bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

A divergent cyclin/cyclin-dependent kinase complex controls the atypical replication of *Plasmodium berghei* during gametogony and parasite transmission.

Aurélia Balestra[1,*], Mohammad Zeeshan[2,*], Edward Rea[2,¦], Carla Pasquarello[1,¦], Natacha Klages[1], Tobias Mourier[3], Amit Kumar Subudhi[3], Patrizia Arboit[1], Lorenzo Brusini[2], Rajan Pandey[2], Declan Brady[2], Sue Vaughan[4], Anthony A Holder[5], Arnab Pain[3], David 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- Computational Bioscience Research Centre, 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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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. We also provide evidence for indirect regulation of the concomitant M-phase progression. 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 unique cyclin/CDK pair fills the functional space of multiple eukaryotic cell-cycle kinases controlling DNA replication and M-phase progression.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Introduction

Progression through the cell cycle critically relies upon post-translational mechanisms including changes in activity of cell cycle kinases and phosphatases, and ubiquitin-mediated degradation of specific components once their function is complete. Central components of these networks are the cyclin-dependent kinases (CDKs). Many different families of CDK exist[1]. In most model organisms however, those primarily regulating DNA replication and mitosis belong to a monophyletic family that includes human CDK1-CDK4/6, Cdc2 and CDC28 in fission and budding yeasts, respectively, and CDKA in plants[2,3]. In malaria parasites, seven CDKs have been described[4-6]. Of these, five are divergent CDK-related kinases (CRK), i.e. they are clearly related to CDKs but have no clear orthologues in the yeast and human kinomes. These include CRK5 (PBANKA_123020), which forms a distinct branch within the CDK cluster with atypical sequence motifs in the activation loop and an additional 148 amino acid C-terminal extension found only in *Plasmodium*[7]. The gene is required, but not essential, for proliferation of asexual blood stages of the human parasite *P. falciparum*[7].

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

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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

Results

CRK5 is a key regulator of gametogony and sporogony in the mosquito. Previous attempts to disrupt *P. berghei crk5* had suggested the gene is essential for asexual blood-stage proliferation[5]. However, using long sequence homology regions to replace *crk5* with a *T. gondii* DHFR/TS resistance marker (Fig. 1A and Fig. S1A), we obtained resistant parasites and cloned them following a two-step enrichment. The gene deletion in the resulting CRK5-knockout (KO) clone was confirmed by PCR (Fig. S1B) and RNAseq analysis (Fig. 1A and Table S1). There was no significant growth defect during erythrocytic asexual multiplication (Fig. 1B), nor an inability to produce morphologically normal gametocytes (Fig. 1C). However, upon XA activation only a few microgametocytes formed active exflagellation centres (Fig. 1D). While no major transcriptional changes were detected (Fig. 1E), a

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

subtle trend in increased expression of multiple regulators of gametogony was observed in CRK5-KO gametocytes (Table S1). Quantitative PCR analysis confirmed these observations but in most cases, changes were not statistically significant (Fig. 1F). To test the possibility that the exflagellation defect was due to multiple subtle transcriptional changes, we used an additional approach to study the effect of rapid CRK5 degradation in mature gametocytes. We tagged the endogenous *crk5* with an AID/HA epitope tag (Fig. S1B) to degrade the fusion protein in presence of auxin in a strain expressing the Tir1 protein[21] (Fig. 1G). Addition of the AID/HA tag to the CRK5 C-terminus imposed a significant fitness cost, with a 2-fold decrease in exflagellation in the absence of auxin, but importantly, depletion of CRK5-AID/HA by auxin treatment of mature gametocytes resulted in a dramatic reduction in exflagellation (Fig. 1H).

Since a residual number of CRK5-KO male gametocytes produced active exflagellation centres, we asked whether the resulting microgametes were fertile by measuring the number of ookinetes per activated female gametocyte. Consistent with the dramatic reduction in male gametogony, CRK5-KO parasites showed a significant reduction in ookinete conversion compared to the parental control line (Fig. 1I). This reduction led to a > 30-fold decrease in the number of CRK5-KO oocysts present on the midgut epithelium of *Anopheles stephensi* mosquitoes allowed to feed on infected mice (Fig. 1J). CRK5-GFP is expressed in early oocysts (Fig. S1C and S1D) and residual CRK5-KO ookinetes produced non-viable oocysts (Fig. 1K), suggesting an additional role during sporogony and important for parasite transmission (Fig. 1L). Altogether, these results demonstrate a critical requirement for CRK5 during male gametogony and an essential function for CRK5 in parasite development and transmission.

Phosphoproteome kinetics point to direct phosphorylation of the pre-replicative complex by CRK5. To elucidate the processes regulated by CRK5 during gametogony, we compared the proteomes and phosphoproteomes of WT and CRK5-KO gametocytes over the first minute of gametogony covering

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

the first S-phase and metaphase of mitosis I (Fig. 2A). Apart from a significant CRK5 reduction in the CRK5-KO line, no significant differences were observed at the proteome level between mutant and parental lines (Table S2). Similarly, non-activated CRK5-KO gametocytes showed limited differences in phosphorylation compared to the parental counterpart. However, 20 phosphopeptides showed significant down-regulation upon activation (Fig. 2B) with most enriched GO terms corresponding to proteins important in DNA replication including the origin recognition complex (ORC - Fig. 2C and Table S3). This group includes a protein with similarities to the licensing factor CDT1 (chromatin licensing and DNA replication factor 1, PBANKA_1356300), a possible orthologue of the DNA replication factor CDC6 (PBANKA_1102900), and two ORC components (ORC2 - PBANKA_0803000 - and ORC4 – PBANKA_1348800). Interestingly, eight down-regulated phosphorylation sites map to the S/T*PxK CDK consensus motif (Fig. 2D). This motif shows a significant > 600 fold enrichment compared to all other phosphorylated sequences identified in this study (p-value<1.10^{-14}).

Multiple phosphopeptides were more abundant following activation (Fig. 2B) suggesting rapid secondary effects on CRK5 phospho-dependent pathways, probably via the regulation of protein phosphatases. Consistently, the phosphatase PPM11 in the mutant line showed a significantly different phosphorylation profile before and after activation. Components of the ribosome were the most significantly enriched (Fig. 2C), suggesting CRK5 can contribute indirectly to the activation of translation that is de-repressed upon gametocyte activation[22]. Interestingly, four phosphopeptides from dynein subunits were also more abundant, suggesting that the microtubule-based movement required during gametogony may also be dependent on CRK5.

The set of enriched GO terms associated with CRK5-regulation is similar to those associated with calcium dependent protein kinase (CDPK) 4 and CRK4 (Fig. S2), two kinases previously shown to be crucial for DNA replication and its initiation during gametogony[19] and schizogony[23], respectively. When compared with a set of human kinase-associated GO terms, the CRK5 set shows similarities to those of CDK1 and CDK2 as well as other cell-cycle kinases including Check1, ATM and Aurora B. Altogether this phosphoproteomic survey suggests that CRK5 is a genuine CDK that likely

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

phosphorylates components of the pre-replicative complex. It may also indirectly regulate phosphorylation of proteins involved in translation and microtubule-based movement.

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

It has been suggested that M-phase is initiated independently of S-phase completion during *Plasmodium* gametogony[19,20]. As multiple dynein-related proteins are phosphorylated in a CRK5-dependent manner, we quantified the formation of spindle-like structures and axonemes in CRK5-KO parasites. Cells depleted of CRK5 are unable to correctly assemble or maintain the spindle, with a greater than two-fold decrease of visible spindles at 1 min pa (Fig. 2F). Tubulin staining at 15 min pa showed that axoneme formation is not affected but axonemal beating is not initiated (Fig. 2G). Ultrastructural analysis by electron microscopy confirms the typical (9+2) axoneme structure however at 6 min pa there appears to be more nuclear poles present in WT compared to mutant parasites (0.56/cell compared to 0.32/cell; random sample of 50 cells of each). The cytoplasm contained electron dense basal bodies, which were often, but not always, associated with nuclear poles (Figs 2Hb and g). CRK5 deletion appeared to prevent activated microgametocytes developing to form multiple nuclear poles and they showed no chromatin condensation, cytokinesis or flagellum formation (Fig. 2H).

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

CRK5 is part of an atypical nuclear cyclin/CDK complex. The binding of cyclins to CDKs is required for progression through cell-cycle phases in model organisms. In the absence of clear S- and M-phase cyclin homologues, we searched for possible CRK5 regulatory subunits. To do this we tagged CRK5 with GFP or 3xHA at its C-terminus by modification of the endogenous gene (Fig. S3A), to facilitate localisation studies and identify any proteins bound to it. The tagged protein was localised to the nucleus in non-activated gametocytes and at 1 min pa showed an additional location at the mitotic spindle (Fig. 3A and Fig. S3D). Mass-spectrometry analyses of affinity-purified tagged protein identified 234 proteins (Table S4). Consistent with a role for CRK5 in the initiation of DNA replication, GO term enrichment analyses of proteins co-purifying with CRK5 highlighted ORC, minichromosome maintenance (MCM), RepC and alpha DNA polymerase primase complexes (Fig. 3B and C).

One of the most abundant proteins co-purified with CRK5 was SOC2 (PBANKA_144220), a protein we had previously found to be important for cell-cycle progression during gametogony[20]. SOC2 has very low homology (<13 %) with a cyclin domain in a protein of the starfish *Asterina pectinifera*[10], but the protein lacks key residues across most of the cyclin box[10] and has no detectable cyclin-like function in *in vitro* biochemical assays[11]. Analysis of CRK5 immunoprecipitates also identified multiple peptides from PBANKA_082440 (CDK regulatory subunit, CDKrs), a protein related to Cks1 and CksHs2, two CDK-associated proteins in *Saccharomyces cerevisiae*[24] and human[25], respectively. To confirm the putative interaction between CRK5, SOC2 and CDKrs, we tagged both SOC2 and CDKrs at the C-terminus by integration of 3xHA[19] or GFP coding sequence into the endogenous gene (Fig. S3B) and affinity purified the tagged proteins. Reciprocal enrichment of CRK5, SOC2 and CDKrs components was observed, consistent with the formation of a complex (Fig. 3C). In addition, each protein also co-purified the ORC, MCM, RepC and alpha DNA polymerase primase complexes (Table S4).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Previous bioinformatics analyses had identified only three *bona fide Plasmodium* cyclins[11]. Here, the identification of SOC2 as a putative CRK5 binding partner suggests it is also a cyclin. We revisited the phylogenetic analyses of apicomplexan cyclins and CDKs, and found that CRK5 and CDKrs are conserved in *Haemosporidia*, *Coccidia* and *Piroplasma*, whilst SOC2 is found only in *Haemosporidia* (Fig. 3D and Fig. S3C).

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

SOC2-KO microgametocytes showed defects in the ploidy transitions during gametogony[19]. In agreement with the notion of complementary roles for CRK5, SOC2 and CDKrs, deletion of both SOC2 and CDKrs encoding genes (Fig. S3E) also resulted in a rapid and severe reduction in exflagellation (Fig. 3G). The defect was already apparent at 1 min pa, with a >3- fold reduction of cells with 2N DNA content (Fig. 3H) and a 5- fold reduction in mitotic spindles (Fig. 3I). Ultrastructural analysis by electron microscopy of SOC2-KO and CDKrs-KO activated gametocytes confirms an early arrest during male gametogony as observed in CRK5-KO parasites (Fig. S3F). Altogether, these data strongly suggest that the CRK5/SOC2/CDKrs complex is a functional CDK/cyclin complex that controls entry into S-phase and possibly controls progression through the following or concurrent M-phase during gametogony. We propose the name 'CSC complex' for this CRK5/SOC2/CDKrs complex.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

SOC2 expression does not follow a temporal cyclin pattern during gametogony and the CSC complex is stable during the first round of mitosis. Cyclin turnover mediated by the ubiquitin-dependent proteasome pathway causes oscillations in CDK activity and therefore we wished to examine the stability of the CSC complex. Western blot and fluorescent microscopy analysis revealed that CRK5, SOC2 and CDKrs are expressed in non-activated gametocytes (Fig. 3A, 3E and 4A) and therefore we examined whether the CSC complex is degraded during the course of gametogony. Western blot analysis of CRK5-HA and SOC2-HA over the 10 minutes of gametogony revealed no evidence for degradation of either protein (Fig. 4A). To confirm that levels of CRK5 and CDKrs are maintained during gametogony in single cells, we monitored by flow cytometry the fluorescence of gametocytes expressing CRK5-GFP or CDKrs-GFP during gametogony. This analysis confirmed that there was no change in the level of CRK5-GFP or CDKrs-GFP fluorescence following male gametocyte activation and gametogony (Fig. S4A).

To examine whether the apparent stable levels of SOC2-HA and CRK5-HA result from a balance of *de novo* translation and protein degradation, we pre-treated non-activated gametocytes with the proteasome inhibitor MG132 for one hour. Although MG132 significantly reduced exflagellation (Fig. S4B), it did not affect DNA replication (Fig. S4C) nor lead to a significant accumulation of SOC2, CRK5, or CDKrs as assessed by western blot (Fig. 4A) and flow cytometry (Fig. S4A). Together, these results suggest no oscillation in levels of the CSC complex resulting from differential temporal translation and degradation during gametogony.

CDKs are usually characterised by low activity in the absence of a bound cyclin, therefore we reasoned that dynamic assembly of the CSC complex might underlie CRK5 regulation. Development of activated gametocytes is highly synchronous and we used this property to investigate the interaction between CRK5 and SOC2 in the first minute after activation. We immunoprecipitated SOC2-HA and CRK5-HA at 0, 15, 30 and 60 seconds pa and measured the relative abundance of immunoprecipitated proteins in the CSC complex by mass-spectrometry

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(Table S4). We detected no difference in the levels of SOC2 or CRK5 protein/protein interactions over this time course (Fig. 4B). However, we noticed that a CDT1-like protein became less abundant in CRK5 and SOC2 immunoprecipitates following gametocyte activation (Fig. 4C). Pre-treatment of gametocytes with BKI-294, a CDPK4 inhibitor prevented this reduction. These results suggest that SOC2-CRK5 binding is not dynamic during the first round of DNA replication. However, a CDT1-like protein is either degraded or released from the complex following CDPK4 activation.

CSC is dynamically phosphorylated during the first round of replication. Reversible phosphorylation of cyclin-CDK complexes regulates kinase activity and therefore we examined whether this mechanism may regulate the activity of the CSC complex. We had previously identified 13 phosphorylated SOC2 residues[20], with S5088/5089 phosphorylated in a CDPK4-dependent manner during early gametogony[19]. Here, S5088 was also phosphorylated in a CDPK4-dependent manner in the immunoprecipitated SOC2 (Fig. S4D), but S5088/5089 substitution with alanine (Fig. S4E) did not cause a significant defect in DNA replication or exflagellation (Fig. S4F and 4G) suggesting that SOC2-S5088/5089 phosphorylation is either not required or insufficient to regulate CSC complex function.

CRK5 is also phosphorylated dynamically during gametogony[20]. We show here that inhibition of CDPK4 with BKI-1294 affected the phosphorylation at positions S654 and S693 of immunoprecipitated CRK5 (Fig. 4D). Interestingly, all eight detected phosphorylation sites in CRK5 are located in the 148-residue C-terminal extension (Fig. 4E), which is only found in *Plasmodium* CRK5[7]. We considered that this extension may have an important role in the regulation of CRK5 activity, and to test this hypothesis we generated a parasite clonal line, $CRK5_{1-562}$, with the last 148 amino acids of CRK5 deleted, leaving an intact catalytic domain (Fig. S4H). This CRK5 truncation produced the same phenotype as that of the CRK5-KO line, with a significant reduction in exflagellation, DNA replication and mitotic spindles (Fig. 4F and G). These results suggest the

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Plasmodium-specific C-terminus extension of CRK5 is essential for its function and may act as a regulatory platform to coordinate cell cycle progression during gametogony.

Discussion

In this study, we have identified an atypical cyclin-CDK pair that is essential for male gametogony in a malaria parasite. The divergent CDK-related kinase, CRK5, interacts with a *Plasmodium*-restricted cyclin, SOC2, and with another putative regulatory subunit to form a complex we have named CSC. Given our poor understanding of the mechanisms that regulate the rapid DNA synthesis and the three rounds of mitosis, followed by chromosome condensation, nuclear division and exflagellation of gametes that comprise male gametogony, it is unclear whether CSC is required for a single checkpoint upstream of S-phase or for two independent checkpoints, necessary to enter S- and M-phases, respectively. We provide evidence of direct phosphorylation by CRK5 of multiple components of the pre-replicative complex, but further studies are necessary to understand the consequences of CRK5-dependent phosphorylation. The abundance of phosphorylated dyneins increased upon CRK5 deletion, suggesting that CRK5-dependent regulation of M-phase may be indirect, affecting microtubule-dependent processes such as formation or segregation of nuclear poles or mitotic spindle organisation[26,27] (Fig. 4H).

CRK5, SOC2 and CDKrs have similar transcriptional profiles during asexual replication in erythrocytes with a peak in late trophozoites[28], suggesting the CSC complex may have conserved control functions in cell cycle progression during schizogony. Consistent with this idea, CRK5 deletion reduced nuclear division during erythrocytic schizogony in *P. falciparum*[7], while SOC2 deletion led to slower proliferation of *P. berghei* in erythrocytes[19]. Given the versatile nature of cyclin-CDK complex function in other organisms, it is possible that other CDKs or cyclins partially compensate for the loss of CRK5 or SOC2 in asexual blood stages. For example, CRK4 activity may compensate for the lack of

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

CRK5 during schizogony, because it is also important for DNA replication during this developmental stage[23].

Usually there are significant oscillations in the abundance of cyclins during the cell cycle. However, SOC2 shows no evidence of oscillation or changes in the level of its interaction with CRK5. Such changes may be undetectable in the experimental conditions and we cannot exclude more subtle changes that may regulate the CSC complex. The observed absence of cycling may also reflect the constraints imposed by the extremely rapid nature of gametogony. No marked Cyc1 oscillations were detected during the slower *P. falciparum* blood stage schizogony and the Cyc1/MRK complex was detected in mature segmented schizonts after Cyc1 is required to complete cytokinesis[29]. Therefore, it is also possible that non-oscillating cell-cycle cyclins and stable cyclin/CDK complexes are a conserved feature of the various cycles of division within *Plasmodium* parasites. However, PbCyc3 showed fluctuating expression patterns during oocyst development[11]. It thus remains unclear whether non-cycling cyclins and stable cyclin/CDK complexes represent conserved features of the various cell cycles of *Plasmodium* parasites.

We provide evidence that the C-terminal extension of CRK5 is important in the control of its function. Interestingly, the C-terminal extensions of both SOC2 and CRK5 are Haemosporidia-specific. For example, they are absent in the related Apicomplexa parasite *Toxoplasma gondii* in which the homologue of CRK5 is also involved in DNA replication[13], highlighting a specific evolution of this cyclin/CDK system in *Plasmodium*. It is tempting to speculate that the CRK5 C-terminal extension and its phosphorylation are important in the regulation of CRK5 activity by controlling its interaction with SOC2, but this hypothesis remains untested. Very few studied CDKs contain additional domains with known function. For example, phosphorylation of the Cdk16 N-terminal domain blocks binding to cyclin Y[30]. Multiple CRK5 and SOC2 residues are phosphorylated in a CDPK4-dependent manner and it is possible that CDPK4 works as a *Plasmodium*-specific link to translate a calcium signal triggering gametogony to prime the CSC complex (Fig. 4H). In addition,

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

CDPK4-independent phosphorylation of CRK5 is also observed[20] and other kinases or phosphatases may control CSC activity during the rounds of DNA replication and mitosis. However, given the absence of canonical checkpoint kinases[5,6,31], it is possible that the successive mitotic cycles of gametogony are separated from each other by timing rather than by checkpoints that rely on signals from one event in the cell cycle to regulate the next.

Plasmodium species lack clear CDK orthologues[3] that are usually highly conserved across Eukarya and indispensable in yeasts and humans[32]. This study indicates that CRK5, despite its divergence from canonical CDKs, acts as a genuine CDK by binding a cyclin, phosphorylating targets at classical CDK motifs and regulating proteins that are targeted by conventional CDKs in human and yeast. However, it remains difficult to identify a specific direct functional orthologue due to functional similarities with multiple CDKs and even non-CDK cell-cycle kinases, and it is tempting to speculate that CRK5 and SOC2 cover the functional space of multiple cell-cycle regulators in other organisms, to satisfy the specificities of the *Plasmodium* cell cycle.

Methods

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

Parasite maintenance and transfection. *P. berghei* ANKA strain[33]-derived clones 2.34[34], 507cl1, 820cl1[35], and 615[6], together with derived transgenic lines, were grown and maintained in CD1

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

outbred mice. Six to ten week-old mice were obtained from Charles River laboratories, and females were used for all experiments. Mice were specific pathogen free and subjected to regular pathogen monitoring by sentinel screening. They were housed in individually ventilated cages furnished with a cardboard mouse house and Nestlet, maintained at 21 ± 2 °C under a 12 hr light/dark cycle, and given commercially prepared autoclaved dry rodent diet and water *ad libitum*. The parasitaemia of infected animals was determined by microscopy of methanol-fixed and Giemsa-stained thin blood smears.

For gametocyte production, parasites were grown in mice that had been phenyl hydrazine-treated three days before infection. One day after infection, sulfadiazine (20mg/L) was added in the drinking water to eliminate asexually replicating parasites. Microgametocyte exflagellation was measured three or four days after infection by adding 4 µl of blood from a superficial tail vein to 70 µl exflagellation medium (RPMI 1640 containing 25 mM HEPES, 4 mM sodium bicarbonate, 5% fetal calf serum (FCS), 100 µM xanthurenic acid, pH 7.4). To calculate the number of exflagellation centres per 100 microgametocytes, the percentage of red blood cells (RBCs) infected with microgametocytes was assessed on Giemsa-stained smears. For gametocyte purification, parasites were harvested in suspended animation medium (SA; RPMI 1640 containing 25 mM HEPES, 5% FCS, 4 mM sodium bicarbonate, pH 7.20) and separated from uninfected erythrocytes on a Histodenz/Nycodenz cushion made from 48% of a Histodenz/Nycodenz stock (27.6% [w/v] Histodenz/Nycodenz [Sigma/Alere Technologies] in 5.0 mM TrisHCl, 3.0 mM KCl, 0.3 mM EDTA, pH 7.20) and 52% SA, final pH 7.2. Gametocytes were harvested from the interface. To induce CRK5-AID/HA degradation, 500 µM auxin dissolved in ethanol was added to purified gametocytes for one hour.

Schizonts for transfection were purified from overnight *in vitro* culture on a Histodenz cushion made from 55% of the Histodenz/Nycodenz stock and 45% PBS. Parasites were harvested from the interface and collected by centrifugation at 500 *g* for 3 min, resuspended in 25 µL Amaxa Basic Parasite Nucleofector solution (Lonza) and added to 10-20 µg DNA dissolved in 10 µl H_2O. Cells were electroporated using the FI-115 program of the Amaxa Nucleofector 4D. Transfected parasites

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

were resupended in 200 µl fresh RBCs and injected intraperitoneally into mice. Parasite selection with 0.07 mg/mL pyrimethamine (Sigma) in the drinking water (pH ~4.5) was initiated one day after infection. Each mutant parasite was genotyped by PCR using three combinations of primers, specific for either the WT or the modified locus on both sides of the targeted region (experimental designs are shown in Supplemental Figures). For allelic replacements, sequences were confirmed by Sanger sequencing using the indicated primers. Controls using wild type DNA were included in each genotyping experiment; parasite lines were cloned when indicated.

Generation of DNA targeting constructs. The oligonucleotides used to generate and genotype the mutant parasite lines are in Table S5.

Restriction/ligation cloning. The C-termini of CRK5 and CDKrs were tagged with GFP by single crossover homologous recombination in the parasite. To generate these GFP lines, a region of the *crk5* or *cdkrs* gene downstream of the ATG start codon was amplified using primers T1851/T1852 and T2461/T2462, respectively, ligated into p277 vector, and transfected in ANKA line 2.34 as described previously[36]. Diagnostic PCRs were performed with primer 1 (IntT185 or intT246) and primer 2 (ol492) to confirm integration of the GFP targeting constructs. Schematic representations of the endogenous *crk5* and *cdkrs* loci, the constructs and the recombined loci are found in Fig. S1 and S3 respectively.

The gene-deletion targeting vector for *crk5* was constructed using the pBS-DHFR plasmid, which contains polylinker sites flanking a *T. gondii* dhfr/ts expression cassette conferring resistance to pyrimethamine, as described previously[5]. PCR primers N1001 and N1002 were used to generate a fragment of *crk5* 5' upstream sequence from genomic DNA, which was inserted into ApaI and HindIII restriction sites upstream of the dhfr/ts cassette of pBS-DHFR. A fragment generated with primers N1003 and N1004 from the 3' flanking region of *crk5* was then inserted downstream of the dhfr/ts cassette using EcoRI and XbaI restriction sites. The linear targeting sequence was released using ApaI/XbaI and the construct was transfected into the ANKA line 507cl1 expressing GFP. Diagnostic

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

PCR was performed with primer 1 (IntN100) and primer 2 (ol248) to confirm integration of the targeting construct, and primer 3 (N100 KO1) and primer 4 (N100 KO2) were used to confirm deletion of the *crk5* gene. A schematic representation of endogenous *crk5*, the constructs and the recombined locus is in Fig. S1.

PlasmoGEM vectors. 3xHA and AID/3xHA tagging of CRK5, CDKrs-KO, $CRK5_{1-562}$ and SOC2 allelic replacement construct were generated using phage recombineering in *Escherichia coli* TSA strain with PlasmoGEM vectors (http://plasmogem.sanger.ac.uk/). For final targeting vectors not available in the PlasmoGEM repository, generation of knock-out and tagging constructs was performed using sequential recombineering and gateway steps as previously described[37,38]. For each gene of interest (goi), the Zeocin-resistance/Phe-sensitivity cassette was introduced using oligonucleotides *goi* HA-F x *goi* HA-R and *goi* KO-F x *goi* KO-R for 3xHA tagging and KO targeting vectors, respectively. Insertion of the GW cassette following gateway reaction was confirmed using primer pairs GW1 x *goi* QCR1 and GW2 x *goi* QCR2. The modified library inserts were then released from the plasmid backbone using NotI. The CRK5-AID/HA targeting vector was transfected into the 615 parasite line[21], the CDKrs-KO, and $CRK5_{1-562}$ vectors into the 820cl1 line and the CRK5-HA vector into the 2.34 line.

Substitutions of $SOC2^{S5088/5089A}$ residues were introduced with a two-step strategy using λ Red-ET recombineering as described in reference[39]. The first step involved the insertion by homologous recombination of a Zeocin-resistance/Phe-sensitivity cassette flanked by 5' and 3' sequences of the codon of interest, which is amplified using the *soc2*-delF x *soc2*-delR primer pair. Recombinant bacteria were then selected on Zeocin. The recombination event was confirmed by PCR and a second round of recombination replaced the Zeocin-resistance/Phe-sensitivity cassette with a PCR product containing the S5088/5089A substitutions amplified using *soc2*-mut-F with *soc2*-mutR primer pairs, respectively. Bacteria were selected on YEG-Cl kanamycin plates. Mutations were confirmed by sequencing vectors isolated from colonies sensitive to Zeocin with primers *soc2*-seqF

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

to *soc2*-seqR. The modified library insert was then released from the plasmid backbone using NotI and the construct transfected into the 801cl1 line.

Parasite phenotype analyses. Blood containing approximately 50,000 parasites of the CRK5-KO line was injected intraperitoneally into mice to initiate infections. Four to five days post infection, exflagellation and ookinete conversion were examined as described previously[36] with a Zeiss AxioImager M2 microscope (Carl Zeiss, Inc) fitted with an AxioCam ICc1 digital camera. When indicated, gametocytes were pre-treated 5 min in SA supplemented with 1 µM BKI-1294 or 60 min in SA supplemented with 1 µM MG132. To analyse mosquito transmission, 30 to 50 *Anopheles stephensi* SD 500 mosquitoes were allowed to feed for 20 min on anaesthetized, infected mice whose asexual parasitaemia had reached 15% and were carrying comparable numbers of gametocytes as determined on Giemsa-stained blood films. To assess mid-gut infection, approximately 15 guts were dissected from mosquitoes on day 7, 14, and 21 post-feeding and oocysts were counted using the Zeiss AxioImager M2 microscope and a 63x oil immersion objective. Mosquito bite-backs were performed 21 days post-feeding using naive mice and blood smears were examined after 3-4 days.

Immunofluorescence labelling. Gametocyte immunofluorescence assays were performed as previously described[40]. For HA and α-tubulin staining, purified cells were fixed with 4% paraformaldehyde and 0.05% glutaraldehyde in PBS for 1 h, permeabilised with 0.1% Triton X-100/PBS for 10 min and blocked with 2% BSA/PBS for 2 h. Primary antibodies were diluted in blocking solution (rat anti-HA clone 3F10, 1:1000; mouse anti-α-tubulin clone DM1A, 1:1000, both from Sigma-Aldrich). Anti-rat Alexa594, anti-mouse Alexa488, anti-rabbit Alexa 488, Anti-rabbit Alexa594 were used as secondary antibodies together with DAPI (all from Life technologies), all diluted 1:1000 in blocking solution. Confocal images were acquired with a LSM700 or a LSM800 scanning confocal microscope (Zeiss).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

The CRK5-KO gametocytes were purified and activated in ookinete medium, then fixed at 15 min pa with 4% paraformaldehyde (PFA, Sigma) diluted in microtubule stabilising buffer (MTSB) for 10-15 min and added to poly-L-lysine coated slides. Immunocytochemistry was performed using primary mouse anti-α-tubulin mAb (Sigma-T9026; 1:1000) and secondary antibody Alexa 488 conjugated anti-mouse IgG (Invitrogen-A11004; 1:1000). The slides were then mounted in Vectashield 19 with DAPI (Vector Labs) for fluorescence microscopy. Parasites were visualised on a Zeiss AxioImager M2 microscope fitted with an AxioCam ICc1 digital camera (Carl Zeiss, Inc).

Flow cytometry analysis of gametocyte DNA content and CRK5/CDKrs-GFP fluorescence. DNA content of microgametocytes was determined by flow cytometry measurement of fluorescence intensity of cells stained with Vybrant dye cycle violet (life Technologies). Gametocytes were purified and resuspended in 100 µl of SA. Activation was induced by adding 100 µl of modified exflagellation medium (RPMI 1640 containing 25 mM HEPES, 4 mM sodium bicarbonate, 5% FCS, 200 µM xanthurenic acid, pH 7.8). To rapidly block gametogony, 800 µl of ice cold PBS was added and cells were stained for 30 min at 4°C with Vybrant dye cycle violet and analysed using a Beckman Coulter Gallios 4. Microgametocytes were selected on fluorescence by gating on GFP positive microgametocytes when the 820cl1 parasite line or its derivatives were used. In this case, cell ploidy was expressed as a percentage of male gametocytes only. When the 2.34 line or its derivatives were analysed, gating was performed on both micro- and macro-gametocytes and cell ploidy was expressed as a percentage of all gametocytes. For each sample, >50,000 cells were analysed.

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Phylogenetic analyses. Annotated protein sequences were downloaded for the following organisms: *P. falciparum* (pfal), *P. reichenowi* (prei), *P. knowlesi* (pkno), *P. vivax* (pviv), *P. berghei* (pber), *P. chabaudi* (pcha), and *P. yoelii* (all from PlasmoDB, [plasmodb.org]). *Theileria annulata* (tann), *T. parva* (tpar), and *Babesia bovis* (bbov) (all from PiroplasmaDB, [piroplasmadb.org]). *Toxoplasma gondii* (tgon), *Neospora caninum* (ncan), and *Eimeria tenella* (eten) (all from ToxoDB, [toxodb.org]). *Cryptosporidium hominis* (chom), *C. parvum* (cpar), *Chromera velia* (cvel), and *Vitrella brassicaformis* (vbra) (all from CryptoDB, [cryptodb.org]). *Perkinsus marinus* (pmar), *Ichthyophthirius multifiliis* (imul), and *Tetrahymena thermophilia* (tthe) (all from EnsemblProtists, [protists.ensembl.org]). *Symbiodinium minutum* (smin) (from OIST, [marinegenomics.oist.jp]).

Reciprocal best hits to *P. berghei* CRK sequences were determined from pair-wise protein BLAST searches. For each CRK protein, reciprocal best BLAST matches were retrieved and aligned using mafft[41]. Alignments were trimmed using trimal[42] and HMM models were produced with HMMer (hmmer.org). These models were then used to search the genomes for which no reciprocal best BLAST hit was detected.

A neighbour-joining tree was then constructed using four combined data sets: 1) the *P. berghei* CRK proteins, 2) all protein sequences with a reciprocal best BLAST hit to *P. berghei*, 3) genes that were best BLAST hit for a *P. berghei* protein, but did not have the CRK protein as best hit in *P. berghei*, 4) all protein sequences with a score below $1x10^{-12}$ to an HMM model. Sequences and identifiers are provided in the source data files of Fig. S3.

Sequence distances were calculated using clustalw2[43,44], which was also used to construct the NJ tree. Protein sequences from 1) and 2) were divided into their CRK groups (*i.e.* CRK1-7, CDKrs, and SOC2), and all pair-wise distances were calculated within each CRK group. For all protein sequences from 3) and 4), the distance to all members in all CRK groups were then measured. A protein sequence was included in a CRK group if the average distance between the protein sequence and all members of a given CRK group did not exceed 40% of all pair-wise distances within that CRK

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

group, or if the CRK group to which a protein sequence had the lowest average distance was lower than the distance to the CRK group with the second-lowest average distance minus 5 standard deviations. Final trees were produced by PhyML[45,46] with the GTR substitution model selected by SMS[47]. Branch support was evaluated with the Bayesian-like transformation of approximate likelihood ratio test, aBayes[48]. Trees were visualized using FigTree (tree.bio.ed.ac.uk/software/figtree/).

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

RNA-seq read alignments were mapped onto the *P. berghei* ANKA genome (May 2015 release in GeneDB—http://www.genedb.org/) using Tophat2 (version 2.0.13) with parameters "-library-type fr-firststrand–no-novel-juncs–r 60". Transcript abundances were extracted as raw read counts using the Python script 'HTseq-count'[49] (model type – union, http://www-huber.embl.de/users/anders/HTSeq/). Counts per million (cpm) values were obtained from count data and gene were filtered if they failed to achieve a cpm value of 1 in at least 30% of samples per condition. Library sizes were normalized by the TMM method using EdgeR software[50] and further subjected to linear model analysis using the voom function in the limma package[51]. Differential expression analysis was performed using DESeq2 in R version 3.2.1[52]. Genes with fold change greater

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

than 2 and p-value less than 0.05 were considered as significantly differentially expressed. *P. berghei* GO terms (Gene Ontology) were downloaded from GeneDB (http://www.genedb.org/; May 2015 release) and gene ontology enrichment analysis was performed for the DEG lists using GOstats R package. All analyses and visualizations were done with R packages- *cummeRbund* and *ggplot2.*

Quantitative Real Time PCR (qRT-PCR) analyses. RNA was isolated from purified gametocytes using an RNA purification kit (Stratagene). cDNA was synthesised using an RNA-to-cDNA kit (Applied Biosystems). Gene expression was quantified from 80 ng of total RNA using SYBR green fast master mix kit (Applied Biosystems). All the primers were designed using primer3 (Primer-blast, NCBI). Analysis was conducted using an Applied Biosystems 7500 fast machine with the following cycling conditions: 95˚C for 20 s followed by 40 cycles of 95˚C for 3 s; 60˚C for 30 s. Three technical replicates and two biological replicates were performed for each assayed gene. The *hsp70* (PBANKA_081890) and *arginyl-t RNA synthetase* (PBANKA_143420) genes were used as endogenous control reference genes. The primers used for qPCR can be found in Table S5.

Protein immunoprecipitation and identification.

Sample preparation. Co-immunoprecipitations (IPs) of protein were performed with purified gametocytes. For HA-based IPs, samples were fixed for 10 min with 1% formaldehyde. Parasites were lysed in RIPA buffer (50 mM Tris HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) and the supernatant was subjected to affinity purification with anti-HA antibody (Sigma) or anti-GFP antibody (Invitrogen) conjugated to magnetics beads. Beads were re-suspended in 100 µl of 6 M urea in 50 mM ammonium bicarbonate (AB). Two µl of 50 mM dithioerythritol (DTE) were added and the reduction was carried out at 37°C for 1h. Alkylation was performed by adding 2 µl of 400 mM iodoacetamide for 1 h at room temperature in the dark. Urea was reduced to 1 M by addition of 500 µl AB and overnight digestion was performed at 37 °C with 5

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

µl of freshly prepared 0.2 µg/µl trypsin (Promega) in AB. Supernatants were collected and completely dried under speed-vacuum. Samples were then desalted with a C18 microspin column (Harvard Apparatus) according to manufacturer's instructions, completely dried under speed-vacuum and stored at -20°C.

Liquid chromatography electrospray ionisation tandem mass spectrometry (LC-ESI-MSMS). Samples were diluted in 20 µl loading buffer (5% acetonitrile [CH_3CN], 0.1% formic acid [FA]) and 2 µl were injected onto the column. LC-ESI-MS/MS was performed either on a Q-Exactive Plus Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific) equipped with an Easy nLC 1000 liquid chromatography system (Thermo Fisher Scientific) or an Orbitrap Fusion Lumos Tribrid mass Spectrometer (Thermo Fisher Scientific) equipped with an Easy nLC 1200 liquid chromatography system (Thermo Fisher Scientific). Peptides were trapped on an Acclaim pepmap100, 3 µm C18, 75 µm x 20 mm nano trap-column (Thermo Fisher Scientific) and separated on a 75 µm x 250 mm (Q-Exactive) or 500 mm (Orbitrap Fusion Lumos), 2µm C18, 100 Å Easy-Spray column (Thermo Fisher Scientific). The analytical separation used a gradient of H_2O/0.1% FA (solvent A) and 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 % A and 35 % B 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 rate was 250 nL/min for a total run time of 90 min.

Data-dependant analysis (DDA) was performed on the Q-Exactive Plus with MS1 full scan at a resolution of 70,000 Full width at half maximum (FWHM) followed by MS2 scans on up to 15 selected precursors. MS1 was performed with an AGC target of 3 x 10^6, a maximum injection time of 100 ms and a scan range from 400 to 2000 m/z. MS2 was performed at a resolution of 17,500 FWHM with an automatic gain control (AGC) target at 1 x 105 and a maximum injection time of 50 ms. Isolation window was set at 1.6 m/z and 27% normalised collision energy was used for higher-energy collisional dissociation (HCD). DDA was performed on the Orbitrap Fusion Lumos with MS1 full scan at a resolution of 120,000 FWHM followed by as many subsequent MS2 scans on selected precursors as possible within a 3 sec maximum cycle time. MS1 was performed in the Orbitrap with

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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. MS2 was performed in the Ion Trap with a rapid scan rate, an AGC target of 1 x 10^4 and a maximum injection time of 35 ms. Isolation window was set at 1.2 m/z and 30% normalised collision energy was used for HCD.

Database searches. Peak lists (MGF file format) were generated from raw data using the MS Convert conversion tool from ProteoWizard. The peak list files were searched against the PlasmoDB_*P.berghei* ANKA database (PlasmoDB.org, release 38, 5076 entries) combined with an in-house database of common contaminants using Mascot (Matrix Science, London, UK; version 2.5.1). Trypsin was selected as the enzyme, with one potential missed cleavage. Precursor ion tolerance was set to 10 ppm and fragment ion tolerance to 0.02 Da for Q-Exactive Plus data and to 0.6 for Lumos data. Variable amino acid modifications were oxidized methionine and deamination (Asn and Gln) as well as phosphorylated serine, threonine and tyrosine. Fixed amino acid modification was carbamidomethyl cysteine. The Mascot search was validated using Scaffold 4.8.4 (Proteome Software) with 1% of protein false discovery rate (FDR) and at least 2 unique peptides per protein with a 0.1% peptide FDR.

Proteomics and phosphoproteomics.

Sample preparation (SDS buffer-FASP procedure). Activated and non-activated purified gametocytes were snap frozen in liquid nitrogen. Cell lysis was performed in 1 ml of 2% SDS, 25 mM NaCl, 2.5 mM EDTA, 20 mM TCEP, 50 mM TrisHCl (pH 7.4), supplemented with 1x Halt™ protease and phosphatase inhibitor (Thermo Fisher). Samples were vortexed and then heated at 95°C for 10 min with mixing at 400 rpm on a thermomixer. DNA was sheared *via* four sonication pulses of 10 s each at 50% power. Samples were then centrifuged for 30 min at 17,000 *g* and the supernatant was collected. Three-hundred µl sample was incubated with 48 µl 0.5 M iodoacetamide for 1h at room temperature. Protein was digested based on the FASP method[53] using Amicon® Ultra-4, 30 kDa cut-off as

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

centrifugal filter units (Millipore). Trypsin (Promega) was added at 1:100 enzyme/protein ratio and digestion was performed at 37°C overnight. The resulting peptides were treated as described above.

TMT10plex-labelling. Peptide concentration was determined using a colorimetric peptide assay (Pierce). A pool made from 1/20 of each sample was created and used as reference for each tandem mass tag (TMT) experiment. Briefly, 100 µg of each sample was labelled with 400 µg of corresponding TMT-10plex reagent previously dissolved in 110 µl 36% CH_3CN, 200 mM 4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid (EPPS, pH 8.5). The reference sample was labelled with TMT11-131C label reagent. The reaction was performed for 1h at room temperature and then quenched by adding hydroxylamine to a final concentration of 0.3% (v/v). Labelled samples of each TMT10 experiment were pooled together, dried and desalted with a peptide desalting spin column (Pierce) according to manufacturer's instructions.

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

High pH Reverse-Phase fractionation. The flow-through fraction of each TMT10 experiment was fractionated into 13 fractions using the Pierce™ High pH Reversed-Phase Peptide Fractionation Kit (Thermo Fisher Scientific) according to manufacturer's instructions.

LC-ESI-MSMS. For all samples, peptide concentration was determined using a colorimetric peptide assay (Pierce). Phosphopeptides were reconstituted in loading buffer (5% CH_3CN, 0.1% FA) and 2 µg injected on to the column. LC-ESI-MS/MS was performed on an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific) equipped with an Easy nLC1200 liquid chromatography system (Thermo Fisher Scientific). Peptides were trapped on a Acclaim pepmap100, C18, 3 µm, 75 µm x 20 mm nano trap-column (Thermo Fisher Scientific) and separated on a 75 µm x 500 mm, C18, 2 µm, 100 Å Easy-Spray column (Thermo Fisher Scientific). The analytical separation was run

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

for 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 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 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 composition. Flow rate was of 250 nL/min. DDA was performed with MS1 full scan at a resolution of 120,000 FWHM followed by as many subsequent MS2 scans on selected precursors as possible within a 3 sec maximum cycle time. MS1 was performed in the Orbitrap with an AGC target of $4x10^5$, a maximum injection time of 50 ms and a scan range from 375 to 1500 *m/z*. MS2 was performed in the Orbitrap using HCD at 38% Normalised collision Energy (NCE). Isolation window was set at 0.7 u with an AGC target of $5x10^4$ and a maximum injection time of 86 ms. A dynamic exclusion of parent ions of 60 s with 10 ppm mass tolerance was applied.

High pH Reversed-Phase Peptide fractions were reconstituted in the loading buffer (5% CH_3CN, 0.1% FA) and 1 µg was injected onto the column. LC-ESI-MS/MS was performed as described above. The analytical separation was run for 90 min using a gradient of 99.9% H_2O/0.1% FA (solvent A) and 80% 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 % 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 composition. Flow rate was 250 nL/min. DDA was performed using the same parameters, as described above.

Database search. Raw data were processed using Proteome Discoverer (PD) 2.3 software (Thermo Fisher). Briefly, spectra were extracted and searched against the *P berghei* ANKA database (PlasmoDB.org, release 38, 5076 entries) combined with an in-house database of common contaminants using Mascot (Matrix Science, London, UK; version 2.5.1). Trypsin was selected as the enzyme, with one potential missed cleavage. Precursor ion tolerance was set to 10 ppm and fragment ion tolerance to 0.02 Da. Carbamidomethyl of cysteine (+57.021) as well as TMT10plex (+229.163) on lysine residues and on peptide N-termini were specified as fixed modifications. Oxidation of methionine (+15.995) as well as phosphorylated serine, threonine and tyrosine were set as variable modifications. The search results were validated with a Target Decoy PSM validator. PSM

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

and peptides were filtered with a FDR of 1%, and then grouped to proteins, again with a FDR of 1% and using only peptides with high confidence level. Both unique and razor peptides were used for quantitation and protein and peptides abundances value were based on signal to noise (S/N) values of reporter ions. Abundances were normalised on "Total Peptide Amount" and then scaled with "On Controls Average" (i.e. using the reference sample channel). All the protein ratios were calculated from the medians of the summed abundances of replicate groups and associated p-values were calculated with an ANOVA test based on individual protein or peptides. GO term enrichment was performed using PlasmoDB. The reference set of GO terms for human kinases was obtained from [23].

Statistical analysis. All statistical analyses were performed using GraphPad Prism 8 (GraphPad Software). Statistical tests are mentioned in the figure legends.

Data availability. All relevant data are available from the authors on request. Mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD017283 (proteomics and phosphoproteomics) and PXD017308 (HA immunoprecipitations). RNAseq data have been deposited to the Gene Expression Omnibus under accession number GSE144743.

Figure legends

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(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).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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 (post-activation) 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, fifteen 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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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. 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.** Overview of the proteins co-immunoprecipitated with CRK5-HA, CRK5-GFP, SOC2-HA and CDKrs-GFP in gametocytes. **D.** Maximum likelihood tree of the CSC complex components detected across Apicomplexa. Branch support values evaluated with aBayes are shown on branches. **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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(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. 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.**

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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)[19]. 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.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Supplementary material

Figure S1. Generation of CRK5-KO, CRK5-AID/HA, and CDKrs-GFP transgenic lines and CRK5-GFP expression in oocysts. A-C. Genetic modification vectors and genotyping data for CRK5-KO (A), CRK5-AID/HA (B), and CRK5-GFP (C). Oligonucleotides used for PCR genotyping are indicated and agarose gels for corresponding PCR products from genotyping reactions are shown. **C.** Western blot detection of CRK5-GFP. **D.** CRK5-GFP expression and localisation in midgut oocysts at day 7, 14, 17, and 21 post-infection. Scale bar = 10 µm.

Figure S2. Percentage of GO terms shared between CRK5 and a set of *Plasmodium* and human kinases. For CRK5 GO terms, proteins with a >2 fold change (*p*-value <0.05) in the CRK5-KO were retained.

Figure S3. Identification and characterisation of the CSC complex components. A-B. Genetic modification vectors and genotyping data for CRK5-HA (A) and CDKrs-GFP (B). Oligonucleotides used for PCR genotyping are indicated and agarose gels for corresponding PCR products from genotyping reactions are shown. **C.** Maximum likelihood tree of the CRK, SOC2 and CDKrs proteins detected across Apicomplexa (see Methods). Branch support values evaluated with aBayes are shown on branches. **D.** Localisation of CRK5-HA in 1 min activated gametocytes. Scale bar = 5 µm. **E.** Genetic modification vectors and genotyping data for CDKrs-KO. **F.** Electron micrographs of the male gametocytes of the SOC2-KO (a, e), CDKrs-KO (b, f), CRK5-KO (c, g) and wild type (d,h) parasites at 15 min pa. The three KO parasites consist of predominately early stages (95-100%) while the WT showed a number of late stages (~40%) with exflagellation and free microgametes. Bars represent 1 µm in a-d and 100 nm in all other micrographs. (a to c) SOC2-KO, CDKrs, and CRK5 male gametocytes showing the early stage of development with a lobed nucleus (N) and cytoplasmic axonemes (A).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Inserts. Detail of cross-sectioned axonemes. (d) WT showing a late stage in development with nucleus (N) containing small clumps of condensed chromatin (Ch). Note the free microgamete flagella (F). Insert. Cross-sectioned axoneme. (e) Detail of a similar stage parasite to that in (a) showing a nuclear pole (NP) but note the disorganized arrangement of spindle microtubules (Mt) and attached kinetochores (K). (f) Detail of the nucleus in (b) showing the nuclear pole (NP) with a closely associated basal body (BB) in the cytoplasm. K – Kinetochore. (g) Detail of a similar stage to that in (c) showing the nuclear pole (NP) and associated basal body (BB). (h) Detail of a similar parasite to that in (d) undergoing exflagellation showing the flagellum (F) protruding from the microgametocyte. Ch – condensed chromatin. Insert. Cross section through a free microgamete showing the flagellum (F) and electron dense nucleus (N).

Figure S4. Study of the CSC complex regulation. A. Flow cytometry analysis of CRK5-GFP and CDKrs-GFP fluorescence over the course of gametogony in haploid or multiploid (>1N) cells in the presence or absence of 1 µM of the proteasome inhibitor MG132 (error bars show standard deviation from the mean; 3 or 4 independent infections; two-way ANOVA). **B.** Pre-treatment of gametocytes with 1 µM MG132 leads to reduced exflagellation (error bars show standard deviation from the mean; technical replicates from 3 independent infections; unpaired two-tailed t-test). **C.** Treatment with 1 µM MG132 has no effect on the number of male gametocytes 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 (>1 nucleus) (error bars show standard deviation from the mean; technical replicates from 3 independent infections; unpaired two-tailed t-test). **D.** SOC2 phosphorylation levels show that S5088 is phosphorylated in a CDPK4-dependent manner 15 sec after activation. ND = not detected. BKI = CDPK4 inhibitor. **E.** Genetic modification vector and genotyping data for SOC2$^{S5088/5089A}$ mutagenesis. Oligonucleotides used for PCR genotyping are indicated and agarose gels for corresponding PCR products from genotyping reactions are shown. **F.**

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

S5088/5089A substitution in SOC2-3xHA does not impair exflagellation (error bars show standard deviation from the mean; technical replicates from 2 independent infections; two-way ANOVA). **G.** There is no reduction in the number of male gametocytes replicating their DNA in SOC2$^{S5088/5089A}$-3xHA gametocytes. 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; from 2 independent infections; two-way ANOVA). **H**. Genetic modification vector and genotyping data for CRK5 mutagenesis. Oligonucleotides used for PCR genotyping are indicated and agarose gels for corresponding PCR products from genotyping reactions are shown.

Table S1. RNAseq data.

Table S2. Peptides detected in WT and CRK5-KO gametocytes.

Table S3. Phosphopeptides detected in WT and CRK5-KO gametocytes.

Table S4. Proteins identified in co-immunoprecipitates by mass-spectrometry.

Table S5. Oligonucleotides used in this study.

Acknowledgements

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

References

1 **Lim S & Kaldis P**. Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development.* **140**, 3079-3093, (2013).

2 **Guo Z & Stiller JW**. Comparative genomics of cyclin-dependent kinases suggest co-evolution of the RNAP II C-terminal domain and CTD-directed CDKs. *BMC Genomics.* **5**, 10.1186/1471-2164-1185-1169, (2004).

3 **Cao L, Chen F, Yang X, Xu W, Xie J & Yu L**. Phylogenetic analysis of CDK and cyclin proteins in premetazoan lineages. *BMC Evol Biol.* **14**, 10.1186/1471-2148-1114-1110, (2014).

4 **Talevich E, Tobin AB, Kannan N & Doerig C**. An evolutionary perspective on the kinome of malaria parasites. *Philos Trans R Soc Lond B Biol Sci.* **367**, 2607-2618, (2012).

5 **Tewari R, Straschil U, Bateman A, Bohme U, Cherevach I, Gong P, Pain A & Billker O**. The systematic functional analysis of *Plasmodium* protein kinases identifies essential regulators of mosquito transmission. *Cell Host Microbe* **8**, 377-387, (2010).

6 **Solyakov L, Halbert J, Alam MM, Semblat JP, Dorin-Semblat D, Reininger L, Bottrill AR, Mistry S, Abdi A, Fennell C, Holland Z, Demarta C, Bouza Y, Sicard A, Nivez MP, Eschenlauer S, Lama T, Thomas DC, Sharma P, Agarwal S, Kern S, Pradel G, Graciotti M, Tobin AB & Doerig C**. Global kinomic and phospho-proteomic analyses of the human malaria parasite *Plasmodium falciparum*. *Nat Commun* **2**, 565, (2011).

7 **Dorin-Semblat D, Carvalho Teresa G, Nivez M-P, Halbert J, Poullet P, Semblat J-P, Goldring D, Chakrabarti D, Mehra P, Dhar S, Paing May M, Goldberg Daniel E, McMillan Paul J, Tilley L & Doerig C**. An atypical cyclin-dependent kinase controls *Plasmodium falciparum* proliferation rate. *Kinome*, 4-16, (2013).

8 **Morgan DO**. Principles of CDK regulation. *Nature.* **374**, 131-134, (1995).

9 **Malumbres M & Barbacid M**. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer.* **9**, 153-166, (2009).

10 **Merckx A, Le Roch K, Nivez MP, Dorin D, Alano P, Gutierrez GJ, Nebreda AR, Goldring D, Whittle C, Patterson S, Chakrabarti D & Doerig C**. Identification and initial characterization of three novel cyclin-related proteins of the human malaria parasite *Plasmodium falciparum*. *J Biol Chem.* **278**, 39839-39850, (2003).

11 **Roques M, Wall RJ, Douglass AP, Ramaprasad A, Ferguson DJ, Kaindama ML, Brusini L, Joshi N, Rchiad Z, Brady D, Guttery DS, Wheatley SP, Yamano H, Holder AA, Pain A, Wickstead B & Tewari R**. *Plasmodium* P-Type Cyclin CYC3 Modulates Endomitotic Growth during Oocyst Development in Mosquitoes. *PLoS Pathog.* **11**, e1005273 (2015).

12 **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**, e00605-00617, (2017).

13 **White MW & Suvorova ES**. Apicomplexa Cell Cycles: Something Old, Borrowed, Lost, and New. *Trends in Parasitology* **34**, 759-771, (2018).

14 **Billker O, Lindo V, Panico M, Etienne AE, Paxton T, Dell A, Rogers M, Sinden RE & Morris HR**. Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito. *Nature* **392**, 289-292, (1998).

15 **Billker O & Alano P**. Gametocytes and Gametes. *W. Sherman (Ed.), Molecular Approaches to Malaria, ASM Press, Washington, D.C.*, 191-219, (2005).

16 **Alvarez CA & Suvorova ES**. Checkpoints of apicomplexan cell division identified in *Toxoplasma gondii*. *PLoS Pathog.* **13**, e1006483, (2017).

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).

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

flagellum formation and is key to malaria transmission. *Life Sci Alliance.* **2(4).** 2/4/e201900488, (2019).

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).

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).

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).

22 **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).

23 **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).

24 **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).

25 **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).

26 **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).

27 **Raaijmakers JA & Medema RH**. Function and regulation of dynein in mitotic chromosome segregation. *Chromosoma* **123**, 407-422, (2014).

28 **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).

29 **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).

30 **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 Spermatogenesis. *Molecular and Cellular Biology* **32**, 868-879, (2012).

31 **Brochet M, Tobin AB, Billker O & doerig C**. The Kinomics of Malaria. *in Kinomics: Approaches and Applications (eds H.-B. Kraatz and S. Martic)*, (2015).

32 **Malumbres M**. Cyclin-dependent kinases. *Genome Biology* **15**, 122, (2014).

33 **Vincke L & Lips M**. Un nouveau *Plasmodium* d'un rongeur sauvage du Congo. *Plasmodium berghei n.sp. Ann Soc Belg Med Trap* **28**, 97-104, (1948).

34 **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).

35 **Mair GR, Lasonder E, Garver LS, Franke-Fayard BM, Carret CK, Wiegant JC, Dirks RW, Dimopoulos G, Janse CJ & Waters AP**. Universal features of post-transcriptional gene regulation are critical for *Plasmodium* zygote development. *PLoS Pathog* **6**, e1000767, (2010).

36 **Guttery DS, Ferguson DJ, Poulin B, Xu Z, Straschil U, Klop O, Solyakov L, Sandrini SM, Brady D, Nieduszynski CA, Janse CJ, Holder AA, Tobin AB & Tewari R**. A putative homologue of

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

CDC20/CDH1 in the malaria parasite is essential for male gamete development. *PLoS Pathog.* **8**, e1002554, (2012).

37 **Pfander C, Anar B, Brochet M, Rayner JC & Billker O**. Recombination-mediated genetic engineering of *Plasmodium berghei* DNA. *Methods Mol Biol* **923**, 127-138, (2013).

38 **Pfander C, Anar B, Schwach F, Otto TD, Brochet M, Volkmann K, Quail MA, Pain A, Rosen B, Skarnes W, Rayner JC & Billker O**. A scalable pipeline for highly effective genetic modification of a malaria parasite. *Nature methods* **8**, 1078-1082, (2011).

39 **Brochet M, Collins MO, Smith TK, Thompson E, Sebastian S, Volkmann K, Schwach F, Chappell L, Gomes AR, Berriman M, Rayner JC, Baker DA, Choudhary J & Billker O**. Phosphoinositide metabolism links cGMP-dependent protein kinase G to essential Ca^{2+} signals at key decision points in the life cycle of malaria parasites. *PLoS Biology* **12**, e1001806, (2014).

40 **Volkmann K, Pfander C, Burstroem C, Ahras M, Goulding D, Rayner JC, Frischknecht F, Billker O & Brochet M**. The alveolin IMC1h is required for normal ookinete and sporozoite motility behaviour and host colonisation in *Plasmodium berghei*. *PLoS One* **7**, e41409, (2012).

41 **Katoh K & Standley DM**. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**, 772-780, (2013).

42 **Capella-Gutiérrez S, Silla-Martínez JM & Gabaldón T**. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973, (2009).

43 **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ & Higgins DG**. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948, (2007).

44 **Thompson JD, Higgins DG & Gibson TJ**. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680, (1994).

45 **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).

46 **Guindon S & Gascuel O**. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology* **52**, 696-704, (2003).

47 **Lefort V, Longueville J-E & Gascuel O**. SMS: Smart Model Selection in PhyML. *Molecular Biology and Evolution* **34**, 2422-2424, (2017).

48 **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).

49 **Anders S, Pyl PT & Huber W**. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169, (2014).

50 **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).

51 **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).

52 **Love MI, Huber W & Anders S**. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550, (2014).

53 **Wiśniewski JR, Zougman A, Nagaraj N & Mann M**. Universal sample preparation method for proteome analysis. *Nature Methods* **6**, 359-362, (2009).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 1

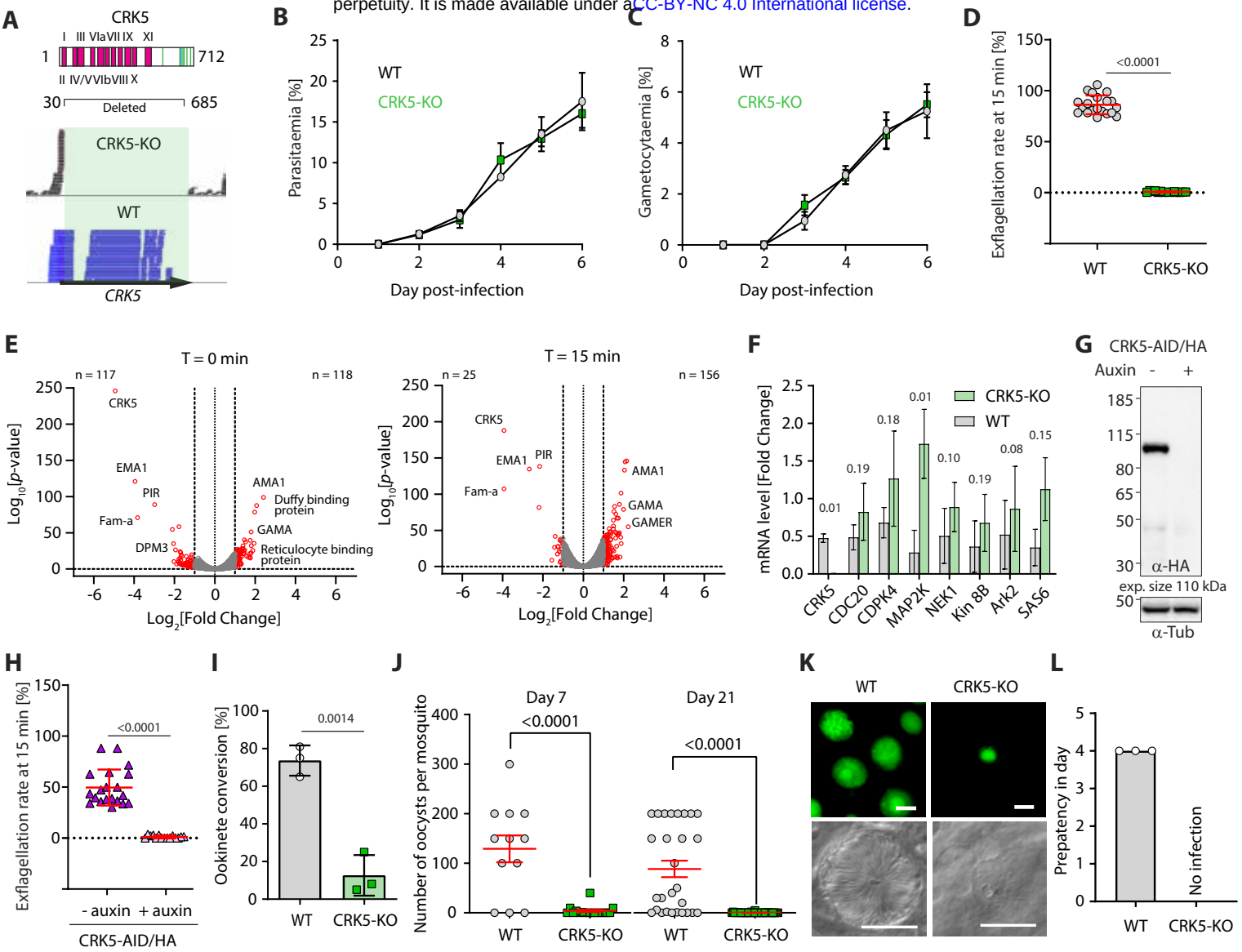

Figure 1. CRK5 is a key regulator of gametogony and sporogony during mosquito transmission.

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).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 2

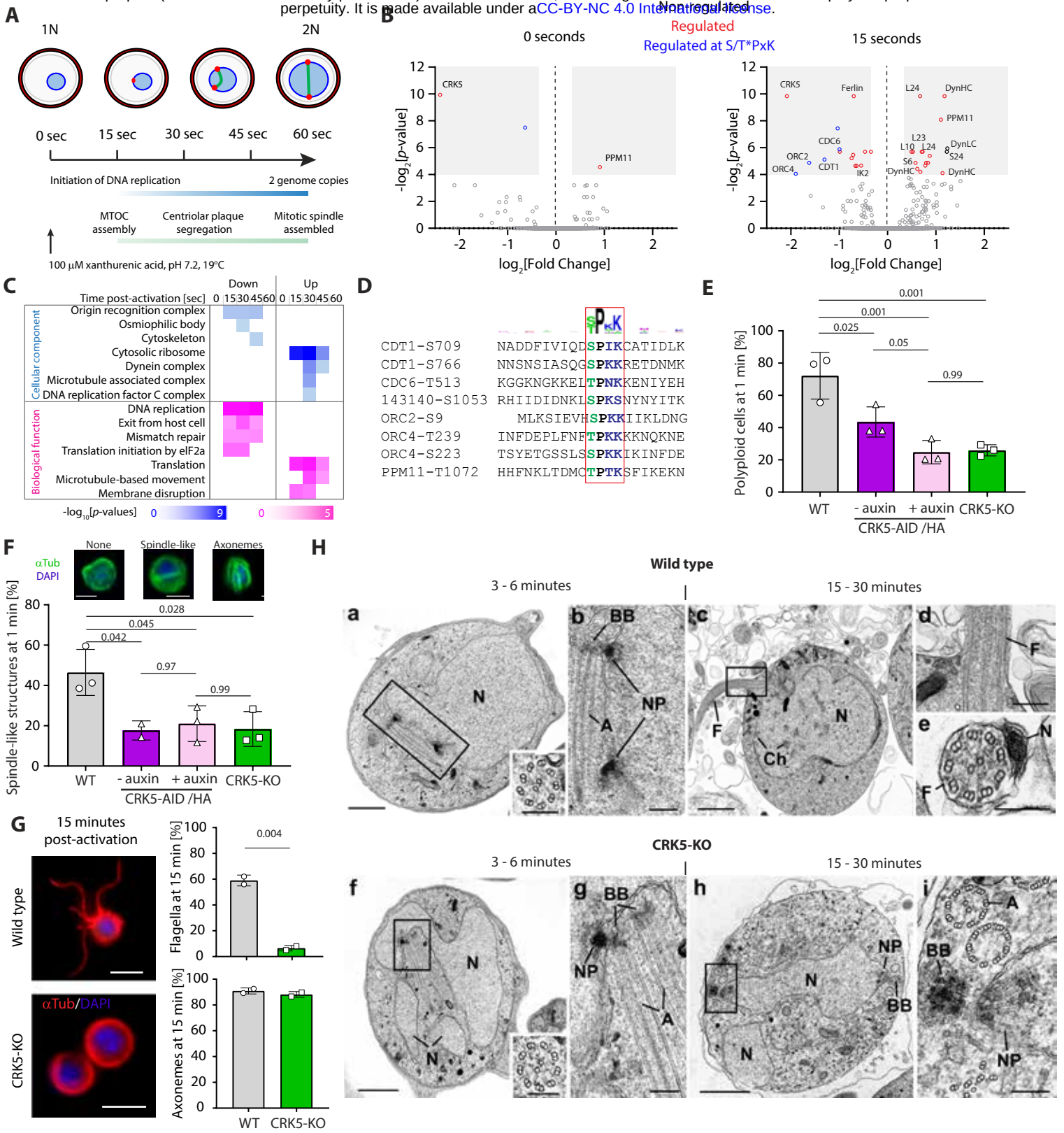

Figure 2. CRK5 shows functional similarities to canonical CDKs and directly regulates DNA replication during gametogony.

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

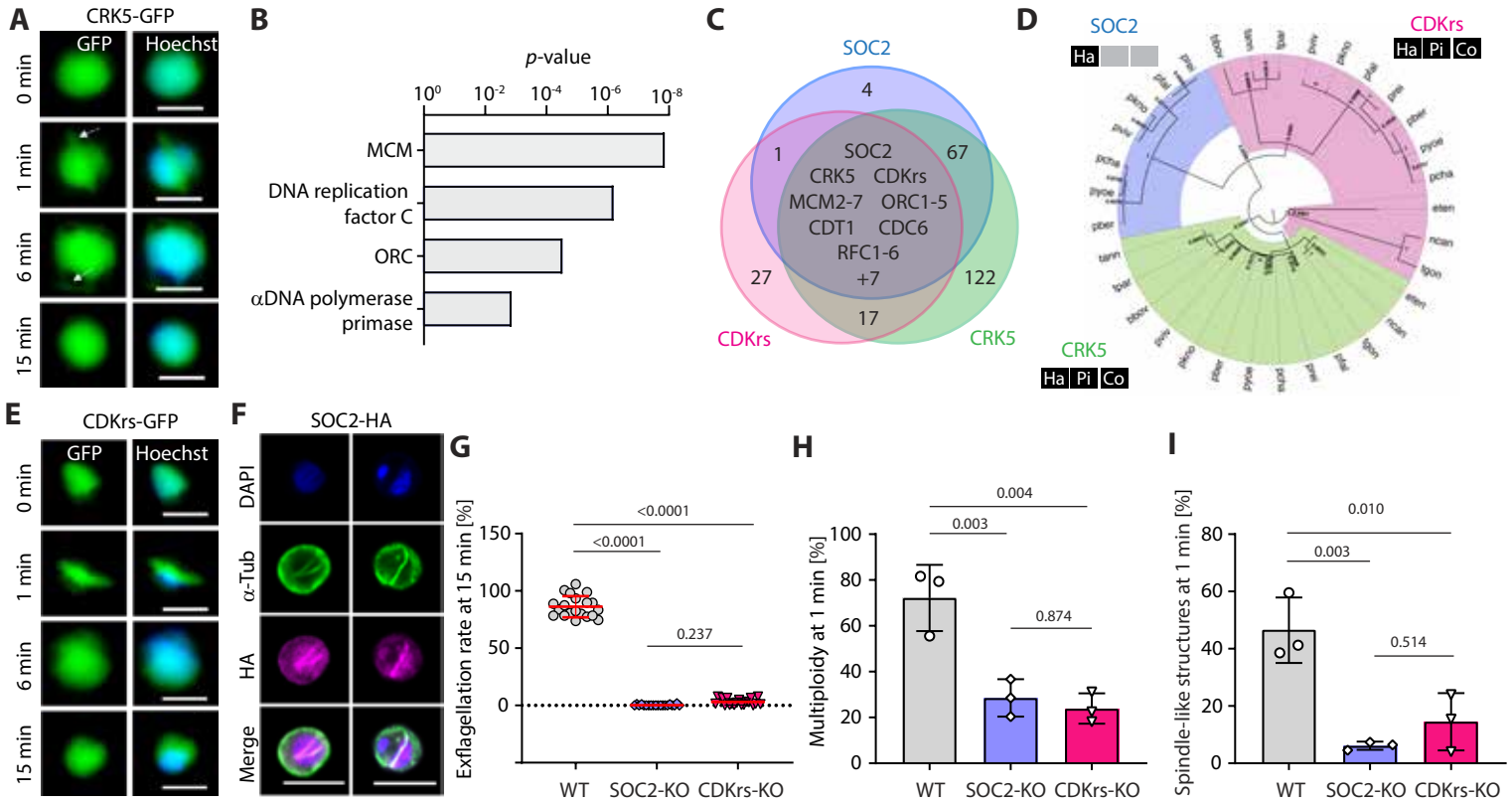

Figure 3. CRK5 interacts with a divergent cyclin and a conserved CDK regulatory subunit.

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. Overview of the proteins co-immunoprecipitated with CRK5-HA, CRK5-GFP, SOC2-HA and CDKrs-GFP in gametocytes.

D. Maximum likelihood tree of the CSC complex components detected across Apicomplexa. Branch support values evaluated with aBayes are shown on branches.

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.14.928432; this version posted February 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

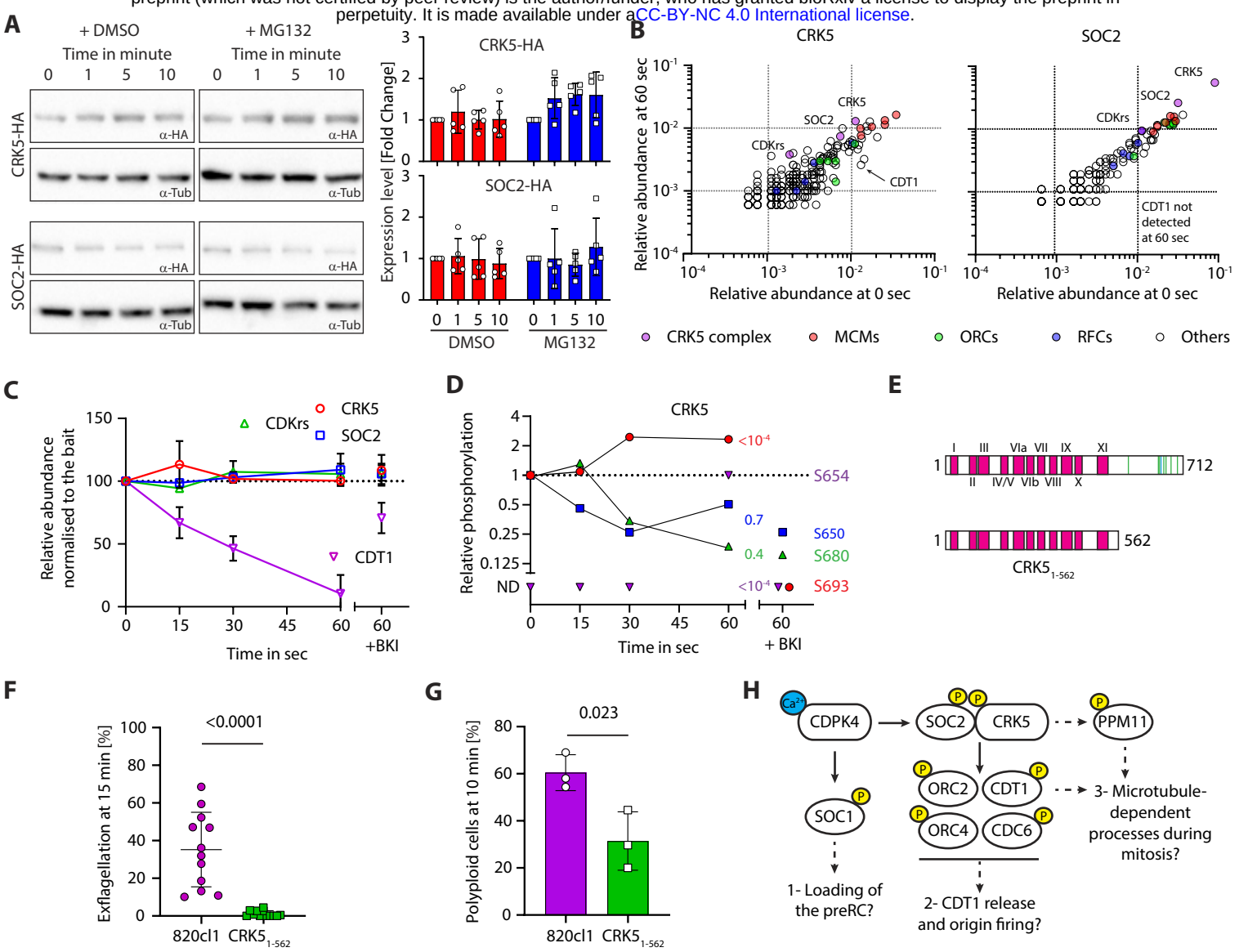

Figure 4. The CSC complex is stable but dynamically phosphorylated during gametogony.

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.