# Sequence elements within the PEXEL motif and its downstream region modulate PTEX dependent protein export in Plasmodium falciparum .

Paul R. Gilson<sup>1</sup>, Mikha Gabriela<sup>1</sup>, Claudia B. G. Barnes<sup>1</sup>, Dickson Leong<sup>1</sup>, Brad E. Sleebs<sup>2</sup>, Molly P. Schneider<sup>1</sup>, Dene R. Littler<sup>3</sup>, Brendan S. Crabb<sup>1</sup>, and Tania F. de Koning-Ward<sup>4</sup>

<sup>1</sup>Burnet Institute <sup>2</sup>Walter and Eliza Hall Institute of Medical Research <sup>3</sup>Monash University Monash Biomedicine Discovery Institute <sup>4</sup>Deakin University School of Medicine

March 23, 2023

## Abstract

The parasite *Plasmodium falciparum* causes the most severe form of malaria and to invade and replicate in red blood cells (RBCs), it exports hundreds of proteins across the encasing parasitophorous vacuole membrane (PVM) into this host cell. The exported proteins help modify the RBC to support rapid parasite growth and avoidance of the human immune system. Most exported proteins possess a conserved *Plasmodium* Export Element (PEXEL) motif with the consensus RxLxE/D/Q amino acid sequence, which acts as a proteolytic cleavage recognition site within the parasite's endoplasmic reticulum (ER). Cleavage occurs after the P<sub>1</sub> L residue and is thought to help release the protein from the ER so it can be putatively escorted by the HSP101 chaperone to the parasitophorous vacuole space surrounding the intraerythrocytic parasite. HSP101 and its cargo are then thought to assemble with the rest of a *Plasmodium* Translocon for Exported proteins (PTEX) complex, that then recognises the xE/D/Q capped N terminus of the exported protein and translocates it across the vacuole membrane into the RBC compartment. Here, we present evidence that supports a dual role for the PEXEL's conserved P<sub>2</sub>' position E/Q/D residue, firstly, for plasmepsin V cleavage in the ER, and secondly, for efficient PTEX mediated export across the PVM into the RBC. We also present evidence that the downstream 'spacer' region separating the PEXEL motif from the folded functional region of the exported protein controls cargo interaction with PTEX as well. The spacer must be of a sufficient length and permissive amino acid composition to engage the HSP101 unfoldase component of PTEX to be efficiently translocated into the RBC compartment.

## Title.

Sequence elements within the PEXEL motif and its downstream region modulate PTEX dependent protein export in *Plasmodium falciparum*.

## Running Title.

Protein export in *Plasmodium falciparum*.

## Authors.

Mikha Gabriela<sup>#1,2</sup>, Claudia B. G. Barnes<sup>#1</sup>, Dickson Leong<sup>1</sup>, Brad E. Sleebs<sup>3,4</sup>, Molly P. Schneider<sup>1</sup>, Dene R. Littler<sup>5</sup>, Brendan S. Crabb<sup>1,4,6,7</sup>, Tania F. de Koning-Ward<sup>2,8</sup>, and Paul R. Gilson<sup>1,6</sup>\*

<sup>1</sup> Malaria Virulence and Drug Discovery Group, Burnet Institute, Melbourne, Australia

<sup>2</sup> School of Medicine, Deakin University, Geelong, Victoria, Australia

<sup>3</sup> The Walter and Eliza Hall Institute of Medical Research, Parkville, Australia

<sup>4</sup> Department of Medical Biology, The University of Melbourne, Parkville, Australia

<sup>5</sup>Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia,

<sup>6</sup> Department of Immunology and Microbiology, University of Melbourne, Parkville, Australia

<sup>7</sup> Department of Immunology and Pathology, Monash University, Melbourne, Australia

<sup>8</sup>Institute for Mental and Physical Health and Clinical Translation (IMPACT), Deakin University, Geelong, Australia

\* Corresponding Author: paul.gilson@burnet.edu.au

# These authors contributed equally

Format for Traffic .

### Synopsis.

The malaria parasite *Plasmodium falciparum*, exports hundreds of proteins into the red blood cells (RBC) it infects to help it grow and replicate. Most exported proteins possess a five amino acid barcode that acts as a proteolytic cleavage site and licenses these proteins for export into the RBC compartment by protein translocating gateways surrounding the parasite. We report here that the conserved fifth amino acid position of the barcode plays an important role in the proteolytic maturation of exported protein and, together with the downstream flexible spacer region preceding the folded region of the exported protein, are important for efficient transport into the RBC via *Plasmodium* Translocon of Exported Proteins (PTEX).

## Abstract

The parasite *Plasmodium falciparum* causes the most severe form of malaria and to invade and replicate in red blood cells (RBCs), it exports hundreds of proteins across the encasing parasitophorous vacuole membrane (PVM) into this host cell. The exported proteins help modify the RBC to support rapid parasite growth and avoidance of the human immune system. Most exported proteins possess a conserved *Plasmodium* Export Element (PEXEL) motif with the consensus RxLxE/D/Q amino acid sequence, which acts as a proteolytic cleavage recognition site within the parasite's endoplasmic reticulum (ER). Cleavage occurs after the  $P_1$  L residue and is thought to help release the protein from the ER so it can be putatively escorted by the HSP101 chaperone to the parasitophorous vacuole space surrounding the intraerythrocytic parasite. HSP101 and its cargo are then thought to assemble with the rest of a *Plasmodium* Translocon for Exported proteins (PTEX) complex, that then recognises the xE/D/Q capped N terminus of the exported protein and translocates it across the vacuole membrane into the RBC compartment. Here, we present evidence that supports a dual role for the PEXEL's conserved  $P_2$ ' position E/Q/D residue, firstly, for plasmepsin V cleavage in the ER, and secondly, for efficient PTEX mediated export across the PVM into the RBC. We also present evidence that the downstream 'spacer' region separating the PEXEL motif from the folded functional region of the exported protein controls cargo interaction with PTEX as well. The spacer must be of a sufficient length and permissive amino acid composition to engage the HSP101 unfoldase component of PTEX to be efficiently translocated into the RBC compartment.

## Keywords.

Plasmodium falciparum, malaria, protein export, PTEX, virulence, PEXEL, chaperone.

# Introduction

The pathogenesis of malaria lies upon the ability of *Plasmodium* parasites, the causative agent of the disease, to propagate asexually within red blood cells (RBCs) while avoiding the immune mechanisms of its host.

To accomplish this, the parasite exports hundreds of its proteins into its host cell to transform it into a hospitable niche and to avoid host immunity <sup>1-5</sup>. The intraerythrocytic stage of the parasite resides within a membranous compartment called the parasitophorous vacuole (PV) <sup>6,7</sup>. Consequently, parasite proteins destined for export must traverse two membranes, the parasite plasma membrane and the encasing parasitophorous vacuole membrane (PVM), to reach the host RBC. The *Plasmodium t* ranslocon of *ex* ported proteins (PTEX) mediates the translocation of proteins across the PVM <sup>8</sup> and is the only protein channel known to reside at the PVM. PTEX is comprised of three core components, HSP101, PTEX150, and EXP2, that are all essential for parasite survival, making it an attractive drug target candidate<sup>8-10</sup>.

Most parasite proteins destined for export into the RBC begin their journey with entry into the endoplasmic reticulum (ER) via the Sec61 translocon, followed by cleavage of the protein within the ER by the aspartyl protease, plasmepsin V (PMV)<sup>11-13</sup>. The cleavage step occurs within a pentameric amino acid motif near the N-terminus of the exported protein destined for export, termed the *Plasmodium* export element (PEXEL)<sup>1,2,14</sup>. The resulting mature proteins then travel via the vesicular transport pathway to the parasite plasma membrane where they are secreted into the parasitophorous vacuole<sup>15</sup>. Here the exported proteins are unfolded and translocated across the PVM and into the host-cell compartment in an ATP-dependent manner by PTEX <sup>16</sup>. It is thought the AAA+ ATPase HSP101 first engages and unfolds cargo proteins in an ATP-dependent manner and threads them through a tetradecamer membrane-spanning channel consisting of the scaffold protein PTEX150 and the pore-forming protein EXP2<sup>17-19</sup>. However, the information that is contained within proteins destined for export that allows HSP101 to specifically recognise these proteins prior to the unfolding step is still unknown<sup>20,21</sup>.

Historically, the N-terminal region of exported proteins (and the protein's transmembrane domain in some cases) have been found to be sufficient to mediate the trafficking of proteins into the RBC compartment  $^{22-26}$ . This N-terminal region usually contains a recessed signal peptide to promote entry of the proteins into the parasite's ER  $^{27}$  or a hydrophobic stretch of amino acids that serve the same function  $^{28}$ . The discovery of the five-amino acid PEXEL motif in the N-terminal region of many exported proteins was the first export-specific signal<sup>1,2</sup> that was proven to be predictive for exported proteins  $^{29}$  and allowed the identification of >450 putative exported proteins in *P. falciparum*<sup>4</sup>. Interestingly, however, the presence of the PEXEL motif alone does not guarantee export, as its location within a protein's primary structure  $^{30}$  and the presence of `12 amino acids downstream of the motif have also been shown to be required to achieve efficient export  $^{24}$ . Furthermore, PEXEL-negative exported proteins (PNEPs) are also present in the *Plasmodium* exportome  $^{5,31,32}$ , suggesting that PEXEL is not strictly required for passage through PTEX.

The function of the PEXEL motif has been relatively well studied and its consensus sequence is RxLxE/D/Q where x can be almost any amino acid. The position  $P_3$  R and  $P_1$  L residues are the most conserved of the PEXEL motif and in *P. falciparum* have been shown to be necessary for efficient cleavage by  $PMV^{12-14}$ . It is assumed therefore that  $P_1$  and  $P_3$  facilitate PEXEL cleavage and thus help to release the exported proteins from the ER membrane where they are originally anchored following ER import <sup>11</sup>. In comparison, the purpose of the last conserved  $P_2$ ' (E/Q/D) residue of PEXEL is less certain. Some evidence suggests that following cleavage, the  $P_1$ ' and  $P_2$ ' resides which now cap the mature PEXEL protein, play a role in promoting export across the PVM <sup>1,14,33</sup>. After cleavage, the exposed xE/Q/D motif becomes N-terminally acetylated<sup>34</sup> resulting in an Ac- xE/Q/D cap that was thought to serve as a 'barcode' or a recognition motif for HSP101 engagement in the PV <sup>35</sup>. Contrary to this, however, other studies have found that  $P_2$ ' amino acid also influences the efficiency of PEXEL processing or does not influence export at all<sup>36-38</sup>.

Here, we present evidence that the  $P_2$ ' residue has a dual function. For some PEXEL proteins,  $P_2$ ' mutations greatly reduce PMV cleavage increasing protein retention in the ER and reducing export. For other proteins,  $P_2$ ' mutations do not inhibit PMV cleavage as much and PEXEL proteins can reach the PV but are translocated less efficiently. We also provide evidence that the length of the region downstream of the PEXEL motif universally regulates the degree of cargo interaction with HSP101 and ultimately affects protein export across the PVM.

Results

2.1 A systematic review of PEXEL motif mutants reveals that a single point mutation of P2' into a positively charged residue produces the most consistent export-blocking effect.

To better understand the contribution of the PEXEL  $P_2$ ' position on protein export, we conducted a review of P<sub>2</sub>' mutagenesis experiments that have been performed on various PEXEL proteins (Table S1). We sorted the mutants based on the amino acids incorporated into the  $P_2$ ' position and linked it to the export phenotype, PEXEL processing pattern and the N-terminal acetylation (Table S1 'PMV cleavage' and 'Nacetylation') reported for the corresponding mutant. This in general revealed that alanine (A) mutations in the  $P_2$ ' position have a limited or uncertain effect on export (Table S1). For example when the  $P_2$ ' position of KAHRP (knob-associated histidine rich protein, PF3D7\_0202000) was exchanged with an A, the protein was not exported and was retained in the parasite  $^{39}$ . This was contrary to the study by  $^{14}$ , where the reporter was mostly trapped in the PV, despite both groups using very similar constructs (i.e. the first 69aa of KAHRP leader sequence fused to GFP). The same discrepancy was also observed with  $P_2$ ' A mutations of STEVOR and GBP130 (PF3D7\_1016300) (Table S1), Rather than regard these findings as contradictory, the different localisation of  $P_2$ ' A mutants could arise because of the continuum from the ER to the PV that might be dependent on the parasite age and the different module(s) appended downstream of the PEXEL motif<sup>36,38</sup>. P<sub>2</sub>' A mutagenesis study on other PEXEL proteins, such as RESA (PF3D7\_0102200), PfEMP3 (PF3D7\_0201900), PF3D7\_0532600, and murine *P. berghei* CP1 (PBANKA\_1246500) did not result in any export defect (Table S1).

Mutation of the  $P_2$ ' PEXEL position to positively charged amino acids appears to produce the most robust export-blocking effect on exported proteins tested in two different studies (Table S1)<sup>36,40</sup>. Mutation to basic amino acids would usually reverse the charge of the  $P_2$ ' position, given that the two most common  $P_2$ ' residues are glutamic and aspartic acids. Curiously, the  $P_2$ ' lysine (K) mutant of REX3 (PF3D7\_0936300) PEXEL also moderately inhibited PMV cleavage of the motif <sup>36</sup>, suggesting that  $P_2$ ' may also be involved in PEXEL processing, albeit to a lesser extent than  $P_1$  and  $P_3$ .

 $P_2$ ' mutations can also cause changes to the N-acetylation profile of the mature protein. However, N-terminal acetylation is not sufficient to determine the fate of exported proteins, as suggested previously <sup>12,36,38</sup>. Whilst it is unclear whether PEXEL proteins without N-terminal acetylation can be exported or not, having an N-terminal acetylation by itself does not guarantee successful export, as can be seen in the case of  $P_2$ ' A mutant of GBP130 and  $P_2$ ' K of STEVOR (PF3D7\_0631900) which fail to be exported despite being shown to be N-terminally acetylated (Table S1)<sup>14</sup>. We therefore conclude that substituting the conserved  $P_2$ ' PEXEL reside with a basic amino acid would likely produce a more pronounced export-blocking phenotype in the parasite.

 $2.2 P_2$ ' Lysine mutations reduces export efficiency in multiple PEXEL proteins through the inhibition of plasmepsin V cleavage

To clarify the effects of  $P_2$ ' mutations of different PEXEL proteins expressed under the same conditions we synthesised several fluorogenic peptides containing wildtype PEXEL motifs and mutations thereof of three different PEXEL proteins and determined how well they were cleaved by recombinantly-expressed *P*. *vivax*plasmepsin V (*Pv* PMV) (Table S2) <sup>41</sup>.*Pv* PMV was employed as this protease was more experimentally amenable than the *P. falciparum* equivalent<sup>41,42</sup>. Our first peptide from KAHRP sequence (RTLAQ) containing P<sub>3</sub> R to A and P<sub>1</sub> L to A mutations (ATAAQ) served as a control and was barely cleaved at all by *Pv* PMV compared to the wildtype sequence (Fig 1A). The next Hyp1 (PF3D7\_0113300) and STEVOR (subtelomeric variable open reading frame, PF3D7\_0200400) fluorogenic peptides (RLLTE and RLLAQ, respectively) had their P<sub>2</sub>' resides exchanged from E to K for Hyp1 and from Q to K for STEVOR. In both cases, the P<sub>2</sub>' K fluorogenic peptides were barely cleaved indicating that P<sub>2</sub>' K strongly inhibits PEXEL processing (Fig 1A).

Previously, a GFP-tagged STEVOR with  $P_2$ ' A mutation was shown to be trapped at the parasite periphery by microscopy<sup>33</sup>. However, the PEXEL processing status of this mutant was not investigated. We therefore included  $P_2$ ' A mutants of both Hyp1 and STEVOR in our cleavage assay. Compared to the wildtype peptides, cleavage of the  $P_2$ ' A mutants of Hyp1 and STEVOR were moderately inhibited, but this was only statistically significant for Hyp1 (58%). The  $P_2$ ' A mutation thus likely represents an intermediate between the WT and the charge reversal  $P_2$ ' K mutation (Fig 1A).

We next sought to determine how the  $P_2$ ' K mutations would affect the trafficking and proteolytic processing of these proteins in parasite infected RBCs. To generate reporters specific for these proteins we fused the first 113aa of Hyp1 containing the PEXEL motif RLLTE and a P<sub>2</sub>' K mutant version, to a reporter cassette comprising nanoluciferase (Nluc), murine dihydrofolate reductase (mDH) and three FLAG epitopes (FL)<sup>42</sup>. For STEVOR, we fused the first 99 amino acids of the protein that includes the PEXEL motif RLLAQ, and the corresponding P<sub>2</sub>' K mutant version to Nluc-mDH-FL. For KAHRP we fused the first 105aa including PEXEL motif RTLAQ and the P<sub>2</sub>' K mutant version to the same reporter. All six constructs were transfected into the HSP101-HA glmS parasite background line to enable direct comparison between the wildtype and P<sub>2</sub>' K mutants. We note that microscopy and western blot data of wildtype (WT) and P<sub>2</sub>' K Hyp1-NlucmDH-FL have been previously published but these are included them here as a comparator to KAHRP and STEVOR<sup>42</sup>. All P<sub>2</sub>' wildtype (P<sub>2</sub>' WT) reporters were exported with most of the signal in the RBC compartment (Fig 1B-D). For the P<sub>2</sub>' K mutants, Hyp1 differed from STEVOR and KAHRP with Hyp1 being largely trapped around the nucleus in the ER and P<sub>2</sub>' K STEVOR and P<sub>2</sub>' K KAHRP being trapped in the ER and around the parasite circumference in the PV (Fig 1B). Quantification of the reporter signals in KAHRP and STEVOR infected RBCs parasites indicated there was more reporter within the parasite with the  $P_2$ ' K reporters than the  $P_2$ ' WT reporters (Fig 1C).

We have previously shown that the Hyp1  $P_2$ ' K reporter was not efficiently cleaved at the PEXEL motif by Pv PMV and that this could be why this reporter was not exported and remained trapped in the ER <sup>42</sup>. To determine if incorrect cleavage of the KAHRP and STEVOR reporters could also be responsible for the increased trapping of the reporters in the parasite we performed western blot analysis on parasite lysates expressing WT and P<sub>2</sub>' K Hyp1-, -STEVOR- and KAHRP-Nluc-mDH-FL constructs. In the western blot analysis of Hyp1 parasite lysates, the predominant correctly cleaved species of WT Hyp1 migrates at 50 kDa and the full-length pre-processed species migrates at about 70 kDa (Fig 1D, lane 1, single and triple asterisks, respectively). In contrast, cleavage of the Hyp1 P<sub>2</sub>' K reporter appears to be upstream of the PEXEL probably near the transmembrane domain (Fig 1D, lane 2, arrow)<sup>42</sup>. Incorrect cleavage correlates with the high ER retention observed by IFA although the mechanism for this is unknown (Fig 1B and D) <sup>42</sup>. In contrast, P<sub>2</sub>' Lys STEVOR and KAHRP constructs migrated predominantly at the same size as their WT counterpart on a western blot (Fig 1D, lanes 4-7), indicating that the P<sub>2</sub>' K STEVOR and KAHRP are mostly processed within their PEXEL which would explain why they visually appear more efficiently trafficked to the PV relative to Hyp1 P<sub>2</sub>' K (Fig 1B). However, we noted an additional low abundance mis-cleaved band for the KAHRP and STEVOR P<sub>2</sub>' K reporters that are approximately 3-4 kDa bigger than the PEXEL cleaved species (Fig 1D, lanes 5 and 7, double asterisks). We were able to detect these mis-cleaved species using anti-FLAG antibody, suggesting they are not C-terminally truncated form of the full-length protein (Fig 1D, Table S3). We therefore conclude that the mis-cleaved species likely represent aberrant N-terminally processed forms of P<sub>2</sub>' K KAHRP and STEVOR reporters that arise from less efficient PEXEL processing, that may account for the small reduction in export. Collectively, both western blot and biochemical analyses suggest that a  $P_2$ ' mutation, particularly to a positively charged residue, can reduce efficient cleavage for Hyp1 and cause ER retention but not for KAHRP and STEVOR. For the latter two proteins, cleavage of the  $P_2$ ' K PEXEL is much more efficient and the proteins traffic to the PV but are translocated less effectively into the RBC than WT reporters.

2.3 ER trapped Hyp1  $P_2$ ' K is retained in a more insoluble form than PV trapped  $P_2$ ' K STEVOR and KAHRP

We have previously observed that the ER-trapped mis-cleaved Hyp1  $P_2$ ' K reporter was poorly soluble which may partly account for its ER retention <sup>42</sup>. To determine if PV trapping of STEVOR and KAHRP  $P_2$ ' K reporter proteins was due to reduced solubility, we performed protein solubility assays on parasite lines expressing the WT and  $P_2$ ' K Nluc-mDH-FL reporters (Fig 2A). We reported previously that the higher MW mis-cleaved species of  $P_2$ ' K Hyp1-Nluc-mDH-FL was mostly concentrated in the TX-100 fraction (Fig 2B lane 7, double asterisk), suggesting that it could be membrane-associated <sup>42</sup>. In contrast, correctly processed  $P_2$ ' K Hyp1-Nluc-mDH-FL (Fig 2B lane 5, single asterisk) was readily extracted by hypotonic lysis (Fig 2B, Tris Sn). The cleaved forms of  $P_2$ ' K KAHRP- (Fig 2B, lanes 13-16, single asterisk) and STEVOR-Nluc-mDH-FL (Fig 2B, lanes 21-24, single asterisk) were largely found in the soluble fraction, although some was present in the other fractions as well (Fig 2B). In contrast to the Hyp1  $P_2$ ' K reporter, the mis-cleaved forms of KAHRP and STEVOR were evenly distributed in all fractions (Fig 2B, double asterisks). For all KAHRP and STEVOR reporters, the larger uncleaved forms were also found in the soluble fraction which was surprising considering these proteins would still retain their signal sequence transmembrane domains (Fig 2B, lanes 9, 13, 17 and 21, triple asterisks). We speculate that mis-cleaved  $P_2$ ' K reporters, particularly Hyp1, may remain trapped in the ER because they are less soluble although the mechanism behind this is not obvious since both size of the mis-cleaved  $P_2$ ' K proteins and previous mass spectrometric analysis suggests they lack their hydrophobic signal peptides<sup>42</sup>. When correctly processed however, the  $P_2$ ' K reporters are more soluble and traffic beyond the ER to at least the PV.

2.4. The length of the spacer region is essential for protein translocation across the parasitophorous vacuole membrane.

Thus far our data have indicated that the amino acid occupying the  $P_2$ ' residue is important for the correct cleavage of Hyp1 but less so for KAHRP and STEVOR indicating that other residues within and/or bordering the PEXEL motif may also be important for accurate cleavage. Earlier work has shown that truncation of the amino acid sequence (termed spacer region) that separates the PEXEL motif from a downstream folded protein such as GFP, influences export<sup>22,24</sup>. Interestingly, the N-terminal regions of PNEPs are functionally exchangeable with this spacer region of PEXEL proteins<sup>38</sup> and replacement of the spacer region with the N-terminal sequence of a PV-resident protein inhibits export<sup>4</sup>, suggesting that this region may comprise abona fide export signal. To investigate whether the spacer region has a role in binding to PTEX, the spacer region of the Hyp1, STEVOR, and KAHRP-Nluc-mDH-FL constructs were C-terminally truncated, from their original lengths of ~50aa, down to 13aa and 3aa preceding the folded domain of Nluc (Fig 3A and 3B). IFAs of trophozoite-infected RBCs expressing the truncation constructs showed reduced export with reducing spacer length in all three constructs. Quantification of the exported signal across the cell population further revealed that truncation from ~50aa to 13aa reduces export by ~10-20%, while export was strongly reduced in 3aa spacer constructs, showing a marked  $\sim 80\%$  reduction in fluorescence signal relative to the control (Figs 3A-C). This observation contrasts with the mutations of the  $P_2$ ' PEXEL motif alone, performed in the previous section (Fig 1) and other studies, which displayed variable export-blocking phenotypes with different PEXEL protein sequences <sup>1,14,33,36</sup>. Co-labelling of microscopy images with EXP2 (PV marker) and Pf ERC (ER marker) further indicated that the Hyp1-Nluc-mDH-FL with a 3aa spacer accumulated mainly in the PV with some signal in the ER overlapping with HSP101-HA (Fig 3A panels 3-7).

IFAs using anti-Nluc with STEVOR-Nluc-mDH-FL and KAHRP- Nluc-mDH-FL parasites showed that these reporters behaved similarly to the Hyp1-reporter (Fig 3B, panels 2, 3 and 5, 6) and also displayed the highest co-localisation with the HA-tagged translocon component HSP101, which we have shown resides within the ER in addition to the  $PV^{42,43}$ .

Truncation of the spacer did not appear to reduce processing of the PEXEL motif in this context as western blots of the Hyp1-Nluc-mDH-FL truncation constructs showed that each reporter protein migrated according to a predicted mass consistent with PMV-processed versions of the proteins (Fig 4A and B, lanes 2-4, Table S4). Taken together, these results show that the length of the spacer mutant is important for export, post-PEXEL processing.

2.5. Truncation of the spacer region reduces cargo binding with HSP101.

The observed co-localisation of all spacer constructs with the translocon components EXP2 and HSP101 in the PV was perplexing as the truncated proteins appeared to have processed PEXEL N-termini suggesting they were unable to engage with PTEX to be exported. We therefore sought to determine if the truncated PEXEL proteins could bind HSP101 by co-immunoprecipitation. To do this, the Hyp1 truncation constructs (51aa, 13aa, and 3aa Hyp1-Nluc-mDH-FL) were transfected into the HSP101-HA glmS parasite line  $^{42}$ . These parasites were grown to the ring stage and treated with WR99210 for 24 hours to stabilise the folding of the murine DHFR, that had previously been demonstrated to stall the cargo unfolding process within PTEX, thereby trapping and stabilising the cargo's interaction with PTEX<sup>16,43-45</sup>. The whole trophozoite infected RBCs were lysed and incubated with anti-HA-IgG agarose to immunoprecipitate the HA-tagged HSP101 from the sample. Western blot analysis of the eluates revealed a significantly reduced amount of Hyp1-Nluc-mDH-FL co-eluted with HSP101 with decreasing length of the spacer, with 60% and 90% reduction (n=3) observed in the 13aa and 3aa spacer, respectively, relative to the 51aa spacer Hyp1-Nluc-mDH-FL (Figs 4A, lanes 6-8 and 4C). Importantly, the experiment was performed in the presence of a stabilising ligand WR99210, suggesting that the 13aa and 3aa spacer truncation mutants did not proceed to the unfolding step within PTEX.

The same samples were also subjected to a reciprocal co-immunoprecipitation using anti-Nluc antibodies to pull-down the Hyp1-Nluc-mDH-FL truncation proteins and we consistently saw a gradual reduction in the amount of HSP101 co-eluted as the length of the spacer was shortened (Figs 4B lanes 6-8 and Fig 4D). We plotted the normalised % exported as observed by IFA and the amount of cargo (Hyp1-Nluc-mDH-FL) bound with HSP101 and we saw a remarkable correlation between these two variables, suggesting that the level of cargo binding with HSP101 determines its exportability (compare Fig 3C with Figs 4B and 4D). Taken together, these results demonstrated that the spacer region regulates cargo engagement with PTEX, particularly with HSP101 which is possibly the first point of contact cargo has with PTEX<sup>19,46</sup>.

2.6 The sequence requirements of the 13aa spacer for export are relatively unconstrained.

Since the 13aa Hyp1 spacer still permitted a reasonable amount of export compared to the full size 51aa spacer we decided to use this as basis for subsequent mutagenesis experiments due its small size. The spacers of PEXEL proteins appear to possess little obvious sequence information for putative PTEX recognition apart from appearing relatively unstructured with no conserved domains. To assess whether the minimal spacer may contain some cryptic trafficking information, we decided to remove it from the original 51aa spacer and fuse this new [?]NT13aa spacer to the Nluc-mDH-FL reporter and express this in parasites (Figs 5A). Western blots of the Hyp1 [?]NT13aa-Nluc-mDH-FL parasites indicated the reporter protein was of a size consistent with correct PMV cleavage (Fig S1). Microscopy of the [?]NT13aa reporter showed it was efficiently exported and quantification of the fluorescence intensity implies a higher degree of export of the [?]NT13aa reporter than the 13aa spacer (Fig 5B and C). This result indicated that the [?]NT13aa spacer for export essential for export of the Hyp1-NlucmDH-FL reporter.

Next, we investigated whether certain amino acids within the minimal spacer were necessary for export. As proteins chaperones tend to recognise misfolded proteins via their hydrophobic residues<sup>47</sup>, the four hydrophobic amino acids Y, L, I and V in the minimal Hyp1 spacer TEY KDTL QI KV EQ were mutated to determine if this reduced export. In the absence of predictive tools for HSP100 chaperones we used a binding predictor for the ER HSP70 chaperone BIP <sup>47</sup> that scans 7aa peptide windows and found that mutation of the hydrophobic amino acids to the polar amino acid serine, produced lower scores than for other amino acids (Figs 5A and S2). We therefore replaced the hydrophobic amino acids in the 13aa spacer with serines and transfected this Hyp1 13aa.Ser reporter into parasites where, by western blot, the Hyp1 13aa.Ser reporter was of a size consistent with correct PMV cleavage (Fig S1). Microscopic analysis of the Hyp1 13aa.Ser parasite-infected RBCs indicated that this mutant was exported less efficiently than the WT minimal 13aa spacer (Figs 5A-C and S1).

Next, the hydrophobic amino acids of the minimal 13aa spacer were mutated to glutamic acids (Hyp1 13aa.Glu) as these mutations were scored most poorly by the BIP binding predictor (Fig S2). Western blots of Hyp1 13aa.Ser-Nluc-mDH-FL parasites indicated PMV processing was correct and microscopic analysis of parasite-infected RBCs expressing the reporter indicated the Hyp1 13aa.Glu protein was exported significantly less well than the Hyp1 13aa parasites (Fig 5B and C). These results indicated that although

there is probably a great deal of flexibility in the amino acid sequences of spacers, some attributes are required such as the presence of hydrophobic amino acids that could be important for general chaperone binding.

As it can be difficult to visibly discern cargo trapped in the PV versus that inside the parasite we employed a recently-developed protein export assay based on the release and detection of Nluc bioluminescence from differentially lysed cellular compartments <sup>43,48</sup>. We saw that the Nluc bioluminescence signal exported into the RBC compartments closely followed the same trend as the spacer reporter parasites whose export was measured visibly by microscopy (compare Fig 5D to 5C). However, the signal for the Nluc bioluminescence export assay was however often 10-20% greater than the microscopy signal probably because bioluminescence is highly sensitive can detect lower protein levels dispersed in the large RBC compartment (Fig 5D). Additionally, microscopy does not measure RBC signal in front or behind the parasite which may underestimate the total signal. Interestingly, the bioluminescence signal of the reporters secreted into the PV compartment increased as export into the RBC compartment decreased indicating that as the spacer length shrunk, or its hydrophobic residues were mutated, the proteins were still able to traverse the parasite plasma membrane but were less efficiently translocated into the RBC by PTEX, leaving them to accumulate in the PV. It is also worth noting that the signal retained in the parasite was relatively constant indicating spacer length or mutations did not reduce secretion to the PV.

## Discussion

Here we sought to understand what element(s) of PEXEL proteins govern recruitment of PTEX to facilitate the export of proteins into the erythrocyte compartment. It was previously hypothesised that the last conserved residue of the PEXEL, which remains on the mature protein, is responsible for interacting with PTEX  $^{1,2,14}$ . Our data partly support this in that mutation of  $P_2$ ' residue to K can lead to increased trapping of some PEXEL proteins in the PV, most notably for the KAHRP and STEVOR reporters. The  $P_2$ ' K mutation can also greatly reduce the efficiency of PMV cleavage in the case of Hyp1 but not so for KAHRP and STEVOR and as such, increase protein trapping in the ER as shown for the Hyp1 reporter. Apart from the efficiency of cleavage of the PEXEL motif for governing PTEX's interaction with the cargo proteins, the spacer region downstream of the PEXEL that separates the PEXEL from the folded functional Nluc region of the reporter protein is also important for trafficking. As the length of spacer increased, so did the reporter's binding to PTEX and the efficiency with which it was exported into the RBC compartment. A small spacer length of 13aa was found to still result in efficient export with hydrophobic residues in the spacer important for export efficiency. Reporters with short or mutated spacers appeared to be still efficiently cleaved by PMV but interacted weakly with PTEX leaving them trapped in the PV compartment and unable to be exported. One caveat of this study is that our reporter had a folded Nluc domain a defined distance from the upstream PEXEL cleavage site, and we are unsure if similar rules will generally apply to native PEXEL proteins many of whose structures and functions are unknown.

Other studies have shown that the amino acid in the last position of export motifs is important for export. For example, reporter proteins containing oomycete effector motif RxLR  $^{30,38}$ , which lack the P<sub>2</sub>' residue, were not cleaved by plasmepsin V and failed to promote export into the host cell  $^{30,38}$ . Of all possible mutations of P<sub>2</sub>', only charge reversal P<sub>2</sub>' mutations, either to arginine or lysine, strongly inhibit plasmepsin V cleavage  $^{36}$  whilst a single alanine mutation was observed to cause a variable export phenotype<sup>4,36</sup>. This was evident from the small reduction of cleavage observed in the *in vitro* cleavage assay with P<sub>2</sub>' A mutation compared to the K mutation. Despite this, the strong inhibition *in vitro* cleavage of STEVOR P<sub>2</sub>' K did not translate to what was observed in parasites as the reporter protein was efficiently cleaved. It seems likely that suboptimal amino acid variants at PEXEL P<sub>2</sub>' are still permissive for PMV cleaving in the cellular context, presumably because of the fast kinetics of the proteolytic reaction *in vivo* or due to concentrated substrate in the ER that favours cleavage even under less cleavable substrate. It is also possible that Pf PMV in parasites is more active against the P<sub>2</sub>' mutants than the Pv PMV used in the peptide cleavage assay<sup>49,50</sup>.

Having previously observed that the ER-trapped Hyp1  $P_2$ ' K mutant, which was mis-cleaved by an unknown protease upstream of the PEXEL, became quite insoluble and remained in the ER<sup>42</sup>, we explored how the

solubility of cleaved PEXEL reporters can influence their export. In contrast, to the mis-cleaved Hyp1 P<sub>2</sub>' K reporter, the solubility of the mis-cleaved KAHRP and STEVOR P<sub>2</sub>' K reporter proteins did not appear to decrease, and it is unclear whether the mis-cleaved species accumulated in the ER or were trafficked to the PV like the correctly cleaved P<sub>2</sub>' K species. We have shown previously that the ER-trapped Hyp1 P<sub>2</sub>' K reporter can still bind to ER-resident HSP101 despite its insolubility <sup>42</sup>. The fact that the KAHRP and STEVOR P<sub>2</sub>' K mutants were efficiently cleaved and trafficked to the PV but were not strongly exported suggests the P<sub>2</sub>' residue might not be important for the initial binding to HSP101 but rather for commitment to stronger downstream interactions with the whole of PTEX required for cargo unfolding and export.

Our results also argue that PTEX, more specifically HSP101, recognises a wider region in the cargo than the PEXEL motif and that this is important for export. Truncation of the spacer region in at least three different reporter constructs used in this study consistently blocked export without apparently affecting PEXEL processing. Microscopic analysis of the 3aa spacer Nluc-mDH-FL constructs clearly showed that the protein is trapped within the PV area where PTEX is located, consistent with previous reports <sup>4,24</sup>. Despite this co-localisation, we have shown that spacer mutants that bind less strongly to HSP101 are exported less efficiently.

While the molecular mechanism of PTEX cargo recognition remains to be elucidated, our data has shed some light into how this process may occur. We found that while a 13aa spacer was sufficient to facilitate modest export, a longer 51aa spacer promoted stronger cargo binding to HSP101, suggesting that perhaps an increase the length of the unstructured N-terminal polypeptide increases the likelihood of the cargo will stably binding to HSP101. Clp/HSP100 chaperones generally require a recognition signal of at least 10-20 broadly diverse amino acids to initiate the polypeptide unfolding and translocation  $^{51,52}$ . Our data is consistent with this model and suggest that HSP101 also requires a large region within the N-terminal portion of the cargo protein to initiate cargo translocation. This could explain why single  $P_2$ ' K point mutation in the mature PEXEL motif is not enough to prevent the interaction altogether and why the N-termini of PNEPs, despite the latter lacking a mature PEXEL motif, can still act as an export signal <sup>38</sup>. Cargo recognition in AAA+ ATPases, particularly the family of Clp/HSP100 chaperones, begins with the pre-unfolding step that is initiated by a low-affinity probabilistic binding of the chaperone to a loosely folded or aggregated region of a protein, followed by a commitment step where the ATPase binds more stably to the cargo protein before unfolding and threading the protein through the chaperone's central cavity<sup>53</sup>. It has been shown for the AAA+ ATPase ClpXP, that the length of the cargo polypeptide bound to the inner cavity of the ClpXP affects the commitment step, such that longer polypeptides seem to promote more successful commitment and subsequent unfolding<sup>54,55</sup>. Clp/HSP100 chaperones, particularly ClpB and HSP104, are thought to have a similar pre-unfolding step $^{52,53,56}$ . We therefore propose a model whereby a longer spacer region may increase accessible areas for the initial probabilistic binding step, or stabilise association of exported proteins to HSP101, subsequently leading to less frequent dissociation from the unfoldase (Fig 6A). Consistently, Hyp1-Nluc-mDH-FL reporter with a short 3aa spacer region exhibited low-level affinity to HSP101 that greatly reduced export, suggesting that the cargo may have initially associated with HSP101 but later dissociated from the unfoldase because there was insufficient net affinity to proceed to the commitment step.

We also explored which amino acids of the spacer region could promote or inhibit HSP101 binding. The subunits of the second nucleotide binding domain (NBD2) of HSP101 contain conserved tyrosine residues which are thought to bind the unfolded cargo protein via hydrophobic interactions to help ratchet the cargo through HSP101<sup>19</sup>. As the HSP101 subunits undergo allosteric changes powered by ATP hydrolysis, the tyrosine residues may interact with the hydrophobic residues in the spacer region, we mutated the four hydrophobic residues in the 13aa Hyp1 spacer and this was shown to reduce export. Export reduction was particularly strong for the mutation of hydrophobics to charged (E) residues compared to polar residues (S). The deletion of the 13aa spacer region from the 51aa acid spacer still resulted in strong export as the next 13aa downstream from the first 13aa still contained four hydrophobic residues. One possible reason why the minimal 13aa spacer preceding the globular Nluc region was exported much better than the 3aa spacer was that the longer spacer could project further into HSP101's central cavity, down into NBD2 where it could

engage the cargo binding tyrosine residues (Fig 6B).

In conclusion, our data suggests dual functions for the  $P_2$ ' position of PEXEL proteins. The first is that it forms part of the PMV recognition sequence for cleavage. In some proteins such as Hyp1, mutation of  $P_2$ ' to K greatly reduces PMV binding and/or successful proteolytic activity but in other PEXEL proteins such as STEVOR and KAHRP  $P_2$ ' is not as critical for PMV cleavage. This indicates that other amino acids within and bordering the PEXEL probably have a bearing on how dependent PMV activity is on the  $P_2$ ' residue. A lack of PMV cleavage probably results in ER retention as evidenced by the substantial retention for poorly cleaved Hyp1 versus efficiently cleaved STEVOR and KAHRP. Successful cleavage of the PEXEL  $P_2$ ' K mutant appears to permit trafficking to the PV as STEVOR and KAHRP more efficiently reach the PV than poorly cleaved Hyp1. Once in the PV, however, the STEVOR and KAHRP  $P_2$ ' K reporters were not efficiently exported into the RBC compartment. This could be due to the reporters not being efficiently recognised by PTEX or because the reporters were in a PTEX-free sub-compartment of the PV. The latter, however, is unlikely, because recent split-GFP experiments indicate that PV-resident protein have full access to PTEX <sup>57</sup>. It is possible therefore that the charge reversal  $P_2$ ' mutant somehow binds to PTEX less efficiently leading to less engagement, unfolding and eventual export into the RBC compartment.

## Materials and Methods

## 4.1. Culture of P. falciparum transfectants

Asexual blood-stage *Plasmodium falciparum* (3D7 or CS2 background) was cultured according to the established protocol <sup>58</sup>. Cultures were routinely maintained in complete RPMI media containing RPMI-1640 base media supplemented with 2.5 mM HEPES, 367  $\mu$ M hypoxanthine, 31.25  $\mu$ g/mL Gentamicin, 25 mM NaHCO<sub>3</sub>, and 0.5% (w/v) Albumax II (Invitrogen). Prior to transfection, 100  $\mu$ g of plasmid DNA was resuspended in TE and cytomix (25 mM HEPES, 120 mM KCl, 0.15 mM CaCl<sub>2</sub>, 2 mM EGTA, 5 mM MgCl<sub>2</sub>, 10 mM K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub>pH 7.6) and mixed with packed RBCs as per <sup>59</sup>. After electroporation using Gene Pulser XCell System (BioRad), the uninfected RBCs were then mixed with 20  $\mu$ L of HSP101-HA*glmS*trophozoite-stage parasites <sup>42</sup> and allowed to invade the transfected RBCs for 2 cell cycles before starting a selection with 2.5  $\mu$ g/mL blasticidin S.

## 4.2 Generation of PEXEL $P_2$ ' mutants and spacer truncation constructs

The expression of Hyp1/ STEVOR/ KAHRP-Nluc-mDH-FL reporters was driven by a bidirectional *Plasmodium berghei* EF1a promoter that also controlled expression of the blasticidin deaminase drug resistance cassette. The plasmid pEF-Hyp1-Nluc-mDH-FL was derived from plasmid pEF-Hyp1-Nluc-DH-APEX <sup>43</sup>. The Hyp1 component of this plasmid contained the first 113aa of Hyp1 (PF3D7\_0113300), including the RLLTE PEXEL motif <sup>43</sup>. A synthetic murine dihydrofolate reductase (mDH) gene fragment with C-terminal 3x FLAG epitopes (Bioneer Pacific) was ligated into the Nluc-DH-APEX plasmid using *SpeI* and *MluI* enzymes to remove the previous mDH-APEX gene cassette.

Generation of the  $P_2$ ' lysine (lys/K) and alanine (ala/A) mutation of the Hyp1-Nluc-mDH-FL was performed as follows: Hyp1 region was first amplified as two overlapping PCR fragments with the primer pairs 1 & 2 and 3 & 4 (Table S4). The overlap region between these two PCR products contained the mutations indicated above. PCR fragments were then sewn together with primer pair 1 & 4 via overlapping PCR, ligated into the pJET1.2/blunt plasmid (ThermoFisher Scientific). Mutagenesis was confirmed via standard Sanger sequencing (service provided by Monash Micromon Genomics) of the isolated plasmids. The mutant  $P_2$ ' K/A Hyp1 fragments were released from pJET1.2/blunt using *XhoI* and *NcoI* and ligated into pEF-Hyp1-Nluc-mDH-FL to replace the wildtype Hyp1 fragment.

Generation of the Hyp1-Nluc-mDH-FL plasmids with a truncated spacer region between the PEXEL motif and the start of the Nluc gene was carried out as follows: PCRs were performed with primer 1 paired with primer 7, 8 or 9 to produce spacers of 3aa, 13aa and 51aa, respectively (Table S4). Note that primers 7, 8 and 9 produced Hyp1 PCR products that contained one additional amino acid between the last Hyp1 residue and the start Met of Nluc. The PCR products were ligated into pJET1.2/blunt and screened as above. To express wildtype STEVOR (PF3D7\_0200400) and KARHP (PF3D7\_0202000) in the Nluc-mDH-FL reporter plasmid, synthetic leader sequences encoding the first 99aa of STEVOR and 105aa of KAHRP containing *XhoI* and *NcoI* restriction sites were first obtained as string oligos from GeneART (ThremoFisher Scientific). P<sub>2</sub>' lysine mutations for both STEVOR and KAHRP leaders were also synthesised as above. The synthetic DNA sequences were ligated into pJET1.2/blunt and validated by sequencing prior to transfer via *XhoI* and *NcoI* sites into the pEF-Hyp1-Nluc-mDH-FL plasmid replacing the Hyp1 sequence.

To produce pEF-STEVOR-Nluc-mDH-FL reporters with truncated spacers, PCRs were performed with primers 10 & 11 and primers 10 & 12 to produce spacer fragments of 3aa, and 13aa respectively, using pEF-STEVOR-Nluc-mDH-FL as a template (Table S5). PCR products were ligated into pJET1.2/blunt and ligated into pEF-Nluc-mDH-FL via*XhoI* and *NcoI* cloning sites to produce gene fusions with only one amino acid between STEVOR/KAHRP leader sequence and Nluc. To produce the pEF-KAHRP-Nluc-mDH-FL with 3aa and 13aa spacers, PCRs were performed with primers 10 & 13 and primers 10 & 14 (Table S4). The PCR products were ligated into the Nluc-mDH-FL reporter as per the STEVOR reporters.

## 4.3 Indirect immunofluorescence analysis

IFA was performed essentially according to  $^{60}$ , where infected RBCs (iRBCs) were settled onto a poly-L-lysine (Sigma, P8920) coated coverslip and fixed with 4% paraformaldehyde/0.0075% glutaraldehyde. Following fixation, the cells were permeabilised with 0.1 M glycine/0.1 % Triton X-100 for 12 minutes at room temperature (RT). Coverslips were probed with rabbit anti-Nluc IgG (12.5 µg/mL), mouse anti-EXP2 (10 µg/mL), rabbit anti-ERC (1:1000), mouse anti-FLAG M2 (Sigma, 10 µg/mL), mouse anti-HA (Sigma clone HA-7; 1:500). After washing, goat anti-rabbit Alexa Fluor 594 nm and goat anti-mouse Alexa Fluor 488 nm (1:2000) secondary antibodies (Invitrogen) were applied for 1 h at RT. Fixed material was mounted in VECTASHIELD with DAPI and imaged on Zeiss Cell Axio-Observer (Carl Zeiss). Image acquisition was performed with Zen Blue imaging software.

## 4.4 Quantification of Imaging Data

Image analysis was carried out according  $to^{42}$ .

# 4.5 Chemical Cross-Linking of P. falciparum Culture and Immunoprecipitation

RBCs infected with the HSP101-HA*glmS* /PEXEL-Nluc-mDH-FL parasites were enriched through magnetic purification and their proteins were solubilised in 20x pellet volume of IP lysis buffer (1% Triton X-100, 0.1% SDS, 150 mM NaCl, 10 mM Tris-HCl pH 7.4) supplemented with cOmpleteTM Protease Inhibitor Cocktail (Roche) and subjected to 2 freeze and thaw cycles. The lysate was clarified by centrifugation and incubated overnight at 4°C with anti-HA monoclonal agarose (Sigma-Aldrich). Following incubation, the agarose beads were washed 5x with 1 mL IP lysis buffer and the proteins were eluted with 50  $\mu$ L 2x NRSB (100 mM Tris-HCl pH 6.8, 4 mM EDTA, 4% SDS, 0.01% bromophenol blue, 20% (v/v) glycerol). For Nluc IgG immunoprecipitations, the parasite lysates were incubated overnight at 4°C with 10  $\mu$ g IgG-purified anti-Nluc antibody. Following incubation, protein A-Sepharose 4B (Invitrogen) was added to bind the immune complexes and samples were incubated for an additional 1 h at RT. The beads were washed, and proteins eluted as above. In all cases, both input and elution fractions were reduced in 200 mM DTT at 70°C for 5-10 mins prior to SDS-PAGE separation, western blot, and immunoblotting.

## 4.6 Western Blotting

Proteins transferred from the gels to a nitrocellulose membrane using iBlot<sup>®</sup> Blotting System (Invitrogen). The blots were blocked in 1% casein in PBS and probed with primary antibody (Table 2.3) diluted in the blocking buffer overnight at 4°C. The blots were washed and probed with goat anti-mouse or goat anti-rabbit fluorescence-based (Alexa Fluor 700 and 800 nm, Invitrogen) antibodies diluted 1/10,000 in blocking buffer for 1 h at room temperature followed by 3 washes with 1x PBS. The fluorescent secondary antibodies were visualized with a LI-COR Odyssey FC imaging system. Densitometry analysis was performed with Image Studio v. 1.0.

## 4.7 Biochemical PEXEL Cleavage Assays

The cleavage assay was performed as described in Hodder et al. (Hodder, Sleebs et al. 2015). 2 nM of *P. vivax* PMV in buffer (25 mM Tris-HCl pH 6.4 and 25 mM MES pH 6.4) was incubated with 5  $\mu$ M FRET peptide substrates representing WT and mutant KAHRP, STEVOR and Hyp1 sequences (Supplementary Table 5) in a total volume of 20  $\mu$ L. Samples were incubated at 20°C for 20 h and measurement was carried out using Envision plate (PerkinElmer) reader (ex. 340 nm; em. 490 nm). Biochemical *P. vivax* PMV inhibitory assays (20  $\mu$ L total volume) were performed using 2 nM *P. vivax* PMV in buffer (25 mM Tris-HCl and 25 mM MES, pH 6.4) with 5  $\mu$ M FRET KAHRP\_WT fluorogenic peptide. Assay reactions were incubated at 37°C for 2 h in the presence of peptides (10 points dose-response, 1 in 2 dilution series starting at 100 nM) representing STEVOR and Hyp1 sequences (Table 2.2). Fluorescence was measured with an Envision plate (PerkinElmer) reader (ex. 340 nm; em. 490 nm). To determine the level of PMV inhibition, "nonlinear regression four-parameter to fit analysis" using Domatics software (version 5.3.1612.8630) was performed.

## 4.8 Protein Solubility Analysis

Protein solubility profiling was performed according to Gruring et al.<sup>38</sup>. 5  $\mu$ L of parasite pellet (obtained through magnetic separation or saponin lysis) was resuspended in 100  $\mu$ L hypotonic buffer (5 mM Tris-HCl pH 8.0) and subjected to 1 cycle of freeze and thawing. The soluble fraction was separated from the pellet by centrifugation at 16,000 x g for 5 minutes at 4°C. The pellet was then incubated sequentially with 100  $\mu$ L 0.1M Na<sub>2</sub>CO<sub>3</sub> pH 11.5 and 1% Triton-X in H<sub>2</sub>O for 30 minutes at 4°C. Soluble fractions from each incubation were transferred into a new tube. Soluble fractions were mixed with 20  $\mu$ L 6x NRSB (300 mM Tris-HCl pH 6.8, 12 mM EDTA, 12% SDS, and 0.03% w/v bromophenol blue, 60% v/v glycerol) and the pellet was resuspended in 120  $\mu$ L 1xNRSB. All samples were kept at -20°C until used.

## 4.9 Statistical Analysis

Numerical data was mainly visualised using Microsoft Excel for Mac (v.15.37) and GraphPad Prism version 9.1.0. Statistical analysis was performed using GraphPad Prism version 9.1.0 unless otherwise stated.

## 4.10 Nanoluciferase protein export assays

The nanoluciferase reporter export assays were performed as per<sup>48</sup>.

## Author Contributions

Mikha Gabriela, Claudia B.G. Barnes, Dickson Leong, Brad E. Sleebs, Molly P. Schneider, Dene R. Littler and Paul R. Gilson performed experimental work. Mikha Gabriela, Claudia G. B. Barnes and Paul R. Gilson wrote and edited manuscript. Brendan S. Crabb and Paul R. Gilson and Tania F. de Koning provided funding and student supervision.

### Acknowledgments

We thank the Lifeblood Biological Resources for providing blood and the WEHI Antibody Facility for the EXP2 and Nluc antibodies. We thank Matthew Dixon and Leann Tilley for kindly providing Pf ERC antibodies.

#### **Funding information**

NHMRC Program Grant and NHMRC Investigator Grant. Grant/ Award numbers: 1092789 and 119780521 and Australian Research Council (ARC) grant DP160102582. This work was supported by funding from the Victorian Operational Infrastructure Support Program received by the Burnet Institute. M.G. was recipient of a Deakin University Postgraduate Research Scholarship.

## Conflict of interest

The authors declare no conflicts of interest.

## Data availability statement

All relevant data are within the manuscript and its Supporting Information files.

## References

1. Marti M, Good RT, Rug M, Knuepfer E, Cowman AF. Targeting Malaria Virulence and Remodeling Proteins to the Host Erythrocyte. *Science*. 2004;306(5703):1930-1933.

2. Hiller NL, Bhattacharjee S, van Ooij C, et al. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science*. 2004;306(5703):1934-1937.

3. Sargeant T, Marti M, Caler E, et al. Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biology*. 2006;7(2):R12.

4. Boddey JA, Carvalho TG, Hodder AN, et al. Role of plasmepsin V in export of diverse protein families from the *Plasmodium falciparum* exportome. *Traffic.* 2013;14(5):532-550.

5. Jonsdottir TK, Gabriela M, Crabb BS, T FdK-W, Gilson PR. Defining the Essential Exportome of the Malaria Parasite. *Trends Parasitol*.2021;37(7):664-675.

6. Lingelbach K, Joiner KA. The parasitophorous vacuole membrane surrounding *Plasmodium* and *Toxoplasma*: an unusual compartment in infected cells. *Journal of Cell Science*.1998;111(11):1467-1475.

7. Ward GE, Miller LH, Dvorak JA. The origin of parasitophorous vacuole membrane lipids in malariainfected erythrocytes. *Journal of Cell Science*. 1993;106(1):237-248.

8. de Koning-Ward TF, Gilson PR, Boddey JA, et al. A newly discovered protein export machine in malaria parasites. *Nature*.2009;459(7249):945-949.

9. Elsworth B, Matthews K, Nie CQ, et al. PTEX is an essential nexus for protein export in malaria parasites. *Nature*.2014;511(7511):587-591.

10. Beck JR, Muralidharan V, Oksman A, Goldberg DE. PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes. *Nature*. 2014;511(7511):592-595.

11. Marapana DS, Dagley LF, Sandow JJ, et al. Plasmepsin V cleaves malaria effector proteins in a distinct endoplasmic reticulum translocation interactome for export to the erythrocyte. *Nat Microbiol.* 2018;3(9):1010-1022.

12. Boddey JA, Hodder AN, Gunther S, et al. An aspartyl protease directs malaria effector proteins to the host cell. *Nature*.2010;463(7281):627-631.

13. Russo I, Babbitt S, Muralidharan V, Butler T, Oksman A, Goldberg DE. Plasmepsin V licenses *Plasmodium* proteins for export into the host erythrocyte. *Nature*. 2010;463(7281):632-636.

14. Boddey JA, Moritz RL, Simpson RJ, Cowman AF. Role of the *Plasmodium* Export Element in Trafficking Parasite Proteins to the Infected Erythrocyte. *Traffic.* 2009;10(3):285-299.

15. Przyborski JM, Nyboer B, Lanzer M. Ticket to ride: export of proteins to the *Plasmodium falciparum* -infected erythrocyte. *Mol Microbiol.* 2016;101(1):1-11.

16. Gehde N, Hinrichs C, Montilla I, Charpian S, Lingelbach K, Przyborski JM. Protein unfolding is an essential requirement for transport across the parasitophorous vacuolar membrane of *Plasmodium falciparum*. *Molecular Microbiology*.2009;71(3):613-628.

17. Ansorge I, Benting J, Bhakdi S, Lingelbach K. Protein sorting in *Plasmodium falciparum* -infected red blood cells permeabilized with the pore-forming protein streptolysin O. . *Biochemical Journal*. 1996;315(1):307-314.

18. Matthews KM, Kalanon M, de Koning-Ward TF. Uncoupling the Threading and Unfoldase Actions of *Plasmodium* HSP101 Reveals Differences in Export between Soluble and Insoluble Proteins. *mBio.* 2019;10(3).

19. Ho C-M, Beck JR, Lai M, et al. Malaria parasite translocon structure and mechanism of effector export. *Nature*. 2018;561:70-75.

20. Beck JR, Ho CM. Transport mechanisms at the malaria parasite-host cell interface. *PLoS Pathog.* 2021;17(4):e1009394.

21. Egea PF. Crossing the Vacuolar Rubicon: Structural Insights into Effector Protein Trafficking in Apicomplexan Parasites. *Microorganisms*. 2020;8(6).

22. Wickham ME, Rug M, Ralph SA, et al. Trafficking and assembly of the cytoadherence complex in *Plasmodium falciparum* -infected human erythrocytes. *EMBO Journal*. 2001;20(20):5636-5649.

23. Lopez-Estran C, Bhattacharjee S, Harrison T, Haldar K. Cooperative domains define a unique host cell-targeting signal in *Plasmodium falciparum* -infected erythrocytes. *Proceedings of the National Academy of Sciences.* 2003;100(21).

24. Knuepfer E, Rug M, Cowman AF. Function of the *Plasmodium* export element can be blocked by green fluorescent protein. *Molecular and Biochemical Parasitology*. 2005;142(2):258-262.

25. Haase S, Herrmann S, Grüring C, et al. Sequence requirements for the export of the *Plasmodium falciparum* Maurer's clefts protein REX2.*Molecular Microbiology*. 2009;71(4):1003-1017.

26. Melcher M, Muhle RA, Henrich PP, et al. Identification of a role for the PfEMP1 semi-conserved head structure in protein trafficking to the surface of *Plasmodium falciparum* infected red blood cells. *Cell Microbiol.* 2010;12(10):1446-1462.

27. Marti M, Baum J, Rug M, Tilley L, Cowman AF. Signal-mediated export of proteins from the malaria parasite to the host erythrocyte. *The Journal of Cell Biology*. 2005;171(4):587-592.

28. Dixon MWA, Hawthorne PL, Spielmann T, Anderson KL, Trenholme KR, Gardiner DL. Targeting of the Ring Exported Protein 1 to the Maurer's Clefts is Mediated by a Two-Phase Process. *Traffic*.2008;9(8):1316-1326.

29. van Ooij C, Tamez P, Bhattacharjee S, et al. The Malaria Secretome: From Algorithms to Essential Function in Blood Stage Infection. *PLoS Pathogens*. 2008;4(6):e1000084.

30. Boddey JA, O'Neill MT, Lopaticki S, et al. Export of malaria proteins requires co-translational processing of the PEXEL motif independent of phosphatidylinositol-3-phosphate binding. *Nature Communications*. 2016;7:10470.

31. Spielmann T, Gilberger T-W. Protein export in malaria parasites: do multiple export motifs add up to multiple export pathways? *Trends in Parasitology*. 2010;26(1):6-10.

32. Spielmann T, Gilberger TW. Critical Steps in Protein Export of *Plasmodium falciparum* Blood Stages. *Trends Parasitol*.2015;31(10):514-525.

33. Przyborski JM, Miller SK, Pfahler JM, et al. Trafficking of STEVOR to the Maurer's clefts in *Plasmodium falciparum* -infected erythrocytes. *EMBO J.* 2005;24(13):2306-2317.

34. Chang HH, Falick AM, Carlton PM, Sedat JW, DeRisi JL, Marletta MA. N-terminal processing of proteins exported by malaria parasites. *Molecular and Biochemical Parasitology*. 2008;160(2):107-115.

35. Crabb BS, de Koning-Ward TF, Gilson PR. Protein export in *Plasmodium* parasites: from the endoplasmic reticulum to the vacuolar export machine. *International Journal For Parasitology*.2010;40(5):509-513.

36. Tarr SJ, Cryar A, Thalassinos K, Haldar K, Osborne AR. The C-terminal portion of the cleaved HT motif is necessary and sufficient to mediate export of proteins from the malaria parasite into its host cell. *Molecular Microbiology*. 2013;87(4):835-850.

37. Haase S, Hanssen E, Matthews K, Kalanon M, de Koning-Ward TF. The exported protein PbCP1 localises to cleft-like structures in the rodent malaria parasite *Plasmodium berghei*. *PloS One*.2013;8(4):e61482.

38. Grüring C, Heiber A, Kruse F, et al. Uncovering Common Principles in Protein Export of Malaria Parasites. *Cell Host & Microbe*.2012;12(5):717-729.

39. Tarr SJ, Cryar A, Thalassinos K, Haldar K, Osborne AR. The C-terminal portion of the cleaved HT motif is necessary and sufficient to mediate export of proteins from the malaria parasite into its host cell. *Molecular Microbiology*. 2012:n/a-n/a.

40. Deil S, Lanzer M. New insights into PEXEL-mediated protein export in Plasmodium falciparum: the role of N-terminal acetylation : Department für Infektiologie Heidelberg, Universität Heidelberg; 2014.

41. Hodder AN, Sleebs BE, Czabotar PE, et al. Structural basis for plasmepsin V inhibition that blocks export of malaria proteins to human erythrocytes. *Nat Struct Mol Biol.* 2015;22(8):590-596.

42. Gabriela M, Matthews KM, Boshoven C, et al. A revised mechanism for how *Plasmodium falciparum* recruits and exports proteins into its erythrocytic host cell. *PLoS Pathog.* 2022;18(2):e1009977.

43. Charnaud SC, Jonsdottir TK, Sanders PR, et al. Spatial organization of protein export in malaria parasite blood stages. *Traffic*.2018;19(8):605-623.

44. Riglar DT, Rogers KL, Hanssen E, et al. Spatial association with PTEX complexes defines regions for effector export into *Plasmodium falciparum* -infected erythrocytes. *Nature Comm.* 2013;4:1415.

45. Mesen-Ramirez P, Reinsch F, Blancke Soares A, et al. Stable Translocation Intermediates Jam Global Protein Export in *Plasmodium falciparum* Parasites and Link the PTEX Component EXP2 with Translocation Activity. *PLoS Pathog.* 2016;12(5):e1005618.

46. Bullen HE, Charnaud SC, Kalanon M, et al. Biosynthesis, Localization, and Macromolecular Arrangement of the *Plasmodiumfalciparum* Translocon of Exported Proteins (PTEX). *J Biol Chem.* 2012;287(11):7871-7884.

47. Schneider M, Rosam M, Glaser M, et al. BiPPred: Combined sequence- and structure-based prediction of peptide binding to the Hsp70 chaperone BiP. *Proteins*. 2016;84(10):1390-1407.

48. Looker O, Dans MG, Bullen HE, Sleebs BE, Crabb BS, Gilson PR. The Medicines for Malaria Venture Malaria Box contains inhibitors of protein secretion in *Plasmodium falciparum* blood stage parasites. *Traffic.* 2022;23(9):442-461.

49. Sleebs BE, Lopaticki S, Marapana DS, et al. Inhibition of Plasmepsin V Activity Demonstrates Its Essential Role in Protein Export, PfEMP1 Display, and Survival of Malaria Parasites. *PLoS Biolo*gy.2014;12(7):e1001897.

50. Bedi RK, Patel C, Mishra V, Xiao H, Yada RY, Bhaumik P. Understanding the structural basis of substrate recognition by *Plasmodium falciparum* plasmepsin V to aid in the design of potent inhibitors. *Sci Rep.* 2016;6:31420.

51. Schlieker C, Weibezahn J, Patzelt H, et al. Substrate recognition by the AAA+ chaperone ClpB. *Nat Struct Mol Biol.* 2004;11(7):607-615.

52. Lum R, Niggemann M, Glover JR. Peptide and protein binding in the axial channel of Hsp104. Insights into the mechanism of protein unfolding. *J Biol Chem.* 2008;283(44):30139-30150.

53. Zolkiewski M, Zhang T, Nagy M. Aggregate reactivation mediated by the Hsp100 chaperones. Arch Biochem Biophys. 2012;520(1):1-6.

54. Kenniston JA, Baker TA, Sauer RT. Partitioning between unfolding and release of native domains during ClpXP degradation determines substrate selectivity and partial processing. *Proc Natl Acad Sci U S* A.2005;102(5):1390-1395.

55. Cordova JC, Olivares AO, Shin Y, et al. Stochastic but highly coordinated protein unfolding and translocation by the ClpXP proteolytic machine. *Cell.* 2014;158(3):647-658.

56. Desantis ME, Shorter J. The elusive middle domain of Hsp104 and ClpB: location and function. *Biochim Biophys Acta*.2012;1823(1):29-39.

57. Garten M, Nasamu AS, Niles JC, Zimmerberg J, Goldberg DE, Beck JR. EXP2 is a nutrient-permeable channel in the vacuolar membrane of *Plasmodium* and is essential for protein export via PTEX.*Nature Microbiology*. 2018;3:1090-1098.

58. Trager W, Jensen J. Human malaria parasites in continuous culture. Science. 1976;193(4254):673-675.

59. Hasenkamp S, Russell KT, Horrocks P. Comparison of the absolute and relative efficiencies of electroporation-based transfection protocols for *Plasmodium falciparum*. *Malar J.* 2012;11:210.

60. Tonkin CJ, van Dooren GG, Spurck TP, et al. Localization of organellar proteins in *Plasmodium falciparum* using a novel set of transfection vectors and a new immunofluorescence fixation method. *Molecular and Biochemical Parasitology.* 2004;137(1):13-21.

## Figures

Figure 1.  $P_2$ ' Ala and Lys mutations reduce with the proteolytic processing of the PEXEL **motif.** (A)5  $\mu$ M fluorogenic peptides were incubated with 2 nM recombinant Pv PMV and assayed at 20°C. Fluorescence data was normalised to the WT substrates (n=3). PEXEL motifs are indicated below the x-axis with mutated resides indicated in red. Statistical significance was determined using ordinary oneway ANOVA. \*\*\*\*, p-value<0.0001, \*\*\*, p-value<0.001). (B)Representative IFA images (n=3 independent replicates) of HSP101-HAgImS parasites expressing WT and P<sub>2</sub>' K of Hyp1, STEVOR, & KAHRP-NlucmDH-FL reporter proteins. Parasite cells were probed with anti-EXP2 and anti-HA antibodies to visualise the PTEX components within the cell. The Nluc-mDH-FL proteins were localised using anti-Nluc antibody. The red bar in the schematic picture of the construct indicates the location of the PEXEL motif. Scale bars, 5µm. DIC, Differential Interference Contrast. DAPI (4',6-diamidino-2-Phenylindole; Blue) was used to stain parasite nuclei. (C) Quantification of the fluorescence signal of STEVOR and KAHRP-Nluc-mDH-FL constructs within the parasite cells from A. Quantification was performed using at least 20 cells expressing each construct. Box and whisker plot represents 25th-75th and 5th-95th percentiles, respectively. Statistical significance was determined using ordinary one-way ANOVA. \*\*\*\*, p-value<0.0001). (D) Representative Western blot (n=3) of lysates made from mid-stage trophozoites expressing Nluc-mDH-FL constructs probed with anti-FLAG and anti-Nluc IgGs. The identity of various protein species was based on their observed sizes. \*\*\*, Full length proteins. , Mis-cleaved P<sub>2</sub>' K Hyp1-Nluc-mDH-FL (58.1  $\pm$  1.2 kDa, n=10). \*\*, Mis-cleaved P<sub>2</sub>' K KAHRP (51.4  $\pm$  1.2 kDa, n=3) and STEVOR (53.8  $\pm$  1.7 kDa, n=3)-Nluc-mDH-FL. \*. correctly PEXEL-cleaved species of the Nluc-mDH-FL proteins. < is a cross-reactive protein. Hyp1-NlucmDH-FL blots are shown separate to KAHRP and STEVOR blots as they were run on different gels. The Neg lane contains protein from HSP101-HAglmS parasites not transfected with the NL-mDH-FL reporter and the blots were probed with anti-PTEX150 antibody as a loading control.

Figure 2. Solubility profile of Nluc-mDH-FL proteins.(A) Predicted full-length and correctly processed sizes of Hyp1, KAHRP and STEVOR-Nluc-mDH-FL reporter proteins. (B)Western blot analysis of infected RBCs (mid-stage trophozoites) sequentially extracted with 5 mM Tris-Cl pH 8.0 (Tris Sn), 0.1M Na<sub>2</sub>CO<sub>3</sub> pH 11.3 (Carb Sn), and 1% Triton X-100 buffer (TX-100 Sn) to partition proteins based on their association with cellular membranes. Insoluble fraction represents the final pellet obtained after Triton X-100 extraction. GAPDH, HSP101, and EXP2 were used as a control for the release of soluble, peripheral, and integral protein, respectively. \*, PEXEL-cleaved species. \*\*, mis-cleaved species. \*\*\*, full-length protein. # WT Hyp1-NL-mDH-FL blot (lanes 1-4) has been published previously<sup>42</sup>, and has been included as a comparator to the KAHRP and STEVOR blots.

Figure 3. Sequential truncation of the spacer region caused trapping of the Nluc-mDH-FL

constructs in the ER and PV of the parasite. (A)Representative IFA images (n=3 independent replicates) of HSP101-HAglmS parasites expressing Hyp1/STEVOR/KAHRP-Nluc-mDH-FL proteins with various spacer lengths (see cartoon representations). Parasite cells were probed with anti-EXP2 and anti-HA antibodies to visualise the PTEX components within the cell. Anti-Pf ERC antibody was used to visualise parasite's ER. The Nluc-mDH-FL proteins were either probed with anti-Nluc or anti-FLAG antibodies. Scale bars, 5 µm. DIC, Differential Interference Contrast. DAPI (4',6-diamidino-2-Phenylindole; Blue) was used to stain parasite nuclei. (B) Quantification of export of the spacer truncation Nluc-mDH-FL constructs presented in A. Export signal of >20 cells were recorded for each construct. Box and whisker plot represents  $25^{\text{th}}$ - $75^{\text{th}}$  and  $10^{\text{th}}$ - $90^{\text{th}}$  percentiles, respectively. Statistical significance was determined using ordinary one-way ANOVA. \*\*\*\*, p-value<0.0001; \*\*, p-value<0.01; \*, p-value = 0.0159).

Figure 4. Truncation of the spacer region of Hyp1-Nluc-mDH-FL reduces its interaction with HSP101. (A) Western blot of anti-HA immunoprecipitation (IP) of HSP101-HA glmS expressing various truncations of the Hyp1 spacer region (n=3) as indicated above each lane. HSP101-HA glmS parasites not transfected with a Hyp1-Nluc-mDH-FL reporter are represented by '-'. Immunoblots were performed to detect other PTEX components (EXP2 and PTEX150) as a positive control and GAPDH as a negative control. Hyp1 spacer mutants were visualised using anti-FLAG antibody. (B) Western blot of the reciprocal anti-Nluc IP of the truncated spacer reporters (n=4). Immunoblots were performed using anti-HA to detect HSP101 and GAPDH as a negative control. The length (C) and (D), Densitometry of the IP performed on A (left) & B (right). For anti-HA IP, the intensity of co-immunoprecipitated 13aa and 3aa spacer. For anti-Nluc IP, the intensity of co-immunoprecipitated HSP101 bands in the 13aa and 3aa spacers were normalised to the 51aa spacer. Error bars,  $\pm$ SD. Statistical significance was determined using an Ordinary one-way ANOVA. \*\*\*\*, p value<0.0001.

Figure 5. The sequence requirements of the minimal 13aa spacer for export are relatively unconstrained. (A) Alignment of the sequences of Hyp1 spacers inclusive of the PEXEL motif who trafficking was investigated. (B) Fluorescence microscopy of Hyp1 spacers detected with rabbit Nluc and mouse EXP2 mAb as a PVM marker that 51aa, [?]NT13aa, 13aa and 13aa.Ser reporter were exported efficiently whereas the Hyp1 13aa.Glu and Hyp1 3aa appeared to be more strongly retained in the parasite. (C) Quantification of the % of exported protein signal beyond the EXP2 boundary into the infected RBC compartment indicated that the degree of export was progressively reduced in the Hyp1 13aa.Ser, Hyp1 13aa.Glu and Hyp1 3aa parasites relative to the Hyp1 13aa parasites. The Hyp1 [?]NT13aa parasites exported more efficiently Hyp1 13aa parasites. Export signal of >60 cells were recorded for each construct. Box and whisker plot represents  $25^{\text{th}}$ - $75^{\text{th}}$  and  $10^{\text{th}}$ - $90^{\text{th}}$  percentiles, respectively. Statistical significance was determined using ordinary one-way ANOVA. P-value, \* <0.05. (D) Nanoluciferase export assay with the same parasite lines produced similar trends as the microscopic imaging but specifically indicated that most of the reduction in export in the mutant Hyp1 13aa and Hyp1 3aa spacer was due to retention in the PV rather than the parasite. Statistical significance was determined using ordinary Dunnett's T3 multiple comparisons test in Graphpad Prism. P-values, \*<0.05, \*<0.01.

Figure 6. Hypothetical model for the HSP101-cargo recognition. (A) Side and top views of scale model of PTEX based on cryoelectron microscopy structure showing how Hyp1spacer with attached Nluc (red) could extend into PTEX's central cavity <sup>19</sup>. (B) Enlarged view of central HSP101 cavity showing the spacer length required to engage the helical protein binding regions of HSP101-nucleotide binding domain 2 (NBD2) which could apply an unfolding force to the cargo proteins for translocation.

Α

PvPMV activity

























C





\*\*\*\*

\*\*\*\*







