A divergent cyclin/cyclin-dependent kinase complex controls the atypical replication of a malaria parasite during gametogony and transmission

Aurélia C Balestra[1,*], Mohammad Zeeshan[2,*], Edward Rea[2,¦], Carla Pasquarello[1,¦], Lorenzo Brusini[1,2], Tobias Mourier[3], Amit Kumar Subudhi[3], Natacha Klages Jemelin[1], Patrizia Arboit[1], Rajan Pandey[2], Declan Brady[2], Sue Vaughan[4], Anthony A Holder[5], Arnab Pain[3], David J P Ferguson[4,6], Alexandre Hainard[1], Rita Tewari[2,+], Mathieu Brochet[1,+]

1- University of Geneva, Faculty of Medicine, Geneva, Switzerland

2- University of Nottingham, School of Life Sciences, Nottingham, United Kingdom

3- Biological and Environmental Sciences and Engineering (BESE) Division, King Abdullah University of Science and Technology, Kingdom of Saudi Arabia

4- Oxford Brookes University, Department Biological & Medical Sciences, United Kingdom

5- The Francis Crick Institute, Malaria Parasitology Laboratory, London, United Kingdom

6- University of Oxford, John Radcliffe Hospital, Nuffield Department of Clinical Laboratory Science, Oxford OX3 9DU, United Kingdom

* Contributed equally

¦ Contributed equally

+ For correspondence

Summary

Cell cycle transitions are generally triggered by variation in the activity of cyclin-dependent kinases (CDKs) bound to cyclins. Malaria-causing parasites have a life cycle with unique cell-division cycles, and a repertoire of divergent CDKs and cyclins of poorly understood function and interdependency. We show that *Plasmodium berghei* CDK-related kinase 5 (CRK5), is a critical regulator of atypical mitosis in the gametogony and is required for mosquito transmission. It phosphorylates canonical CDK motifs of components in the pre-replicative complex and is essential for DNA replication. During a replicative cycle, CRK5 stably interacts with a single *Plasmodium*-specific cyclin (SOC2), although we obtained no evidence of SOC2 cycling by transcription, translation or degradation. Our results provide evidence that during *Plasmodium* male gametogony, this divergent cyclin/CDK pair fills the functional space of other eukaryotic cell-cycle kinases controlling DNA replication.

37 **Introduction**

38 Progression through the cell cycle critically relies upon post-translational mechanisms including
39 changes in activity of cell cycle kinases and phosphatases, and ubiquitin-mediated degradation of
40 specific components once their function is complete. Central components of these networks are the
41 cyclin-dependent kinases (CDKs). Many different families of CDK exist[1]. In most model organisms
42 however, those primarily regulating DNA replication and mitosis belong to a monophyletic family
43 that includes human CDK1-CDK4/6, Cdc2 and Cdc28 in fission and budding yeasts, respectively, and
44 CDKA in plants[2,3]. In malaria parasites, seven CDKs have been described[4-6]. Of these, five are
45 divergent CDK-related kinases (CRK), i.e. they are clearly related to CDKs but have no clear
46 orthologues in the yeast or human kinomes. These include CRK5 (PBANKA_123020), which forms a
47 distinct branch within the CDK cluster with atypical sequence motifs in the activation loop and an
48 additional 148 amino acid C-terminal extension found only in *Plasmodium*[7]. The gene is required, but
49 not essential, for proliferation of asexual blood stages of the human parasite *P. falciparum*[7].

50 The primary regulator of CDK activity is the cyclin subunit. Cyclins were originally named
51 because of their oscillation in level that reach a threshold required to drive cell-cycle transitions[8,9].
52 Cyclins are now defined as a family of evolutionarily related proteins encoding a cyclin box motif that
53 is required for binding to the CDK catalytic subunit[3]. *Plasmodium* genomes contain no sequence-
54 identifiable G_1-, S-, or M-phase cyclins, and only three proteins have sequence homology with cyclin
55 family members in other eukaryotes[10,11]. Cyc1 is important for cytokinesis in *P. falciparum* blood
56 stage replication, possibly regulating the CDK7 homolog MRK[12]. *P. berghei* Cyc3 is dispensable for
57 blood-stage replication but important for oocyst maturation in the mosquito midgut[11]. The paucity
58 of putative cyclins and the diversity of CDKs, has led to suggestions that some of these kinases
59 function without a cyclin partner[13].

60 The malaria parasite has several proliferative phases in its life cycle. Male gametogony or
61 gametogenesis is a proliferative sexual stage in the mosquito vector that is essential for parasite

62 transmission. Circulating mature male gametocytes in the vertebrate host are arrested in a G_0-like
63 phase and resume development in the mosquito midgut following a blood meal, activated by the
64 presence of xanthurenic acid (XA) and a drop in temperature[14]. In about ten minutes the haploid
65 male gametocyte completes three rounds of genome replication and closed endomitosis, assembles
66 the component parts of eight axonemes, and following nuclear division, produces eight flagellated
67 motile male gametes in a process called exflagellation[15]. The organisation and regulation of the cell
68 cycle during male gametogony is unclear. Current evidence suggests that certain canonical cell-cycle
69 checkpoints are absent[16]. For example, compounds that interfere with mitotic spindle formation do
70 not prevent DNA replication from proceeding[17,18], while spindle formation is not affected in a mutant
71 that is unable to replicate DNA[18,19]. Recently, we observed that proteins involved in DNA replication
72 and cytoskeletal reorganisation are similarly phosphorylated during the first seconds of
73 gametogony[20]. Interestingly, CRK5 was linked to both groups of proteins, suggesting it has a key
74 regulatory role during male gametogony[20]. Here, we provide evidence suggesting CRK5 is part of a
75 unique and divergent CDK/cyclin complex required for progression through male gametogony and
76 essential for parasite transmission.

77

78 **Results**

79 **CRK5 is a key regulator of gametogony and sporogony in the mosquito.** Previous attempts to
80 disrupt *P. berghei crk5* had suggested the gene is essential for asexual blood-stage proliferation[5].
81 However, using long sequence homology regions to replace *crk5* with a *T. gondii* DHFR/TS resistance
82 marker (Figure 1A and Figure 1 - figure supplement 1A), we obtained resistant parasites and cloned
83 them following a two-step enrichment. The gene deletion in the resulting CRK5-knockout (KO) clone
84 was confirmed by PCR (Figure 1 - figure supplement 1A) and RNAseq analysis (Figure 1A and
85 Supplementary file 1). There was no significant growth defect during erythrocytic asexual
86 multiplication (Figure 1B), nor an inability to produce morphologically normal gametocytes

87 (Figure 1C). However, upon XA activation only a few microgametocytes formed active exflagellation
88 centres (Figure 1D). While no major transcriptional changes were detected (Figure 1E), a significant
89 (p-value < 10^{-2}) but limited increase in expression ($0 < \log_2$[Fold Change] < 1) of multiple regulators of
90 gametogony (including *cdpk4*, *map2k*, *nek1* and *ark2*) was observed in CRK5-KO gametocytes
91 (Supplementary file 1). This could indicate that CRK5 plays a role during gametocytogenesis prior to
92 activation of gametocytes or that deletion of its coding sequence may be compensated by
93 upregulation of other kinases. Quantitative PCR analysis confirmed these observations but in most
94 cases, changes were not statistically significant (Figure 1F). To test the possibility that the
95 exflagellation defect was due to multiple subtle transcriptional changes prior to activation, we used
96 an additional approach to study the effect of rapid CRK5 degradation in mature gametocytes. We
97 tagged the endogenous *crk5* with an AID/HA epitope tag (Figure 1 - figure supplement 1B) to
98 degrade the fusion protein in presence of auxin in a strain expressing the Tir1 protein[21] (Figure 1G).
99 Addition of the AID/HA tag to the CRK5 C-terminus imposed a significant fitness cost, with a 2-fold
100 decrease in exflagellation in the absence of auxin, but importantly, depletion of CRK5-AID/HA by one
101 hour of auxin treatment of mature gametocytes prior to activation resulted in a dramatic reduction
102 in exflagellation (Figure 1H).

103 Since a residual number of CRK5-KO male gametocytes produced active exflagellation
104 centres, we asked whether the resulting microgametes were fertile by measuring the number of
105 ookinetes per activated female gametocyte. Consistent with the dramatic reduction in male
106 gametogony, CRK5-KO parasites showed a significant reduction in ookinete conversion compared to
107 the parental control line that expresses GFP under the strong *eef1α* promoter[22] (Figure 1I). This
108 reduction led to a > 30-fold decrease in the number of CRK5-KO oocysts present on the midgut
109 epithelium of *Anopheles stephensi* mosquitoes allowed to feed on infected mice (Figure 1J). To
110 further test for a possible role of CRK5 during sporogony, we endogenously tagged CRK5 with GFP at
111 its C-terminus (Figure 1 - figure supplement 1C). Consistently, CRK5-GFP is expressed in early oocysts
112 (Figure 1 - figure supplement 1D), which corroborates that residual CRK5-KO ookinetes produced

113 non-viable oocysts (Figure 1K), suggesting an additional role during sporogony and important for
114 parasite transmission (Figure 1L). Altogether, these results demonstrate a critical requirement for
115 CRK5 during male gametogony and an essential function for CRK5 in parasite development and
116 transmission.

117

118 **Phosphoproteome kinetics point to direct phosphorylation of the pre-replicative complex by CRK5.**
119 To elucidate the processes regulated by CRK5 during gametogony, we compared the proteomes and
120 phosphoproteomes of WT and CRK5-KO gametocytes using TMT labelling. We focus our analysis on
121 the first minute of gametogony that covers the first S-phase and metaphase of mitosis I (Figure 2A).
122 Apart from a significant CRK5 reduction in the CRK5-KO line, no significant differences among the
123 2598 quantified proteins were observed between mutant and parental lines (Supplementary file 2).
124 Similarly, non-activated CRK5-KO gametocytes showed limited differences in phosphorylation
125 compared to the parental counterpart. However, out of the 5901 quantified peptides, 20
126 phosphopeptides showed significant down-regulation upon activation (Benjamini-Hochberg adjusted
127 *p*-values <0.05 - Figure 2B) with most enriched GO terms corresponding to proteins important in
128 DNA replication including the origin recognition complex (ORC - Figure 2C and Supplementary file 3).
129 This group includes a protein with similarities to the licensing factor CDT1 (chromatin licensing and
130 DNA replication factor 1, PBANKA_1356300), a possible orthologue of the DNA replication factor
131 CDC6 (PBANKA_1102900), and two ORC components (ORC2 - PBANKA_0803000 - and ORC4 –
132 PBANKA_1348800). Interestingly, eight down-regulated phosphorylation sites map to the S/T*PxK
133 CDK consensus motif (Figure 2D). This motif shows a significant > 600 fold enrichment compared to
134 all other phosphorylated sequences identified in this study (*p*-value<1.10^{-14}).
135 Multiple phosphopeptides were more abundant following activation (Figure 2B) suggesting
136 rapid secondary effects on CRK5 phospho-dependent pathways, probably via the regulation of
137 protein phosphatases, as well as the inactivation or non-activation of other kinases. For example, the
138 phosphatase PPM11 in the mutant line showed a significantly different phosphorylation profile

139 before and after activation, although the functional relevance of PPM11 phosphorylation remains to
140 be determined. Components of the ribosome were the most significantly enriched (Figure 2C),
141 suggesting CRK5 can contribute indirectly to the activation of translation that is de-repressed upon
142 gametocyte activation[23]. Interestingly, four phosphopeptides from dynein subunits were also more
143 abundant, suggesting that the microtubule-based movement required during gametogony may also
144 be dependent on CRK5.

145 The set of enriched GO terms associated with CRK5-regulation is similar to those associated
146 with calcium dependent protein kinase (CDPK) 4 and CRK4 (Figure 2 - figure supplement 1), two
147 kinases previously shown to be crucial for DNA replication and its initiation during gametogony[19] and
148 schizogony[24], respectively. When compared with a set of human kinase-associated GO terms, the
149 CRK5 set shows similarities to those of CDK1 and CDK2 as well as other cell-cycle kinases including
150 Chk1, ATM and Aurora B. Altogether this phosphoproteomic shows that components of the pre-
151 replicative complex are phosphorylated on a canonical CDK motif in a CRK5-dependent manner
152 suggesting that CRK5 functions as a CDK. CRK5 may also indirectly regulate phosphorylation of
153 proteins involved in translation and microtubule-based movement.

154

155 **CRK5 is required for both S- and M- phases during *P. berghei* gametogony.** Given CRK5 similarity to
156 canonical CDKs and its possible role in initiating DNA replication, we quantified the DNA content by
157 flow cytometry of microgametocytes lacking *crk5* or upon CRK5-AID/HA degradation mediated by
158 auxin. In both cases, we observed an average 2-fold fewer polyploid (≥2N) gametocytes 1 min post-
159 activation (pa) with XA (Figure 2E), demonstrating a dependency on CRK5 for DNA replication during
160 male gametogony.

161 It has been suggested that M-phase is initiated independently of S-phase completion during
162 *Plasmodium* gametogony[19,20]. As multiple dynein-related proteins are phosphorylated in a CRK5-
163 dependent manner, we quantified the formation of spindle-like structures and axonemes in CRK5-

164 KO parasites. Cells depleted of CRK5 are unable to correctly assemble or maintain the spindle, with a
165 greater than two-fold decrease of visible spindles at 1 min pa (Figure 2F) in absence of CRK5. It is
166 important to note that this decrease is also observed in absence of Auxin in the CRK5-AID/HA line
167 suggesting that the defect in exflagellation observed for this line is also associated with a slower
168 gametogenesis. Tubulin staining at 15 min pa showed that axoneme formation is not affected but
169 axonemal beating is not initiated (Figure 2G). Ultrastructural analysis by electron microscopy
170 confirms the typical (9+2) axoneme structure; however at 6 min pa there appears to be more
171 nuclear poles present in WT compared to mutant parasites (0.56/cell compared to 0.32/cell; random
172 sample of individual sections from 50 cells of each population) consistent with reduced or slower
173 mitotic progression in the mutant. The cytoplasm contained electron dense basal bodies, which
174 were often, but not always, associated with nuclear poles (Figure 2Hb and g). Deletion of *crk5*
175 appeared to prevent activated microgametocytes developing to form multiple nuclear poles and
176 they showed no chromatin condensation, cytokinesis or flagellum formation (Figure 2H).

177 These results indicate that CRK5 is a critical primary regulator of DNA replication initiation
178 and progression and, possibly, an indirect regulator of microtubule-based processes required for
179 spindle and nuclear pole formation in the nucleus, but not for axoneme assembly in the cytosol.

180

181 **CRK5 is part of an atypical nuclear cyclin/CDK complex.** The binding of cyclins to CDKs is required
182 for progression through cell-cycle phases in model organisms. In the absence of clear S- and M-
183 phase cyclin homologues, we searched for possible CRK5 regulatory subunits. To identify any
184 proteins bound to CRK5, we used the CRK5-GFP line and additionally tagged the protein with 3xHA
185 at its C-terminus by modification of the endogenous gene (Figure 3 - figure supplement 1A). The
186 tagged protein was localised to the nucleus in non-activated gametocytes (Figure 3A) and at 1 min
187 pa showed an additional location at the mitotic spindle as suggested by co-localisation with tubulin
188 (Figure 3 - figure supplement 1D). Mass-spectrometry analyses of affinity-purified tagged CRK5-3xHA

189 (0, 15, 30 and 60 sec pa) and CRK5-GFP (at 6 and 15 min pa) identified 234 proteins (Supplementary
190 file 4). Consistent with a role for CRK5 in the initiation of DNA replication, GO term enrichment
191 analyses of proteins co-purifying with CRK5 highlighted ORC, minichromosome maintenance (MCM),
192 RepC and alpha DNA polymerase primase complexes (Figure 3B and C).

193 One of the most abundant proteins co-purified with CRK5 was SOC2 (PBANKA_144220), a
194 protein we had previously found to be important for cell-cycle progression during gametogony[20].
195 SOC2 has very low homology (<13 %) with a cyclin domain in a protein of the starfish *Asterina*
196 *pectinifera*[10], but the protein lacks key residues across most of the cyclin box[10] and has no detectable
197 cyclin-like function in *in vitro* biochemical assays[11]. Analysis of CRK5 immunoprecipitates also
198 identified multiple peptides from PBANKA_082440 (CDK regulatory subunit, CDKrs), a protein related
199 to Cks1 and CksHs2, two CDK-associated proteins in *Saccharomyces cerevisiae*[25] and human[26],
200 respectively. To confirm the putative interaction between CRK5, SOC2 and CDKrs, we tagged both
201 SOC2 and CDKrs at the C-terminus by integration of 3xHA[19] or GFP coding sequence into the
202 endogenous gene (Figure 3 - figure supplement 1B) and affinity purified SOC2-3xHA (0, 15, 30 and 60
203 sec pa) or CDKrs-GFP (at 60 sec, 6 min and 30 min pa). Reciprocal detection of CRK5, SOC2 and CDKrs
204 components was observed (Supplementary file 4). In addition, each protein also co-purified the ORC,
205 MCM, RepC and alpha DNA polymerase primase complexes (Supplementary file 4).

206 To better characterise the specificity of CRK5, SOC2 and CDKrs co-immunoprecipitations, we
207 compared the relative abundance of proteins recovered in CRK5, SOC2 and CDKrs
208 immunoprecipitates at 1 min post-activation with those of four proteins involved in DNA replication
209 or gametogenesis. These included CDPK4 that is critical to initiate DNA replication[19,27] as well as
210 MCM5. We also used two substrates of CDPK4, SOC1 that is important for DNA replication and SOC3
211 that is required later in gametogenesis prior to axoneme activation[19]. Seven-dimension principal-
212 component analysis (PCA) of all detected proteins confirmed the clustering of CRK5 with SOC2 and

213 CDKrs consistent with the formation of a complex (Figure 3C). As previously proposed, CDPK4 was
214 associated with MCM proteins while ORC and RepC proteins clustered in between the two groups.

215 Previous bioinformatics analyses had identified only three *bona fide Plasmodium* cyclins[11].
216 Here, the identification of SOC2 as a putative CRK5 binding partner suggests it is also a cyclin. We
217 revisited the phylogenetic analyses of apicomplexan cyclins and CDKs, and found that CRK5 and
218 CDKrs are conserved in *Haemosporidia*, *Coccidia* and *Piroplasma*, whilst SOC2 is found only in
219 *Haemosporidia* (Figure 3D and Figure 3 - figure supplement 1C).

220

221 **CRK5, SOC2 and CDKrs have a similar location and complementary functions during gametogony.**
222 Given the biochemical evidence that CRK5, SOC2 and CDKrs form a complex, we wanted to compare
223 their subcellular location. Live fluorescence microscopy identified CDKrs-GFP in the nucleus of non-
224 activated microgametocytes and throughout male gametogony, with a spindle-like localisation
225 during mitoses, as found for CRK5 (Figure 3E). By indirect immunofluorescence of fixed cells, we
226 confirmed the pattern of SOC2-3xHA distribution (Figure 3F), as previously described[19].

227 SOC2-KO microgametocytes showed defects in the ploidy transitions during gametogony[19].
228 In agreement with the notion of complementary roles for CRK5, SOC2 and CDKrs, deletion of either
229 SOC2 and CDKrs encoding genes (Figure 3 - figure supplement 1E) also resulted in a severe reduction
230 in exflagellation (Figure 3G). The defect was already apparent at 1 min pa, with a >3- fold reduction
231 of cells with ≥2N DNA content (Figure 3H) and a 3- fold reduction in mitotic spindles (Figure 3I).
232 Ultrastructural analysis by electron microscopy of SOC2-KO and CDKrs-KO activated gametocytes
233 confirms an early arrest during male gametogony as observed in CRK5-KO parasites (Figure 3 - figure
234 supplement 1F). Altogether, these data strongly suggest that the CRK5/SOC2/CDKrs complex is a
235 functional CDK/cyclin complex that controls early stages of S-phase and possibly controls
236 progression through the following or concurrent M-phase during gametogony. We propose the
237 name 'CSC complex' for this CRK5/SOC2/CDKrs complex.

238

239 **SOC2 expression does not follow a temporal cyclin pattern during gametogony and the CSC**
240 **complex is stable during the first round of mitosis.** Cyclin turnover mediated by the ubiquitin-
241 dependent proteasome pathway causes oscillations in CDK activity and therefore we wished to
242 examine the stability of the CSC complex. Western blot and fluorescent microscopy analysis revealed
243 that CRK5, SOC2 and CDKrs are expressed in non-activated gametocytes (Figure 3A, 3E and 4A) and
244 therefore we examined whether the CSC complex is degraded during the course of gametogony.
245 Western blot analysis of CRK5-3xHA and SOC2-3xHA over the 10 minutes of gametogony revealed no
246 evidence for degradation of either protein (Figure 4A). To confirm that levels of CRK5 and CDKrs are
247 maintained during gametogony in single cells, we monitored by flow cytometry the fluorescence of
248 gametocytes expressing CRK5-GFP or CDKrs-GFP during gametogony. This analysis confirmed that
249 there was no change in the level of CRK5-GFP or CDKrs-GFP fluorescence following male gametocyte
250 activation and gametogony (Figure 4 - figure supplement 1A).

251 To examine whether the apparent stable levels of SOC2-3xHA and CRK5-3xHA result from a
252 balance of *de novo* translation and protein degradation, we pre-treated non-activated gametocytes
253 with the proteasome inhibitor MG132 for one hour. Although MG132 significantly reduced
254 exflagellation (Figure 4 - figure supplement 1B), it did not affect DNA replication (Figure 4 - figure
255 supplement 1C) nor lead to a significant accumulation of SOC2, CRK5, or CDKrs as assessed by
256 western blot (Figure 4A) and flow cytometry (Figure 4 - figure supplement 1A). Together, these
257 results suggest no oscillation in levels of the CSC complex resulting from differential temporal
258 translation and degradation during gametogony.

259 CDKs are usually characterised by low activity in the absence of a bound cyclin, therefore
260 we reasoned that dynamic assembly of the CSC complex might underlie CRK5 regulation.
261 Development of activated gametocytes is highly synchronous and we used this property to
262 investigate the interaction between CRK5 and SOC2 in the first minute after activation. We

263 immunoprecipitated SOC2-3xHA and CRK5-3xHA at 0, 15, 30 and 60 seconds pa and measured the
264 relative abundance of immunoprecipitated proteins in the CSC complex by label-free mass-
265 spectrometry (Supplementary file 4). We detected no difference in the levels of SOC2 or CRK5
266 protein/protein interactions over this time course (Figure 4B). However, we noticed that a CDT1-like
267 protein became less abundant in CRK5 and SOC2 immunoprecipitates following gametocyte
268 activation (Figure 4C). Pre-treatment of gametocytes with BKI-294, a CDPK4 inhibitor prevented this
269 reduction. This suggests that the protein is degraded or less accessible to the complex following
270 CDPK4 activation. These results suggest that SOC2-CRK5 binding is not dynamic during the first
271 round of DNA replication.

272

273 **CSC is dynamically phosphorylated during the first round of replication.** Reversible phosphorylation
274 of cyclin-CDK complexes regulates kinase activity and therefore we examined whether this
275 mechanism may regulate the activity of the CSC complex. We had previously identified 13
276 phosphorylated SOC2 residues[20], with S5088/5089 phosphorylated in a CDPK4-dependent manner
277 during early gametogony[19]. Here, S5088 was also phosphorylated in a CDPK4-dependent manner in
278 the immunoprecipitated SOC2 (Figure 4 - figure supplement 1D), but S5088/5089 substitution with
279 alanine (Figure 4 - figure supplement 1E) did not cause a significant defect in DNA replication or
280 exflagellation (Figure 4 - figure supplement 1F and 1G) suggesting that SOC2-S5088/5089
281 phosphorylation is either not required or insufficient to regulate CSC complex function.

282 CRK5 is also phosphorylated dynamically during gametogony[20]. We show here that
283 inhibition of CDPK4 with BKI-1294 affected the phosphorylation at positions S654 and S693 of
284 immunoprecipitated CRK5 (Figure 4D). Interestingly, all eight detected phosphorylation sites in CRK5
285 are located in the 148-residue C-terminal extension (Figure 4E), which is only found in *Plasmodium*
286 CRK5[7]. We considered that this extension may have an important role in the regulation of CRK5
287 activity, and to test this hypothesis we generated a parasite clonal line, $CRK5_{1-562}$, with the last 148

288 amino acids of CRK5 deleted, leaving an intact catalytic domain (Figure 4 - figure supplement 1H).
289 This CRK5 truncation produced the same phenotype as that of the CRK5-KO line, with a significant
290 reduction in exflagellation, DNA replication and mitotic spindles (Figure 4F and G). These results
291 suggest the *Plasmodium*-specific C-terminus extension of CRK5 is essential for its function and may
292 act as a regulatory platform to coordinate cell cycle progression during gametogony.

293

294 **Discussion**

295 In this study, we have identified an atypical cyclin-CDK pair that is essential for male gametogony in a
296 malaria parasite. The divergent CDK-related kinase, CRK5, interacts with a *Plasmodium*-restricted
297 cyclin, SOC2, and with another putative regulatory subunit to form a complex we have named CSC.
298 We provide evidence of CRK5-dependent phosphorylation of multiple components of the pre-
299 replicative complex at a conserved eukaryotic CDK motif. Altogether, this suggests a very early
300 requirement for CRK5 during DNA replication. Further studies are necessary to link phosphorylation
301 of these proteins with CRK5 activity and to characterise the functional relevance of their
302 phosphorylation on the activity of the pre-replicative complex. Given our poor understanding of the
303 mechanisms that regulate the rapid DNA synthesis and the three rounds of mitosis during
304 gametogony, it is unclear whether components of the CSC complex are required for a single
305 checkpoint at the onset of S-phase or for two independent checkpoints, necessary to enter S- and M-
306 phases, respectively. In any case, the abundance of phosphorylated dyneins increased upon CRK5
307 deletion suggesting that CRK5-dependent regulation of M-phase would likely be indirect, possibly
308 affecting microtubule-dependent processes such as formation or segregation of nuclear poles or
309 mitotic spindle organisation[28,29] (Figure 4H).

310 CRK5, SOC2 and CDKrs have similar transcriptional profiles during asexual replication in
311 erythrocytes with a peak in late trophozoites[30], suggesting the CSC complex may have conserved
312 control functions in cell cycle progression during schizogony. Consistent with this idea, CRK5 deletion

313 reduced nuclear division during erythrocytic schizogony in *P. falciparum*[7], while SOC2 deletion led to
314 slower proliferation of *P. berghei* in erythrocytes[19]. Given the versatile nature of cyclin-CDK complex
315 function in other organisms, it is possible that other CDKs or cyclins partially compensate for the loss
316 of CRK5 or SOC2 in asexual blood stages. For example, CRK4 activity may compensate for the lack of
317 CRK5 during schizogony, because it is also important for DNA replication during this developmental
318 stage[24].

319 Usually there are significant oscillations in the abundance of cyclins during the cell cycle.
320 However, SOC2 shows no evidence of oscillation or changes in the level of its interaction with CRK5.
321 Such changes may be undetectable in the experimental conditions and we cannot exclude more
322 subtle changes that may regulate the CSC complex. The observed absence of cycling may also reflect
323 the constraints imposed by the extremely rapid nature of gametogony. No marked Cyc1 oscillations
324 were detected during the slower *P. falciparum* blood stage schizogony. Furthermore, the Cyc1/MRK
325 complex was detected in mature segmented schizonts after Cyc1 is required to complete
326 cytokinesis[31]. Therefore, it is also possible that non-oscillating cell-cycle cyclins and stable cyclin/CDK
327 complexes are a conserved feature of the various cycles of division within *Plasmodium* parasites.
328 However, PbCyc3 showed fluctuating expression patterns during oocyst development[11]. It thus
329 remains unclear whether non-cycling cyclins and stable cyclin/CDK complexes represent conserved
330 features of the various cell cycles of *Plasmodium* parasites.

331 We provide evidence that the C-terminal extension of CRK5 is important in the control of its
332 function. Interestingly, SOC2 and the C-terminal extension of CRK5 are both Haemosporidia-specific.
333 For example, they are absent in the related Apicomplexa parasite *Toxoplasma gondii* in which the
334 homologue of CRK5 is also involved in DNA replication[13], highlighting a specific evolution of this
335 cyclin/CDK system in *Plasmodium*. It is tempting to speculate that the CRK5 C-terminal extension and
336 its phosphorylation are important in the regulation of CRK5 activity via conformational changes
337 within the complex, but this hypothesis remains untested as the exact role of SOC2 in CRK5

338 regulation. Very few studied CDKs contain additional domains with known function. For example,
339 phosphorylation of the Cdk16 N-terminal domain blocks binding to cyclin Y[32]. Multiple CRK5 and
340 SOC2 residues are phosphorylated in a CDPK4-dependent manner and it is possible that CDPK4
341 works as a *Plasmodium*-specific link to translate a calcium signal triggering gametogony to prime the
342 CSC complex (Figure 4H). In addition, CDPK4-independent phosphorylation of CRK5 is also
343 observed[20] and other kinases or phosphatases may control CSC activity during the rounds of DNA
344 replication and mitosis. However, given the absence of canonical checkpoint kinases[5,6,33], it is
345 possible that the successive mitotic cycles of gametogony are separated from each other by a clock
346 mechanism rather than by checkpoints that rely on signals from one event in the cell cycle to
347 regulate the next.

348 *Plasmodium* species lack clear CDK orthologues[3] that are usually highly conserved across
349 Eukarya and indispensable in yeasts and humans[34]. This study indicates that CRK5, despite its
350 divergence from canonical CDKs, possibly acts as a CDK by binding a divergent cyclin and
351 phosphorylating classical CDK motifs on proteins that are targeted by conventional CDKs and
352 involved in the initiation of DNA replication in human and yeast. However, it remains difficult to
353 identify a specific direct functional orthologue due to functional similarities with multiple CDKs and
354 even non-CDK cell-cycle kinases, and it is tempting to speculate that CRK5 and SOC2 cover the
355 functional space of multiple cell-cycle regulators in other organisms, to satisfy the specificities of the
356 *Plasmodium* cell cycle. Future functional analyses will have to confirm that CRK5 activity is regulated
357 in a SOC2-dependent manner and the direct phosphorylation of the pre-replicative complex by
358 CRK5.

359

360

361 **Methods**

Key Resources Table				
Reagent type (species) or resource	**Designation**	**Source or reference**	**Identifiers**	**Additional information**
strain, strain background (*Escherichia coli*)	BigEasy®-TSA	Lucigen	Cat# 60224-1	Electrocompetent cells
transfected constructs (*P. berghei*)	PlasmoGEM vectors	https://plasmogem.sanger.ac.uk/	Method section and supplementary figures	
cell line (*P. berghei*)	ANKA 2.34	Billker et al, 2004		
cell line (*P. berghei*)	507cl1	Janse et al, 2006		
cell line (*P. berghei*)	820cl1	Mair et al, 2010		
cell line (*P. berghei*)	615	Philip et al, 2015		
cell line (*P. berghei*)	507cl1 CRK5-KO	This study		Available from the Nottingham laboratory
cell line (*P. berghei*)	615 CRK5-AID/HA	This study		Available from the Geneva laboratory
cell line (*P. berghei*)	2.34 CRK5-GFP	This study		Available from the Nottingham laboratory
cell line (*P. berghei*)	2.34 CRK5-3xHA	This study		Available from the Geneva laboratory
cell line (*P. berghei*)	2.34 CDKrs-GFP	This study		Available from the Nottingham laboratory
cell line (*P. berghei*)	820cl1 CDKrs-KO	This study		Available from the Geneva laboratory

cell line (*P. berghei*)	820cl1 SOC2-KO	Fang et al, 2017		Available from the Geneva laboratory
cell line (*P. berghei*)	820cl1 SOC2-3xHA	Fang et al, 2017		Available from the Geneva laboratory
cell line (*P. berghei*)	820cl1 SOC2$^{S5088/89A}$ -3xHA	This study		Available from the Geneva laboratory
cell line (*P. berghei*)	820cl1 CRK5$_{1-562}$	This study		Available from the Geneva laboratory
antibody	Anti-HA High Affinity from rat IgG1 (3F10)	Sigma	Cat# 000001186742 3001	IF(1:1000), WB (1:1000), IP (1:1000)
antibody	Anti-c-Myc antibody produced in rabbit	Sigma	Cat# C3956	WB (1:1000)
antibody	Anti-α-Tubulin antibody from mouse (DM1A)	Sigma	Cat# T6199	IF(1:1000), WB (1:1000)
antibody	Goat anti-Rat IgG (H+L), Alexa Fluor 594	Thermofisher	Cat# A-11007	IF(1:1000)
antibody	Goat anti-Mouse IgG (H+L), Alexa Fluor 488	Thermofisher	Cat# A-11029	IF(1:1000)
antibody	Goat anti-Rabbit IgG (H+L), Alexa Fluor 488	Thermofisher	Cat# A-11034	IF(1:1000)
sequence-based reagent	PCR primers	Microsynth	Method section and	

			supplementary file 5	
chemical compound, drug	BKI-1294	Ojo et al, 2014		CDPK4 inhibitor (1 µM)

362

363 **Ethics statement.** The animal work performed in the United Kingdom passed an ethical review
364 process and was approved by the UK Home Office. Work was carried out under UK Home Office
365 Project Licenses (40/3344 and 30/3248) in accordance with the United Kingdom 'Animals (Scientific
366 Procedures) Act 1986'. All animal experiments performed in Switzerland were conducted with the
367 authorisation numbers GE/82/15 and GE/41/17, according to the guidelines and regulations issued
368 by the Swiss Federal Veterinary Office.

369

370 **Parasite maintenance and transfection.** *P. berghei* ANKA strain[35]-derived clones 2.34[27], 507cl1[36],
371 820cl1[37], and 615[6], together with derived transgenic lines, were grown and maintained in CD1
372 outbred mice. Six to ten week-old mice were obtained from Charles River laboratories, and females
373 were used for all experiments. Mice were specific pathogen free (including mycoplasma pulmonis)
374 and subjected to regular pathogen monitoring by sentinel screening. They were housed in
375 individually ventilated cages furnished with a cardboard mouse house and Nestlet, maintained at 21
376 ± 2 °C under a 12 hr light/dark cycle, and given commercially prepared autoclaved dry rodent diet
377 and water *ad libitum*. The parasitaemia of infected animals was determined by microscopy of
378 methanol-fixed and Giemsa-stained thin blood smears.

379 For gametocyte production, parasites were grown in mice that had been phenyl hydrazine-
380 treated three days before infection. One day after infection, sulfadiazine (20 mg/L) was added in the
381 drinking water to eliminate asexually replicating parasites. Microgametocyte exflagellation was
382 measured three or four days after infection by adding 4 µl of blood from a superficial tail vein to 70
383 µl exflagellation medium (RPMI 1640 containing 25 mM HEPES, 4 mM sodium bicarbonate, 5% fetal

384 calf serum (FCS), 100 µM xanthurenic acid, pH 7.4). To calculate the number of exflagellation centres
385 per 100 microgametocytes, the percentage of red blood cells (RBCs) infected with microgametocytes
386 was assessed on Giemsa-stained smears. For gametocyte purification, parasites were harvested in
387 suspended animation medium (SA; RPMI 1640 containing 25 mM HEPES, 5% FCS, 4 mM sodium
388 bicarbonate, pH 7.20) and separated from uninfected erythrocytes on a Histodenz/Nycodenz
389 cushion made from 48% of a Histodenz/Nycodenz stock (27.6% [w/v] Histodenz/Nycodenz
390 [Sigma/Alere Technologies] in 5.0 mM TrisHCl, 3.0 mM KCl, 0.3 mM EDTA, pH 7.20) and 52% SA, final
391 pH 7.2. Gametocytes were harvested from the interface. To induce CRK5-AID/HA degradation, 1 mM
392 auxin dissolved in ethanol (0.2% final concentration) was added to purified gametocytes for one
393 hour prior to activation by XA.

394 Schizonts for transfection were purified from overnight *in vitro* culture on a Histodenz
395 cushion made from 55% of the Histodenz/Nycodenz stock and 45% PBS. Parasites were harvested
396 from the interface and collected by centrifugation at 500 *g* for 3 min, resuspended in 25 µL Amaxa
397 Basic Parasite Nucleofector solution (Lonza) and added to 10-20 µg DNA dissolved in 10 µl H_2O. Cells
398 were electroporated using the FI-115 program of the Amaxa Nucleofector 4D. Transfected parasites
399 were resupended in 200 µl fresh RBCs and injected intraperitoneally into mice. Parasite selection
400 with 0.07 mg/mL pyrimethamine (Sigma) in the drinking water (pH ~4.5) was initiated one day after
401 infection. Each mutant parasite was genotyped by PCR using three combinations of primers, specific
402 for either the WT or the modified locus on both sides of the targeted region (experimental designs
403 are shown in Supplemental Figures). For allelic replacements, sequences were confirmed by Sanger
404 sequencing using the indicated primers. Controls using wild type DNA were included in each
405 genotyping experiment; parasite lines were cloned when indicated.

406

407 **Generation of DNA targeting constructs.** The oligonucleotides used to generate and genotype the
408 mutant parasite lines are in Supplementary file 5.

409 Restriction/ligation cloning. The C-termini of CRK5 and CDKrs were tagged with GFP by single
410 crossover homologous recombination in the parasite. To generate these GFP lines, a region of the
411 *crk5* or *cdkrs* gene downstream of the ATG start codon was amplified using primers T1851/T1852
412 and T2461/T2462, respectively, ligated into p277 vector, and transfected in ANKA line 2.34 as
413 described previously[38]. Diagnostic PCRs were performed with primer 1 (IntT185 or intT246) and
414 primer 2 (ol492) to confirm integration of the GFP targeting constructs. Schematic representations of
415 the endogenous *crk5* and *cdkrs* loci, the constructs and the recombined loci are found in Figure 1 -
416 figure supplement 1 and S3 respectively.

417 The gene-deletion targeting vector for *crk5* was constructed using the pBS-DHFR plasmid,
418 which contains polylinker sites flanking a *T. gondii* dhfr/ts expression cassette conferring resistance
419 to pyrimethamine, as described previously[5]. PCR primers N1001 and N1002 were used to generate a
420 fragment of *crk5* 5' upstream sequence from genomic DNA, which was inserted into ApaI and HindIII
421 restriction sites upstream of the dhfr/ts cassette of pBS-DHFR. A fragment generated with primers
422 N1003 and N1004 from the 3' flanking region of *crk5* was then inserted downstream of the dhfr/ts
423 cassette using EcoRI and XbaI restriction sites. The linear targeting sequence was released using
424 ApaI/XbaI and the construct was transfected into the ANKA line 507cl1 expressing GFP. Diagnostic
425 PCR was performed with primer 1 (IntN100) and primer 2 (ol248) to confirm integration of the
426 targeting construct, and primer 3 (N100 KO1) and primer 4 (N100 KO2) were used to confirm
427 deletion of the *crk5* gene. A schematic representation of endogenous *crk5*, the constructs and the
428 recombined locus is in Figure 1 - figure supplement 1.

429

430 PlasmoGEM vectors. 3xHA and AID/HA tagging of CRK5, CDKrs-KO, $CRK5_{1-562}$ and SOC2 allelic
431 replacement construct were generated using phage recombineering in *Escherichia coli* TSA strain
432 with PlasmoGEM vectors (http://plasmogem.sanger.ac.uk/). For final targeting vectors not available
433 in the PlasmoGEM repository, generation of knock-out and tagging constructs was performed using
434 sequential recombineering and gateway steps as previously described[39,40]. For each gene of interest

435 (goi), the Zeocin-resistance/Phe-sensitivity cassette was introduced using oligonucleotides *goi* HA-F x
436 *goi* HA-R and *goi* KO-F x *goi* KO-R for 3xHA tagging and KO targeting vectors, respectively. Insertion
437 of the GW cassette following gateway reaction was confirmed using primer pairs GW1 x *goi* QCR1
438 and GW2 x *goi* QCR2. The modified library inserts were then released from the plasmid backbone
439 using NotI. The CRK5-AID/HA targeting vector was transfected into the 615 parasite line[21], the CDKrs-
440 KO, and $CRK5_{1-562}$ vectors into the 820cl1 line and the CRK5-3xHA vector into the 2.34 line.
441 Substitutions of $SOC2^{S5088/5089A}$ residues were introduced with a two-step strategy using λ
442 Red-ET recombineering as described in reference 41. The first step involved the insertion by
443 homologous recombination of a Zeocin-resistance/Phe-sensitivity cassette flanked by 5' and 3'
444 sequences of the codon of interest, which is amplified using the *soc2*-delF x *soc2*-delR primer pair.
445 Recombinant bacteria were then selected on Zeocin. The recombination event was confirmed by
446 PCR and a second round of recombination replaced the Zeocin-resistance/Phe-sensitivity cassette
447 with a PCR product containing the S5088/5089A substitutions amplified using *soc2*-mut-F with *soc2*-
448 mutR primer pairs, respectively. Bacteria were selected on YEG-Cl kanamycin plates. Mutations were
449 confirmed by sequencing vectors isolated from colonies sensitive to Zeocin with primers *soc2*-seqF
450 to *soc2*-seqR. The modified library insert was then released from the plasmid backbone using NotI
451 and the construct transfected into the 801cl1 line.
452
453 **Parasite phenotype analyses.** Blood containing approximately 50,000 parasites of the CRK5-KO line
454 was injected intraperitoneally into mice to initiate infections. Four to five days post infection,
455 exflagellation and ookinete conversion were examined as described previously[38] with a Zeiss
456 AxioImager M2 microscope (Carl Zeiss, Inc) fitted with an AxioCam ICc1 digital camera. When
457 indicated, gametocytes were pre-treated 5 min in SA supplemented with 1 µM BKI-1294[42] or 60 min
458 in SA supplemented with 1 µM MG132. To analyse mosquito transmission, 30 to 50 *Anopheles*
459 *stephensi* SD 500 mosquitoes were allowed to feed for 20 min on anaesthetized, infected mice
460 whose asexual parasitaemia had reached 15% and were carrying comparable numbers of

461 gametocytes as determined on Giemsa-stained blood films. To assess mid-gut infection,
462 approximately 15 guts were dissected from mosquitoes on day 7, 14, and 21 post-feeding and
463 oocysts were counted using the Zeiss AxioImager M2 microscope and a 63x oil immersion objective.
464 Mosquito bite-backs were performed 21 days post-feeding using naive mice and blood smears were
465 examined after 3-4 days.
466
467 **Immunofluorescence labelling.** Gametocyte immunofluorescence assays were performed as
468 previously described[43]. For HA and α-tubulin staining, purified cells were fixed with 4%
469 paraformaldehyde and 0.05% glutaraldehyde in PBS for 1 h, permeabilised with 0.1% Triton X-
470 100/PBS for 10 min and blocked with 2% BSA/PBS for 2 h. Primary antibodies were diluted in
471 blocking solution (rat anti-HA clone 3F10, 1:1000; mouse anti-α-tubulin clone DM1A, 1:1000, both
472 from Sigma-Aldrich). Anti-rat Alexa594, anti-mouse Alexa488, anti-rabbit Alexa 488, Anti-rabbit
473 Alexa594 were used as secondary antibodies together with DAPI (all from Life technologies), all
474 diluted 1:1000 in blocking solution. Confocal images were acquired with a LSM700 or a LSM800
475 scanning confocal microscope (Zeiss).
476 The CRK5-KO gametocytes were purified and activated in ookinete medium, then fixed at
477 15 min pa with 4% paraformaldehyde (PFA, Sigma) diluted in microtubule stabilising buffer (MTSB)
478 for 10-15 min and added to poly-L-lysine coated slides. Immunocytochemistry was performed using
479 primary mouse anti-α-tubulin mAb (Sigma-T9026; 1:1000) and secondary antibody Alexa 488
480 conjugated anti-mouse IgG (Invitrogen-A11004; 1:1000). The slides were then mounted in
481 Vectashield 19 with DAPI (Vector Labs) for fluorescence microscopy. Parasites were visualised on a
482 Zeiss AxioImager M2 microscope fitted with an AxioCam ICc1 digital camera (Carl Zeiss, Inc).
483
484 **Flow cytometry analysis of gametocyte DNA content and CRK5/CDKrs-GFP fluorescence.** DNA
485 content of microgametocytes was determined by flow cytometry measurement of fluorescence
486 intensity of cells stained with Vybrant dye cycle violet (life Technologies). Gametocytes were purified

487 and resuspended in 100 µl of SA. Activation was induced by adding 100 µl of modified exflagellation
488 medium (RPMI 1640 containing 25 mM HEPES, 4 mM sodium bicarbonate, 5% FCS, 200 µM
489 xanthurenic acid, pH 7.8). To rapidly block gametogony, 800 µl of ice cold PBS was added and cells
490 were stained for 30 min at 4°C with Vybrant dye cycle violet and analysed using a Beckman Coulter
491 Gallios 4. Microgametocytes were selected on fluorescence by gating on GFP positive
492 microgametocytes when the 820cl1 parasite line or its derivatives were used. In this case, cell ploidy
493 was expressed as a percentage of male gametocytes only. When the 2.34 line or its derivatives were
494 analysed, gating was performed on both micro- and macro-gametocytes and cell ploidy was
495 expressed as a percentage of all gametocytes. For each sample, >5,000 cells were analysed.
496

497 **Electron microscopy.** Gametocytes samples at 1, 3, 6, 15 and 30 min pa were fixed in 4%
498 glutaraldehyde in 0.1 M phosphate buffer. Samples were post fixed in osmium tetroxide, treated *en*
499 *bloc* with uranyl acetate, dehydrated and embedded in Spurr's epoxy resin. Thin sections were
500 stained with uranyl acetate and lead citrate prior to examination in a JEOL1200EX electron
501 microscope (Jeol UK Ltd).

502

503 **Phylogenetic analyses.** To search for orthologs of *Plasmodium* CRK sequences in alveolates,
504 annotated protein sequences were downloaded for the following organisms: *P. falciparum* (pfal), *P.*
505 *reichenowi* (prei), *P. knowlesi* (pkno), *P. vivax* (pviv), *P. berghei* (pber), *P. chabaudi* (pcha), and *P.*
506 *yoelii* (all from PlasmoDB, [plasmodb.org]). *Theileria annulata* (tann), *T. parva* (tpar), and *Babesia*
507 *bovis* (bbov) (all from PiroplasmaDB, [piroplasmadb.org]). *Toxoplasma gondii* (tgon), *Neospora*
508 *caninum* (ncan), and *Eimeria tenella* (eten) (all from ToxoDB, [toxodb.org]). *Cryptosporidium hominis*
509 (chom), *C. parvum* (cpar), *Chromera velia* (cvel), and *Vitrella brassicaformis* (vbra) (all from
510 CryptoDB, [cryptodb.org]). *Perkinsus marinus* (pmar), *Ichthyophthirius multifiliis* (imul), and
511 *Tetrahymena thermophilia* (tthe) (all from EnsemblProtists, [protists.ensembl.org]). *Symbiodinium*

512 *minutum* (smin) (from OIST, [marinegenomics.oist.jp]).

513 Reciprocal best hits to *P. berghei* CRK sequences were determined from pair-wise protein
514 BLAST searches. For each CRK protein, reciprocal best BLAST matches were retrieved and aligned
515 using mafft[44]. Alignments were trimmed using trimal[45] and HMM models were produced with
516 HMMer (hmmer.org). These models were then used to search the genomes for which no reciprocal
517 best BLAST hit was detected. A neighbour-joining tree was then constructed using four combined
518 data sets: 1) the *P. berghei* CRK proteins, 2) all protein sequences with a reciprocal best BLAST hit to
519 *P. berghei*, 3) genes that were best BLAST hit for a *P. berghei* protein, but did not have the CRK
520 protein as best hit in *P. berghei*, 4) all protein sequences with a score below $1x10^{-12}$ to an HMM
521 model. Sequences and identifiers are provided in the source data files of Figure 3 - figure
522 supplement 1.

523 Sequence distances were calculated using clustalw2[46,47], which was also used to construct
524 the NJ tree. Protein sequences from 1) and 2) were divided into their CRK groups (*i.e.* CRK1-7, CDKrs,
525 and SOC2), and all pair-wise distances were calculated within each CRK group. For all protein
526 sequences from 3) and 4), the distance to all members in all CRK groups were then measured. A
527 protein sequence was included in a CRK group if the average distance between the protein sequence
528 and all members of a given CRK group did not exceed 40% of all pair-wise distances within that CRK
529 group, or if the CRK group to which a protein sequence had the lowest average distance was lower
530 than the distance to the CRK group with the second-lowest average distance minus 5 standard
531 deviations. Final trees were produced by PhyML[48,49] with the GTR substitution model selected by
532 SMS[50]. Branch support was evaluated with the Bayesian-like transformation of approximate
533 likelihood ratio test, aBayes[51]. Trees were visualized using FigTree
534 (tree.bio.ed.ac.uk/software/figtree/).

535

536 **Transcriptomics.** Gametocytes (non-activated -0 min- and activated -15 min-) were collected from
537 CRK5-KO and WT lines. Total RNA was isolated from purified parasites using an RNeasy purification
538 kit (Qiagen). RNA was vacuum concentrated (SpeedVac) and transported using RNA stable tubes
539 (Biomatrica). Strand-specific mRNA sequencing was performed from total RNA using TruSeq
540 Stranded mRNA Sample Prep Kit LT (Illumina) according to the manufacturer's instructions. Briefly,
541 polyA+ mRNA was purified from total RNA using oligo-dT dynabead selection. First strand cDNA was
542 synthesised primed with random oligos followed by second strand synthesis where dUTPs were
543 incorporated to achieve strand-specificity. The cDNA was adapter-ligated and the libraries amplified
544 by PCR. Libraries were sequenced in an Illumina Hiseq machine with paired-end 100bp read
545 chemistry.

546 RNA-seq read alignments were mapped onto the *P. berghei* ANKA genome (May 2015
547 release in GeneDB—http://www.genedb.org/) using Tophat2 (version 2.0.13) with parameters "-
548 library-type fr-firststrand–no-novel-juncs–r 60". Transcript abundances were extracted as raw read
549 counts using the Python script 'HTseq-count'[52] (model type – union, http://www-
550 huber.embl.de/users/anders/HTSeq/). Counts per million (cpm) values were obtained from count
551 data and gene were filtered if they failed to achieve a cpm value of 1 in at least 30% of samples per
552 condition. Library sizes were normalized by the TMM method using EdgeR software[53] and further
553 subjected to linear model analysis using the voom function in the limma package[54]. Differential
554 expression analysis was performed using DESeq2 in R version 3.2.1[55]. Genes with fold change greater
555 than 2 and p-value less than 0.05 were considered as significantly differentially expressed.
556 *P. berghei* GO terms (Gene Ontology) were downloaded from GeneDB (http://www.genedb.org/;
557 May 2015 release) and gene ontology enrichment analysis was performed for the DEG lists using
558 GOstats R package. All analyses and visualizations were done with R packages-
559 *cummeRbund* and *ggplot2.*
560

561 **Quantitative Real Time PCR (qRT-PCR) analyses**. RNA was isolated from purified gametocytes using
562 an RNA purification kit (Stratagene). cDNA was synthesised using an RNA-to-cDNA kit (Applied
563 Biosystems). Gene expression was quantified from 80 ng of total RNA using SYBR green fast master
564 mix kit (Applied Biosystems). All the primers were designed using primer3 (Primer-blast, NCBI).
565 Analysis was conducted using an Applied Biosystems 7500 fast machine with the following cycling
566 conditions: 95˚C for 20 s followed by 40 cycles of 95˚C for 3 s; 60˚C for 30 s. Three technical
567 replicates and two biological replicates were performed for each assayed gene. The *hsp70*
568 (PBANKA_081890) and *arginyl-t RNA synthetase* (PBANKA_143420) genes were used as endogenous
569 control reference genes. The primers used for qPCR can be found in Supplementary file 5.
570
571 **Protein immunoprecipitation and identification.**
572 Sample preparation. Co-immunoprecipitations (IPs) of proteins were performed with purified
573 gametocytes. The following IPs were performed: SOC2-3xHA (0, 15, 30 and 60 sec pa), CDKrs-GFP (60
574 sec, 6 min and 30 min pa), CRK5-3xHA (0, 15, 30 and 60 sec pa), and CRK5-GFP (at 6 and 15 min pa).
575 IP from wild type non-activated gametocytes lacking an epitope tag or constitutively expressing GFP
576 were used as controls.
577 For HA-based IPs, samples were fixed for 10 min with 1% formaldehyde. Parasites were lysed in RIPA
578 buffer (50 mM Tris HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) and the
579 supernatant was subjected to affinity purification with anti-HA antibody (Sigma) or anti-GFP
580 antibody (Invitrogen) conjugated to magnetics beads. Beads were re-suspended in 100 µl of 6 M
581 urea in 50 mM ammonium bicarbonate (AB). Two µl of 50 mM dithioerythritol (DTE) were added and
582 the reduction was carried out at 37°C for 1h. Alkylation was performed by adding 2 µl of 400 mM
583 iodoacetamide for 1 h at room temperature in the dark. Urea was reduced to 1 M by addition of 500
584 µl AB and overnight digestion was performed at 37 °C with 5 µl of freshly prepared 0.2 µg/µl trypsin
585 (Promega) in AB. Supernatants were collected and completely dried under speed-vacuum. Samples

586 were then desalted with a C18 microspin column (Harvard Apparatus) according to manufacturer's
587 instructions, completely dried under speed-vacuum and stored at -20°C.
588 Liquid chromatography electrospray ionisation tandem mass spectrometry (LC-ESI-MSMS). Samples
589 were diluted in 20 µl loading buffer (5% acetonitrile [CH_3CN], 0.1% formic acid [FA]) and 2 µl were
590 injected onto the column. LC-ESI-MS/MS was performed either on a Q-Exactive Plus Hybrid
591 Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific) equipped with an Easy nLC 1000
592 liquid chromatography system (Thermo Fisher Scientific) or an Orbitrap Fusion Lumos Tribrid mass
593 Spectrometer (Thermo Fisher Scientific) equipped with an Easy nLC 1200 liquid chromatography
594 system (Thermo Fisher Scientific). Peptides were trapped on an Acclaim pepmap100, 3 µm C18, 75
595 µm x 20 mm nano trap-column (Thermo Fisher Scientific) and separated on a 75 µm x 250 mm (Q-
596 Exactive) or 500 mm (Orbitrap Fusion Lumos), 2µm C18, 100 Å Easy-Spray column (Thermo Fisher
597 Scientific). The analytical separation used a gradient of H_2O/0.1% FA (solvent A) and CH_3CN/0.1 % FA
598 (solvent B). The gradient was run as follows: 0 to 5 min 95 % A and 5 % B, then to 65 % A and 35 % B
599 for 60 min, then to 10 % A and 90 % B for 10 min and finally for 15 min at 10 % A and 90 % B. Flow
600 rate was 250 nL/min for a total run time of 90 min.
601 Data-dependant analysis (DDA) was performed on the Q-Exactive Plus with MS1 full scan at a
602 resolution of 70,000 Full width at half maximum (FWHM) followed by MS2 scans on up to 15
603 selected precursors. MS1 was performed with an AGC target of 3 x 10^6, a maximum injection time of
604 100 ms and a scan range from 400 to 2000 m/z. MS2 was performed at a resolution of 17,500
605 FWHM with an automatic gain control (AGC) target at 1 x 10^5 and a maximum injection time of 50
606 ms. Isolation window was set at 1.6 m/z and 27% normalised collision energy was used for higher-
607 energy collisional dissociation (HCD). DDA was performed on the Orbitrap Fusion Lumos with MS1
608 full scan at a resolution of 120,000 FWHM followed by as many subsequent MS2 scans on selected
609 precursors as possible within a 3 sec maximum cycle time. MS1 was performed in the Orbitrap with
610 an AGC target of 4 x 10^5, a maximum injection time of 50 ms and a scan range from 400 to 2000 m/z.
611 MS2 was performed in the Ion Trap with a rapid scan rate, an AGC target of 1 x 10^4 and a maximum

612 injection time of 35 ms. Isolation window was set at 1.2 m/z and 30% normalised collision energy
613 was used for HCD.
614 Database searches. Peak lists (MGF file format) were generated from raw data using the MS Convert
615 conversion tool from ProteoWizard. The peak list files were searched against the
616 PlasmoDB_*P.berghei* ANKA database (PlasmoDB.org, release 38, 5076 entries) combined with an in-
617 house database of common contaminants using Mascot (Matrix Science, London, UK; version 2.5.1).
618 Trypsin was selected as the enzyme, with one potential missed cleavage. Precursor ion tolerance
619 was set to 10 ppm and fragment ion tolerance to 0.02 Da for Q-Exactive Plus data and to 0.6 for
620 Lumos data. Variable amino acid modifications were oxidized methionine and deamination (Asn and
621 Gln) as well as phosphorylated serine, threonine and tyrosine. Fixed amino acid modification was
622 carbamidomethyl cysteine. The Mascot search was validated using Scaffold 4.8.4 (Proteome
623 Software) with 1% of protein false discovery rate (FDR) and at least 2 unique peptides per protein
624 with a 0.1% peptide FDR.
625 PCA analysis. Enrichment and principal component analysis was performed in the statistical
626 programming package 'R' (www.r-project.org). Quantitative values were analysed as log-
627 transformed spectral count values and displayed in principal components with greatest degrees of
628 variance.
629 Label-free comparison of peptide and phosphopeptides in SOC2 or CRK5 immunoprecipitates. Label-
630 free comparison of peptide and phosphopeptides in SOC2 or CRK5 immunoprecipitates during the
631 first minute of gametogony was determined using Proteome Discoverer 2.2 (Thermo Fisher
632 Scientific). Peptide-spectrum matches were validated using Percolator validator node with a target
633 FDR of 0.01 and a Delta Cn of 0.5. For label-free quantification, features and chromatographic peaks
634 were detected using the "Minora Feature Detector" Node with the default parameters. PSM and
635 peptides were filtered with a false discovery rate (FDR) of 1%, and then grouped to proteins with
636 again a FDR of 1% and using only peptides with high confidence level. Both unique and razor
637 peptides were used for quantitation and protein abundances are calculated by summing samples

638 abundances of the connected peptides group. The abundances were normalised on the "Total
639 Peptide Amount" and then "The pairwise Ratio Based" option was selected for protein ratio
640 calculation and associated p-values were calculated with an ANOVA test based on background.
641
642 **Proteomics and phosphoproteomics.**
643 Sample preparation (SDS buffer-FASP procedure). Activated and non-activated purified gametocytes
644 were snap frozen in liquid nitrogen at 0, 15, 30, 45 and 60 sec pa. For each time point, two
645 independent biological replicates were analysed. Cell lysis was performed in 1 ml of 2% SDS, 25 mM
646 NaCl, 2.5 mM EDTA, 20 mM TCEP, 50 mM TrisHCl (pH 7.4), supplemented with 1x Halt™ protease
647 and phosphatase inhibitor (Thermo Fisher). Samples were vortexed and then heated at 95°C for 10
648 min with mixing at 400 rpm on a thermomixer. DNA was sheared *via* four sonication pulses of 10 s
649 each at 50% power. Samples were then centrifuged for 30 min at 17,000 *g* and the supernatant was
650 collected. Three-hundred µl sample was incubated with 48 µl 0.5 M iodoacetamide for 1h at room
651 temperature. Protein was digested based on the FASP method[56] using Amicon® Ultra-4, 30 kDa cut-
652 off as centrifugal filter units (Millipore). Trypsin (Promega) was added at 1:100 enzyme/protein ratio
653 and digestion was performed at 37°C overnight. The resulting peptides were treated as described
654 above.
655 TMT11plex-labelling.Peptide concentration was determined using a colorimetric peptide assay
656 (Pierce). A pool made from 1/20 of each sample was created and used as reference for each tandem
657 mass tag (TMT) experiment. Briefly, 100 µg of each sample was labelled with 400 µg of
658 corresponding TMT-11plex reagent previously dissolved in 110 µl 36% CH_3CN, 200 mM 4-(2-
659 hydroxyethyl)-1-piperazinepropanesulfonic acid (EPPS, pH 8.5). The reference sample was labelled
660 with TMT11-131C label reagent. The reaction was performed for 1h at room temperature and then
661 quenched by adding hydroxylamine to a final concentration of 0.3% (v/v). Labelled samples of each
662 TMT11 experiment were pooled together, dried and desalted with a peptide desalting spin column
663 (Pierce) according to manufacturer's instructions.

664 Phosphopeptide enrichment. Phosphopeptides were enriched using the High-Select Fe-NTA
665 Phosphopeptide Enrichment Kit (Thermo Fisher Scientific) following manufacturer's instructions. The
666 phosphopeptide fraction as well as the flow-through fraction were desalted with a C18 macrospin
667 column (Harvard Apparatus) according to manufacturer's instruction and then completely dried
668 under speed-vacuum.

669 High pH Reverse-Phase fractionation. The flow-through fraction of each TMT11 experiment was
670 fractionated into 13 fractions using the Pierce™ High pH Reversed-Phase Peptide Fractionation Kit
671 (Thermo Fisher Scientific) according to manufacturer's instructions.

672 LC-ESI-MSMS. For all samples, peptide concentration was determined using a colorimetric peptide
673 assay (Pierce). Phosphopeptides were reconstituted in loading buffer (5% CH_3CN, 0.1% FA) and 2 µg
674 injected on to the column. LC-ESI-MS/MS was performed on an Orbitrap Fusion Lumos Tribrid mass
675 spectrometer (Thermo Fisher Scientific) equipped with an Easy nLC1200 liquid chromatography
676 system (Thermo Fisher Scientific). Peptides were trapped on a Acclaim pepmap100, C18, 3 µm,
677 75 µm x 20 mm nano trap-column (Thermo Fisher Scientific) and separated on a 75 µm x 500 mm,
678 C18, 2 µm, 100 Å Easy-Spray column (Thermo Fisher Scientific). The analytical separation was run for
679 125 min using a gradient of 99.9% H_2O/ 0.1% FA (solvent A) and 80% CH_3CN/0.1% FA (solvent B). The
680 gradient was run as follows: 0 to 2 min 92 % A and 8 % B, then to 72 % A and 28 % B for 105 min, to
681 58% A and 42% B for 20 min, and finally to 5 % A and 95 % B for 10 min with a final 23 min at this
682 composition. Flow rate was of 250 nL/min. DDA was performed with MS1 full scan at a resolution of
683 120,000 FWHM followed by as many subsequent MS2 scans on selected precursors as possible
684 within a 3 sec maximum cycle time. MS1 was performed in the Orbitrap with an AGC target of $4x10^5$,
685 a maximum injection time of 50 ms and a scan range from 375 to 1500 *m/z*. MS2 was performed in
686 the Orbitrap using HCD at 38% Normalised collision Energy (NCE). Isolation window was set at 0.7 u
687 with an AGC target of $5x10^4$ and a maximum injection time of 86 ms. A dynamic exclusion of parent
688 ions of 60 s with 10 ppm mass tolerance was applied.

689 High pH Reversed-Phase Peptide fractions were reconstituted in the loading buffer (5% CH_3CN, 0.1%
690 FA) and 1 µg was injected onto the column. LC-ESI-MS/MS was performed as described above. The
691 analytical separation was run for 90 min using a gradient of 99.9% H_2O/0.1% FA (solvent A) and 80%
692 CH_3CN/0.1%FA (solvent B). The gradient was run as follows: 0 to 5 min 95 % A and 5 % B, then to 65
693 % A and 35 % B for 60 min, and finally to 5% A and 95% B for 10 min with a final 15 min at this
694 composition. Flow rate was 250 nL/min. DDA was performed using the same parameters, as
695 described above.
696 Database search. Raw data were processed using Proteome Discoverer 2.3 software (Thermo Fisher
697 Scientific). Briefly, spectra were extracted and searched against the *P. berghei* ANKA database
698 (PlasmoDB.org, release 38, 5076 entries) combined with an in-house database of common
699 contaminants using Mascot (Matrix Science, London, UK; version 2.5.1). Trypsin was selected as the
700 enzyme, with one potential missed cleavage. Precursor ion tolerance was set to 10 ppm and
701 fragment ion tolerance to 0.02 Da. Carbamidomethyl of cysteine (+57.021) as well as TMT10plex
702 (+229.163) on lysine residues and on peptide N-termini were specified as fixed modifications.
703 Oxidation of methionine (+15.995) as well as phosphorylated serine, threonine and tyrosine were set
704 as variable modifications. The search results were validated with a Target Decoy PSM validator. PSM
705 and peptides were filtered with a FDR of 1%, and then grouped to proteins, again with a FDR of 1%
706 and using only peptides with high confidence level. Both unique and razor peptides were used for
707 quantitation and protein and peptides abundances value were based on signal to noise (S/N) values
708 of reporter ions. Abundances were normalised on "Total Peptide Amount" and then scaled with "On
709 Controls Average" (i.e. using the reference sample channel). All the protein ratios were calculated
710 from the medians of the summed abundances of replicate groups and associated p-values were
711 calculated with an ANOVA test based on individual protein or peptides. GO term enrichment was
712 performed using PlasmoDB. The reference set of GO terms for human kinases was obtained from [24].
713

714 **Statistical analysis.** All statistical analyses were performed using GraphPad Prism 8 (GraphPad
715 Software) assuming normal distribution. Statistical tests are mentioned in the figure legends.
716

717 **Data availability.** Mass spectrometry proteomics data have been deposited to the
718 ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier
719 PXD017283 (proteomics and phosphoproteomics), PXD017308 (HA immunoprecipitations) and
720 PXD017622 (GFP immunoprecipitations). RNAseq data have been deposited to the Gene Expression
721 Omnibus under accession number GSE144743. The authors declare that all other relevant data
722 generated or analysed during this study are included in the Article or its Supplementary Information.
723 Raw data are available for each figure. Materials are available from the corresponding authors on
724 reasonable request.
725

726 **Figure legends**

727 **Figure 1. CRK5 is a key regulator of gametogony and sporogony during mosquito transmission. A.**
728 Schematic representation of CRK5 showing the eleven regions that constitute the catalytic domain in
729 pink, and the *Plasmodium*-specific C-terminal extension that is phosphorylated at eight sites (green
730 bars). Amino acids 31 to 685 were replaced by the selection marker in the CRK5-KO line. The lower
731 panel shows the sequence reads mapping on *crk5* in the KO and WT lines, confirming deletion of the
732 gene in the KO. **B.** Deletion of *crk5* has no effect on the growth of asexual blood stage parasites
733 (error bars show standard deviation from the mean; 3 and 2 independent infections for the CRK5-KO
734 and WT lines, respectively). **C.** Deletion of *crk5* has no effect on gametocytaemia (error bars show
735 standard deviation from the mean; 3 and 2 independent infections for the CRK5-KO and WT lines,
736 respectively). **D.** Deletion of *crk5* leads to a strong defect in exflagellation as measured by the
737 percentage of active exflagellation centres per microgametocyte (error bars show standard deviation
738 from the mean; technical replicates from 3 independent infections; two-way ANOVA). **E.** Volcano

739 plots show limited transcriptional changes between the CRK5-KO and wild type (WT) parasite lines,
740 comparing non-activated gametocytes (T=0 min) with gametocytes 15 min after activation (T=15
741 min). **F.** The levels of mRNA for multiple critical regulators of gametogony do not change significantly
742 in CRK5-KO compared to WT parasites, despite a trend for up-regulation in the mutant (error bars
743 show standard deviation from the mean; 2 independent biological replicates, two-tailed unpaired t-
744 test). **G.** Depletion of CRK5-AID/HA protein upon addition of auxin to mature purified gametocytes,
745 as measured by western blotting; α-tubulin serves as a loading control. **H.** Depletion of CRK5-AID/HA
746 upon auxin treatment leads to a profound defect in exflagellation (error bars show standard
747 deviation from the mean; technical replicates from 3 independent infections; two-way ANOVA). **I.**
748 Effect of *crk5* deletion on ookinete formation as a measure of fertile male gamete formation.
749 Conversion is reported as the percentage of female gametes forming ookinetes (error bars show
750 standard deviation from the mean; 3 independent infections, two-tailed unpaired t-test). **J.** *crk5*
751 deletion leads to a dramatic reduction in oocyst formation when mature gametocytes are fed to
752 mosquitoes. Midguts were dissected and GFP-positive oocysts were counted at 7 and 21 days post-
753 infection (error bars show standard deviation from the mean; 3 independent infections; two tailed
754 unpaired t-test). **K.** The residual CRK5-KO oocysts show abnormal development; scale bar = 20 µm. **L.**
755 No viable CRK5-KO sporozoite develop. Mosquitoes were infected and 21 days later allowed to feed
756 on naive mice. For the WT line, blood stage parasites were readily observed after a 4-day prepatent
757 period, while no mice became infected with CRK5-KO parasites (error bars show standard deviation
758 from the mean; 3 independent infections). Source files for Figure 1 are available in Figure 1 - source
759 data 1 to 3.

760

761 **Figure 2. CRK5 shows functional similarities to canonical CDKs and directly regulates DNA**
762 **replication during gametogony. A.** Schematic of the first mitosis following activation of male
763 gametogony. Red dots represent the nuclear poles and the green line the mitotic-spindle. Time

764 points selected for a high-time resolution phosphoproteome survey of CRK5-KO parasites are
765 indicated. **B.** Volcano plots showing the extent of differentially phosphorylated peptides in non-
766 activated and activated gametocytes after 15 sec, comparing CRK5-KO and WT parasite lines.
767 Significantly regulated sites (p-value <0.05) with a fold change >2 are highlighted in red, and in blue
768 when corresponding to the CDK S/T*PxK motif. **C.** GO term enrichment analysis for down- and up-
769 regulated phosphopeptides in CRK5-KO gametocytes during the first minute of gametogony. **D.**
770 Alignment of eight down-regulated phosphopeptides showing the classical S/T*PxK CDK motif. **E.**
771 Reduction in the number of male gametocytes replicating their DNA upon *crk5* deletion or CRK5-
772 AID/HA depletion. Proportion of male gametocytes undergoing DNA replication was determined at 1
773 min pa (post-activation) and is expressed as a percentage of polyploid (>1N) cells (error bars show
774 standard deviation from the mean; 3 independent infections; two-way ANOVA). **F.** Reduction in the
775 number of male gametocytes with spindle-like structures upon *crk5* deletion or CRK5-AID/HA
776 depletion, as assessed by α-tubulin staining 1 min pa (error bars show standard deviation from the
777 mean; 3 independent infections; two-way ANOVA). Inset pictures show representative typical tubulin
778 distribution patterns (green), observed in WT microgametocytes counter stained with DAPI (blue);
779 scale bar = 2 µm. **G.** CRK5-KO gametocytes, fifteen min after activation, have fully assembled
780 axonemes that do not initiate flagellar motility, in contrast to their WT counterparts (error bars show
781 standard deviation from the mean; 2 independent infections; two tailed unpaired t-test). Scale bar =
782 5 µm. **H.** Electron micrographs of male gametogony in wild type parasites at 3-6 minutes (a, b) and
783 15-30 minutes (c, d, e) and CRK5-KO mutant parasites at 3-6 minutes (f, g) and 15 – 30 minutes (h, i).
784 Bars represent 1 µm (a, c, f, h) and 100 nm in other images. (a) Low power micrograph an early
785 gametocyte showing the large nucleus (N) and two nuclear poles. Insert: detail of a cross-sectioned
786 axoneme. (b) Detail of the enclosed area showing the nuclear poles (NP) with an adjacent basal body
787 (BB) and elongated axoneme (A). (c) Late microgametocyte undergoing exflagellation showing the
788 nucleus (N) with some clumped heterochromatin (Ch) and a flagellum (F) protruding from the
789 surface. (d) Detail of the enclosed area showing the point of exit for the microgamete flagellum (F).

790 (e) Cross-section through a free microgamete showing the 9+2 flagellum (F) and electron dense
791 nucleus (N). (f) Low power of an early microgametocyte showing an irregular lobated nucleus (N) and
792 the presence of a nuclear pole and axonemes unusually centrally located in the cytoplasm. Insert:
793 detail of a cross-sectioned axoneme. (g) Detail of the enclose area showing the nuclear pole (NP)
794 closely associated with a basal body (BB) and longitudinally running axonemes. (h) At a late time
795 point, the parasite still resembles an early stage with lobated nucleus (N) which displays nuclear
796 poles (NP) and basal bodies (BB). (i) Detail of the enclosed area showing the nuclear pole (NP) and
797 associated cross section basal body (BB). Note the normal and abnormal axonemes (A). Source files
798 for Figure 2 are available in Figure 2 - source data 1.

799

800 **Figure 3. CRK5 interacts with a divergent cyclin and a conserved CDK regulatory subunit. A.**
801 Localisation of CRK5-GFP during male gametogony. The protein colocalises with Hoechst DNA-
802 binding dye and shows a mitotic spindle distribution during M-phases. Scale bar = 5 µm. **B.** GO term
803 enrichment analysis of proteins co-immunoprecipitated with CRK5 from gametocytes. Bonferroni
804 corrected *p*-values are indicated. **C.** Spectral count values as identified by mass spectrometry for
805 proteins co-purifying with CRK5, CDKrs, SOC2, CDPK4, SOC1, SOC3 and MCM5 (highlighted in bold)
806 following immunoprecipitation, and displayed in first and second principal components. Proteomic
807 sets (including CRKs, cyclins, MCMs, ORCs and RepC complex proteins) are additionally highlighted.
808 **D.** Phylogenetic distribution of the CSC complex components detected in alveolate genomes. Dark
809 boxes with X denote the presence of genes. **E.** Localisation of CDKrs-GFP during male gametogony.
810 The protein colocalises with Hoechst dye and shows a mitotic spindle distribution during M-phases.
811 Scale bar = 5 µm. **F.** Localisation of SOC2-3xHA by immunofluorescence in activated male
812 gametocytes. The protein colocalises with DAPI and shows overlap with mitotic spindles stained by
813 α-tubulin antibodies during M-phase. Scale bar = 5 µm. **G.** Deletions of *soc2* or *cdkrs* lead to a
814 profound defect in exflagellation (error bars show standard deviation from the mean; 3 independent

815 infections; two-way ANOVA). **H.** Reduction in the number of male gametocytes replicating their DNA
816 in parasites with *soc2* or *cdkrs* deletions. The proportion of male gametocytes undergoing DNA
817 replication was determined at 1 min pa and is expressed as the percentage of cells that are polyploid
818 (>1N) (error bars show standard deviation from the mean; 3 independent infections; two-way
819 ANOVA). **I.** Reduction in the number of male gametocytes with spindle-like structures following *soc2*
820 or *cdkrs* deletion, as assessed by α-tubulin staining 1 min pa (error bars show standard deviation
821 from the mean; 3 independent infections; two-way ANOVA). Source files for Figure 3 are available in
822 Figure 3 - source data 1.

823

824 **Figure 4. The CSC complex is stable but dynamically phosphorylated during gametogony. A.**
825 Western blot analysis of CRK5-3xHA and SOC2-3xHA over the course of gametogony shows that the
826 level of each protein is stable in absence or presence of the proteasome inhibitor MG132; α-tubulin
827 serves as a loading control. Quantification from five independent biological samples is shown to the
828 right (error bars show standard deviation from the mean). **B.** Dot blots showing the relative
829 abundance of proteins immunoprecipitated (% of all spectra) from gametocytes at 0 or 60 sec
830 activation, with CRK5-3xHA (left panel) and SOC2-3xHA (right panel). No major changes were
831 observed between the two conditions; CDT1 is however recovered less in both IPs at 60 sec. **C.**
832 Relative abundance of CRK5, SOC2, CDKrs and CDT1 in SOC2-3xHA and CRK5-3xHA IPs over the first
833 min of gametogony and in gametocytes pre-treated with the CDPK4 inhibitor BKI-1294[42] (1 µM).
834 Abundance are relative to the bait and normalised to the 0 sec time point (error bars show standard
835 deviations from the mean; data is extracted from single SOC2-3xHA and CRK5-3xHA
836 immunoprecipitations). **D.** Phosphorylation profile of immunoprecipitated CRK5 during the first
837 minute of gametogony in presence or absence of the CDPK4 inhibitor BKI-1294 (1 µM). ND = Not
838 detected. **E.** Schematic representation of the $CRK5_{1-562}$ mutant generated in this study. **F.** $CRK5_{1-562}$
839 gametocytes show a strong defect in exflagellation (error bars show standard deviation from the

840 mean; 3 independent infections; two-tailed t-test). **G.** Reduction in the number of male gametocytes
841 with a *crk5* 3' deletion replicating their DNA. The proportion of male gametocytes undergoing DNA
842 replication was determined at 10 min pa and is expressed as a percentage of cells that are polyploid
843 (>1N) (error bars show standard deviation from the mean; 3 independent infections; two-tailed t-
844 test). **H.** Possible signalling cascades regulating entry and progression through the first replicative
845 cycle during *P. berghei* gametogony. Early calcium mobilisation upon gametocyte activation leads to
846 CDPK4-dependent phosphorylation of SOC1 and the CSC complex. SOC1 may be involved in the
847 loading of the pre-replicative complex (preRC)[19]. Here, we propose that CDPK4-dependent
848 phosphorylation of SOC2 and/or CRK5 activates the kinase activity to phosphorylate components of
849 the pre-replicative complex allowing initiation of DNA replication. Indirect regulation of dyneins
850 through phosphatases such as PPM11 could possibly regulate progression into M-phase. Yellow
851 circles represent phosphorylation. Source files for Figure 4 are available in Figure 4 - source data 1 to
852 3.

853 **Supplementary material**

854 **Figure 1 - figure supplement 1. Generation of CRK5-KO and CRK5-AID/HA transgenic lines and**
855 **CRK5-GFP expression in oocysts. A-C.** Genetic modification vectors and genotyping data for CRK5-
856 KO (A), CRK5-AID/HA (B), and CRK5-GFP (C). Oligonucleotides used for PCR genotyping are indicated
857 and agarose gels for corresponding PCR products from genotyping reactions are shown. **C.** Western
858 blot detection of CRK5-GFP. **D.** CRK5-GFP expression and localisation in midgut oocysts at day 7, 14,
859 17, and 21 post-infection. Scale bar = 10 µm.

860

861 **Figure 2 - figure supplement 1. Percentage of GO terms shared between CRK5 and a set of**
862 ***Plasmodium* and human kinases.** For CRK5 GO terms, proteins with a >2 fold change (*p*-value <0.05)
863 in the CRK5-KO were retained. Source files for Figure 2 - figure supplement 1 are available in Figure 2
864 - figure supplement 1 - source data 1.

865

866 **Figure 3 - figure supplement 1. Identification and characterisation of the CSC complex**
867 **components. A-B.** Genetic modification vectors and genotyping data for CRK5-3xHA (A) and CDKrs-
868 GFP (B). Oligonucleotides used for PCR genotyping are indicated and agarose gels for corresponding
869 PCR products from genotyping reactions are shown. **C.** Phylogenetic distribution of the CSC complex
870 components and of *Plasmodium* CRKs in selected alveolate genomes. Dark boxes with X denote the
871 presence of genes. **D.** Localisation of CRK5-3xHA in 1 min activated gametocytes. Scale bar = 5 µm. **E.**
872 Genetic modification vectors and genotyping data for CDKrs-KO. **F.** Electron micrographs of the male
873 gametocytes of the SOC2-KO (a, e), CDKrs-KO (b, f), CRK5-KO (c, g) and wild type (d,h) parasites at 15
874 min pa. The three KO parasites consist of predominately early stages (95-100%) while the WT
875 showed a number of late stages (~40%) with exflagellation and free microgametes. Bars represent 1
876 µm in a-d and 100 nm in all other micrographs. (a to c) SOC2-KO, CDKrs, and CRK5 male gametocytes

877 showing the early stage of development with a lobed nucleus (N) and cytoplasmic axonemes (A).
878 Inserts. Detail of cross-sectioned axonemes. (d) WT showing a late stage in development with
879 nucleus (N) containing small clumps of condensed chromatin (Ch). Note the free microgamete
880 flagella (F). Insert. Cross-sectioned axoneme. (e) Detail of a similar stage parasite to that in (a)
881 showing a nuclear pole (NP) but note the disorganized arrangement of spindle microtubules (Mt)
882 and attached kinetochores (K). (f) Detail of the nucleus in (b) showing the nuclear pole (NP) with a
883 closely associated basal body (BB) in the cytoplasm. K – Kinetochore. (g) Detail of a similar stage to
884 that in (c) showing the nuclear pole (NP) and associated basal body (BB). (h) Detail of a similar
885 parasite to that in (d) undergoing exflagellation showing the flagellum (F) protruding from the
886 microgametocyte. Ch – condensed chromatin. Insert. Cross section through a free microgamete
887 showing the flagellum (F) and electron dense nucleus (N). Source files for Figure 3 - figure
888 supplement 1 are available in Figure 3 - figure supplement 1 - source data 1 and 2.

889

890 **Figure 4 - figure supplement 1. Study of the CSC complex regulation. A.** Flow cytometry analysis of
891 CRK5-GFP and CDKrs-GFP fluorescence over the course of gametogony in haploid or multiploid (>1N)
892 cells in the presence or absence of 1 µM of the proteasome inhibitor MG132 (error bars show
893 standard deviation from the mean; 3 or 4 independent infections; two-way ANOVA). **B.** Pre-
894 treatment of gametocytes with 1 µM MG132 leads to reduced exflagellation (error bars show
895 standard deviation from the mean; technical replicates from 3 independent infections; unpaired two-
896 tailed t-test). **C.** Treatment with 1 µM MG132 has no effect on the number of male gametocytes
897 replicating their DNA. The proportion of male gametocytes undergoing DNA replication was
898 determined at 10 min pa and is expressed as a percentage of cells that are polyploid (>1N) (error
899 bars show standard deviation from the mean; technical replicates from 3 independent infections;
900 unpaired two-tailed t-test). **D.** SOC2 phosphorylation levels show that S5088 is phosphorylated in a
901 CDPK4-dependent manner 15 sec after activation. ND = not detected. BKI = CDPK4 inhibitor. **E.**

902 Genetic modification vector and genotyping data for SOC2$^{S5088/5089A}$ mutagenesis. Oligonucleotides
903 used for PCR genotyping are indicated and agarose gels for corresponding PCR products from
904 genotyping reactions are shown. **F.** S5088/5089A substitution in SOC2-3xHA does not impair
905 exflagellation (error bars show standard deviation from the mean; technical replicates from 2
906 independent infections; two-way ANOVA). **G.** There is no reduction in the number of male
907 gametocytes replicating their DNA in SOC2$^{S5088/5089A}$-3xHA gametocytes. The proportion of male
908 gametocytes undergoing DNA replication was determined at 1 min pa and is expressed as a
909 percentage of cells that are polyploid (>1N) (error bars show standard deviation from the mean;
910 from 2 independent infections; two-way ANOVA). **H**. Genetic modification vector and genotyping
911 data for CRK5 mutagenesis. Oligonucleotides used for PCR genotyping are indicated and agarose gels
912 for corresponding PCR products from genotyping reactions are shown. Source files for Figure 4 -
913 figure supplement 1 are available in Figure 4 - figure supplement 1 - source data 1.

914

915 **Supplementary file 1. RNAseq data.**

916 **Supplementary file 2. Peptides detected in WT and CRK5-KO gametocytes.**

917 **Supplementary file 3. Phosphopeptides detected in WT and CRK5-KO gametocytes.**

918 **Supplementary file 4. Spectral count values of proteins identified in co-immunoprecipitates by**
919 **mass-spectrometry.**

920 **Supplementary file 5. Oligonucleotides used in this study.**

921 **Supplementary file 6. Correspondence between TMT labels and samples.**

922

923 **Acknowledgements**

924 We thank Julie Rodger (Nottingham University) for her assistance in the insectary maintenance and
925 Zineb Rchiad (KAUST) for RNAseq library preparation. We thank the excellent service at the
926 bioimaging and flow-cytometry core facilities at the Faculty of Medicine of the University of Geneva.
927 We also would like to thank Nisha Philip (University of Edinburgh) for sharing the 615 Tir1-expressing
928 line as well as Wesley Van Voorhis and Kayode Ojo (University of Washington) for sharing compound
929 BKI-1294. We thank Markus Ganter for sharing the reference GO term set used in this work.

930 This work was supported by the Swiss National Science Foundation grant BSSGI0_155852
931 and 31003A_179321 to MB. MB is an INSERM and EMBO young investigator. The work in RT lab is
932 supported by Medical Research Council UK (G0900109, G0900278, MR/K011782/1) and
933 Biotechnology and Biological Sciences Research Council (BB/N017609/1). The work in AP lab is
934 supported by a faculty baseline fund (BAS/1/1020-01-01) and a Competitive Research Grant (CRG)
935 award from OSR (OSR-2018-CRG6-3392) from the King Abdullah University of Science and
936 Technology.

937 **References**

938 1 **Lim S & Kaldis P**. Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development.*
939 **140**, 3079-3093, 2013, doi: 10.1242/dev.091744.
940 2 **Guo Z & Stiller JW**. Comparative genomics of cyclin-dependent kinases suggest co-evolution
941 of the RNAP II C-terminal domain and CTD-directed CDKs. *BMC Genomics.* **5**, 10.1186/1471-
942 2164-1185-1169, 2004, doi: 10.1186/1471-2164-5-69.
943 3 **Cao L, Chen F, Yang X, Xu W, Xie J & Yu L**. Phylogenetic analysis of CDK and cyclin proteins in
944 premetazoan lineages. *BMC Evol Biol.* **14**, 10.1186/1471-2148-1114-1110, 2014, doi:
945 10.1186/1471-2148-14-10.
946 4 **Talevich E, Tobin AB, Kannan N & Doerig C**. An evolutionary perspective on the kinome of
947 malaria parasites. *Philos Trans R Soc Lond B Biol Sci.* **367**, 2607-2618, 2012, doi:
948 10.1098/rstb.2012.0014.
949 5 **Tewari R, Straschil U, Bateman A, Bohme U, Cherevach I, Gong P, Pain A & Billker O**. The
950 systematic functional analysis of *Plasmodium* protein kinases identifies essential regulators
951 of mosquito transmission. *Cell Host Microbe* **8**, 377-387, 2010, doi:
952 10.1016/j.chom.2010.09.006.
953 6 **Solyakov L, Halbert J, Alam MM, Semblat JP, Dorin-Semblat D, Reininger L, Bottrill AR,**
954 **Mistry S, Abdi A, Fennell C, Holland Z, Demarta C, Bouza Y, Sicard A, Nivez MP, Eschenlauer**
955 **S, Lama T, Thomas DC, Sharma P, Agarwal S, Kern S, Pradel G, Graciotti M, Tobin AB &**
956 **Doerig C**. Global kinomic and phospho-proteomic analyses of the human malaria parasite
957 *Plasmodium falciparum*. *Nat Commun* **2**, 565, 2011, doi: 10.1038/ncomms1558.
958 7 **Dorin-Semblat D, Carvalho Teresa G, Nivez M-P, Halbert J, Poullet P, Semblat J-P, Goldring**
959 **D, Chakrabarti D, Mehra P, Dhar S, Paing May M, Goldberg Daniel E, McMillan Paul J, Tilley**
960 **L & Doerig C**. An atypical cyclin-dependent kinase controls *Plasmodium falciparum*
961 proliferation rate. *Kinome*, 4-16, 2013, doi: 10.2478/kinome-2013-0001.
962 8 **Morgan DO**. Principles of CDK regulation. *Nature.* **374**, 131-134, 1995, doi:
963 10.1038/374131a0.
964 9 **Malumbres M & Barbacid M**. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev*
965 *Cancer.* **9**, 153-166, 2009, doi: 10.1038/nrc2602.
966 10 **Merckx A, Le Roch K, Nivez MP, Dorin D, Alano P, Gutierrez GJ, Nebreda AR, Goldring D,**
967 **Whittle C, Patterson S, Chakrabarti D & Doerig C**. Identification and initial characterization
968 of three novel cyclin-related proteins of the human malaria parasite *Plasmodium falciparum*.
969 *J Biol Chem.* **278**, 39839-39850, 2003, doi: 10.1074/jbc.M301625200.
970 11 **Roques M, Wall RJ, Douglass AP, Ramaprasad A, Ferguson DJ, Kaindama ML, Brusini L,**
971 **Joshi N, Rchiad Z, Brady D, Guttery DS, Wheatley SP, Yamano H, Holder AA, Pain A,**
972 **Wickstead B & Tewari R**. *Plasmodium* P-Type Cyclin CYC3 Modulates Endomitotic Growth
973 during Oocyst Development in Mosquitoes. *PLoS Pathog.* **11**, e1005273 2015, doi:
974 10.1371/journal.ppat.1005273.
975 12 **Robbins JA, Absalon S, Streva VA & Dvorin JD**. The Malaria Parasite Cyclin H Homolog
976 PfCyc1 Is Required for Efficient Cytokinesis in Blood-Stage *Plasmodium falciparum*. *mBio* **8**,
977 e00605-00617, 2017, doi: 10.1128/mBio.00605-17.
978 13 **White MW & Suvorova ES**. Apicomplexa Cell Cycles: Something Old, Borrowed, Lost, and
979 New. *Trends in Parasitology* **34**, 759-771, 2018, doi: 10.1016/j.pt.2018.07.006.
980 14 **Billker O, Lindo V, Panico M, Etienne AE, Paxton T, Dell A, Rogers M, Sinden RE & Morris**
981 **HR**. Identification of xanthurenic acid as the putative inducer of malaria development in the
982 mosquito. *Nature* **392**, 289-292, 1998, doi: 10.1038/32667.
983 15 **Billker O & Alano P**. Gametocytes and Gametes. *W. Sherman (Ed.), Molecular Approaches to*
984 *Malaria, ASM Press, Washington, D.C.*, 191-219, 2005, doi: 10.1128/9781555817558.ch10.
985 16 **Alvarez CA & Suvorova ES**. Checkpoints of apicomplexan cell division identified in
986 *Toxoplasma gondii*. *PLoS Pathog.* **13**, e1006483, 2017, doi: 10.1371/journal.ppat.1006483.

17 **Billker O, Shaw MK, Jones IW, Ley SV, Mordue AJ & Sinden RE**. Azadirachtin disrupts formation of organised microtubule arrays during microgametogenesis of *Plasmodium berghei*. *J Eukaryot Microbiol* **49**, 489-497, 2002, doi: 10.1111/j.1550-7408.2002.tb00234.x.

18 **Zeeshan M, Ferguson DJ, Abel S, Burrrell A, Rea E, Brady D, Daniel E, Delves M, Vaughan S, Holder AA, Le Roch KG, Moores CA & Tewari R**. Kinesin-8B controls basal body function and flagellum formation and is key to malaria transmission. *Life Sci Alliance.* **2(4).** 2/4/e201900488, 2019, doi: 10.26508/lsa.201900488.

19 **Fang H, Klages N, Pardo M, Yu L, Choudhary J & Brochet M**. Multiple short windows of CDPK4 activity regulate distinct cell cycle events during *Plasmodium* gametogenesis. *Elife*, e26524, 2017, doi: 10.7554/eLife.26524.

20 **Invergo BM, Brochet M, Yu L, Choudhary J, Beltrao P & Billker O**. Sub-minute Phosphoregulation of Cell Cycle Systems during *Plasmodium* Gamete Formation. *Cell Rep.* **21**, 2017-2029, 2017, doi: 10.1016/j.celrep.2017.10.071.

21 **Philip N & Waters AP**. Conditional Degradation of *Plasmodium* Calcineurin Reveals Functions in Parasite Colonization of both Host and Vector. *Cell Host Microbe.* **18**, 122-131, 2015, doi: 10.1016/j.chom.2015.05.018.

22 **Janse CJ, Franke-Fayard B & Waters AP**. Selection by flow-sorting of genetically transformed, GFP-expressing blood stages of the rodent malaria parasite, *Plasmodium berghei*. *Nat Protoc* **1**, 614-623, 2006, doi.

23 **Mair GR, Braks JA, Garver LS, Wiegant JC, Hall N, Dirks RW, Khan SM, Dimopoulos G, Janse CJ & Waters AP**. Regulation of sexual development of *Plasmodium* by translational repression. *Science* **313**, 667-669, 2006, doi: 10.1126/science.1125129.

24 **Ganter M, Goldberg JM, Dvorin JD, Paulo JA, King JG, Tripathi AK, Paul AS, Yang J, Coppens I, Jiang RH, Elsworth B, Baker DA, Dinglasan RR, Gygi SP & Duraisingh MT**. *Plasmodium falciparum* CRK4 directs continuous rounds of DNA replication during schizogony. *Nat Microbiol.* **2:17017.**, 10.1038/nmicrobiol.2017.1017., 2017, doi: 10.1038/nmicrobiol.2017.17.

25 **Bourne Y, Watson MH, Arvai AS, Bernstein SL, Reed SI & Tainer JA**. Crystal structure and mutational analysis of the *Saccharomyces cerevisiae* cell cycle regulatory protein Cks1: implications for domain swapping, anion binding and protein interactions. *Structure* **8**, 841-580, 2000, doi: 10.1016/s0969-2126(00)00175-1.

26 **Parge HE, Arvai AS, Murtari DJ, Reed SI & Tainer JA**. Human CksHs2 atomic structure: a role for its hexameric assembly in cell cycle control. *Science.* **262**, 387-395, 1993, doi: 10.1126/science.8211159.

27 **Billker O, Dechamps S, Tewari R, Wenig G, Franke-Fayard B & Brinkmann V**. Calcium and a calcium-dependent protein kinase regulate gamete formation and mosquito transmission in a malaria parasite. *Cell* **117**, 503-514., 2004, doi: 10.1016/s0092-8674(04)00449-0.

28 **Agircan FG, Schiebel E & Mardin BR**. Separate to operate: control of centrosome positioning and separation. *Philos Trans R Soc Lond B Biol Sci* **369**, 20130461, 2014, doi: 10.1098/rstb.2013.0461.

29 **Raaijmakers JA & Medema RH**. Function and regulation of dynein in mitotic chromosome segregation. *Chromosoma* **123**, 407-422, 2014, doi: 10.1007/s00412-014-0468-7.

30 **Lopez-Barragan MJ, Lemieux J, Quinones M, Williamson KC, Molina-Cruz A, Cui K, Barillas-Mury C, Zhao K & Su XZ**. Directional gene expression and antisense transcripts in sexual and asexual stages of *Plasmodium falciparum*. *BMC Genomics* **12**, 10.1186/1471-2164-1112-1587, 2011, doi: 10.1186/1471-2164-12-587.

31 **Robbins JA, Absalon S, Streva VA & Dvorin JD**. The Malaria Parasite Cyclin H Homolog PfCyc1 Is Required for Efficient Cytokinesis in Blood-Stage *Plasmodium falciparum*. *MBio.* **8(3).** e00605-00617, 2017, doi: 10.1128/mBio.00605-17.

32 **Mikolcevic P, Sigl R, Rauch V, Hess MW, Pfaller K, Barisic M, Pelliniemi LJ, Boesl M & Geley S**. Cyclin-Dependent Kinase 16/PCTAIRE Kinase 1 Is Activated by Cyclin Y and Is Essential for

1038 Spermatogenesis. *Molecular and Cellular Biology* **32**, 868-879, 2012, doi:
1039 10.1128/mcb.06261-11.
1040 33 **Brochet M, Tobin AB, Billker O & doerig C**. The Kinomics of Malaria. *in Kinomics: Approaches*
1041 *and Applications (eds H.-B. Kraatz and S. Martic)*, 2015, doi: 10.1002/9783527683031.ch5.
1042 34 **Malumbres M**. Cyclin-dependent kinases. *Genome Biology* **15**, 122, 2014, doi:
1043 10.1186/gb4184.
1044 35 **Vincke L & Lips M**. Un nouveau *Plasmodium* d'un rongeur sauvage du Congo. *Plasmodium*
1045 *berghei n.sp. Ann Soc Belg Med Trap* **28**, 97-104, 1948, doi.
1046 36 **Janse CJ, Ramesar J & Waters AP**. High-efficiency transfection and drug selection of
1047 genetically transformed blood stages of the rodent malaria parasite *Plasmodium berghei*.
1048 *Nat Protoc* **1**, 346-356, 2006, doi.
1049 37 **Mair GR, Lasonder E, Garver LS, Franke-Fayard BM, Carret CK, Wiegant JC, Dirks RW,**
1050 **Dimopoulos G, Janse CJ & Waters AP**. Universal features of post-transcriptional gene
1051 regulation are critical for *Plasmodium* zygote development. *PLoS Pathog* **6**, e1000767, 2010,
1052 doi: 10.1371/journal.ppat.1000767.
1053 38 **Guttery DS, Ferguson DJ, Poulin B, Xu Z, Straschil U, Klop O, Solyakov L, Sandrini SM, Brady**
1054 **D, Nieduszynski CA, Janse CJ, Holder AA, Tobin AB & Tewari R**. A putative homologue of
1055 CDC20/CDH1 in the malaria parasite is essential for male gamete development. *PLoS Pathog.*
1056 **8**, e1002554, 2012, doi: 10.1371/journal.ppat.1002554.
1057 39 **Pfander C, Anar B, Brochet M, Rayner JC & Billker O**. Recombination-mediated genetic
1058 engineering of *Plasmodium berghei* DNA. *Methods Mol Biol* **923**, 127-138, 2013, doi:
1059 10.1007/978-1-62703-026-7_8.
1060 40 **Pfander C, Anar B, Schwach F, Otto TD, Brochet M, Volkmann K, Quail MA, Pain A, Rosen B,**
1061 **Skarnes W, Rayner JC & Billker O**. A scalable pipeline for highly effective genetic
1062 modification of a malaria parasite. *Nature methods* **8**, 1078-1082, 2011, doi:
1063 10.1038/nmeth.1742.
1064 41 **Brochet M, Collins MO, Smith TK, Thompson E, Sebastian S, Volkmann K, Schwach F,**
1065 **Chappell L, Gomes AR, Berriman M, Rayner JC, Baker DA, Choudhary J & Billker O**.
1066 Phosphoinositide metabolism links cGMP-dependent protein kinase G to essential Ca^{2+}
1067 signals at key decision points in the life cycle of malaria parasites. *PLoS Biology* **12**,
1068 e1001806, 2014, doi: 10.1371/journal.pbio.1001806.
1069 42 **Ojo KK, Eastman RT, Vidadala R, Zhang Z, Rivas KL, Choi R, Lutz JD, Reid MC, Fox AM,**
1070 **Hulverson MA, Kennedy M, Isoherranen N, Kim LM, Comess KM, Kempf DJ, Verlinde CL, Su**
1071 **XZ, Kappe SH, Maly DJ, Fan E & Van Voorhis WC**. A specific inhibitor of PfCDPK4 blocks
1072 malaria transmission: chemical-genetic validation. *J Infect Dis.* **209**, 275-284, 2014, doi:
1073 10.1093/infdis/jit522.
1074 43 **Volkmann K, Pfander C, Burstroem C, Ahras M, Goulding D, Rayner JC, Frischknecht F,**
1075 **Billker O & Brochet M**. The alveolin IMC1h is required for normal ookinete and sporozoite
1076 motility behaviour and host colonisation in *Plasmodium berghei*. *PLoS One* **7**, e41409, 2012,
1077 doi: 10.1371/journal.pone.0041409.
1078 44 **Katoh K & Standley DM**. MAFFT Multiple Sequence Alignment Software Version 7:
1079 Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**, 772-780,
1080 2013, doi: 10.1093/molbev/mst010.
1081 45 **Capella-Gutiérrez S, Silla-Martínez JM & Gabaldón T**. trimAl: a tool for automated
1082 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973,
1083 2009, doi: 10.1093/bioinformatics/btp348.
1084 46 **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F,**
1085 **Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ & Higgins DG**. Clustal W and Clustal
1086 X version 2.0. *Bioinformatics* **23**, 2947-2948, 2007, doi: 10.1093/bioinformatics/btm404.
1087 47 **Thompson JD, Higgins DG & Gibson TJ**. CLUSTAL W: improving the sensitivity of progressive
1088 multiple sequence alignment through sequence weighting, position-specific gap penalties

and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680, 1994, doi: 10.1093/nar/22.22.4673.

48 **Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W & Gascuel O**. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology* **59**, 307-321, 2010, doi: 10.1093/sysbio/syq010.

49 **Guindon S & Gascuel O**. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology* **52**, 696-704, 2003, doi: 10.1080/10635150390235520.

50 **Lefort V, Longueville J-E & Gascuel O**. SMS: Smart Model Selection in PhyML. *Molecular Biology and Evolution* **34**, 2422-2424, 2017, doi: 10.1093/molbev/msx149.

51 **Anisimova M, Gil M, Dufayard J-F, Dessimoz C & Gascuel O**. Survey of Branch Support Methods Demonstrates Accuracy, Power, and Robustness of Fast Likelihood-based Approximation Schemes. *Systematic Biology* **60**, 685-699, 2011, doi: 10.1093/sysbio/syr041.

52 **Anders S, Pyl PT & Huber W**. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169, 2014, doi: 10.1093/bioinformatics/btu638.

53 **McCarthy DJ, Chen Y & Smyth GK**. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**, 4288-4297, 2012, doi: 10.1093/nar/gks042.

54 **Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W & Smyth GK**. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47-e47, 2015, doi: 10.1093/nar/gkv007.

55 **Love MI, Huber W & Anders S**. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550, 2014, doi: 10.1186/s13059-014-0550-8.

56 **Wiśniewski JR, Zougman A, Nagaraj N & Mann M**. Universal sample preparation method for proteome analysis. *Nature Methods* **6**, 359-362, 2009, doi: 10.1038/nmeth.1322.

Figure 1

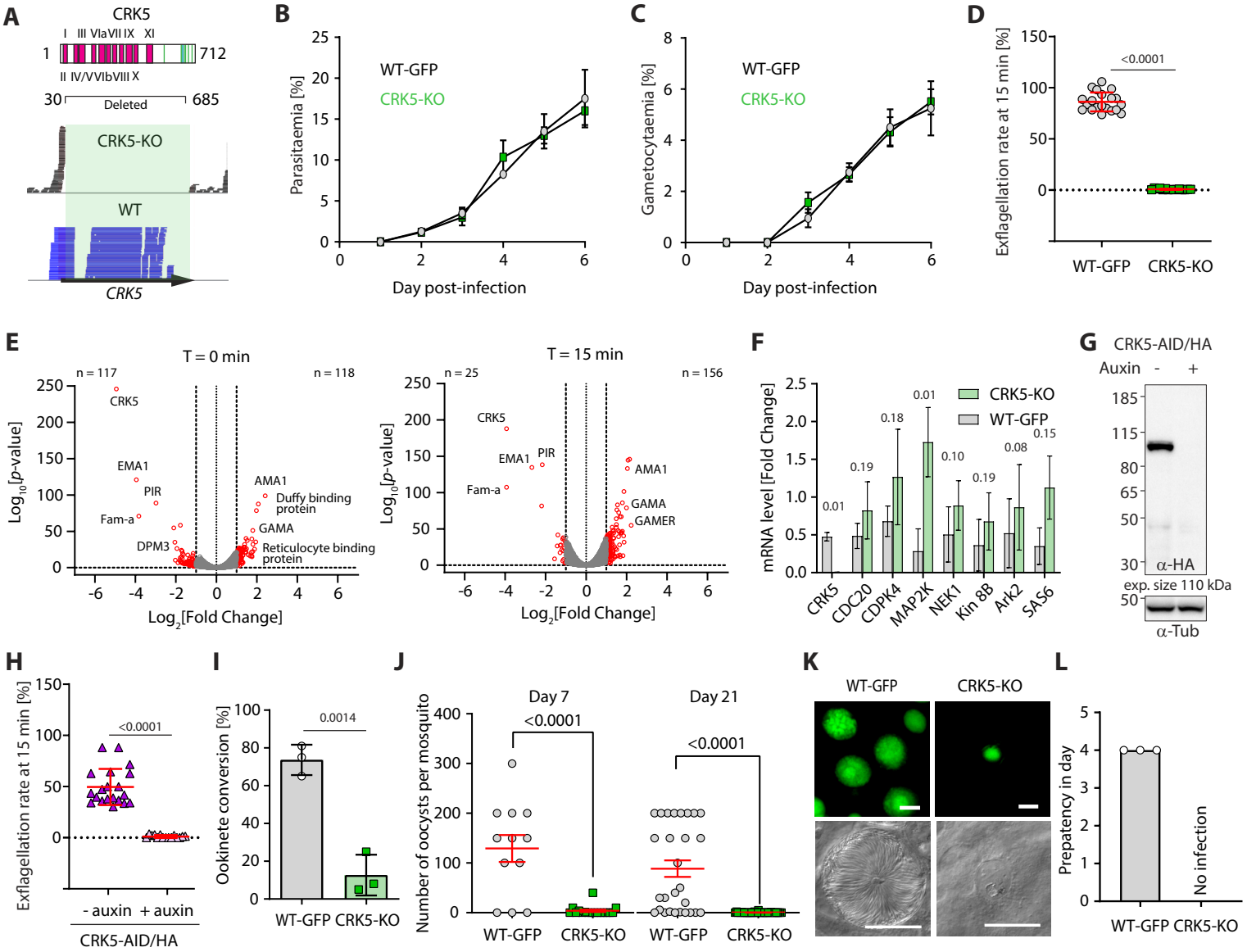

Figure 1. CRK5 is a key regulator of gametogony and sporogony during mosquito transmission.

A. Schematic representation of CRK5 showing the eleven regions that constitute the catalytic domain in pink, and the *Plasmodium*-specific C-terminal extension that is phosphorylated at eight sites (green bars). Amino acids 31 to 685 were replaced by the selection marker in the CRK5-KO line. The lower panel shows the sequence reads mapping on *crk5* in the KO and WT lines, confirming deletion of the gene in the KO.

B. Deletion of *crk5* has no effect on the growth of asexual blood stage parasites (error bars show standard deviation from the mean; 3 and 2 independent infections for the CRK5-KO and WT lines, respectively).

C. Deletion of *crk5* has no effect on gametocytaemia (error bars show standard deviation from the mean; 3 and 2 independent infections for the CRK5-KO and WT lines, respectively).

D. Deletion of *crk5* leads to a strong defect in exflagellation as measured by the percentage of active exflagellation centres per microgametocyte (error bars show standard deviation from the mean; technical replicates from 3 independent infections; two-way ANOVA).

E. Volcano plots show limited transcriptional changes between the CRK5-KO and wild type (WT) parasite lines, comparing non-activated gametocytes (T=0 min) with gametocytes 15 min after activation (T=15 min).

F. The levels of mRNA for multiple critical regulators of gametogony do not change significantly in CRK5-KO compared to WT parasites, despite a trend for up-regulation in the mutant (error bars show standard deviation from the mean; 2 independent biological replicates, two-tailed unpaired t-test).

G. Depletion of CRK5-AID/HA protein upon addition of auxin to mature purified gametocytes, as measured by western blotting; α-tubulin serves as a loading control.

H. Depletion of CRK5-AID/HA upon auxin treatment leads to a profound defect in exflagellation (error bars show standard deviation from the mean; technical replicates from 3 independent infections; two-way ANOVA).

I. Effect of *crk5* deletion on ookinete formation as a measure of fertile male gamete formation. Conversion is reported as the percentage of female gametes forming ookinetes (error bars show standard deviation from the mean; 3 independent infections, two-tailed unpaired t-test).

J. *crk5* deletion leads to a dramatic reduction in oocyst formation when mature gametocytes are fed to mosquitoes. Midguts were dissected and GFP-positive oocysts were counted at 7 and 21 days post-infection (error bars show standard deviation from the mean; 3 independent infections; two tailed unpaired t-test).

K. The residual CRK5-KO oocysts show abnormal development; scale bar = 20 µm.

L. No viable CRK5-KO sporozoite develop. Mosquitoes were infected and 21 days later allowed to feed on naive mice. For the WT line, blood stage parasites were readily observed after a 4-day prepatent period, while no mice became infected with CRK5-KO parasites (error bars show standard deviation from the mean; 3 independent infections).

Figure 2

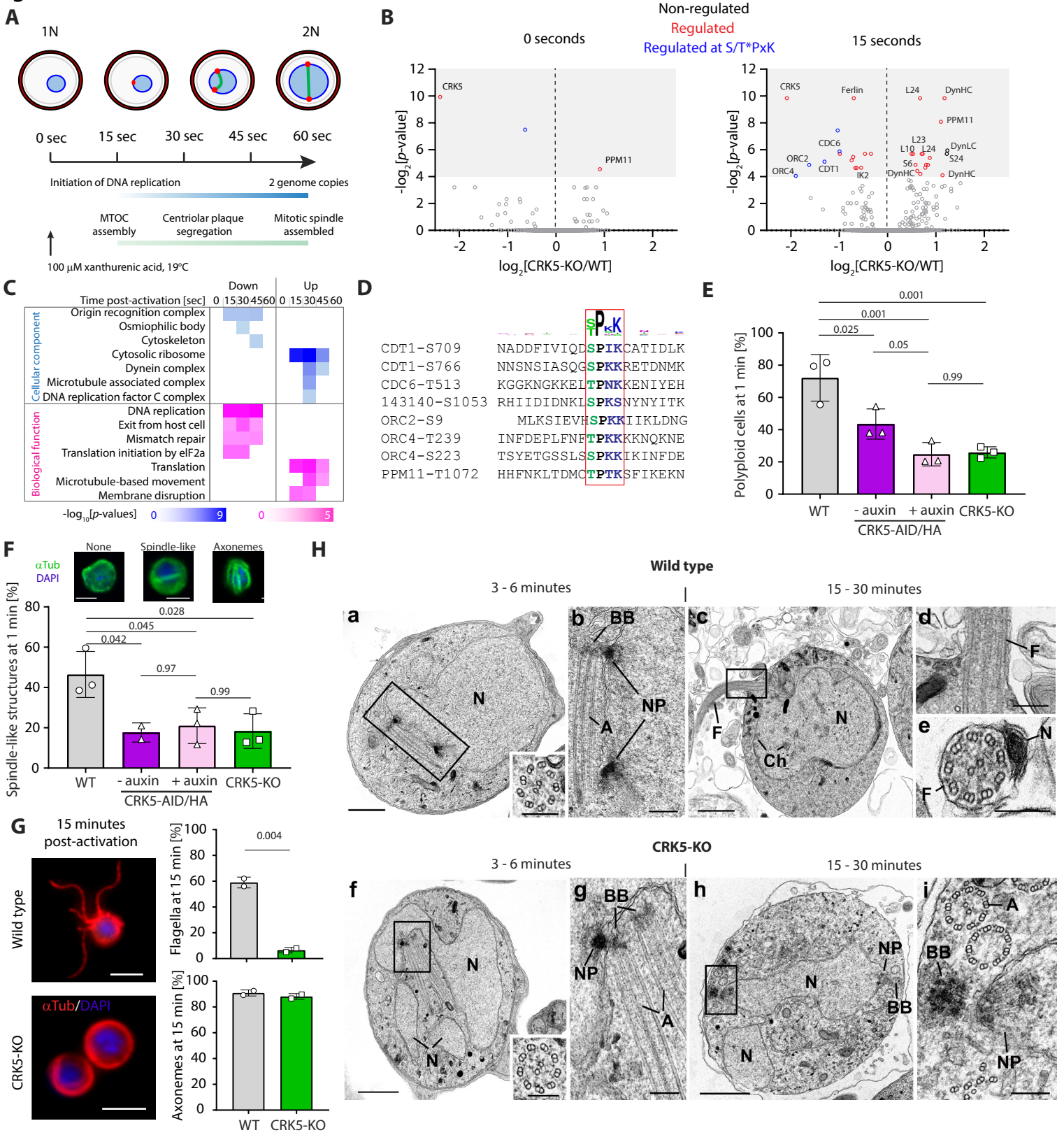

Figure 2. CRK5 shows functional similarities to canonical CDKs and directly regulates DNA replication during gametogony.

A. Schematic of the first mitosis following activation of male gametogony. Red dots represent the nuclear poles and the green line the mitotic-spindle. Time points selected for a high-time resolution phosphoproteome survey of CRK5-KO parasites are indicated.

B. Volcano plots showing the extent of differentially phosphorylated peptides in non-activated and activated gametocytes after 15 sec, comparing CRK5-KO and WT parasite lines. Significantly regulated sites (p-value <0.05) with a fold change >2 are highlighted in red, and in blue when corresponding to the CDK S/T*PxK motif.

C. GO term enrichment analysis for down- and up-regulated phosphopeptides in CRK5-KO gametocytes during the first minute of gametogony.

D. Alignment of eight down-regulated phosphopeptides showing the classical S/T*PxK CDK motif.

E. Reduction in the number of male gametocytes replicating their DNA upon *crk5* deletion or CRK5-AID/HA depletion. Proportion of male gametocytes undergoing DNA replication was determined at 1 min pa and is expressed as a percentage of polyploid (>1N) cells (error bars show standard deviation from the mean; 3 independent infections; two-way ANOVA).

F. Reduction in the number of male gametocytes with spindle-like structures upon crk5 deletion or CRK5-AID/HA depletion, as assessed by α-tubulin staining 1 min pa (error bars show standard deviation from the mean; 3 independent infections; two-way ANOVA). Inset pictures show representative typical tubulin distribution patterns (green), observed in WT microgametocytes counter stained with DAPI (blue); scale bar = 2 µm.

G. CRK5-KO gametocytes, 15 min after activation, have fully assembled axonemes that do not initiate flagellar motility, in contrast to their WT counterparts (error bars show standard deviation from the mean; 2 independent infections; two tailed unpaired t-test). Scale bar = 5 µm.

H. Electron micrographs of male gametogony in wild type parasites at 3-6 minutes (a, b) and 15-30 minutes (c, d, e) and CRK5-KO mutant parasites at 3-6 minutes (f, g) and 15 – 30 minutes (h, i). Bars represent 1 µm (a, c, f, h) and 100 nm in other images. (a) Low power micrograph an early gametocyte showing the large nucleus (N) and two nuclear poles. Insert: detail of a cross-sectioned axoneme. (b) Detail of the enclosed area showing the nuclear poles (NP) with an adjacent basal body (BB) and elongated axoneme (A). (c) Late microgametocyte undergoing exflagellation showing the nucleus (N) with some clumped heterochromatin (Ch) and a flagellum (F) protruding from the surface. (d) Detail of the enclosed area showing the point of exit for the microgamete flagellum (F). (e) Cross-section through a free microgamete showing the 9+2 flagellum (F) and electron dense nucleus (N). (f) Low power of an early microgametocyte showing an irregular lobated nucleus (N) and the presence of a nuclear pole and axonemes unusually centrally located in the cytoplasm. Insert: detail of a cross-sectioned axoneme. (g) Detail of the enclose area showing the nuclear pole (NP) closely associated with a basal body (BB) and longitudinally running axonemes. (h) At a late time point, the parasite still resembles an early stage with lobated nucleus (N) which displays nuclear poles (NP) and basal bodies (BB). (i) Detail of the enclosed area showing the nuclear pole (NP) and associated cross section basal body (BB). Note the normal and abnormal axonemes (A).

Figure 3

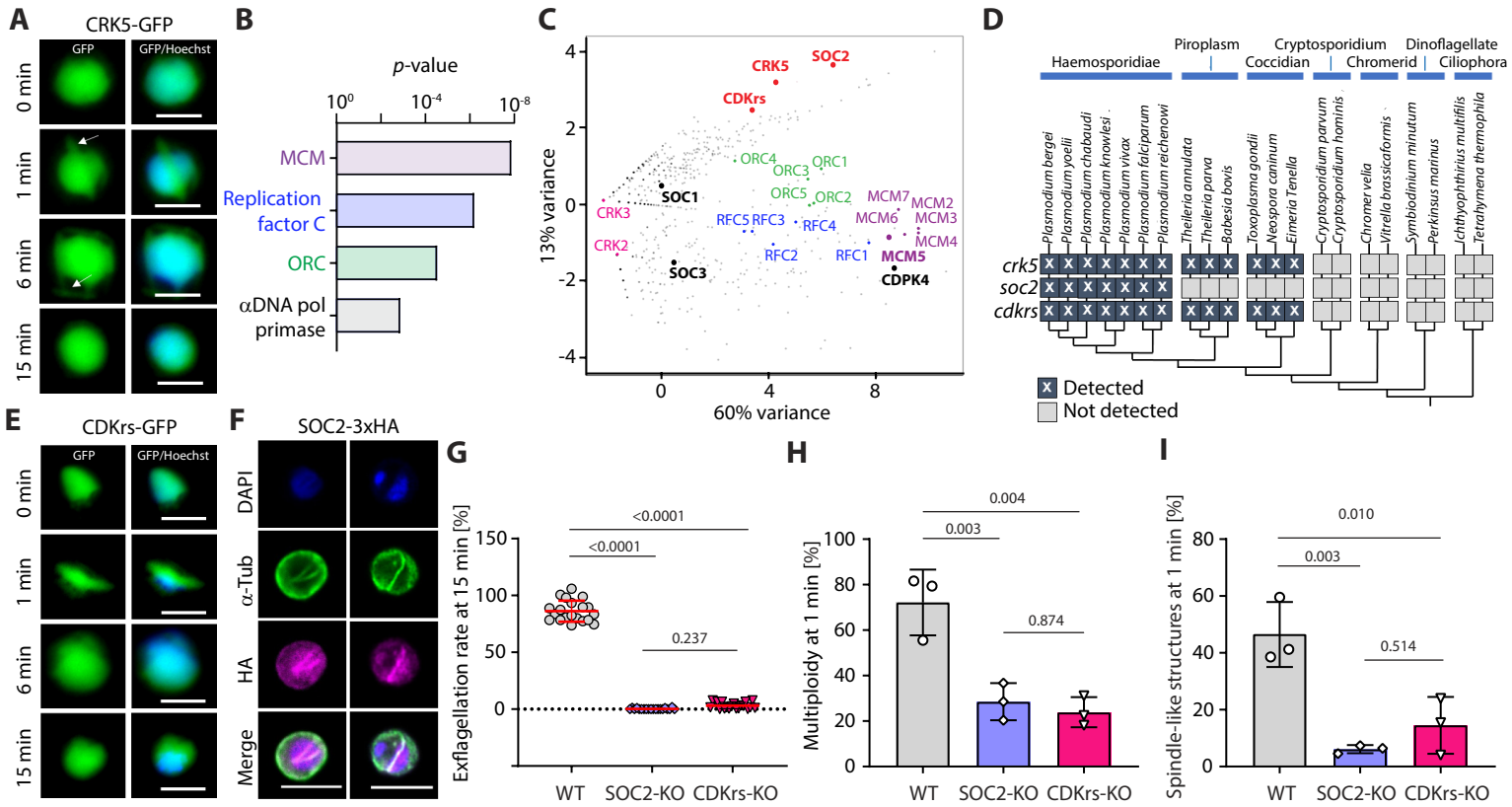

Figure 3. CRK5 interacts with a divergent cyclin and a conserved CDK regulatory subunit.
A. Localisation of CRK5-GFP during male gametogony. The protein colocalises with Hoechst DNA-binding dye and shows a mitotic spindle distribution during M-phases. Scale bar = 5 µm.

B. GO term enrichment analysis of proteins co-immunoprecipitated with CRK5 from gametocytes. Bonferroni corrected p-values are indicated.

C. Spectral count values as identified by mass spectrometry for proteins copurifying with CRK5, CDKrs, SOC2, CDPK4, SOC1, SOC3 and MCM5 following immunoprecipitation, and displayed in first and second principal components. Proteomic sets (including CRKs, cyclins, MCMs, ORCs and RepC complex proteins) are additionally highlighted.

D. Phylogenetic distribution of the CSC complex components detected in alveolate genomes. Dark boxes with X denote the presence of genes.

E. Localisation of CDKrs-GFP during male gametogony. The protein colocalises with Hoechst dye and shows a mitotic spindle distribution during M-phases. Scale bar = 5 µm.

F. Localisation of SOC2-HA by immunofluorescence in activated male gametocytes. The protein colocalises with DAPI and shows overlap with mitotic spindles stained by α-tubulin antibodies during M-phase. Scale bar = 5 µm.

G. Deletions of *soc2* or *cdkrs* lead to a profound defect in exflagellation (error bars show standard deviation from the mean; 3 independent infections; two-way ANOVA).

H. Reduction in the number of male gametocytes replicating their DNA in parasites with *soc2* or *cdkrs* deletions. The proportion of male gametocytes undergoing DNA replication was determined at 1 min pa and is expressed as the percentage of cells that are polyploid (>1N) (error bars show standard deviation from the mean; 3 independent infections; two-way ANOVA).

I. Reduction in the number of male gametocytes with spindle-like structures following *soc2* or *cdkrs* deletion, as assessed by α-tubulin staining 1 min pa (error bars show standard deviation from the mean; 3 independent infections; two-way ANOVA).

Figure 4

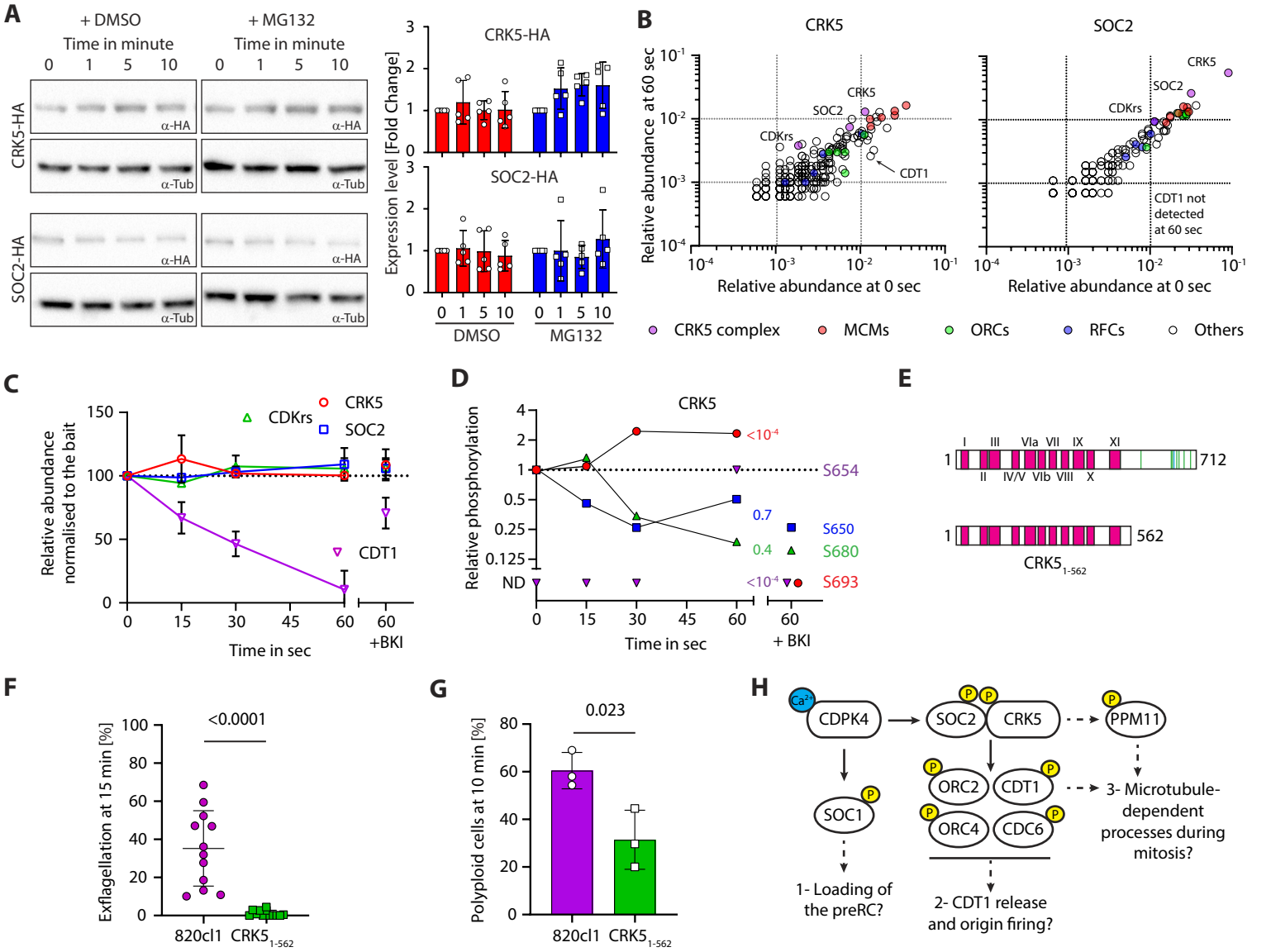

Figure 4. The CSC complex is stable but dynamically phosphorylated during gametogony.

A. Western blot analysis of CRK5-3xHA and SOC2-3xHA over the course of gametogony shows that the level of each protein is stable in absence or presence of the proteasome inhibitor MG132; α-tubulin serves as a loading control. Quantification from five independent biological samples is shown to the right (error bars show standard deviation from the mean).

B. Dot blots showing the relative abundance of proteins immunoprecipitated (% of all spectra) from gametocytes at 0 or 60 sec activation, with CRK5-3xHA (left panel) and SOC2-3xHA (right panel). No major changes were observed between the two conditions; CDT1 is however recovered less in both IPs at 60 sec.

C. Relative abundance of CRK5, SOC2, CDKrs and CDT1 in SOC2-HA and CRK5-HA IPs over the first min of gametogony and in gametocytes pre-treated with the CDPK4 inhibitor BKI-1294 (1 µM). Abundance are relative to the bait and normalised to the 0 sec time point (error bars show standard deviations from the mean; data is extracted from single SOC2-HA and CRK5-HA immunoprecipitations).

D. Phosphorylation profile of CRK5 during the first minute of gametogony in presence or absence of the CDPK4 inhibitor BKI-1294 (1 µM). ND = Not detected

E. Schematic representation of the $CRK5_{1-562}$ mutant generated in this study.

F. $CRK5_{1-562}$ gametocytes show a strong defect in exflagellation (error bars show standard deviation from the mean; 3 independent infections; two-tailed t-test).

G. Reduction in the number of male gametocytes with a *crk5* 3′ deletion replicating their DNA. The proportion of male gametocytes undergoing DNA replication was determined at 1 min pa and is expressed as a percentage of cells that are polyploid (>1N) (error bars show standard deviation from the mean; 3 independent infections; two-tailed t-test).

H. Possible signalling cascades regulating entry and progression through the first replicative cycle during *P. berghei* gametogony. Early calcium mobilisation upon gametocyte activation leads to CDPK4-dependent phosphorylation of SOC1 and the CSC complex. SOC1 may be involved in the loading of the pre-replicative complex (preRC). Here, we propose that CDPK4-dependent phosphorylation of SOC2 and/or CRK5 activates the kinase activity to phosphorylate components of the pre-replicative complex allowing initiation of DNA replication. Indirect regulation of dyneins through phosphatases such as PPM11 could possibly regulate progression into M-phase. Yellow circles represent phosphorylation.