2 homologues outside the *Plasmodium* genus show that this PH domain is typically predicted to contains two sets of anti-parallel beta-sheets separated by a short alpha-helix, followed by a C-terminal helix (Supplementary Figure 6). By contrast, *Plasmodium* homologues of CERLI2 are predicted to possess an expanded helix between the beta-sheets that disrupts their interaction (Supplementary Figure 6). Additionally, PfCERLI2 is predicted to contain a shorter C-terminal helix and a disordered region between the beta-sheets (Supplementary Figure 6). Disruption of the interaction between the two sets of anti-parallel beta-sheets suggests that CERLI2 homologues in *Plasmodium* have a degenerated and therefore non-functional PH domain.

While the overall sequence identity between PfCERLI1, PfCERLI2, and their homologues is low, all contain the highly conserved PHIS motif with the consensus sequence PHPSECxP (when comparing all apicomplexan/chromerid homologues; Supplementary Figure 4b) or PHIS[-]xxP (when comparing *P. falciparum* sequences; Supplementary Figure 7). Notably, the first proline is fully conserved across all analysed sequences, and the histidine residue is mutated only in *Gregarina niphandrodes*. A second proline lies at position 8 in all CERLI homologues in *Plasmodium* spp., while in all other apicomplexans outside *Plasmodium*, the second proline lies at position 3. Such high sequence conservation across highly divergent organisms suggests this motif is likely to be critical for the physiological function of PfCERLI1, PfCERLI2 and their homologues. While the function of this motif is unknown, similar motifs exist in other *P. falciparum* proteins; including the ring-expressed surface antigens (RESA, RESA2, RESA3, and RESA-like), Iron superoxide dismutase (FeSoD) and ATP-dependent Clp protease proteolytic subunit (ClpR) (Supplementary Table 4; Supplementary Figure 7). Although the PHIS motif is a signature of PfCERLI1 and PfCERLI2, and their apicomplexan homologues, nothing is currently known about the function of this motif or whether it has a related role in the other *P. falciparum* PHIS-containing proteins.

PfCERLI2 contains a variable copy number decapeptide tandem repeat

All PfCERLI2 homologues in *Laverania* contain a C-terminal decapeptide tandem repeat, with the consensus sequence QTIEIkNDhi in *P. falciparum*, which is not present in any homologues outside the subgenus (Supplementary Table 1) [15]. The repeat number is variable both between *P. falciparum* isolates (e.g. 10 repeats in isolate 7G8 and 20 repeats in isolate Dd2), and between *Laverania* (9 repeats in *P. praefalciparum* and 4 repeats in *P. adleri*). It is not clear what the function of this repeat is, or what drives changes in its number. However,

we noticed that all sequenced *P. falciparum* isolates have a higher number of repeats than other *Laverania*, with organisms more closely related to *P. falciparum* also displaying higher repeat numbers than those more distantly related (Supplementary Figure 8; Supplementary Table 1). When comparing the amino acid sequences of PfCERLI2 and its Laveranian homologues without the repeat region, their identity parallels the evolutionary relatedness of different *Laverania* (Supplementary Figure 8). The full sequence of *Pfcerli2* is under strong negative selection across *Laverania*, with a Ka/Ks ratio of 0.3. By contrast, the repeat region of *Pfcerli2* when comparing across *Laverania* has a Ka/Ks ratio of 1.43 indicating that positive selection (favouring sequence diversity) is occurring in the decapeptide repeat region of PfCERLI2 and Laveranian homologues.

PfCERLI2 is essential for blood-stage growth.

PfCERLI2 has previously been suggested to be essential by a saturation mutagenesis screen[13], and its homologue in *P. berghei* has also been shown to be essential for blood-stage growth[14]. To confirm this in *P. falciparum*, we targeted *Pfcerli2* by selection linked integration targeted gene disruption (SLI-TGD)[16]. Multiple attempts to knockout *Pfcerli2* failed confirming that *Pfcerli2* has an important role in blood-stage growth (Figure 2a). Next, we placed *Pfcerli2* under control of the glucosamine (GLCN) inducible *GlmS* ribozyme, whereby addition of GLCN leads to specific degradation of *Pfcerli2* mRNA (Figure 2 b&c). Additionally, we introduced a haemaglutinnin (HA) tag onto the C-terminus of PfCERLI2, allowing detection using anti-HA antibodies. Using these PfCERLI2^{HAGlmS} parasites, we harvested parasites at either ring, trophozoite or schizont stages and probed for PfCERLI2 by Western blot (Figure 2d). Concordant with previously published transcriptomic data[17], PfCERLI2 was most highly expressed in schizonts. When PfCERLI2^{HAGlmS} parasites were treated from ring stages to schizonts with 2.5 mM GLCN, the expression of PfCERLI2 was

reduced by 72% (± 3.8% SEM) relative to untreated controls (Figure 2e), showing that PfCERLI2 protein expression is tunable using the *GlmS* system.

To determine whether knockdown of PfCERLI2 altered parasite growth, we treated ring-stage PfCERLI2^{HAGlmS} and 3D7 WT parasites with increasing concentrations of GLCN and measured trophozoite-stage parasitaemia by flow cytometry 72 hours later (Figure 2f). At 5 mM GLCN there was ~77% (± 5.1% SEM) decrease in the growth of PfCERLI2^{HAGlmS} parasites. However, at this high concentration of GLCN growth of 3D7 WT parasites was also affected. At a concentration of 2.5 mM GLCN, growth of PfCERLI2^{HAGlmS} parasites was inhibited by ~53% (± 6.0% SEM), with 3D7 WT parasites showing negligible off-target growth defects.

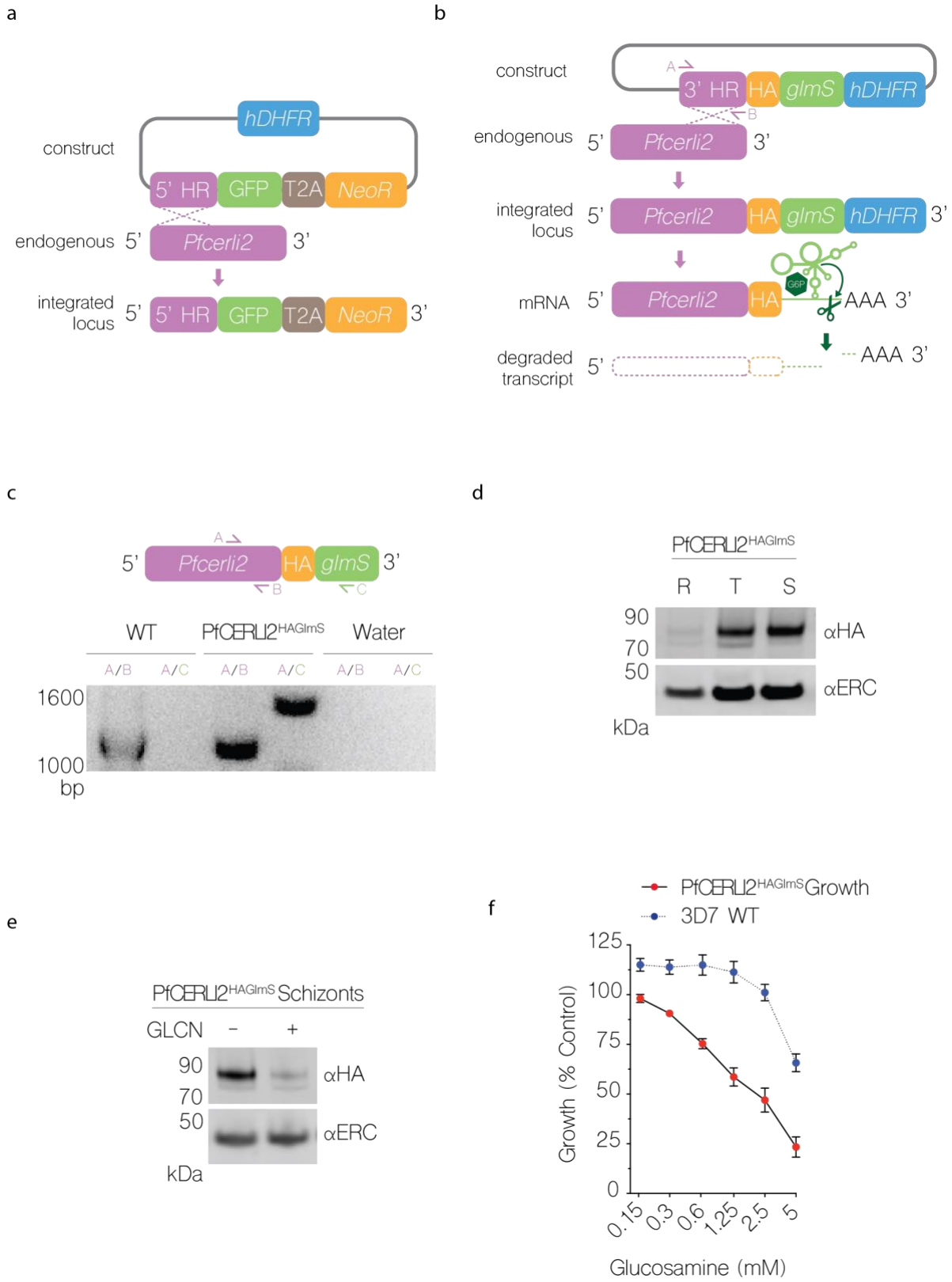


Figure 2: PfCERLI2 is essential for blood-stage growth.

(a) *Schematic representation of the selection linked integration targeted gene disruption (SLI-TGD) system used for the attempted gene knockout of Pfcerli2.* ***(b)*** *Schematic representation of the haemaglutinnin (HA) tag, and GlmS ribozyme system used to identify and knockdown PfCERLI2. A plasmid that contained a 3' Pfcerli2 flank (2154bp-2912bp) with an HA-tag and GlmS ribyzome was transfected into 3D7 parasites by single crossover recombination. Following mRNA production, glucosamine-6-phosphate (G6P) binds to the GlmS ribozyme, leading to the cleavage of Pfcerli2 mRNA and subsequent protein knockdown.* ***(c)*** *To confirm integration of the Pfcerli2^{HAGlmS} plasmid, gDNA was harvested from transfected parasites. A Pfcerli2 specific forward primer, A, with either a Pfcerli2 specific, B (WT), or GlmS specific, C (integrated), reverse primer was used to confirm WT DNA sequence or integration of the Pfcerli2^{HAGlmS} plasmid.* ***(d)*** *Western blot of ring, trophozoite, or schizont stage PfCERLI2^{HAGlmS} lysates probed with either anti-HA (PfCERLI2) or anti-ERC (loading control) antibodies.* ***(e)*** *Western blot of PfCERLI2^{HAGlmS} schizont lysates either GLCN treated (+) or untreated (-) and probed with anti-HA (PfCERLI2) or anti-ERC (loading control) antibodies.* ***(f)*** *PfCERLI2^{HAGlmS} or 3D7 WT parasites were treated with increasing concentrations of GLCN for 72 hours, with the trophozoite-stage parasitaemia determined by flow cytometry. Growth is expressed as a percentage of an untreated media control. n=4, error bars = SEM. X-axis presented as a log 2 scale for viewing purposes.*

PfCERLI2 knockdown inhibits merozoite invasion

As PfCERLI2 was most highly expressed at schizont-stages, and its knockdown inhibited growth, we hypothesised that PfCERLI2 could be involved in schizont development, rupture, or merozoite invasion. To quantify invasion, we transfected PfCERLI2^{HAGlmS} parasites with a plasmid that expresses cytosolic GFP (PfCERLI2$^{HAGlmS/GFP}$), which allows reliable detection of newly invaded ring stages by flow cytometry. Treating PfCERLI2$^{HAGlmS/GFP}$ parasites with increasing concentrations of GLCN from trophozoite-stages for 24 hours until early ring-stages the following cycle, we saw that PfCERLI2 knockdown inhibited invasion in a dose-dependent manner (Figure 3a). The extent of invasion inhibition caused by PfCERLI2 knockdown was lower than growth inhibition over 72 hours, however, it is likely that this difference is caused by the shorter GLCN treatment time used for assessing inhibition of invasion.

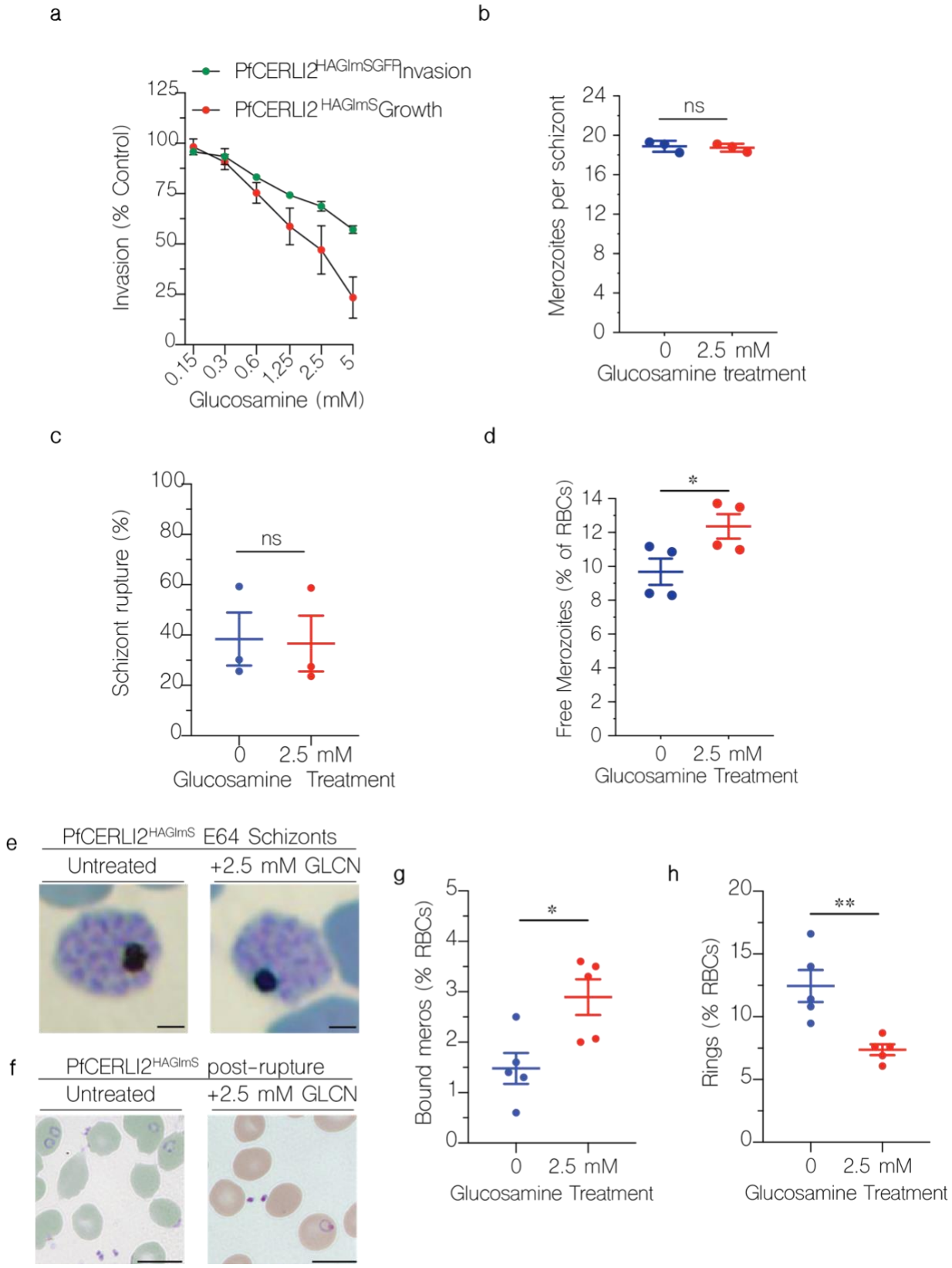


Figure 3: PfCERLI2 knockdown inhibits merozoite invasion.

(a) *GFP-expressing PfCERLI2$^{HAGlmS/GFP}$ parasites were treated with a range of concentrations of GLCN from early trophozoite stages until early rings the following cycle (~24 hours), with invasion assessed by flow cytometry as the number of newly invaded rings as a percentage of an untreated media control. n=4, error bars = SEM. PfCERLI2^{HAGlmS} growth (~72 hours) is replicated from Figure 2f for direct comparison between growth and invasion inhibition. X-axis presented as log 2 scale for viewing purposes.* ***(b)*** *Early schizonts were treated with the egress inhibitor E64 for ~5 hours. Following treatment, cultures were smeared and the number of merozoites per schizont was determined by blinded microscopy analysis of Giemsa-stained blood smears. Each data point represents the mean number of merozoites per schizont from 20 schizonts. n=3, error bars = SEM, ns=p>0.05* ***(c)*** *Percentage of schizonts that ruptured over a 6-hour window in either the presence or absence of GLCN, as determined by flow cytometry. n=3, error bars = SEM, ns=p>0.05* ***(d)*** *During the invasion assay described in 3a, the number of free merozoites was also quantified by flow cytometry, with results presented as % of total RBCs. n=4, error bars = SEM, *=p<0.05 by unpaired t-test.* ***(e)*** *PfCERLI2^{HAGlmS} schizonts, either in the presence or absence of GLCN, were matured in the presence of the schizont rupture inhibitor E64 before being fixed, Giemsa-stained and imaged by light microscopy. Scale bar = 2 µm. PfCERLI2^{HAGlmS} parasites were treated with 2.5 mM GLCN or left untreated, with giemsa smears made following schizont rupture* ***(f)*** *and the number of merozoites bound to RBCs* ***(g)*** *and newly-invaded rings* ***(h)*** *quantified. Scale bar = 5 µm, n=5, error bars =SEM, *=p<0.05, **=p<0.01 by unpaired t-test.*

To determine if this invasion inhibition was specifically due to a defect in merozoite development, PfCERLI2^{HAGlmS} schizonts either GLCN treated or untreated, were matured in the presence of the schizont egress inhibitor E64. Knockdown of PfCERLI2 did not result in a change in the number of fully formed merozoites per schizont (Figure 3b) and the merozoites within the schizonts appeared morphologically normal (Figure 3e). To determine whether PfCERLI2 knockdown altered the rate of schizont rupture, the ability of both GLCN treated and untreated PfCERLI2^{HAGlmS} synchronised schizonts to rupture over a period of 6 hours was quantified. PfCERLI2 knockdown had no influence on the rate of schizont rupture (Figure 3c), suggesting knockdown has no influence on schizont development or egress. To test if PfCERLI2 knockdown directly inhibited the invasion of merozoites, we measured the number of free merozoites in the culture supernatant following invasion in both GLCN treated and untreated PfCERLI2$^{HAGlmS/GFP}$ parasites. PfCERLI2 knockdown lead to an increase in the

number of free merozoites, suggesting that a direct defect in the ability of merozoites to invade is the cause of PfCERLI2 knockdown-mediated growth inhibition.

The flow cytometry-based assay does not allow discrimination between free merozoites inhibited prior to or after tight-junction formation. Therefore, we analysed Giemsa-stained thin smears of GLCN treated and untreated PfCERLI2^{HAGlmS} ring-stage parasites and quantified the number of bound merozoites following schizont rupture (Figure 3 f&g). PfCERLI2 knockdown led to a 49% (± 12% SEM) increase in the percentage of merozoites bound to RBCs, suggesting that knockdown inhibits invasion following tight-junction formation. Newly invaded ring-stage parasites (Figure 3h) were also quantified, with GLCN treatment leading to a 41% (± 5.9% SEM) reduction in newly invaded ring stages. These data indicate that the GLCN inducible knockdown of PfCERLI2 inhibits merozoite invasion both at and prior to formation of the tight-junction.

PfCERLI2 localises to the rhoptry bulb

The homology to PfCERLI1 as well as its functional implications in RBC invasion suggested PfCERLI2 may have a rhoptry localisation. After failing to get consistent PfCERLI2 localisation using conventional antibody staining, likely due to the low peak expression level of *Pfcerli2*[17], we investigated the cellular distribution of PfCERLI2 using a high affinity anti-HA biotin system[18]. Despite its peak transcription late in schizont development, we observed that PfCERLI2 foci were brightest in early schizonts (~44 hrs post invasion) (Figure 4a), but largely absent in fully mature E64-treated schizonts (~48 hrs post invasion) (Supplementary Figure 10). Harvesting early schizonts (~44hr post invasion), we analysed subcellular organisation of CERLI2 by comparing it to known markers of the rhoptry bulb (RAP1), rhoptry neck (RON4), micronemes (AMA1), and merozoite surface (MSP1-19) (Figure 4 a&b). In these parasites, PfCERLI2 showed the strongest colocalisation with the rhoptry bulb marker

RAP1, suggesting PfCERLI2 likely localises to the rhoptry bulb (Figure 4a&b). For a more detailed view of PfCERLI2 and its association with the rhoptry, we imaged double labelled parasites in three-dimensions using the super-resolution microscopy platform Airyscan, which showed that PfCERLI2 forms the donut shape characteristic of membrane associated rhoptry bulb proteins (Figure 4c).

Repeat immunofluorescence experiments consistently showed that mature schizonts had little PfCERLI2 staining. Parallel western blots, however, showed no significant decrease in PfCERLI2 detection between early and late schizonts (Supplementary Figure 11). Given the lack of change in detectability by Western blot, it is possible that the HA epitope is masked and therefore undetectable using anti-HA antibodies in microscopy with mature schizont stage parasites.

PfCERLI2 is peripherally associated with the cytosolic face of the rhoptry bulb membrane.

The absence of a signal peptide suggested that PfCERLI2 is a peripheral membrane protein that localises to the rhoptries but is exposed to the merozoite cytosol. To test this, we subjected PfCERLI2^{HAGlmS} schizont lysates to a proteinase protection assay. Following lysis with saponin and digitonin, which lyse the RBC membrane, PVM, and PPM, but not organellar membranes, PfCERLI2 was sensitive to proteinase K degradation (Figure 5a); suggesting it is exposed to the cytosol. To assess membrane association of PfCERLI2, we performed a solubility assay using lysates from PfCERLI2^{HAGlmS} schizonts (Figure 5b), whereby proteins are solubilised into cytosolic (hypotonic lysis), peripheral membrane (carbonate treatment), integral membrane (triton-x-100 soluble), or covalently lipid-linked (triton-x-100 insoluble) fractions. PfCERLI2 predominantly solubilised in the carbonate fraction, a distinctive feature of peripheral membrane proteins[19]. Combining the localisation analyses with the proteinase

protection and solubility assays suggest that PfCERLI2 localises to the cytosolic face of the rhoptry bulb membrane, a subcellular localisation akin to that reported for PfCERLI1[9].

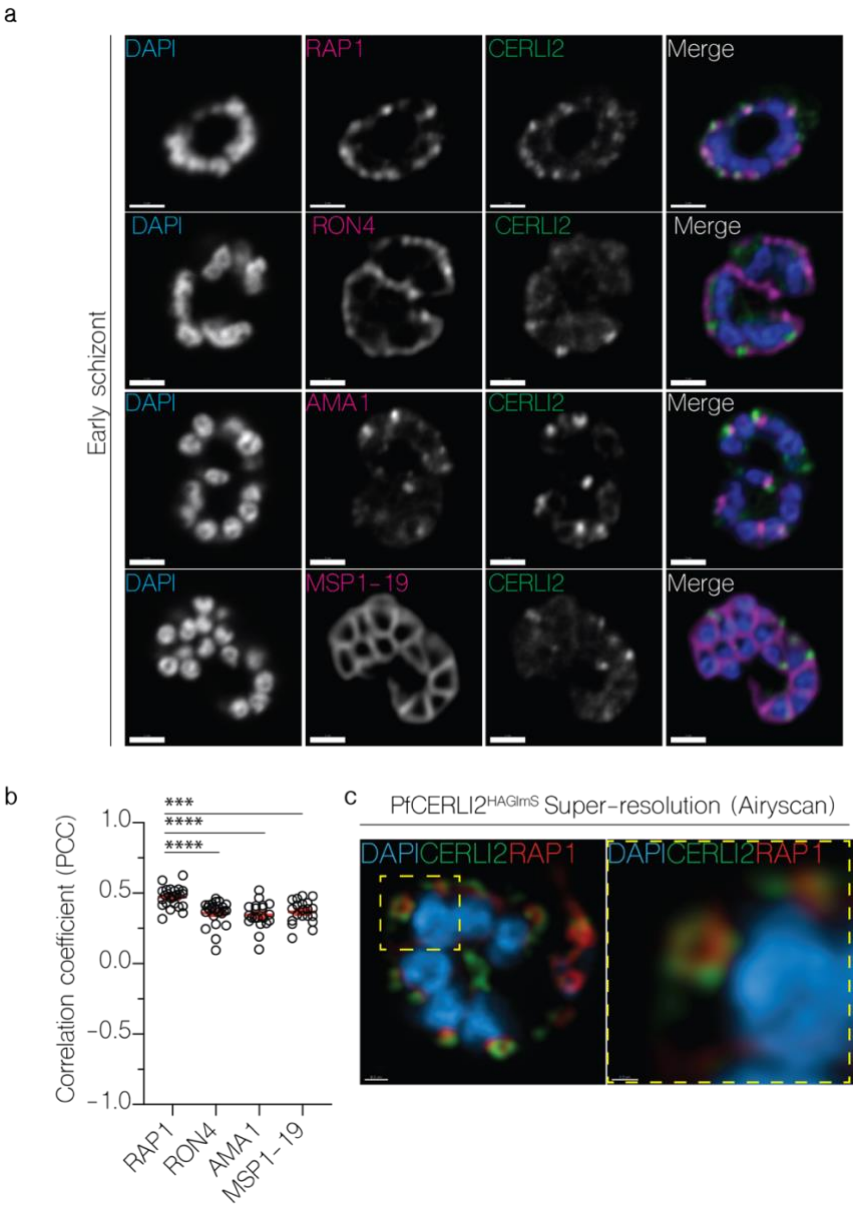


Figure 4: PfCERLI2 localises to the rhoptry bulb.

(a) *2D Confocal immunofluorescence microscopy of early PfCERLI2^{HAGlmS} schizonts stained with DAPI (nucleus) and anti-HA (PfCERLI2) antibodies, along with antibodies to either RAP1 (rhoptry bulb), RON4 (rhoptry neck), AMA1 (micronemes), or MSP1-19 (merozoite surface). Scale bar = 2 µm* ***(b)*** *Quantification of colocalisation between PfCERLI2 and the merozoite organelle markers RAP1, RON4, AMA1 and MSP1-19. Colocalisation quantified as Pearson's correlation coefficient (PCC) when the PfCERLI2 signal was defined as the region of interest. n= 3 biological replicates, 6 schizonts per replicate. Error bars = SEM,*

*** = p<0.001, **** = p<0.0001. **(c)** *Representative maximum-intensity projection of PfCERLI2^{HAGlmS} parasites stained with antibodies to RAP1 (rhoptry bulb) and HA (CERLI2) and imaged using the super-resolution microscopy platform Airyscan. Yellow box indicates the zoom area for the right-hand panel.*

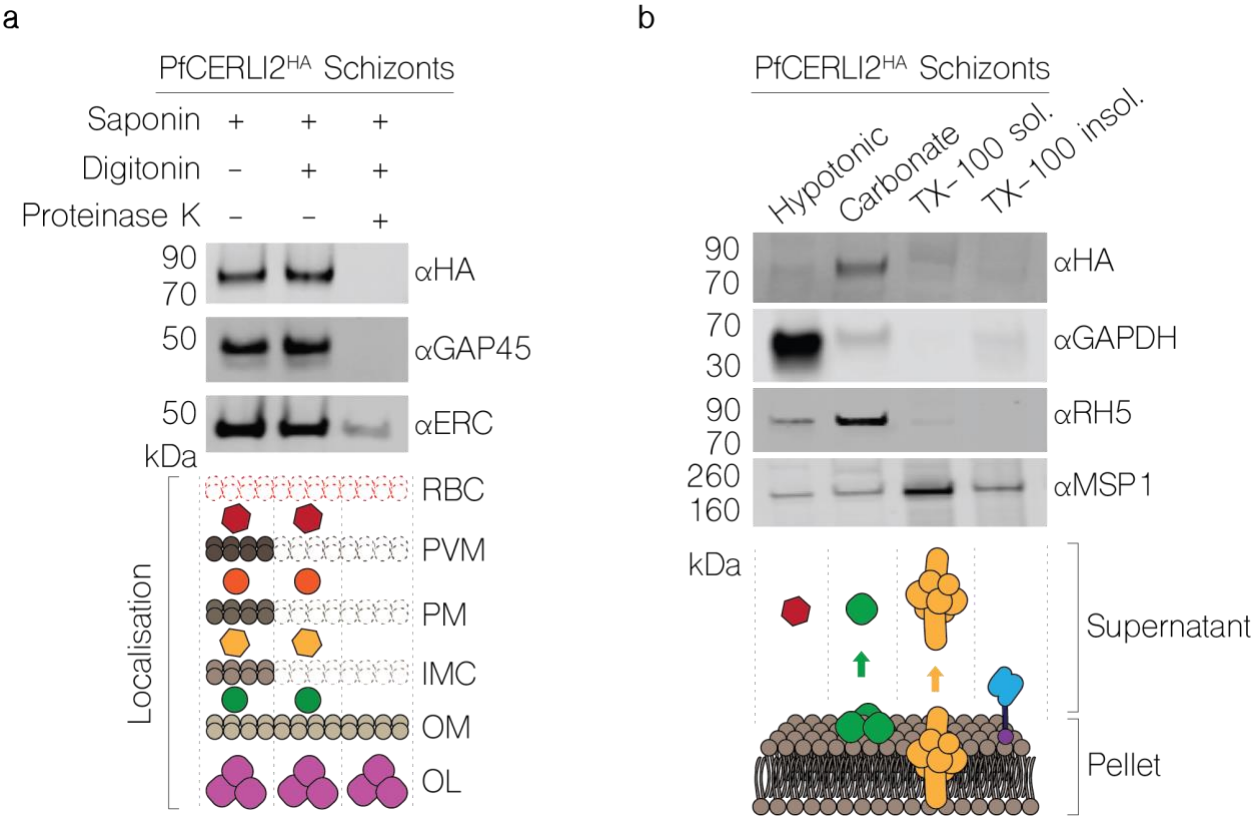


Figure 5: PfCERLI2 is peripherally associated with the cytosolic face of the rhoptry bulb membrane.

(a) *PfCERLI2^{HAGlmS} Schizont lysates were treated with either saponin alone, saponin and digitonin, or saponin, digitonin, and proteinase K. Saponin lyses the RBC membrane and PVM, while digitonin lyses the parasite plasma membrane but leaves the membrane of internal organelles intact. Treated lysates were then probed with anti-HA (PfCERLI2), anti-ERC (endoplasmic reticulum internal protein), or anti-GAP45 (inner-membrane complex, exposed to cytosol) antibodies with ERC and GAP45 serving as negative and positive controls, respectively, for proteinase K digestion. Images representative of 3 independent experiments. RBC = RBC membrane, PVM = PV membrane, PM = parasite plasma membrane, IMC = inner-membrane complex, OM= organellar membrane, OL = organelle lumen.* ***(b)*** *PfCERLI2^{HAGlmS} schizont lysates were subjected to differential lysis and solubilisation to determine the membrane solubility of PfCERLI2. Lysates were first hypotonically lysed, before being treated sequentially with sodium carbonate and triton-x-100. The supernatant was collected following each treatment, along with a final triton-x-100 soluble fraction. Each fraction was then probed with anti-HA (PfCERLI2), anti-GAPDH*

(cytosolic), anti-RH5 (peripheral) and anti-MSP1 (transmembrane domain containing) antibodies. Images representative of 3 independent experiments.

PfCERLI2 knockdown inhibits rhoptry bulb antigen processing but not secretion of rhoptry neck proteins

PfCERLI1 has previously been shown to be involved with rhoptry secretion[9,10], and given the similarities between PfCERLI1 and PfCERLI2 we sought to determine whether PfCERLI2 was also involved in rhoptry secretion. To test the effect of PfCERLI2 knockdown on rhoptry secretion, ring-stage PfCERLI2^{HAGlmS} parasites were treated with 2.5 mM GLCN or left untreated. The RBCs of each culture had been treated with trypsin, chymotrypsin, and neuraminidase to prevent reinvasion. Parasites were cultured until schizont rupture when the supernatant was collected (containing secreted proteins) and a parasite lysate (containing merozoites that had failed to invade) were prepared for Western blot (Figure 6a). As expected, there was no change in secretion of the micronemal antigen erythrocyte binding antigen 175 (EBA-175) with knock-down of the rhoptry associated PfCERLI2 (Figure 6b). There was also no noticeable defect in the secretion of the rhoptry neck antigens reticulocyte binding homologue 4 (RH4) and rhoptry neck protein 4 (RON4). These data indicate that PfCERLI2 knock-down does not block secretion of rhoptry neck antigens in merozoites that fail to invade enzyme treated RBCs.

Parasite lysates from the secretion assay, containing free merozoites, were also probed with antibodies to the rhoptry bulb antigen rhoptry associated protein 1 (RAP1) (Figure 6a). RAP1 has previously been shown to be processed by Plasmepsin IX and subtilisin-like protease 1 (SUB1) from an 84 kDa precursor in intact schizonts into a 67 kDa mature product in free merozoites[20-23]. Diminished PfCERLI2 expression leads to the inhibition of RAP1 processing, corresponding to an accumulation of the unprocessed form of the protein (Figure 6c). When

PfCERLI2^{HAGlmS} schizonts either with, or without, 2.5 mM GLCN were arrested prior to parasitophorous vacuole membrane (PVM) rupture with the protein kinase G inhibitor compound 1 (C1) [24] and probed for RAP1 (Supplementary Figure 12a), both 2.5 mM GLCN treated and untreated cultures showed equally little processing of RAP1, supporting that the defect in RAP1 processing occurs only after schizont rupture and merozoite release (Supplementary Figure 12b). Collectively, this suggests that knockdown of PfCERLI2 inhibits the processing of RAP1 between PVM rupture and merozoite invasion, a phenotype previously demonstrated for PfCERLI1 knockdown[9].

PfCERLI2 knockdown alters rhoptry morphology

While PfCERLI2 knockdown schizonts showed no gross morphological defects, we aimed to determine if PfCERLI2 knockdown altered the morphology or biogenesis of the rhoptries. We matured GLCN treated or untreated 3D7 WT and PfCERLI2^{HAGlmS} schizonts in the presence of C1 and subsequently fixed and imaged these mature schizonts using thin-section transmission electron microscopy (TEM) (Figure 6d). PfCERLI2^{HAGlmS} schizonts that had not been treated with GLCN were morphologically indistinguishable to both GLCN treated and untreated 3D7 WT parasites. By contrast, GLCN treated PfCERLI2^{HAGlmS} parasites appeared to contain noticeably elongated rhoptry necks. We measured rhoptry length in randomly selected sections (38 untreated, 33 GLCN treated) and the measured rhoptry lengths were on average ~90 nm (± 3.4% SEM) longer in GLCN treated PfCERLI2^{HAGlmS} parasites than those of untreated controls (Figure 6e).

Dimensional analysis of organelles is partially compromised in thin section TEM, as XY measurements of an organelle are influenced by differences in sectioning. To overcome this, we performed serial block face scanning electron microscopy (SBF-SEM) array tomography of both GLCN treated and untreated PfCERLI2^{HAGlmS} schizonts (Figure 7a-f) to

analyse changes in rhoptry morphology. Rhoptries of GLCN treated PfCERLI2^{HAGlmS} parasites (*n* = 100) were on average 130 nm (± 10 nm SEM) (Figure 7g) longer than rhoptries from untreated parasites (*n* = 100) representing a 24.17% rhoptry lengthening following PfCERLI2 knockdown. Additionally, rhoptries of GLCN treated PfCERLI2^{HAGlmS} parasites exhibited similar increases in surface area (Figure 7h) and volume (Supplementary Figure 13a) but no change in surface area to volume ratio (Supplementary Figure 13b). By contrast to their lengthening, GLCN treated PfCERLI2^{HAGlmS} parasites showed a small decrease in rhoptry width (Supplementary Figure 13c) by an average of 12 nm (± 5.5 nm SEM), leading to overall longer and narrower rhoptries following PfCERLI2 knockdown (Supplementary Figure 13d).

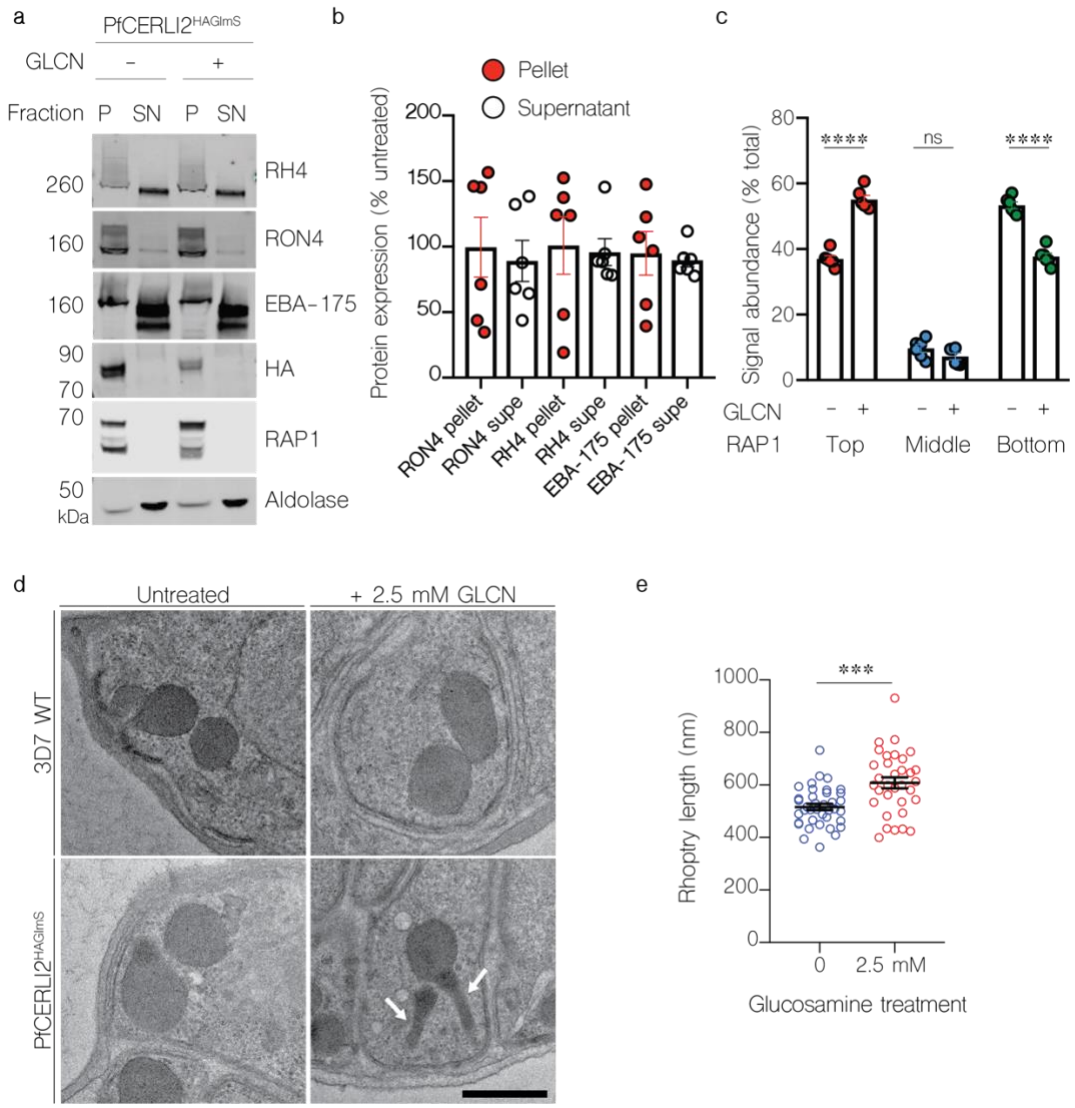


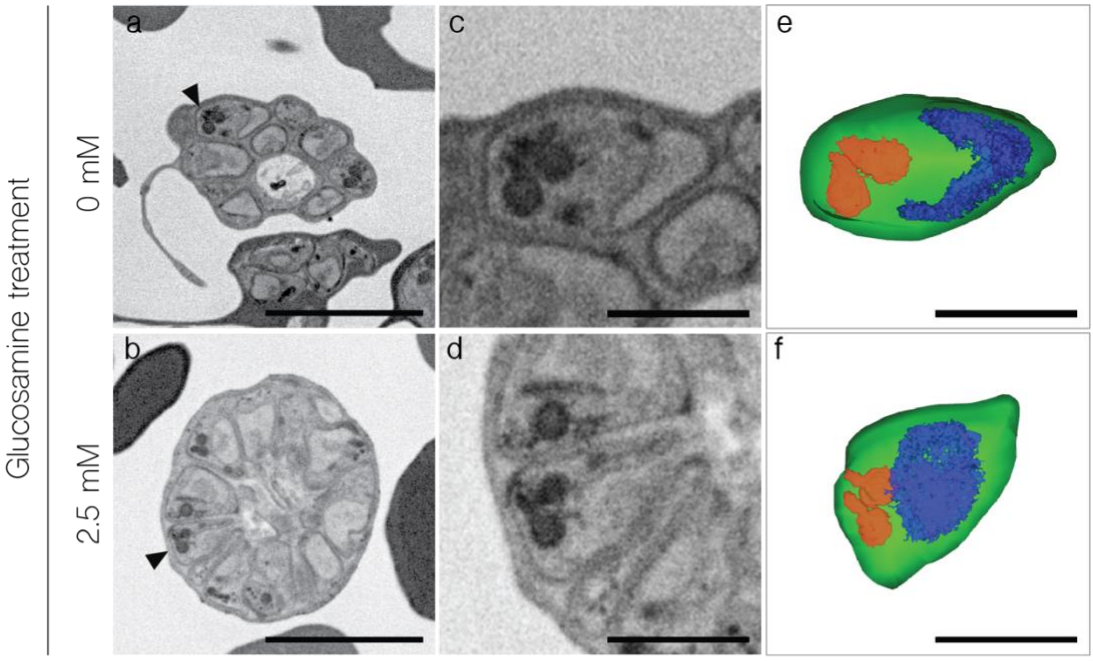
Figure 6: Knockdown of PfCERLI2 inhibits rhoptry antigen processing, but not secretion, and is associated with rhoptry malformation.

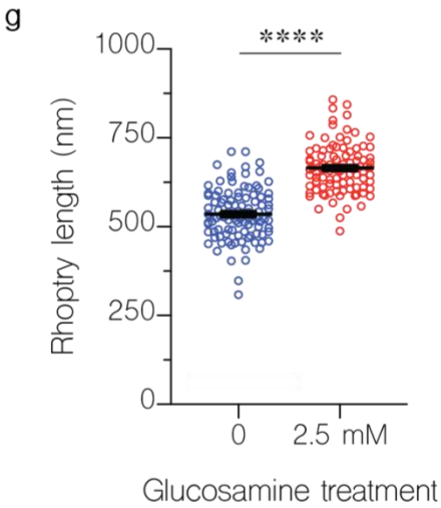
(a) *PfCERLI2^{HAGlmS} ring-stage parasites were GLCN treated (+) or left untreated (-), before enzyme treatment of RBCs to prevent reinvasion. Lysates (P) and culture supernatants (SN) were harvested following schizont rupture and probed with anti-RH4, anti-RON4 (rhoptry neck), anti-EBA-175 (microneme), anti-HA (PfCERLI2), anti-RAP1 (rhoptry bulb), or anti-aldolase (loading control) antibodies. Representative blots of 5 biological replicates shown.* ***(b)*** *Western blots were normalised to the loading control, aldolase, and quantified, with results displayed as protein expression in GLCN treated sample as a percentage of the untreated signal. n=5 biological replicates.* ***(c)*** *Quantification of individiual band intensities, as a percentage of the total signal, from RAP1 signals using the parasite lysates from the secretion assay. n=5 biological replicates. ns= p>0.05, **** = p<0.0001 by two-way ANOVA* ***(d)*** *Transmission electron microscopy of C1 schizont arrested PfCERLI2 and 3D7 WT parasites, either in the presence or absence of GLCN. Representative images from 3 biological replicates.* ***(e)*** *Quantification of rhoptry length from PfCERLI2^{HAGlmS} TEM images. n=3 biological replicates. *** = p<0.001. All error bars = SEM.*

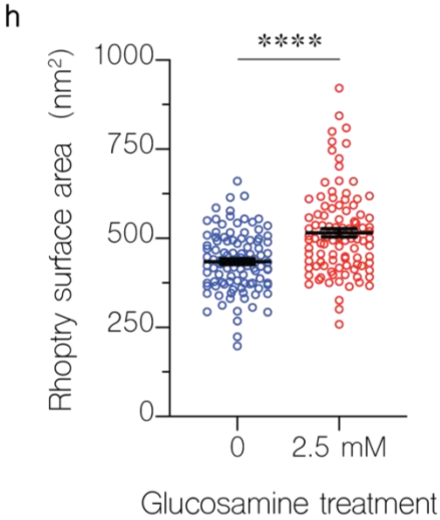
To further confirm the rhoptry morphological changes we observed by electron microscopy, we performed super-resolution (airyscan) microscopy on either untreated or 2.5 mM GLCN treated PfCERLI2^{HAGlmS} parasites, probed with rhoptry bulb and neck markers RAP1 and RON4. We then measured the distance between the basal end of the nucleus and the centre of the RAP1 or RON4 (Supplementary Figure 14) foci, determining the average distance between nucleus and rhoptry. In GLCN treated parasites, we consistently observed a shorter distance between the nucleus and both RAP1 (Figure 7i) and RON4 (Figure 7j) foci (Supplementary Figure 14). This closer proximity of rhoptry markers to the nucleus supports the observation of rhoptry lengthening following PfCERLI2 knockdown.

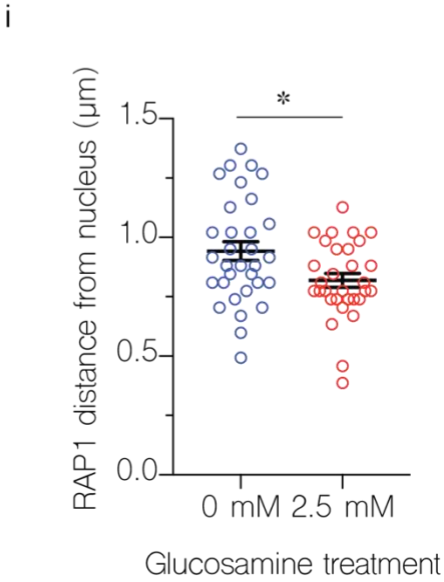
To determine if distribution of these rhoptry markers changed following PfCERLI2 knockdown, we also quantified the area, volume, and sphericity of RAP1 and RON4 foci in both untreated and 2.5 mM GLCN treated PfCERLI2^{HAGlmS} parasites (Supplementary Figure 15). Consistent with the observation of an elongated rhoptry neck following PfCERLI2 knockdown, RON4 foci of GLCN treated parasites displayed a much larger surface area (Supplementary Figure 15a), larger volume (Supplementary Figure 15b) and were less

spherical (Supplementary Figure 15c). Additionally, RAP1 foci in 2.5 mM GLCN treated parasites displayed a lower surface area and volume (Supplementary Figure 13d&e), although the magnitude of this change was small, but no change in sphericity (Supplementary Figure 13f). To determine if these minor changes in RAP1 foci shape following PfCERLI2 knockdown were indicative of a change to the size of the rhoptry bulb, RAP1 foci diameter was measured; but no change in rhoptry bulb diameter was observed in GLCN treated parasites (Supplementary Figure 15g). Thus, PfCERLI2 knockdown alters rhoptry morphology and is associated with incomplete processing of the rhoptry bulb protein RAP1, suggesting that PfCERLI2 plays an important role in normal rhoptry physiology.


Glucosamine treatment
0 mM
2.5 mM
a
b
c
d
e
f


g
****
Rhoptry length (nm)
1000
750
500
250
0
0
2.5 mM
Glucosamine treatment


h
****
Rhoptry surface area (nm²)
1000
750
500
250
0
0
2.5 mM
Glucosamine treatment


i
*
RAP1 distance from nucleus (µm)
1.5
1.0
0.5
0.0
0 mM
2.5 mM
Glucosamine treatment

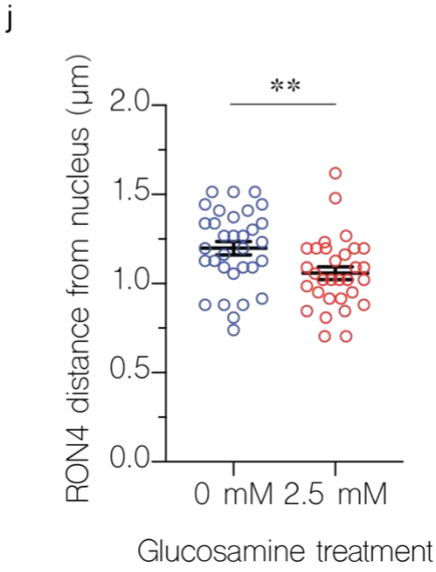

j
**
RON4 distance from nucleus (µm)
2.0
1.5
1.0
0.5
0.0
0 mM
2.5 mM
Glucosamine treatment

Figure 7: Knockdown of PfCERLI2 alters rhoptry length and positioning.

*PfCERLI2^{HAGlmS} schizonts were matured in the presence of E64 and imaged using serial block face scanning electron microscopy (SBF-SEM) array tomography. Inverted backscattered images of untreated **(a)** and GLCN treated **(b)** PfCERLI2^{HAGlmS} schizonts. **(c&d)** Higher magnification images revealing individual merozoites and associated rhoptry pairs indicated by black arrow heads in a & b. **(e&f)** 3D renderings corresponding to the images in c & d. Colour legend: Merozoite plasma membrane, green; nucleus, blue; rhoptries, red. Scale bars: a & b = 4 µm; c-f = 1 µm. Using SBF-SEM array tomography, 100 rhoptries for each treatment were were analysed to determine their length **(g)** and surface area **(h)**. PfCERLI2^{HAGlmS} schizonts were matured in the presence of E64, stained with antibodies against RAP1 (rhoptry bulb) and RON4 (rhoptry neck), and imaged by Airyscan super-resolution microscopy. The flourescence intensity from maximum-intensity projections of RAP1 **(i)** and RON4 **(j)** signals were then measured from the basal end of the nucleus. 32 merozoites for untreated and 31 merozoites for 2.5 mM GLCN treated were measured from a total of 12 schizont images across three biological replicates. Error bars = SEM. *=p<0.05, **=p<0.01, ****=p<0.0001 by unpaired t-test.*

DISCUSSION

We and Suarez *et al*. had previously identified the rhoptry bulb protein PfCERLI1 and determined that it played an essential role in merozoite invasion and rhoptry secretion[9,10]. Bioinformatic searches identified PfCERLI2 as a protein potentially related to PfCERLI1 based on their amino acid sequence homology. Here, we show that *cerli2* likely arose through ancestral gene duplication from *cerli1*. Furthermore, we show that like PfCERLI1, PfCERLI2 also localises to the cytosolic face of the rhoptry bulb, is essential for merozoite invasion and has important functions in rhoptry biology.

A previous study identified a growth inhibitory monoclonal antibody, known as M26-32, which was suggested to bind to PfCERLI2 and was reported to localise this protein to the merozoite surface[15]. However, it had also been shown that MG26-32 reacted with multiple *P. falciparum* proteins, and that the peptide it recognised was contained in many proteins[15]. Our results indicate, using endogenous tagging of PfCERLI2, that it localises to the rhoptry bulb of developing schizonts. Therefore, it is highly unlikely to be merozoite surface exposed and targetable with invasion inhibitory antibodies.

PfCERLI2 solubilised mainly in the carbonate treatment, suggesting it is perhipherally associated with the rhoptry bulb membrane. PfCERLI1 was present in both the carbonate and triton-x100 insoluble fractions, which we hypothesised may be a result of strong cooperative membrane association through its C2 domain, PH domain, and palmitoylation site[9]. By contrast, PfCERLI2 appears to lack a functional PH domain and the palmitoylation site possessed by PfCERLI1. In the most likely *Toxoplasma gondii* homologue of PfCERLI2 (TGME49_315160, also known as TgRASP2), the PH domain (degenerated in PfCERLI2) is involved in, but not essential for, rhoptry localisation and secretion[10]. Based on evidence from its *T. gondii* homologue and the lack of a functional PH domain and palmitoylation site, we suggest that membrane attachment of PfCERLI2 is mediated by its C2 domain, a structure involved in targeting many proteins to various membranes[25].

Knockdown of PfCERLI2 led to reduced parasite growth and merozoite invasion. Morphology and rate of schizont rupture in PfCERLI2 knockdown parasites were indistinguishable compared to the controls, indicating parasite development was not impacted by knockdown of PfCERLI2. Instead, we observed increased numbers of both free merozoites in the culture medium and merozoites bound to the RBC surface in PfCERLI2 knockdown parasites. Based on this evidence, the growth inhibition seen with PfCERLI2 knockdown is most likely to be a result of merozoite invasion being interrupted at or before formation of the irreversible tight-junction with the RBC surface. The increase in RBC bound merozoites following PfCERLI2 knockdown suggests that rhoptry neck and micronemal protein contents were secreted for PfCERLI2 deficient merozoites, and the TJ formed for some of these, but these merozoites were then unable to proceed with invasion. The fact that we saw no reduction in secretion of the rhoptry neck proteins RON4 and RH4 following CERLI2 knockdown supports the idea that in many cases PfCERLI2 knockdown parasites can secrete enough protein to form the TJ, but these merozoites fail to invade due to a later-occurring defect.

The presence of a prominent decapeptide repeat region at the C-terminus of PfCERLI2 is a differentiating feature compared to PfCERLI1. Based on its differential evolution relative to the rest of the PfCERLI2 protein, and its exclusive presence in *Laverania*, we hypothesise that the decapeptide tandem repeat number may be under positive selection, whereas amino acids outside the repeat region are under negative selection. It has previously been reported that the decapeptide tandem repeat contains highly immunogenic antigens that bind malaria-exposed serum antibodies. A potential mechanism for the positive selection of the PfCERLI2 repeat could therefore be classical antibody selection. However, our data using endogenously tagged PfCERLI2 shows that the protein localises to the cytosolic face of the rhoptry bulb membrane and has a role in the function of this organelle, making it unlikely that PfCERLI2 is significantly exposed to the host immune system and antibody selection. We therefore suggest that the PfCERLI2 repeat has an as yet unidentified function in *P. falciparum* whose evolution is under positive selection based on its Ka/Ks value of 1.43 across *Laverania*. Amongst *P. falciparum* isolates, repeat number varies and it is currently unclear whether the same occurs among different isolates of all *Laverania*, or whether this is a unique property of PfCERLI2. Given the conservation of the repeat sequence in *P. falciparum*, but the variation in repeat number, this variation in repeat number may have an important role in this protein's function between isolates.

Due to the presence of two paralogous genes in all coccidians (e.g. *Toxoplasma gondii* and *Besnoitia* spp.) and haematozoa (e.g *Plasmodium* spp.), it is likely that an ancestral *cerli1* sequence was duplicated in the most recent common ancestor of these related parasite lineages. By contrast, all more divergent Apicomplexans and Chromerids contain only a single *cerli* gene, suggesting that there was originally only a single *cerli* gene in the more distant ancestors of coccidia and haematozoa. *Cryptosporidium* is more divergent than coccidia and haematozoa and contains only a single CERLI homologue. As the gene structure of the *cerli* homologue of

Cryptosporidium is highly similar to that of *cerli1* across other Apicomplexa, we suggest that *cerli1* is most similar to the ancestral gene that underwent duplication and gave rise to the more divergent *cerli2* gene. From studies of both PfCERLI1 and PfCERLI2, it is clear that both proteins have functional similarities due to their stage of expression and localisation, along with their knockout and knockdown phenotypes[9,10,13]. Additionally, in *Toxoplasma gondii* it has been reported that TgRASP1 and TgRASP2 form a complex[10], but it not clear if this occurs in other Apicomplexa. For the coccidia and haematozoa containing *cerli* paralogues, it is likely that the duplicated *cerli2* locus that arose from gene duplication of an ancestral *cerli1* locus underwent subfunctionalisation, likely contributing to the function of an ancestral CERLI protein but also recent CERLI1 homologous sequences. This is supported by shared characteristics of PfCERLI1 and PfCERLI2, along with the strong negative selection CERLI2 has been under across Apicomplexa. A potential driver for this subfunctionalisation of the paralogous genes may be a direct binding of CERLI1 to CERLI2, which has been observed for RASP1 and RASP2 in *T. gondii*[10] and could constrain sequence variation of the paralogues.

CERLI2 homologues in coccidia and haematozoa possess a PH domain towards their C-terminus that is under negative selection. The exception to this is the genus *Plasmodium*, where CERLI homologues instead contain a degenerated and likely non-functional PH domain. Additionally, PfCERLI2 homologues in the *Laverania* subgenus of *Plasmodium* also have a decapeptide tandem repeat that is under positive selection. While it is not clear what the function of this repeat region is, the likely absence of the PH domain's lipid-binding function [10] and presence of the decapeptide repeat region may indicate that CERLI2 in *Laverania* has undergone neosubfunctionalisation, where following subfunctionalisation one paralogue adopts a new function[26]. We can therefore speculate that CERLI2 homologues in *Laverania* may have modified functions or different protein-protein/protein-lipid interactions compared to homologues outside this subgenus.

METHODS

Bioinformatic analyses

PfCERLI2 (Pf3D7_0405200), PfCERLI1 (Pf3D7_0210600) and orthologous sequences used for comparison *of Pfcerli1* and *Pfcerli2* gene structure from *Babesia bigemina* (BBBOND_0209090, BBBOND_0311360), *B. bovis* (BBOV_IV005720, BBOV_III003780), *Cryptosporidium andersoni* (cand_006180), *C. hominis* (Chro.60252), *C. muris* (CMU_007720), *C. parvum* (cgd6_2143), *C. tyzzeri* (CTYZ_00001395), *C. ubiquitum* (cubi_02338), *Chromera Velia* (Cvel_21914), *Cyclospora Cayetanensis* (cyc_01372), *Cystoisopora Suis* (CSUI_005102), *Cytauxzoon Felis* (CF002482, CF003293), *Eimeria brunetti* (EBH_0041390), *G. niphandrodes* (GNI_053280), *N. caninum* (NCLIV_057920, NCLIV_049720), *P. reichenow*i (PRCDC_0402800, PRCDC_0209500), *P. malariae* (PmUG01_03015300, PmUG01_04021700), *P. ovale* (PocGH01_03012700, PocGH01_04019500), *P. knowlesi* (PKNH_0303300, PKNH_0410600), *P. vivax* (PVP01_0304600, PVP01_0414300), *P. berghei* (PBANKA_1002900, PBANKA_0307500), *P. yoelii* (PY17X_1004300, PY17X_0308100) and *P. chabaudi* (PCHAS_1003800, PCHAS_0309700), *S. neurona* (SN3_01800355, SN3_00600405), *T. annulata* (TA21020, TA12010), *T. equi* (BEWA_030720, BEWA_019330), *T. gondii* (TGME49_315160, TGME49_235130), *T. orientalis* (TOT_010000334, TOT_020000207), *V. brassicaformis* (Vbra_14966) were obtained by searching within the EuPathDB.org database[27]. Additionally, homologues of PfCERLI2 from the Laveranian parasites *P. adleri* (PADL01_0404800), *P. billcollinsi* (PBILCG01_0404100), *P. blacklocki* (PBLACG01_0402700), *P. gaboni* (PGABG01_0403400) *and P. praefalciparum* (PFRFG01_0404700) were obtained by searching within the PlasmoDB.org database[11].

Both pairwise and multiple sequence alignments were generated using Geneious 9.1.3 (Biomatters) using a Geneious global alignment with free end gaps (gap open penalty = 12, gap extension penalty = 3) algorithm with the Blosum62 cost matrix. Multiple sequence alignments were also constructed using Muscle[28] in an R environment[29]. Phylogenetic trees were constructed from nucleotide and amino acid sequences using the Jukes-Cantor[30] DNA or LG[31] amino acid substitution models, respectively, genetic distance model and were built using the unweighted pair group method with arithmetic mean (UPGMA) tree build method using Geneious or the R package phangorn v. 2.5.5[32]. A tanglegram visualisation was generated using the cophyloplot function of the R package ape[33].

To identify *P. falciparum* proteins that contained a PHIS, or PHIS-like motif, the PHIS sequence from PfCERLI1, PfCERLI2, or the consensus sequence generated from their Apicomplexan homologues was used in a BLASTp search against the *P. falciparum* 3D7 proteome.

Ka/Ks ratios were calculated using the online Ka/Ks Calculation tool (http://services.cbu.uib.no/tools/kaks)[34,35] using the maximum likelihood tree method, discrete_Grantham submatrix, and a LI rate of moderate. Where Ka/Ks ratios are reported, they were generated using the same coding sequences as the corresponding phylogenetic tree.

Protein structure prediction

The protein structure of PfCERLI1, PfCERLI2, along with their homologues in *T. gondii* (TGME49_235130, TGME49_31516) and *C. muris* (CMU_007720) was predicted using the online protein structure prediction tool Phyre2 on intensive mode[36]. All predicted structures were visualised, presented and imaged using either Jmol[37] or EzMol[38].

Continuous culture of asexual stage *P. falciparum*

P. falciparum 3D7 parasites were cultured in O^+ red blood cells (Red Cross Blood Service Australia) as previously described[39]. Parasites were grown in RPMI-HEPES media (Sigma-Aldrich) supplemented with 0.5% v/v Albumax (Gibco), 52 μM gentamycin (Gibco), 367 μM hypoxanthine (Sigma-Aldrich), 2 mM L-Glutamax (Gibco) and 2 mM sodium bicarbonate (Thermo Fisher Scientific), adjusted to a pH of 7.4. Cultures were maintained at 37ºC in sealed acrylic boxes with the following gas composition: 1% O2, 5% CO2 and 94% N2.

Plasmid construction and transfection

List of genetic constructs in this study: *Pfcerli2*^{HAGlmS}, *Pfcerli2*$^{HAGlmS/GFP}$, *Pfcerli2*$^{SLI-TGD}$.

Pfcerli2^{HAGlmS} riboswitch transfection vectors were prepared from a PTEX150^{HAGlmS} plasmid backbone [40]. A 758 base pair homology region of *Pfcerli2* was PCR amplified from the 3' end of the genomic sequence of 3D7 gDNA using the primers *Pfcerli2* 5' F RBW (GGT**AGATCT**GAAGGAATCTTTTGGAGATGAGC) and *Pfcerli2* 3' R RBW (GGT**CTGCAGC**TATATTTTGTATGGTATTTTCTAATTGTGC). Each of the resulting PCR products were digested with the restriction enzymes BglII and PstI (indicated in bold in primer sequence) and cloned into the PTEX150^{HAGlmS} vector.

To generate the *Pfcerli2*$^{HAGlmS/GFP}$ parasite line, which expresses cytosolic GFP, the pHGBrHrBl-1/2 GFP plasmid was used, without modification[41].

For disruption of *Pfcerli2* using selection linked integration – targeted gene disruption (SLI-TGD) system, a source SLI-TGD vector (Pf3D7_1463000 SLI-TGD) was used[16]. A 772 bp homology region of *Pfcerli2* was PCR amplified using the primers *Pfcerli2* SLI-TGD F

(GGT**GCGGCCGC**TCCACATATAAGTGATTTCGAGCC) and *Pfcerli2* SLI-TGD R (GGT**ACGCGT**CCTTGCACACTTCTTGTCC). Both PCR product, and the SLI-TGD vector were then digested with NotI and MluI (restriction sites in bold) and ligated together.

P. falciparum 3D7 parasites were transfected by RBC loading[42] (*Pfcerli2*^{HAGlmS}, *Pfcerli2*$^{HAGlmS/GFP}$) or schizont transfection[43] (*Pfcerli2*$^{SLI-TGD}$). For RBC loading, uninfected RBCs were centrifuged at 440 rcf for 1 min, washed in culture media and then washed in cytomix (0.895% KCl, 0.0017% $CaCl_2$, 0.076% EGTA, 0.102% $MgCl_2$, 0.0871% K_2HPO_4, 0.068% KH_2PO_4, 0.708% HEPES). Washed RBCs were then resuspended in cytomix containing 200 µg of the plasmid of interest. The combined RBCs and DNA were transferred to a 0.2 cm cuvette (Bio-Rad) to be electroporated (Bio-Rad) at 0.31 kV with a capacitance of 960 µF. Electroporated RBCs were rinsed from the cuvette with culture media, centrifuged at 440 rcf for 1 min, and washed twice before being introduced to cultures of gelatin-purified schizont stage parasites. Alternatively for schizont transfection, late stage schizont cultures were Percoll-purified[44] and resuspended in 90µl transfection buffer (90mM Na_2HPO_4; 5mM KCl; 50mM HEPES; 0.15mM $CaCl_2$; pH 7.3) plus 10µl of TE containing 50 µg of the plasmid of interest, and electroporated using a 2D Nucleofector (Lonza) on program U-033. Electroporated parasites were transferred into 1.5 mL tubes with 1mL culture media and 200 µL uninfected RBCs and shaken at 37ºC for 30 min to promote merozoites to invade, before being placed back into culture. To generate *Pfcerli2*^{HAGlmS}, 5 nM WR99210 (Jacobus Pharmaceuticals) drug treatment was used to select for integration of the HA-*GlmS* construct over 3 cycles. For *Pfcerli2*$^{HAGlmS/GFP}$, the HA-*GlmS* line was transfected with the pHGBrHrBl-1/2 construct in order to episomally express cytosolic GFP then selected for using 5 µg/mL blasticidin-S-deaminase HCl (Merck Millipore). To generate *Pfcerli2*$^{SLI-TGD}$, maintenance of the SLI-TGD construct was selected using 4 nM WR99210 and integrants were selected using 400 µg/mL G418 (geneticin, Sigma). *Pfcerli2*$^{SLI-TGD}$ plasmids were transfected twice into 3D7

parasite culture, which after they grew on WR99210 and were selected each four times independently with G418 for eight weeks.

Assessment of *in vitro* blood stage growth and invasion

To investigate the effects of *Pfcerli2* knockdown on parasite growth 3D7 and *Pfcerli2*^{HAGlmS} cultures were synchronised to ring-stages using sorbitol lysis and adjusted to 1% parasitaemia at 1% haematocrit in 96 well U-bottom plates (Corning) as previously described[45]. From this adjusted culture, 45 μL was then suspended in 5 μL of a 10 x stock of D-(+)-glucosamine hydrochloride (hereby referred to as GLCN, Sigma-Aldrich) or media to a final volume of 50 μL. Plates were then incubated in standard culture conditions for 72 hours, until trophozoite stages the following cycle. Completed assays were then stained in 10 μg/mL ethidium bromide (Bio-Rad) in PBS and the final parasitaemia was assessed using the BD Accuri™ C6 Plus flow cytometer (PE-H, FSC-H filters for growth, PE-H, FITC-H for invasion to visualise GFP-fluorescent ring stages). All flow cytometry files were analysed using FlowJo Version 10 (Tree Star).

To investigate the effect of *Pfcerli2* knockdown on invasion, *Pfcerli2*^{HAGlmS} cultures were synchronised to early trophozoite stages using sorbitol lysis and adjusted to 1% parasitaemia at 1% haematocrit and placed in 96 well U-bottom plates and set up as described for growth assays. Plates were subsequently incubated in standard culture conditions for 24 hours, until early ring stages the following cycle. Completed assays were stained with ethidium bromide and analysed by flow cytometry as described for growth assays.

To observe whether PfCERLI2 knockdown had an impact on merozoite quantity per schizont or merozoite morphology, *Pfcerli2*^{HAGlmS} parasites were treated with or without 2.5 mM GLCN at ring stages until matured into 40 hr schizonts. To ensure full merozoite development, schizonts were arrested with E64 for 4 hrs to prevent rupture and then smeared

as a thin blood film. Smears were fixed in 100% v/v methanol and stained in 10% v/v Giemsa (Merck Millipore) in water and then blindly assessed by light microscopy (n = 20 individual schizonts). To count free merozoites, *Pfcerli2*$^{HAGlmS/GFP}$ schizonts +/- 2.5 mM GLCN without E64 were permitted to incubate until rupture and gated for merozoites by flow cytometry as described above.

Schizont rupture assay

To observe whether PfCERLI2 knockdown affected schizont rupture and merozoite egress, synchronous *Pfcerli2*^{HAGlmS} ring stages were treated in duplicate with or without 2.5 mM GLCN in 96-well U bottom plates as described above. Parasites were permitted to grow into schizonts prior to rupture and the parasitaemia was recorded by flow cytometry. Plates were then incubated for an additional 6 hrs to allow for schizont rupture followed by a second round of parasitaemia measurement by flow cytometry. To calculate the percentage of schizont rupture within this time window, the following equation was used:

$$\text{\% schizont rupture} = \left(\frac{post-rupture\ schizontaemia}{pre-rupture\ schizontaemia}\right) \times 100$$

Bound merozoite assay

To determine whether PfCERLI2 knockdown inhibited merozoite invasion prior to, or following, tight junction formation PfCERLI2^{HAGlmS} ring-stage parasites were either treated with 2.5 mM GLCN or left untreated and incubated until schizont rupture. Following schizont rupture, Giemsa-stained smears were made of these cultures without washing or removing the culture medium, blinded and counted to determine the number of bound merozoites and newly invaded rings in each treatment. For each treatment >1000 RBCs were counted per replicate.

Saponin lysis and Western blot

To prepare protein samples, high parasitaemia cultures were lysed in 0.15% w/v saponin (ThermoFisher Scientific) on ice for 10 min, centrifuged at 16000 rcf, washed once in 0.075% w/v saponin and twice in PBS. All lysis reagents included protease inhibitors (CØmplete, Roche). Parasite lysates were DNase I treated (Qiagen) for 5 min at room temperature before being resuspended in reducing sample buffer (0.125 M Tris-HCl pH 7, 20% v/v glycerol, 4% v/v SDS, 10% v/v β-mercaptoethanol (Sigma-Aldrich), 0.002% w/v bromophenol blue (Sigma-Aldrich)). Protein lysates were separated by size on SDS-PAGE 4-12% Bis-tris Gels (Bolt, Invitrogen) at 110 V for 80 min then transferred onto a nitrocellulose membrane (iBlot, Invitrogen) at 20 V for 7 min. Membranes were blocked in 1% w/v skim milk 0.05% v/v Tween20 (Sigma-Aldrich) in PBS (hereby referred to as 1% milk PBS-T) for 1 hr at room temperature. Primary (mouse 12CA5 anti-HA (1:5000 Roche), rabbit anti-aldolase (1:5000 abcam), rabbit anti-GAP45 [46] (1:5000), mouse anti-RAP1[47] (1:10000), rabbit anti-RON4[48] (1:5000), rabbit anti-Rh4[49] (1:5000), rabbit anti-EBA175[50] (1:10000) and secondary (IRDye ® 800CW goat anti-mouse (1:4000, LI-COR Biosciences), IRDye® 680RD goat anti-rabbit (1:4000, LI-COR Biosciences) antibodies were prepared in 1% milk PBS-T and incubated on the membranes for 1 hr each at room temperature while rocking. Primary antibodies were washed three times in 0.05% v/v PBS-Tween20 (hereby referred to as PBS-T), while secondary antibodies were washed twice in PBS-T and once in PBS. Completed blots were dried on filter paper (Whatman) and visualised on the Odyssey Infrared Imaging System (LI-COR Biosciences). All band quantification was performed in Image Studio Lite 2.5 (LI-COR Biosciences).

Proteinase K protection assay

To determine whether PfCERLI2 was cytosolically exposed or compartmentalised in an organelle, a proteinase K protection assay modified from a previous study was performed[7,9]. Three 10 mL aliquots of high schizontaemia cultures were centrifuged at 440 rcf for 5 min, supernatant removed and lysed in 0.15% w/v saponin as described above. One set of schizonts were treated in SOTE (0.6 M sorbitol, 20 mM Tris HCl pH 7.5, 2 mM EDTA) alone, the second treated in SOTE with 0.02% w/v digitonin (Sigma-Aldrich) left incubating for 10 min at 4ºC before being washed in SOTE, and the third digested with 0.1 μg/μL Proteinase K (Sigma-Aldrich) in SOTE for 30 min at 4ºC following the aforementioned digitonin treatment. Proteinase K was inactivated with 50 μL 100% v/v trichloroacetic acid followed by PBS. All samples were resuspended in 500 μL acetone to ensure full Proteinase K deactivation, centrifuged then washed twice in 500 μL MilliQ H_2O. The final pellets were used for Western blot analysis of PfCERLI2's sensitivity to Proteinase K.

Protein solubility assay

To biochemically determine whether PfCERLI2 was membrane-associated and by what mechanism of association, a protein solubility assay was performed as previously described[9]. Briefly, a high schizontaemia culture was saponin-lysed as described above, resuspended in 100 μL MilliQ H_2O, snap-frozen in dry ice four times, passed through a 29 gauge needle 5 times to disrupt parasite membranes, centrifuged at 16000 rcf for 10 min before reserving the water-soluble fraction containing cytosolic proteins. The remaining pellet was washed twice in MilliQ H_2O and once in PBS before being resuspended in 100 μL 0.1 M Na_2CO_3 for 30 min at 4ºC, centrifuged before reserving the carbonate-soluble fraction containing peripherally-associated proteins. The pellet was washed in the same manner above before being resuspended in 0.1% v/v Triton X-100 for 30 min at 4 ºC, centrifuged and the final supernatant reserved

containing integral proteins. The remaining pellet was washed twice in PBS and was used to represent Triton X-100-insoluble proteins. All fractions were analysed by Western blot to determine PfCERLI2's membrane solubility profile.

Sample preparation for fixed-cell immunofluorescence microscopy

Pfcerli2^{HAGlmS} cultures of 3% E64 arrested schizonts were centrifuged at 440 rcf for 3 min, washed in PBS, resuspended in fixative (4% v/v paraformaldehyde (PFA, Sigma-Aldrich), 0.0075% v/v glutaraldehyde (Electron Microscopy Sciences), pH 7.4) and then left gently rocking for 30 min at room temperature. Samples were centrifuged at 440 rcf, fixative removed, washed twice in PBS, then adjusted to 1% haematocrit in PBS. #1.5H high-precision coverslips (Carl Zeiss, Oberkochen, Germany) were soaked in methanol, airdried, coated in 0.01% v/v poly-L-lysine (Sigma-Aldrich) for 30 min at room temperature then washed in MilliQ H_2O. Fixed cells were then laid on top at room temperature for 30 min, with non-adherent cells being gently aspirated off. Cells were permeabilised with 0.1% v/v Triton X-100 for 10 minutes then incubated in 3% bovine serum albumin (BSA) in PBS-T for 1 hr. Primary antibodies (anti-HA biotin conjugate 1:1000 (Roche), mouse anti-RAP1[47] 1:500, mouse anti-AMA1[51]1:500, rabbit anti-RON4[48] 1:500, rabbit anti-MSP1-19[52] 1:500) were diluted in 1% w/v BSA in PBS-T as specified above and applied to the coverslips overnight at 4 ºC. The next day, coverslips were washed three times in PBS-T, before being incubated with Alexa Fluor-conjugated secondary antibodies (streptavidin 488 nm, mouse 594 nm, rabbit 647 nm, Life Technologies) for 1 hour in the dark at room temperature. Coverslips were again washed three times in PBS-T before being dehydrated in ethanol (70% v/v 3 min, 90% v/v 3 min, 100 v/v 3 min), air dried and then mounted on glass slides with 20 µL Prolong® Gold antifade solution (refractive index 1.4) with 4', 6-diamidino-2-phenylindole dihydrochloride (DAPI,

Thermofisher Scientific). The mountant was allowed to cure overnight and coverslips were analysed on an Olympus FV3000 confocal microscope.

Confocal microscopy and colocalisation analysis

Confocal microscopy was performed using an Olympus FV3000 fluorescence microscope (Olympus) equipped with a ×100 MPLAPON oil objective (NA 1.4) using the 405 nm, 488 nm, 561 nm, and 633 nm lasers. Colocalisation analysis was performed using Imaris Coloc Suite (v9.0, Bitplane Inc., Switzerland) as described in detail previously [9]. Briefly, images were exported from the Olympus FV3000 in .oir format, before conversion into .ims format. Following conversion, colocalisation analysis was performed whereby the image dataset was masked to the thresholded PfCERLI2 (anti-HA, 488nm laser) signal, defining the PfCERLI2 signal as the region of interest. Subsequently the other thresholded channels, corresponding to either RAP1, RON4, AMA1, or MSP1-19, were colocalised with the PfCERLI2 signal. Colocalisation statistics reported in this study are Pearson's correlation coefficient between the PfCERLI2 signal and the signal of the corresponding organelle marker, when PfCERLI2 is defined as the region of interest. Colocalisation analyses were performed over three biological replicates. For each of these biological replicates, threshold values were determined for each of the channels used in the colocalisation analysis to differentiate between true signal and background. Threshold values were fixed across all channels for each image in that biological replicate.

Rhoptry and microneme ligand secretion assay

Secretion assays were modified from a previous protocol[9]. Briefly, *Pfcerli2*^{HAGlmS} synchronous ring stages at 15% haematocrit were grown in 6-well plates (Corning) with or without 2.5 mM GLCN for 24 hrs. Once trophozoite stages had formed, cultures were

centrifuged at 440 rcf, supernatant removed and resuspended in an enzyme mix (0.067 U/mL neuraminidase (Sigma-Aldrich), 1 mg/mL chymotrypsin (Worthington Biochemical Corporation), 1 mg/mL trypsin (Sigma-Aldrich)) to be incubated at 37ºC for 45 minutes to cleave RBC surface receptors. Parasites were then washed twice in culture media and returned to culture for a further 24 hrs until schizont rupture. Once ruptured, cultures were centrifuged at 16000 rcf for 10 min at 4ºC to pellet uninvaded merozoites. Supernatants containing secreted protein was collected and kept on ice, and the pellet subject to saponin lysis to represent non-secreted protein. Each fraction was analysed by Western blot and quantified in Image Studio Lite 2.5 (LI-COR Biosciences).

Transmission electron microscopy

Pfcerli2^{HAGlmS} synchronous ring stage cultures were treated with or without 2.5 mM GLCN and grown until early schizonts had formed before being treated with 2 μM C1 for 4 hrs. Late schizonts were then enriched using Percoll then fixed in 2.5% v/v glutaraldehyde (Electron Microscopy Sciences) in PBS overnight at 4ºC. Fixed schizonts were post-fixed in 1% osmium tetroxide (v/v) and 1.5% potassium ferrocyanide (w/v) in 0.1 M Sorensen's phosphate buffer (PB) pH 7.4 for 30 min at room temperature. Cells were washed in 0.1 M PB (3 × 5 min) and incubated in 1% tannic acid (w/v) in 0.1 M PB for 30 min. Samples were washed in ultrapure water (3 × 5 min) and *en bloc* stained in 1% aqueous uranyl acetate (w/v) for 1 h at room temperature. Cells were washed in ultrapure water (3 × 5 min) and incubated in Walton's Lead Aspartate for 30 min at 60ºC. Cells were then washed in ultrapure water (3 × 5 min), dehydrated in increasing concentrations of ethanol-H_2O and absolute acetone and finally embedded in medium-grade Procure 812 resin. 70 nm ultrathin sections were generated using a Leica EM UC7 ultramicrotome (Leica Microsystems) and attached to 100-mesh Cu

formvar-carbon coated grids. Images were acquired using a Tecnai G2 F30 transmission electron microscope (Field Electron and Ion Company, FEI) operating at 200 kV.

Array tomography and rhoptry morphometric analysis

Array tomography[53] was performed in order to obtain 2-D and 3-D morphometric data. Blocks processed for TEM were sectioned on a Leica EM UC7 ultramicrotome (Leica Microsystems), generating 100 consecutive sections of 100 nm-thickness (depth sectioned = 10 µm/sample). Sections were collected on a hydrophilized silicon wafer and allowed to dry at room temperature. An aluminium stub was adhered to the underside of wafers using conductive double-sided adhesive carbon-tabs.

Inverted backscattered field emission scanning electron microscopy was performed using a FEI Teneo SEM (Field Electron and Ion Company, FEI) operating at 5 kV at a working distance of 7.3 mm. Images (6144 × 4096 pixel array, 8-bit, 9.76 × 9.76 × 100 nm voxel dimensions, pixel dwell time 2 µs) were acquired at 3,500× yielding a *XY* field of view of 59.97 × 39.98 µm).

Image histogram stack normalisation was performed, and images were aligned using the StackReg plugin for Fiji[54]. For 3-D modelling and visualisation, datasets were processed using 3dmod, a program bundled with the IMOD software package[53]. Rhoptries were segmented by means of manual tracing of high-contrast lines. 3-D information including surface area and volume was obtained using the “get info” command. Rhoptry length was quantified from SEM micrographs where rhoptry necks could be seen contacting the merozoite plasma membrane and the bottom of the bulb could be clearly discerned. Using blinded assessment, a vector was drawn from the tip of the necks to the end of the rhoptry bulbs in 3dmod to provide a length in nm.

Quantitative analysis of rhoptry length in thin-section schizonts

Rhoptry length was quantified from transmission electron micrographs for rhoptries where necks could be seen contacting the merozoite plasma membrane and the bottom of the bulb could be clearly discerned. Using blinded assessment, a vector was drawn from the tip of the necks to the end of the rhoptry bulbs in ImageJ to provide a length in nm. Micrographs where the rhoptries had been cut sideways such that it removed part of the neck or bulb were excluded from this analysis.

Airyscan microscopy, merozoite measurement, and foci analysis

Airyscan microscopy was performed using a Zeiss LSM800 AxioObserver Z1 microscope (Carl Zeiss, Oberkochen, Germany) that had an Airyscan detector and a Plan-Apochromat 63x M27 oil objective (NA 1.4). Z-stacks were acquired using a 0.04 µm XY and 0.16 µm Z pixel resolution. Images were acquired sequentially as follows: Channel 1 (RON4) – 633 nm laser, channel 2 (RAP1) – 594 nm laser, channel 3 (HA) – 488 nm laser, channel 4 (DAPI) – 405 nm laser.

The distance between the basal end of the nucleus (DAPI signal) and the RAP1 or RON4 foci was measured as described previously[55]. Images were first blinded and then using the Plot Profile plugin on FIJI, a line was drawn from the basal extremity of the DAPI signal through the RAP1 and RON4 foci to generate intensity profiles for the three markers as a function of distance of the line. Data generated from these intensity plots was then exported, and graphed using GraphPad PRISM 7, with the distance corresponding to the highest RAP1 or RON4 signal intensity representing the distance of that signal from the basal end of the DAPI signal.

RON4 and RAP1 foci were analysed using an object analysis pipeline on Imaris (v9.8.0, BitPlane AG, Zurich, Switzerland) that had previously been established for these markers[9]. RAP1 or RON4 signals were pre-processed to segment and split individual foci, and then thresholded to remove background signal. Threshold values for each marker were kept consistent across all replicates and treatments. Identified objects were then filtered on the basis of size, with only objects > 0.1 µm included in the analysis. This analysis was performed in batch mode, so all images were processed and quantified identically. Once the surfaces had been established, their surface area, volume, and sphericity values were exported as .csv files.

RAP1 diameter was measured using the line function on ImageJ. Briefly, images of PfCERLI2^{HAGlmS} parasites that had either been treated with 2.5 mM GLCN or left untreated and stained with anti-RAP1 antibodies were blinded. Rhoptry bulb diameters were then manually quantified by drawing a vector across the diameter of the RAP1 signal. Only rhoptries with visible 'donut-shapes' were measured, with the diameter measured on the Z-slice between the points with the greatest signal intensity. Merozoites that were densely packed inside the schizont, or RAP1 signals where it was unclear which merozoite they belong to, were excluded from the analysis.

Statistical analysis

All graphs and statistical analyses were completed in Graphpad PRISM 7 (GraphPad Software Inc.). All figure legends note the type of statistical test, the definition of significance for various p-values, and the number of biological replicates (n) for each experiment.

ACKNOWLEDGEMENTS

We thank Prof. Alan Cowman for provision of RON4, EBA175 and GAP45 antibodies, A/Prof. Wai-Hong Tham for RH4 antibodies. We also thank Dr. Paul Gilson for the PTEX150^{HAGlmS} transfection vector. Electron microscopy was performed at the Ian Holmes Imaging Centre, Bio21, The University of Melbourne (www.microscopy.unimelb.edu.au). For provision of the SLI-TGD vector, we thank Dr. Tobias Spielmann. We thank Dr. Brad Sleebs for Compound 1. We thank the Australian Red Cross Blood Bank for the provision of human blood.

This work was supported by funding from the NHMRC (Project Grant APP1143974, DW), University of Adelaide Beacon Fellowship and Hospital Research Foundation Fellowship (DW), DAAD/Universities Australia joint research co-operation scheme (TG, DW, BL, JB), Australian Government Research Training Program Scholarship (BL, JB), South Australian Commonwealth Scholarship (BL), DFG BA5213/3-1 (JSW) and Jürgen Manchot-Stiftung fellowship (AA).

AUTHOR CONTRIBUTIONS

Study design and planning: D.W.W., B.L., J.B., M.W.A.D., L.T., and T.G. Performed experiments and generated reagents: B.L., J.B., D.W.W., G.J.S., J.S., G.K.H., A.A, and J.S.W. Data analysis: B.L., J.B., G.J.S., J.S., S.F., and D.W.W. Manuscript writing: B.L, J.B., and D.W.W. Manuscript was drafted with input from all authors.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

1 World Health Organization. *World malaria report 2019*, <www.who.int/malaria/publications/world-malaria-report-2019/en/> (2019).

2 Cowman, A. F., Healer, J., Marapana, D. & Marsh, K. Malaria: Biology and Disease. *Cell* **167**, 610-624, doi:10.1016/j.cell.2016.07.055 (2016).

3 Cowman, A. F., Tonkin, C. J., Tham, W. H. & Duraisingh, M. T. The Molecular Basis of Erythrocyte Invasion by Malaria Parasites. *Cell Host and Microbe* **22**, 232-245, doi:10.1016/j.chom.2017.07.003 (2017).

4 Weiss, G. E., Crabb, B. S. & Gilson, P. R. Overlaying Molecular and Temporal Aspects of Malaria Parasite Invasion. *Trends Parasitol* **32**, 284-295, doi:10.1016/j.pt.2015.12.007 (2016).

5 Counihan, N. A., Kalanon, M., Coppel, R. L. & de Koning-Ward, T. F. *Plasmodium* rhoptry proteins: why order is important. *Trends in Parasitology* **29**, 228-236, doi:10.1016/j.pt.2013.03.003 (2013).

6 Hanssen, E. *et al.* Electron tomography of *Plasmodium falciparum* merozoites reveals core cellular events that underpin erythrocyte invasion. *Cellular Microbiology* **15**, 1457-1472, doi:doi:10.1111/cmi.12132 (2013).

7 Cabrera, A. *et al.* Dissection of Minimal Sequence Requirements for Rhoptry Membrane Targeting in the Malaria Parasite. *Traffic* **13**, 1335-1350, doi:10.1111/j.1600-0854.2012.01394.x (2012).

8 Geiger, M. *et al.* Structural Insights Into PfARO and Characterization of its Interaction With PfAIP. *Journal of Molecular Biology*, doi:https://doi.org/10.1016/j.jmb.2019.12.024 (2019).

9 Liffner, B. *et al.* PfCERLI1 is a conserved rhoptry associated protein essential for *Plasmodium falciparum* merozoite invasion of erythrocytes. *Nature Communications* **11**, 1411, doi:10.1038/s41467-020-15127-w (2020).

10 Suarez, C. *et al.* A lipid-binding protein mediates rhoptry discharge and invasion in *Plasmodium falciparum* and *Toxoplasma gondii* parasites. *Nature Communications* **10**, 4041, doi:10.1038/s41467-019-11979-z (2019).

11 Aurrecoechea, C. *et al.* PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Research* (2008).

12 Woo, Y. H. *et al.* Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. *eLife* **4**, e06974, doi:10.7554/eLife.06974 (2015).

13 Zhang, M. *et al.* Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science* **360** (2018).

14 Bushell, E. *et al.* Functional Profiling of a Plasmodium Genome Reveals an Abundance of Essential Genes. *Cell* **170**, 260-272.e268, doi:10.1016/j.cell.2017.06.030 (2017).

15 Gao, Y. H. *et al.* Identification of a vaccine candidate antigen, PfMAg-1, from *Plasmodium falciparum* with monoclonal antibody M26-32. *Parasitology research* **105**, 1723-1732, doi:10.1007/s00436-009-1617-4 (2009).

16 Birnbaum, J. *et al.* A genetic system to study *Plasmodium falciparum* protein function. *Nature Methods* **14**, 450, doi:10.1038/nmeth.4223 (2017).

17 Lopez-Barragan, M. J. *et al.* Directional gene expression and antisense transcripts in sexual and asexual stages of *Plasmodium falciparum*. *BMC Genomics* (2011).

18 Yao, Z. *et al.* Improved targeting of radiolabeled streptavidin in tumors pretargeted with biotinylated monoclonal antibodies through an avidin chase. *J Nucl Med* **36**, 837-841 (1995).

19 Fujiki, Y., Hubbard, A. L., Fowler, S. & Lazarow, P. B. Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. *J Cell Biol* **93**, 97-102, doi:10.1083/jcb.93.1.97 (1982).

20 Silmon de Monerri, N. C. *et al.* Global identification of multiple substrates for *Plasmodium falciparum* SUB1, an essential malarial processing protease. *Infection and immunity* **79**, 1086-1097, doi:10.1128/IAI.00902-10 (2011).

21 Howard, R. F., Narum, D. L., Blackman, M. & Thurman, J. Analysis of the processing of *Plasmodium falciparum* rhoptry-associated protein 1 and localization of Pr86 to schizont rhoptries and p67 to free merozoites. *Molecular and Biochemical Parasitology* **92**, 111-122, doi:10.1016/S0166-6851(97)00238-7 (1998).

22 Nasamu, A. S. *et al.* Plasmepsins IX and X are essential and druggable mediators of malaria parasite egress and invasion. *Science* **358**, 518, doi:10.1126/science.aan1478 (2017).

23 Favuzza, P. *et al.* Dual Plasmepsin-Targeting Antimalarial Agents Disrupt Multiple Stages of the Malaria Parasite Life Cycle. *Cell Host & Microbe*, doi:10.1016/j.chom.2020.02.005 (2020).

24 Taylor, H. M. *et al.* The Malaria Parasite Cyclic GMP-Dependent Protein Kinase Plays a Central Role in Blood-Stage Schizogony. *Eukaryotic Cell* **9**, 37, doi:10.1128/EC.00186-09 (2010).

25 Corbalan-Garcia, S. & Gómez-Fernández, J. C. Signaling through C2 domains: More than one lipid target. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1838**, 1536-1547, doi:10.1016/j.bbamem.2014.01.008 (2014).

26 Innan, H. & Kondrashov, F. The evolution of gene duplications: classifying and distinguishing between models. *Nature Reviews Genetics* **11**, 97-108, doi:10.1038/nrg2689 (2010).

27 Aurrecoechea, C. *et al.* EuPathDB: the eukaryotic pathogen genomics database resource. *Nucleic Acids Research* (2016).

28 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797, doi:10.1093/nar/gkh340 (2004).

29 R Core Team. *R: A language and environment for statistical computing*, <www.R-project.org/> (2020).

30 Jukes, T. H. & Cantor, C. R. in *Mammalian Protein Metabolism* (ed H.N. Munro) 21-132 (Academic Press, 1969).

31 Le, S. Q. & Gascuel, O. An Improved General Amino Acid Replacement Matrix. *Molecular Biology and Evolution* **25**, 1307-1320, doi:10.1093/molbev/msn067 (2008).

32 Schliep, K. P. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592-593, doi:10.1093/bioinformatics/btq706 (2010).

33 Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526-528, doi:10.1093/bioinformatics/bty633 (2019).

34 Liberles, D. A. Evaluation of Methods for Determination of a Reconstructed History of Gene Sequence Evolution. *Molecular Biology and Evolution* **18**, 2040-2047, doi:10.1093/oxfordjournals.molbev.a003745 (2001).

35 Siltberg, J. & Liberles, D. A. A simple covarion-based approach to analyse nucleotide substitution rates. *Journal of Evolutionary Biology* **15**, 588-594, doi:10.1046/j.1420-9101.2002.00416.x (2002).

36 Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**, 845, doi:10.1038/nprot.2015.053 (2015).

37 Jmol. *Jmol: an open-source Java viewer for chemical structures in 3D*, <http://www.jmol.org>

38 Reynolds, C. R., Islam, S. A. & Sternberg, M. J. E. EzMol: A Web Server Wizard for the Rapid Visualization and Image Production of Protein and Nucleic Acid Structures.

Journal of Molecular Biology **430**, 2244-2248, doi:https://doi.org/10.1016/j.jmb.2018.01.013 (2018).

39 Trager, W. & Jensen, J. B. Human malaria parasites in continuous culture. *Science* **193**, 673 (1976).

40 Elsworth, B. *et al.* PTEX is an essential nexus for protein export in malaria parasites. *Nature* **511**, 587-591, doi:10.1038/nature13555 (2014).

41 Wilson, D. W., Crabb, B. S. & Beeson, J. G. Development of fluorescent *Plasmodium falciparum* for *in vitro* growth inhibition assays. *Malaria journal* **9**, 152, doi:10.1186/1475-2875-9-152 (2010).

42 Deitsch, K., Driskill, C. & Wellems, T. Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. *Nucleic acids research* **29**, 850-853 (2001).

43 Moon, R. W. *et al.* Adaptation of the genetically tractable malaria pathogen Plasmodium knowlesi to continuous culture in human erythrocytes. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 531-536, doi:10.1073/pnas.1216457110 (2013).

44 Rivadeneira, E. M., Wasserman, M. & Espinal, C. T. Separation and concentration of schizonts of *Plasmodium falciparum* by Percoll gradients. *The Journal of protozoology* **30**, 367-370 (1983).

45 Wilson, D. W., Langer, C., Goodman, C. D., McFadden, G. I. & Beeson, J. G. Defining the timing of action of antimalarial drugs against *Plasmodium falciparum*. *Antimicrobial agents and chemotherapy* **57**, 1455-1467 (2013).

46 Baum, J. *et al.* A Conserved Molecular Motor Drives Cell Invasion and Gliding Motility across Malaria Life Cycle Stages and Other Apicomplexan Parasites. *Journal of Biological Chemistry* **281**, 5197-5208, doi:10.1074/jbc.M509807200 (2006).

47 Schofield, L. *et al.* A rhoptry antigen of *Plasmodium falciparum* contains conserved and variable epitopes recognized by inhibitory monoclonal antibodies. *Molecular and Biochemical Parasitology* **18**, 183-195, doi:10.1016/0166-6851(86)90037-X (1986).

48 Richard, D. *et al.* Interaction between *Plasmodium falciparum* Apical Membrane Antigen 1 and the Rhoptry Neck Protein Complex Defines a Key Step in the Erythrocyte Invasion Process of Malaria Parasites. *Journal of Biological Chemistry* **285**, 14815-14822, doi:10.1074/jbc.M109.080770 (2010).

49 Tham, W.-H. *et al.* Complement receptor 1 is the host erythrocyte receptor for *Plasmodium falciparum* PfRh4 invasion ligand. *Proceedings of the National Academy of Sciences* **107**, 17327, doi:10.1073/pnas.1008151107 (2010).

50 Lopaticki, S. *et al.* Reticulocyte and Erythrocyte Binding-Like Proteins Function Cooperatively in Invasion of Human Erythrocytes by Malaria Parasites. *Infection and Immunity* **79**, 1107, doi:10.1128/IAI.01021-10 (2011).

51 Coley, A. M. *et al.* The most polymorphic residue on *Plasmodium falciparum* apical membrane antigen 1 determines binding of an invasion-inhibitory antibody. *Infection and immunity* **74**, 2628-2636, doi:10.1128/IAI.74.5.2628-2636.2006 (2006).

52 O'Donnell, R. A., Saul, A., Cowman, A. F. & Crabb, B. S. Functional conservation of the malaria vaccine antigen MSP-119across distantly related *Plasmodium* species. *Nat Med* **6**, 91-95, doi:10.1038/71595 (2000).

53 Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer Visualization of Three-Dimensional Image Data Using IMOD. *Journal of Structural Biology* **116**, 71-76, doi:https://doi.org/10.1006/jsbi.1996.0013 (1996).

54 Thevenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans Image Process* **7**, 27-41, doi:10.1109/83.650848 (1998).

55 Ebrahimzadeh, Z. *et al.* A pan-apicomplexan phosphoinositide-binding protein acts in malarial microneme exocytosis. *EMBO reports* **20**, e47102, doi:10.15252/embr.201847102 (2019).