The mature N termini of *Plasmodium* effector proteins confer specificity of export

Muhammad M. Hasan, Alexander J. Polino, Sumit Mukherjee, Barbara Vaupel, Daniel E. Goldberg

Division of Infectious Diseases, Department of Medicine, and Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri, USA

Address correspondence: Daniel E. Goldberg

Email: dgoldberg@wustl.edu

Abstract

The intraerythrocytic malaria parasite *Plasmodium falciparum* exports hundreds of proteins into the host red blood cell (RBC). Most are targeted to the ER by a stretch of hydrophobic amino acids and cleaved further downstream at a conserved motif called the Protein Export Element (PEXEL) by the ER protease plasmepsin V (PM V). The mature effectors then travel through the secretory pathway to the parasitophorous vacuole (PV) that surrounds the parasite. There, PEXEL proteins are somehow recognized as export-destined proteins, as opposed to PV-resident proteins, and are selectively translocated out into the RBC. The mature N terminus appears to be important for export. There is conflicting data on whether PM V cleavage is needed for proper export, or whether any means of generating the mature N terminus would suffice. We replaced the PEXEL-containing N-terminal sequence of an exported GFP reporter with a signal peptide sequence and showed that precise cleavage by signal peptidase, generating the proper mature N terminus, yields export competence. Expressing a construct with only the native ER targeting signal without the PM V cleavage site dramatically decreased the amount of a mature PEXEL reporter, indicating that the hydrophobic stretch lacks an

efficient cleavage signal. Therefore, the PEXEL motif functions as a specialized signal cleavage site when appropriately located after an ER targeting sequence. Our data suggest that PM V cleavage and RBC export are two independent events for PEXEL proteins. We also tested and rejected the hypothesis that an alpha-helical mature N terminus is necessary for export.

Importance

Malaria parasites export hundreds of proteins to the cytoplasm of the host red blood cells for their survival. A five amino acid sequence, called the PEXEL motif, is conserved among many exported proteins and is thought to be a signal for export. However, the motif is cleaved inside the endoplasmic reticulum of the parasite and mature proteins starting from the fourth PEXEL residue travel to the parasite periphery for export. We showed that the PEXEL motif is dispensable for export as long as identical mature proteins can be efficiently produced via alternative means in the ER. We also showed that the exported and non-exported proteins are differentiated at the parasite periphery based on their mature N termini, however, any discernible export signal within that region remained cryptic. Our study resolves a longstanding paradox in PEXEL protein trafficking.

Introduction

The malaria parasite *Plasmodium falciparum* asexually replicates in the human Red Blood Cell (RBC). An infected RBC is extensively modified by the parasite to make it compatible for growth. Hundreds of parasite proteins are exported into the RBC for this purpose and as such, protein export is critical for parasite survival (1, 2).

Exported proteins are first loaded into the parasite endoplasmic reticulum (ER) and then progress along the secretory pathway by vesicular trafficking to be released outside the parasite plasma membrane (PPM) (3-7). A membrane-bound space called the parasitophorous vacuole (PV) surrounds the PPM throughout the parasite's replication cycle (8-11). To reach the RBC

cytoplasm and beyond, exported proteins, as opposed to PV-resident proteins, must translocate through a multiprotein complex in the PV membrane called Plasmodium translocon of exported proteins (PTEX) (12-16).

In eukaryotic cells, secretory proteins generally have a hydrophobic signal peptide at their N terminus, which targets the protein to the ER membrane. It is then cleaved by the ER protease signal peptidase (SP) to release the secretory protein into the ER lumen. The mature protein can reside in the ER or travel to other organelles depending on the presence of diverse organelle-targeting signals in its sequence. The absence of any such signal generally takes the protein to the plasma membrane to be released outside the cell (17). Similarly, *Plasmodium* proteins with a signal peptide travel to the PV, and for further export into the RBC, an additional signal is necessary (18-20).

A pentameric motif was found to be conserved in most exported proteins of *Plasmodium* and was named the *Plasmodium* Export Element (PEXEL)/Host Targetting motif (21, 22). These proteins are cotranslationally loaded into the ER by a stretch of hydrophobic residues resembling a signal peptide, and the PEXEL motif is generally located after a non-conserved "spacer" region following that signal (5, 23, 24). Minimal constructs capable of export must include around 10 residues downstream of the PEXEL motif as well. However, there is very little primary sequence conservation in these critical residues, and replacing them with a stretch of alanine suffices for export (6, 23, 25).

The PEXEL motif is composed of arginine and leucine in the first and third positions and aspartate, glutamine or glutamate in the fifth position, often expressed as RxLxE/Q/D. Interestingly, the motif is cleaved between the third and the fourth residues by an ER resident aspartic protease called plasmepsin V (PM V), and the mature proteins starting with a comparatively less unique signal of xE/Q/D are then exported (26-28). This raises the question that if the PEXEL motif's most conserved part (RxL) is cleaved off in the ER, how can it convey

73 export-specificity at the PV? Multiple hypotheses have been offered in this regard. One model is
74 that PM V hands over the cleaved PEXEL proteins to specific chaperones that guide them in the
75 secretory pathway till their delivery to PTEX (27). A recently published study proposed that
76 HSP101, a component of the PTEX translocon, binds to nascent PEXEL proteins at the ER and
77 takes them to the PTEX following PM V cleavage (29). Another model is that PEXEL proteins
78 enter the ER through a PM V-containing translocon that is distinct from the SP-containing
79 translocon, which specifies a selective route to the PTEX complex (4). Mature PEXEL proteins
80 are N-terminally acetylated, which has also been speculated to be a postmark for their eventual
81 export, though acetylation alone is insufficient to achieve export (30, 31). None of these theories
82 has been validated convincingly.

83 Taking a step back, several experiments have tested the assumption that the PEXEL motif
84 confers export capacity to *Plasmodium* proteins with disparate conclusions. Mutating the
85 conserved R and L, or deleting the PEXEL motif altogether in reporter constructs inhibited their
86 export to the RBC (23, 26, 32). However, these alterations affected the cleavage by PM V, and
87 thus only verified the requirement of the PEXEL motif for cleavage at the ER and not for RBC
88 export per se. More pertinent experiments in this regard would be to check the exportability of
89 mature reporters that are identical to PEXEL-cleaved reporters, albeit in the absence of the
90 PEXEL motif itself. One such construct was designed by removing the region following the
91 hydrophobic stretch to the RxL sequence of a PEXEL protein so that the mature reporter started
92 with the sequence xE after the cleavage by SP at the end of the hydrophobic stretch (27). It did
93 not get exported into the RBC, indicating that the PEXEL motif is required in nascent PEXEL
94 proteins for export. On the other hand, another study tested reporter constructs where the
95 mature parts of PEXEL proteins were preceded by a signal anchor (an ER loading signal,
96 similar to a signal peptide but without the signal cleavage site) and a self-cleaving viral capsid
97 protease sequence that cleaves the reporter right before the xE/Q/D sequences (33).

Interestingly, these reporters were exported into the RBC, suggesting that the signal for export resides downstream of the PEXEL cleavage site and the motif itself is not required for export. Reporters starting with mature N terminal residues of PEXEL proteins were also exported when targeted to the ER by an internal transmembrane domain, supporting the later proposition (34).

In this study, we tested the essentiality of the PEXEL motif for *Plasmodium* protein export with reporter constructs that generate identical mature proteins in the ER, albeit differing in the presence or absence of a PEXEL motif. Our data support a model where the PEXEL motif is required for the cleavage of a protein in the ER but is irrelevant for export to the RBC. Rather, the mature N-terminal domain of PEXEL proteins alone determines their localization once they reach the PV.

Results

Serial fractionation can differentiate the localization of PV-resident and RBC-exported protein.

We expressed minimal constructs (containing the hydrophobic stretch, spacer, PEXEL motif and short mature N termini of variable lengths) of three PEXEL proteins as well as a construct containing the first 62 residues of the known PV-resident protein SERA5 whose signal peptide cleavage site was previously determined (35). All were tagged C-terminally with eGFP (Fig. 1A, S1). We integrated them for expression at the *attB* locus in the *P. falciparum* NF54^{attb} strain (36). To determine their localization, we harvested 30h old parasite-infected RBCs and carried out a cell fractionation/anti-GFP western blot strategy (Fig. 1B).

Cells were treated with tetanolysin, a bacterial toxin that selectively perforates the RBC membrane (32, 37). Following centrifugation, we collected the supernatant as the infected RBC cytoplasm fraction and further treated the pellet with saponin. Saponin permeabilizes the PV membrane more efficiently than the PPM (38, 39). Supernatant from this treatment was

collected as the PV fraction and the pellet was then fully lysed in RIPA lysis buffer as the intracellular parasite fraction. Western blotting of these fractions revealed the efficient export of the mature PEXEL reporters into the RBC and the retention of the SERA5 reporter in the PV (Fig 1C). We also probed the fractions with an anti-SERA5 antibody that does not recognize the reporter fragment (40). It revealed an enrichment of the native protein in the PV. The ER protein PM V was only detected in the intracellular parasite fraction (Fig 1C). These two proteins were also probed in subsequent experiments as markers of consistent fractionation. In our blots, we saw substantial free GFP in the intracellular parasite and PV fractions (marked with asterisks in Fig. 1C). This steered us away from using fluorescent microscopy and cautioned against the quantitative overinterpretation of previous studies.

Mature N-terminal sequences of PEXEL proteins are sufficient for their export into the RBC.

To test the necessity of the PEXEL motif for export, we constructed fusion reporters of SERA5 and KAHRP, in which modules N-terminal to the cleavage site are combined with modules C-terminal to the cleavage site (Fig. 2A, S1). The mature KAHRP reporter was efficiently cleaved (presumably by SP) and exported when placed after the SERA5 signal peptide (Fig. 2B). This suggests that its export signal resides C-terminal to the cleavage site and that neither the PEXEL motif nor processing specifically by PM V is required for export. On the other hand, the mature SERA5 reporter placed after the KAHRP RxL was cleaved (presumably by PM V), yet the conserved PEXEL residues were not sufficient to turn a PV resident reporter into an exported reporter. Therefore, a properly placed RxL is sufficient for cleavage in the ER but not for export. Taken together, these results demonstrate that if a protein travels to the PV from the ER, its mature N terminus determines its export capacity and not the pre-cleavage sequence.

It can be argued from the above experiments that mature PEXEL reporters were exported due to the remnant part of the PEXEL, i.e. the xE/Q/D sequence at their N termini and the mature SERA5 reporter (starting with TG) was not exported because it did not have this signature at its N terminus (Fig. S1). We tested the importance of the semi-conserved 5th residue of PEXEL by substituting it with alanine in both KAHRP and GBP-130 reporters. These substitutions did not inhibit their export, indicating that the 5th residue of PEXEL is not required for RBC export (Fig. 2C).

Substituting the hydrophobic stretch of the EMP3 reporter with the SERA5 signal peptide restores export.

Previously, a case was made in favour of the essentiality of the PEXEL motif for export with the observation that the export of the minimal mature EMP3 reporter was abrogated when residues between the hydrophobic stretch and the mature N terminus of EMP3, including the RxL, were replaced with a single alanine. From this construct, the same mature reporter was produced, presumably by the action of SP after the alanine at the end of the hydrophobic stretch, yet the reporter was not exported (26). As this result is paradoxical to our conclusion, we decided to test a similar EMP3 reporter ourselves (Fig. 3A(ii), S1) and found, as previously observed, that the reporter did not pass the PV (Fig. 3B). Interestingly though, the level of the total mature reporter in this deletion construct was significantly lower compared to the original reporter (Fig. 3C). We then fused the mature EMP3 reporter after the SERA5 signal peptide (Fig. 3A(iii)). The total quantity and the export of the mature reporter from this construct were comparable to those of the full EMP3 reporter (Fig. 3C). Using mass spectrometry, we confirmed that the N termini of all three of these mature reporters were identical (Fig. S2).

Efficient export of the SERA5$^{1-22}$-EMP3$^{63-82}$-eGFP fusion reporter again confirmed that the PEXEL motif is not essential for export and we hypothesized that the perplexing lack of export of the EMP3$^{1-36}$ -A-EMP3$^{63-82}$-eGFP construct had arisen from an inefficient cleavage after the

hydrophobic stretch by the SP, which led to the degradation of the reporter. Protein level and export of the mature reporter were maintained when the nascent N-terminal residues of SERA5 were replaced with those from EMP3 (Fig. 3A (iv), B, C). When we did the same replacement in the full-length, PEXEL-containing EMP3 reporter (fig 3A(v)), in addition to the exported mature reporter, we also observed a higher molecular weight band in the parasite fraction whose size corresponds to an alternative mature reporter starting from the SP cleavage site (Fig. 3B, C). Because we do not observe this higher molecular weight band in the original full-length reporter, it supports the argument that the signal cleavage site of EMP3 is non-functional. In any case, it was clear from our tested EMP3 reporters that the export deficiency of the mature reporter from the EMP3$^{1-36}$-A-EMP3$^{63-82}$-eGFP construct could be reversed by changing the hydrophobic stretch of EMP3 without reintroducing the PEXEL motif.

Putative structures of PEXEL protein mature N-termini resemble each other and are different than those of PV-resident proteins.

As the RBC-export signal of the PEXEL proteins seemingly resides at their mature N terminus, we searched for commonalities in the first 10 amino acids following the cleavage sites of several PEXEL proteins for which there was experimental proof of RBC export (Fig. 4A, Table S1). As expected, there was barely any primary sequence conservation, except for the semi-conserved second position, which we already found to be not essential for export (Fig. 2C). We then looked at the AlphaFold structural predictions for that region of the selected proteins. Interestingly, most of them are predicted to form an alpha-helical structure (Fig. 4B, Table S1), albeit with low confidence scores. On the other hand, when we looked at the structures of the cognate residues following the signal peptide cleavage sites of a few known soluble PV resident proteins (Fig. 4C, Table S1), the majority of them had predicted random coil structures (Fig. 4D, Table S1). Based on this, we hypothesized that an alpha-helical structure at the mature N terminus of the PEXEL proteins is required for their export.

196 **Helix-breaking proline insertion did not abrogate the export of reporters to the RBC.**

197 To test this hypothesis, we designed minimal reporter constructs of KAHRP and EMP3 where
198 we inserted proline at the 3rd or 6th position of the mature reporters (Table S1). However, these
199 insertions did not abrogate the export of the reporter constructs (Fig. 5A). Although proline is
200 known to break alpha-helical structures, it might not work as such in the context of the KAHRP
201 and EMP3 mature N terminal sequences (41-44). Therefore, we tested two other reporters, one
202 where we placed a known alpha-helical sequence after the PEXEL cleavage site and another
203 where a proline insertion has been experimentally determined to break the alpha helix
204 conformation (Fig. S1) (45, 46). Using CD spectrometry, we verified the supposed
205 conformations of these two peptides in vitro (Fig. 5B). However, in vivo, both of these reporters
206 were exported into the RBC with similar efficiency (Fig. 5C, S1).

207

208 **Discussion**

209 In this study, we have reevaluated the function of the PEXEL motif as the signal for the export of
210 *Plasmodium* proteins into the RBC and found that it is not a direct prerequisite for export. We
211 tested two PEXEL reporter constructs that conferred export as long as their export-competent
212 mature N termini were exposed either by SP cleavage or by PEXEL cleavage. On the other
213 hand, introducing a PEXEL cleavage site in a PV resident reporter allowed proper cleavage
214 without altering its PV localization.

215 We also tested a previously published reporter (Fig. 3A(ii)) whose export was blocked upon
216 removal of the PEXEL motif. We made variants of this reporter to test if the absence of the
217 PEXEL motif was the cause of its export deficiency and found that replacing its ER targeting
218 signal remedied the defect. The addition of the SERA5 signal peptide before the mature part of
219 this reporter (Fig. 3A(iv)) possibly directed it to the SP-containing ER translocon, whereas the

EMP3$^{1-13}$-SERA5$^{3-22}$-EMP3$^{37-82}$-eGFP ((Fig. 3A(v)) construct was likely targeted to both PM V and SP-containing translocons, maturing into two alternative reporters. In the absence of both SP and PM V cleavage sites, the EMP3$^{1-36}$-EMP3$^{63-82}$-eGFP (Fig. 3A(ii)) construct was likely degraded by Endoplasmic Reticulum Associated Protein Degradation (ERAD) or proteasomal-degradation pathway. It is not completely clear why the small amount of the mature reporter liberated from this construct was not exported despite having an export-competent mature N terminus. We suspect that it might be related to the ER entry of the reporter because the knock-down of PfSPC25 and PfSec62, two components of the ER translocon required for protein import into the ER, also exhibited a decrease in protein levels for some native PEXEL proteins (4). Overall, more pertinent to the scope of this study, we showed that export could be re-established by introducing a functional SP cleavage site in this PEXEL-less reporter.

In light of these observations, two obvious questions emanate. One, if the PEXEL motif is not directly involved in protein export to RBC, what is its function and why is it conserved in so many exported proteins? And two, what is the real signal for export in the PEXEL proteins? The experiments presented here and other published experiments indicate that the PEXEL motif serves as a very potent cleavage signal at the ER of the parasite (26, 27, 30). We have shown that the hydrophobic stretch of the PEXEL protein EMP3 lacks a strong signal cleavage site, and that is also the case for several other PEXEL proteins tested before. For example, removing the RxL from PEXEL protein KAHRP results in more full-length reporter accumulation than for reporters cleaved after the hydrophobic stretch (5, 32). Another PEXEL protein HRPII accumulated as the full-length protein when PM V activity was inhibited (27). Pharmacological inhibition of PM V resulted in the accumulation of full-length EMP3 reporters, which strongly supports our conclusion (47, 48). Therefore, we think the PEXEL motif can be considered as a specialized signal cleavage site located distally from a hydrophobic signal anchor sequence of the PEXEL proteins and PM V acts as a non-canonical signal peptidase in *Plasmodium*. As the

exported proteins are first loaded into the ER, it is unsurprising that a lot of them possess this motif.

In our attempt to address the second question, we searched for commonalities at the mature N terminus of several experimentally validated exported PEXEL proteins by analyzing their AlphaFold structure. We found an abundance of alpha-helical conformational predictions, which were absent from the corresponding sequences of several PV-resident proteins. A huge caveat of this analysis is that this structural prediction is based on the whole protein sequence where the mature N terminus is not yet liberated. Though the alpha helix is a very common secondary structure in proteins and not all the exported N termini had alpha-helical predictions, we decided to test its worth as the export signal because it is clear from the nonconserved nature of the primary sequences of exported N-termini that the export signal would not be a very unique one. Our results argue against the hypothesis that a simple alpha-helical structure at the mature N terminus would suffice as an export signal. However, this approach needs further refinements, and the experimental determination of the structure of the mature N termini of multiple PEXEL proteins could be enlightening. N termini of several PEXEL-negative exported proteins (PNEPs) also function as efficient export signals, indicating a common mechanism of selection of exported proteins (34). The role of N terminal acetylation is another interesting potential export signal. However, there are also examples of N terminally acetylated PV resident reporters (33), which indicates that it is not sufficient in itself as an export signal. Our data also supports this view as our export-deficient mature EMP3 reporter was acetylated. For now, the real signal for export is still unknown except that it is very promiscuous at the level of primary sequence and functions as an export signal only in the context of the N-terminal end of a protein.

As discussed in the introduction, multiple models have been proposed to connect two events in PEXEL protein trafficking, the cleavage by Plasmepsin V at the ER and the export through the PTEX channel at the PV membrane. Our conclusion indicates that these two events are

independent of each other. Therefore, at least theoretically, there is no need for special sorting or chaperoning of the export-destined proteins from the ER; the selection can take place at the PV in its entirety. In the simplest scenario, an export-competent N terminus is recognized and differentiated from the export-incompetent N termini of PV-resident proteins by the PTEX complex (Fig. 6).

Another important implication of our conclusion is that there might be PEXEL proteins that do not get exported into the RBC, but rather travel to other compartments from the ER. Examples of such proteins are rare but not nonexistent. For example, RESA is a dense granule protein that is cleaved by PM V in vitro and in vivo (23, 49, 50). Plasmepsin IX has an appropriately located PEXEL motif but localizes to the rhoptries (51). Therefore, it is important to experimentally determine the localization of PEXEL proteins rather than assuming that they are exported.

Materials and Methods

Maintenance of parasite cultures.

P. falciparum strain NF54^{attB} was cultured in RPMI1640 (Gibco) media supplemented with 0.25% (w/v) Albumax (Gibco), 15 mg/l hypoxanthine, 110 mg/l sodium pyruvate, 1.19 g/l Hepes, 2.52 g/l sodium bicarbonate, 2 g/l glucose, and 10 mg/l gentamicin. Hematocrit concentration was maintained at 2%. Parasites expressing the reporters were maintained in 5nM WR99210. Human RBCs were collected from St Louis Children's Hospital blood bank. Cultures were kept inside gas (5% O2, 5% CO2, and 90% N2) chambers at 37°C.

Plasmid construction and transfection.

A donor plasmid containing an attP site integrates into the *cg6 attB* locus of the NF54^{attB} strain when co-transfected with the pINT plasmid coding for Bxb1 integrase (52). All our reporter

constructs were integrated into the genome using this strategy. They were expressed under the control of HSP86 (PF3D7_0708400) promoter and 3'UTR.

We first cloned the eGFP sequence between the AvrII and EagI sites of the pEOE-attP vector (53) using In-Fusion cloning (Clontech). KAHRP, GBP130, EMP3 and SERA5 minimal regions were amplified from NF54 mRNA isolated with TRIzol (ThermoFisher) using the SuperScript RT-PCR kit (Invitrogen). They were cloned into the XhoI-AvrII site of the pEOE-attP-eGFP vector using In-Fusion cloning. All the fusion constructs were made from the appropriate backbone vector with the QuikChange Lightning Multi Site Directed Mutagenesis kit (Agilent Technologies). Reporters were sequenced from Genewiz before transfecting the parasites. All the primers used for cloning and sequencing were purchased from IDT and their sequences are listed in Table S2.

Plasmids were isolated from bacterial clones using Nucleobond Xtra Midi (MN) kit and electroporated into the parasite as previously described (53). Successfully integrated clones were selected with media containing 5nM WR99210 from 36h post-transfection onwards as the donor plasmid codes for human dihydrofolate reductase (hDHFR) as the selection marker (54).

Culture synchronization.

An asynchronous parasite culture was washed in RPMI medium and then passed through a MACS LD magnet column (Miltenyi Biotec). NF54^{attB} parasites complete a replication cycle in 44 to 48h under our culture condition and older parasites (>28h old) are captured on the column due to the presence of paramagnetic hemozoin crystal (55). They were eluted in a prewarmed 2% hematocrit culture and incubated for 3 hours for egress and invasion. This culture was then treated with 5% sorbitol at 37°C for 10 minutes to osmotically lyse older parasites (due to the establishment of the new permeability pathway), leaving the newly invaded rings intact (56, 57).

316 Synchronized parasites were thereafter maintained by constantly shaking at 80 RPM under 5%
317 parasitemia to maintain the synchrony.

318 **Compartment fractionation.**

319 Around 30h old synchronous parasite cultures were passed through the magnetic columns to
320 harvest only infected RBCs. This step was critical because otherwise, haemoglobin from
321 uninfected RBCs mask the western blot signals from a sample. Infected RBCs were washed
322 twice in PBS and then treated with 50HU tetanolysin (Biological Laboratory Inc) in 60µl PBS
323 plus HALT-Protease Inhibitor (PI) Cocktail (Thermo Fisher Scientific) for 10 minutes at room
324 temperature. Following centrifugation at 1500g for 2 minutes, the supernatants were collected
325 as the RBC fractions. The pellets were washed twice in PBS before treating with 60µl of 0.035%
326 saponin in PBS-PI for 5 minutes on ice. Supernatants were collected as PV fractions and the
327 pellets were washed twice before adding 60µl RIPA lysis buffer with PI. These were rapidly
328 frozen and thawed using liquid nitrogen and a 42°C water bath three times and the supernatants
329 were collected as the parasite fractions after 15 minutes of centrifugation at 4°C for 10 minutes.
330 20µl of 4X sample buffer with β-mercaptoethanol as the reducing agent was mixed with each
331 sample and boiled for 5 minutes before storage at -20°C.

332 **SDS PAGE and western blotting.**

333 15 µl samples from each fraction were run in a 4-15% gradient gel (Biorad) and then transferred
334 to a PVDF membrane for western blotting. We used mouse anti-GFP (Takara) at 1:1000
335 dilution, mouse anti-PM V (58) at 1:250 dilution and rabbit anti-SERA5 (40) at 1:1000 dilution as
336 primary antibodies and IRDye conjugated goat secondary antibodies (LICOR) at 1:15000
337 dilution. Blots were incubated with primary antibodies O/N at 4°C and with secondary antibodies
338 for 1 hour at room temperature. Licor Odessey blocking buffer was used for blocking and
339 primary antibody dilutions and PBS plus 1% tween-20 was used to prepare secondary antibody

dilutions as well as in all the washing steps. Blots were imaged in a Licor Odyssey imager and images were prepared (and quantified when required) using Image Studio Lite 5.2 (Licor). The protein level of EMP3 reporters was calculated by adding the intensities of reporter GFP bands (excluding the free GFP or any higher molecular weight band) from all three fractions and then normalizing that value for the PM V band intensity from the parasite fraction. Statistical analyses were performed in GraphPad Prism.

Preparation of samples for mass spectrometry.

Infected RBCs were harvested using magnetic columns from 300ml 5% parasitemia cultures and then directly lysed with 500ul GFP-trap lysis buffer (10 mM Tris/Cl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5 % Nonidet™ P40 Substitute) by freeze-thaw. Supernatants were incubated with GFP-trap magnetic agarose (ChromoTek) at 4°C for 1 hour with continuous rotation. The beads were washed 3 times with wash buffer (10 mM Tris/Cl pH 7.5, 150 mM NaCl, 0.05 % Nonidet™ P40 Substitute, 0.5 mM EDTA) and then 50 µl 2x sample buffer with β-mercaptoethanol was added to the beads and boiled for 5 minutes for elution. All the eluted samples were run in a Biorad gradient gel. Specific bands were visualized with Coomassie blue staining and cut out of the gel for submission to the Mass Spectrometry Technology Access Center.

Proteomics and data analysis.

The protein gel bands were subjected to in-gel digestion. Each gel band was washed in 100 mM Ammonium Bicarbonate (AmBic)/Acetonitrile (ACN), reduced with 10 mM dithiothreitol, and cysteines were alkylated with 100mM iodoacetamide. Gel bands were washed in 100mM AmBic/ACN prior to adding 1 µg trypsin for overnight incubation at 37°C. The supernatant containing peptides was saved into a new tube. Gel was washed at room temperature for ten minutes with gentle shaking in 50% ACN/5% FA, and the supernatant was saved to peptide

364 solution. The wash step was repeated each by 80% ACN/5% FA, and 100% ACN, and all
365 supernatant was saved and then subject to the speedvac dry. After lyophilization, peptides were
366 reconstituted with 0.1% FA in water. Peptides were injected onto a Neo trap cartridge coupled
367 with an analytical column (75 µm ID x 50 cm PepMap™ Neo C18, 2 µm). Samples were separated
368 using a linear gradient of solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in
369 ACN) using a Vanquish Neo UHPLC System coupled to an Orbitrap Eclipse Tribrid Mass
370 Spectrometer with FAIMS Pro Duo interface (Thermo Fisher Scientific).

371 The resulting tandem MS data was queried for protein identification against the custom database,
372 Plasmodium falciparum 3D7 database plus the 3 custom proteins (The cleaved form of EMP3(i),
373 (ii), and (iii)), using Mascot v.2.8.0 (Matrix Science). The following modifications were set as
374 search parameters: peptide mass tolerance at 20 ppm, trypsin enzyme, 3 allowed missed
375 cleavage sites, carbamidomethylated cysteine (static modification), and oxidized methionine,
376 deaminated asparagine/glutamine, and protein N-term acetylation (variable modification). The
377 search results were validated with 1% FDR of protein threshold and 90% of peptide threshold
378 using Scaffold v5.2.1 (Proteome Software). Data are available via ProteomeXchange with
379 identifier PXD041451. Please use the following credential for review: Username:
380 reviewer_pxd041451@ebi.ac.uk Password: m25z5Zvm.

381 **CD Spectrometry.**

382 Custom peptides with N-terminal acetylation were purchased from Biomatik. Peptides were
383 diluted in 10 nM potassium phosphate buffer with 40% TFE with a final concentration of 0.2 mg/ml.
384 CD spectra were recorded on a JASCO-J715 polarimeter (JASCO, Tokyo, Japan) over the
385 wavelength range 190-250 nm in a 1-mm path length quartz cuvette using a step size of 0.1 nm.
386 For each wavelength, three scans were performed. AVIV software was used for background
387 subtraction. Mean residual ellipticity [θ] vs Wavelength plots were generated using CDtoolX (59)
388 and online server K2D3 (60) was used to determine the alpha-helical composition of the peptides.

Analysis of PEXEL and PV-resident protein sequences.

Experimentally validated exported PEXEL proteins were selected from the list of the PEXEL proteins from Jonsdottir TK et al. (1). If multiple proteins had the same primary sequence at their mature N termini (mainly from exported protein families like RESA, RIFIN, STEVOR etc.), only one sequence was included in the list. PV-resident proteins were manually selected by reviewing several publications. The relevant part of their protein sequences and AlphaFold structures were taken from the Plasmodb database (61). A, B and C were used as codes if a particular residue was part of an alpha-helix, beta-sheet or random coil structure. Frequency plots of the amino acids or structures were created using the Weblogo tool (62).

Acknowledgements

Mass Spectrometry analyses were performed by the Mass Spectrometry Technology Access Center at McDonnell Genome Institute (MTAC@MGI) at Washington University School of Medicine. We thank the group of Dr Michael Blackman from The Francis Crick Institute for the anti-SERA5 antibody and the group of Dr Sergej Djuranovic from the department of Cell Biology & Physiology at Washington University School of Medicine for the eGFP construct. This work was supported by NIH grant AI17106201.

Figure legends

Figure 1: Minimal reporter constructs and their localization. (A) Schematic representation of minimal reporter constructs. (i) (ii) and (iii) are PEXEL reporters and (iv) is a PV resident reporter. Different segments of PEXEL reporters are delineated by dashed lines and labelled at the bottom for (iv). Residue numbers starting from the nascent N terminus are printed within each segment and the PEXEL motif residues are highlighted in bold red. (B) Compartment fractionation strategy. RBCM: Red blood cell membrane; PVM: Parasitophorous vacuole

membrane; PPM: Parasite plasma membrane. (C) Representative western blots of different fractions from RBCs infected with *P. falciparum* expressing the reporters from A. Primary antibodies that were used to probe the blots are labelled at the left. The bottom band (marked with asterisks) in the anti-GFP blots is the free GFP band devoid of mature reporter portions. Each construct was tested at least 3 times.

Figure 2: Investigation of the role of the PEXEL motif in protein export to the RBC. (A) Schematic representation of KAHRP and SERA5 fusion reporters (ii and iii). Original reporters (i and iv) are also shown for reference. (B) Representative western blots of different fractions from RBCs infected with *P. falciparum* expressing the reporters from A. Primary antibodies that were used to probe the blots are labelled at the left. The experiment was performed twice. (C) Representative western blots of different fractions from RBCs infected with *P. falciparum* expressing KAHRP and GBP130 reporters with alanine substituted semi-conserved 5th positions of the PEXEL motif. Each construct was tested twice.

Figure 3: Processing, export and comparison of the protein level of different EMP3 reporters. (A) Schematic representation of the EMP3 reporter constructs. An alternative mature N terminus for the last construct is marked at the bottom (B) Representative western blots of different fractions from RBCs infected with *P. falciparum* expressing the reporters from A. Primary antibodies that were used to probe the blots are labelled at the left. Note the presence of an alternative mature form (marked with a purple arrow) for the last reporter in the parasite fraction. (C) Normalized western blot quantification of the standard mature forms of the reporters. Signals were combined from each fraction and then normalized to the PM V signal from the parasite fraction. * denotes a $P \leq 0.05$ and ** denotes $P \leq 0.01$ in Fisher's LSD test. The P-value for the one-way ANOVA was 0.0113. Mean and standard deviations from 2 or 3 biological replicates are shown along with individual data points.

Figure 4: Primary sequences and putative structures of PEXEL and PV-resident proteins. (A) Amino acid frequency plots of 59 experimentally validated exported PEXEL proteins starting from the first position of the PEXEL motif to the 10th position of the mature N terminus. (B) Frequency plot of AlphaFold structural predictions of the same residues shown in panel (A), where “A” denotes alpha-helical, “B” denotes beta-sheet and “C” denotes random coil. (C) Amino acid frequency plots of 13 experimentally validated PV-resident proteins starting from the -3 position of the signal peptide cleavage site to the 10th position of the mature N terminus. (D) Frequency plot of AlphaFold structural predictions of the same residues shown in panel (A), with the letters denoting the same structural conformation as in panel (B).

Figure 5: Investigation of the role of alpha-helical mature N terminus in protein export. (A) Representative western blots of different fractions of the KAHRP and EMP3 reporters with proline insertions at the 3rd or the 6th position of their mature N terminal region. The experiment was performed twice. (B) CD spectra of two small peptides, one of which takes an alpha-helical conformation in vitro whereas the other one loses the conformation due to alanine to proline substitution. (C)Representative western blots of different fractions from RBCs infected with *P. falciparum* expressing two artificially designed PEXEL reporters with the mature N terminal sequences shown in the construct name. The experiment was performed twice.

Figure 6: A model showing the trafficking of *P. falciparum* secretory proteins. Secretory proteins are targetted to the ER by their hydrophobic stretch and then cleaved by PM V at the PEXEL motif or by SP at the signal cleavage site, followed by acetylation. The cleavage liberates mature N termini that can be export-competent (shown in red) or incompetent (shown in black). There might be other organelle-targeting signals also in the mature proteins that direct them to their respective target organelle. Mature proteins secreted into the PV are recognized and loaded into the PTEX translocon based on the export competency of their mature N terminus.

Supplementary Figure 1: Full sequence of all the reporters in this study. eGFP sequence is not shown as well as the first 29 residues of GBP130. Bold arrows denote cleavage sites. Sequences from PEXEL proteins are printed in red. Substitutions and insertions are highlighted in bold purple.

Supplementary Figure 2: Coverage map and most N terminal peptide spectrum of EMP3 reporters. Construct names are on top. The coverage map highlights detected peptides at the 95% threshold. Green shades in the coverage map denote post-translational modification.

Supplementary table 1: List of PEXEL proteins and PV resident proteins along with their mature N terminal sequences and AlphaFold structural predictions. These sequences were used to generate the frequency plots shown in Figure 4. For the AlphaFold structural predictions, “a” denotes alpha-helical, “b” denotes beta-sheet and “c” denotes random coil.

Supplementary table 2: List of primers used in this study

References:

1. Jonsdottir TK, Gabriela M, Crabb BS, T FdK-W, Gilson PR. Defining the Essential Exportome of the Malaria Parasite. Trends in parasitology. 2021;37(7):664-75.
2. Spillman NJ, Beck JR, Goldberg DE. Protein export into malaria parasite-infected erythrocytes: mechanisms and functional consequences. Annual review of biochemistry. 2015;84:813-41.
3. 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):509-13.
4. Marapana DS, Dagley LF, Sandow JJ, Nebl T, Triglia T, Pasternak M, et al. Plasmepsin V cleaves malaria effector proteins in a distinct endoplasmic reticulum translocation interactome for export to the erythrocyte. Nature microbiology. 2018;3(9):1010-22.
5. Boddey JA, O'Neill MT, Lopaticki S, Carvalho TG, Hodder AN, Nebl T, 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.
6. 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):258-62.
7. Wickham ME, Rug M, Ralph SA, Klonis N, McFadden GI, Tilley L, et al. Trafficking and assembly of the cytoadherence complex in Plasmodium falciparum-infected human erythrocytes. The EMBO journal. 2001;20(20):5636-49.

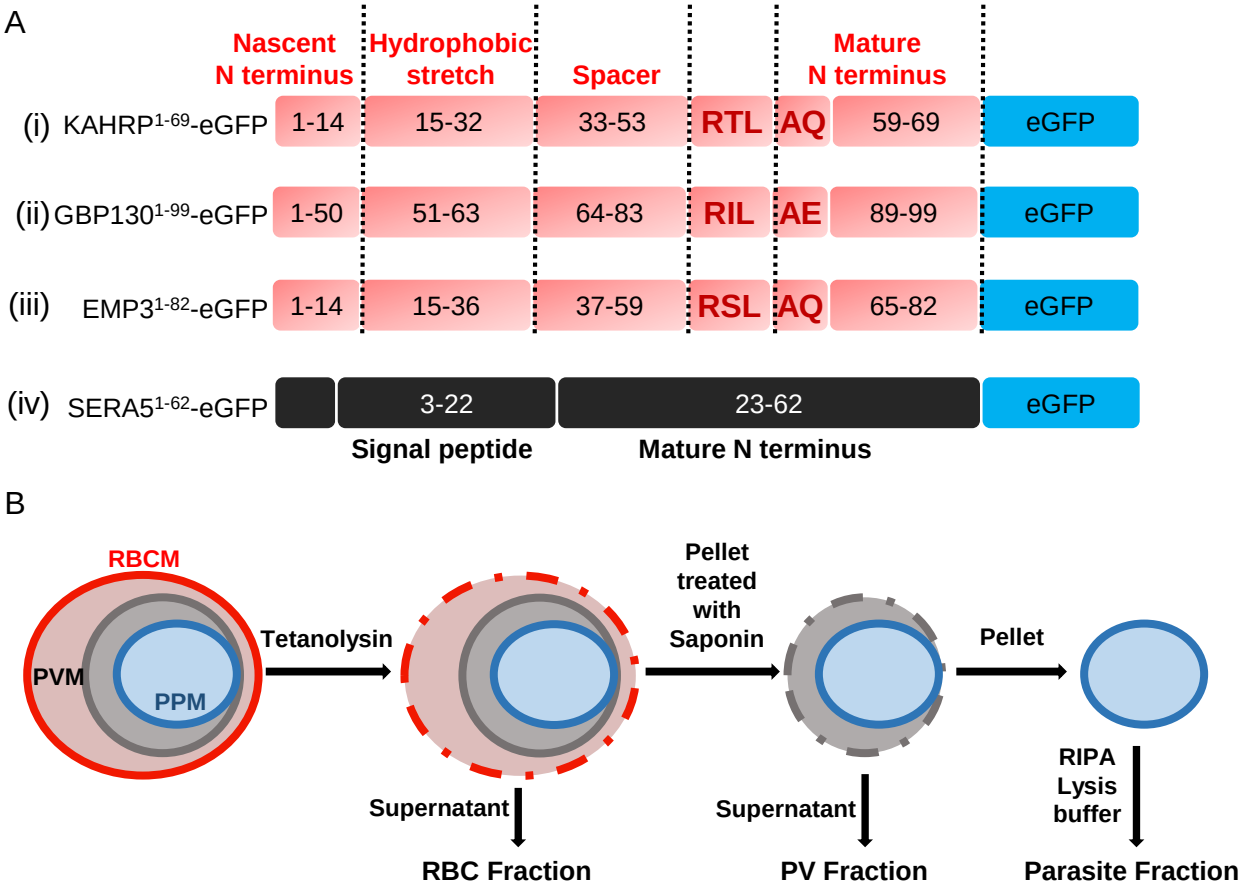
494 8. Goldberg DE, Zimmerberg J. Hardly Vacuous: The Parasitophorous Vacuolar Membrane of
495 Malaria Parasites. Trends in parasitology. 2020;36(2):138-46.
496 9. Lingelbach K, Joiner KA. The parasitophorous vacuole membrane surrounding Plasmodium and
497 Toxoplasma: an unusual compartment in infected cells. Journal of cell science. 1998;111 (Pt 11):1467-
498 75.
499 10. Bannister LH, Hopkins JM, Fowler RE, Krishna S, Mitchell GH. A brief illustrated guide to the
500 ultrastructure of Plasmodium falciparum asexual blood stages. Parasitology today (Personal ed).
501 2000;16(10):427-33.
502 11. Garten M, Beck JR, Roth R, Tenkova-Heuser T, Heuser J, Istvan ES, et al. Contacting domains
503 segregate a lipid transporter from a solute transporter in the malarial host-parasite interface. Nature
504 communications. 2020;11(1):3825.
505 12. Ho CM, Beck JR, Lai M, Cui Y, Goldberg DE, Egea PF, et al. Malaria parasite translocon structure
506 and mechanism of effector export. Nature. 2018;561(7721):70-5.
507 13. Chisholm SA, Kalanon M, Nebl T, Sanders PR, Matthews KM, Dickerman BK, et al. The malaria
508 PTEX component PTEX88 interacts most closely with HSP101 at the host-parasite interface. The FEBS
509 journal. 2018;285(11):2037-55.
510 14. Elsworth B, Sanders PR, Nebl T, Batinovic S, Kalanon M, Nie CQ, et al. Proteomic analysis reveals
511 novel proteins associated with the Plasmodium protein exporter PTEX and a loss of complex stability
512 upon truncation of the core PTEX component, PTEX150. Cellular microbiology. 2016;18(11):1551-69.
513 15. Elsworth B, Matthews K, Nie CQ, Kalanon M, Charnaud SC, Sanders PR, et al. PTEX is an essential
514 nexus for protein export in malaria parasites. Nature. 2014;511(7511):587-91.
515 16. Beck JR, Muralidharan V, Oksman A, Goldberg DE. PTEX component HSP101 mediates export of
516 diverse malaria effectors into host erythrocytes. Nature. 2014;511(7511):592-5.
517 17. Viotti C. ER to Golgi-Dependent Protein Secretion: The Conventional Pathway. Methods in
518 molecular biology (Clifton, NJ). 2016;1459:3-29.
519 18. Cheresh P, Harrison T, Fujioka H, Haldar K. Targeting the malarial plastid via the parasitophorous
520 vacuole. The Journal of biological chemistry. 2002;277(18):16265-77.
521 19. Waller RF, Reed MB, Cowman AF, McFadden GI. Protein trafficking to the plastid of Plasmodium
522 falciparum is via the secretory pathway. The EMBO journal. 2000;19(8):1794-802.
523 20. Adisa A, Rug M, Klonis N, Foley M, Cowman AF, Tilley L. The signal sequence of exported
524 protein-1 directs the green fluorescent protein to the parasitophorous vacuole of transfected malaria
525 parasites. The Journal of biological chemistry. 2003;278(8):6532-42.
526 21. Marti M, Good RT, Rug M, Knuepfer E, Cowman AF. Targeting malaria virulence and remodeling
527 proteins to the host erythrocyte. Science (New York, NY). 2004;306(5703):1930-3.
528 22. Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-Estraño C, et al. A host-
529 targeting signal in virulence proteins reveals a secretome in malarial infection. Science (New York, NY).
530 2004;306(5703):1934-7.
531 23. Boddey JA, Carvalho TG, Hodder AN, Sargeant TJ, Sleebs BE, Marapana D, et al. Role of
532 plasmepsin V in export of diverse protein families from the Plasmodium falciparum exportome. Traffic
533 (Copenhagen, Denmark). 2013;14(5):532-50.
534 24. Sargeant TJ, Marti M, Caler E, Carlton JM, Simpson K, Speed TP, et al. Lineage-specific expansion
535 of proteins exported to erythrocytes in malaria parasites. Genome biology. 2006;7(2):R12.
536 25. Rhiel M, Bittl V, Tribensky A, Charnaud SC, Strecker M, Müller S, et al. Trafficking of the exported
537 P. falciparum chaperone PfHsp70x. Scientific reports. 2016;6:36174.
538 26. Boddey JA, Hodder AN, Günther S, Gilson PR, Patsiouras H, Kapp EA, et al. An aspartyl protease
539 directs malaria effector proteins to the host cell. Nature. 2010;463(7281):627-31.
540 27. Russo I, Babbitt S, Muralidharan V, Butler T, Oksman A, Goldberg DE. Plasmepsin V licenses
541 Plasmodium proteins for export into the host erythrocyte. Nature. 2010;463(7281):632-6.

542 28. Osborne AR, Speicher KD, Tamez PA, Bhattacharjee S, Speicher DW, Haldar K. The host targeting
543 motif in exported Plasmodium proteins is cleaved in the parasite endoplasmic reticulum. Molecular and
544 biochemical parasitology. 2010;171(1):25-31.
545 29. Gabriela M, Matthews KM, Boshoven C, Kouskousis B, Jonsdottir TK, Bullen HE, et al. A revised
546 mechanism for how Plasmodium falciparum recruits and exports proteins into its erythrocytic host cell.
547 PLoS pathogens. 2022;18(2):e1009977.
548 30. Chang HH, Falick AM, Carlton PM, Sedat JW, DeRisi JL, Marletta MA. N-terminal processing of
549 proteins exported by malaria parasites. Molecular and biochemical parasitology. 2008;160(2):107-15.
550 31. Polino AJ, Hasan MM, Floyd K, Avila-Cruz Y, Yang Y, Goldberg DE. An essential endoplasmic
551 reticulum-resident N-acetyltransferase ortholog in Plasmodium falciparum. Journal of cell science.
552 2023;136(6).
553 32. Boddey JA, Moritz RL, Simpson RJ, Cowman AF. Role of the Plasmodium export element in
554 trafficking parasite proteins to the infected erythrocyte. Traffic (Copenhagen, Denmark).
555 2009;10(3):285-99.
556 33. Tarr SJ, Cryar A, Thalassinos K, Haldar K, Osborne AR. The C-terminal portion of the cleaved HT
557 motif is necessary and sufficient to mediate export of proteins from the malaria parasite into its host
558 cell. Molecular microbiology. 2013;87(4):835-50.
559 34. Grüring C, Heiber A, Kruse F, Flemming S, Franci G, Colombo SF, et al. Uncovering common
560 principles in protein export of malaria parasites. Cell host & microbe. 2012;12(5):717-29.
561 35. Debrabant A, Maes P, Delplace P, Dubremetz JF, Tartar A, Camus D. Intramolecular mapping of
562 Plasmodium falciparum P126 proteolytic fragments by N-terminal amino acid sequencing. Molecular
563 and biochemical parasitology. 1992;53(1-2):89-95.
564 36. Adjalley SH, Johnston GL, Li T, Eastman RT, Ekland EH, Eappen AG, et al. Quantitative assessment
565 of Plasmodium falciparum sexual development reveals potent transmission-blocking activity by
566 methylene blue. Proceedings of the National Academy of Sciences of the United States of America.
567 2011;108(47):E1214-23.
568 37. Blumenthal R, Habig WH. Mechanism of tetanolysin-induced membrane damage: studies with
569 black lipid membranes. Journal of bacteriology. 1984;157(1):321-3.
570 38. Bhatnagar S, Nicklas S, Morrisey JM, Goldberg DE, Vaidya AB. Diverse Chemical Compounds
571 Target Plasmodium falciparum Plasma Membrane Lipid Homeostasis. ACS infectious diseases.
572 2019;5(4):550-8.
573 39. Elandalloussi LM, Smith PJ. Preparation of pure and intact Plasmodium falciparum plasma
574 membrane vesicles and partial characterisation of the plasma membrane ATPase. Malaria journal.
575 2002;1:6.
576 40. Collins CR, Hackett F, Strath M, Penzo M, Withers-Martinez C, Baker DA, et al. Malaria parasite
577 cGMP-dependent protein kinase regulates blood stage merozoite secretory organelle discharge and
578 egress. PLoS pathogens. 2013;9(5):e1003344.
579 41. O'Neil KT, DeGrado WF. A thermodynamic scale for the helix-forming tendencies of the
580 commonly occurring amino acids. Science (New York, NY). 1990;250(4981):646-51.
581 42. Woolfson DN, Williams DH. The influence of proline residues on alpha-helical structure. FEBS
582 letters. 1990;277(1-2):185-8.
583 43. Strehlow KG, Robertson AD, Baldwin RL. Proline for alanine substitutions in the C-peptide helix
584 of ribonuclease A. Biochemistry. 1991;30(23):5810-4.
585 44. Li SC, Goto NK, Williams KA, Deber CM. Alpha-helical, but not beta-sheet, propensity of proline
586 is determined by peptide environment. Proceedings of the National Academy of Sciences of the United
587 States of America. 1996;93(13):6676-81.

45. Chakrabartty A, Kortemme T, Baldwin RL. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. Protein science : a publication of the Protein Society. 1994;3(5):843-52.
46. Rohl CA, Chakrabartty A, Baldwin RL. Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. Protein science : a publication of the Protein Society. 1996;5(12):2623-37.
47. Gazdik M, Jarman KE, O'Neill MT, Hodder AN, Lowes KN, Jousset Sabroux H, et al. Exploration of the P3 region of PEXEL peptidomimetics leads to a potent inhibitor of the Plasmodium protease, plasmepsin V. Bioorganic & medicinal chemistry. 2016;24(9):1993-2010.
48. Sleebs BE, Lopaticki S, Marapana DS, O'Neill MT, Rajasekaran P, Gazdik M, et al. Inhibition of Plasmepsin V activity demonstrates its essential role in protein export, PfEMP1 display, and survival of malaria parasites. PLoS biology. 2014;12(7):e1001897.
49. Culvenor JG, Day KP, Anders RF. Plasmodium falciparum ring-infected erythrocyte surface antigen is released from merozoite dense granules after erythrocyte invasion. Infection and immunity. 1991;59(3):1183-7.
50. Aikawa M, Torii M, Sjölander A, Berzins K, Perlmann P, Miller LH. Pf155/RESA antigen is localized in dense granules of Plasmodium falciparum merozoites. Experimental parasitology. 1990;71(3):326-9.
51. Nasamu AS, Glushakova S, Russo I, Vaupel B, Oksman A, Kim AS, et al. Plasmepsins IX and X are essential and druggable mediators of malaria parasite egress and invasion. Science (New York, NY). 2017;358(6362):518-22.
52. Nkrumah LJ, Muhle RA, Moura PA, Ghosh P, Hatfull GF, Jacobs WR, Jr., et al. Efficient site-specific integration in Plasmodium falciparum chromosomes mediated by mycobacteriophage Bxb1 integrase. Nature methods. 2006;3(8):615-21.
53. Polino AJ, Miller JJ, Bhakat S, Mukherjee S, Bobba S, Bowman GR, et al. The nepenthesin insert in the Plasmodium falciparum aspartic protease plasmepsin V is necessary for enzyme function. The Journal of biological chemistry. 2022;298(9):102355.
54. de Koning-Ward TF, Fidock DA, Thathy V, Menard R, van Spaendonk RM, Waters AP, et al. The selectable marker human dihydrofolate reductase enables sequential genetic manipulation of the Plasmodium berghei genome. Molecular and biochemical parasitology. 2000;106(2):199-212.
55. Egan TJ. Recent advances in understanding the mechanism of hemozoin (malaria pigment) formation. Journal of inorganic biochemistry. 2008;102(5-6):1288-99.
56. Kutner S, Breuer WV, Ginsburg H, Aley SB, Cabantchik ZI. Characterization of permeation pathways in the plasma membrane of human erythrocytes infected with early stages of Plasmodium falciparum: association with parasite development. Journal of cellular physiology. 1985;125(3):521-7.
57. Lambros C, Vanderberg JP. Synchronization of Plasmodium falciparum erythrocytic stages in culture. The Journal of parasitology. 1979;65(3):418-20.
58. Banerjee R, Liu J, Beatty W, Pelosof L, Klemba M, Goldberg DE. Four plasmepsins are active in the Plasmodium falciparum food vacuole, including a protease with an active-site histidine. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(2):990-5.
59. Miles AJ, Wallace BA. CDtoolX, a downloadable software package for processing and analyses of circular dichroism spectroscopic data. Protein science : a publication of the Protein Society. 2018;27(9):1717-22.
60. Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. Proteins. 2012;80(2):374-81.
61. Amos B, Aurrecoechea C, Barba M, Barreto A, Basenko EY, Bażant W, et al. VEuPathDB: the eukaryotic pathogen, vector and host bioinformatics resource center. Nucleic acids research. 2022;50(D1):D898-d911.

635 62. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome
636 research. 2004;14(6):1188-90.

637


A
Nascent N terminus
Hydrophobic stretch
Spacer
Mature N terminus
(i) KAHRP1-69-eGFP
1-14
15-32
33-53
RTL
AQ
59-69
eGFP
(ii) GBP1301-99-eGFP
1-50
51-63
64-83
RIL
AE
89-99
eGFP
(iii) EMP31-82-eGFP
1-14
15-36
37-59
RSL
AQ
65-82
eGFP
(iv) SERA51-62-eGFP
3-22
23-62
eGFP
Signal peptide
Mature N terminus
B
RBCM
PVM
PPM
Tetanolysin
Pellet treated with Saponin
Pellet
RIPA Lysis buffer
Supernatant
RBC Fraction
Supernatant
PV Fraction
Parasite Fraction

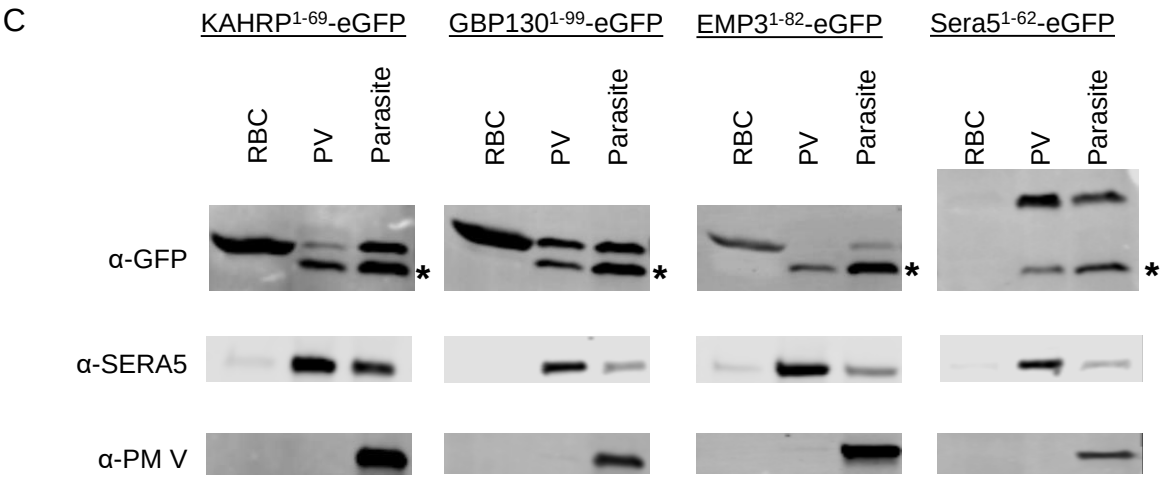

C
KAHRP1-69-eGFP
GBP1301-99-eGFP
EMP31-82-eGFP
Sera51-62-eGFP
RBC
PV
Parasite
α-GFP
α-SERA5
α-PM V
*

Figure 1

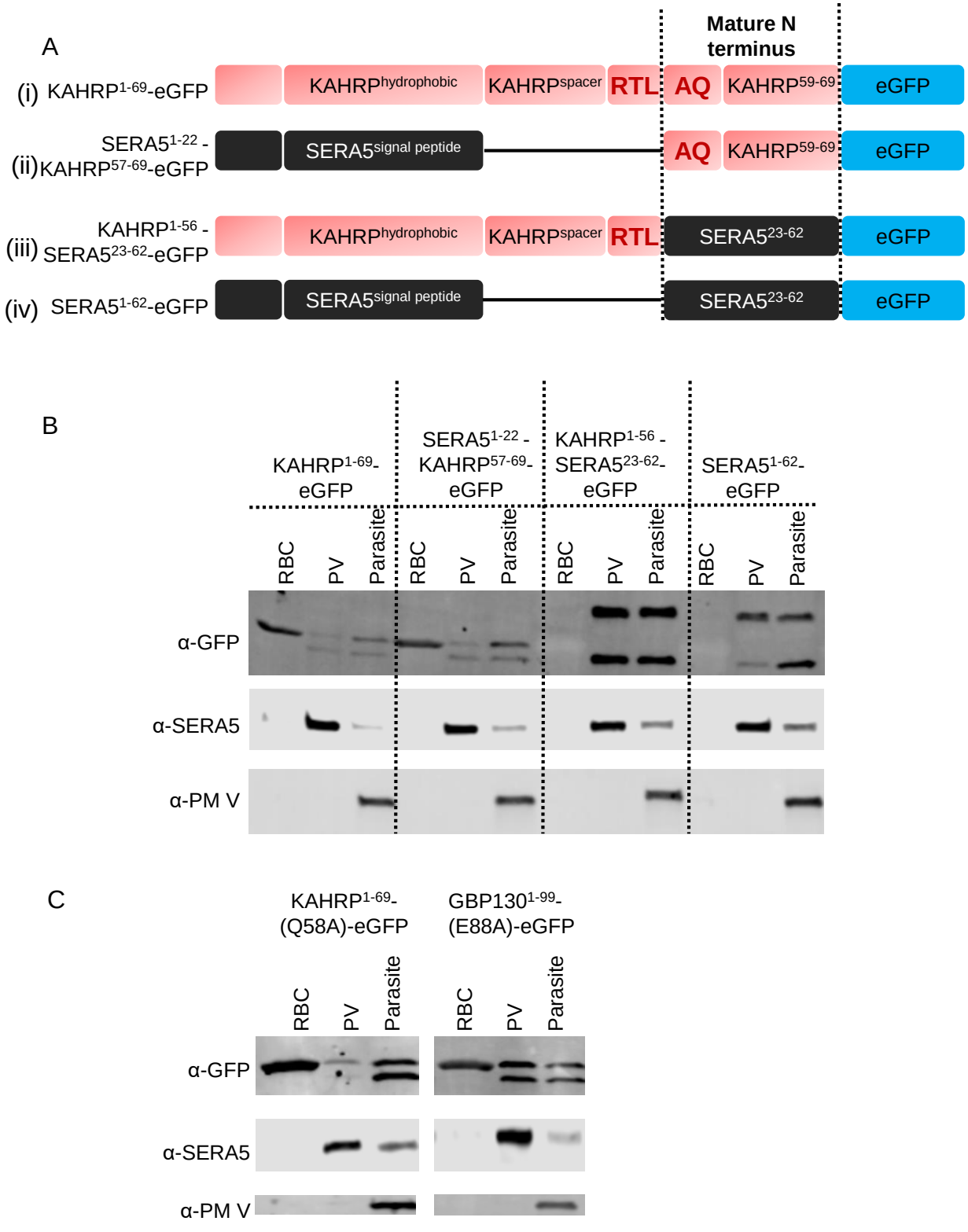

A
Mature N terminus
(i) KAHRP1-69-eGFP
KAHRPhydrophobic
KAHRPspacer
RTL
AQ
KAHRP59-69
eGFP
(ii) SERA51-22 - KAHRP57-69-eGFP
SERA5signal peptide
AQ
KAHRP59-69
eGFP
(iii) KAHRP1-56 - SERA523-62-eGFP
KAHRPhydrophobic
KAHRPspacer
RTL
SERA523-62
eGFP
(iv) SERA51-62-eGFP
SERA5signal peptide
SERA523-62
eGFP
B
KAHRP1-69-eGFP
SERA51-22 - KAHRP57-69-eGFP
KAHRP1-56 - SERA523-62-eGFP
SERA51-62-eGFP
RBC
PV
Parasite
α-GFP
α-SERA5
α-PM V
C
KAHRP1-69-(Q58A)-eGFP
GBP1301-99-(E88A)-eGFP
RBC
PV
Parasite
α-GFP
α-SERA5
α-PM V

Figure 2

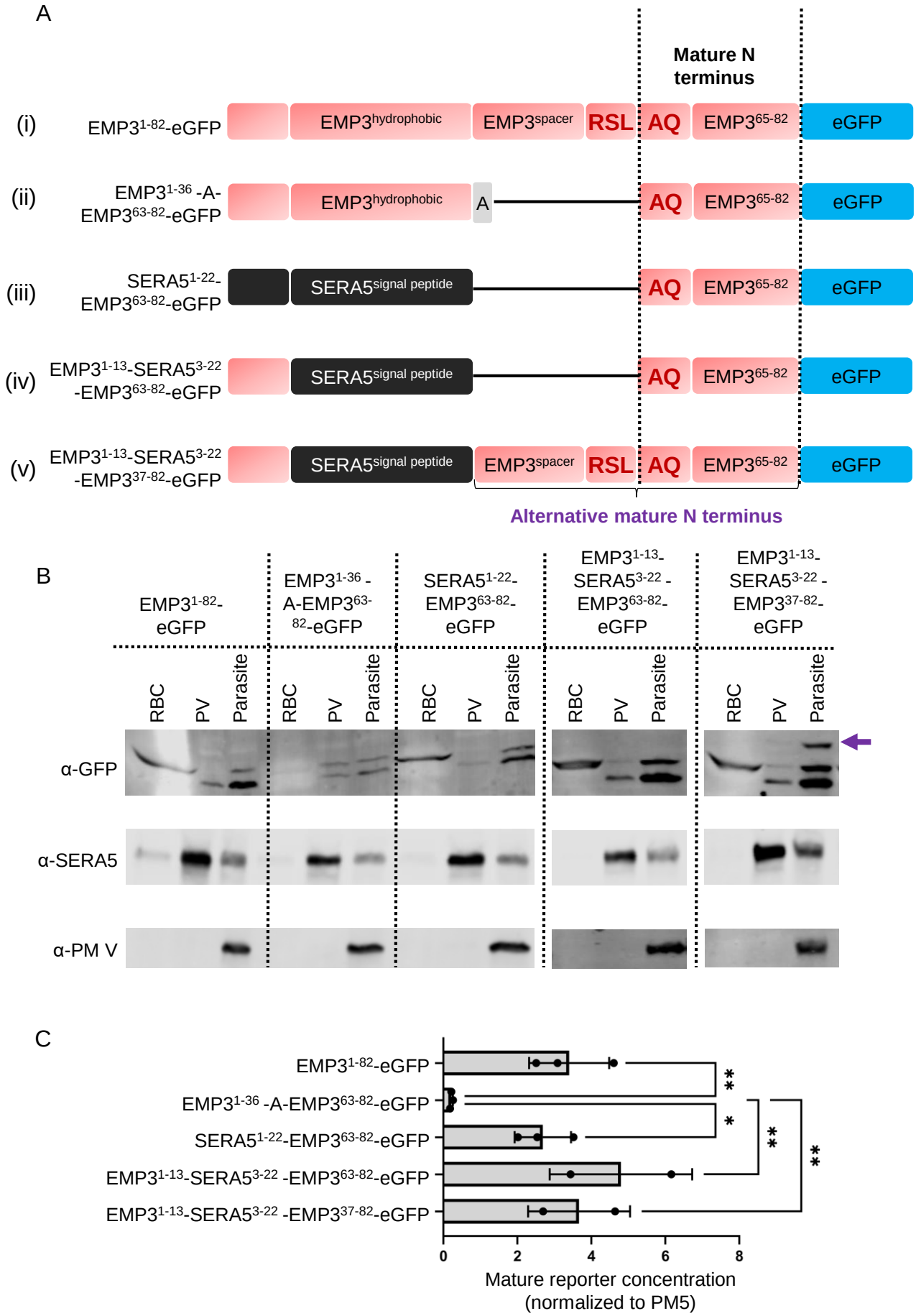

A
Mature N terminus
(i)
EMP3^1-82-eGFP
EMP3^hydrophobic
EMP3^spacer
RSL
AQ
EMP3^65-82
eGFP
(ii)
EMP3^1-36 -A-EMP3^63-82-eGFP
EMP3^hydrophobic
A
AQ
EMP3^65-82
eGFP
(iii)
SERA5^1-22-EMP3^63-82-eGFP
SERA5^signal peptide
AQ
EMP3^65-82
eGFP
(iv)
EMP3^1-13-SERA5^3-22-EMP3^63-82-eGFP
SERA5^signal peptide
AQ
EMP3^65-82
eGFP
(v)
EMP3^1-13-SERA5^3-22-EMP3^37-82-eGFP
SERA5^signal peptide
EMP3^spacer
RSL
AQ
EMP3^65-82
eGFP
Alternative mature N terminus
B
EMP3^1-82-eGFP
EMP3^1-36 -A-EMP3^63-82-eGFP
SERA5^1-22-EMP3^63-82-eGFP
EMP3^1-13-SERA5^3-22 -EMP3^63-82-eGFP
EMP3^1-13-SERA5^3-22 -EMP3^37-82-eGFP
RBC
PV
Parasite
α-GFP
α-SERA5
α-PM V
C
EMP3^1-82-eGFP
EMP3^1-36 -A-EMP3^63-82-eGFP
SERA5^1-22-EMP3^63-82-eGFP
EMP3^1-13-SERA5^3-22 -EMP3^63-82-eGFP
EMP3^1-13-SERA5^3-22 -EMP3^37-82-eGFP
**
*
**
**
0
2
4
6
8
Mature reporter concentration
(normalized to PM5)

Figure 3

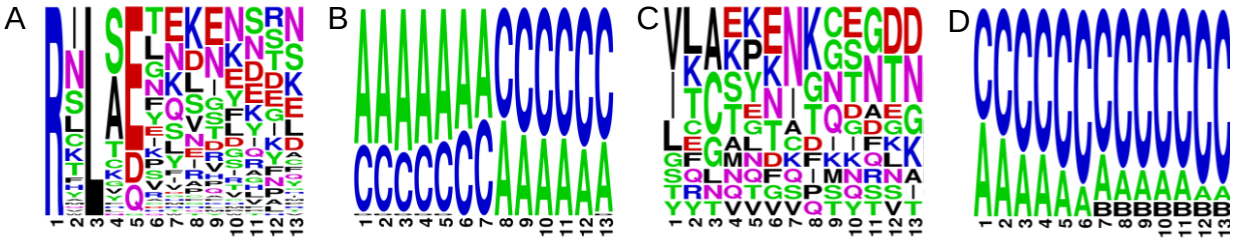


Figure 4

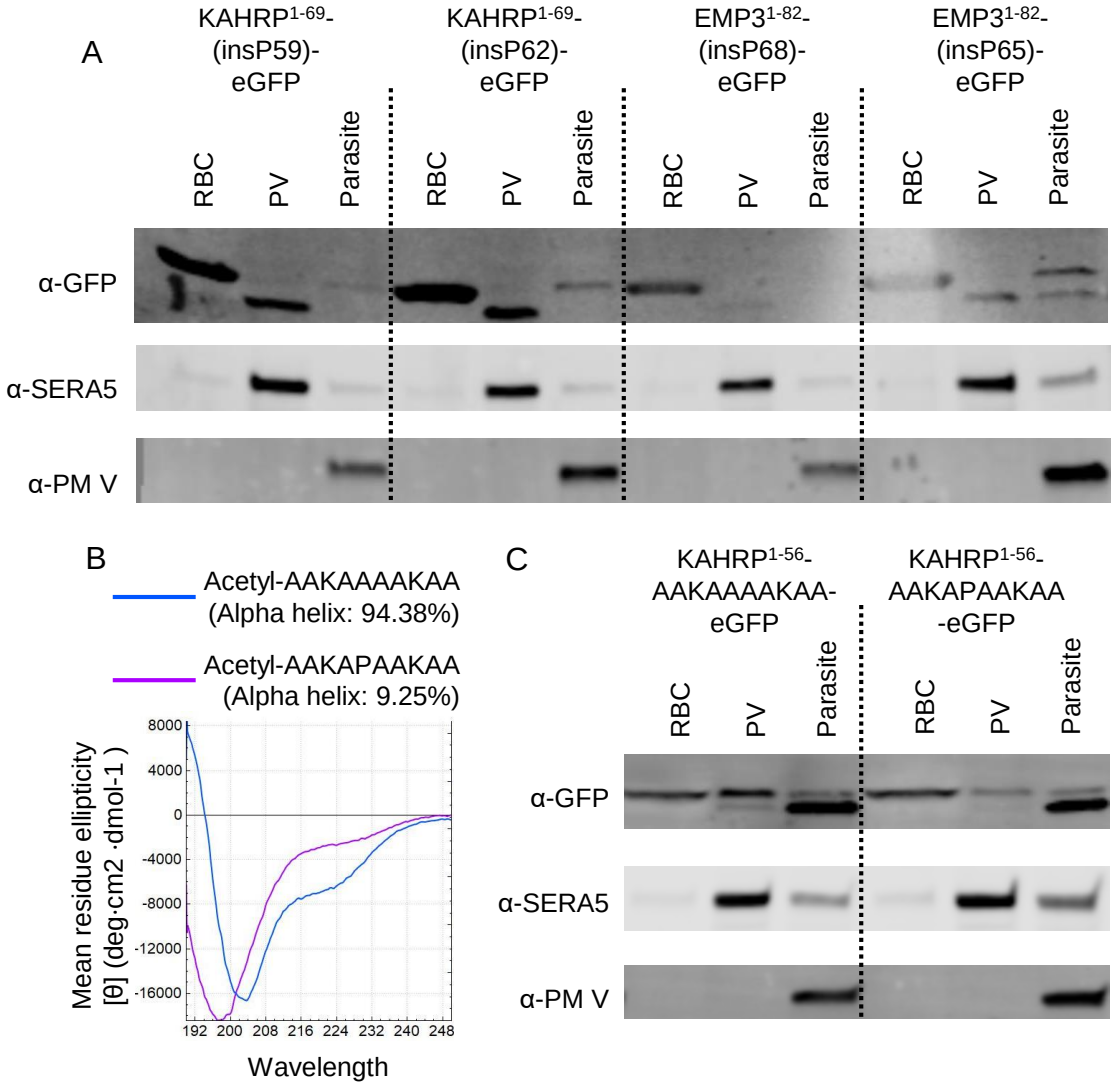


Figure 5

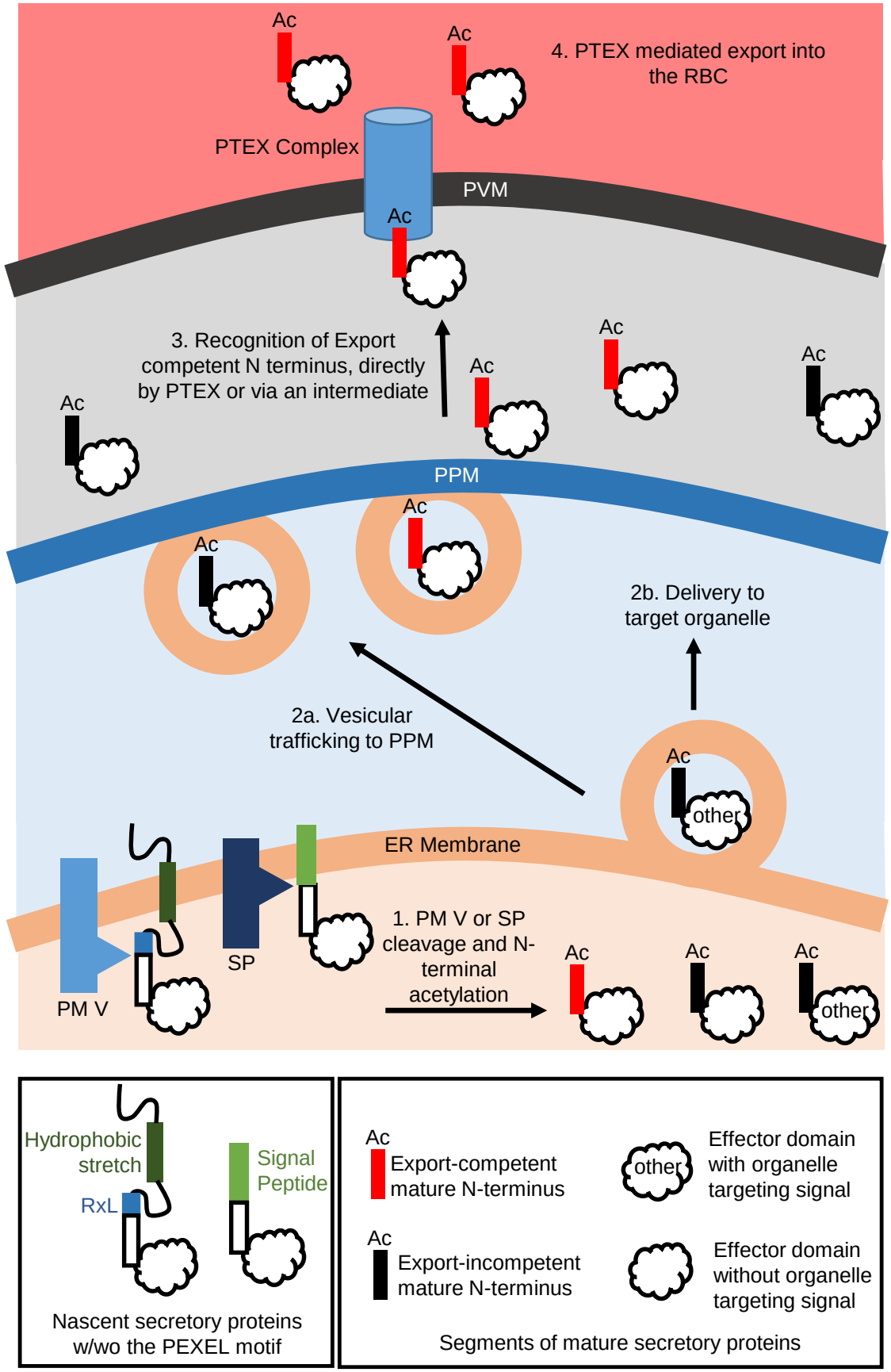
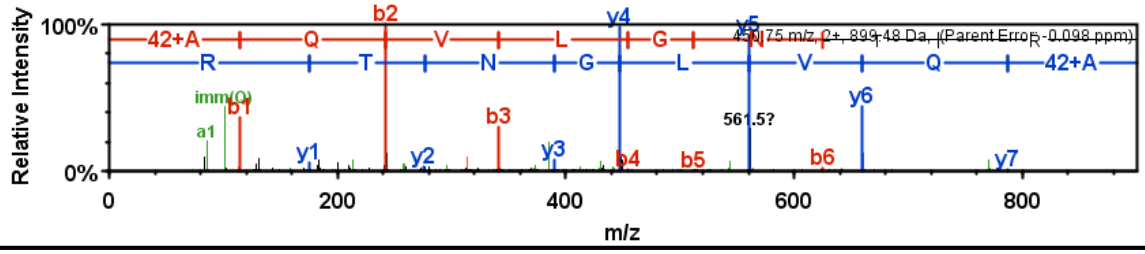

Ac
Ac
4. PTEX mediated export into the RBC
PTEX Complex
PVM
Ac
3. Recognition of Export competent N terminus, directly by PTEX or via an intermediate
Ac
Ac
Ac
Ac
PPM
Ac
Ac
2b. Delivery to target organelle
2a. Vesicular trafficking to PPM
Ac
other
ER Membrane
SP
PM V
1. PM V or SP cleavage and N-terminal acetylation
Ac
Ac
Ac
other
Hydrophobic stretch
RxL
Signal Peptide
Nascent secretory proteins w/wo the PEXEL motif
Ac
Export-competent mature N-terminus
Ac
Export-incompetent mature N-terminus
other
Effector domain with organelle targeting signal
Effector domain without organelle targeting signal
Segments of mature secretory proteins

Figure 6

Supplementary Figure 1

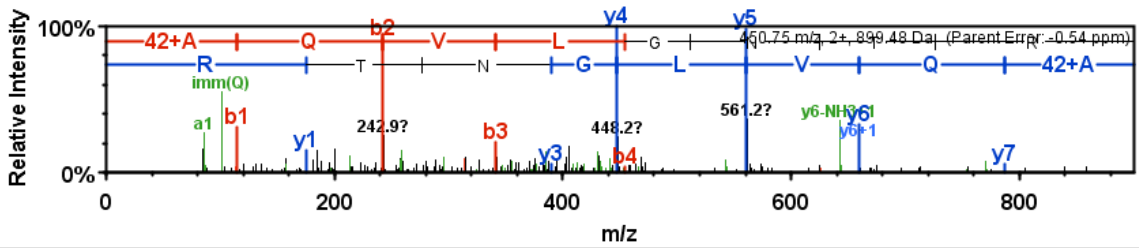
EMP3$^{1-82}$-eGFP

12 exclusive unique peptides, 35 exclusive unique spectra, 275 total spectra, 156/260 amino acids (60% coverage)



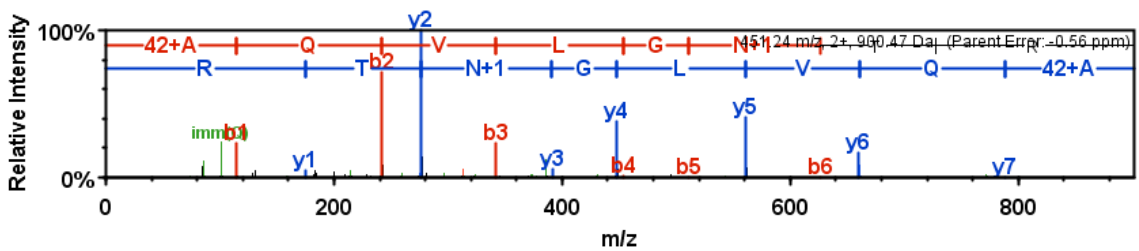
EMP3$^{1-36}$-EMP3$^{63-82}$-eGFP

10 exclusive unique peptides, 24 exclusive unique spectra, 161 total spectra, 111/260 amino acids (43% coverage)



SERA5$^{1-22}$-EMP3$^{63-82}$-eGFP

12 exclusive unique peptides, 28 exclusive unique spectra, 120 total spectra, 154/260 amino acids (59% coverage)



Supplementary Figure 2