

## Multiple charged amino acids of *Plasmodium falciparum* SURFIN<sub>4.1</sub> N-terminal region are important for efficient export to the red blood cell

Ben-Yeddy Abel Chitama<sup>a,b,1</sup>, Shinya Miyazaki<sup>b,c,1</sup>, Xiaotong Zhu<sup>b,d,1</sup>, Wataru Kagaya<sup>b,e,f</sup>, Kazuhide Yahata<sup>b</sup>, Osamu Kaneko<sup>a,b,\*</sup>

<sup>a</sup> Leading Program, Graduate School of Biomedical Sciences, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan; <sup>b</sup> Department of Protozoology, Institute of Tropical Medicine (NEKKEN), Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan; <sup>c</sup> Leiden Malaria Research Group, Parasitology, Leiden University Medical Center (LUMC), Leiden, The Netherlands; <sup>d</sup> Department of Immunology, College of Basic Medical Science, China Medical University, Shenyang, Liaoning 110122, China; <sup>e</sup> Department of Environmental Parasitology, Graduate School of Tokyo Medical and Dental University, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo, Tokyo 113-0034, Japan; <sup>f</sup> Department of Parasitology, Graduate School of Medicine, Osaka City University, 1-4-3 Asahimachi, Abeno, Osaka 545-8585, Japan

<sup>1</sup> These authors contributed equally to this work

\* Correspondence should be addressed to O.K. (email: okaneko@nagasaki-u.ac.jp).

**Keywords:** erythrocyte; malaria; *Plasmodium falciparum*; protein export; SURFIN

Ben-Yeddy Abel Chitama, benchitama@gmail.com; Shinya Miyazaki, shinya841212@gmail.com; Xiaotong Zhu, xtzhu@cmu.edu.cn; Wataru Kagaya, wataru.kagaya@gmail.com; Kazuhide Yahata, kyahata@nagasaki-u.ac.jp; Osamu Kaneko, okaneko@nagasaki-u.ac.jp.

### Abstract

*Plasmodium falciparum*, an obligate intracellular protozoan parasite which causes the severe form of human malaria, exports numerous proteins to the infected red blood cell that are important for its survival and of severe pathological effect to its host. These proteins and their export mechanisms are candidates for drug and vaccine development, and among them is the *Plasmodium* SURFIN family of proteins. Previously we showed that the N-terminal region along with the sequence surrounding the transmembrane domain of SURFIN<sub>4.1</sub> is essential for its export to Maurer's clefts in the red blood cell cytoplasm. We proposed that this region is recognized by a machinery responsible for protein translocation across the parasitophorous vacuole membrane surrounding the parasite. To understand the export mechanism further, we utilized a fluorescent protein-tagged mini-SURFIN<sub>4.1</sub> consisting of the minimum essential components for export. Alanine scanning of all charged amino acids within the N-terminal region revealed that replacement of 3 glutamic acid and 2 lysine residues significantly impairs the export efficiency of this protein across the parasitophorous vacuole membrane. In addition, N-terminally Myc-tagged mini-SURFIN<sub>4.1</sub> and mini-SURFIN<sub>4.2</sub> with similar architectures were detected with anti-Myc antibody at Maurer's clefts, indicating that elements required for export to Maurer's clefts are conserved between SURFIN<sub>4.1</sub> and SURFIN<sub>4.2</sub>, and that N-terminal sequences of these SURFIN members are not cleaved during export. Our results implicate a conserved nature of SURFIN export to the red blood cell, particularly an important role of multiple glutamic acid and lysine residues in the SURFIN N-terminal region.

### 1. Introduction

Malaria remains a barrier to the economic progress of developing countries, particularly due to *Plasmodium falciparum*, the causative agent of the most severe form of malaria [1]. Of the total malaria cases occurring worldwide, half occur in sub-Saharan African, which also contains a majority of the least developed countries [2]. Malaria is a disease of the red blood cell (RBC), and the pathology is largely due to parasite-encoded proteins that the parasite exports to the RBC cytoplasm and surface [3]. These proteins are involved in RBC remodeling essential to parasite survival, while consequently mediating pathophysiological reaction in the host [4]. In trafficking to the RBC surface majority of these proteins pass through parasite-generated membranous structures called Maurer's clefts which act as a sorting platforms within the RBC cytoplasm [5].

A conserved N-terminal region sequence, termed the *Plasmodium* EXport ELEMENT (PEXEL) or Vacuolar Targeting Sequence (VTS) motif, was shown

to be essential for export of a diverse group of proteins to the RBC cytoplasm. The PEXEL/VTS motif is a pentameric amino acid motif with the consensus sequence RxLxE/Q/D; and identifies a repertoire of over 300 *Plasmodium* predicted exported proteins, termed the exportome [6, 7, 8]. During transport within the endoplasmic reticulum (ER), this motif is cleaved by the ER-resident protease plasmepsin V [9, 10]. The exposed amino acid residues (xE/Q/D) after cleavage are essential for export to the RBC cytoplasm [8]. Many studies have identified a second group of exported proteins which lack obvious PEXEL/VTS motifs; and are termed PEXEL negative exported proteins (PNEPs). Examples include SBP1, REX1, REX2, MAHRP1, MAHRP2, MSRP6, and MSRP7 [11, 12, 13, 14, 15, 16]. The N-terminal regions of some PNEPs such as SBP1, REX2 and MAHRP2 were shown to be essential for export to the RBC [13, 15, 17], indicating that the N-terminal sequence requirement for RBC export is conserved in PNEPs [18]. The N-terminal regions of PNEPs are potentially cleaved by an unknown

mechanism, as shown for REX2 [13]. However, it is not known whether N-terminal processing of PNEPs is required for export to the RBC.

SURFIN<sub>4.1</sub> (Surface-associated interspersed protein 4.1) is a PNEP exported to the Maurer's clefts and is encoded by the *surf* gene family consisting of 10 members in *P. falciparum* [19, 20]. The N terminal region (18 amino acids (aa)), transmembrane domain (23 aa), and following short cytosolic region (17 aa) of SURFIN<sub>4.1</sub> contain sufficient information for export to Maurer's clefts [21]; however, specific amino acid residues in the N-terminal region that contribute to this export have not been evaluated. Herein we report the presence of five charged amino acid residues (3 glutamic acid and 2 lysine residues) in the N-terminal sequence of SURFIN<sub>4.1</sub> that independently contribute to efficient export of recombinant mini-SURFIN<sub>4.1</sub> to Maurer's clefts. We also show that elements required for export of SURFIN<sub>4.1</sub> are conserved in the paralog SURFIN<sub>4.2</sub>, for which infected RBC (iRBC) surface exposure has been shown [19, 22]. We additionally determined that the N terminal sequences of SURFIN<sub>4.1</sub> and SURFIN<sub>4.2</sub> are not processed during export to Maurer's clefts.

## 2. Materials and Methods

### 2.1. *P. falciparum* cultivation and transfection

*P. falciparum* MS822 and 3D7S8 lines were cultured basically as described [21, 23]. The parasites were maintained with O<sup>+</sup> RBCs at 2% hematocrit in RPMI-1640 medium supplemented with 5% heat-inactivated pooled type AB<sup>+</sup> human serum, 0.25% AlbuMaxI (Invitrogen, Carlsbad, CA), 200 mM hypoxanthine (Sigma, St. Louis, MO), and 10 µg/mL gentamicin (Sigma).

Transfection of *P. falciparum* was performed as described [24]. Briefly, 100 µg of plasmids were electroporated to uninfected RBCs (Gene Pulser Xcell; Bio-Rad, Hercules, CA), mixed with *P. falciparum*-iRBCs and cultured for 3 days without drug. At day 4, 5 nM WR99210 (a gift from Dr. D. Jacobus) was added to the culture for the selection of drug-resistant transfectants. The MS822 line was used to evaluate Myc-mini-SURFIN<sub>4.1</sub> and Myc-mini-SURFIN<sub>4.2</sub>. After transfectants were obtained, the concentration of WR99210 was gradually increased to 20 nM before analysis.

### 2.2. Plasmid construction

Plasmids were constructed using the Multisite Gateway system (Thermo Fisher Scientific, Waltham, MA) and In-Fusion cloning system (Takara Bio Inc., Japan). The pENT12 plasmids for the Ala replacement (pENT12\_SURFIN4.1<sup>N4-D34A-T-cyt</sup>, pENT12\_SURFIN4.1<sup>N4-K35A-T-cyt</sup>, pENT12\_SURFIN4.1<sup>N4-D37A-T-cyt</sup>, pENT12\_SURFIN4.1<sup>N4-E38A-T-cyt</sup>,

pENT12\_SURFIN4.1<sup>N4-D40A-T-cyt</sup>, pENT12\_SURFIN4.1<sup>N4-K42A-T-cyt</sup>, pENT12\_SURFIN4.1<sup>N4-E45A-T-cyt</sup>, pENT12\_SURFIN4.1<sup>N4-K49A-T-cyt</sup> and pENT12\_SURFIN4.1<sup>N4-E50A-T-cyt</sup>) were modified based on a plasmid to express SURFIN4.1<sup>N4-T-Cyt</sup> [21] by site-directed mutagenesis using KOD -Plus- Mutagenesis Kit (Toyobo, Japan) with oligonucleotide primers listed in supplementary Table 1. The expression plasmids of N-terminal Ala-replaced mini-SURFIN<sub>4.1</sub> were constructed by the Gateway MultiSite LR recombination reaction with pENT4/1-*Pf*CRT5' (as a promoter component), pENT23\_3×Ty-GFP, and pCHDR-3/4 plasmids [25]. The open reading frame sequences of all plasmids were confirmed before transfection.

For Myc-mini-SURFIN<sub>4.1</sub>, pENT12\_SURFIN4.1<sup>2Myc-N-T-cyt</sup> was constructed from a plasmid to express SURFIN4.1<sup>N-T-Cyt</sup> [21] by inserting a DNA fragment encoding double Myc tags using a KOD -Plus- Mutagenesis Kit. For Myc-mini-SURFIN<sub>4.2</sub>, a DNA fragment encoding the N-terminal region was PCR-amplified from *P. falciparum* genomic DNA and introduced into the EcoRV site of pB13 plasmid to yield pB13\_SURFIN4.2<sup>N</sup>. pB13 has been made from pB12 [26] by replacing the sequence between NheI and PstI with the oligonucleotides CTAGCAGATCTACTAGTAGGCCTCTGCA and GAGGCCTACTAGTAGATCTG and by modifying the attB2 site to attB3 site by site-directed mutagenesis with the oligonucleotides CAAAGTTGTAGCTCGAATTCCTGGCC and TATAATAAAGTTGGTTACTCGAGCAGCTCG. SURFIN<sub>4.2</sub> sequence spanning from the variable region 2 (Var2 region) to a part of the cytoplasmic region was amplified and inserted into the XbaI site of pB13\_SURFIN4.2<sup>N</sup> to yield pB13\_SURFIN4.2<sup>N-V2-T-C</sup>. A frameshift was then corrected, and the intron and Var2 region were removed, to yield pB13\_SURFIN4.2<sup>N-T-C</sup>. pB13\_SURFIN4.2<sup>2Myc-N-T-C</sup> was constructed by inserting a DNA fragment encoding double Myc tags and subjected to the Gateway BP reaction followed by a Gateway Multisite LR recombination reaction with pENT41-*Pf*CRT5' and pCHD43(II) [27].

### 2.3. Live cell imaging and data analysis

To capture live green fluorescent protein (GFP) signals from transgenic parasites, parasites were synchronized by the 5% D-sorbitol method [28] and allowed to grow until the trophozoite stage. Images were visualized using a fluorescence microscope (Axio Imager Z2; Carl Zeiss, Germany) with a 100x/1.40 oil-immersion lens and captured with a charge-coupled device camera (AxioCam MRm; Carl Zeiss). Images were processed using Adobe Photoshop CS3 (Adobe systems Inc., San Jose, CA).

To quantify the intensity of signals located in the iRBC cytoplasm, 20 images were selected for each transfectant. Area and mean values of the signal were

measured from the entire parasite-iRBC (signals from the parasite and in the iRBC cytoplasm), signals in the iRBC cytoplasm, and signals from the nearby external area surrounding the parasite (background) using Fiji [29]. Background signal mean values were subtracted from the test mean values and multiplied by the area values to produce integrated density values for each test area. The proportion of the signals in the iRBC cytoplasm was obtained by dividing integrated signal density values in the iRBC cytoplasm by those from the entire parasite-iRBC. Statistical differences of the proportion of the signals in the iRBC cytoplasm were evaluated by Kruskal-Wallis one-way analysis of variance followed by Dunn's multiple comparison test using GraphPad PRISM6 (GraphPad, San Diego, CA) and graphs were drawn using Microsoft Excel 2016.

#### 2.4. Indirect immunofluorescence assay (IFA)

Thin blood smears on glass slides were air-dried and stored at  $-80^{\circ}\text{C}$ . Retrieved blood smears were fixed with 4% paraformaldehyde/0.075% glutaraldehyde at room temperature for 30 min, followed by quenching with 50 mM glycine for 15 min at room temperature. The slides were blocked with 10% normal goat serum (Invitrogen) at  $37^{\circ}\text{C}$  for 60 min and probed with mouse anti-GFP monoclonal antibody (1:250; 3E6, Invitrogen), mouse anti-Ty1 monoclonal antibody (1:500; MAb-054-050; Diagenode, Belgium) or mouse anti-Myc monoclonal antibody (1:1000; 9B11, Cell Signaling Technology, Danvers, MA), and rabbit anti-SBP1 serum (1:2000; a kind gift from T. Tsuboi). After washing with phosphate-buffered saline (PBS), the smears were incubated with PBS containing Alexa Fluor® 488-conjugated goat anti-mouse IgG antibody and Alexa Fluor® 594-conjugated goat anti-rabbit IgG antibody (1:500, Invitrogen). Parasite nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI). Images were captured using a fluorescence microscope and camera and processed as described above.

#### 2.5. Western blotting

Parasite-iRBCs were exposed to 5% D-sorbitol at hours 0 and 6, then harvested at hour 30 to obtain synchronized parasite material. Mature trophozoite/schizonts-iRBCs were treated with 0.15% saponin in PBS containing a mixture of protease inhibitors (PI; cOmplete, EDTA-free; Roche, Basel, Switzerland), and the pellets after centrifugation were washed with PBS-PI, and proteins were dissolved in SDS-loading buffer with 2% 2-mercaptoethanol. After passing through a 27G needle to shear parasite DNA, samples were centrifuged and supernatants were collected. Protein from an estimated  $1.0 \times 10^5$  parasites was subjected to electrophoresis on 5-20% SDS-polyacrylamide gradient mini gels (ATTO, Japan). The protein bands were transferred from gels to PVDF membranes (Millipore, Billerica, MA). The membranes were then probed with mouse anti-Ty1 monoclonal

antibody (1:500) or anti-*Plasmodium berghei* HSP70 mouse monoclonal antibody (obtained from J. Sattabongkot, which cross-reacts with HSP70 orthologs of other *Plasmodium* spp) [30] followed by a secondary incubation with horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG antibodies (Promega, Madison, WI). Bands were visualized with Immobilon™ Western Chemiluminescent HRP substrate (Millipore) and detected using a chemiluminescence detection system (LAS-4000EPUVmini; Fujifilm, Japan). The relative molecular sizes of the proteins were calculated based on molecular size standards (Precision Plus Dual Color Standards; Bio-Rad).

### 3. Results

#### 3.1. Identification of SURFIN<sub>4.1</sub> N-terminal amino acid residues that are important for efficient export to the iRBC cytoplasm

To evaluate the influence of charged amino acid residues within the SURFIN<sub>4.1</sub> N-terminal region on export to the iRBC cytoplasm, we utilized sequences consisting of amino acid residues 34 – 50 (termed N4 region), which was found to contain sufficient information for export to Maurer's clefts [21]. Initially we used the MS822 parasite line to express a panel of recombinant mini-SURFIN<sub>4.1</sub> proteins consisting of the N4 region fused to the transmembrane domain and a short cytoplasmic region (N4TC). In these proteins, six negatively charged amino acid residues (D34A, D37A, E38A, D40A, E45A, and E50A) and three positively charged amino acid residues (K35A, K42A, and K49A) were independently replaced to Ala (Fig. 1A). To monitor protein export we fused a Ty1 peptide tag and GFP to the C-terminus (Fig. 1A). Western blot analysis confirmed that all recombinant mini-SURFIN<sub>4.1</sub> proteins were expressed in a similar quantity and fashion with a few extra smaller bands of possible proteolysed products (Fig. 1B). Live cell imaging for MS822 line transfectants revealed that E38A, K42A, E45A, K49A and E50A showed sparse signal in the iRBC cytoplasm (Fig. 1C) and had significantly lower signal in the iRBC cytoplasm per total signal compared to N4TC which served as a background ( $p < 0.0001$ ; Fig. 1D). Signal in the iRBC cytoplasm of D34A, K35A, D37A, and D40A replacements were similar to N4TC (Fig. 1C) and no significant difference was detected for the signal in the iRBC cytoplasm per total signal compared to N4TC (Fig. 1D).

To confirm these results we further examined the influence of amino acid residues in the N4 region using 3D7S8-background transfectants; focusing in part on amino acid residues that by the replacement to Ala showed significant reduction of the signal in the iRBC cytoplasm per total signal in MS822-background. We used two different lines to add more confirmatory value to our findings. Western blot analysis indicated that all generated recombinant mini-SURFIN<sub>4.1</sub> proteins were

expressed with a similar quantity and fashion with a few extra smaller bands of possible proteolysed products (Fig. 2A). Live imaging indicated that D37A, for which reduction was not observed in the MS822-background transfectant, showed signals in the iRBC cytoplasm similar to N4TC (Fig. 2B) and the signal in the iRBC cytoplasm per total signal was not significantly different from that of N4TC (Fig. 2C). Five replacements that reduced the export efficacy in MS822-background (E38A, K42A, E45A, K49A, and E50A) again showed sparse signal in the iRBC cytoplasm (Fig. 2B) and the ratios of the iRBC cytoplasm per total signal were significantly lower than that of the 3D7S8-background N4TC parasite (Fig. 2C). Amino acid residues for which Ala replacement resulted in a significant loss of signal export to iRBC cytoplasm in both MS822 and 3D7S8 parasite lines are shown in the figure 2D.

The signal in the iRBC cytoplasm, although faint, were still visible for some parasites in 5 transfectants. To assess if the observed GFP signals in the iRBC cytoplasm in these parasites localized to Maurer's clefts, we performed IFA using anti-GFP antibody because of its superior sensitivity compared to GFP live imaging. Detected signals in the iRBC cytoplasm in these parasites colocalized with the Maurer's cleft marker SBP1 (Fig. 3), indicating that Ala replacement significantly reduced the export efficacy in these transgenic parasites, but the proteins were able to reach Maurer's clefts.

These results, based on transfection to two genetic backgrounds, consistently showed that Ala replacement of negatively (E<sub>38</sub>, E<sub>45</sub>, and E<sub>50</sub>) and positively (K<sub>42</sub> and K<sub>49</sub>) charged amino acid residues in the N4 region significantly reduced the signal intensity in the iRBC cytoplasm per total signal, suggesting that these five amino acid residues are independently involved in efficient export of mini-SURFIN<sub>4.1</sub> to the iRBC cytoplasm. This effect is independent because charged amino acids tested were able to affect protein export individually, that is without depending on each other.

### 3.2. N-terminal sequences of SURFIN<sub>4.1</sub> protein was not cleaved during export to Maurer's clefts

To assess whether the SURFIN<sub>4.1</sub> N-terminal sequence is cleaved during export to Maurer's clefts, we generated a transfectant expressing N-terminally Myc-tagged and C-terminally Ty1 and GFP-tagged mini-SURFIN<sub>4.1</sub> consisting of a short N terminal region, transmembrane domain, and the adjacent short cytosolic region (Fig. 4A). Both anti-Ty1 and anti-Myc antibodies detected a punctate pattern of expression in the iRBC cytoplasm which co-localized with the Maurer's cleft marker protein SBP1 (Fig. 4B), indicating that the N-terminal sequence of SURFIN<sub>4.1</sub> is not processed during export to Maurer's clefts.

### 3.3. Regions required for export to Maurer's clefts are conserved between SURFIN<sub>4.1</sub> and SURFIN<sub>4.2</sub>

To determine if regions of SURFIN<sub>4.1</sub> required for export to Maurer's clefts are conserved in another SURFIN family member, we generated, termed Myc-mini-SURFIN<sub>4.2</sub>, a transfectant expressing N-terminally Myc-tagged and C-terminally Ty1 and GFP-tagged mini-SURFIN<sub>4.2</sub> whose composition is the same as Myc-mini-SURFIN<sub>4.1</sub>; specifically, containing the N-terminal region (51 aa), transmembrane domain (23 aa) with the adjacent short cytosolic region (20 aa) (Fig. 4A). IFA with anti-Ty1 or anti-Myc antibodies showed that both signals co-localized with SBP1 (Fig. 4C), indicating that Myc-mini-SURFIN<sub>4.2</sub> is exported to Maurer's clefts and at that location Myc-mini-SURFIN<sub>4.2</sub> retains the N-terminal Myc sequence. This result indicates that the sequence information for export to the RBC is conserved between SURFIN<sub>4.1</sub> and SURFIN<sub>4.2</sub> and confirms that the processing of the N-terminal sequence of SURFIN is not required for this process.

## 4. Discussion

Protein export to the iRBCs plays an essential role for malaria parasite development and proliferation within the RBC and consequently virulence in the human host. In this study we demonstrate the following: *i*) at least five charged amino acid residues in the N-terminal region of SURFIN<sub>4.1</sub> (E<sub>38</sub>, K<sub>42</sub>, E<sub>45</sub>, K<sub>49</sub>, and E<sub>50</sub>) are involved in its efficient export to the iRBC cytoplasm, *ii*) regions required for export are conserved between SURFIN<sub>4.1</sub> and SURFIN<sub>4.2</sub> and *iii*) processing of the N-terminal region of SURFIN<sub>4.1</sub> and SURFIN<sub>4.2</sub> is not required for export to the iRBC cytoplasm.

The importance of the charged amino acids for export of *P. falciparum*-encoded proteins to the iRBC cytoplasm has been shown in other studies; for example, the role of Glu has been also observed for export of another PNEP, REX2, whereby mutation of Glu at position 7 (MKMYLAEIFSSGK) was sufficient to abolish export of this protein to the iRBC cytoplasm [13]. Ala replacement of Glu residues regarding export of PfHsp70x (SNNAEES) was also found to be essential as it led to the accumulation of this protein in the parasitophorous vacuole (PV) [31]. In the case of SBP1, Ala replacement of the N-terminal DEPTQLQDAVP (amino acid positions 16-26), which contains three negatively charged residues, stopped export of this protein [17]. A net negative charge of the N-terminal region at a neutral pH was proposed to be important for SBP1 export to Maurer's clefts (pI of N-terminal 35 amino acid residues of SBP1 is 4.50) [17]. The pI of the N-terminal 50 amino acid residues of SURFIN<sub>4.1</sub> and N-terminal 18 amino acid residues of mini-SURFIN<sub>4.1</sub> N4 region are 4.27 and 4.30, respectively. Ala replacement at E<sub>38</sub>, E<sub>45</sub>, and E<sub>50</sub> in the mini-SURFIN<sub>4.1</sub> N4 region all only slightly increased the pI (4.44) but had a strong negative impact on export. We found that K42A and K49A, both of which decreased the pI to 4.02, also reduced the protein

transport efficacy, suggesting that rather than a net negative charge the correct positioning of negatively and positively charged amino acid residues are critical for mini-SURFIN<sub>4.1</sub> export to the iRBC cytoplasm. The structure was recently resolved for a translocon on the parasitophorous vacuole, termed *Plasmodium* translocon of exported proteins (PTEX) [32]. Because PTEX translocates both PEXEL-containing proteins and PNEPs [33-35], these charged amino acid residues may interact directly with PTEX or may be responsible for a correct conformation to be recognized by this translocation machinery. Future computational modeling of the interaction between PTEX and mini-SURFIN<sub>4.1</sub> may provide insight into how SURFIN<sub>4.1</sub> and PNEPs are recognized by the translocon.

Marti et al. (2004) demonstrated Maurer's cleft transport of a mini-PfEMP1 consisting of the N-terminal aa 1-32 (MGPPSTAPDYSSAKDAKELLDKIGQQVHDKVK, "PEXEL-like" residues are underlined) combined with the transmembrane and cytoplasmic domains (aa 1913-2360) of PfEMP1 (PF3D7\_1240600). This is a similar domain composition to our mini-SURFIN proteins. For the mini-PfEMP1 the simultaneous Ala replacement of K<sub>17</sub>, L<sub>19</sub>, and D<sub>21</sub> (MGPPSTAPDYSSAKDAAEALAKI<sub>19</sub>QVHDKVK), which alters the PEXEL-like motif, abolished export [35]. Because PfEMP1 does not possess an N-terminal ER-signal sequence, the proposed "PEXEL-like" motif is likely recognized by PTEX in a similar fashion to our mini-SURFIN proteins. If this is the case, other charged amino acids in the PfEMP1 N-terminal region (K<sub>14</sub>, D<sub>15</sub>, E<sub>18</sub>, K<sub>22</sub>, D<sub>29</sub>, K<sub>30</sub> and K<sub>32</sub>) may also be critical for its export. We speculate that the "PEXEL-like" motif in the PfEMP1 N-terminus is not cleaved, based upon the lack of processing of the N-terminal SURFIN-derived sequences in mini-SURFINs.

Because we focused on the signal detected within Maurer's clefts in the iRBC cytoplasm, we did not assay possible inhibitions at prior trafficking steps, such as the stability of the proteins in their transit from the ER to incorporation in Maurer's cleft structures. IFA signal can be seen around the parasite nucleus even in the transfectant with parental construct N4TC (e.g. Fig 2A), suggesting that a portion of the recombinant protein is retained in the ER at varying levels in all transfectants. However, all transfectants, including five transfectants with reduced signal in the iRBC cytoplasm, showed IFA signal surrounding the parasite, indicating that proteins were able to reach the parasite plasma membrane. We interpret that translocation across the parasitophorous vacuole membrane is disturbed and mini-SURFIN<sub>4.1</sub> were retained on the parasite plasma membrane in the five transfectants with reduced signals in the iRBC cytoplasm, likely due to insufficient recognition by PTEX.

Mini-SURFIN<sub>4.1</sub> and mini-SURFIN<sub>4.2</sub> proteins containing N-terminal Myc tags were detected within Maurer's clefts, indicating that N-terminal processing of SURFIN sequence is not required for export of these

proteins to the iRBC cytoplasm. To our knowledge this is the first report with unambiguous data showing that the N-terminal region is not cleaved during this process for PNEPs. Thus, although another PNEP, REX2, was shown to be processed at its N-terminal region like PEXEL-dependent proteins [13], the N-terminal processing is clearly not a general feature of PNEPs.

In summary, we characterized the N-terminus of *P. falciparum* SURFIN<sub>4.1</sub> and SURFIN<sub>4.2</sub> for their role in export to Maurer's clefts in the iRBC cytoplasm. We found that five amino acid residues in the N-terminus N4 region are involved in efficient export of mini-SURFIN<sub>4.1</sub> and that not only a net negative charge, but the correct positioning of negatively as well as positively charged amino acid residues is important for export. We showed that the SURFIN N-terminal sequence in the Myc-mini-SURFIN proteins is not processed during export to Maurer's clefts, indicating that N-terminal processing is not required for SURFIN export to the iRBC cytoplasm. Our results implicate a conserved nature of SURFIN transport and contribute to understanding export of PNEPs to Maurer's clefts

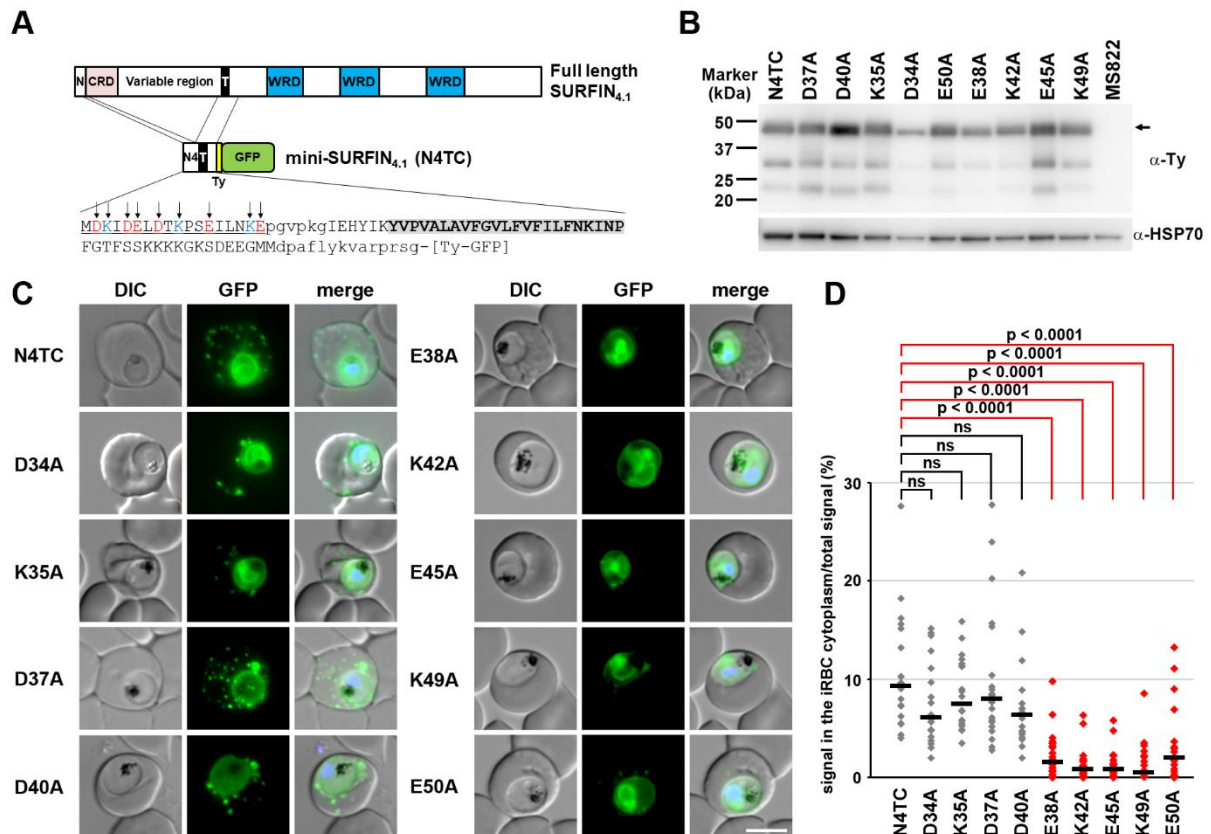
## Acknowledgements

We thank Takeda M, Tanaka R, Ogoshi (Sakura) M, Takahama S, Matsumoto M and Kinoshita M for their expertise. We thank Templeton TJ for his critical comments. We are also grateful to the Nagasaki Red Cross Blood Center for providing human RBCs and plasma. This study was supported in part by Grants-in-Aids for Scientific Research on Innovative Areas 23117008 (OK). BYAC was supported by the leading program, Nagasaki University, Japan. This study was conducted at the Joint Usage / Research Center for Tropical Disease, Institute of Tropical Medicine, Nagasaki University, Japan.

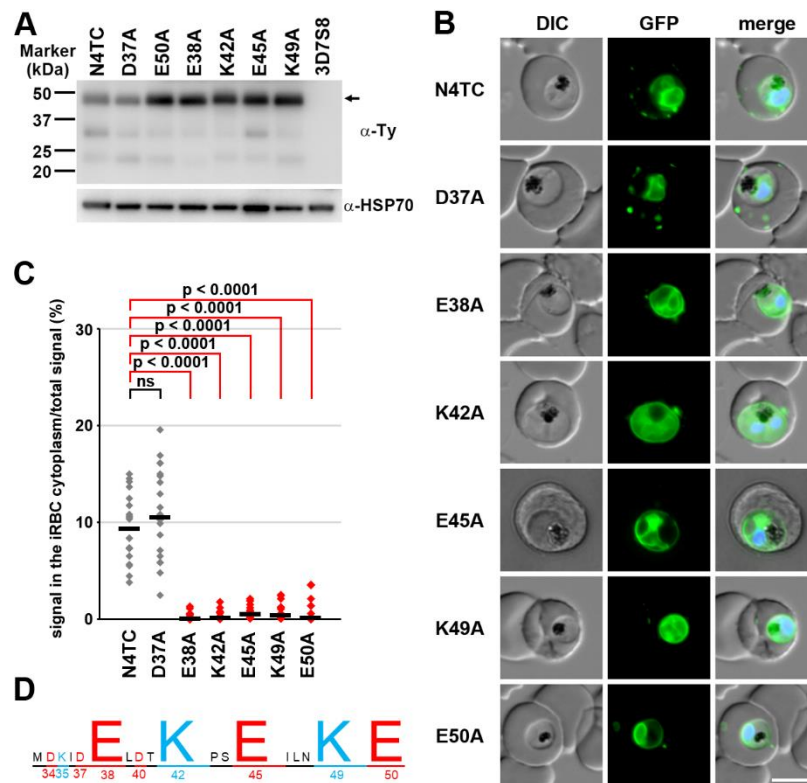
## References

1. J. Sachs, P. Malaney, The economic and social burden of malaria, *Nature* 415 (2002) 680–685. <https://doi.org/10.1038/415680a>
2. World Health Organization, World Malaria Report 2018, World Health Organization, 2018, Geneva, Switzerland.
3. T.F. de Koning-Ward, M.W.A. Dixon, L. Tilley, P.R. Gilson, *Plasmodium* species: master renovators of their host cells, *Nat. Rev. Microbiol.* 14 (2016) 494–507. <https://doi.org/10.1038/nrmicro.2016.79>.
4. A.G. Maier, M. Rug, M.T. O'Neill, M. Brown, S. Chakravorty, T. Szeszak, J. Chesson, Y. Wu, K. Hughes, R.L. Coppel, C. Newbold, J.G. Beeson, A. Craig, B.S. Crabb, A.F. Cowman, Exported proteins required for virulence and rigidity of *Plasmodium falciparum*-infected human erythrocytes, *Cell* 134 (2008) 48–61. <https://doi.org/10.1016/j.cell.2008.04.051>.
5. E. Mundwiler-Pachlatko, H-P. Beck, Maurer's clefts, the enigma of *Plasmodium falciparum*, *Proc. Natl. Acad. Sci. U.S.A.* 110 (2013) 19987–19994. <https://doi.org/10.1073/pnas.1309247110>.
6. P. Horrocks, D. Muhia, Pexel/VTS: a protein-export motif in erythrocytes infected with malaria parasites, *Trends Parasitol.* 21 (2005) 396–399. <https://doi.org/10.1016/j.pt.2005.07.004>.
7. J. Schulze, M. Kwiatkowski, J. Borner, H. Schlüter, I. Bruchhaus, T. Burmester, T. Spielmann, C. Pick, The *Plasmodium falciparum* exportome contains non-canonical PEXEL/HT proteins, *Mol. Microbiol.* 97 (2015) 301–314. <https://doi.org/10.1111/mmi.13024>.

8. C. Grüring, A. Heiber, F. Kruse, S. Flemming, G. Franci, S.F. Colombo, E. Fasana, H. Schoeler, N. Borgese, H.G. Stunnenberg, J.M. Przyborski, T.W. Gilberger, T. Spielmann, Uncovering common principles in protein export of malaria parasites, *Cell Host Microbe* 12 (2012) 717–729. <https://doi.org/10.1016/j.chom.2012.09.010>.
9. J.A. Boddey, A.N. Hodder, S. Günther, P.R. Gilson, H. Patsiouras, E.A. Kapp, J.A. Pearce, T.F. de Koning-Ward, R.J. Simpson, B.S. Crabb, A.F. Cowman, An aspartyl protease directs malaria effector proteins to the host cell, *Nature* 463 (2010) 627–631. <https://doi.org/10.1038/nature08728>.
10. I. Russo, S. Babbitt, V. Muralidharan, T. Butler, A. Oksman, D.E. Goldberg, Plasmepsin V licenses *Plasmodium* proteins for export into the host erythrocyte, *Nature* 463 (2010) 632–636. <https://doi.org/10.1038/nature08726>.
11. T. Blisnick, M.E.M. Betoulle, J.C. Barale, P. Uzureau, L. Berry, S. Desroses, H. Fujioka, D. Mattei, C. Brown Benton, Pfsbp1, a Maurer's cleft *Plasmodium falciparum* protein, is associated with the erythrocyte skeleton, *Mol. Biochem. Parasitol.* 111 (2000) 107–121. [https://doi.org/10.1016/S0166-6851\(00\)00301-7](https://doi.org/10.1016/S0166-6851(00)00301-7).
12. M.W.A. Dixon, P.L. Hawthorne, T. Spielmann, K.L. Anderson, K.R. Trenholme, D.L. Gardiner, Targeting of the ring exported protein 1 to the Maurer's clefts is mediated by a two-phase process, *Traffic* 9 (2008) 1316–1326. <https://doi.org/10.1111/j.1600-0854.2008.00768.x>.
13. S. Haase, S. Herrmann, C. Grüring, A. Heiber, P.W. Jansen, C. Langer, M. Treeck, A. Cabrera, C. Bruns, N.S. Struck, M. Kono, K. Engelberg, U. Ruch, H.G. Stunnenberg, T.W. Gilberger, T. Spielmann, Sequence requirements for the export of the *Plasmodium falciparum* Maurer's clefts protein REX2, *Mol. Microbiol.* 71 (2009) 1003–1017. <https://doi.org/10.1111/j.1365-2958.2008.06582.x>.
14. C. Spycher, N. Klönis, T. Spielmann, E. Kump, S. Steiger, L. Tilley, H.P. Beck, MAHRP-1, a novel *Plasmodium falciparum* histidine-rich protein, binds ferriprotoporphyrin IX and localizes to the Maurer's clefts, *J. Biol. Chem.* 278 (2003) 35373–35383. <https://doi.org/10.1074/jbc.M305851200>.
15. E. Pachlatko, S. Rusch, A. Müller, A. Hemphill, L. Tilley, E. Hanssen, H-P. Beck, MAHRP2, an exported protein of *Plasmodium falciparum*, is an essential component of Maurer's cleft tethers, *Mol. Microbiol.* 77 (2010) 1136–1152. <https://doi.org/10.1111/j.1365-2958.2010.07278.x>.
16. A. Heiber, F. Kruse, C. Pick, C. Grüring, S. Flemming, A. Oberli, H. Schoeler, S. Retzlaff, P. Mesén-Ramírez, J.A. Hiss, M. Kadekoppala, L. Hecht, A.A. Holder, T.W. Gilberger, T. Spielmann, Identification of new PNEPs indicates a substantial non-PEXEL exportome and underpins common features in *Plasmodium falciparum* protein export, *PLoS Pathog.* 9 (2013) e1003546. <https://doi.org/10.1371/journal.ppat.1003546>.
17. T. Saridaki, K.S. Fröhlich, C. Braun-Bretton, M. Lanzer, Export of PfSBP1 to the *Plasmodium falciparum* Maurer's clefts, *Traffic* 10 (2009) 137–152. <https://doi.org/10.1111/j.1600-0854.2008.00860.x>.
18. C. Grüring, A. Heiber, F. Kruse, S. Flemming, G. Franci, S.F. Colombo, E. Fasana, H. Schoeler, N. Borgese, H.G. Stunnenberg, J.M. Przyborski, T.W. Gilberger, T. Spielmann, Uncovering common principles in protein export of malaria parasites, *Cell Host Microbe* 12 (2012) 717–729. <https://doi.org/10.1016/j.chom.2012.09.010>.
19. G. Winter, S. Kawai, M. Haeggström, O. Kaneko, A. von Euler, S. Kawazu, D. Palm, V. Fernandez, M. Wahlgren, SURFIN is a polymorphic antigen expressed on *Plasmodium falciparum* merozoites and infected erythrocytes, *J. Exp. Med.* 201 (2005) 1853–1863. <https://doi.org/10.1084/jem.20041392>.
20. F.A. Mphande, U. Ribacke, O. Kaneko, F. Kironde, G. Winter, M. Wahlgren, SURFIN<sub>4.1</sub>, a schizont-merozoite associated protein in the SURFIN family of *Plasmodium falciparum*, *Malar. J.* 7 (2008) 116. <https://doi.org/10.1186/1475-2875-7-116>.
21. X. Zhu, K. Yahata, J.S.F. Alexandre, T. Tsuboi, O. Kaneko, The N-terminal segment of *Plasmodium falciparum* SURFIN<sub>4.1</sub> is required for its trafficking to the red blood cell cytosol through the endoplasmic reticulum, *Parasitol. Int.* 62 (2013) 215–229. <https://doi.org/10.1016/j.parint.2012.12.006>.
22. W. Kagaya, S. Miyazaki, K. Yahata, N. Ohta, O. Kaneko, The cytoplasmic region of *Plasmodium falciparum* SURFIN<sub>4.2</sub> is required for transport from Maurer's clefts to the red blood cell surface, *Trop. Med. Health* 43 (2015) 265–272. <https://doi.org/10.2149/tmh.2015-38>.
23. W. Trager, J.B. Jensen, Human malaria parasites in continuous culture, *Science* 193 (1976) 673–675. <https://doi.org/10.1126/science.781840>.
24. K. Deitsch, C. Driskill, T. Wellem, Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes, *Nucleic Acids Res.* 29 (2001) 850–853. <https://doi.org/10.1093/nar/29.3.850>.
25. G.G. van Dooren, M. Marti, C.J. Tonkin, L.M. Stimmler, A.F. Cowman, G.I. McFadden, Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of *Plasmodium falciparum*, *Mol. Microbiol.* 57 (2005) 405–419. <https://doi.org/10.1111/j.1365-2958.2005.04699.x>.
26. J.S.F. Alexandre, K. Yahata, S. Kawai, M. Torii, O. Kaneko, PEXEL-independent trafficking of *Plasmodium falciparum* SURFIN<sub>4.2</sub> to the parasite-infected red blood cell and Maurer's clefts, *Parasitol. Int.* 60 (2011) 313–320. <https://doi.org/10.1016/j.parint.2011.05.003>.
27. T. Sakura, K. Yahata, O. Kaneko, The upstream sequence segment of the C-terminal cysteine-rich domain is required for microneme trafficking of *Plasmodium falciparum* erythrocyte binding antigen 175, *Parasitol. Int.* 62 (2013) 157–164. <https://doi.org/10.1016/j.parint.2012.12.002>.
28. C. Lambros, J.P. Vanderberg, Synchronization of *Plasmodium falciparum* erythrocytic stages in culture, *J. Parasitol.* 65 (1979) 418–420.
29. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis, *Nature Methods* 9 (2012) 676–682. <https://doi.org/10.1038/nmeth.2019>.
30. M. Tsuji, D. Mattei, R.S. Nussenzweig, D. Eichinger, F. Zavala, Demonstration of heat-shock protein 70 in the sporozoite stage of malaria parasites, *Parasitol. Res.* 80 (1994) 16–21.
31. M. Rhiel, V. Bittl, A. Tribensky, S.C. Charnaud, M. Strecker, S. Müller, M. Lanzer, C. Sanchez, C. Schaeffer-Reiss, B. Westermann, B.S. Crabb, P.R. Gilson, S. Külzer, J.M. Przyborski, Trafficking of the exported *P. falciparum* chaperone PfHsp70x, *Sci. Rep.* 6 (2016) 36174. <https://doi.org/10.1038/srep36174>.
32. C.M. Ho, J.R. Beck, M. Lai, Y. Cui, D.E. Goldberg, P.F. Egea, Z.H. Zhou, Malaria parasite translocon structure and mechanism of effector export, *Nature* 561 (2018) 70–75. <https://doi.org/10.1038/s41586-018-0469-4>.
33. B. Elsworth, K. Matthews, C.Q. Nie, M. Kalanon, S.C. Charnaud, P.R. Sanders, S.A. Chisholm, N.A. Counihan, P.J. Shaw, P. Pino, J.A. Chan, M.F. Azevedo, S.J. Rogerson, J.G. Beeson, B.S. Crabb, P.R. Gilson, T.F. de Koning-Ward, PTEX is an essential nexus for protein export in malaria parasites, *Nature* 511 (2014) 587–591. <https://doi.org/10.1038/nature13555>.
34. J.R. Beck, V. Muralidharan, A. Oksman, D.E. Goldberg, PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes, *Nature* 511 (2014) 592–595. <https://doi.org/10.1038/nature13574>.
35. M. Marti, R.T. Good, M. Rug, E. Knuepfer, A.F. Cowman, Targeting malaria virulence and remodeling proteins to the host erythrocyte, *Science* 306 (2004) 1930–1933. <https://doi.org/10.1126/science.1102452>.
36. J.N. Gitaka, M. Takeda, M. Kimura, Z. Md Idris, C.W. Chan, J. Kongere, K. Yahata, F.W. Muregi, Y. Ichinose, A. Kaneko, O. Kaneko, Selections, frameshift mutations, and copy number variation detected on the *surf*<sub>4.1</sub> gene in the western Kenyan *Plasmodium falciparum* population, *Malar. J.* 16 (2017) 98. <https://doi.org/10.1186/s12936-017-1743-x>.

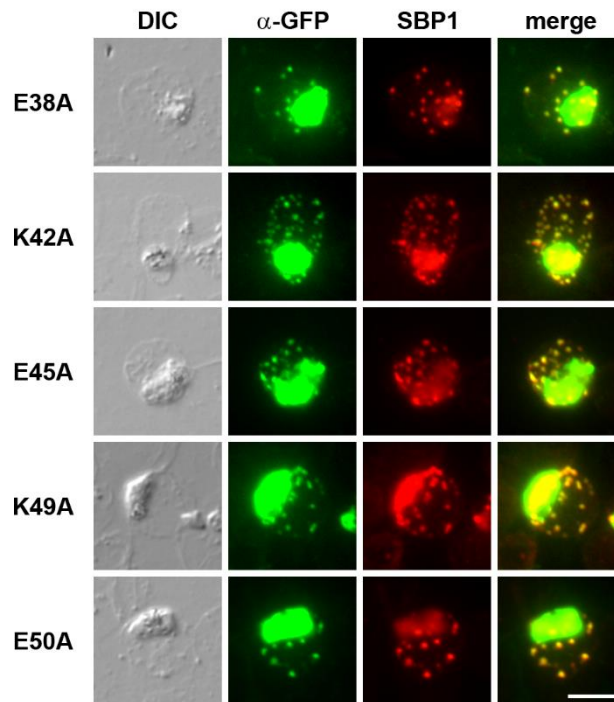


**Figure 1. The effect of Ala replacements within the SURFIN<sub>4.1</sub>-NTC N-terminal region for export in the *P. falciparum* line MS822.** (A) Schematic of full length SURFIN<sub>4.1</sub> with three tryptophan-rich domains (WRD) [36] and SURFIN<sub>4.1</sub><sup>N4-T-Cyt</sup> (N4TC) protein with corresponding amino acid sequences. The N4 region is underlined and the predicted transmembrane domain (T, 23 aa) is boxed. Lower case letters indicate amino acid residues translated from the plasmid backbone sequence. Purple letters indicate the cytosolic region (C, 19 aa). Six negatively (red) and 3 positively (blue) charged amino acid residues in the N4 region (17 aa) are indicated. N, N-terminal region; CRD, cysteine-rich domain; and GFP, green fluorescent protein. Downward arrows indicate charged amino acids mutated to Ala. (B) Western blot of SURFIN<sub>4.1</sub>-NTC variants with mouse anti-Ty1 antibody. Anti-HSP70 was used as a loading control marker. Arrow indicates the expected size of the N4TC protein. (C) Live cell imaging of MS822-background transfectants expressing modified N4TC protein. N4TC serves as a background parasite. Scale bar = 5 μm. (D) Quantification of signal in the iRBC cytoplasm. Significance was evaluated by Kruskal-Wallis one way analysis of variance followed by Dunn's multiple comparison test (p values are indicated). Twenty infected RBCs were evaluated for each line.

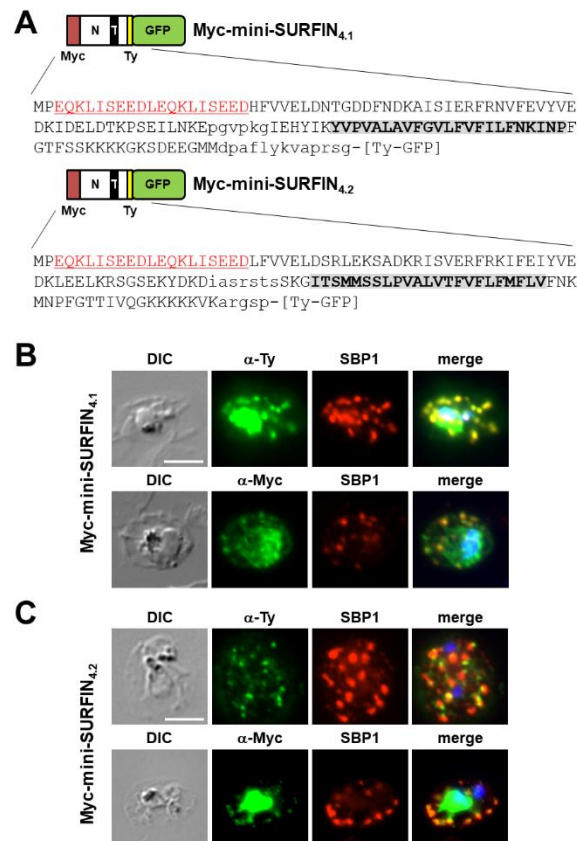


**Figure 2. The effect of Ala replacements within the SURFIN<sub>4.1</sub> N-terminal region for export in the *P. falciparum* line 3D7S8.** (A) Western blot of N4TC variants with mouse anti-Ty1 antibody. HSP70 was used as a loading control marker. Arrow indicates the expected size of N4TC protein. (B) Live cell imaging of MS822-background transfectants expressing modified SURFIN<sub>4.1</sub><sup>N4-T-Cyt</sup> (N4TC) protein. Scale bar = 5  $\mu$ m. (C) Quantification of signal in the iRBC cytoplasm. Significance was evaluated by Kruskal-Wallis one way analysis of variance followed by Dunn's multiple comparison test (p values are indicated). Twenty infected RBCs were evaluated for each line. (D) Amino acid residues for which Ala replacement made a significant loss of signal export are shown with enlarged letters. Blue and red letters indicate negatively and positively charged amino acids, respectively.





**Figure 3. IFA of the five transgenic parasites that showed low exported signal by live cell imaging.** Parasites were dual-stained with anti-GFP (green) and anti-SBP1 (red) antibodies. Green signals in the iRBC cytoplasm colocalized with red signals of the Maurer's cleft marker, SBP1. The signal intensities of the green color were intentionally increased to clearly show the location of the stained-dots in the iRBC cytoplasm. Scale bar = 5  $\mu$ m.



**Figure 4. Myc-tag fused to the N-terminus of mini-SURFIN<sub>4.1</sub> and mini-SURFIN<sub>4.2</sub> proteins are detected in Maurer's clefts.** (A) Schematics of Myc-mini-SURFIN<sub>4.1</sub> and Myc-mini-SURFIN<sub>4.2</sub> proteins. Myc tag sequences are underlined with red color and predicted transmembrane domain (T) are boxed. N-terminal regions (N, 49 aa) of SURFIN<sub>4.1</sub> and SURFIN<sub>4.2</sub> were used. (B and C) IFA images of the transfectants expressing Myc-mini-SURFIN<sub>4.1</sub> (B) or Myc-mini-SURFIN<sub>4.2</sub> (C) with anti-Ty1 or anti-Myc antibodies co-stained with anti-SBP1 Maurer's cleft marker antibody. Scale bar = 5  $\mu$ m.

**Supplementary Table S1.** Oligonucleotides used in this study for *surf*<sub>4.1</sub> and *surf*<sub>4.2</sub> genes

Name	Sequence
<b><i>For site-directed mutagenesis</i></b>	
SURF4.1-NC4-D34A_F	caAAAATAGATGAATTAGATACAAAAC
SURF4.1-NC4-D34A_R	cCATTGTTTATTAGCCTGCTTTTTTG
SURF4.1-NC4-K35Ar_F	caATAGATGAATTAGATACAAAAC
SURF4.1-NC4-K35Ar_R	cGTCCATTGTTTATTAGCCTGC
SURF4.1-NC4-D37A_F	caGAATTAGATACAAAACCATCAG
SURF4.1-NC4-D37A_R	cTATTTTGTCCATTGTTTATTAGC
SURF4.1-NC4-E38A_F	caTTAGATACAAAACCATCAGAAATAC
SURF4.1-NC4-E38A_R	cATCTATTTTGTCCATTGTTTATTAGC
SURF4.1-NC4-D40A_F	cTACAAAACCATCAGAAATACTTAATAAAG
SURF4.1-NC4_D40A_R2	cTAATTCATCTATTTTGTCCATTG
SURF4.1-NC4-K42A_F	caCCATCAGAAATACTTAATAAAG
SURF4.1-NC4-K42A_R	cTGTATCTAATTCATCTATTTTGTCC
SURF4.1-NC4-E45A_F	caATACTTAATAAAGAACCCGGG
SURF4.1-NC4-E45A_R	cTGATGGTTTTGTATCTAATTCATC
SURF4.1-NC4-K49A_F	caGAACCCGGGGTCCCAAAGGC
SURF4.1-NC4-K49A_R	cATTAAGTATTTCTGATGGTTTTG
SURF4.1-NC4-E50A_F	caCCCGGGGTCCCAAAGGCATTGAAC
SURF4.1-NC4-E50A_R	cTTTATTAAGTATTTCTGATGGTTTTG
<b><i>For N-terminal myc-tag</i></b>	
SURF4.1myc_F	ctggaacagaagttaataagtgaggaagacCATTTTGTAGTTGAATTGG
SURF4.1myc_R	atcctctctgagatgagttttgttcgggcatTGTTTATTAGCCTGC
SURF4.2myc_F	ctggaacagaagttaataagtgaggaagacCTTTTTGTTGTTGAGCTCGACAGCAGATTG
SURF4.2myc_R	atcctctctgagatgagttttgttcgggCATattatcaATCTAAGCCTGCTTTTTTTGTAC
SURF4.2Nter_F	ATGCTTTTTGTTGTTGAGCTCGACAGCA
SURF4.2Nter_R	ATCCTTATCATACTTTTCAGATCCTGACC
SURF4.2Var2_F	<u>TCTAGATTTGATGGAAGATATATAAATGTTGATATATAAGAGGT</u>
SURF4.2Cyt.R	<u>TCTAGAATTTTACTTTTTTTTTTTTCTTCCCTTGAACAATTGTT</u>
SURF4.2intronF	ATGAATCCTTTTGGTACAACAATTGTTCAA
SURF4.2intronR	CTTATTAAATACCAAAAACATAAAAAGAAATACAAATG
SURF4.2miniF	TCAAAGGGATAACATCAATGATGTCGTCTC
SURF4.2miniR	ACTAGTAGATCTGCTAGCGATATCCTTATCA
SURF4.2.corF	gCTAGAGGATCCCCGGAAGTACATACTAAT
SURF4.2.corR	TTTTACTTTTTTTTTTTTCTTCCCTTGAACAATTGTT

Lower case letters indicate codons replaced for alanine or inserted nucleotides. Underlined letters indicate XbaI sites.