

1 **The pseudogene SURFIN 4.1 is vital for merozoite formation in blood stage *P. falciparum***

2

3 Tatiane Macedo-Silva¹, Rosana Beatriz Duque Araujo¹, and Gerhard Wunderlich^{1,*}

4

5 1 Department of Parasitology, Institute for Biomedical Sciences, University of São Paulo,
6 Avenida Professor Lineu Prestes, 1374, 05508-000 São Paulo, Brazil

7 * correspondence to: Gerhard Wunderlich, gwunder@usp.br

8

9 keywords: *surf* genes, *Plasmodium*, stop codon readthrough, pseudogenes

10

11 **Abstract**

12 The *surf* gene family of the human malaria parasite *Plasmodium falciparum* encodes for
13 antigens with largely unknown functions. Three of the ten *surf* genes found in the *P. falciparum*
14 3D7 genome are annotated as pseudogenes, and one of these – *surf4.1* (PF3D7_0402200) -
15 was continuously transcribed in *P. falciparum* 3D7 blood stage forms. GFP-tagging revealed
16 that despite several stop codons a full-length protein was expressed, which localized to
17 developing merozoites. Analysis of cDNAs showed that no specific editing occurred pointing to
18 readthrough of stop codons during translation. Intriguingly, attempts to generate parasite lines
19 containing an additional artificial stop codon failed. Transcript knockdown revealed that *surf4.1*
20 is essential for merozoite formation in late trophozoite/schizont stages while DNA replication
21 seemed not to be influenced. SURFIN4.1 is the first example of a plasmodial multigene family
22 member of which a knockout is deleterious and may pose as a novel target for anti-malarial
23 therapy.

24

25 **Introduction**

26 Malaria is still one of major infectious diseases in the world and despite huge advances in the
27 control of the disease, a recent stalling in the decrease of cases is alarming (World Health
28 Organization, 2018). Malaria is caused by apicomplexan protozoans of the genus *Plasmodium*
29 and the most lethal species for humans is *Plasmodium falciparum*. It is estimated that only in
30 2017 435,000 people died of malaria and 219 million new cases were reported. Most of the
31 victims occur in sub-Saharan Africa and are children younger than 5 years (World Health
32 Organization, 2018).

33 In the vertebrate host, parasites invade red blood cells and go through intracellular multiplication
34 by a process termed schizogony which culminates in the lysis of the infected red blood cell
35 (IRBC) and the liberation of up to 32 merozoites which instantly reinvade erythrocytes. The
36 release of toxic byproducts and cytoadherence of infected red blood cell to other uninfected red
37 blood cells or receptors in deep venules are the main triggers for the severe outcomes of *P.*
38 *falciparum* malaria in individuals without adequate immunity against infection (revised in (van
39 der Heyde et al., 2006; Milner, 2018)). In the parasite genome, a number of multigene families
40 are present, and in many cases, these encode variant proteins which are involved in host
41 immune response evasion, and also in pathogenic processes, such as PfEMP1 (Baruch et al.,
42 1995; Smith et al., 1995; Su et al., 1995), RIFINs (Fernandez et al., 1999), STEVOR (Cheng et
43 al., 1998) and others (revised in (Wahlgren et al., 2017)).

44 Another, smaller gene family comprises *surf* genes (surface-associated interspersed family)
45 (Winter et al., 2005), which appear in different numbers in the *P. falciparum* strains sequenced
46 so far. In contrast to *rif* and *var* genes which encode RIFINs and PfEMP1 antigens, respectively,
47 *surf* genes are apparently more conserved between different strains and probably do not
48 undergo frequent ectopic recombination as shown for *var* genes (Freitas-Junior et al., 2000).
49 Additionally, in *Plasmodium vivax*, the PvSTP gene family (del Portillo et al., 2001) shows
50 significant similarities to *surf* genes, considering the domain structure of a small variant
51 ectodomain-encoding part and a larger, tryptophan-rich regions-encoding domain (Winter et al.,
52 2005). SURFINs are large antigens with more than 200 kDa. In *P. falciparum* strain 3D7, ten
53 genes encode *surf* genes and three of these are annotated as pseudogenes. The presence of a
54 modified export motif (PEXEL/VTS (Hiller et al., 2004; Marti et al., 2004)) in the N-terminal
55 region of SURFINs in some of the alleles suggested that SURFINs are exported to the IRBC
56 surface and first results analyzing the allele SURFIN 4.2 (Winter et al., 2005) revealed that this
57 protein was present on merozoites and partially on IRBCs, colocalizing with PfEMP1. In a recent
58 study, it was observed that SURFIN 4.2 from the FCR3 line forms complexes with RON4
59 (rhoptry neck protein 4) and GLURP (Glutamate-rich protein). Also, antibody-mediated inhibition
60 of RBC invasion was documented, suggesting a role in parasitophorous vacuole formation.
61 Intriguingly, *surf4.2* from the CS2 line can be knocked out without any growth defect (Maier et
62 al., 2008) and recent genome-wide mutational analysis indicated that *surf4.2* may be
63 dispensable with a small fitness cost upon insertion-mutation (Zhang et al., 2018). Few data are
64 available regarding other *surf* alleles. Mphande and colleagues confirmed the localization of
65 SURFIN 4.1 on the merozoite surface but not on the surface of IRBC (Mphande et al., 2008).
66 The same authors also detected that *surf4.1* occurred in six copies in the FCR3 strain, while
67 only one copy was present in the strain 3D7 and the Brazilian isolate 7G8. Later, Ochola and
68 colleagues detected in a genome-wide approach that *surf4.1* is under selective pressure in
69 circulating field isolates meaning that the protein is a target of the naturally acquired immune
70 response (Ochola et al., 2010). A similar analysis examining *surf4.1* ectodomain-encoding
71 sequences from Thai isolates showed diversifying selection, although the frequency of allele
72 distribution was stable over years in the analyzed isolates (Xangsayarath et al., 2012). More
73 recently, Gitaka and authors analyzed the surprisingly high occurrence of frameshift mutations
74 in the *surf4.1* open reading frame (ORF) in field samples and hypothesized that this may lead to
75 truncated versions of the protein (Gitaka et al., 2017). Similar to *surf4.2*, *surf4.1* was deemed
76 dispensable in genome-wide insertion-mutation assays (Zhang et al., 2018), although insertion
77 was only observed near the 3' end of the *surf4.1* open reading frame. During an effort to

78 analyze the mode of transcription of the *surf* gene family and eventually observe transcriptional
79 switching, we generated NF54-based parasite lines (isogenic with strain 3D7) with either a
80 green fluorescent protein-hemagglutinin tag or a degrading domain, variant DD24 (de Azevedo
81 et al., 2012), fused to tagged to the SURFIN4.1 polypeptide. Although there are a number of
82 stop codons in the NF54/3D7 *surf4.1*-ORF, green fluorescence was readily observed and in
83 western blots, a full-length protein was detected (Macedo-Silva et al., 2017). Also, all parasites
84 were positive for GFP presence, ruling out parasites with only truncated versions of the protein.
85 In the same study, significant transcriptional activity was found from the *surf4.1* locus, and other
86 loci were only sporadically activated during multiple reinvasions. Further, although an 80%
87 protein knockdown of SURFIN4.1 led to no discernible growth defect, a significant increase in
88 the steady-state levels of the *surf4.1* transcript was observed during knockdown which returned
89 to normal values when the protein was reestablished (Macedo-Silva et al., 2017). This
90 phenomenon may be interpreted that the parasite counterbalances the lack of viable protein by
91 increasing transcription. If so, a further decrease of SURFIN 4.1 may be deleterious for parasite
92 survival. Coincident with this, a successful knockout of *surf4.1*, if attempted, was never reported.
93 Together with all previous observations, these results point to a specific role of *surf4.1* at least in
94 the NF54 or 3D7 line. Here, we addressed the biological function of *surf4.1* by introducing an
95 efficient knockdown at the transcriptional level and provide evidence that *surf4.1* is indeed an
96 essential protein, probably involved in merozoite formation during red blood cell schizogony.

97 **Material and Methods**

98

99 **Plasmid constructs**

100 DNA fragments encoding parts of the *surf4.1* gene (PlasmoDB ID PF3D7_0402200) were
101 amplified by PCR (See Supplementary Table 1 for sequences) using Elongase proofreading
102 enzyme (Invitrogen). The amplicons were cloned in pGEM T-easy vectors (Promega) and
103 sequenced. The SURFIN4.1-3' ORF encoding fragment was excised using Bgl2 and Pst1 and
104 transferred to a modified pRESA-GFP-HA vector (de Azevedo et al., 2012) digested with the
105 same enzymes. This vector had the glmS element inserted downstream of the stop codon,
106 resulting in pS4GFPHAglmS. The plasmid pS4GFPHAglmSTK was cloned using two homology
107 regions for double crossover recombination. For this, the *surf4.1* 3' ORF fragment was excised
108 together with the GFP-HAglmS encoding region plus its terminator via BglIII and EcoRV and
109 inserted in the same site in pHHTK (Duraisingh et al., 2002). Afterwards, the 3'UTR was
110 amplified, cloned in pGEM T easy, sequenced and transferred to the pHHTK vector using NcoI

111 and ClaI. The plasmid pS4(TAA)GFPHAgImS TK was constructed similarly with the exception
112 that a fusion-PCR was performed to introduce a single nucleotide exchange (T→A) at position
113 5663. The fused, mutated 3'-ORF fragment was transferred into pS4GFPHAgImS (via BglIII and
114 NheI), resulting in pS4(TAA)GFPHAgImS, from which it was transferred via BglIII/EcoRV to
115 pHHTK already containing the 3'UTR homology region to create pS4(TAA)GFPHAgImS TK. The
116 construction of pS4KO for 5' and 3' ORF was done using the pHHTK backbone inserting the
117 respective fragments into the BglIII/EcoRI and SpeI/NcoI restriction sites. Recombinant
118 plasmids were grown to high quantities using the Maxiprep protocol (Sambrook, 1991) and used
119 for transfections. All oligonucleotides used in the amplification and cloning steps are informed in
120 Supplementary Table 1.

121 **Parasite culture and transfection**

122 *Plasmodium falciparum* lineage NF54 (Walliker et al., 1987), kindly provided by Mats Wahlgren
123 (Karolinska Institutet, Sweden) was used throughout the experiments. Blood stage parasites
124 were maintained in RPMI supplemented with 0,23% NaHCO₃, 0.5% Albumax 1 (Gibco,
125 Rockville MD) and human B+ erythrocytes in a defined gas mixture (90% N₂, 5% O₂ and 5%
126 CO₂) (Trager and Jensen, 1976). The synchronization of parasites was done by plasmagel
127 flotation (Lelievre et al., 2005) of mature trophozoites followed by sorbitol lysis (Lambros and
128 Vanderberg, 1979) of ring stage forms. Transfection of schizont stage parasites was done using
129 the protocol published by Hasenkamp and colleagues (Hasenkamp et al., 2012). Transfected
130 parasites were grown using 2.5 nM WR99210 (a gift from Jacobus Inc., USA). For the
131 integration via single crossover recombination, transfected parasite lines were cultivated for 14-
132 20 days without WR99210, after which the drug was added again. Normally, after three cycles
133 locus-integrated parasite lines were obtained. These were cloned by limiting dilution.
134 Experiments were done in biological duplicates or triplicates. For the integration via double
135 recombination, following the outgrowth of transfected parasites, these were cultivated for 15
136 days without WR99210 and after re-adding the drug, the parasitemias were adjusted to 1%.
137 Then, 2.2 μM ganciclovir was added into the culture for 6 days, followed by 4 days at ganciclovir
138 concentrations of 4.4 μM and 4 days at 8.8 μM. Afterwards, genomic DNA of parasites was
139 extracted to check for successful integration and elimination of episomal plasmid forms.

140

141 **Knockdown assay**

142 To knockdown protein expression in SURFIN4.1GFPHAgImS parasites, D-glucosamine 6-
143 phosphate (Sigma-Aldrich G5509) was added to a final concentration of 2.2 mM to highly
144 synchronous ring stage parasites (6 h post reinvasion, hpi). Parasites were kept in culture until

145 the schizont phase (38-46 hpi), monitored by microscopy. For protein extraction, pelleted red
146 blood cells were treated with 0.1% saponin in PBS (supplemented with Complete Protease
147 Cocktail Inhibitors (Roche)) to remove hemoglobin. Proteins were electrophoresed on standard
148 discontinuous SDS-polyacrylamide gels, transferred to Hybond C membranes (Amersham) and
149 analyzed as described below. The same saponin lysis process was done for RNA extraction
150 applying 1 ml of Trizol reagent (Life Technologies) to pre-purified pelleted parasites and
151 subsequent storage at -80°C until use.

152

153 **Sequencing of genomic DNA and complementary DNA**

154 Total genomic DNA of parasites was extracted using the Wizard® Genomic DNA Purification Kit
155 (Promega) following the manufacturer's instructions. cDNA was synthesized from total RNA
156 purified from Trizol-treated samples (Life Technologies) following the manufacturer's
157 instructions (see below). From resulting cDNA and gDNAs, fragments of *surf4.1* spanning the
158 sequence of interest were amplified by PCR using Elongase (Invitrogen) with primers specified
159 in Supplementary Table 1 and cloned into vector pGEM-T Easy (Promega). A number of
160 resulting clones was Sanger-sequenced and analyzed using the BLAST tool at PlasmoDB
161 (www.plasmodb.org). Sequence alignments were also done using ClustalX2.1 (Larkin et al.,
162 2007).

163

164 **Immunoprecipitation (IP) and mass spectrometry (MS)**

165 For IP, parasite protein extracts were prepared as described above. The Pierce Co-IP kit
166 (MACS® MicroBeads) was used following the manufacturer's instructions. In brief, anti-GFP
167 coupled to beads was added to parasite extracts from late schizonts and incubated for 2 h at
168 4°C. Afterwards, the samples were passed through a magnetic column followed by three
169 washes and two elution steps. Immunoprecipitated fractions were visualized by SDS-PAGE and
170 Coomassie blue staining. MS analyses were performed on eluted proteins on an LTQ-Orbitrap
171 Velos ETD (Thermo) coupled with Easy nanoLC II (Thermo). The peptides were separated on a
172 C18 RP column on a 115 min gradient. The instrumental conditions were checked using 100
173 fmol of a tryptic digest of BSA as standard. Peptides were identified using the ProteinDiscovery
174 software tool (Thermo) against a *P. falciparum* databank from Uniprot.

175

176 **Realtime qPCR**

177 For *surf* transcript quantification, 9 oligo pairs corresponding to the 3D7 *surf* genes available in
178 PlasmoDB (version 8) were used (Supplementary Table 1). Notably, *surf* genes 3 and 8 are

179 identical. Whole RNA was purified from synchronized stages (ring stage, directly after sorbitol
180 treatment, trophozoite stage, 20 h after sorbitol treatment, and schizont stage, 30 h after sorbitol
181 treatment) by the Trizol protocol and dissolved in pure water. RNAs were then treated with
182 DNase1 (Fermentas) and cDNA synthesis was done using RevertAid reverse transcriptase
183 (Fermentas) using random oligos as published earlier (Gölnitz et al., 2008). As an endogenous
184 control transcript, the plasmodial serine tRNA ligase transcript (PlasmoDB PF3D7_0717700),
185 was used.

186

187 **Immunoblotting**

188 Whole parasite protein extracts were prepared from saponin-lysed IRBCs as described in
189 Methods in Malaria Research (Ljungström et al., 2008). Proteins were loaded on standard
190 discontinuous SDS-polyacrylamide gels and transferred to Hybond C membranes (Amersham).
191 After blocking with 5% skimmed milk in 1xPBS/0.1% Tween20, HA-tagged proteins were
192 recognized using a murine antiHA antibody (Sigma-Aldrich) and then an antiMouse IgG-
193 peroxidase antibody (KPL). Blots were exhaustively washed with PBS/Tween between
194 incubations and finally incubated with ECL substrate (GE). As a loading control, a murine anti-
195 PTEX150 (generously provided by Dr. Mauro Azevedo) or an anti-Histone 3 antibody (Cell
196 Signaling Inc.) was used. Chemoluminescent signals were captured in an ImageQuant (GE)
197 apparatus and/or X-ray films and intensities were quantified using ImageJ software (NIH). The
198 obtained values were normalized using the PTEX150 or Histone 3 signals.

199

200 **Microscopy for GFP Fluorescence and Immunofluorescence**

201 To analyze GFP fluorescence, synchronous parasites were collected immediately before use.
202 After incubation with DAPI (5 µg/ml) for nuclear DNA staining, 20 µl of erythrocytes were placed
203 on a glass slide, covered with a coverslip, and analyzed on a Zeiss Observer Axio Imager M2
204 microscope with 1000x magnification.

205 For immunofluorescence, the samples were collected and prepared using the protocol
206 published by Tonkin and colleagues (Tonkin et al., 2004).

207

208 **Southern Blot**

209 10 µg genomic DNA of lineages NF54 (control) and NF54::pS4GFPHAgIms parasites, isolated
210 using the Wizard® Genomic DNA Purification Kit (Promega), and 25 ng of plasmid DNA, were
211 digested using BamHI (Fermentas/Thermo). The probe was amplified by PCR with digoxigenin-
212 dUTP from the DIG High Prime DNA Labeling and Detection Starter Kit I (Roche Diagnostics),

213 using the glms sequence as a template, amplified with oligos 5'-
214 TAATTATAGCGCCCGAACTAAGC-3' and 5'- AGATCATGTGATTTCTCTTTG-3'. Hybridization
215 and visualization were done following the manufacturer's instructions (Roche), using Hybond N
216 membranes (Amersham/GE Healthcare) at a hybridization temperature of 45 °C.

217

218 **Results**

219 **The genomic sequence and the reverse translated message of *surf4.1* have stop codons**

220 In order to confirm if the predicted stop codons were indeed present in the NF54 parasite line
221 used throughout our experiments, a number of plasmid clones containing amplicons from a
222 *surf4.1* region that contain the splicing site and stop codons were sequenced. The same was
223 done for plasmid clones made from cDNA-originated amplicons. Of note, a proofreading
224 enzyme was included in PCRs in order to avoid sequence artifacts. The used NF54 parasite line
225 constitutively transcribes *surf4.1* (Supplementary Figure 1).

226 Translating from the predicted start ATG, we observed that a first stop codon occurred at
227 nucleotides 2488-2490. The only difference found was a six-thymidine deletion (2447-2452)
228 inside a highly repetitive region in the intron, which may indicate that the used NF54 strain
229 differs at this point from the deposited 3D7 strain sequence in PlasmoDB (data not shown).
230 When analyzing the sequences from cDNA-derived amplicons, the splicing of the predicted
231 intron at nt 2371 to nt 2464 was observed, however, another predicted non-canonical two-base-
232 pair intron (2530-2531) was not spliced out (Supplementary Figure 2). When starting from the
233 predicted translation start ATG, this insertion leads to a stop codon (TGA) at this point and
234 several others from this point on. Moreover, all other possible ORFs resulted in stop codons.
235 This means that the parasite must employ read-through of stop codons when translating the
236 *surf4.1* into protein.

237

238 **Tagging of *surf4.1* with *gfp* leads to fluorescent forms of full length SURFIN4.1-GFP-HA**

239 To examine the presence of protein SURFIN4.1-tagged GFP in blood stage parasites, we
240 created an NF54 parasite line which had its *surf4.1* gene genetically tagged with a GFP-HA tag
241 (Figure 1). After drug on/off-cycling and cloning by limiting dilution, the transfectant parasite
242 lines containing an integrated version of the plasmid pSURF4-GFP-HA were PCR tested and no
243 amplification product was obtained with an oligo pair which detects episomal forms (Figure 1B).
244 When blotting a whole parasite extract of the transfectant line, SURFIN4.1-GFP-HA was
245 detected in its correct size using an antiHA antibody in Western blots (Figure 1C). The parasite
246 line containing an integrated version of the plasmid showed green fluorescence in late schizonts

247 in cytometry analyses (data not shown) as well as in fluorescence microscopy (Figure 1D). No
248 increase in abnormal parasite forms or extension or decrease of the blood stage cycle duration
249 was noted. This reinforces that a read-through of stop codons must occur during translation of
250 the *surf4.1*-gfp-ha transcript.

251

252 **A novel stop codon abrogates production of SURFIN4.1-GFP-HA**

253 To examine if the presence of an artificially introduced stop codon (TAA) could be introduced in
254 the *surf4.1* ORF, additional transfectant lines were prepared (Figure 2). The first one was
255 transfected with the plasmid pS4(TAA)GFPHAgImS-TK and the second with pS4GFPHAgImS-
256 TK. Both plasmids differ in i) one single base pair which introduces a stop codon in the middle of
257 the first homology region which contains the 3'-ORF of *surf4.1* (Figure 2 and Supplementary
258 Figure 3) and ii) the 3'-ORF homology region of pS4GFPHAgImS-TK is shorter. Following
259 outgrowth of the parasite lines and one cycling without/with WR99210 as described above, the
260 resulting parasite lines were treated with ganciclovir. Integration into the genome would then
261 insert a novel stop codon in the predicted translated ORF. However, after ganciclovir treatment
262 no parasites were recovered from the pS4(TAA)GFPHAgImS-TK transfected line, suggesting
263 that the insertion of a new stop codon into the genomic sequence of *surf4.1* severely interfered
264 with the survival of the parasite during blood stage. In contrast, the transfectant line bearing the
265 plasmid pS4GFPHAgImS-TK was readily recovered and showed rapid integration into the
266 *surf4.1* locus (Figure 2), as shown by PCR with integration-specific oligonucleotides. Of note,
267 while the stop-codon containing parasite line failed to show a knockin genotype (Figure 2, seen
268 as an amplicon with oligo pair III, gels with amplicons from genomic DNA of
269 NF54::pS4(TAA)GFPHAgImS-TK+Gcv, with and without WR), the specific presence of a
270 knockin genotype with the absence of an amplicon for the TK cassette indicated a successful
271 knockin resulting in the NF54::pS4GFPHAgImS-TK line. As controls, genomic DNA from NF54
272 and the transfection plasmid pS4GFPHAgImS-TK resulted solely in the expected PCR products
273 for the presence of the 1303 bp amplicon (oligos 3 and 4) or the TK cassette (931 bp, oligos 1
274 and 2). Southern blot analysis using BamH1-digested genomic DNA from the ganciclovir-treated
275 NF54::pS4GFPHAgImS-TK line confirmed the successful knockin and the expected 7786 bp
276 digestion product, stemming from a BamH1 site upstream of the knockin fragment of *surf4.1*
277 and a second site in the hDHFRcassette (Figure 2 C). In contrast, the transfection plasmid was
278 visualized in its linearized form, detected as a 11380 bp fragment.

279 To further confirm if *surf4.1* is essential, a conventional knockout by gene disruption was
280 attempted, using transfecting plasmid pS4KO (Supplementary Figure 3). Again, the transfected

281 parasite line did not survive ganciclovir treatment even after three drug-on/off cycles. It was also
282 impossible to detect any signal for integration in *surf4.1* by PCR on genomic DNA of the
283 transfected parasite line (data not shown).

284 **glmS-Glucosamine-mediated knockdown of *surf4.1* strongly inhibits schizont** 285 **development**

286 Given its probable importance for the survival of parasites during blood stages, we tested if a
287 transcript knockdown led to a measurable growth phenotype. In previous studies, the
288 SURFIN4.1 protein knockdown to approximately 20% of the normal SURFIN4.1 level resulted
289 only in an increase of the *surf4.1* transcript, but no apparent growth phenotype (Macedo-Silva et
290 al., 2017). To test for a growth phenotype, the integrated parasite line used as a control line in
291 the previous experiment (Figure 3) was submitted to glucosamine-6-phosphate treatment. First,
292 the efficiency of knockdown was monitored and RT-qPCR showed an approximately 10fold
293 decrease in the *surf4.1* message (Figure 3) while other surf transcripts were either silenced or
294 not altered (Supplementary Figure 4). Glucosamine treatment led also to a decrease in the
295 quantity of the SURFIN4.1-GFP-HA protein, while the PTEX polypeptide (loading control) was
296 not strongly influenced. Image analysis indicated a 70% knockdown of SURFIN4.1-GFP-HA
297 after 24 h treatment with glucosamine (Figure 3 C). The viability of parasites during knockdown
298 was also tested. During a 96 h treatment with glucosamine-6-phosphate, a significant growth
299 inhibition started apparently at the time point of schizont lysis and reinvasion at around 40 h
300 post treatment initiation (Figure 3 D), while treatment of untransfected NF54 parasites only
301 slightly suppressed parasite multiplication (Figure 3 E).

302

303 **Disruption of *surf4.1* impairs schizogony and parasite proliferation**

304 We then monitored the morphologic effects of parasites submitted to glmS-mediated
305 SURFIN4.1 knockdown. For this, SURFIN4.1-GFP-HA-expressing parasites were cultivated in
306 the presence of 2.2 mM of glucosamine starting in ring stage (6 hpi). A stalling of parasite
307 growth was observed (Figure 4) and the knocked down parasite seems to be unable to perform
308 a proper merozoite development, especially in late trophozoites/schizonts. Longer treatment
309 indicated that the stalled parasites did not complete schizogony and thus were unable to
310 reinvade. When parasite forms were counted, strong and significant inhibition of the schizont
311 development was observed after knockdown of the *surf4.1* transcript. Longer cultivation in the
312 presence of glucosamine-6-phosphate was impossible and parasites disappeared while
313 untreated parasites proliferated normally (data not shown).

314 **SURFIN4.1 probably interacts with tubulin and hsp90**

315 Quintana et al. detected that SURFIN4.2 from the *P. falciparum* IT strain forms a complex
316 between RON4 and GLURP and are probably involved in the formation of the parasite vacuole
317 at reinvasion (Quintana et al., 2018). If SURFIN4.2 from the IT/FCR3 strain and SURFIN4.1 of
318 the NF54 line are functionally equivalent, this interaction may be confirmed by mass
319 spectrometry analysis. Using antiGFP-sepharose, co-immunoprecipitating proteins were trypsin-
320 digested and analyzed by Orbitrap mass spectrometry. As a result, a number of proteins were
321 detected, including tubulin and heat shock protein 70 and 90 (Figure 5), but not RON4 or
322 GLURP. To support a possible interaction between β -tubulin and SURFIN4.1,
323 immunofluorescence microscopy was employed and a partial overlay of the fluorescence of
324 SURFIN4.1-GFP-HA and β -tubulin was visualized (Figure 5). Interestingly, the knockdown led to
325 the absence not only of GFP but also β -tubulin fluorescence. Also, in parasites under
326 knockdown conditions, a decrease in the production of β -tubulin is notable compared to
327 parasites without glucosamine 6 phosphate treatment (Figure 5 D). In the same samples, the
328 quantity of histone 3 is not influenced. Beta-tubulin is an essential part of microtubules which, in
329 turn, form an integral part of the cytoskeleton that is a dynamic set of long and thin fibers that
330 contribute to the transport of intracellular components and in cell division. The lack of
331 SURFIN4.1 seems to lead to a breakdown of the merozoite-forming process and a concomitant
332 decrease of DAPI fluorescence. This may point to a lack of mitotic activity. In order to establish
333 an influence of glucosamine treatment and SURFIN4.1 knockdown on the replication process,
334 genomic DNA of an identical number of glucosamine-treated and untreated parasites was
335 quantified. As shown in Figure 6, 20 h after the onset of glucosamine treatment (starting in ring
336 stage parasites), an increase of the DNA quantity could be identified compared to the initial
337 sample. However, the DNA quantities in glucosamine-treated or untreated parasites were not
338 significantly different, pointing to the view that replication itself may not have been influenced.
339 Taken together, the lack of SURFIN4.1 severely interferes with schizont development and this
340 may occur by influencing intracellular trafficking.

341 **Discussion**

342 The *surf* gene family and their orthologs are present in many primate species of *Plasmodium*.
343 Specifically, *surf4.1* is localized in a highly syntenic region in different *Plasmodium falciparum*
344 strains, but also in *Plasmodium gaboni* and *Plasmodium reichenowi*, where it is localized
345 between the genes encoding PHISTb and Reticulocyte Binding protein 1 homolog. Intriguingly,
346 and despite its apparent conservation, *surf4.1* is annotated in several sequenced strains as a

347 pseudogene, showing stop codons along its predicting open reading frame. Also, slightly
348 different patterns of splicing sites were predicted: While there is apparently no splicing in the
349 *surf4.1* variant of the *P. falciparum* GA01 strain, two introns were predicted for the SN01 strain
350 (PlasmoDB 41). Another strain (ML01 strain), also with no predicted intron, shows *surf4.2*
351 instead of *surf4.1* in the syntenic region. Here, we confirmed that the stop codons were also
352 present in the transcripts of *surf4.1*. Together with the observation that in the 3D7 strain there
353 are in total 160 genes annotated as pseudogenes of which for 29 exist proteomic evidence
354 (including SURFIN4.1), we conclude that readthrough of stop codons is a frequent event.
355 However, which amino acid is incorporated at the place where a stop codon occurs is unknown
356 and was not further addressed. Intriguingly, there must be an adaptation to the read-through of
357 stop codons, since an artificially introduced stop codon resulted in a non-viable phenotype. A
358 more detailed analysis, perhaps including a three-dimensional structure of the transcript, may
359 reveal patterns which allow the ribosome to read through stop codons. In many organisms, the
360 sequence surrounding stop codons clearly plays a role in the decision if a read through occurs
361 or not (Dabrowski et al., 2015). In several organisms, stop codons are translated into seleno-
362 cysteines (reviewed in (Rodnina et al., 2017)). In *P. falciparum*, a selenocysteine tRNA which
363 recognizes the UGA stop codon was described (Mourier et al., 2005). Later, Lobanov and
364 colleagues predicted four seleno-proteins, none of which is SURFIN4.1 (Lobanov et al., 2006).
365 Regarding *surf4.1* splicing variants, other authors also already documented uncommon splicing
366 variants of this gene (Zhu et al., 2013), and one resulted in the deletion of the tryptophan-rich
367 domains at the C-terminus of the 3D7 line. These authors also tested different domains for their
368 contribution to the subcellular trafficking of SURFIN4.1. However, in none of their constructs, the
369 complete, stop-codon containing open reading frame of *surf4.1* was tested, as was
370 successfully done in this work. Interestingly, Zhu and colleagues found a truncated variant of
371 SURFIN4.1 using an antibody raised against the N-terminal domain, but no full-length protein
372 was shown in their analyses (Zhu et al., 2013) using the untransfected 3D7A line. However, the
373 corresponding blot appears not to show very large proteins. In the same study, SURFIN4.1 was
374 detected in different sites depending on which domains were included for transgene expression
375 in *Plasmodium*, while the herein used GFP-tagged variant localized clearly and exclusively to
376 merozoites. The reason for this discrepancy is not clear and perhaps strain differences are
377 responsible for the observed differences. It is probable that we missed truncated variants in our
378 experiments since only full-length proteins were detectable in the experimental setup. In the
379 same way, it is possible that both the full length and truncated versions coexist and play

380 different biological roles in the parasite dependent on their localization, which may be
381 determined by the presence or absence of the C-terminal domain of the protein.

382 Regarding the function of the antigen SURFIN4.1, we have shown in a previous study that full-
383 length SURFIN4.1 could be knocked down without triggering any growth phenotype (Macedo-
384 Silva et al., 2017) and the only effect observed was the increase of its transcript. Initially, and to
385 further explore this phenomenon, we created a parasite line where the transcript could be
386 knocked down. Surprisingly, the knockdown of the *surf4.1* transcript had profound effects on the
387 schizont development and the observed phenotypes showed a stalling of the parasite in the late
388 trophozoite/early schizont stage, which was not observed in the SURFIN4.1-protein knockdown.
389 When comparing the knockdown efficiency of the full-length protein using either the degra-
390 (destabilizing domain variant “24” (de Azevedo et al., 2012; Macedo-Silva et al., 2017)) or the
391 transcript knockdown (Prommana et al., 2013) applied here, the overall knockdown efficiency
392 observed in western blots was comparable. One explanation may be that truncated forms as
393 observed by Zhu and colleagues have an important role in the effects observed here, while the
394 full-length protein has perhaps only a partial and accessory effect. The transcript knockdown
395 would suppress all resulting forms of the SURF4.1 protein, while the degra-mediated
396 knockdown only interferes with the full-length protein, which possesses the destabilizing domain
397 while truncated forms do not. The use of an antibody against an N-terminal region – not
398 available in this study - of SURFIN4.1 may elucidate this question.

399 We then addressed the function of SURFIN4.1. In a previous study, Quintana and colleagues
400 provided evidence that SURFIN4.2 interacts with GLURP and RON4, and blockage of SURFIN
401 by antibodies against its ectodomain led to a small but measurable decrease in the reinvasion of
402 merozoites into erythrocytes (Quintana et al., 2018). No data, however, are available what
403 missing SURFIN4.1 or IT/FCR SURFIN4.2 would cause in the formation of the postulated
404 complex. By knocking down the *surf4.1* transcript, a profound perturbation of merozoite
405 formation and stalling of schizont development was observed, reminiscent of a block in cell
406 cycle progression. However, and documented in Figure 6, no significant interference in DNA
407 replication was found in ~24 h post reinvasion parasites treated or not with glucosamine-
408 phosphate, indicating that the observed stalling in early schizont stage was possibly due to the
409 structural or metabolic hindrance of merozoite formation, and not interference in genome
410 replication. In a recent study, genome-wide integration of insertions using the piggy-bac
411 approach was applied to characterize genes which were possibly essential in *P. falciparum*
412 (Zhang et al., 2018). In this approach, *surf4.1* was deemed amenable to insertion without
413 hampering the growth of the parasite. This apparent conflict may be explained by the fact that

414 piggyBac-mediated insertion occurred in sequences at 400 nt near the 3' end of the *surf4.1*
415 reading frame and therefore after the stop codons which lead to the possibly essential truncated
416 forms of SURFIN4.1, and also after the inserted stop codon introduced in
417 pS4(TAA)GFPHAglmS-TK. No other successful insertion was observed in the remaining
418 approximately 6000 base pairs upstream ((Zhang et al., 2018), compare supplementary data,
419 Table S1). RNAseq results deposited by different authors may give a hint at what stage different
420 SURFINs possibly exert functions. In the dataset published by Otto and colleagues (Otto et al.,
421 2010) the highest transcript quantities in trophozoites and schizonts were also found for *surf4.1*,
422 while another *surf* gene (PF3D7_0113100) was strongly detected in ookinetes. In another study
423 which approached transcript bound to polysomes, *surf4.1* was also the most abundantly present
424 surf transcript in schizonts (Bunnik et al., 2013). In yet another experiment monitoring transcripts
425 in the 3D7 strain, *surf4.1* was also the most abundantly found transcript in trophozoites and
426 schizonts (López-Barragán et al., 2011). Collectively, this reinforces a role for SURFIN4.1 at
427 least in the 3D7/NF54 parasite lines.

428 In order to detect interacting proteins, mass spectrometry was used to identify proteins which
429 interact with GFP-tagged SURFIN4.1. While heat shock proteins 70, 90 and beta-tubulin were
430 detected, neither RON4 nor GLURP could be identified. This may be due to the fact that the
431 complex between these factors is built only in later stages which are not formed when
432 SURFIN4.1 is knocked down, or else, full-sized SURFIN4.1 does not interact with these
433 proteins, differently from IT/FCR3's SURFIN4.2. We are currently addressing the question of
434 why enolase is possibly interacting with SURFIN4.1. In other studies, plasmodial enolase was
435 identified as a multifunctional protein, found in different subcellular localizations (Pal Bhowmick
436 et al., 2009) and possibly interacting with a number of different factors, apparently including
437 SURFIN4.1. When we tried to confirm the colocalization of SURFIN4.1 and beta-tubulin,
438 knocked-down parasites did not reveal any signal of beta-tubulin at all, while a partially
439 overlapping fluorescence signal was observed for schizont stage parasites kept in the absence
440 of glucosamine phosphate. In immunoblots, beta-tubulin also appeared in decreased amounts,
441 indicating that less SURFIN4.1 also somehow leads to the suppression of beta-tubulin
442 production. Alternatively, knocked down parasites are unable to reach the phase when beta-
443 tubulin is strongly expressed. Of note, the expression of beta-tubulin varies during the blood
444 stage parasite development and beta-tubulin transcription is strongest in stages older than 20 h
445 post reinvasion (Otto et al., 2010) and the protein is weakly visible in younger forms (Fennell et
446 al., 2008). In two hybrid-studies, SURFIN4.1 appeared to interact predominantly with DNA
447 binding proteins such as zinc-finger nucleases, DNA binding chaperones, SET10 – a histone

448 lysine N-methyl transferase, putative kinetochore-suppressor protein, but also a Maurer's cleft
449 protein (ETRAMP) and a putative ABC transporter, besides others (LaCount et al., 2005). These
450 may not have been detected in our proteomic analyses due to their low abundance.

451 Taken together, while the exact function of SURFIN4.1 remains elusive we have shown that this
452 unusual protein possesses essential functions during the early schizogony phase of
453 intraerythrocytic development. Taking in account the differential expression of *surf* alleles in
454 different stages of *P. falciparum* forms including sporozoites and ookinetes, the functions of
455 each *surf* allele area probably non-redundant unlike other multigene families such as *var* and *rif*
456 and may reveal targets of intervention against this still devastating disease.

457

458 **Acknowledgments**

459 This work was supported by FAPESP grants 2015/17174-7 and 2017/24267-7. TMS and RBDA
460 were supported by CNPq fellowships. GW is a CNPq research fellow. The proteomic analysis
461 was performed at the CEFAP at ICB-USP.

462

463 **Authors contributions**

464 TMS and RBDA performed experiments. GW and TMS conceived the experimental outline and
465 wrote the manuscript.

466 **Literature**

467 de Azevedo, M.F., Gilson, P.R., Gabriel, H.B., Simões, R.F., Angrisano, F., Baum, J., Crabb,
468 B.S., and Wunderlich, G. (2012). Systematic Analysis of FKBP Inducible Degradation Domain
469 Tagging Strategies for the Human Malaria Parasite Plasmodium falciparum. PLoS One 7,
470 e40981.

471 Baruch, D.I., Pasloske, B.L., Singh, H.B., Bi, X., Ma, X.C., Feldman, M., Taraschi, T.F., and
472 Howard, R.J. (1995). Cloning the *P. falciparum* gene encoding PfEMP1, a malarial variant
473 antigen and adherence receptor on the surface of parasitized human erythrocytes. Cell 82, 77–
474 87.

475 Bunnik, E.M., Chung, D.-W., Hamilton, M., Ponts, N., Saraf, A., Prudhomme, J., Florens, L., and
476 Le Roch, K.G. (2013). Polysome profiling reveals translational control of gene expression in the
477 human malaria parasite Plasmodium falciparum. Genome Biol. 14, R128.

478 Cheng, Q., Cloonan, N., Fischer, K., Thompson, J., Waine, G., Lanzer, M., and Saul, A. (1998).
479 *stevor* and *rif* are Plasmodium falciparum multicopy gene families which potentially encode
480 variant antigens. Mol Biochem Parasitol 97, 161–176.

481 Dabrowski, M., Bukowy-Bieryllo, Z., and Zietkiewicz, E. (2015). Translational readthrough
482 potential of natural termination codons in eucaryotes--The impact of RNA sequence. RNA Biol.
483 12, 950–958.

484 Duraisingh, M.T., Triglia, T., and Cowman, A.F. (2002). Negative selection of Plasmodium
485 falciparum reveals targeted gene deletion by double crossover recombination. Int. J. Parasitol.

- 486 32, 81–89.
- 487 Fennell, B.J., Al-shatr, Z.A., and Bell, A. (2008). Isotype expression, post-translational
488 modification and stage-dependent production of tubulins in erythrocytic *Plasmodium falciparum*.
489 *Int. J. Parasitol.* 38, 527–539.
- 490 Fernandez, V., Hommel, M., Chen, Q., Hagblom, P., and Wahlgren, M. (1999). Small, clonally
491 variant antigens expressed on the surface of the *Plasmodium falciparum*-infected erythrocyte
492 are encoded by the rif gene family and are the target of human immune responses. *J. Exp. Med.*
493 190, 1393–1404.
- 494 Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., Nehrbass,
495 U., Wellems, T.E., and Scherf, A. (2000). Frequent ectopic recombination of virulence factor
496 genes in telomeric chromosome clusters of *P. falciparum*. *Nature* 407, 1018–1022.
- 497 Gitaka, J.N., Takeda, M., Kimura, M., Idris, Z.M., Chan, C.W., Kongere, J., Yahata, K., Muregi,
498 F.W., Ichinose, Y., Kaneko, A., et al. (2017). Selections, frameshift mutations, and copy number
499 variation detected on the surf 4.1 gene in the western Kenyan *Plasmodium falciparum*
500 population. *Malar. J.* 16, 98.
- 501 Gölnitz, U., Albrecht, L., and Wunderlich, G. (2008). Var transcription profiling of *Plasmodium*
502 *falciparum* 3D7: assignment of cytoadherent phenotypes to dominant transcripts. *Malar J* 7, 14.
- 503 Hasenkamp, S., Russell, K.T., and Horrocks, P. (2012). Comparison of the absolute and relative
504 efficiencies of electroporation-based transfection protocols for *Plasmodium falciparum*. *Malar. J.*
505 11, 210.
- 506 van der Heyde, H.C., Nolan, J., Combes, V., Gramaglia, I., and Grau, G.E. (2006). A unified
507 hypothesis for the genesis of cerebral malaria: sequestration, inflammation and hemostasis
508 leading to microcirculatory dysfunction. *Trends Parasitol* 22, 503–508.
- 509 Hiller, N.L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estraño, C., and
510 Haldar, K. (2004). A host-targeting signal in virulence proteins reveals a secretome in malarial
511 infection. *Science* (80-.). 306, 1934–1937.
- 512 LaCount, D.J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J.R., Schoenfeld,
513 L.W., Ota, I., Sahasrabudhe, S., Kurschner, C., et al. (2005). A protein interaction network of the
514 malaria parasite *Plasmodium falciparum*. *Nature* 438, 103–107.
- 515 Lambros, C., and Vanderberg, J.P. (1979). Synchronization of *Plasmodium falciparum*
516 erythrocytic stages in culture. *J Parasitol* 65, 418–420.
- 517 Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H.,
518 Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., et al. (2007). Clustal W and Clustal X version
519 2.0. *Bioinformatics* 23, 2947–2948.
- 520 Lelievre, J., Berry, A., and Benoit-Vical, F. (2005). An alternative method for *Plasmodium* culture
521 synchronization. *Exp Parasitol* 109, 195–197.
- 522 Ljungström, I., Perlmann, H., Schlichtherle, M., Scherf, A., and Wahlgren, M. (2008). *Methods in*
523 *Malaria Research* (Manassas, VA: MR4/ATCC, BioMalPar).
- 524 Lobanov, A. V, Delgado, C., Rahlfs, S., Novoselov, S. V, Kryukov, G. V, Gromer, S., Hatfield,
525 D.L., Becker, K., and Gladyshev, V.N. (2006). The *Plasmodium* selenoproteome. *Nucleic Acids*
526 *Res.* 34, 496–505.
- 527 López-Barragán, M.J., Lemieux, J., Quiñones, M., Williamson, K.C., Molina-Cruz, A., Cui, K.,

- 528 Barillas-Mury, C., Zhao, K., and Su, X. (2011). Directional gene expression and antisense
529 transcripts in sexual and asexual stages of *Plasmodium falciparum*. *BMC Genomics* 12, 587.
- 530 Macedo-Silva, T., Araujo, R.B.D., Meissner, K.A., Fotoran, W.L., Medeiros, M.M., de Azevedo,
531 M.F., and Wunderlich, G. (2017). Knockdown of the *Plasmodium falciparum* SURFIN4.1 antigen
532 leads to an increase of its cognate transcript. *PLoS One* 12, e0183129.
- 533 Maier, A.G., Rug, M., O'Neill, M.T., Brown, M., Chakravorty, S., Szeszak, T., Chesson, J., Wu,
534 Y., Hughes, K., Coppel, R.L., et al. (2008). Exported proteins required for virulence and rigidity
535 of *Plasmodium falciparum*-infected human erythrocytes. *Cell* 134, 48–61.
- 536 Marti, M., Good, R.T., Rug, M., Knuepfer, E., and Cowman, A.F. (2004). Targeting malaria
537 virulence and remodeling proteins to the host erythrocyte. *Science* (80-.). 306, 1930–1933.
- 538 Milner, D.A. (2018). *Malaria Pathogenesis*. Cold Spring Harb. Perspect. Med. 8, a025569.
- 539 Mourier, T., Pain, A., Barrell, B., and Griffiths-Jones, S. (2005). A selenocysteine tRNA and
540 SECIS element in *Plasmodium falciparum*. *RNA* 11, 119–122.
- 541 Mphande, F.A., Ribacke, U., Kaneko, O., Kironde, F., Winter, G., and Wahlgren, M. (2008).
542 SURFIN4.1, a schizont-merozoite associated protein in the SURFIN family of *Plasmodium*
543 *falciparum*. *Malar. J.* 7, 116.
- 544 Ochola, L.I., Tetteh, K.K.A., Stewart, L.B., Riitho, V., Marsh, K., and Conway, D.J. (2010). Allele
545 Frequency-Based and Polymorphism-Versus-Divergence Indices of Balancing Selection in a
546 New Filtered Set of Polymorphic Genes in *Plasmodium falciparum*. *Mol. Biol. Evol.* 27, 2344–
547 2351.
- 548 Otto, T.D., Wilinski, D., Assefa, S., Keane, T.M., Sarry, L.R., Böhme, U., Lemieux, J., Barrell, B.,
549 Pain, A., Berriman, M., et al. (2010). New insights into the blood-stage transcriptome of
550 *Plasmodium falciparum* using RNA-Seq. *Mol. Microbiol.* 76, 12–24.
- 551 Pal Bhowmick, I., Kumar, N., Sharma, S., Coppens, I., and Jarori, G.K. (2009). *Plasmodium*
552 *falciparum* enolase: stage-specific expression and sub-cellular localization. *Malar. J.* 8, 179.
- 553 del Portillo, H.A., Fernandez-Becerra, C., Bowman, S., Oliver, K., Preuss, M., Sanchez, C.P.,
554 Schneider, N.K., Villalobos, J.M., Rajandream, M.A., Harris, D., et al. (2001). A superfamily of
555 variant genes encoded in the subtelomeric region of *Plasmodium vivax*. *Nature* 410, 839–842.
- 556 Prommana, P., Uthaipibull, C., Wongsombat, C., Kamchonwongpaisan, S., Yuthavong, Y.,
557 Knuepfer, E., Holder, A.A., and Shaw, P.J. (2013). Inducible Knockdown of *Plasmodium* Gene
558 Expression Using the glmS Ribozyme. *PLoS One* 8, e73783.
- 559 Quintana, M.D.P., Ch'ng, J.-H., Zandian, A., Imam, M., Hultenby, K., Theisen, M., Nilsson, P.,
560 Qundos, U., Moll, K., Chan, S., et al. (2018). SURGE complex of *Plasmodium falciparum* in the
561 rhoptry-neck (SURFIN4.2-RON4-GLURP) contributes to merozoite invasion. *PLoS One* 13,
562 e0201669.
- 563 Rodnina, M. V, Fischer, N., Maracci, C., and Stark, H. (2017). Ribosome dynamics during
564 decoding. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 372, 20160182.
- 565 Sambrook, J. (1991). *Molecular cloning: a laboratory manual* (CSHL Press).
- 566 Smith, J.D., Chitnis, C.E., Craig, A.G., Roberts, D.J., Hudson-Taylor, D.E., Peterson, D.S.,
567 Pinches, R., Newbold, C.I., and Miller, L.H. (1995). Switches in expression of *Plasmodium*
568 *falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of
569 infected erythrocytes. *Cell* 82, 101–110.

- 570 Su, X.Z., Heatwole, V.M., Wertheimer, S.P., Guinet, F., Herrfeldt, J.A., Peterson, D.S., Ravetch,
571 J.A., and Wellems, T.E. (1995). The large diverse gene family var encodes proteins involved in
572 cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell* **82**,
573 89–100.
- 574 Tonkin, C.J., van Dooren, G.G., Spurck, T.P., Struck, N.S., Good, R.T., Handman, E., Cowman,
575 A.F., and McFadden, G.I. (2004). Localization of organellar proteins in *Plasmodium falciparum*
576 using a novel set of transfection vectors and a new immunofluorescence fixation method. *Mol.*
577 *Biochem. Parasitol.* **137**, 13–21.
- 578 Trager, W., and Jensen, J.B. (1976). Human malaria parasites in continuous culture. *Science*
579 (80-). **193**, 673–675.
- 580 Wahlgren, M., Goel, S., and Akhouri, R.R. (2017). Variant surface antigens of *Plasmodium*
581 *falciparum* and their roles in severe malaria. *Nat. Rev. Microbiol.* **15**, 479–491.
- 582 Walliker, D., Quakyi, I.A., Wellems, T.E., McCutchan, T.F., Szarfman, A., London, W.T.,
583 Corcoran, L.M., Burkot, T.R., and Carter, R. (1987). Genetic analysis of the human malaria
584 parasite *Plasmodium falciparum*. *Science* **236**, 1661–1666.
- 585 Winter, G., Kawai, S., Haeggström, M., Kaneko, O., von Euler, A., Kawazu, S., Palm, D.,
586 Fernandez, V., and Wahlgren, M. (2005). SURFIN is a polymorphic antigen expressed on
587 *Plasmodium falciparum* merozoites and infected erythrocytes. *J. Exp. Med.* **201**, 1853–1863.
- 588 World Health Organization (2018). WHO | World malaria report 2018. WHO.
- 589 Xangsayarath, P., Kaewthamasorn, M., Yahata, K., Nakazawa, S., Sattabongkot, J.,
590 Udomsangpetch, R., and Kaneko, O. (2012). Positive diversifying selection on the *Plasmodium*
591 *falciparum* surf4.1 gene in Thailand. *Trop. Med. Health* **40**, 79–89.
- 592 Zhang, M., Wang, C., Otto, T.D., Oberstaller, J., Liao, X., Adapa, S.R., Udenze, K., Bronner,
593 I.F., Casandra, D., Mayho, M., et al. (2018). Uncovering the essential genes of the human
594 malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science* **360**, eaap7847.
- 595 Zhu, X., Yahata, K., Alexandre, J.S.F., Tsuboi, T., and Kaneko, O. (2013). The N-terminal
596 segment of *Plasmodium falciparum* SURFIN4.1 is required for its trafficking to the red blood cell
597 cytosol through the endoplasmic reticulum. *Parasitol. Int.* **62**, 215–229.
- 598
- 599

600 **Legends to Figures**

601 **Figure 1: Tagging of SURFIN 4.1 with GFP and HA leads to green fluorescent parasites.**

602 In **A**, the proposed model for single crossover recombination of the plasmid pS4-GFP-HA is
603 shown. The arrows indicate the localization of oligonucleotides used for PCR. Grey arrows
604 indicate the oligonucleotides which were used to create the knockin-homology region in pS4-
605 GFP-HA. In **B**, PCR with the primers indicated above show knockin of the construct into the
606 *surf4.1* locus. The primer combinations used in PCRs and the amplicon sizes are indicated.
607 Note that primer pair II results in an amplicon solely in knocked-in parasites. In **C**, Western blot
608 of NF54::pS4-GFP-HA and NF54 show the full-length, tagged SURFIN4.1 in schizont stage in
609 the transfectant line but not NF54 wildtype parasites. On the right, the loading of each lane is
610 demonstrated by Ponceau-staining. In **D**, Fluorescence microscopy with a schizont from the
611 parasite line NF54::pS4GFPHA I, bright field, II, nuclear staining using DAPI, III, GFP-tagged
612 SURFIN4.1 and IV, overlay of I to III.

613

614 **Figure 2: A construct containing a novel stop codon in the 3' part of the *surf4.1* ORF is refractory to knockin.**

615 In **A**, the proposed model for double crossover recombination of the
616 plasmid pS4(TAA)-GFPHAglmS-TK is shown. The arrows indicate the localization of
617 oligonucleotides used for PCR and each red star indicates stop codons predicted in PlasmODB
618 and a black star indicates the position of the novel introduced stop codon. In **B**, genomic DNA
619 from transfectant lineages treated or not with Ganciclovir (Gcv) is characterized by PCR by the
620 indicated primer combinations. Note that primer pair I amplifies fragments from wild type and
621 knocked-in constructs. Amplification products using material from the parasite line transfected
622 with pS4(TAA)-GFPHAglmS-TK and treated with Gcv and afterwards with WR for 4 days are
623 shown in the 3rd panel from the left and appear devoid of any amplifiable material. The panels
624 on the right show amplicons from either untransfected NF54 genomic DNA, or transfectants
625 using the pS4-GFPHAglmS-TK before and after Gcv treatment. In **C**, Southern blot analysis
626 using the glmS specific probe (black bar in **A**) shows integration as expected of the constructs
627 and the virtual absence of episomal material. "P" contains BamHI-linearized plasmid and "T"
628 contains 10 µg BamHI restricted genomic DNA from NF54::pS4-GFPHAglmS-TK.

629

630 **Figure 3: Transcript knockdown of *surf4.1* results in a decrease of *surf4.1* transcripts, SURFIN4.1-GFP-HA protein and impairs parasite growth.**

631 In **A**, knockdown of the *surf4.1*
632 transcript measured by RT-qPCR in NF54::pS4-GFPHAglmS-TK transfectants treated or not
633 with 2.2 mM glucosamine-phosphate in relation to the endogenous control seryl-tRNA ligase. In
634 **B**- Western blot of NF54::pS4-GFPHAglmS and NF54 in schizont stage and in the presence and
635 absence of glucosamine-phosphate, showing SURFIN4.1-GFP-HA (representative of 2
636 individual experiments) and untransfected NF54 (no detectable signal in 260 kDa, due to the
637 absence of the HA tag). The loading control below was done with anti-PTEX150. In **C**,
638 densitometry analysis of the observed signals in B using ImageJ, normalized against the

639 PTEX150 signal. Three asterisks mean highly significant differences ($p= 0.0006$, Student's T-
640 test, with a mean of the difference of 1.472 relative units). **D**: Effect on the parasitemia of GlcN
641 and subsequent absence of SURFIN4.1 is shown. Error bars indicate the standard deviation of
642 transcription of each surf gene in individual triplicates. Three asterisks mean highly significant
643 differences ($p < 0.0001$, Two-Way ANOVA). **E**: A slight negative influence of Glucosamine
644 phosphate on the growth of wildtype NF54 cultures is shown.

645

646 **Figure 4: Parasites with knocked-down SURFIN4.1 show a strongly reduced capacity of**
647 **parasites to progress to mature schizonts.** **A**: Highly synchronized parasites (2 sequential
648 cycles of floating/sorbitol treatment) were submitted to Glucosamine phosphate treatment or
649 not. The progression of intraerythrocytic parasite development was examined by thin blood
650 smears stained with Giemsa (representative data of 3 experiments). **B**: Counting of slides
651 resulted in the observation that parasites stop to develop in late-stage trophozoites.

652

653 **Figure 5: Co-IP by Pull-down and mass spectrometry of SURFIN4.1 and**
654 **Immunofluorescence showing the effect on the tubulin structure and nuclear increase**
655 **and division (DAPI mark) after disruption of SURFIN4.1.** In **A**, left side, Colloidal Coomassie
656 blue stained SDS-PAGE (8%) was loaded with a third of the eluted material from
657 immunoprecipitated lysed parasites (2 ml compacted red blood cells at 8% parasitemia late
658 trophozoites/schizonts) and a second third was run on the same gel, transferred onto
659 nitrocellulose and later GFP-containing proteins were detected with antiGFP as described. The
660 third part of the eluted material was trypsin digested and analyzed by mass spectrometry
661 resulting in the co-precipitating protein species as shown. Listed are only proteins which
662 produced peptide fragments coincident with unique peptides from the *P. falciparum* databank
663 (**B**). The complete curated output from ProteinDiscovery can be accessed in Supplementary
664 Table 2. In **C**, parasites were examined for SURFIN4.1-GFP and beta-tubulin colocalization.
665 Note that treatment with glucosamine phosphate (GlcN) strongly decreases both SURFIN4.1-
666 GFP and beta-tubulin detection. In **D**, western blot detection with the indicated antibodies
667 revealed a strong decrease in beta-tubulin presence while Histone 3 is not influenced.
668 Quantification using ImageJ (**E**) resulted in only around 30% of the beta-tubulin signal in GlcN
669 treated parasites.

670

671 **Figure 6: Knockdown of SURFIN4.1-GFP does not interfere with DNA replication.**
672 Parasites were treated or not with 2.5 mM GlcN for 24 h starting in early ring stage. An equal
673 number of parasites had their genomic DNA extracted at timepoint 0 and 24 h. Ct values were
674 obtained for the single copy gene (seryl-tRNA ligase) and compared between each other,
675 assuming that each Ct unit difference equals the double amount of input DNA. Experiments
676 were done in triplicate.

677

Figure 1

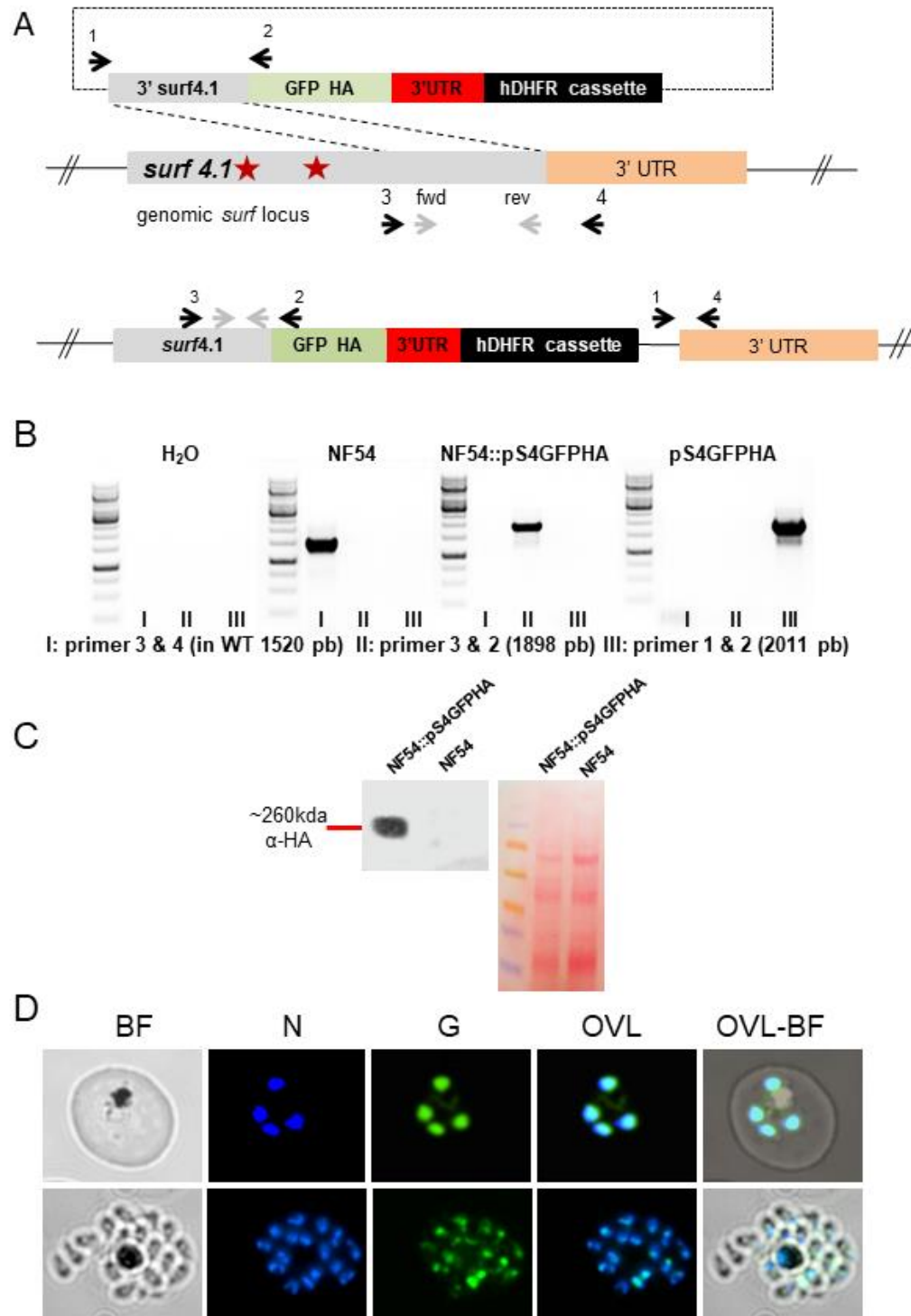


Figure 2

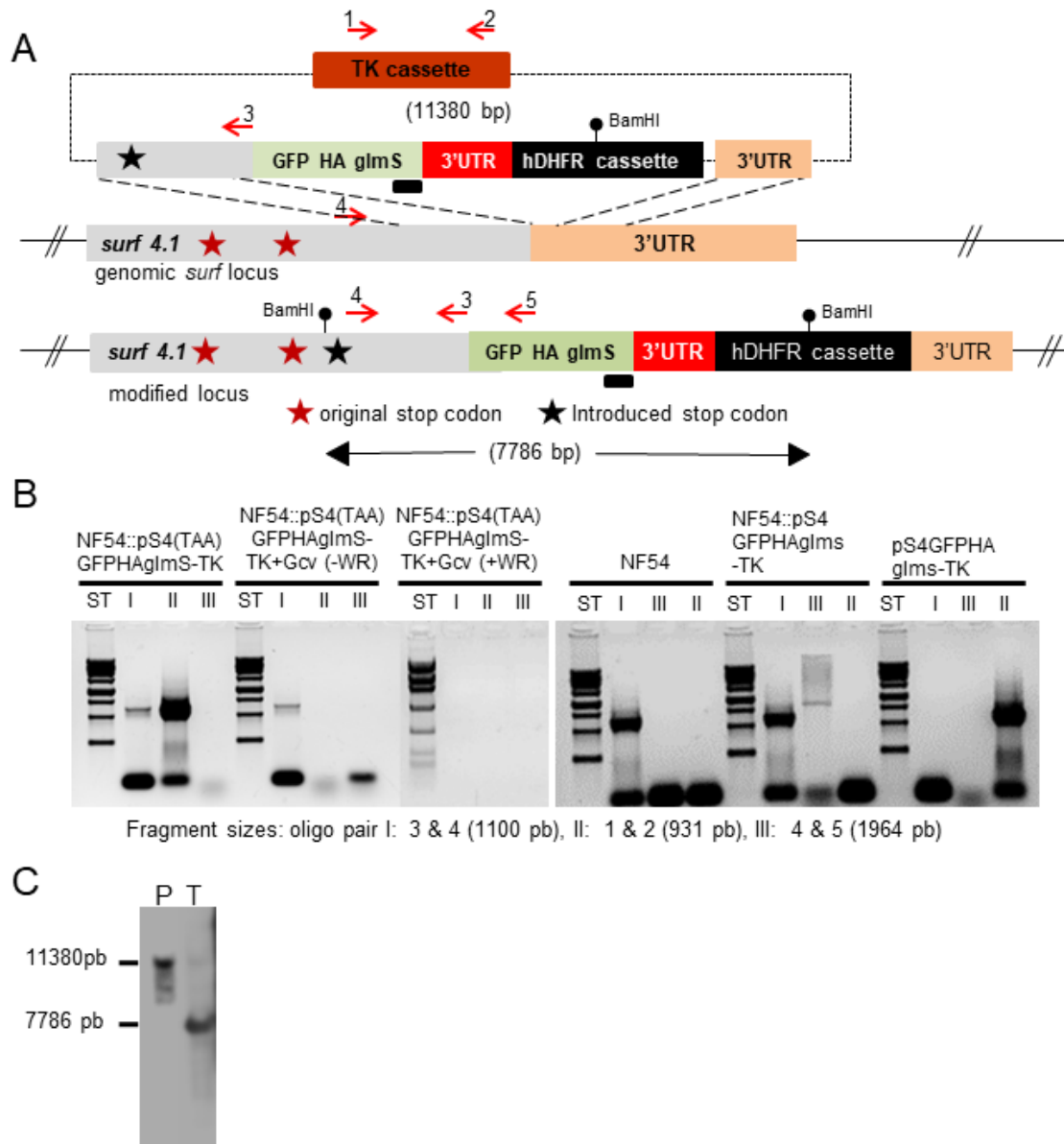
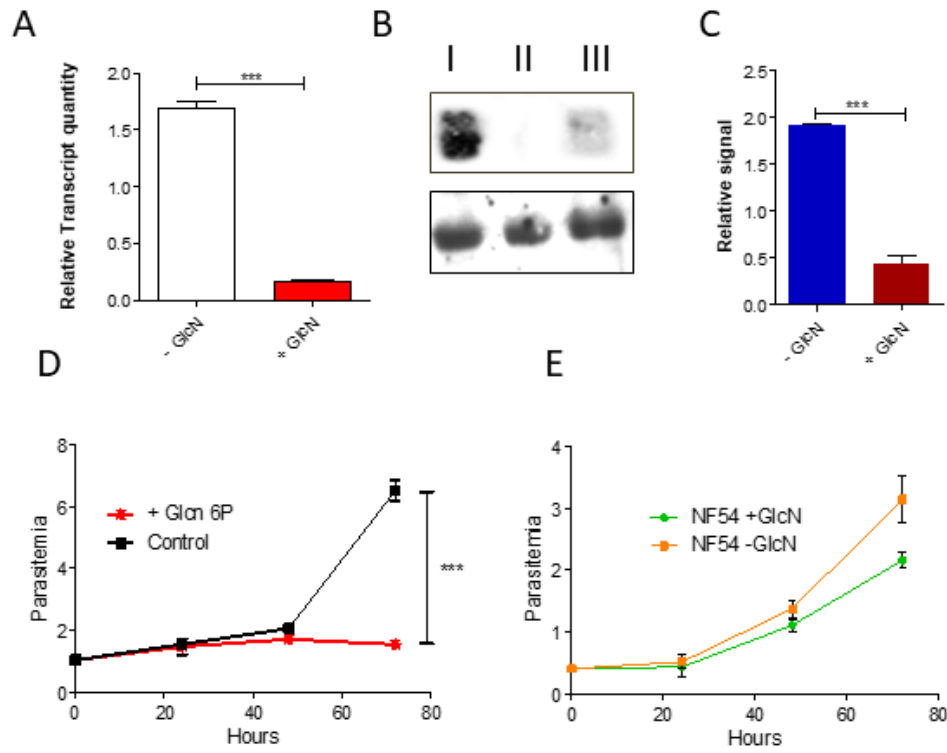


Figure 3



682

683

Figure 4

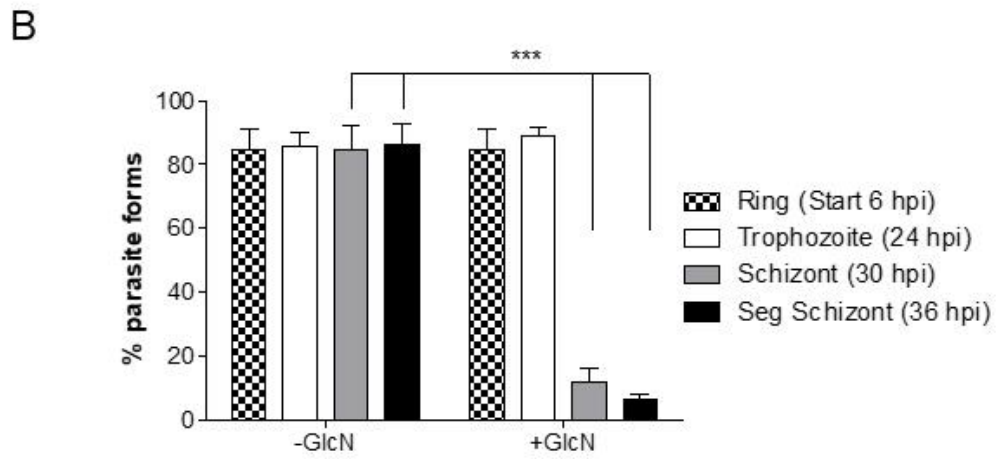
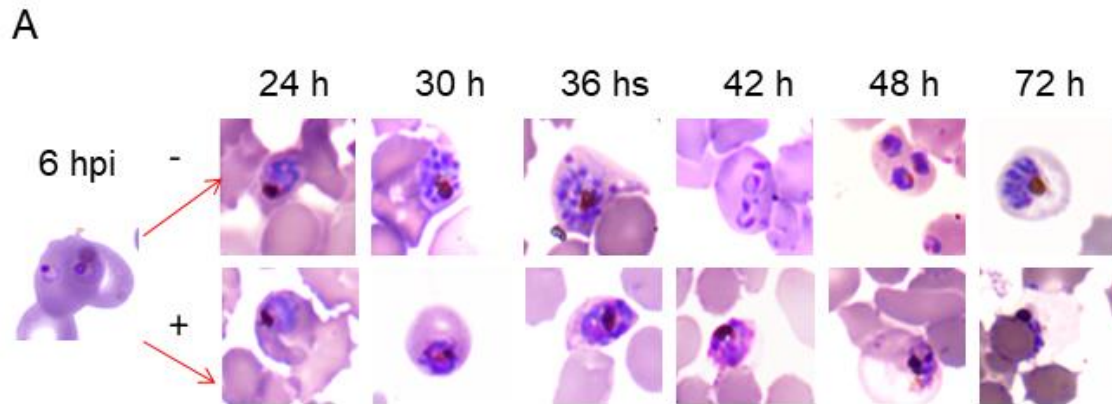


Figure 5

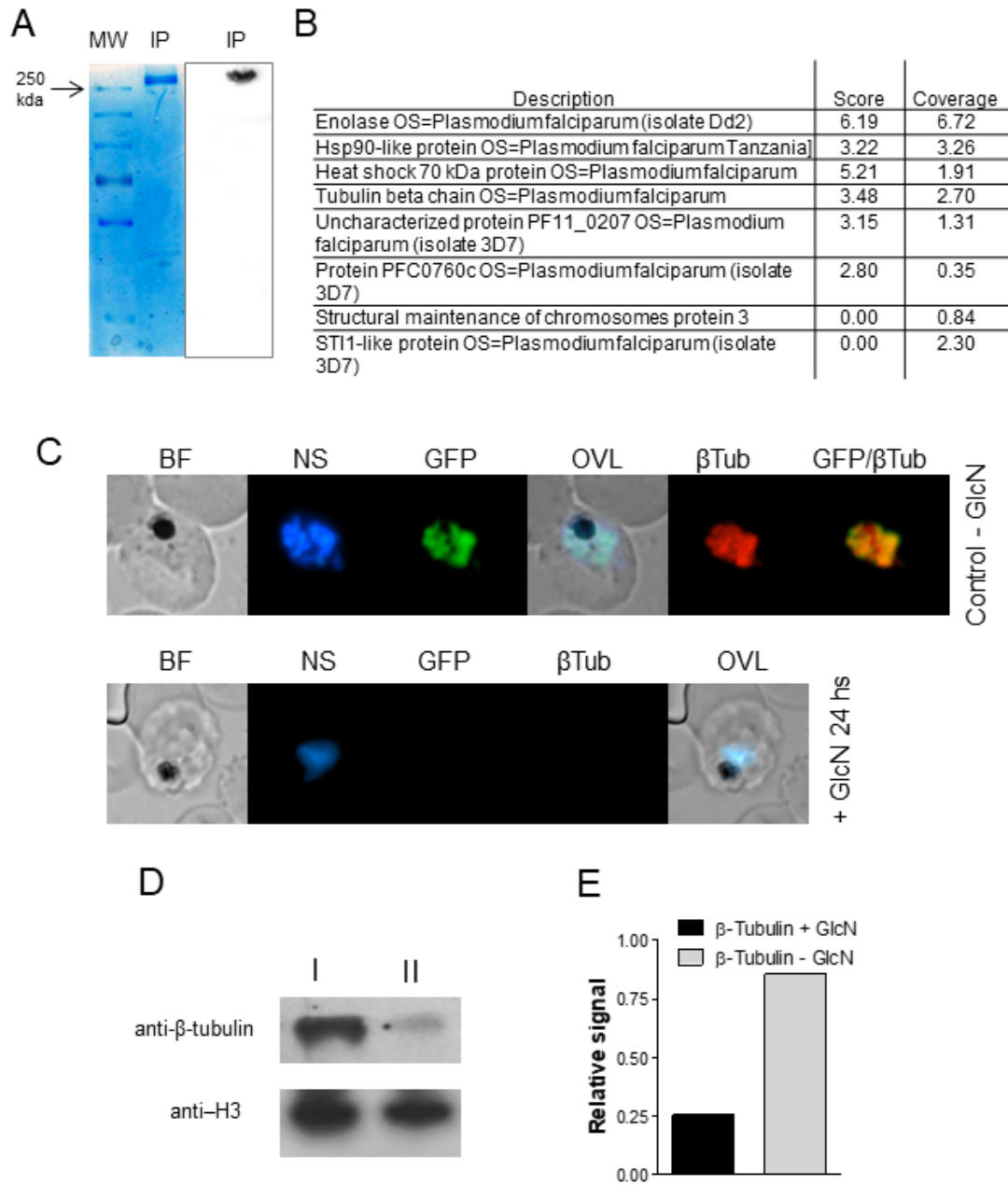


Figure 6

