

# 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

Sarah J. Tarr,<sup>1</sup> Adam Cryar,<sup>1</sup>  
Konstantinos Thalassinos,<sup>1</sup> Kasturi Haldar<sup>2</sup> and  
Andrew R. Osborne<sup>1\*</sup>

<sup>1</sup>*Institute of Structural and Molecular Biology, Division of Biosciences, Birkbeck and University College London, London, UK.*

<sup>2</sup>*Department of Biological Sciences, Center for Rare and Neglected Diseases, University of Notre Dame, Notre Dame, IN 46556, USA.*

## Summary

The malaria parasite exports proteins across its plasma membrane and a surrounding parasitophorous vacuole membrane, into its host erythrocyte. Most exported proteins contain a Host Targeting motif (HT motif) that targets them for export. In the parasite secretory pathway, the HT motif is cleaved by the protease plasmepsin V, but the role of the newly generated N-terminal sequence in protein export is unclear. Using a model protein that is cleaved by an exogenous viral protease, we show that the new N-terminal sequence, normally generated by plasmepsin V cleavage, is sufficient to target a protein for export, and that cleavage by plasmepsin V is not coupled directly to the transfer of a protein to the next component in the export pathway. Mutation of the fourth and fifth positions of the HT motif, as well as amino acids further downstream, block or affect the efficiency of protein export indicating that this region is necessary for efficient export. We also show that the fifth position of the HT motif is important for plasmepsin V cleavage. Our results indicate that plasmepsin V cleavage is required to generate a new N-terminal sequence that is necessary and sufficient to mediate protein export by the malaria parasite.

## Introduction

According to the World Health Organization, there were approximately 781 000 malaria-related deaths in 2009

Accepted 17 December, 2012. \*For correspondence. E-mail a.osborne@ucl.ac.uk; Tel. (+44) 0207 679 3155; Fax (+44) 0207 679 7046.

(W.H.O., 2009). The disease is caused by unicellular eukaryotes, such as *Plasmodium falciparum*. The parasite life cycle is complex, but symptoms of the disease occur when the parasite replicates inside erythrocytes. Within the erythrocyte, the parasite is contained within a vacuole, referred to as the parasitophorous vacuole (PV). By exporting proteins across its plasma membrane, the surrounding PV membrane, and into the erythrocyte, the parasite alters properties of its host cell (Haldar and Mohandas, 2007; Goldberg and Cowman, 2010). The mechanism of protein export is poorly understood, but many exported proteins are essential for parasite survival (Maier *et al.*, 2008) and virulence (Kraemer and Smith, 2006; Maier *et al.*, 2009).

The first step in the export of most proteins is targeting to the parasite endoplasmic reticulum, mediated by an N-terminal transmembrane domain (Knuepfer *et al.*, 2005c; Boddey *et al.*, 2010). An additional sequence, referred to as a Host Targeting (HT) motif (also referred to as a PEXEL), is also required for proteins to be transported across the PV membrane (Hiller *et al.*, 2004; Marti *et al.*, 2004; Przyborski *et al.*, 2005). The HT motif is located approximately 35 residues downstream of the N-terminal transmembrane domain and corresponds to the consensus sequence RxLxE/D/Q, (x is any amino acid) (Hiller *et al.*, 2004; Marti *et al.*, 2004; Przyborski *et al.*, 2005; Sargeant *et al.*, 2006; van Ooij *et al.*, 2008). Prior to export, the HT motif is proteolytically cleaved after the Leu residue to release a C-terminal fragment with a new N-terminal sequence that starts xE/D/Q (Chang *et al.*, 2008; Boddey *et al.*, 2009). Cleavage occurs in the parasite endoplasmic reticulum (Chang *et al.*, 2008; Osborne *et al.*, 2010) and is mediated by the protease plasmepsin V (Boddey *et al.*, 2010; Russo *et al.*, 2010). The new N-terminus is N-acetylated by an unknown enzyme (Chang *et al.*, 2008). Cleaved and N-acetylated proteins are trafficked to the PV where they are unfolded and translocated across the PV membrane (Gehde *et al.*, 2009). It is thought that the latter step is mediated by the PTEX complex that includes an ATPase and a putative protein conducting channel located in the PV membrane (de Koning-Ward *et al.*, 2009).

Several models have been put forward to explain the contribution of the HT motif to protein export (Boddey *et al.*, 2010; Bhattacharjee *et al.*, 2012b). Mutation of the

Arg or Leu residues to Ala blocks protein export (Hiller *et al.*, 2004; Marti *et al.*, 2004; Przyborski *et al.*, 2005; Boddey *et al.*, 2009). Mutation of the Glu/Asp/Gln residue to Ala either has little effect on export (Grüning *et al.*, 2012; Bhattacharjee *et al.*, 2012a) or partially blocks export (Boddey *et al.*, 2009). It has been proposed that the information required for export resides predominantly in the RxL sequence and that targeting of a protein for export occurs before cleavage of the HT motif, or is directly coupled to the cleavage event. Consistent with this a previous study showed that a protein whose N-terminus is generated by signal peptidase and resembles the released C-terminal fragment of the cleaved HT motif, starting xE/D/Q, is not exported (Boddey *et al.*, 2010). Hence, it has been proposed that cleavage by plasmepsin V may be directly coupled to an obligatory handoff to the next component in the export pathway (Boddey *et al.*, 2010).

A receptor for the uncleaved HT motif has not been identified. However, the motif binds to phosphatidylinositol-3-phosphate (Bhattacharjee *et al.*, 2012b). It is thought that sorting in the parasite ER is mediated by the interaction of the intact HT motif with PI(3)P on the luminal side of the ER membrane (Bhattacharjee *et al.*, 2012b). Nonetheless, protein chimeras containing only the C-terminal fragment of the cleaved HT motif were exported but this was proposed to be via an HT-motif independent export pathway (Bhattacharjee *et al.*, 2012b) that can act in conjunction with the HT signal (Bhattacharjee *et al.*, 2012a).

We propose a model in which the HT motif is cleaved by plasmepsin V and that the newly generated N-terminal sequence starting xE/D/Q is recognized late in the secretory pathway and is sufficient to mediate export. We show that a protein targeted to the ER lumen, lacking the RxL portion of the HT motif but whose N-terminus starts with the sequence xE/D/Q, is efficiently exported by the parasite. Notably, we bypass the requirement for the RxL portion of the HT motif, and plasmepsin V cleavage, by using an exogenous protease to generate an N-terminal xE/D/Q sequence in an exported protein. This shows that the role of plasmepsin V is to cleave the HT motif and reveal a new N-terminal sequence that is sufficient to mediate export. Using multiple HT motif-containing model proteins, we show that the identity of the amino acids in the fourth and fifth positions of the HT motif, as well as amino acids further downstream, influence the efficiency of protein export. Importantly, these mutations behave in a similar manner in proteins whose new N-terminal xE/D/Q sequence is generated by plasmepsin V or an exogenous protease indicating that these proteins are exported by the same export pathway. Together these data show that the role of plasmepsin V cleavage is to generate the C-terminal fragment of the cleaved HT motif

that is both necessary and sufficient to target a protein for export.

## Results

### *The C-terminal fragment of the cleaved HT motif is sufficient to mediate protein export*

We tested whether the C-terminal portion of the cleaved HT motif is sufficient to mediate export of PFI1755c (REX3) (Spielmann *et al.*, 2006), which has the HT motif <sup>46</sup>RQLSE<sup>50</sup>. A protein comprising residues 1–61 of PFI1755c fused to GFP (PFI1755c<sub>1–61</sub>:GFP) (Sargeant *et al.*, 2006) was efficiently exported to the host cell (Fig. 1A). Cleavage of the HT motif generates a new N-terminal sequence starting with Ser and Glu as the first and second residues respectively. By convention, the two residues located on the C-terminal side of a protease cleavage site are referred to as P1' and P2' residues; the P1' residue is one amino acid from the cleavage site and the P2' residue is two amino acids from the cleavage site (Schechter and Berger, 1967); hereafter we use this convention.

To test whether the C-terminal fragment of the cleaved HT motif is sufficient to mediate export, we exploited the protease domain from the Semliki forest virus capsid protein. This domain is a self-cleaving protease; once the protease folds, its C-terminus loops into the active site and is cleaved in *cis* after a specific Trp residue (Choi *et al.*, 1997; Kowarik *et al.*, 2002). The region following this Trp residue is released from the protein. To test whether the cleaved HT motif is sufficient to mediate protein export, we generated a fusion protein comprising an ER-targeting signal sequence, followed by the capsid protease domain, residues 49–61 of PFI1755c, and a C-terminal GFP tag (SS:CapProt:PFI1755c<sub>49–61</sub>:GFP; Fig. 1B). Self-cleavage by the capsid protease will release a C-terminal fragment that starts with the N-terminal P1' P2' sequence, Ser Glu, which is identical to the N-terminus of PFI1755c<sub>1–61</sub>:GFP generated by plasmepsin V cleavage (Fig. 1B).

SS:CapProt:PFI1755c<sub>49–61</sub>:GFP was expressed in *P. falciparum*. Consistent with the idea that the new N-terminus generated by cleavage of the HT motif is sufficient to mediate protein export, the GFP protein was efficiently exported into the host erythrocyte (Fig. 1C). Quantification of the GFP fluorescence localized in the erythrocyte cytoplasm showed that export efficiency was similar in parasites expressing SS:CapProt:PFI1755c<sub>49–61</sub>:GFP and PFI1755c<sub>1–61</sub>:GFP (Fig. 1D). Analysis by Western blotting showed that a GFP-tagged fragment was generated by capsid protease cleavage, the size of which matched that generated by plasmepsin V mediated cleavage of PFI1755c<sub>1–61</sub>:GFP (Fig. 1E, closed arrow, compare lanes 2 and 3). Mass spectrometry analysis of tryptic peptides derived from immunoprecipi-

**Table 1.** Analysis of protein N-terminal sequences by mass spectrometry.

Sample (mutation)	N-terminal peptides	Acetylation	Experiment 1			Experiment 2		
			Error (ppm)	Error SD	Percent matched products	Error (ppm)	Error SD	Percent matched products
PFI1755c <sub>1-61</sub> :GFP (wild-type)	SEPVVEEQDLK	N-acetyl	1.91	1.9	51.5	1.11	2.7	100
	SEPVVEEQDLKK	N-acetyl	0.98	2.23	75	-0.5	1.71	100
	PVVEEQDLK		0.56	2.13	48.1	-0.8	1.76	66.7
PFI1755c <sub>1-61</sub> :GFP (P1' to Asp)	DEPVVEEQDLK		-0.66	1.37	54.5	-1.86	1.15	54.5
	DEPVVEEQDLKK		2.44	2.67	47.2	0.47	2.32	45.8
	DEPVVEEQDLKK	N-acetyl	0.54	3.12	83.3	-1.39	0.1	54.2
	PVVEEQDLKK		4.29	1.56	90	1.07	2.34	95
PFI1755c <sub>1-61</sub> :GFP (P1' and P2' to Tyr-Gly)	YGPVVEEQDLK		1.79	2.31	100	2.74	0.87	100
	YGPVVEEQDLKK		2.8	1.59	100	1.29	1.46	97.2
	PVVEEQDLKK		2.31	0.76	90	1.18	3.1	90
	EEQDLKK		-0.99	0.78	57.1	0.02	2.19	52.4
	VEEQDLKK		1.07	2.5	83.3	-0.53	1.6	70.8
	VVEEQDLKK		1.09	1.86	66.7	-0.38	0.68	63
	GPVVEEQDLKK		1.78	1.51	87.9	1.47	1.54	84.8
SS:CapProt:PFI1755c <sub>49-61</sub> :GFP	SEPVVEEQDLK	N-acetyl	0.67	2.61	87.9	1.37	1.09	63.6
	SEPVVEEQDLKK	N-acetyl	1.3	1.46	97.2	3.69	0.47	83.3
	PVVEEQDLK		0.47	1.12	70.4	0.68	1.5	77.8
SS:CapProt:PFI1755c <sub>49-61</sub> :GFP (P1' to Asp)	DEPVVEEQDLK		2.56	1.95	84.8	1.53	2.58	78.8
	DEPVVEEQDLKK		2.18	1.14	66.7	2.02	0.09	45.8
	DEPVVEEQDLKK	N-acetyl	2	1.43	52.8	0.16	0.13	50
	PVVEEQDLK		0.61	2.1	77.8	-0.42	0.68	33.3
	PVVEEQDLKK		3.17	2.94	90	3.11	0.5	90
SS:CapProt:PFI1755c <sub>49-61</sub> :GFP (P1' and P2' to Tyr-Gly)	YGPVVEEQDLK		4.07	2.99	100	3.39	2.05	100
	YGPVVEEQDLKK		0.93	0.83	100	3.99	2.1	100
	PVVEEQDLKK		2.4	2.95	83.3	-0.34	0.93	93.3
	EEQDLKK		-2.23	3.85	52.4	-1.54	2.46	61.9
	VEEQDLKK		0.78	1.46	66.7	1.67	1.53	75
	VVEEQDLKK		1.08	0.73	66.7	-0.34	1.24	63
	GPVVEEQDLKK		1.21	0.81	84.8	1.84	1.59	84.8
SS:CapProt:PFI1755c <sub>49-61</sub> :GFP (P1' to Ala)	SAPVVEEQDLK	N-acetyl	2.95	3.03	90.9	2.28	3.26	100
	SAPVVEEQDLKK	N-acetyl	0.79	0.71	100	0.5	1.09	95.8
	PVVEEQDLK		1.41	1.93	70.4	1.99	4.65	77.8
	EEQDLK		-2.68	3.12	33.3	-0.84	6.03	50
PFI1755c <sub>1-61</sub> :GFP (sequence downstream of cleavage site: SAPVVAAAALK)	SAPVVAAAALKK	N-acetyl	2.51	1.01	58.3	1.6	2.3	88.9
	VAAAALK		0.77	1.16	64.3	-0.48	0.64	47.6
PFI1755c <sub>1-61</sub> :GFP (sequence downstream of cleavage site: SAPVVSTLTK)	SAPVVSTLTK	N-acetyl	1.85	0.95	75.8	-0.91	1.44	63.6
	SAPVVSTLKK	N-acetyl	0.55	1.67	86.1	2.37	1.53	77.8

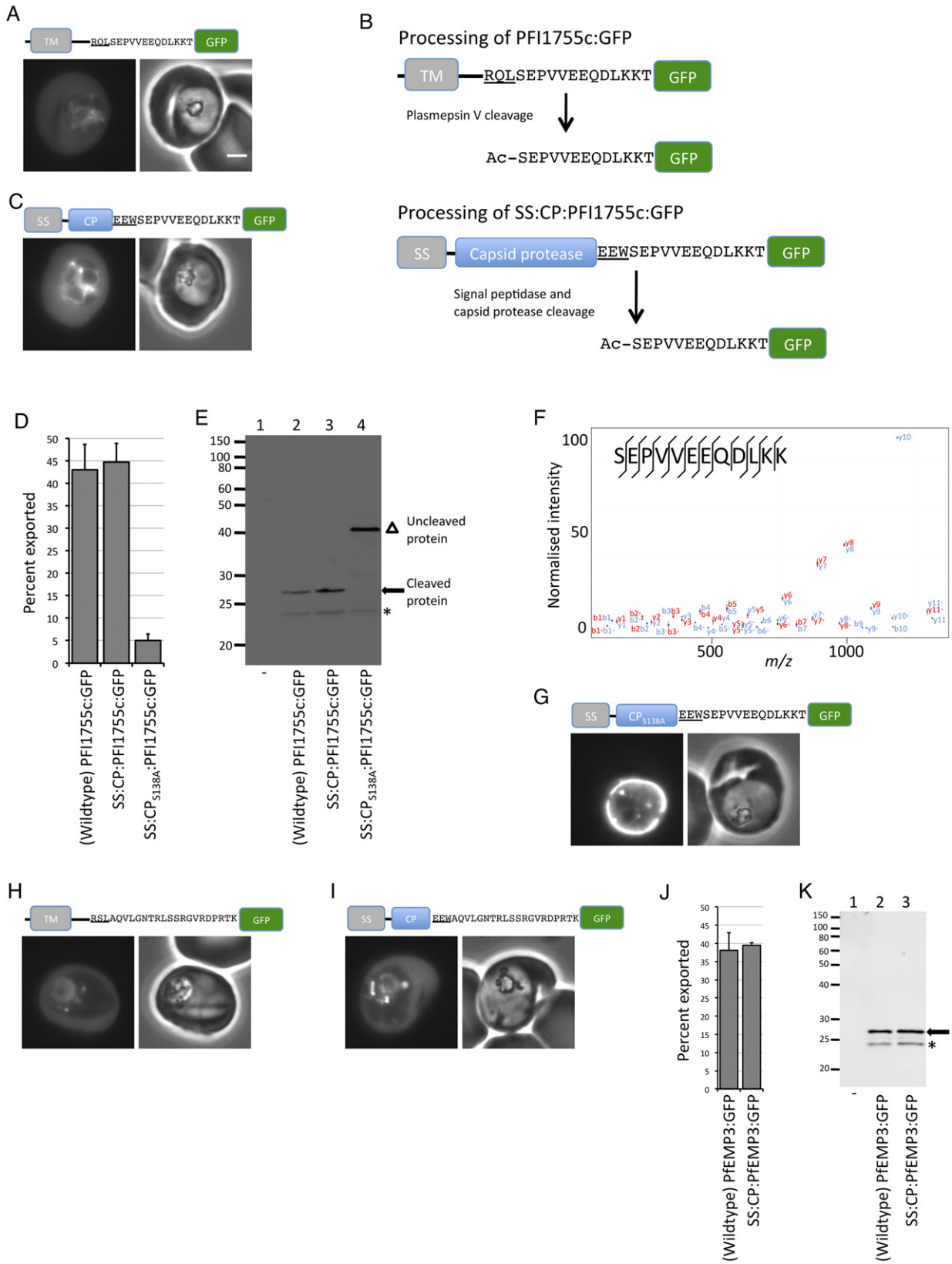
N-terminal peptides derived from the PFI1755c fragment of the indicated proteins are shown from two independent experiments. No peptides N-terminal to the protease cleavage sites of proteins processed by plasmepsin V or capsid protease were identified. GFP-derived peptides are not shown. ppm, parts per million mass accuracy of the precursor ion; percent matched products, percentage of expected b and y ions obtained. The ppm error refers to the standard deviation of three technical replicates within each experiment. Averaged tandem mass spectra for the most N-terminal peptides derived from PFI1755c<sub>1-61</sub>:GFP with the linker region mutated to SAPVVSTLTKT, and SS:CapProt:PFI1755c<sub>49-61</sub>:GFP are shown in Fig. S4. SD, standard deviation.

tated SS:CapProt:PFI1755c<sub>49-61</sub>:GFP and PFI1755c<sub>1-61</sub>:GFP indicated that these proteins indeed have identical P1' P2' N-terminal sequences. Although the N-termini are not completely homogeneous, the peptide derived from the cleaved and N-acetylated HT motif, N-Ac-SEPVVEEQDK(K), was the most N-terminal peptide identified in both samples (Table 1 and Fig. 1F).

To confirm that the capsid protease does not contain a cryptic export sequence, the active site Ser138 in the capsid protease domain of SS:CapProt:PFI1755c<sub>49-61</sub>:GFP was mutated to Ala (Choi *et al.*, 1997). This mutation prevented protease self-cleavage and release of the C-terminal GFP fragment from the protease domain (Fig. 1E, triangle, lane 4). Importantly, the mutation

blocked export of GFP to the host cell; GFP accumulated at the periphery of the parasite (Fig. 1G and D). Immunofluorescence analysis showed colocalization of the GFP signal with the PV marker Exp2 (Fig. S1C). These results show that the capsid protease does not contain a cryptic export sequence and that cleavage of SS:CapProt:PFI1755c<sub>49-61</sub>:GFP is mediated by the capsid protease, not by plasmepsin V. Additionally, this indicates that the ability of the P1', P2' and following residues to support export is blocked if they are preceded by an N-terminal folded domain, in this case the capsid protease.

We also tested whether the C-terminal portion of the cleaved HT motif of the exported protein *P. falciparum* Erythrocyte Membrane Protein 3 (PfEMP3), was sufficient



**Fig. 1.** Cleavage of the HT motifs in PFI1755c and PfEMP3 releases a C-terminal fragment that is sufficient to mediate protein export.

A. Images of a parasite expressing wild-type PFI1755c<sub>1-61</sub>:GFP. The left- and right-hand images show GFP localization and a phase contrast image respectively. A diagram illustrating the construct type is also shown. The amino acid sequence of the protein adjacent to the protease cleavage site is written using the single letter amino acid code. The three residues preceding the protease cleavage site are underlined. TM, transmembrane segment; GFP, green fluorescent protein. Scale bar, 2 µm.

B. Diagram illustrating the proteolytic cleavage of PFI1755c<sub>1-61</sub>:GFP by plasmepsin V and the self-cleavage of SS:CapProt:PFI1755c<sub>49-61</sub>:GFP. Illustrations are as in Fig. 1A. SS, signal sequence.

C. Images of a parasite expressing SS:CapProt:PFI1755c<sub>49-61</sub>:GFP. The left- and right-hand images show GFP localization and a phase contrast image respectively. A diagram illustrating the construct type is also shown (as for Fig. 1A). CP, capsid protease.

D. Quantification of GFP export of PFI1755c<sub>1-61</sub>:GFP and SS:CapProt:PFI1755c<sub>49-61</sub>:GFP constructs. The percentage of GFP fluorescence localized in the erythrocyte is shown. The GFP-tagged proteins expressed in each parasite are indicated. Error bars show standard deviation.

E. Anti-GFP Western blot of parasites expressing PFI1755c<sub>1-61</sub>:GFP, SS:CapProt:PFI1755c<sub>49-61</sub>:GFP or a similar construct in which the active site of the capsid protease is mutated (Ser138 to Ala). The constructs are indicated below the blot. Triangle: SS:CapProt:PFI1755c<sub>49-61</sub>:GFP that has not been processed by the capsid protease. Closed arrow: protein that has been processed by either plasmepsin V or capsid protease. Asterisk: GFP degradation product. Lane 1 contains material from untransfected parasites.

F. Averaged tandem mass spectrum for the peptide N-acetyl-SEPVVEEQDLKK, derived from PFI1755c<sub>1-61</sub>:GFP. Mean b and y ion intensities for two experiments are in blue and red respectively. Standard deviation of triplicate measurements for each ion intensity, normalized to the most abundant ion, is shown. \*, neutral loss of ammonia; Δ, neutral loss of water.

G. Images of a parasite expressing SS:CapProt:PFI1755c<sub>49-61</sub>:GFP in which the capsid protease domain contains the S138A active site mutation.

H-I. Images of parasites expressing GFP-tagged PfEMP3. GFP image panels, phase contrast panels and illustrations are arranged as in Fig. 1A. Wild-type PfEMP3<sub>1-82</sub>:GFP (H); SS:CapProt:PfEMP3<sub>63-82</sub>:GFP (I).

J. Quantification of GFP export of PfEMP3 constructs. The percentage of GFP fluorescence localized in the erythrocyte is shown. The GFP-tagged proteins expressed in each parasite are indicated. Error bars show the standard deviation.

K. Anti-GFP Western blot of PfEMP3<sub>1-82</sub>:GFP and SS:CapProt:PfEMP3<sub>63-82</sub>:GFP; labelled as in (E). Lane 1 contains material from untransfected parasites.

to mediate protein export. The HT motif of PfEMP3 (<sup>60</sup>RSLAQ<sup>64</sup>) is cleaved by plasmepsin V to yield P1' P2' residues, Ala Gln (Boddey *et al.*, 2010). A protein comprising residues 1–82 of PfEMP3 fused to GFP (PfEMP3<sub>1-82</sub>:GFP) (Knuepfer *et al.*, 2005b) was efficiently exported to the host erythrocyte (Fig. 1H). An ER-targeted capsid protease fused to residues 63–82 of PfEMP3, and tagged C-terminally with GFP (SS:CapProt:PfEMP3<sub>63-82</sub>:GFP) was also exported into the host erythrocyte with similar efficiency (Fig. 1I and J). This shows that the PfEMP3 C-terminal fragment normally generated by HT motif cleavage is sufficient to mediate protein export. Western blotting indicated that both PfEMP3<sub>1-82</sub>:GFP and SS:CapProt:PfEMP3<sub>63-82</sub>:GFP were proteolytically processed as expected (Fig. 1K, closed arrow, lanes 2 and 3 respectively).

Together, these results show that the C-terminal fragments of the cleaved HT motifs of PFI1755c and PfEMP3 are sufficient to mediate efficient protein export.

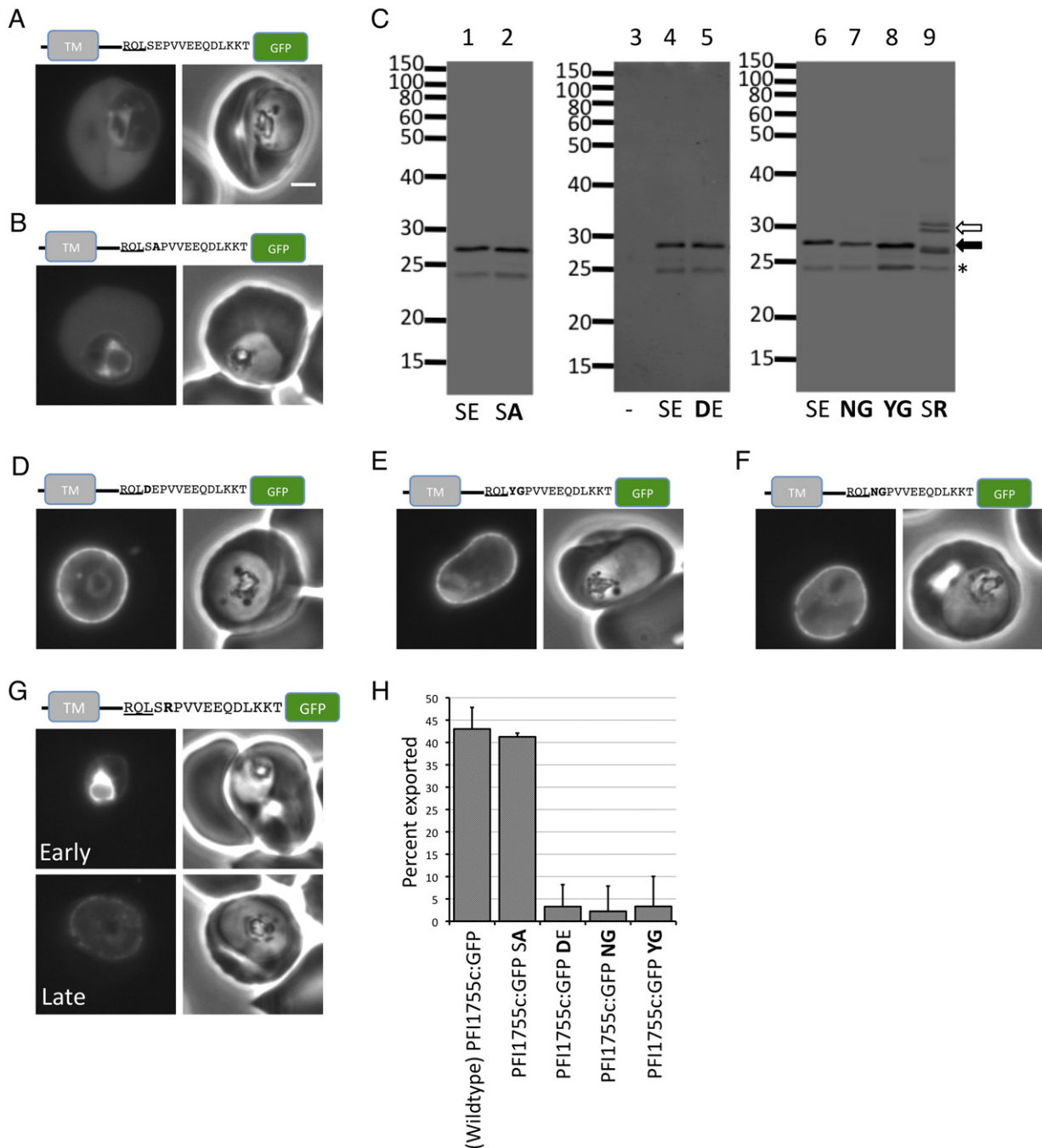
*The identity of the P1' and P2' residues of the HT motif influences the efficiency of protein export and plasmepsin V cleavage*

We subsequently tested whether the new N-terminus generated by HT motif cleavage is necessary for protein export. Specifically, we investigated whether mutations in the P1' and P2' positions of PFI1755c<sub>1-61</sub>:GFP affected protein export efficiency.

Mutation of the P2' position in PFI1755c<sub>1-61</sub>:GFP from Glu to Ala did not block protein export (Fig. 2B and H) and did not significantly reduce the efficiency of plas-

mepsin V processing (Fig. 2C, closed arrow, lane 2). Mutation of the P1' position of the HT motif in PFI1755c to Asp blocked protein export and led to protein accumulation in the PV (Figs 2D, H and S1D). Plasmepsin V processing was not affected by this mutation as high molecular weight unprocessed protein was not detected (Fig. 2C, lane 5). Tryptic peptides from PFI1755c<sub>1-61</sub>:GFP with the P1' position mutated to Asp were analysed by mass spectrometry. The most N-terminal peptide identified was DEPVVEEQDLK(K), indicating that the protein was correctly processed by plasmepsin V. Both N-acetylated and non-N-acetylated peptides were identified (Table 1).

We also tested a double mutant in which the P1' and P2' positions were mutated to Tyr and Gly. This mimics the P1' and P2' amino acids in the HT motif of RESA. Although RESA is exported via a dense granule pathway in early parasites, it is not exported when expressed from a trophozoite stage promoter (Rug *et al.*, 2004). In the context of PFI1755c<sub>1-61</sub>:GFP, expressed from the calmodulin promoter, this mutation indeed blocked protein export (Fig. 2E and H), leading to an accumulation of GFP in the PV (Fig. S1E). No high molecular weight bands were detected by Western blotting suggesting that this mutation did not affect processing by plasmepsin V (Fig. 2C, lane 8). Tryptic peptides, derived from PFI1755c<sub>1-61</sub>:GFP with the P1' and P2' positions mutated to Tyr–Gly, were analysed by mass spectrometry. The most N-terminal peptide identified corresponded to the sequence YGPVVEEQDSK(K). In this case, only non-N-acetylated peptide was identified. As with other proteins analysed the N-terminus was not entirely homogeneous (Table 1). An efficient block in protein export



**Fig. 2.** Mutations in the P1' and P2' positions in the cleaved HT motif influence the efficiency of protein export.

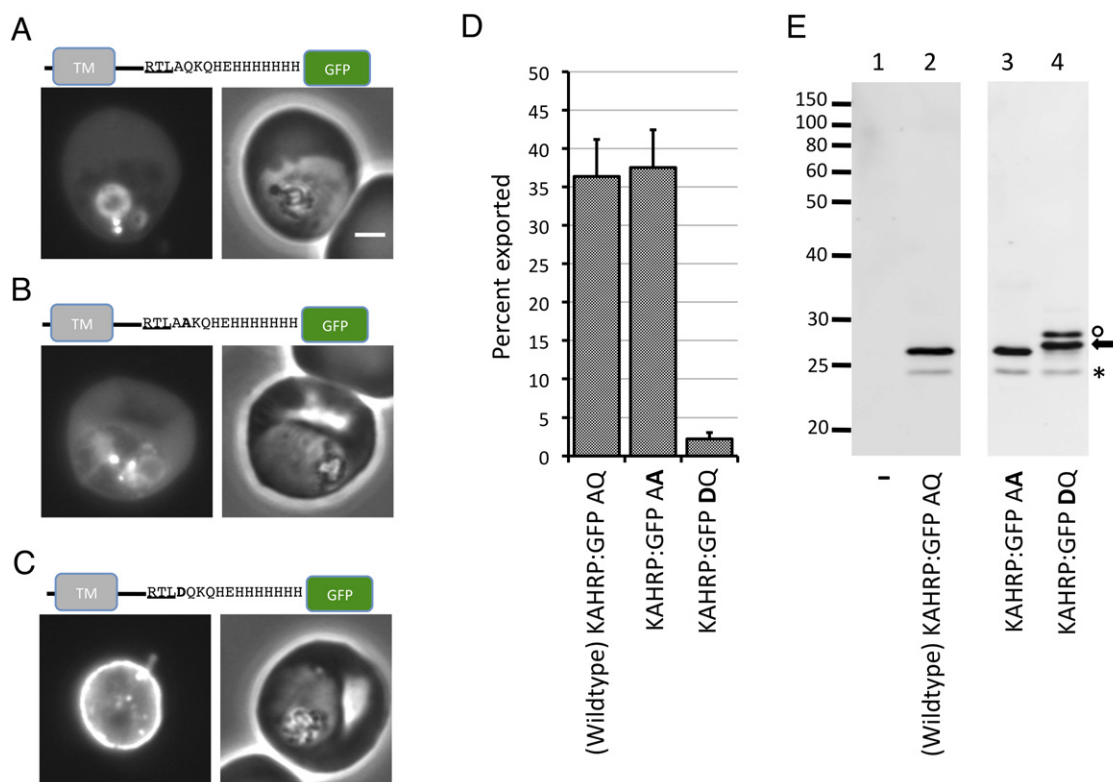
A–B. Localization of PFI1755c<sub>1–61</sub>:GFP proteins. Image panels and illustrations are as in Fig. 1. Scale bar, 2 µm. Wild-type PFI1755c<sub>1–61</sub>:GFP (A); P2' position is mutated to Ala (B).

C. Anti-GFP Western blot of PFI1755c<sub>1–61</sub>:GFP proteins containing mutations in the P1' and P2' positions. The amino acid sequence of the P1' and P2' positions of each protein is indicated below each lane; mutated residues are in bold. Open arrow: doublet of high molecular weight protein bands that have not been processed by plasmepsin V. Closed arrow: protein that has been processed by plasmepsin V. Asterisk: GFP degradation product. Lane 3 contains material from untransfected parasites.

D–F. Localization of PFI1755c<sub>1–61</sub>:GFP proteins containing mutations in the P1' and P2' positions. Image panels and illustrations are as in Fig. 1. P1' position mutated to Asp (D); P1' and P2' positions are mutated to Tyr–Gly (E) or Asn–Gly (F).

G. Localization of PFI1755c<sub>1–61</sub>:GFP with the P2' position mutated to Arg. In early parasites GFP accumulated within the parasite (upper panel). In late parasites GFP was distributed around the parasite periphery but the fluorescence intensity was low (lower panel).

H. Quantification of GFP export in parasites expressing PFI1755c<sub>1–61</sub>:GFP with either wild-type or mutated HT motifs. The percentage of GFP fluorescence localized in the erythrocyte cytoplasm is shown. The GFP-tagged proteins expressed in each parasite are indicated. Error bars show standard deviation.



**Fig. 3.** Mutations in the P1' and P2' positions of GFP-tagged KAHRP block protein export.

A–C. Localization of KAHRP<sub>1–69</sub>:GFP proteins containing mutations in the P1' and P2' positions. Image panels and illustrations are as in Fig. 1. Scale bar, 2  $\mu$ m. Wild-type KAHRP<sub>1–69</sub>:GFP (A); P2' position mutated from Gln to Ala (B); P1' position mutated to Asp (C).

D. Quantification of GFP export in parasites expressing KAHRP<sub>1–69</sub>:GFP with either wild-type or mutated HT motifs. The percentage of GFP fluorescence localized in the erythrocyte cytoplasm is shown. The GFP-tagged protein expressed in each parasite is indicated. Error bars show standard deviation.

E. Anti-GFP Western blot of KAHRP<sub>1–69</sub>:GFP containing mutations in the P1' and P2' positions. Below the blot, the construct type is indicated and followed by the residues in the P1' and P2' positions; mutated residues are in bold. Open circle: alternately processed form of the protein. Closed arrow: protein that has been processed by plasmepsin V. Asterisk: GFP degradation product. Lane 1 contains material from untransfected parasites.

was also observed if the P1' and P2' positions were mutated to Asn–Gly (Figs 2F, H and S1F). This protein was processed by plasmepsin V, as judged by Western blotting (Fig. 2C, closed arrow, lane 7).

Although mutation of the P2' position to Ala did not affect plasmepsin V processing, mutation to Arg did reduce plasmepsin V cleavage efficiency as several higher molecular weight bands were observed by Western blotting of this mutant (Fig. 2C, open arrow, lane 9). In early parasites GFP fluorescence was found within the parasite (Fig. 2G, upper panel), likely in the ER, as seen previously for HT motif cleavage mutants (Boddey *et al.*, 2009). In later parasites, GFP accumulated in the parasite periphery (Fig. 2G, lower panel). Exported GFP was not detectable but fluorescence levels were very low in late parasites.

In summary, we have shown that single mutation of the P1' residue to Asp, and double mutation of the P1' P2' residues to Tyr–Gly or Asn–Gly, blocked protein export at a step downstream of plasmepsin V processing. Further-

more, we have shown that the P2' position can influence cleavage of the HT motif by plasmepsin V. Together these data show that the P1' and P2' residues are necessary for efficient protein export.

#### *The P1' and P2' positions in the HT motifs of KAHRP and PfEMP3 are important for efficient protein export*

We subsequently tested the effects of P1' and P2' mutations in other model proteins. Specifically, we tested whether the P1' and P2' residues in the HT motifs of the exported proteins, Knob-Associated Histidine Rich Protein (KAHRP) and PfEMP3, were also important for protein export. KAHRP has the HT motif <sup>54</sup>RTLAQ<sup>58</sup> and as shown previously (Knuepfer *et al.*, 2005a), a chimera of residues 1–69 fused to GFP was efficiently exported (Fig. 3A). Mutation of the KAHRP P2' Gln to Ala had no effect on protein export (Fig. 3B and D). Proteolytic processing of this mutant by plasmepsin V was similar to that of the

wild-type protein (Fig. 3E, closed arrow, compare lanes 2 and 3). When the P1' Ala residue was mutated to Asp, GFP accumulated in the PV (Figs 3C, D and S2C). Western blotting of this mutant revealed protein that had been correctly processed by plasmepsin V but an additional higher molecular weight band was also observed (Fig. 3E, lane 4, open circle). The mass of this band appears smaller than would be expected for the unprocessed protein. Therefore, we postulate that it may be the result of a cleavage event at an alternate site. Nonetheless, the effects of these P1' mutations are consistent with our observations for equivalent mutations in the HT motif of PFI1755c.

Mutations were also introduced into the P1' and P2' positions of PfEMP3<sub>1-82</sub>:GFP. PfEMP3<sub>1-82</sub>:GFP was efficiently exported into the host erythrocyte (Fig. 4A). Mutation of the P1' Ala to Asp (Figs 4B, E and S3C) or mutation of the P1' and P2' positions to Tyr-Gly (Figs 4C, E and S3D) led to GFP being predominantly localized in the PV. As seen with other exported GFP proteins, PfEMP3<sub>1-82</sub>:GFP with the P2' Gln mutated to Ala was efficiently exported to the host erythrocyte cytosol (Fig. 4D and E).

Some accumulation of high molecular weight bands, consistent with reduced plasmepsin V processing, was seen in Western blots of several PfEMP3 mutants, most notably when the P1' amino acid is mutated to Asp (Fig. 4F, open arrow, lane 3). However, in all cases the predominant protein species appeared to be correctly processed.

Our analyses show that the effects of mutations of the P1' and P2' residues are consistent across a range of model proteins. Together with our experiments using PFI1755c, these experiments also show that the identity of the amino acids in the P1' and P2' positions is important for a key targeting step which is likely downstream of plasmepsin V processing.

*Proteins with N-terminal sequences generated by plasmepsin V or capsid protease cleavage are exported via the same pathway*

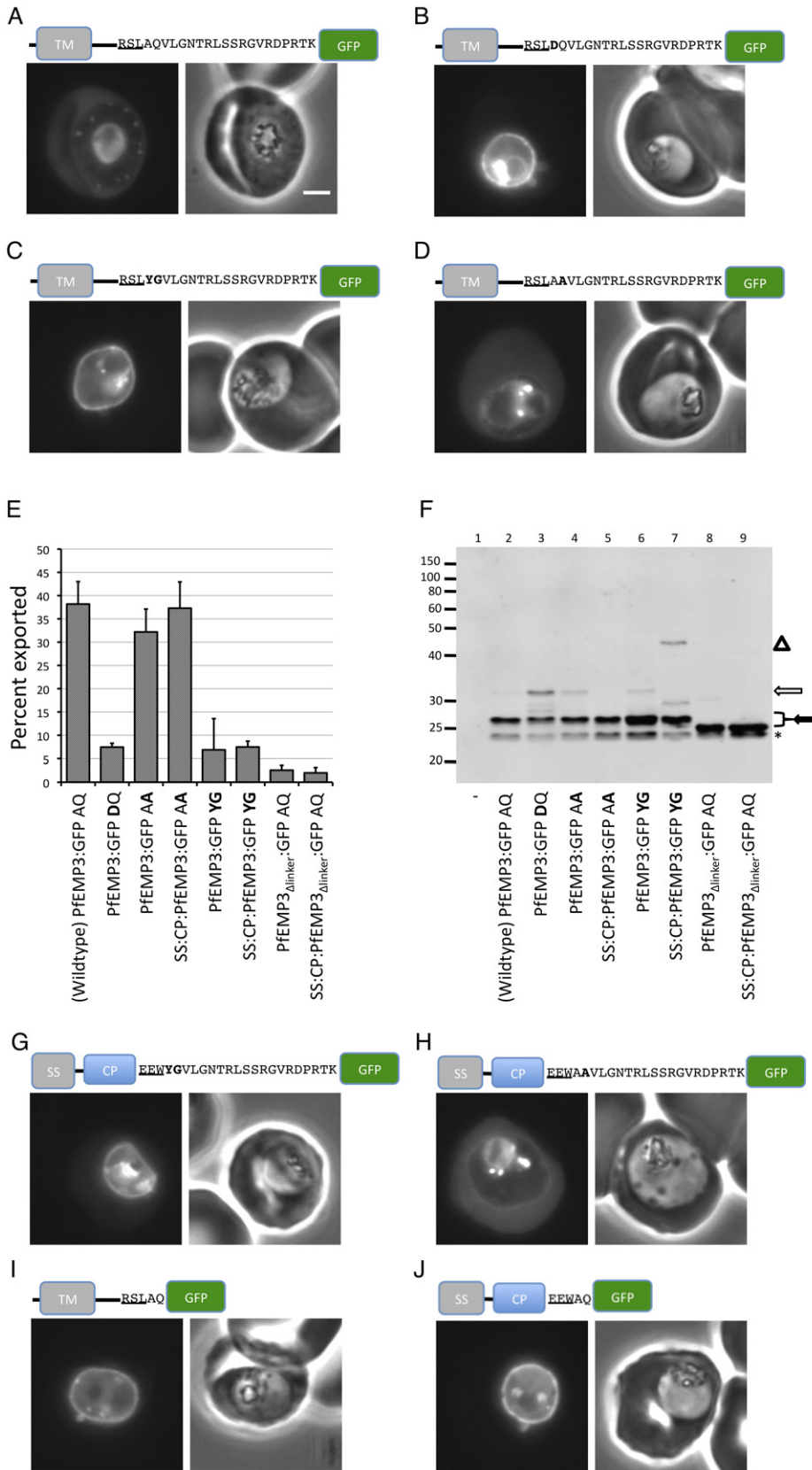
In the following experiments, P1' and P2', mutations were introduced into model proteins in which the N-terminus of an exported protein was generated by capsid protease cleavage. First, this allowed us to establish whether the capsid protease-based model proteins and exported proteins processed by plasmepsin V were similarly affected by P1' and P2' mutations and hence, exported via the same export pathway. This is a key conclusion as it is possible that the capsid protease-based constructs are exported via a different export pathway from that used by proteins processed by plasmepsin V. Only if the capsid protease-based constructs are exported by the same export

pathway as proteins processed by plasmepsin V can we conclude that the cleaved C-terminal fragment of the HT motif is sufficient to mediate export of HT motif-containing proteins. Second, as the capsid protease-based proteins are not processed by plasmepsin V, we were able to clearly establish whether the amino acids in the P1' and P2' positions in the HT motif are important for steps in the export pathway that occur downstream of plasmepsin V cleavage.

Mutations were introduced into the P1' and P2' positions in the context of SS:CapProt:PFI1755c<sub>49-61</sub>:GFP. Mutation of the P1' Ser of SS:CapProt:PFI1755c<sub>49-61</sub>:GFP to Asp efficiently blocked export of this construct leading to an accumulation of GFP in the PV (Figs 5A and S1G). Likewise, mutation of the P1' P2' sequence to Tyr-Gly, also blocked protein export (Fig. 5B) leading to PV accumulation of GFP (Fig. S1H). Quantification of the block in export that occurred for each of these mutations shows that they behave similarly in proteins processed by both capsid protease and plasmepsin V (Fig. 5D). Western blotting of the proteins indicated that they were correctly processed and cleaved efficiently by the capsid protease (Fig. 5E, closed arrow, P1' P2' Tyr-Gly, lane 3; P1' Asp, lane 4); only weak bands corresponding to uncleaved protein were observed (Fig. 5E, triangle). When the P2' Glu was mutated to Ala in the context of SS:CapProt:PFI1755c<sub>49-61</sub>:GFP, the protein was efficiently exported (Fig. 5C and D). The protein appeared to be correctly processed as assessed by Western blotting (Fig. 5E, closed arrow, lane 5). We performed mass spectrometry analysis of tryptic peptides derived from SS:CapProt:PFI1755c<sub>49-61</sub>:GFP in which the P1' position was mutated to Asp, or the P1' and P2' positions were mutated to Tyr-Gly. As with the equivalent mutations in the context of PFI1755c<sub>49-61</sub>:GFP, the most N-terminal peptides corresponded to DEPVVEEQDLK(K) and YGPVVEEQDSK(K) respectively (Table 1). N-acetylated and non-N-acetylated versions of the DEPVVEEQDLK(K) peptide were identified but only the non-acetylated YGPVVEEQDSK(K) peptide was detected. Similarly, analysis of the SS:CapProt:PFI1755c<sub>49-61</sub>:GFP with the P2' position mutated to Ala revealed that the most N-terminal peptide corresponded to the peptide N-acetyl-SAPVVEEQDLK(K). Some heterogeneity of the N-terminal sequences was seen particularly when the P1' and P2' residues were mutated to Tyr-Gly (Table 1).

We also introduced mutations that blocked export of PfEMP3<sub>1-82</sub>:GFP into SS:CapProt:PfEMP3<sub>63-82</sub>:GFP. An efficient block in export was observed when the P1' and P2' residues were mutated to Tyr and Gly respectively (Fig. 4G). The mutant protein accumulated in the PV (Fig. S3E). Quantification of the amount of GFP exported into the erythrocyte cytoplasm showed that this mutation behaved similarly in constructs cleaved by either capsid





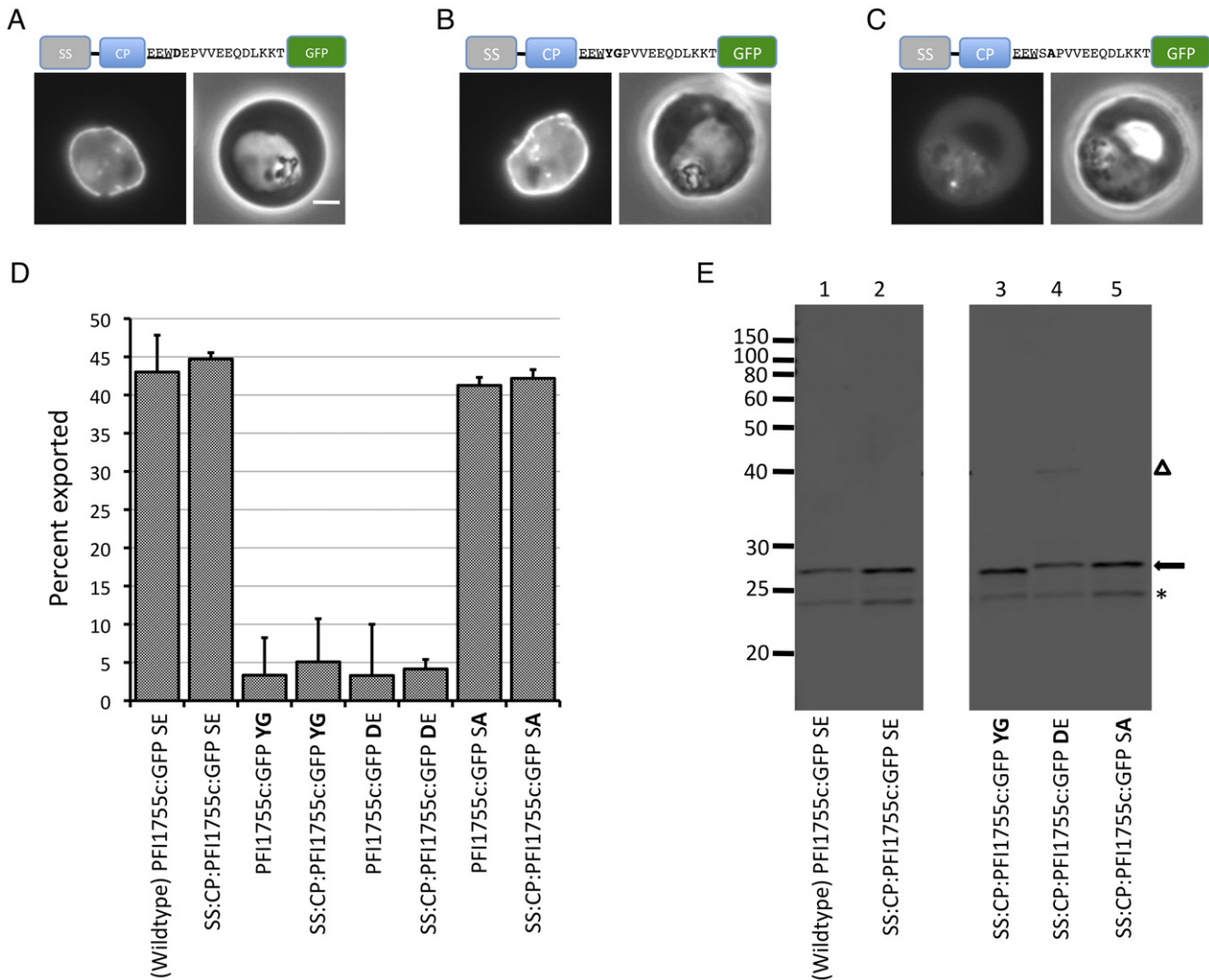
**Fig. 4.** Mutations in the P1' and P2' position of PfEMP3 block protein export.

A–D. Localization of PfEMP3<sub>1–82</sub>:GFP proteins containing mutations in the P1' and P2' positions. Image panels and illustrations are as in Fig. 1. Scale bar, 2 μm. Wild-type PfEMP3<sub>1–82</sub>:GFP (A); P1' position mutated to Asp (B); P1' and P2' residues mutated to Tyr–Gly (C); P2' position mutated to Ala (D). E. Quantification of GFP export in parasites expressing PfEMP3<sub>1–82</sub>:GFP or SS:CapProt: PfEMP3<sub>63–82</sub>:GFP proteins. The percentage of GFP fluorescence localized in the erythrocyte cytoplasm is shown. The GFP-tagged proteins expressed in each parasite are indicated. Error bars show standard deviation.

F. Anti-GFP Western blot of GFP-tagged PfEMP3 proteins, annotated as in Fig. 3E. Triangle: SS:CapProt: PfEMP3<sub>63–82</sub>:GFP that has not been processed by the capsid protease. Open arrow: high molecular weight protein bands that have not been processed by plasmepsin V. Closed arrow: protein that has been processed by either plasmepsin V or capsid protease. Asterisk: GFP degradation product. Lane 1 contains material from untransfected parasites.

G–H. Localization of SS:CapProt: PfEMP3<sub>63–82</sub>:GFP proteins containing mutations in the P1' or both the P1' and P2' positions. Image panels and illustrations are as in Fig. 1. P1' and P2' residues mutated to Tyr–Gly (G); P1' residue mutated to Ala (H).

I–J. Localization of PfEMP3 constructs in which the linker following the HT motif is deleted. PfEMP3<sub>63–64</sub>:GFP (I); SS:CapProt: PfEMP3<sub>63–64</sub>:GFP (J).



**Fig. 5.** Block in export by mutations in the P1' and P2' positions of proteins cleaved by the capsid protease.

A–C. Localization SS:CapProt:PFI1755c<sub>49–61</sub>:GFP containing mutations in the P1' and P2' positions. Image panels and illustrations are as in Fig. 1. Scale bar, 2 µm. P1' position mutated to Asp (A); P1' and P2' positions mutated to Tyr–Gly (B); P2' position is mutated to Ala (C).

D. Quantification of GFP export in parasites expressing wild-type and mutated versions of PFI1755c<sub>49–61</sub>:GFP or SS:CapProt:PFI1755c<sub>49–61</sub>:GFP. The percentage of GFP fluorescence localized in the erythrocyte cytoplasm is shown. The GFP-tagged proteins expressed in each parasite are indicated (for comparison the efficiency of protein export of similar mutations introduced into PFI1755c<sub>49–61</sub>:GFP are also shown, as in Figs 1 and 2). Error bars show standard deviation.

E. Anti-GFP Western blot of parasites expressing PFI1755c<sub>49–61</sub>:GFP or SS:CapProt:PFI1755c<sub>49–61</sub>:GFP and similar constructs in which the P1' and P2' residues are mutated. The constructs, followed by the residues in the P1' and P2' positions, are indicated below the blot; mutated residues are in bold. Triangle: SS:CapProt:PFI1755c<sub>49–61</sub>:GFP that has not been processed by the capsid protease. Closed arrow: protein that has been processed by either plasmepsin V or capsid protease. Asterisk: GFP degradation product.

protease or plasmepsin V (Fig. 4E). Western blotting of SS:CapProt:PfEMP3<sub>63–82</sub>:GFP with the P1' P2' Tyr–Gly mutation indicated that the majority of the protein was correctly processed by the capsid protease; only weak bands corresponding to uncleaved protein were observed (Fig. 4F, triangle, lane 7). Robust export of SS:CapProt:PfEMP3<sub>63–82</sub>:GFP was observed when the P2' Gln was mutated to Ala (Fig. 4H and E); this mutation therefore behaved in a similar manner in substrates cleaved by either capsid protease or plasmepsin V. Western blotting of SS:CapProt:PfEMP3<sub>63–82</sub>:GFP with P2'

Gln mutated to Ala indicated efficient processing of the mutant protein (Fig. 4F, lane 5).

It has previously been shown that close proximity of the HT motif to a tightly folded GFP domain can prevent protein export (Knuepfer *et al.*, 2005b). PfEMP3<sub>1–82</sub>:GFP, has an 18 residue linker between the HT motif and the GFP domain. A GFP-tagged PfEMP3 construct was generated that lacked this linker, such that the N-terminal P1' and P2' residues (Ala Gln) directly preceded the GFP domain (PfEMP3<sub>1–64</sub>:GFP). This protein contained the wild-type HT motif but was not exported (Fig. 4I), and

accumulated in the PV (Fig. S3F). When the same N-terminus was generated in the context of the capsid protease SS:CapProt:PFEMP3<sub>63-64</sub>:GFP, a block in export was also observed (Fig. 4J and E) leading to an accumulation of GFP in the PV (Fig. S3G). Both PFEMP3<sub>1-64</sub>:GFP and SS:CapProt:PFEMP3<sub>63-64</sub>:GFP were proteolytically processed, as assessed by Western blotting (Fig. 4F, lanes 8 and 9 respectively). These results show that the block in export observed as a consequence of proximity of the wild-type HT motif to the GFP domain, occurs when the P1' P2' residues are generated by plasmepsin V or the capsid protease.

In summary, diverse mutations in the HT motif P1' and P2' positions and the following linker region, behaved in a similar manner in the context of two model proteins processed by either plasmepsin V or by the capsid protease. This shows that whether their new N-termini were generated by plasmepsin V or by the capsid protease, the proteins were trafficked via the same export pathway. Importantly, this confirms that the capsid protease constructs are not exported by a non-HT motif dependent export pathway. Furthermore, this allows us to conclude that following plasmepsin V cleavage, it is the C-terminal portion of the HT motif that mediates protein export.

#### *The context of the HT motif can influence protein export*

The preceding results show that the C-terminal fragment of the cleaved HT motif is both necessary and sufficient to mediate export and that the P1' and P2' residues of the cleaved HT motif are important for protein export. We sought to determine whether additional amino acids within the linker region between the HT motif and GFP domain of PFI1755c<sub>1-61</sub>:GFP also influenced protein export. To do this we used a PFI1755c<sub>1-61</sub>:GFP construct that contained the P2' position mutated from E to A, and introduced multiple mutations into the following linker region. As we show later, this model protein was more sensitive to mutations within the linker region, and hence yielded more robust mutant phenotypes, than the model protein with an E in the P2' position. These experiments allowed us to show that the linker sequence following the HT motif can affect export efficiency but also that the identity of the P2' amino acid of the HT motif can be particularly important for efficient export in certain sequence contexts.

In the model protein PFI1755c<sub>1-61</sub>:GFP, mutation of the P2' Glu to Ala did not affect protein export (Fig. 6B and H). After HT motif cleavage, the N-terminus of PFI1755c<sub>1-61</sub>:GFP with P2' Glu to Ala has the sequence **SAPVVEE**QDLKKT; the mutated residue is in bold. In the context of this mutant, we mutated multiple residues downstream of the P1' and P2' position of the HT motif. Export was not greatly affected in mutants whose N-terminus, after plasmepsin V cleavage, corresponded to the

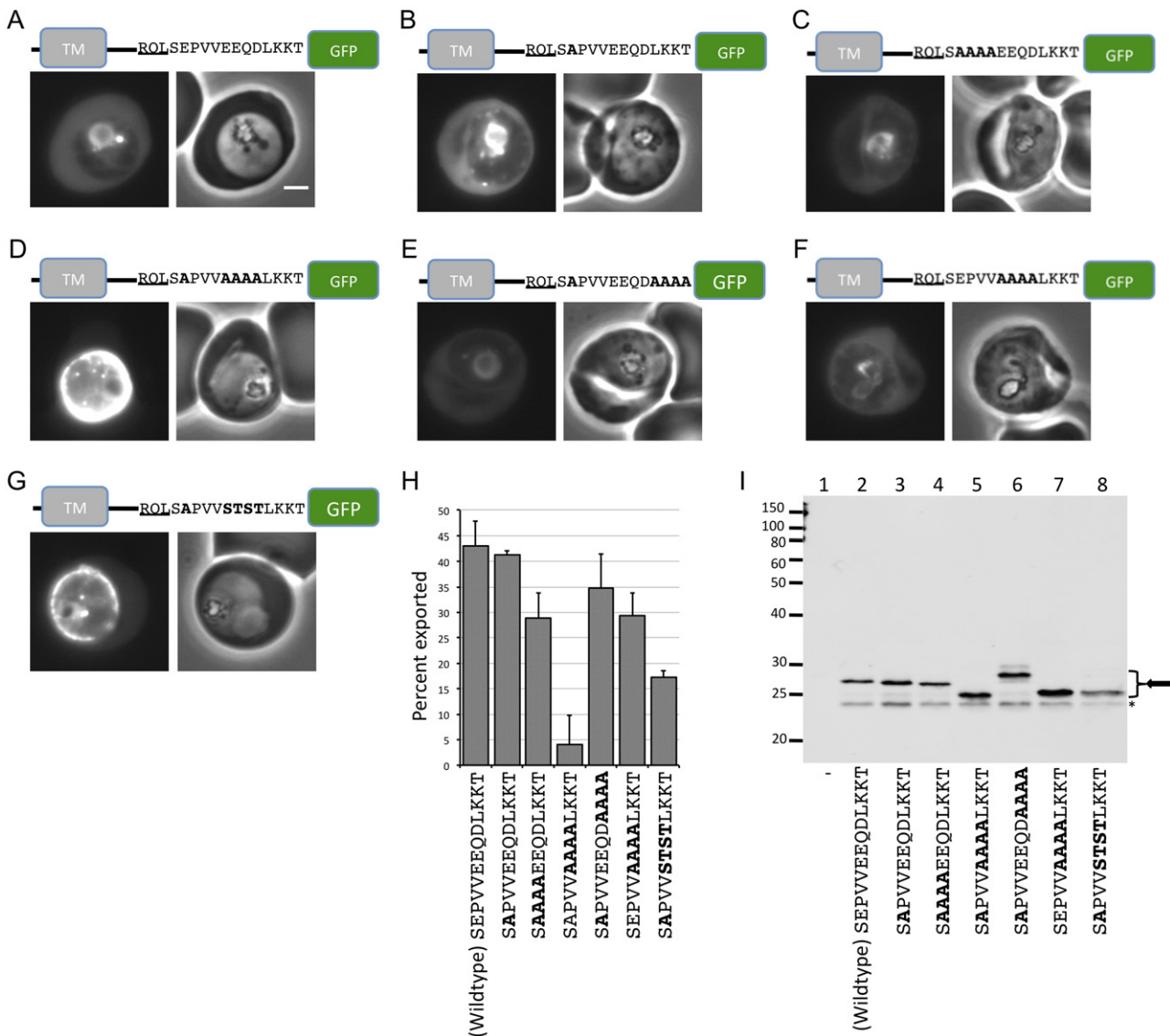
sequence **SAAAAE**QDLKKT or **SAPVVEE**Q**AAAA** (Fig. 6C and E respectively, and Fig. 6H). However, a PFI1755c<sub>1-61</sub>:GFP protein whose new N-terminus after HT motif processing was **SAPVVA**AAALKKT, was not efficiently exported (Fig. 6D and H) and accumulated in the PV (Fig. S1J). When the same mutation was introduced into a construct with a wild-type HT motif i.e. with the new N-terminal sequence **SEPVV**AAALKKT, GFP was robustly exported (Fig. 6F and H). A small amount of GFP was retained within these parasites. However, the robust export of this construct indicates that the EEQD to AAAA mutation alone does not generate a protein refractory to export. We also tested an alternative construct with the linker region **SAPVVST**STLKKT, in which these residues were mutated to polar amino acids. In this mutant, a low level of GFP fluorescence was detectable in the host cell cytosol but the protein was predominantly localized in the PV (Figs 6G, H and S1K). When analysed by Western blotting, it was clear that many of the PFI1755c<sub>1-61</sub>:GFP proteins with AAAA mutations within the linker region exhibited altered electrophoretic mobility relative to the wild-type PFI1755c<sub>1-61</sub>:GFP and P2' Glu to Ala mutants (Fig. 6I, closed arrow).

To determine the identity of the N-terminal residues, we immunoprecipitated proteins with the linker sequences **SAPVVA**AAALKKT and **SAPVVST**STLKKT and analysed peptides using mass spectrometry. For the PFI1755c<sub>1-61</sub>:GFP with the linker sequence **SAPVVA**AAALKKT, the most N-terminal peptide detected had the sequence **SAPVVA**AAALK(K). Only N-acetylated peptides were identified (Table 1). The most N-terminal peptides identified for PFI1755c<sub>1-61</sub>:GFP with the linker region **SAPVVST**STLKKT started with the sequence **SAPVVST**STLK(K). Again, only N-acetylated peptides were identified (Table 1). These data confirm that in both cases, HT motif cleavage yielded proteins with the expected N-termini.

These data show that the amino acid sequence downstream of the HT motif can influence protein export and that the importance of the amino acid in the P2' position of the HT motif is dependant upon the downstream sequence. As the protein with the new N-terminus **SAPVVA**AAALKKT is processed by plasmepsin V and N-acetylated we also conclude that neither plasmepsin V cleavage *per se*, nor an N-acetylated N-terminus is sufficient for efficient protein export.

## Discussion

The malaria parasite exports many proteins into its host erythrocyte. The majority of exported proteins contain a HT motif that has the consensus RxLxE/D/Q, which is cleaved in the ER by the protease plasmepsin V. A number of models have been proposed to explain the contribution of the HT motif to protein export. The only



**Fig. 6.** Mutations beyond the P1' and P2' positions influence the export efficiency of GFP-tagged PF1755c.

A–G. Images of parasites expressing PF1755c<sub>1–61</sub>:GFP with mutations in the P2' position and in the linker sequence. Image panels and illustrations are as in Fig. 1. Scale bar, 2 µm. PF1755c<sub>1–61</sub>:GFP (A); P2' position mutated to Ala (B); linker region following the HT motif cleavage site mutated to SAAAAEEQDLKKT (C), SAPVAAAALKKT (D), SAPVVEEQDAAAA (E); SEPVAAAALKKT (F) or SAPVSTSTLKKKT (G).

H. Quantification of GFP export in parasites expressing PF1755c<sub>1–61</sub>:GFP containing mutations in the P2' residue and following sequence. The percentage of GFP fluorescence localized in the erythrocyte cytoplasm is shown. The GFP-tagged proteins expressed in each parasite are indicated. Error bars show standard deviation.

I. Anti-GFP Western blot of GFP-tagged PF1755c<sub>1–61</sub>:GFP protein containing mutations in the P2' residue and following sequence. For each construct, the protein sequence following the HT motif cleavage site is indicated below the blot. Closed arrows: protein that has been processed by plasmepsin V. Asterisk: GFP degradation product. Lane 1 contains material from untransfected parasites.

mutations shown to completely block protein export are in the first and third positions of the HT motif (Arg and Leu respectively) and block processing by plasmepsin V.

We show that the C-terminal portion of the cleaved HT motif is sufficient to mediate protein export. Proteins lacking a HT motif but whose new N-termini resemble the cleaved HT motif, generated by the capsid protease, not plasmepsin V, were efficiently exported. Our results differ

from previous work showing that a protein whose N-terminus resembled the cleaved HT motif was not exported (Boddey *et al.*, 2010). In the latter experiments, signal peptidase was used to generate a protein whose N-terminus resembled the cleaved HT motif. In the context of these experiments it is likely that the capsid protease generates the N-terminus with greater accuracy than signal

peptidase cleavage is uncomplicated and cleavage may occur at multiple sites proximal to the signal peptide (Kotia and Raghani, 2010). By using an exogenous protease we have bypassed the requirement for plasmepsin V cleavage. This shows that plasmepsin V cleavage is not directly coupled to the hand-off of a protein to the next step in the export pathway as proposed previously (Boddey *et al.*, 2010). Rather, our data support a model in which the role of plasmepsin V is to cleave off the N-terminal transmembrane domain and generate a new N-terminal sequence that is sufficient to function as a signal for protein export. Notably, the ability of this sequence to mediate export is blocked if it is preceded by a folded domain (such as the capsid protease domain with an active site mutation). This suggests that the free N-terminus of this sequence is also an important feature of the signal for export.

Our exported proteins whose N-terminal export sequences were generated by the capsid protease lacked part of the HT motif (RxL). This region, with some contribution from the P2' position, has been implicated in PI(3)P-binding in the ER lumen (Bhattacharjee *et al.*, 2012b). Robust export of proteins whose N-terminal export sequences were generated by the capsid protease suggests that recognition of the new N-terminal export sequence by putative export machinery may occur downstream of the PI(3)P sorting event.

Mutagenesis of the P1' P2' amino acids showed that this region is necessary for protein export. Mutation of the P1' position to Asp blocked protein export. Furthermore, mutation of the P1' and P2' positions to Tyr–Gly or Asn–Gly also blocked export. Likely a number of other double mutations in the P1' and P2' positions of the HT motif will also block protein export. Although our experiments clearly show that the identity of the P2' amino acid is important for export, surprisingly mutation of the P2' amino acid to Ala alone, did not block protein export; a result also seen by others (Grüring *et al.*, 2012; Bhattacharjee *et al.*, 2012a). Some studies have seen a decrease in export of reporter proteins after mutation of the P2' position to Ala (Przyborski *et al.*, 2005; Boddey *et al.*, 2009). The reason for this discrepancy is unclear; it is possible that the effects of the P2' Ala mutations are influenced by the sequences that follow the HT motifs of different model proteins. Notably, we show that mutating the PFI1755c P2' residue to Ala reduces the efficiency of protein export but only when combined with further mutations in the linker region following the HT motif. It would be interesting to investigate the contribution to export of the regions downstream of the HT motifs of other exported proteins. Additionally, expression of model proteins from different promoters might also influence whether mutation of the P2' position affects export (Boddey *et al.*, 2009).

We also showed that in the context of PFI1755c<sub>1–61</sub>:GFP, mutation of the P2' residue to Arg reduced the

efficiency of cleavage of the HT motif by plasmepsin V. This is the first instance where a HT motif residue, other than those in the first and third positions, has been shown to influence processing of the HT motif by plasmepsin V. Overall our data show that the P2' residue is important for efficient proteolytic processing of the HT motif and that both the P1' and P2' residues, upon cleavage of the HT motif, form part of the signal that is recognized in a step downstream of plasmepsin V processing. However, as for the signals that mediate protein secretion by bacterial type III and IV secretion systems, there does not appear to be an extensive and rigid consensus sequence for the export signal (Galán and Wolf-Watz, 2006; Alvarez-Martinez and Christie, 2009). Resident PV proteins might avoid engaging with the export apparatus by having structured N-termini, N-terminal residues that do not favour interaction with the export apparatus, or having a tightly folded structure (Gehde *et al.*, 2009).

Importantly, we showed that P1' and P2' mutations had the same effect on export of proteins whose N-termini were generated by plasmepsin V or by the capsid protease. Likewise, in both contexts removing the linker sequence that separates the P1' and P2' residues from the folded GFP domain blocked export. These are key observations as they show that proteins whose N-termini are generated by capsid protease, and only contain the C-terminal portion of the HT motif, follow the same export pathway as proteins that are processed by plasmepsin V. This is important as it allows us to conclude that the C-terminal fragment of the HT motif is both necessary and sufficient to mediate the normal export of proteins that contain a HT motif once they are processed by plasmepsin V. While this paper was under review, Grüring and colleagues (2012), using a different approach, published a study that strongly supports this model. These data indicate that the new N-terminus of exported proteins, generated by plasmepsin V cleavage, is a key feature of the export signal in proteins that contain a HT motif.

A linker region between the HT motif and GFP domain is required for protein export (Knuepfer *et al.*, 2005a). There is little sequence conservation downstream of the HT motif and the linker sequence can be replaced by a stretch of alanines in the context of a wild-type HT motif. However, we have shown that in the context of a mutation in the P2' position of the HT motif, mutation of acidic residues in the sequence downstream of the HT motif can influence protein export efficiency. Therefore, in addition to its role in separating the HT motif from the GFP domain of model proteins, the linker region sequence can also contribute to the export signal. It has previously been shown that mutation of charged residues within exported proteins can influence protein export (Spielmann and Gilberger, 2010). This effect may be particularly important for the efficient export of proteins with suboptimal HT motifs.

Our findings also address the relationship between N-acetylation and protein export. As shown previously (Boddey *et al.*, 2009) we find that N-acetylation of the N-terminus generated by HT motif cleavage is not sufficient to mediate protein export. The PFI1755c mutant with the cleaved N-terminal sequence SAPVVAAAALKKT:GFP was not efficiently exported although it was N-acetylated, localized in the parasite secretory pathway, and had an N-terminal sequence that is of similar length to the N-terminal sequences of exported proteins. It remains unclear whether N-acetylation is required for export. However, we identified non-acetylated peptides from P1' and P2' mutants that were significantly blocked in protein export. Identification of the N-acetyltransferase for exported proteins will be crucial to gain a better understanding of the signal that directs proteins for export.

Export of proteins into the infected erythrocyte is essential for parasite survival and virulence. Our data show that cleavage of the HT motif of exported proteins exposes an N-terminal sequence that, although lacking a strong consensus, contains sufficient information to mediate protein export.

## Experimental procedures

### Plasmids

DNA encoding PFI1755c, PfEMP3 and KAHRP was PCR amplified from *P. falciparum* 3D7 cDNA and cloned into a *P. falciparum* expression plasmid, in frame with 3' GFP. Expression was under the control of pFCAM 5' and pbDT 3' regions. The constructs were named PFI1755c<sub>1-61</sub>:GFP (GenBank ID: JX279925), PfEMP3<sub>1-82</sub>:GFP (GenBank ID: JX279927) and KAHRP<sub>1-69</sub>:GFP (GenBank ID: JX279926) respectively. Similar expression plasmids were used to generate SS:CapProt:PFI1755c<sub>49-61</sub>:GFP (GenBank ID: JX279928), and SS:CapProt:PfEMP3<sub>63-82</sub>:GFP (GenBank ID: JX279929). Semliki forest virus capsid protease gene includes the mutation V149S that does not affect protease activity (numbering as in UniProt entry P03315). The active site capsid protease mutant S138A corresponds to residue 219 in the full-length viral protein (Choi *et al.*, 1997). DNA fusions and mutations were generated using overlapping PCR.

### Parasite culture and transfection

PFI1755c linker mutants, PFI1755c capsid protease, PFI1755c with P1' and P2' amino acids mutated to Asn–Gly, and KAHRP mutants were transfected into 3D7 *attB* parasites using pINT (Nkrumah *et al.*, 2006) as described previously (Deitsch *et al.*, 2001). Other PFI1755c P1', P2' mutants and PfEMP3 constructs were transfected into 3D7 parasites using pHTH (Balu *et al.*, 2005).

### Microscopy

Parasites were imaged by placing a drop of culture material between a coverslide and coverslip. Fluorescence and phase

contrast images were acquired using a Zeiss Axiovert S100 microscope, equipped with a HBO100 lamp, an FDI Photonic Sciences camera, and ImagePro Plus software. Fluorescence images were acquired using identical exposure times; equivalent brightness and contrast settings were applied to images using ImageJ. Parasites were fed one day before imaging. Immunofluorescence imaging was carried out as described in the supplementary information.

Z-stacks of images used for quantification of export efficiency were acquired using a Zeiss Axiovert 200 M microscope and Axiovision software (parasites were randomly selected under phase contrast illumination). Images were deconvolved using Volocity software. Z-stack sum-projections were assembled using ImageJ. Within each projection, the area encompassing parasite and the area encompassing the cytoplasm of the infected erythrocyte, were manually defined. GFP fluorescence intensities of each region were determined using ImageJ. The fraction of protein export was calculated by dividing the amount of GFP fluorescence in the red cell only by the amount of GFP fluorescence in the red cell and the parasite. This fraction is expressed as a percentage. Quantification of export in each parasite line was performed in two independent experiments involving a total of 30–35 randomly selected parasites. The data presented are the combined mean of both experiments and error bars show a combined standard deviation for both experiments.

### Protein immunoprecipitation and Western blotting

For analysis by Western blotting, cells from non-synchronous cultures were lysed in 50 mM Tris pH 8.5, 200 mM NaCl, 2 mM EDTA containing a protease inhibitor cocktail (Roche). Culture volumes were adjusted to obtain approximately equivalent amounts of mid- to late-stage parasites in each sample. Lysates were clarified by centrifugation at 18 500 *g*. Proteins were immunoprecipitated using anti-GFP agarose (MBL), analysed by Western blotting using rabbit anti-GFP antibody (Invitrogen), and a LI-COR Odyssey imager.

### Mass spectrometry

Proteins were immunoprecipitated from 30 ml cultures, processed as above, and SDS-PAGE gels were Coomassie stained. Bands were digested as described previously (Shevchenko *et al.*, 2007). Peptides were separated using a nanoACQUITY UPLC system coupled to a Synapt HDMS mass spectrometer (Pringle *et al.*, 2007) (Waters Corporation). The mass spectrometer was operated in a data-independent mode (LC-MS<sup>E</sup>) of acquisition (Silva *et al.*, 2005). Raw data were processed using PLGS v2.5. Identified peptides that were present in two out of three technical replicates, had a mass error of less than 10 ppm, had more than 30% matched product coverage and were present in replicate experiments are listed in Table 1. Detailed methods are described in the supplementary information.

## Acknowledgements

This work was supported by a Wellcome Trust Grant to AO (09095/Z/10/Z), ISMB start-up grants to AO and KT, and NIH grants (R01 HL069630 and P01 HL078826) to KH. AC is

funded by a BBSRC PhD studentship. MR4 provided malaria parasites (MRA-102, MRA-845) and pINT plasmid (MRA-847) contributed by Daniel Carucci and David Fidock respectively. Anti-Exp2 antibody (7.7) was kindly provided by David Cavanagh and Jana McBride. Capsid protease DNA was kindly provided by Ari Helenius. We thank Anoop Kumar for assistance with microscopy, and Louise Gavigan for preliminary characterization of the PFI1755c<sub>1-61</sub>:GFP P2' Arg mutant.

## References

- Alvarez-Martinez, C.E., and Christie, P.J. (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* **73**: 775–808.
- Balu, B., Shoue, D.A., Fraser, M.J., and Adams, J.H. (2005) High-efficiency transformation of *Plasmodium falciparum* by the lepidopteran transposable element piggyBac. *Proc Natl Acad Sci USA* **102**: 16391–16396.
- Bhattacharjee, S., Speicher, K.D., Stahelin, R.V., Speicher, D.W., and Haldar, K. (2012a) PI(3)P-independent and -dependent pathways function together in a vacuolar translocation sequence to target malarial proteins to the host erythrocyte. *Mol Biochem Parasitol* **185**: 106–113.
- Bhattacharjee, S., Stahelin, R.V., Speicher, K.D., Speicher, D.W., and Haldar, K. (2012b) Endoplasmic reticulum PI(3)P lipid binding targets malaria proteins to the host cell. *Cell* **148**: 201–212.
- Boddey, J.A., Moritz, R.L., Simpson, R.J., and Cowman, A.F. (2009) Role of the *Plasmodium* export element in trafficking parasite proteins to the infected erythrocyte. *Traffic* **10**: 285–299.
- Boddey, J.A., Hodder, A.N., Gunther, S., Gilson, P.R., Patsiouras, H., Kapp, E.A., *et al.* (2010) An aspartyl protease directs malaria effector proteins to the host cell. *Nature* **463**: 627–631.
- Chang, H.H., Falick, A.M., Carlton, P.M., Sedat, J.W., DeRisi, J.L., and Marletta, M.A. (2008) N-terminal processing of proteins exported by malaria parasites. *Mol Biochem Parasitol* **160**: 107–115.
- Choi, H.K., Lu, G., Lee, S., Wengler, G., and Rossmann, M.G. (1997) Structure of Semliki Forest virus core protein. *Proteins* **27**: 345–359.
- Deutsch, K., Driskill, C., and Wellems, T. (2001) Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. *Nucleic Acids Res* **29**: 850–853.
- Galán, J.E., and Wolf-Watz, H. (2006) Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **444**: 567–573.
- Gehde, N., Hinrichs, C., Montilla, I., Charpian, S., Lingelbach, K., and Przyborski, J.M. (2009) Protein unfolding is an essential requirement for transport across the parasitophorous vacuolar membrane of *Plasmodium falciparum*. *Mol Microbiol* **71**: 613–628.
- Goldberg, D.E., and Cowman, A.F. (2010) Moving in and renovating: exporting proteins from *Plasmodium* into host erythrocytes. *Nat Rev Microbiol* **8**: 617–621.
- Grüning, C., Heiber, A., Kruse, F., Flemming, S., Franci, G., Colombo, S.F., *et al.* (2012) Uncovering common principles in protein export of malaria parasites. *Cell Host Microbe* **12**: 717–729.
- Haldar, K., and Mohandas, N. (2007) Erythrocyte remodeling by malaria parasites. *Curr Opin Hematol* **14**: 203–209.
- Hiller, N.L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estrano, C., and Haldar, K. (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306**: 1934–1937.
- Knuepfer, E., Rug, M., and Cowman, A.F. (2005a) Function of the *Plasmodium* export element can be blocked by green fluorescent protein. *Mol Biochem Parasitol* **142**: 258–262.
- Knuepfer, E., Rug, M., Klonis, N., Tilley, L., and Cowman, A.F. (2005b) Trafficking determinants for PfEMP3 export and assembly under the *Plasmodium falciparum*-infected red blood cell membrane. *Mol Microbiol* **58**: 1039–1053.
- Knuepfer, E., Rug, M., Klonis, N., Tilley, L., and Cowman, A.F. (2005c) Trafficking of the major virulence factor to the surface of transfected *P. falciparum*-infected erythrocytes. *Blood* **105**: 4078–4087.
- de Koning-Ward, T.F., Gilson, P.R., Boddey, J.A., Rug, M., Smith, B.J., Papenfuss, A.T., *et al.* (2009) A newly discovered protein export machine in malaria parasites. *Nature* **459**: 945–949.
- Kotia, R.B., and Raghani, A.R. (2010) Analysis of monoclonal antibody product heterogeneity resulting from alternate cleavage sites of signal peptide. *Anal Biochem* **399**: 190–195.
- Kowarik, M., Kung, S., Martoglio, B., and Helenius, A. (2002) Protein folding during cotranslational translocation in the endoplasmic reticulum. *Mol Cell* **10**: 769–778.
- Kraemer, S.M., and Smith, J.D. (2006) A family affair: var genes, PfEMP1 binding, and malaria disease. *Curr Opin Microbiol* **9**: 374–380.
- Maier, A.G., Rug, M., O'Neill, M.T., Brown, M., Chakravorty, S., Szeszak, T., *et al.* (2008) Exported proteins required for virulence and rigidity of *Plasmodium falciparum*-infected human erythrocytes. *Cell* **134**: 48–61.
- Maier, A.G., Cooke, B.M., Cowman, A.F., and Tilley, L. (2009) Malaria parasite proteins that remodel the host erythrocyte. *Nat Rev Microbiol* **7**: 341–354.
- Marti, M., Good, R.T., Rug, M., Knuepfer, E., and Cowman, A.F. (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**: 1930–1933.
- Nkrumah, L.J., Muhle, R.A., Moura, P.A., Ghosh, P., Hatfull, G.F., Jacobs, W.R., Jr, and Fidock, D.A. (2006) Efficient site-specific integration in *Plasmodium falciparum* chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat Methods* **3**: 615–621.
- van Ooij, C., Tamez, P., Bhattacharjee, S., Hiller, N.L., Harrison, T., Liolios, K., *et al.* (2008) The malaria secretome: from algorithms to essential function in blood stage infection. *PLoS Pathog* **4**: e1000084.
- Osborne, A.R., Speicher, K.D., Tamez, P.A., Bhattacharjee, S., Speicher, D.W., and Haldar, K. (2010) The host targeting motif in exported *Plasmodium* proteins is cleaved in the parasite endoplasmic reticulum. *Mol Biochem Parasitol* **171**: 25–31.
- Pringle, S.D., Giles, K., Wildgoose, J.L., Williams, J., Slade, S.E., Thalassinou, K., *et al.* (2007) An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int J Mass Spectrom* **261**: 1–12.

- Przyborski, J.M., Miller, S.K., Pfahler, J.M., Henrich, P., Rohrbach, P., Crabb, B.S., and Lanzer, M. (2005) Trafficking of STEVOR to the Maurer's clefts in *Plasmodium falciparum*-infected erythrocytes. *EMBO J* **24**: 2306–2317.
- Rug, M., Wickham, M.E., Foley, M., Cowman, A.F., and Tilley, L. (2004) Correct promoter control is needed for trafficking of the ring-infected erythrocyte surface antigen to the host cytosol in transfected malaria parasites. *Infect Immun* **72**: 6095–6105.
- Russo, I., Babbitt, S., Muralidharan, V., Butler, T., Oksman, A., and Goldberg, D.E. (2010) Plasmepsin V licenses *Plasmodium* proteins for export into the host erythrocyte. *Nature* **463**: 632–636.
- Sargeant, T., Marti, M., Caler, E., Carlton, J., Simpson, K., Speed, T., and Cowman, A. (2006) Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol* **7**: R12.
- Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27**: 157–162.
- Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., and Mann, M. (2007) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* **1**: 2856–2860.
- Silva, J.C., Denny, R., Dorschel, C.A., Gorenstein, M., Kass, I.J., Li, G.Z., et al. (2005) Quantitative proteomic analysis by accurate mass retention time pairs. *Anal Chem* **77**: 2187–2200.
- Spielmann, T., and Gilberger, T.-W. (2010) Protein export in malaria parasites: do multiple export motifs add up to multiple export pathways? *Trends Parasitol* **26**: 6–10.
- Spielmann, T., Hawthorne, P.L., Dixon, M.W., Hannemann, M., Klotz, K., Kemp, D.J., et al. (2006) A cluster of ring stage-specific genes linked to a locus implicated in cytoadherence in *Plasmodium falciparum* codes for PEXEL-negative and PEXEL-positive proteins exported into the host cell. *Mol Biol Cell* **17**: 3613–3624.
- W.H.O. (2009) *World Malaria Report 2009*. Geneva: World Health Organisation.

### Supporting information

Additional supporting information may be found in the online version of this article.