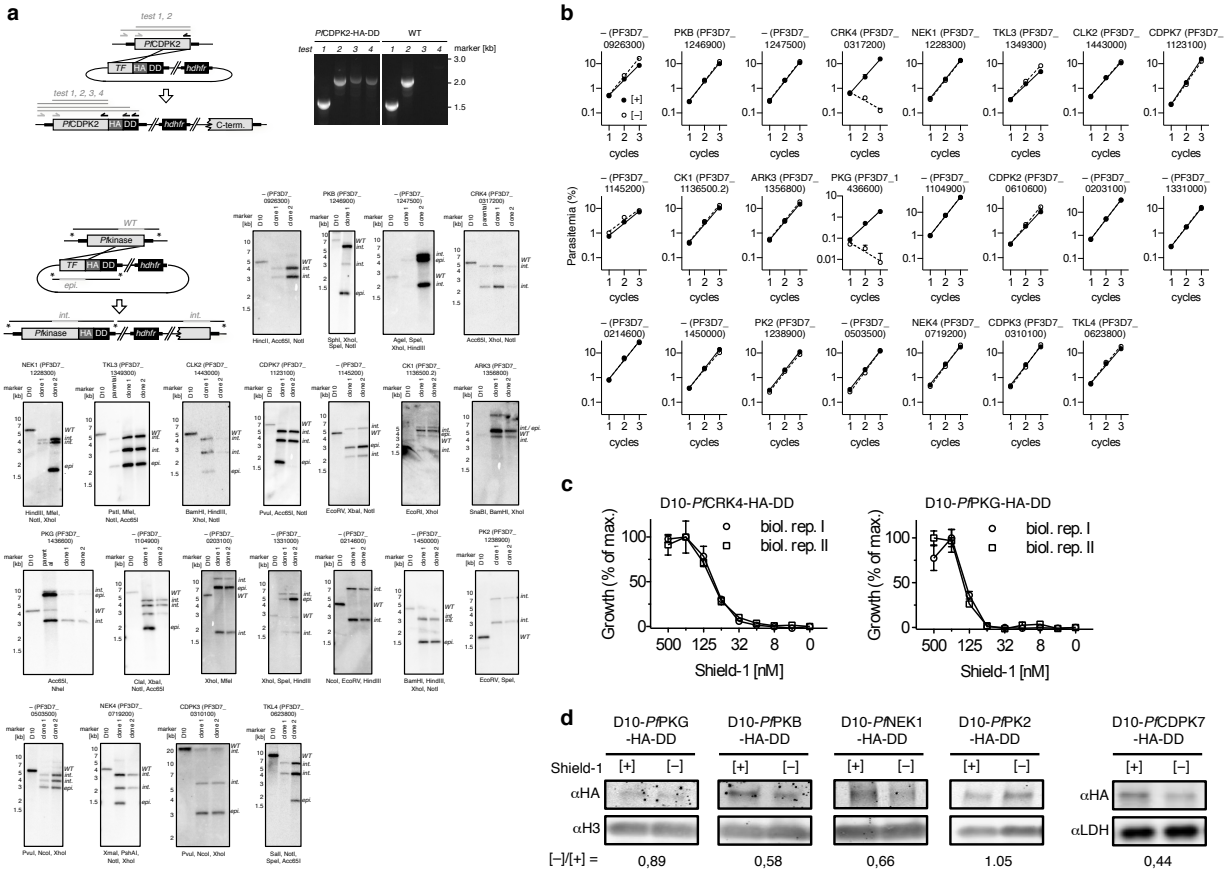
# Supplementary Information

# *Plasmodium falciparum* CRK4 directs continuous rounds of DNA replication during schizogony

Markus Ganter, Jonathan M. Goldberg, Jeffrey D. Dvorin, Joao A. Paulo, Jonas G. King, Abhai K. Tripathi, Aditya S. Paul, Jing Yang, Isabelle Coppens, Rays H.Y. Jiang, Brendan Elsworth, David A. Baker, Rhoel R. Dinglasan, Steven P. Gygi, Manoj T. Duraisingh*

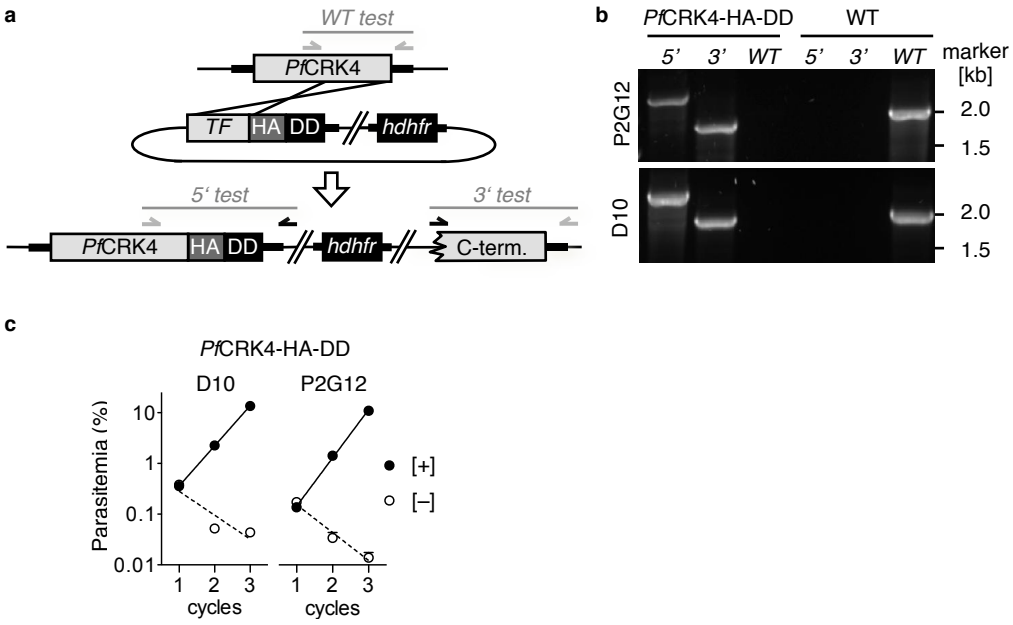
*Correspondence to:
Manoj. T. Duraisingh
Harvard T.H. Chan School of Public Health
651 Huntington Avenue, FXB, Rm. 202
Boston, MA 02115, USA
Email: mduraisi@hsph.harvard.edu

# Supplementary Figures

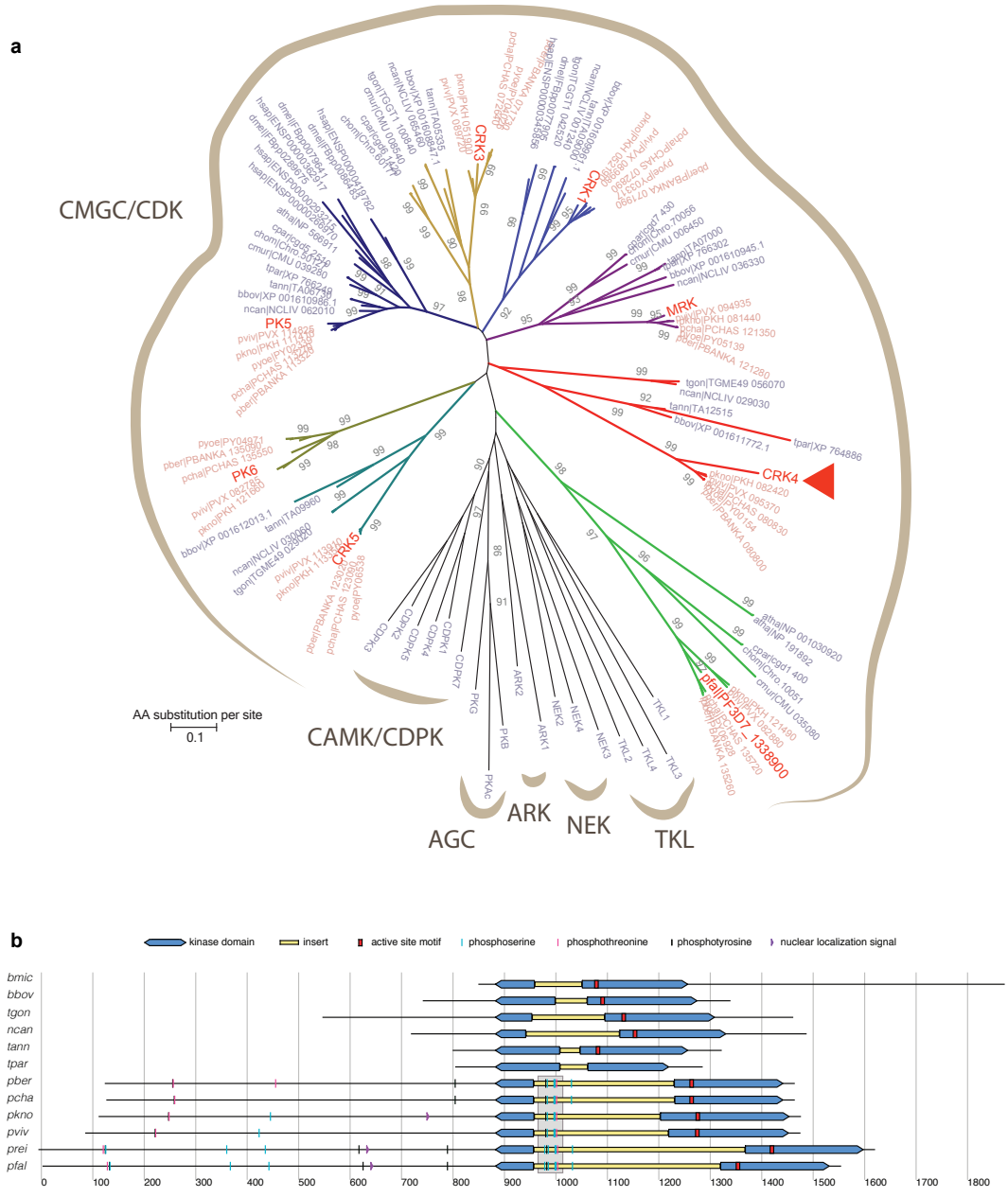



**Supplementary Figure 1 | Conditional destabilization screen of *P. falciparum* schizont-stage kinases. a,** Endogenous hemagglutinin (HA) and destabilization domain (DD) fusions of 23 schizont-stage kinases in the *P. falciparum* D10 strain. Top panel: schematic representation of the 3' replacement strategy targeting *Pf*CDPK2. The vector contained a targeting fragment (*TF*), an HA tag, a DD tag, and a selection cassette (*hdhfr*), arrows indicate test primer combinations and expected PCR products are shown as grey lines. Successful integration was confirmed by diagnostic PCR using genomic DNA from a clonal *Pf*CDPK2-HA-DD line and wild type (WT) parasites. Bottom panel: schematic representation of the 3' replacement strategy and Southern blot analysis. The respective kinase open reading frame was targeted with a 3' replacement vector. Asterisks indicate restriction sites; black bars indicate the expected DNA fragments after restriction digest; grey bars indicate hybridization sites for the Southern blot probes. Successful integration of the respective targeting plasmid was shown for 22 kinases, using genomic DNA from clonal *Pf*kinase-HA-DD lines and the D10 parental line. The respective targeting fragment was used as a probe; enzymes used for the restriction digest of genomic DNA are shown below each

blot. The expected signal for the wild type locus (WT), episomal plasmid (epi.), and integrated plasmid (int.) are indicated. **b,** Growth curves of *P. falciparum* candidate kinases HA-DD fusions cultured [+] Shield-1, filled symbols, black line; and [–] Shield-1, open symbols, dashed line; mean ± SD of triplicates. **c,** Dose response curves of D10-*Pf*CRK4-HA-DD and D10-*Pf*PKG-HA-DD parasites presented as percentage of maximal (max.) growth; shown are mean ± SD of triplicates; biol. rep., biological replicate. **d,** Western blot detecting protein levels of HA-DD-tagged kinases in parasites [+] and [–] Shield. The level of knockdown is indicated below (data from a single biological replicate, full raw Western blots are shown in Supplementary Fig. 9c, d); H3, histone H3; LDH, lactate dehydrogenase.

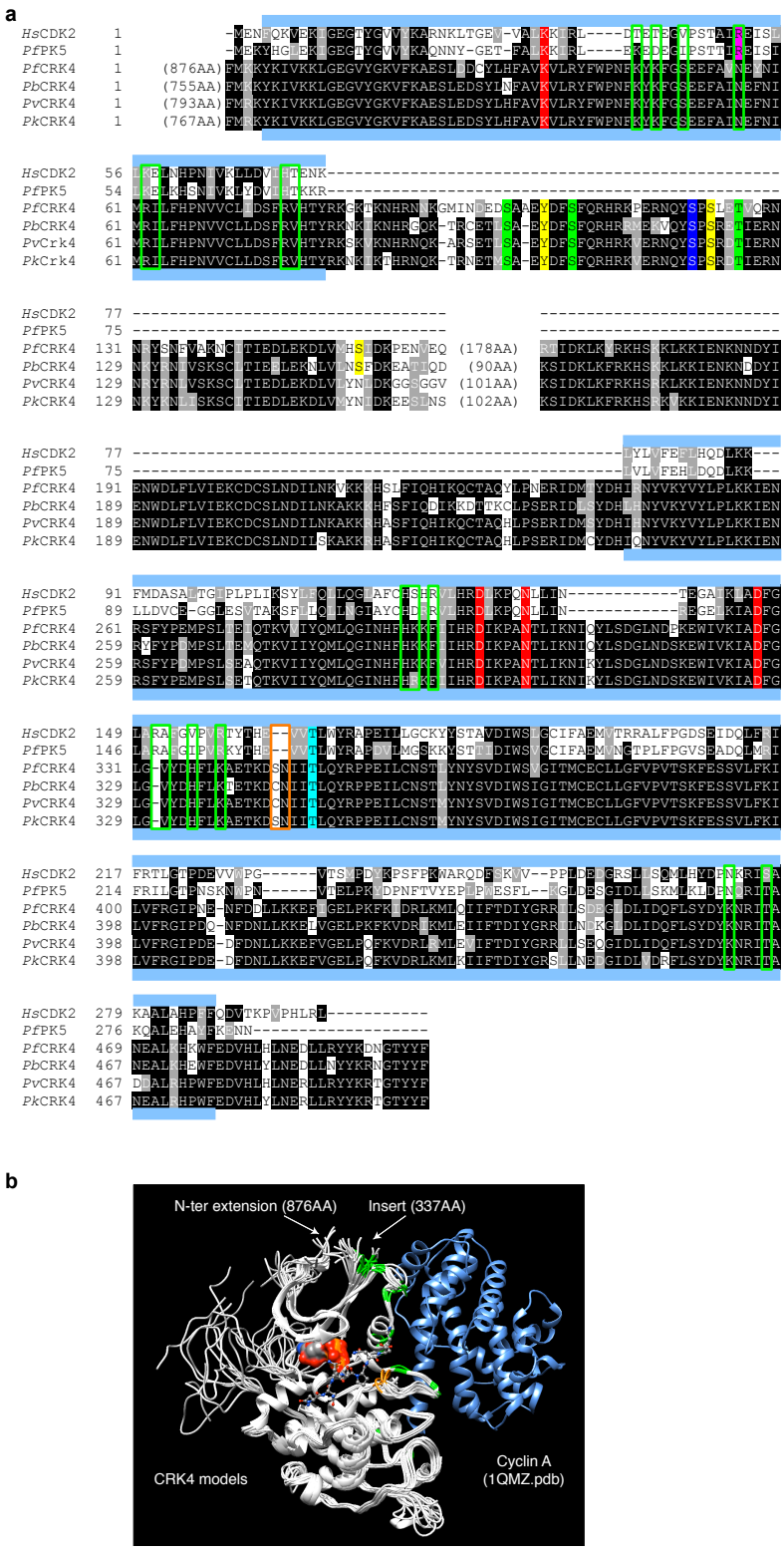



**Supplementary Figure 2 | Generation of a *Pf*CRK4-HA-DD parasite line in the P2G12 parental background. a,** Schematic representation of the 3' replacement strategy. The *Pf*CRK4 open reading frame was targeted with a 3' replacement vector containing a targeting fragment (*TF*), HA tag, a DD tag, and a selection cassette (*hdhfr*). Arrows indicate replacement-specific test primer combinations and expected PCR products are shown as grey lines. **b,** Diagnostic PCR genotyping of a clonal P2G12-*Pf*CRK4-HA-DD line obtained by limiting dilution cloning (top), compared to clonal D10-*Pf*CRK4-HA-DD parasites (bottom). Genomic DNA from wild type (WT) parasites served as controls, data representative of two replicates. **c,** Growth curves of the D10- and P2G12-*Pf*CRK4-HA-DD lines cultured [+] Shield-1, filled symbols, black line; and [–] Shield-1, open symbols, dashed line; mean ± SD of triplicates. Data are representative of at least three biological replicates.



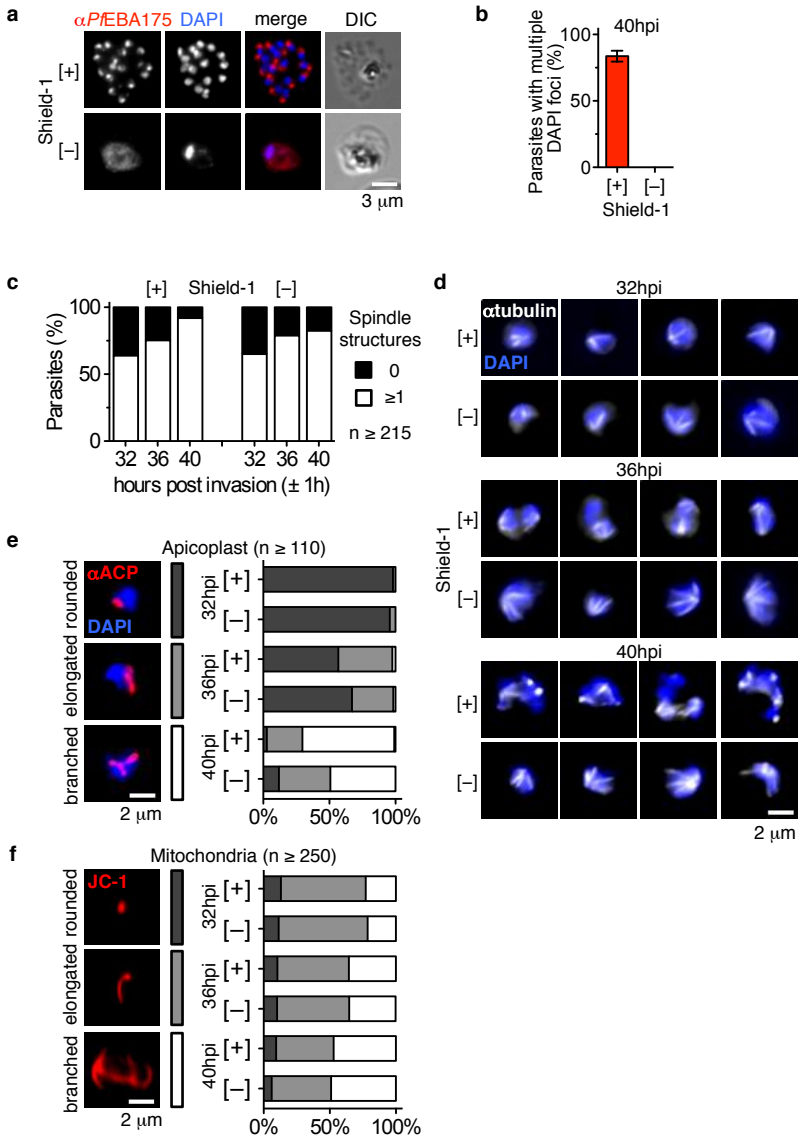

**Supplementary Figure 3 | Evolutionary relationships of apicomplexan CDK-like and reference protein kinase domains, and the domain structures of apicomplexan *Pf*CRK4 homologues. a,** Unrooted maximum likelihood tree (RAXML, 100 bootstraps) of CDK-like protein kinase domains from Apicomplexa and reference species, and non-CDK-like kinases from *Plasmodium falciparum*. *Plasmodium* CDKs, orange text; other CDKs, purple text; non-CDK *Plasmodium falciparum* kinases, black text; arrow head highlights *Pf*CRK4. **b,** Domain diagrams of apicomplexan *Pf*CRK4 homologues. The light grey-shaded area indicates a region of the *Pf*CRK4 insert with phosphorylation sites detected in both this study and in[68]. Abbreviations: CDK, cyclin dependent kinase; aory, *Aspergillus oryzae* RIB40; bbov, *Babesia bovis* T2Bo; bmic, *Babesia microti* RI; chom, *Cryptosporidium hominis* TU502;

cmur, *Cryptosporidium muris* RN66; cpar, *Cryptosporidium parvum* Iowa II; dmel, *Drosophila melanogaster*; hsap, *Homo sapiens*; lbra, *Leishmania braziliensis*; linf, *Leishmania infantum*; lmaj, *Leishmania major* strain Friedlin; lmex, *Leishmania mexicana*; ncan, *Neospora caninum*; pber, *Plasmodium berghei* ANKA; pcha, *Plasmodium chabaudi chabaudi*; pfal, *Plasmodium falciparum* 3D7; pkno, *Plasmodium knowlesi* H; pram, *Phytophthora ramorum*; prei, *Plasmodium reichenowi*; pviv, *Plasmodium vivax* SaI-1; pyoe, *Plasmodium yoelii yoelii* 17XNL; tann, *Theileria annulata* strain Ankara; tbrg, *Trypanosoma brucei gambiense*; tbru, *Trypanosoma brucei*; tcon, *Trypanosoma congolense*; tcru, *Trypanosoma cruzi* strain CL Brener; tgon, *Toxoplasma gondii*; tpar, *Theileria parva* strain Muguga; tviv, *Trypanosoma vivax*.



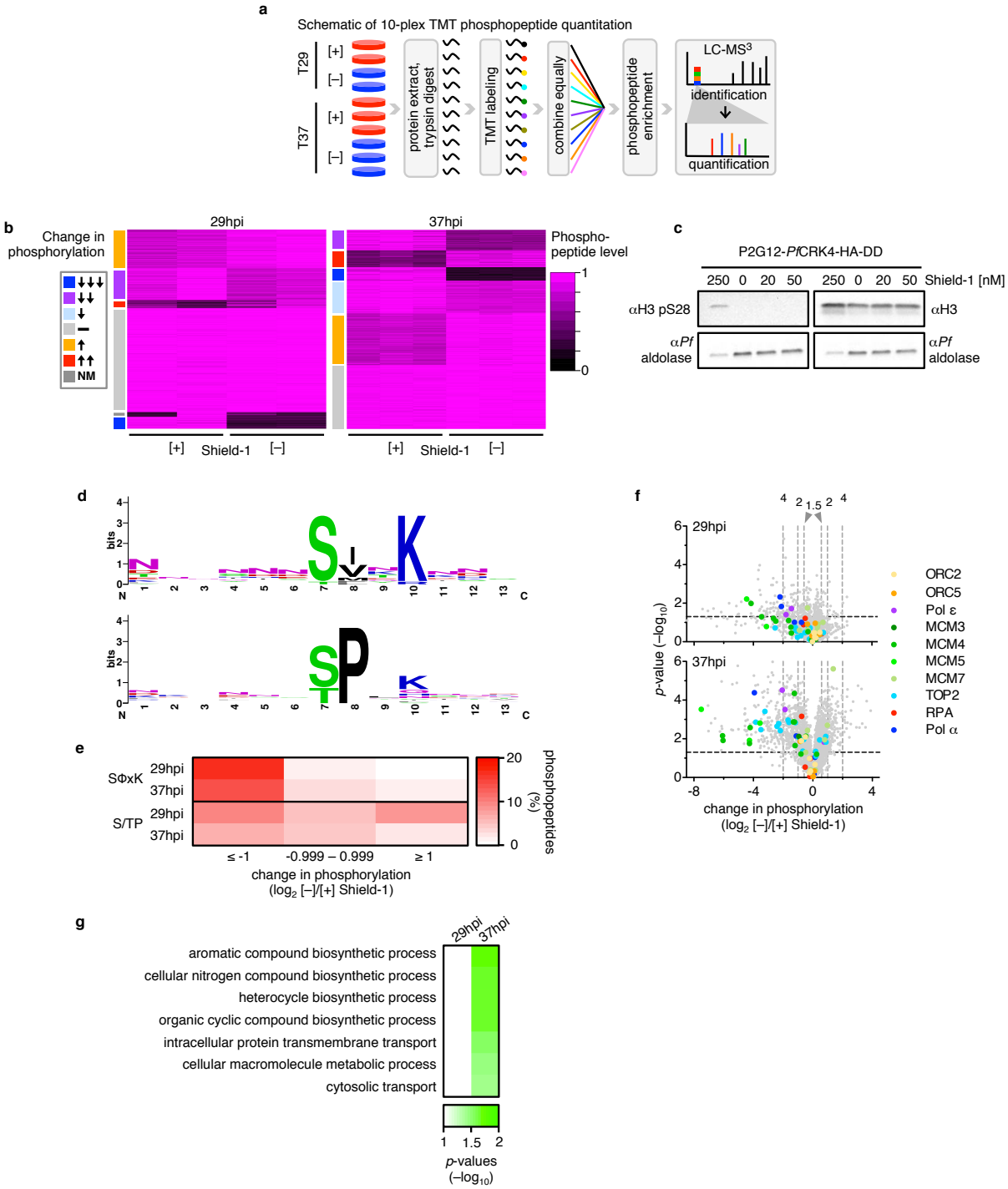

**Supplementary Figure 4 | Sequence alignment of *Plasmodium* CRK4 orthologues with *Pf*PK5 and human CDK2 and homology models of *Pf*CRK4. a,** The N-terminus extensions and variable regions of the inserts were omitted from the *Pf*CRK4 sequences. The protein kinase domain region (Pfam: Pkinase) is indicated by light blue bars above and below the alignment. Sequence text background fill indicates the following: black, identical

residues; grey, similar residues; red, essential active site residues[69]; yellow, residues phosphorylated in[68] only; green, residues phosphorylated in[68] and this study; blue, residues phosphorylated in[68], this study, and[70]; cyan, activation loop threonine phosphorylated in CDK2, but not observed to be phosphorylated in *Pf*CRK4; magenta, arginine residue that interacts with the activation loop phosphate of CDK2. Green rectangles indicate alignment with *Hs*CDK2 regions within 3 Å of Cyclin A in the *Hs*CDK2-Cyclin A-substrate peptide complex (1QMZ.pdb)[25]. Note that the *Pf*CRK4 sequence corresponding to these regions is highly divergent. The orange rectangle shows the location of the two-residue insertion in the activation loop of *Pf*CRK4. **b,** *Pf*CRK4 structures were modelled using the co-crystal structure of the human CDK2-Cyclin A-substrate peptide complex (1QMZ.pdb)[25]. Ten alternate models (white) are shown, whose variation indicates the locations of loops and extensions that do not align with CDK2. Cyclin A (blue) and ATP (solvent accessible surface representation) were mapped to the model as rigid bodies. A modelled substrate candidate (NNNSPNK) is shown as a ball-and-stick representation. The positions of the large N-terminal extension and the large insert of *Pf*CRK4 are indicated. *Pf*CRK4 regions that align with CDK2 regions within 3 Å of Cyclin A are shown in green. The two-residue insertion in the activation loop of *Pf*CRK4 is shown in orange.



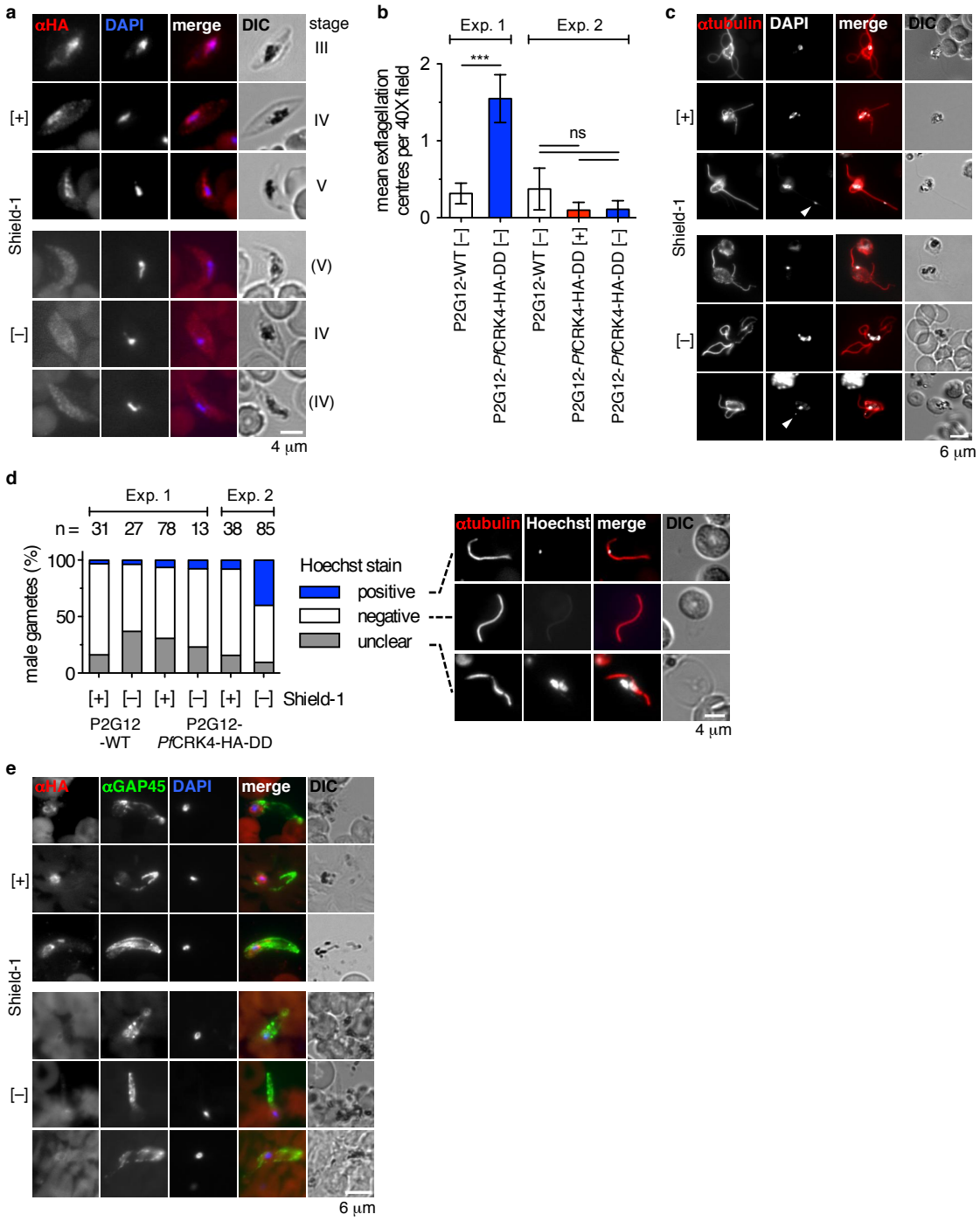

**Supplementary Figure 5 | *Pf*CRK4 is critical for trophozoite-to-schizont transition and nuclear division but not for mitochondria and apicoplasts development. a,** Immunofluorescence to detect the micronemal protein *Pf*EBA175 in D10-*Pf*CRK4-HA-DD parasites at 48 hpi cultured [+] or [–] Shield-1 (data from a single experiment). **b,** Quantification of DAPI-stained nuclei in D10-*Pf*CRK4-HA-DD parasites cultured [+] and [–] Shield-1, three biological replicates, mean ± SEM. **c,** Quantification of tubulin structures in parasites cultured [+] and [–] Shield-1. **d,** Immunofluorescence detecting tubulin structures in P2G12-*Pf*CRK4-HA-DD parasites cultured [+] and [–] Shield-1; nuclei were stained with DAPI; data representative of two biological replicates. **e,** Apicoplast development of P2G12-*Pf*CRK4-HA-DD parasites was assessed employing an anti-acyl carrier protein (ACP) antibody and DAPI nuclear stain. Left, three categories were defined based on the apicoplast morphology, *i.e.*, rounded, elongated, and branched. Right, quantification of apicoplast

morphology in parasites cultured [+] or [–] Shield-1 (data from a single experiment). **f,** Mitochondrial development of D10-*Pf*CRK4-HA-DD parasites was assessed using the cationic mitochondrial membrane potential sensor JC-1. Left, three categories were defined based on the mitochondria morphology, *i.e.*, rounded, elongated, and branched. Right, quantification of mitochondria morphology in parasites cultured [+] or [–] Shield-1 (data from a single experiment).



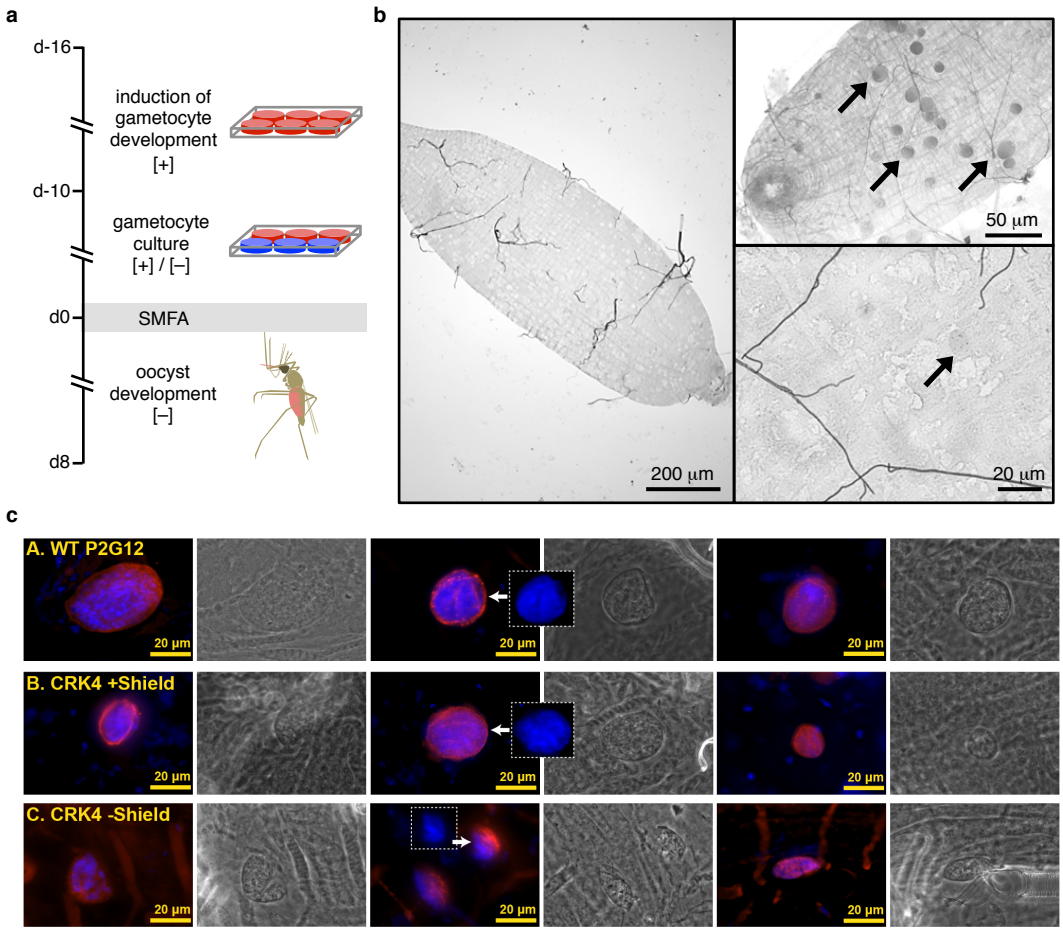

**Supplementary Figure 6 | Phosphoproteomic profiling of D10-*Pf*CRK4-HA-DD parasites. a,** Illustration of the workflow for the phosphoproteomic profiling of D10-*Pf*CRK4-HA-DD parasites. Note, 29 hpi was done in duplicate and 37 hpi in triplicate. **b,** Heat map showing changes in phosphopeptide levels upon *Pf*CRK4 depletion at 29 and 37 hpi. Phosphopeptide levels (Supplementary Data table 4) were adjusted based on scaled protein levels (Supplementary Data table 3) and normalized by the channel sum; rows with

sums below 100 were omitted from the analysis. Scaled from 0 - 1, phosphopeptides were clustered using k-means; the change in phosphorylation per cluster is indicated at the left of each heat map; NM, no match of replicates. **c,** Immunoblot confirming the reduced phosphorylation of histone H3 serine 28 (H3 pS28) upon *Pf*CRK4 depletion as seen by phosphoproteomics (Supplementary Data table 4), data from a single experiment, full raw Western blots are shown in Supplementary Fig. 9b. Note the P2G12 parental background was used in this experiment. **d,** Sequence logos of the two motifs found in phosphopeptides with ≥2-fold reduced phosphorylation and *p*-values < 0.05 in *Pf*CRK4 depleted parasites. Top, motif shows a preference for serine in the P0 position, a hydrophobic residue (ϕ) in the P+1 position and a basic residue in the P+3 position (SϕxK motif). Asparagine is common in the P+2 position of this motif, but any non-hydrophobic amino acid is acceptable. Bottom, motif prefers serine or threonine in the P0 position, requires proline in the P+1 position, and has some preference for a basic residue in the P+3 position (S/TP motif). **e,** Heat map showing the percentage of peptides (*p*-value < 0.05) with SϕxK or S/TP motifs from data collected at 29 and 37 hpi, phosphopeptides containing the motifs are annotated in Supplementary Data table 4. **f,** Differential phosphorylation in D10-*Pf*CRK4-HA-DD parasites [–] Shield-1 relative to [+] Shield-1; highlighted are phosphopeptides of *P. falciparum* homologues of *S. cerevisiae* factors (or subunits thereof) required for origin of replication activation *in vitro*[26]; black dashed line indicates a *p*-value of 0.05; grey dashed lines indicate a 4-fold, 2-fold, and 1.5-fold change in phosphorylation, respectively; Pol, DNA polymerase; MCM, mini chromosome maintenance complex; TOP2, topoisomerase 2; RPA, replication protein A; ORC, origin recognition complex; *p*-values by Student`s t-test. **g,** GO term enrichment analysis of proteins with increased phosphorylation (≥2-fold and $p<0.05$) in *Pf*CRK4-depleted parasites. Note, at 29 hpi no GO term reached statistical significance; *p*-values for GO term enrichment by Modified Fisher's Exact Test.

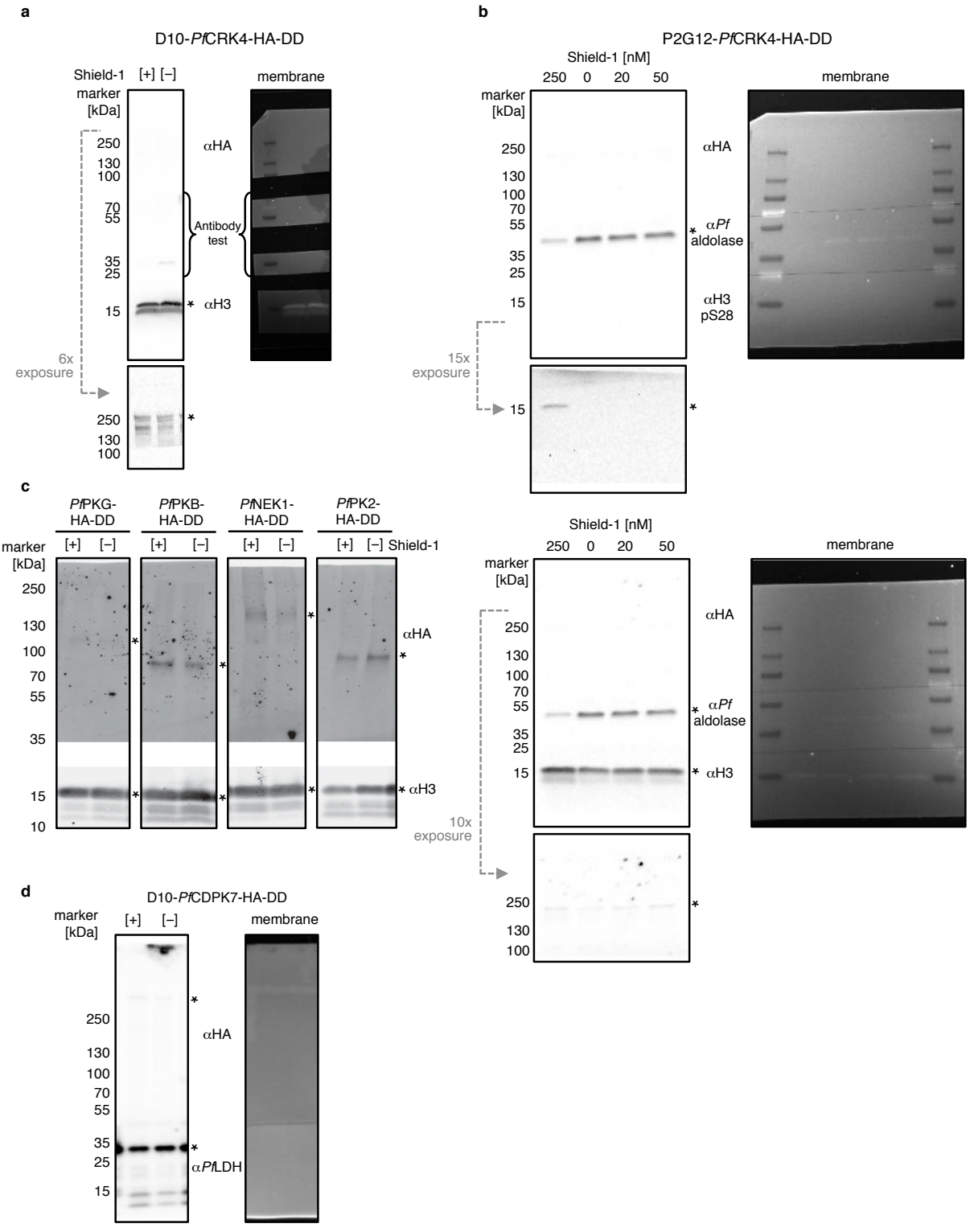



**Supplementary Figure 7 | *Pf*CRK4 depletion does not impair microgamete exflagellation. a,** Immunofluorescence detecting *Pf*CRK4-HA-DD (P2G12 parent) in mid and late stage gametocytes when cultured [+] and [–] Shield-1; representative of two biological replicates; in parenthesis, staging with less confidence. **b,** Quantification of exflagellation centers of P2G12 wild type (P2G12-WT) and P2G12-*Pf*CRK4-HA-DD gametocytes cultured [+] and [–] Shield-1 by microscopy; mean ± SEM of three replicate measurements, each quantifying exflagellation in six 40X fields; ***, $p \leq 0.001$; ns, not significant (two-tailed paired t test). **c,** Immunofluorescence detects exflagellating male

gametocytes of P2G12-*Pf*CRK4-HA-DD parasites maintained [+] and [–] Shield-1, detecting parasites with single and multiple nuclei; nuclei were stained with DAPI; arrowheads, nuclei of newly formed male gametes; representative of two biological replicates. **d,** Quantification of Hoechst 33342-positive male gametes of wild type (WT) and *Pf*CRK4-HA-DD parasites (P2G12 line). Male gametes were identified by anti-tubulin immunofluorescence staining; representative parasite images of the three categories, *i.e.*, positive, negative, and unclear, are shown on the right; data from two independent experiments (Exp.) are shown. **e,** Immunofluorescence detection of *Pf*CRK4-HA-DD in ookinetes from the blood bolus, which was isolated from mid guts of female *An. gambiae* mosquitoes approximately 24 hours post-feeding. Mosquitos were fed with late stage *Pf*CRK4-HA-DD gametocyte cultures, maintained [+] and [–] Shield-1 from day 6 post-induction (data from a single experiment).




**Supplementary Figure 8 | *Pf*CRK4-depleted parasites poorly infect *An. gambiae* mosquitoes. a,** Cartoon illustrating the typical timeline of a Standard Membrane Feeding Assay (SMFA) [+] or [–] Shield-1. Note, once fed to mosquitoes parasites were [–] Shield-1. **b,** Bright-field images of Mercurochrome stained oocysts at day 9 post infectious blood meal. Left, uninfected midgut; top right, midgut from an infection with P2G12-*Pf*CRK4-HA-DD gametocytes cultured [+] Shield-1; bottom right, blow-up from a midgut infected with P2G12-*Pf*CRK4-HA-DD gametocytes cultured [–] Shield-1; arrows mark parasites; data from a single experiment. **c,** Immunofluorescence detects the circumsporozoite protein (red) on midgut oocysts, nuclei were stained using DAPI (blue). Top panel, P2G12 WT; middle panel, P2G12-*Pf*CRK4-HA-DD [+] Shield-1 during gametocytogenesis, and bottom panel [–] Shield-1 during gametocytogenesis.




**Supplementary Figure 9 | Original Western blot data. a,** Raw Western blot data related to Fig. 1d (note, the middle sections of the membrane were used to test antibodies). **b,** Raw Western blot data related to Fig. 1d and Supplementary Fig. 7c. **c, d,** Raw Western blot data related to Supplementary Fig. 1d. Asterisks indicate bands relevant to main or Supplementary figures.

## Supplementary Table

**Supplementary Table 1 | Gametocytemia and sex ratio quantification of *P. falciparum* P2G12 wild type (WT) and P2G12-*Pf*CRK4-HA-DD parasites ([+] and [–] Shield-1) at day 16 post-induction of gametocytogenesis.**

| | P2G12 WT | P2G12-*Pf*CRK4-HA-DD | |
|---|---|---|---|
| | | [+] | [–] |
| Gametocytemia | 13/1058 (1.22%) | 8/1104 (0.73%) | 13/1016 (1.27%) |
| Males | 57/142 (40%) | 99/261 (38%) | 54/166 (33%) |
| Females | 85/142 (60%) | 162/261 (62%) | 112/166 (67%) |
| Male:Female Ratio | 0.67 | 0.61 | 0.48 |

## Supplementary References


68. Pease, B. N. *et al.* Global Analysis of Protein Expression and Phosphorylation of Three Stages of Plasmodium falciparum Intraerythrocytic Development. *J. Proteome Res.* **12,** 4028–4045 (2013).

69. Hanks, S. K. & Hunter, T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9,** 576–596 (1995).

70. Lasonder, E. *et al.* The Plasmodium falciparum schizont phosphoproteome reveals extensive phosphatidylinositol and cAMP-protein kinase A signaling. *J. Proteome Res.* **11,** 5323–5337 (2012).