Title

Identification of Antimalarial Compounds that Inhibit Apicomplexan AP2 Proteins in the Human Malaria Parasite *Plasmodium falciparum*

Word Count: 9239

Authors

Timothy Russell[1,2,3,4], Erandi K. De Silva[6], Valerie Crowley[1], Kathryn Shaw-Saliba[7], Namita Dube[1], Gabrielle Josling[1,3,4], Charisse Flerida A. Pasaje[8], Irene Kouskoumvekaki[9], Gianni Panagiotou[10,11], Jacquin C. Niles[8], Marcelo Jacobs-Lorena[7], C. Denise Okafor[1,5], Francisco-Javier Gamo[12], Manuel Llinás[1,2,3,4,5]

1. Department of Biochemistry and Molecular Biology, Pennsylvania State University.
2. Huck Institutes Center for Eukaryotic Gene Regulation (CEGR), Pennsylvania State University.
3. Huck Institutes Center for Malaria Research (CMaR), Pennsylvania State University.
4. Huck Institutes Center for Infectious Disease Dynamics, Pennsylvania State University.
5. Department of Chemistry, Pennsylvania State University.
6. Lewis-Singler Institute for Integrative Genomics, Princeton University.
7. Department of Molecular Biology and Immunology, Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health.
8. Department of Biological Engineering, Massachusetts Institute of Technology.
9. Department of Systems Biology, Technical University of Denmark, Kgs. Lyngby, Denmark.
10. Systems Biology and Bioinformatics, Leibniz Institute for Natural Products Research and Infection Biology, Hans Knöll Institute.
11. Department of Medicine, the University of Hong Kong, Hong Kong SAR, China
12. Global Health Medicines, GlaxoSmithKline, Tres Cantos, Spain.

Abstract

Plasmodium parasites are reliant on the Apicomplexan AP2 (ApiAP2) transcription factor family to regulate gene expression programs. AP2 DNA binding domains have no homologs in the human or mosquito host genomes, making them potential antimalarial drug targets. Using an *in-silico* screen to dock thousands of small molecules into the crystal structure of the AP2-EXP (Pf3D7_1466400) AP2 domain (PDB:3IGM), we identified compounds that interact with this domain. Four compounds were found to compete for DNA binding with AP2-EXP and at least one additional ApiAP2 protein. Our top ApiAP2 competitor compound perturbs the transcriptome of *P. falciparum* trophozoites and results in a decrease in abundance of $\log_2$ fold change > 2 for 50% (46/93) of AP2-EXP target genes. Additionally, two ApiAP2 competitor compounds have multi-stage anti-*Plasmodium* activity against blood and mosquito stage parasites. In summary, we describe a novel set of antimalarial compounds that are targeted against the ApiAP2 family of proteins. These compounds may be used for future chemical genetic interrogation of ApiAP2 proteins or serve as starting points for a new class of antimalarial therapeutics.

Author Summary

Plasmodium parasites are the causative agent of malaria, which resulted in over 600,000 deaths in 2021. Due to resistance arising for every antimalarial therapeutic deployed to date, new drug targets and druggable pathways must be explored. To address this concern, we used a molecular docking screen to predict competitors of DNA binding by the parasite specific family of Apicomplexan AP2 (ApiAP2) transcription factor proteins for testing *in vitro* and *in vivo*. We find that ApiAP2 competing

53 compounds have antimalarial activity consistent with the disruption of gene regulation.
54 This work will further our understanding of both the biological role and targetability of
55 parasite transcriptional regulation.
56

57 **Introduction**

58 Malaria is a disease caused by intracellular parasites from the genus *Plasmodium* that
59 represents a significant health and economic burden worldwide(1). The most virulent of
60 the human infectious malaria parasites is *Plasmodium falciparum*, which caused more
61 than 200 million cases of malaria and resulted in over 600 thousand deaths in 2021(1).
62 Resistance has been reported for every antimalarial therapeutic deployed to date,
63 necessitating the need for malaria drugs that target new parasite processes(2) . To
64 successfully proliferate, *P. falciparum* parasites must develop through a complex
65 lifecycle that includes intracellular and extracellular stages in both the human and
66 *Anopheles* mosquito hosts(3). The clinical symptoms of malaria are caused by the
67 intraerythrocytic development cycle (IDC), a 48-hour cyclic asexual proliferation that
68 results in the destruction of red blood cells (RBCs). During the asexual IDC, malaria
69 parasites progress through three morphological stages referred to as ring, trophozoite,
70 and schizont. Parasites transmit from human to mosquito following differentiation and
71 maturation into sexual stage gametocytes. Once ingested by the mosquito, gametes
72 sexually reproduce and ultimately develop into sporozoites, which can be transmitted
73 back to the human host to initiate a liver stage infection which precedes the asexual
74 blood stages(3).
75

76 Up to 80% of *P. falciparum* protein coding transcripts are developmentally regulated
77 during the IDC(4–8) as part of a 'just in time' cascade(5, 8, 9). This is predicted to be
78 principally driven by the 27 Apicomplexan Apetala AP2 (ApiAP2) proteins(10–12), which
79 are the major family of sequence specific transcription factors encoded in the
80 *Plasmodium* genome. ApiAP2 proteins contain one to three AP2 DNA binding domains
81 that are analogous to the plant APETALA2/Ethylene Responsive Factor (AP2/ERF)
82 domain and therefore have no homologs in the human or mosquito genome(13, 14).
83 Over half of the ApiAP2 proteins are predicted to be essential for the IDC in *P.*
84 *falciparum*(15–17). In-depth studies of ApiAP2 proteins to date have demonstrated roles
85 in transcriptional activation or repression affecting diverse processes including
86 invasion(18, 19), heterochromatin maintenance(17, 20–23), sexual and mosquito stage
87 differentiation(16, 24–34), and heat stress tolerance(35). Despite these properties which
88 make ApiAP2 proteins potential drug targets, no efforts to date have focused on
89 targeting the sequence specific DNA binding ApiAP2 transcription factors.
90
91 In this study we conducted an *in silico* chemical screen against the crystal structure of
92 the AP2-EXP AP2 domain(36) to predict competitors of DNA binding activity. AP2-EXP
93 is predicted to be essential for the *P. falciparum* asexual blood stage,(16, 17, 21) and its
94 orthologue in the rodent malaria parasite *P. berghei* (PbAP2-Sp) is a master regulator of
95 sporogony(32). Several compounds identified in the screen can compete for DNA
96 binding activity with the AP2-EXP AP2 domain *in vitro*. We also find that modifications to
97 the substitution state of an AP2 domain competitor compound can modulate its ability to
98 compete DNA binding. Our lead compound kills *P. falciparum* parasites in concordance

99 with the stage specific expression of AP2-EXP. We measured AP2-EXP genomic
100 occupancy by Chromatin immunoprecipitation followed by high throughput sequencing
101 (ChIP-seq) and found high correlation between dysregulated mRNA transcripts in the
102 presence of our lead compound and AP2-EXP gene targets, corroborating the
103 hypothesis that AP2-EXP is inhibited *in vivo*. Finally, two AP2 domain competitor
104 compounds are active against *P. berghei* development in the mosquito stage. Overall,
105 our results demonstrate the potential to chemically compete the DNA binding activity of
106 ApiAP2 proteins.
107

108 **Results**

109 ***In silico* Prediction of AP2-EXP Competitors**

110 To identify competitors of DNA binding by an ApiAP2 protein, we used AP2-EXP
111 (PF3D7_1466400), which is the only ApiAP2 protein whose AP2 domain is structurally
112 characterized (PDB accession: 3IGM)(36). *In silico* molecular docking was run on
113 AUTODOCK(37) using over ten thousand small molecules from both the Tres Cantos
114 Antimalarial Set (TCAMS)(38) and the Drug Bank(39) against AP2-EXP. Docking hits
115 were prioritized based on the free energy of interaction and proximity to amino acids
116 necessary for DNA contact **(Fig S1A)**. In some cases, the compound that generated a
117 high docking score was not available for purchase, so alternative choices with >0.9
118 Tanimoto similarity score were identified, computationally docked, and used for further
119 evaluation **(All compounds listed in Figure S1B).** These docking simulations resulted
120 in a final list of nine top scoring predicted competitors of DNA binding by AP2-EXP.
121 Each compound was assigned an alphabetical identifier A-I **(Fig 1A, Fig S2A, B)** for

122 use in this study. We determined that all nine compounds kill asexual *P. falciparum*
123 parasites at micromolar concentrations (range 11-170μM) in a growth inhibition assay
124 **(Fig 1A, Fig S3)**.
125
126 **Four Compounds Have Activity Against AP2-EXP *in vitro***
127 The nine predicted competitors of DNA binding by AP2-EXP were tested in a
128 competitive electrophoretic mobility shift assay (EMSA). The minimum mass of AP2-
129 EXP required to visualize DNA binding **(Fig S4A)** was mixed with each compound and
130 incubated prior to adding the DNA oligonucleotide **(Table S1)**. Four of the nine
131 compounds (A, B, C and I) **(Fig 1B)** were able to effectively compete for DNA binding
132 by AP2-EXP *in vitro*. Three of the four AP2 domain competitors (Compounds A, B and
133 C) have a benzoxazole core moiety, while Compound I is made up of planar rings **(Fig**
134 **1C)**.
135
136 To assess whether DNA binding competition was specific to the AP2-EXP AP2 domain,
137 we repeated competitive EMSAs with three different purified *P. falciparum* AP2
138 domains: AP2-I Domain 3 (AP2-I D3)(12, 18) **(Fig S4B)**, AP2-HS Domain 1 (AP2-HS
139 D1)(12, 35) **(Fig S4C)**, and PfSIP2 Domain 1 (PfSIP2 D1)(20) **(Fig S4D)**, again using
140 the minimum amount of protein needed to visualize interaction with the DNA
141 oligonucleotide **(Table S1)**. We found that Compounds A, B, C, and I can compete DNA
142 binding by AP2-I **(Fig S5A)**, Compounds B and I compete DNA binding by AP2-HS D1
143 **(Fig S5B)** and Compounds B, C and I compete DNA binding by PfSIP2 D1 **(Fig S5C)**.
144

145 As a control for compound specificity, we performed competitive EMSAs using two non-
146 *Plasmodium* AP2 domain proteins: the *Arabidopsis thaliana* Ethylene Response Factor
147 1(40) (AtERF1) **(Fig S4E)** and the human High Mobility Group Box (HMGB) domain
148 protein SOX2 **(Fig S4F)**(41). Only Compound I was able to compete AtERF1 **(Fig S6A)**,
149 while SOX2 was competed by Compound A only **(Fig S6B)**. Therefore, Compounds B
150 and C demonstrate specificity for Apicomplexan AP2 domains *in vitro*. (EMSA Results
151 Summarized in **Table S2**).
152
153 To eliminate the possibility that ApiAP2 competing compounds interfere non-specifically
154 with DNA binding proteins by intercalating with DNA, we conducted an ethidium
155 bromide exclusion assay(42). Compared to the positive control DRAQ5(42), only
156 Compound I was able to effectively displace ethidium bromide from DNA **(Fig S7)**.
157 Based on these results, we focused on Compounds B and C for further study.
158
159 **Structural Analogs of Compound B Differ in DNA Binding Competition Activity**
160 We tested the four closest available analogs to Compound B from the TCAMS based on
161 similarity score(38) (Designated as Compounds B-1, B-2, B-3 and B-4) for competition
162 against AP2-EXP. Compounds B-1 and B-4 compete DNA binding by AP2-EXP **(Fig**
163 **2A)** and AP2-I D3 **(Fig S8A)**, while Compounds B-2 and B-3 do not compete DNA
164 binding by either AP2-EXP or AP2-I D3 **(Fig 2A, S8A)**. Compound B-1 is the least
165 effective competitor relative to Compounds B and B-4 **(Fig 2B)**. Non-AP2 domain
166 competitor Compounds B-2 and B-3 each have halogen atoms substituted on the
167 benzene ring. Conversely, Compounds B, B-1, and B-4, have methyl or ethyl groups
168 substituted onto the benzene ring and can compete DNA binding **(Fig 2C)**. Compound

169 B-1 has both a methyl group and a chlorine atom substitution. If the halogen atom
170 substitution decreases DNA binding competition, this mix of substitutions is consistent
171 with its lower efficacy of DNA binding competition relative to Compounds B and B-4**.**
172 Compounds B, B-1 and B-4 have no difference in DNA intercalation ability measured by
173 ethidium bromide exclusion **(Fig S8B)**. A molecular dynamics simulation of AP2-EXP
174 docked with Compound B and each analogue predicts that Compounds B, B-1, B-2, and
175 B-4 dock stably with AP2-EXP, while non-AP2 competitor Compound B-3 moves away
176 from the protein **(Supplemental Movies 1-5)**. In summary, substituting bulky,
177 electronegative groups onto Compound B can abolish competition for DNA binding with
178 AP2-EXP.
179
180 **Compound C Activity Coincides With AP2-EXP Expression**
181 To determine the stage in the IDC at which AP2-EXP is expressed, we used Selection
182 Linked Integration (SLI)(43) to create a parasite line with endogenously GFP tagged
183 AP2-EXP (AP2-EXP::GFP) **(Fig S9A)**. We observed maximum abundance of AP2-
184 EXP::GFP in 30 hpi trophozoites by both western blot and live fluorescence microscopy
185 **(Fig 3A, Fig S10, S11)**. Maximum protein abundance of AP2-EXP in mid-trophozoites
186 was independently confirmed in a parasite line expressing AP2-EXP tagged with 2XHA
187 and the TetR:DOZI mRNA repression aptamer (AP2-EXP::HA) **(Fig S9B, Fig S12)**.
188 AP2-EXP was not detected in maturing Stage III gametocytes **(Fig 3A)**.
189
190 To assess whether Compounds B and C are likely to share a common ApiAP2 target,
191 we determined their parasite killing phenotypes against *P. falciparum* during the IDC.

192 Compound H was used as a control since it does not compete AP2-EXP **(Fig 1B)**. Each
193 compound was spiked (40μM) into synchronous asexual blood stage *P. falciparum*
194 parasites at 24 hpi and parasite morphology was determined at 24 and 48 hours post
195 spike-in **(Fig S13A)**. At 24 hours post spike-in, parasites in the presence of either
196 Compounds B or C had failed to progress past the mid-trophozoite stage and did not
197 morphologically progress any further by 48 hours post spike-in **(Fig 3B)**. In contrast,
198 parasites spiked with Compound H progressed through one IDC before failing to
199 reinvade in the next cycle **(Fig 3B)**. As expected, each compound exposure prevented
200 further parasite growth within 48 hours of spike-in **(Fig S13B)**.
201
202 Due to its commercial availability, we continued with Compound C as the lead
203 compound for phenotypic and molecular characterization. To determine the precise
204 timing of Compound C antimalarial activity, synchronous parasites were monitored
205 throughout the 48-hour IDC in the presence of either 40μM Compound C or DMSO
206 control after spiking in Compound C at 6 hpi. Compound C spiked parasites failed to
207 progress beyond the mid-trophozoite stage at 24-30 hours post invasion **(Fig S14)**. To
208 further narrow down the timing of the anti-*Plasmodium* action of Compound C, 40μM
209 Compound C was washed out of a parasite culture at 8-hour intervals beginning at 14
210 hpi. Growth was rescued, albeit with diminishing effectiveness, until 30 hpi **(Fig 3C)**.
211 Spiking 40μM Compound C into culture at 8-hour intervals beginning at 14 hpi
212 effectively killed the parasite culture at each time point except for 46 hpi. **(Fig 3D)**.
213 Taken together, these results suggest that Compound C maximally inhibits the IDC
214 progression of *Plasmodium* parasites between 30-46 hpi. These results show that the

215 maximum abundance of AP2-EXP at 30 hpi coincides with the timing of action for
216 Compound C.
217
218 **Compound C Disrupts the *P. falciparum* Transcriptome with Stage Specificity**
219 Since AP2-EXP is predicted to act as a transcription factor(17, 21) we hypothesized that
220 changes in parasite transcript abundance should occur when AP2-EXP DNA binding is
221 competed by Compound C. We cultured parasites in 12μM (0.66xIC_{50}) Compound C or
222 DMSO vehicle and measured RNA abundance at seven time points during the asexual
223 blood stage. As a quality control for normal IDC progression, we found high correlation
224 for a set of "control" genes(44, 45) with an established 48-hour periodicity in the IDC
225 between our DMSO and Compound C samples **(Fig S15A-D)**. Since 12μM Compound
226 C kills fewer than 50% of parasites within one IDC **(Fig S16A, B)**, this demonstrates
227 that there is not an overall perturbation to the IDC transcriptome upon Compound C
228 incubation. The Spearman Correlation between total transcriptome samples deviated
229 most significantly at 30 hpi, (Corr. = -0.16) **(Fig 4A)** as expected from the compound
230 activity measured **(Fig 3C, D, Gene Expression values in Table S3)**.
231
232 Next, we quantified differences in transcript abundance that are unique to the
233 progression from 24 to 30 hpi for Compound C vs. DMSO vehicle control populations
234 using Linear Models for Microarray Data (LIMMA(46)) **(Differentially Expressed genes**
235 **in Table S3)**. Overall, 463 RNA transcripts are decreased in abundance by log_2 fold
236 change >2 between 24-30 hpi in the presence of Compound C. Gene ontology (GO)
237 analysis(47) **(Full GO Terms in Table S3)** revealed an enrichment for genes that
238 encode proteins important for the parasite to invade or modify red blood cells among

239 decreased transcripts (e.g., RON3, GAP45, RhopH2, RhopH3, MSP1, MSP6). We
240 independently determined differential transcript abundance over the entire time course
241 using the RNA time course specific software Rnits(48) and found high overlap with our
242 LIMMA analysis **(Overlaps in Table S3)**. Therefore, the maximal perturbation of the
243 transcriptome **(Fig 4A)**, action of Compound C **(Fig 3C, D)**, and maximal abundance of
244 AP2-EXP **(Fig 3A)** all occur at roughly 30 hpi in the IDC.
245

246 **AP2-EXP Gene Targets Correlate with Differentially Abundant Transcripts**

247 We used ChIP-seq to determine the genome wide binding occupancy of AP2-EXP. A
248 total of three samples from highly synchronous trophozoites were collected at 30 hpi
249 using both the AP2-EXP::GFP (2 replicates) and AP2-EXP::HA (1 replicate) parasite
250 lines **(Peaks in Table S4)**. As a quality control, we determined that ChIP recovers intact
251 AP2-EXP **(Fig S17A)**, and a co-immunoprecipitation blot confirmed that AP2-EXP
252 interacts with Histone H3 in the nucleus **(Fig S17B)**. In aggregate, AP2-EXP binds 240
253 genomic loci, corresponding to 101 total target genes **(Genes in Table S4)**. Enrichment
254 of AP2-EXP was well conserved between the GFP and HA tagged parasite lines, as
255 indicated by nearly identical coverage in a metagene plot of AP2-EXP target genes **(Fig
256 4B)**. A 'no epitope control' ChIP done on a wild type Pf3D7 parasite line with the anti-
257 GFP antibody detected only 2 peaks, neither of which overlapped with those of AP2-
258 EXP::GFP or AP2-EXP::HA **(Figure 4B, Table S4)**. AP2-EXP peaks were highly
259 enriched for the known AP2-EXP and PbAP2-Sp DNA sequence motif CATGCA(17, 49)
260 **(Fig 4B, Fig S18A-C)**.
261

262 Out of the 463 transcripts decreased in abundance by log_2 fold change > 2 in the
263 presence of Compound C, 46 are AP2-EXP targets **(Table S3)**. This represents 50%
264 (46/93) of AP2-EXP target genes detected in the RNA time course **(Fig 4C, Figure**
265 **S19)**. In general, the transcripts with the greatest decrease in abundance are AP2-EXP
266 targets. We found that 7/14 total transcripts decreased in abundance by log_2 fold
267 change > 5, and 16/33 decreased in abundance by log_2 fold change > 4, are AP2-EXP
268 targets **(Fig 4C, Table S3)**. AP2-EXP targets that are dysregulated at 24-30 hpi have
269 functions related to red blood cell invasion and host remodeling **(Fig 4D, Table S3)**. To
270 assess whether AP2-EXP DNA binding is impacted by Compound C, parasite cultures
271 were spiked with either 40μM Compound C or DMSO vehicle control at 30 hpi and AP2-
272 EXP occupancy was measured by ChIP-quantitative PCR (ChIP-qPCR). AP2-EXP DNA
273 binding decreases at five specific peaks of occupancy (GAP45, SIP2, RON3, RALP,
274 and AMAI) in the presence of Compound C **(Fig S20)**.
275
276 We then compared AP2-EXP genomic occupancy with several published datasets in
277 order to further evaluate its function as a sequence specific transcription factor. AP2-
278 EXP occupancy is correlated with a nucleosome depleted region **(Fig S21A, B)**(50) and
279 activating chromatin marks(51, 52) **(Fig S22A, B)**, and is proximal to target gene
280 transcription start sites (TSS)(53) **(Fig S23)**. The majority of target genes increase in
281 abundance starting at 32 hpi **(Fig S24)**(54). AP2-EXP target gene function is enriched
282 for invasion and red blood cell modification, **(Table S3)** as was found for the genes that
283 decrease in abundance in the presence of Compound C **(Fig 4D, Table S3)**. In

284 aggregate, these findings suggest that Compound C inhibits the function of AP2-EXP as
285 a sequence specific transcription factor.
286
287 **ApiAP2 Competitor Compounds B and C are Active Against Mosquito Stage**
288 ***Plasmodium* Parasites**
289 Elucidation of the role of essential *P. falciparum* ApiAP2 proteins in the IDC is
290 challenging due to the inability to knock out asexual blood stage essential genes. We
291 did not recover transgenic parasites after attempting to completely disrupt the coding
292 sequence of AP2-EXP using the targeted gene disruption (pSLI-TGD) system(43) in
293 three independent attempts, suggesting that AP2-EXP is essential for the IDC (not
294 shown). We then attempted to disrupt AP2-EXP protein abundance with the conditional
295 knockdown approaches Knock Sideways(43), TetR:DOZI mRNA repression(55), and
296 *glms* ribozyme mediated cleavage(56). Each of these genetic systems failed to mediate
297 protein knockdown or mislocalization **(Fig S9A, B, Fig S25, Fig S26)**. To account for
298 this limitation, we tested the ApiAP2 competitor Compounds B and C against rodent
299 infectious *Plasmodium berghei* parasites due to the well-characterized genetic
300 phenotype of the AP2-EXP orthologue PbAP2-Sp (89% amino acid identity to the AP2-
301 EXP AP2 domain)(32, 49) as a master regulator of sporogenesis.
302
303 We injected the measured IC_{50} **(Figure 1A)** of Compounds B, C, or F into the midgut of
304 *Anopheles* mosquitoes infected with *P. berghei* parasites to determine their effect on
305 *Plasmodium* mosquito stage development **(Fig S27)**. Compound F **(Fig 1B)** was used
306 as a non AP2 domain competing control. Mosquitoes injected with Compounds F or C

307 developed comparable oocyst numbers to the DMSO control **(Fig 6A)** while Compound
308 B prevented the development of midgut oocysts entirely **(Fig 6A)**. Compound C treated
309 mosquitoes developed 10-fold fewer midgut sporozoites than the control **(Fig 6B)**.
310 Therefore, consistent with their hypothesized *in vitro* activity against ApiAP2 proteins,
311 Compounds B and C are strong inhibitors of *P. berghei* sporozoite development *in vivo*
312 **(**parasite counts in **Table S5)**.
313

314 **Discussion**

315 ApiAP2 transcription factors are unique to Apicomplexan parasites due to their plant-like
316 domain architecture(13, 14) and many are essential to asexual blood stage
317 development, making them valuable as potential drug targets. It is therefore desirable to
318 discover chemical scaffolds that can target essential ApiAP2 proteins. Furthermore,
319 chemical inhibition of ApiAP2 proteins may be used in future studies to uncover details
320 about their true biological functions during parasite development. Using a combination
321 of *in silico*, biochemical, and genetic approaches, we identified a series of compounds
322 (Compounds A, B, C, and I) that compete DNA binding by ApiAP2 proteins. Compounds
323 A, B, and C all have the same core benzoxazole moiety but vary in their measured IC_{50},
324 indicating that changes which do not affect ApiAP2 DNA binding competition can alter
325 the potency against parasites. Furthermore, Compounds B and C can inhibit *P. berghei*
326 parasite development without killing mosquitoes, suggesting that ApiAP2 competitor
327 compounds may also be tolerated by the host and may serve as transmission blocking
328 agents(57).
329

330 Compounds B and C have no activity against the plant encoded AtERF1 or human
331 encoded SOX2, supporting their selectivity for Apicomplexan AP2 domains. Although
332 the mode of DNA binding is shared between AP2-EXP and AtERF1(40), the specific
333 amino acids that contact DNA are not strictly conserved(36). Lack of activity against
334 AtERF1 may be the result of filtering molecular docking hits based on proximity to DNA
335 base contacting amino acids. The DNA binding competition capacity of Compound B
336 analogs varies *in vitro* based on substitutions to the benzene ring. In a molecular
337 dynamics simulation, Compound B remains stably associated with AP2-EXP, while the
338 non ApiAP2 competitor Compound B-3 does not. This suggests that a steric clash
339 between AP2-EXP and the bromine atom on Compound B-3 is responsible for its lack of
340 DNA binding competition. By extension, this may explain the lower efficacy of DNA
341 binding competition observed for Compounds B-1, B-2, and B-3 compared to B and B-4.
342
343 Based on the description of its genome-wide DNA binding sites and transcript
344 dysregulation in the presence of Compound C, new inferences can be made about the
345 biological function of AP2-EXP. This demonstrates the utility of combining chemical and
346 traditional genetics, because previous attempts to genetically characterize AP2-EXP
347 have not provided a complete picture of its function in the asexual blood stage. The
348 *ap2-exp* AP2 domain coding region has previously been shown to be essential for the
349 asexual blood stage by both saturating mutagenesis and targeted deletion attempts(15–
350 17, 21). Conversely, the coding region beyond the AP2 domain was truncated in two
351 studies(15, 21). Transcriptomic profiling of a truncated AP2-EXP parasite line revealed
352 that many transcripts encoding exported proteins were dysregulated(21). Unexpectedly,

353 AP2-EXP target genes identified in this study overlap poorly with the differentially
354 regulated transcripts reported as a result of truncation of AP2-EXP **(Table S4)**(21).
355 There is greater overlap between our study and AP2-EXP ChIP-seq recently described
356 by Shang *et al*(17), with 50/101 target genes conserved **(Table S4)**. This may explain
357 the apparent essentiality of the AP2 domain, while the full-length AP2-EXP has roles
358 that impact a different subset of non-essential genes. A comparison of transcripts
359 differentially abundant in the presence of Compound C to the DNA binding occupancy
360 of AP2-I(18) revealed that 54/85 AP2-I target genes are decreased in abundance by
361 $\log_2$ fold change >2 at 24-30 hpi. Interestingly, 22/55 AP2-I target genes that decrease
362 in abundance when parasites are exposed to Compound C are also AP2-EXP targets,
363 implying the potential for co-regulation of certain gene subsets. This overlap is
364 consistent with our EMSA results and corroborates the hypothesis that Compound C
365 competes both AP2-EXP and AP2-I *in vivo*. Only 3/22 PfSIP2 target genes are
366 decreased in abundance in the presence of Compound C **(ApiAP2 Target Gene**
367 **Results Summarized in Table S3)**. This may reflect the cryptic relationship between
368 PfSIP2 DNA binding and transcriptional control, as was noted when PfSIP2 was
369 originally characterized as playing a role in genome integrity and heterochromatin
370 formation(20). Several ApiAP2 proteins have been implicated in oocyst and sporozoite
371 development in a *P. berghei* ApiAP2 knockout screen(24). Since Compounds B and C
372 both inhibit mosquito stage development of *P. berghei* with different phenotypes, it is
373 possible that they target different sets of ApiAP2 proteins *in vivo*. Since PbAP2-Sp is
374 required for sporogony, Compounds B and C should minimally inhibit sporozoite
375 development, which is consistent with our results. Overall, these data support a

376 potential for multi-AP2 domain competition by Compounds B and C *in vivo*. Due to their
377 inclusion in the TCAMS, Compounds B, B-1, B-2, B-3, and B-4 have all been tested for
378 activity against the human HEPG2 cell line(38). ApiAP2 competitor Compounds B and
379 B-4 inhibit HEPG2 cell growth by just 4 and 8%, respectively(38). Therefore, either drug
380 may potentially be prioritized for further development based on selectivity for
381 *Plasmodium* parasites.
382
383 AP2-EXP remains the only ApiAP2 protein for which the structure has been solved.
384 Surprisingly, over a decade later, the field still lacks a clear understanding of the
385 biological role of AP2-EXP or insights into the druggability of the AP2 domain. Our study
386 has enabled us to make inferences about the role of AP2-EXP and set a proof of
387 principle for targeting the highly unique Apicomplexan AP2 DNA binding proteins as a
388 new antimalarial strategy.
389

390 **Materials and Methods**

391 **Parasite lines**

392 Parasites were grown at 37°C, 5% O2, 7% CO2 using RPMI1640 media supplemented
393 with hypoxanthine and .5% Albumax II (Thermo). The parasite lines used in this study
394 were AP2-EXP::GFP, AP2-EXP::HA, AP2-EXP::*glms*, and wild Type Pf3D7 (Malaria
395 Research and Reference Reagents Repository). The human biological samples were
396 sourced ethically and their research use was in accord with the terms of the informed
397 consents under an IRB/EC approved protocol.
398

399 To create AP2-EXP::GFP the C-Terminal coding region of AP2-EXP was cloned **(Table**
400 **S1)** into the plasmid pSLI::2xFKBP(43) for endogenous tagging by single homologous
401 recombination **(Fig S9A)**. Transgenic parasites with the correct C-Terminal tag were
402 further transfected with pLyn-FRB-mCherry(43) for inducible mislocalization using the
403 same method. Parasites were maintained in media with 2.5nM WR99210. All ChIP and
404 western blot experiments using AP2-EXP::GFP except **Figure S26C** were collected
405 using AP2-EXP::GFP without the pLyn mislocalization plasmid.
406
407 To create AP2-EXP::HA we used the pSN054 vector system(58). The right homology
408 region (RHR) was amplified by PCR **(Table S1)**, and the recodonized left homology
409 region (LHR) was synthesized using the BioXP™ 3200 System. Single guide RNA
410 fragments synthesized using the BioXP™ 3200 System **(Table S1)** were cloned into the
411 linearized pSN054 donor vector(58) The parental parasite line used for transfection
412 expresses Cas9 and T7 RNA polymerase(59) **(Fig S9B)**. Cell cultures were maintained
413 in 500nM anhydrotetracycline (aTc, Sigma-Aldrich 37919) and 2.5 mg/mL of Blasticidin
414 S.
415
416 To create AP2-EXP::*glms*::HA, the C-terminal coding region of AP2-EXP **(Table S1)**
417 was cloned into the plasmid pSLI::3xHA::*glms*(60) **(Figure S25)**. Parasite cultures were
418 maintained in media with 2.5nM WR99210.
419
420 All *P. falciparum* transfection were performed as described(43, 61). All parasite strains
421 were cloned by limiting dilution and genotyped using PCR. AP2-EXP::GFP and AP2-

422 EXP::HA were also genotyped using whole genome sequencing **(NCBI SRA:**
423 **PRJNA818769)**.
424
425 **Genomic DNA Isolation**
426 Parasite cultures were lysed with .1% Saponin in 1xPBS and collected by centrifugation
427 at 1500 RPM, then resuspended in 1xPBS. Genomic DNA was then isolated using the
428 Qiagen DNeasy nucleic acid isolation kit according to the manufacturer's instructions.
429
430 **AP2-EXP Knockdown Assays**
431 For AP2-EXP::GFP + pLyn, synchronous parasites were split into two populations.
432 250nM Rapalog was added to one parasite group for 48 hours as described(43) while
433 the second group was used as a control. Addition of Rapalog did not affect parasite
434 growth (not shown). After 48 hours parasite protein was collected in nuclear and
435 cytosolic fractions and AP2-EXP::GFP, Histone H3, and Pf Aldolase were detected by
436 western blot.
437
438 For AP2-EXP::HA parasites were grown routinely in the presence of 500nM aTc. aTc
439 was washed out of one group, while the second group was maintained with aTc as a
440 control for 120 hours before protein harvest. Parasite growth was not affected by
441 removal of aTc (not shown). AP2-EXP::HA and Histone H3 were detected by western
442 blot.
443

444 For AP2-EXP::*glms*::HA, synchronous parasites were split into two populations. 5mM
445 glucosamine was added to one parasite group for 72 hours as described(56) while the
446 second group was used as a control. Addition of glucosamine did not affect parasite
447 growth (not shown). After 72 hours parasite protein was harvested for detection of AP2-
448 EXP::*glms*::HA and Histone H3 by Western blot.
449

450 ***In silico* docking screen**

451 Three dimensional structures of every available molecule from the Tres Cantos
452 Antimalarial Set(38) and the DrugBank(39) were created using BALLOON and the
453 Merck Molecular Force Field. The crystal structure of AP2-EXP (PDB:3IGM) was
454 modeled using AutoDockTools(62) and a molecular docking screen was run using
455 AutoDock(37). Details about the preparation of ligands, the protein macromolecule, and
456 the docking screen criteria are provided in the SI text.
457

458 **IC_{50} Determination**

459 The IC_{50} of each putative ApiAP2 competing compound was determined using a 48-
460 hour growth inhibition assay as described(63). Parasites were seeded at .5%
461 parasitemia and 4% hematocrit on a 96 well plate. Drug or DMSO vehicle control was
462 added in triplicate to each well and parasites were incubated in standard culture
463 conditions for 48 hours. Growth values were normalized to the vehicle control. IC_{50}
464 values were calculated and growth inhibition curves were plotted using GraphPad
465 Prism.
466

467 **Recombinant Protein Expression**

468 Recombinant AP2 domains AP2-EXP, AP2-I D3, AP2-HS D1, and PfSIP2 D1 were
469 overexpressed and purified from BL21 PlysS *E. coli* as described previously(12).
470 Recombinant AtERF1 was purified using the same method after cloning the AtERF1
471 AP2 domain(40) into the pGex4t-1 overexpression vector **(Table S1)** Recombinant
472 protein was quantified using Braford reagent (Pierce). Input, flowthrough, and eluate
473 fractions were analyzed by SDS-PAGE to ensure recovery of the full- length
474 recombinant protein. Recombinant full-length SOX2 protein was purchased from Abcam
475 (ab169843).

476

477 **Electrophoretic Mobility Shift Assays**

478 EMSAs were performed using the Thermo Light Shift EMSA kit with recommended
479 buffer components. Gel shifts were visualized using the Light Shift detection reagents
480 and imaged using a BioRad Chemiluminescence imager. Additional details are provided
481 in the SI text.

482

483 **Ethidium Bromide Exclusion Assay**

484 10μM of double stranded DNA (sequence TGCATGCA, purchased from IDT) in .01mM
485 EDTA, 9.4mM NaCl, 2mM HEPES buffer pH 7.9 was incubated for 10 minutes in the
486 presence of 100nM ethidium bromide. Fluorescence was measured at
487 excitation/emission 546nm/595nm for each well. Following the baseline reading, each
488 putative ApiAP2 competing compound was added. DRAQ5 nucleic acid dye was used
489 as a positive control for knockdown of ethidium bromide fluorescence by DNA

490 intercalation. Ethidium bromide exclusion assays were performed in technical triplicate
491 using a 96 well plate.
492
493 **Molecular Dynamics Simulations**
494 Five Compounds (B, B-1, B-2, B-3, and B-4) were simulated using AMBER(64) for
495 100ns of interaction with the AP2-EXP AP2 domain. The initial position of each
496 compound used was identical to the predicted docking conformation of Compound B
497 **(Figure S2A)**. Additional details are provided in the SI text.
498
499 **AP2 Domain Competitor Phenotyping Assays**
500 For Compound B, C, and H spike in phenotyping **(Fig 3B)**, highly synchronous Pf3D7
501 wild type parasites were spiked with 40μM of each drug at 24 hpi. An equivalent volume
502 of DMSO was used as a control. Parasites were morphologically assessed by Giemsa
503 staining at 24 and 48 hours post spike in (48 and 72 hpi).
504
505 For continuous Compound C exposure phenotyping **(Fig S14)**, highly synchronous
506 Pf3D7 wild type parasites were spiked with 40μM Compound C or equivalent DMSO
507 vehicle control at 6 hpi. Parasite progression was then monitored through the IDC by
508 Giemsa staining at 6-hour intervals.
509
510 For Compound C spike in and washout assays **(Fig 3C, D)**, 40μM of Compound C was
511 spiked into highly synchronous Pf3D7 wild type parasites at 8-hour intervals starting at
512 14 hpi. For Compound C washout assays, 40μM of Compound C was washed out of

513 highly synchronous culture at 8-hour intervals starting at 14 hpi. For each assay an
514 equivalent volume of DMSO vehicle was used as a control for normal growth.
515
516 40μM of compound was used for all phenotyping assays because it was the maximum
517 concentration achievable in culture after resuspending each compound in 100% DMSO.
518 Higher concentrations of compound caused parasite mortality due to the high volume of
519 DMSO required (>1% volume/volume). All assays were performed in triplicate and
520 parasitemia was counted by Giemsa-stained slides. Error bars represent the standard
521 deviation of the mean.
522
523 **Preparation of RNA for DNA Microarray**
524 DNA microarrays were prepared using the protocol described in(44). After collecting a
525 control timepoint at 6 hpi, parasites were cultured with either 12μM Compound C
526 (0.66xIC_{50}) or DMSO vehicle control. Parasite RNA was then harvested at 12, 18, 24,
527 30, 42, and 48 hpi. cDNA was synthesized using SuperScript II, hybridized onto Agilent
528 DNA microarrays(44), and scanned using an Axon 4200A Scanner. Agilent Feature
529 Extraction Software version 11.0.1.1 with the protocol GE2-
530 v5_95_Feb07_nospikein was used to extract signal intensities. Raw signal intensities for
531 all microarrays are available in **Table S6**.
532
533 **Analysis of DNA Microarray**s
534 Microarray data were extracted and normalized using the R packages LIMMA(46) and
535 Rnits(48). LIMMA was used to analyze differential transcript abundance specific to the

536 24-30 hpi time points, while Rnits was used to assess differential transcript abundance
537 across the entire IDC. Additional details are provided in the SI text.
538
539 **ChIP Sample Preparation**
540 ChIP-seq was performed as described in(19). Highly synchronous parasites were grown
541 up to 5-10% parasitemia and AP2-EXP ChIP was performed at 30 hpi for all replicates.
542 Immunoprecipitation was performed overnight with .5mg/mL 3F10 anti-HA (Sigma) or
543 .1mg/mL Ab290 (Abcam) anti-GFP.
544
545 **DNA Library Preparation**
546 ChIP-seq DNA libraries were prepared as described in(19) using the NEBNext II DNA
547 library kit (New England Biolabs) according to the manufacturer's instructions. Quality
548 was assessed using an Agilent 2100 Bioanalyzer or TapeStation. Libraries were
549 sequenced using a HiSeq 2500 (Replicate GFP1) or NextSeq 550 (Replicates GFP2
550 and HA3) Illumina sequencer. AMPure XP beads (Beckmann Coulter) were used to size
551 select and purify DNA between NEBNext II library preparation steps. Whole genome
552 sequencing DNA libraries were prepared using the Illumina Tru-Seq PCR free DNA
553 library kit and sequenced using a HiSeq 2500.
554
555 **qPCR**
556 ChIP-qPCR samples were collected from 30-35 hpi trophozoites following two hours of
557 Compound C (40μM) or DMSO vehicle control spike in. Primer pairs to be used for
558 ChIP-qPCR were first evaluated to check for 80-110% efficiency using sonicated

559 genomic DNA. RT-qPCR was carried out using Sybr Green Polymerase master mix
560 (Thermo) with the specified primer concentration **(Table S1)**. The Ct was calculated
561 using SDSv1.4 (Applied Biosystems) software, averaged over technical triplicate. The
562 percent of input per immunoprecipitated DNA fraction was calculated using the delta Ct
563 method. Each assay was performed in biological triplicate, with the exception of the
564 *ama1* primer pair, where n =2. Error bars represent standard deviation of the mean.
565 Data was obtained using an Applied Biosystems 7300 Real-Time PCR Machine.
566

567 **ChIP-seq data analysis**

568 ChIP-seq reads were mapped to generate bam and bigwig coverage files as
569 described(19). Peaks of occupancy were called using MACS2(65) and peak intervals
570 were analyzed for overlap and gene proximity using BedTools(66). DeepTools(67) and
571 cegr-tools: https://github.com/seqcode/cegr-tools were used to compare AP2-EXP
572 genomic occupancy to previously published datasets. BedTools(66) was used to
573 compare AP2-EXP peaks determined in this study to AP2-EXP occupancy determined
574 in Shang *et al*(17). Additional details are available in the SI text.
575

576 **Western blot**

577 Full parasite protein western blot samples were collected by lysing RBCs with .1%
578 saponin and boiling protein in Loading Buffer (50mM Tris-Cl pH 8.0, 20% SDS, 1%
579 Bromophenol Blue). Fractionated parasite protein was prepared as described in(68).
580 Blots were performed as described(19). Primary antibodies used were: 1/1000 rat ant-
581 HA (Roche 3F10), 1/1000 mouse anti-GFP (Roche), 1/3000 rabbit anti-aldolase

582 conjugated to HRP (Abcam ab38905), or 1/3000 mouse anti-H3 (Abcam ab10799).
583 Secondary antibody concentrations used were 1/3000 goat anti-rat HRP conjugate
584 (Millipore), 1/3000 goat anti-mouse HRP conjugate, or 1/10,000 (Pierce) goat anti-rabbit
585 HRP conjugate (Millipore). ECL reagent (Pierce) was used to detect HRP signal. Blots
586 were exposed to autoradiography film (VWR) and visualized using an autoradiography
587 developer.
588

589 **Protein Pulldown**

590 Parasite nuclear protein was isolated as described(68). Following the final centrifugation
591 step, the supernatant was collected and diluted by 1:3 in Dilution Buffer (30% glycerol,
592 20mM HEPES pH 7.8). GFP tagged AP2-EXP was pulled down using Chromotek anti-
593 GFP or mock immunoprecipitated using Chromotek negative control magnetic beads.
594 Beads were washed twice in Wash Buffer prior to use (20mM HEPES pH 7.4, 250mM
595 NaCl, 1mM EDTA, 1mM TCEP, .05% NP-40). Protein and beads were incubated
596 together for 1 hour at 4°C. The beads were then washed twice with Wash Buffer and
597 bound protein was collected in Loading Buffer by boiling at 95°C for 10 minutes.
598

599 **Fluorescent Microscopy**

600 Samples were prepared by incubating packed infected red blood cells with DRAQ5 dye
601 (Thermo) for 15 minutes. Parasites were then washed in 1xPBS to remove excess dye
602 and immediately placed on a glass slide for imaging. Fluorescent microscopy images
603 were acquired using an Olympus Bx61 fluorescent microscope. All images were
604 processed using SlideBook 5.0.

605

606 **Mosquito Stage *P. berghei* AP2 Competitor Growth Assay**

607 A female Swiss Webster mouse was inoculated with Plasmodium berghei ANKA 2.34
608 from frozen stock. Once the parasitemia reached 15%, the blood was harvested by
609 heart puncture, washed twice with 1xPBS and resuspended to 10 mL in 1xPBS. 5
610 female Swiss Webster mice were infected with 500µL of the resuspended blood. 3-4
611 days after blood passage, exflagellation of the *P. berghei* gametocytes was assayed.
612 Briefly, a drop of tail vein blood was incubated in RPMI 1640 (Invitrogen) containing
613 1µM xanthurenic acid (Sigma) for ~12-15 minutes. The mixture was added to a slide
614 and observed under a light microscope at 40x. Ten or more fields were observed. Mice
615 with 0.3-0.7 exflagellations per field were anesthetized and fed on 3-day post
616 emergence *Anopheles stephensi* mosquitoes. Mosquitoes were maintained at 19°C and
617 fed 10% sucrose. Bloodfed mosquitoes were separated 30 hours post-bloodfeeding. 10
618 days post-bloodfeeding, 5-10 mosquitoes were dissected and oocysts were counted by
619 mercurochrome staining and light microscopy to ensure *P. berghei* oocyst development.
620 The remaining mosquitoes were injected by standard mouth pipette technique with
621 different small molecule inhibitors or a mixture of PBS-DMSO for control. The surviving
622 injected mosquitoes were dissected 14 days post-bloodfeeding. Oocyst numbers were
623 counted as before. Pooled groups of midguts were ground using a pestle, centrifuged at
624 7,000 rpm for 5 minutes and resuspended in 20µL 1xPBS to release developing mid-gut
625 sporozoites. Sporozoites were counted on a haemocytometer. All *in vivo* studies were
626 conducted in accordance with the GSK Policy on the Care, Welfare and Treatment of

627 Laboratory Animals and were reviewed by the ethical review process at the institution
628 where the work was performed.
629
630 **Data Availability**
631 All sequencing data for AP2-EXP ChIP-seq and whole genome sequencing for
632 transgenic parasite lines AP2-EXP::GFP and AP2-EXP::HA have been deposited in the
633 SRA under **PRJNA818769**.
634
635 Raw microarray data for the Compound C and DMSO control RNA time course is
636 available in **Table S6**.
637
638 **Author Contributions**
639 T.J.R., E.K.D.S., K.S.S., N.D., C.F.A.P., G.P., M.J.L., C.D.O., J.C.N. and M.L. designed
640 experiments. T.J.R., E.K.D.S., V.C., K.S.S., N.D., G.J., C.F.A.P., and I.K. performed the
641 experiments. T.J.R. and N.D. analyzed the data. T.J.R. generated figures. T.J.R. and
642 M.L. wrote the manuscript in collaboration with the other authors
643
644 **Acknowledgements**
645 This work was funded through NIH/NIAID R01AI076276 (M.L.), R01AI125565 (M.L.),
646 and with support from the Center for Quantitative Biology (P50 GM071508) (M.L.).
647 T.J.R. was supported by NIH T32 Predoctoral Training Grant (5T32GM125592-01)
648 awarded to the Center for Eukaryotic Gene Regulation (CEGR) at The Pennsylvania
649 State University. G.A.J. is a recipient of the Sir Keith Murdoch Fellowship from the

American Australian Association and a Postdoctoral Research Grant from the American Heart Association (16POST26420067).

Conflict of Interest Statement

F.J.G. is a GlaxoSmithKline employee and own shares of the company

REFERENCES

1. WHO, *Word Malaria Report 2021* (2021).
2. M. M. Ippolito, K. A. Moser, J.-B. B. Kabuya, C. Cunningham, J. J. Juliano, Antimalarial Drug Resistance and Implications for the WHO Global Technical Strategy. *Curr. Epidemiol. Reports* **8**, 46–62 (2021).
3. A. F. Cowman, J. Healer, D. Marapana, K. Marsh, Malaria: Biology and Disease. *Cell* **167**, 610–624 (2016).
4. G. Zanghì, *et al.*, A Specific PfEMP1 Is Expressed in P. falciparum Sporozoites and Plays a Role in Hepatocyte Infection. *Cell Rep.* **22**, 2951–2963 (2018).
5. Z. Bozdech, *et al.*, The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. *PLoS Biol.* **1**, 85–100 (2003).
6. X. M. Lu, *et al.*, Nascent RNA sequencing reveals mechanisms of gene regulation in the human malaria parasite Plasmodium falciparum. *Nucleic Acids Res.* **45**, 7825–7840 (2017).
7. H. J. Painter, M. Carrasquilla, M. Llinás, Capturing in vivo RNA transcriptional dynamics from the malaria parasite Plasmodium falciparum. *Genome Res.* **27**, 1074–1086 (2017).
8. K. G. Le Roch, *et al.*, Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**, 1503–1508 (2003).
9. B. J. Foth, *et al.*, Quantitative time-course profiling of parasite and host cell proteins in the human malaria parasite Plasmodium falciparum. *Mol. Cell. Proteomics* **10**, 1–16 (2011).
10. T. Hollin, K. G. Le Roch, From Genes to Transcripts, a Tightly Regulated Journey in Plasmodium. *Front. Cell. Infect. Microbiol.* **10**, 1–13 (2020).
11. H. J. Painter, T. L. Campbell, M. Llinás, The Apicomplexan AP2 family: Integral factors regulating Plasmodium development. **176**, 1–7 (2012).
12. T. L. Campbell, E. K. de Silva, K. L. Olszewski, O. Elemento, M. Llinás, Identification and Genome-Wide Prediction of DNA Binding Specificities for the ApiAP2 family of regulators from the malaria parasite. *PLoS Pathog.* **6**, e1001165 (2010).
13. S. Balaji, M. Madan Babu, L. M. Iyer, L. Aravind, Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res.* **33**, 3994–4006 (2005).
14. M. Jeninga, J. Quinn, M. Petter, ApiAP2 Transcription Factors in Apicomplexan Parasites. *Pathogens* **8** (2019).
15. M. Zhang, *et al.*, Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science* **360**, eaap7847 (2018).
16. X. Shang, *et al.*, A cascade of transcriptional repression determines sexual commitment and

692 development in Plasmodium falciparum. *Nucleic Acids Res.* **49**, 9264–9279 (2021).
693 17. X. Shang, *et al.*, Genome-wide landscape of ApiAP2 transcription factors reveals a
694 heterochromatin-associated regulatory network during Plasmodium falciparum blood-stage
695 development. *Nucleic Acids Res.* **50**, 3413–3431 (2022).
696 18. J. M. Santos, *et al.*, Red Blood Cell Invasion by the Malaria Parasite Is Coordinated by the PfAP2-I
697 Transcription Factor. *Cell Host Microbe* **21**, 731–741 (2017).
698 19. G. A. Josling, *et al.*, Dissecting the role of PfAP2-G in malaria gametocytogenesis. *Nat. Commun.*
699 **11**, 1503 (2020).
700 20. C. Flueck, *et al.*, A major role for the Plasmodium falciparum ApiAP2 protein PfSIP2 in
701 chromosome end biology. *PLoS Pathog.* **6**, e1000784 (2010).
702 21. R. M. Martins, *et al.*, An ApiAP2 member regulates expression of clonally variant genes of the
703 human malaria parasite Plasmodium falciparum. *Sci. Rep.* **7**, 14042 (2017).
704 22. E. Carrington, *et al.*, The ApiAP2 factor PfAP2-HC is an integral component of heterochromatin in
705 the malaria parasite Plasmodium falciparum. *iScience* **24**, 102444 (2021).
706 23. M. Sierra-Miranda, *et al.*, PfAP2Tel, harbouring a non-canonical DNA-binding AP2 domain, binds
707 to Plasmodium falciparum telomeres. *Cell. Microbiol.* **19**, e12742 (2017).
708 24. K. Modrzynska, *et al.*, A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting
709 Transcriptional Regulators Controlling the Plasmodium Life Cycle. *Cell Host Microbe* **21**, 11–22
710 (2017).
711 25. Z. Li, *et al.*, Plasmodium transcription repressor AP2-O3 regulates sex-specific identity of gene
712 expression in female gametocytes. *EMBO Rep.*, e51660 (2021).
713 26. Y. Xu, *et al.*, PfAP2-G2 Is Associated to Production and Maturation of Gametocytes in Plasmodium
714 falciparum via Regulating the Expression of PfMDV-1. *Front. Microbiol.* **11**, 631444 (2021).
715 27. S. Singh, *et al.*, The PfAP2-G2 transcription factor is a critical regulator of gametocyte maturation.
716 *Mol. Microbiol.* **115**, 1005–1024 (2021).
717 28. C. Zhang, *et al.*, Systematic CRISPR-Cas9-Mediated ApiAP2 Genes Reveal Functional Insights into
718 Parasite Development. *MBio* **8**, 1–17 (2017).
719 29. A. Sinha, *et al.*, Europe PMC Funders Group Europe PMC Funders Author Manuscripts A cascade
720 of DNA binding proteins for sexual commitment and development in Plasmodium. **507**, 253–257
721 (2014).
722 30. B. F. C. Kafsack, *et al.*, A transcriptional switch underlies commitment to sexual development in
723 malaria parasites. *Nature* **507**, 248–52 (2014).
724 31. S. Iwanaga, I. Kaneko, T. Kato, M. Yuda, Identification of an AP2-family Protein That Is Critical for
725 Malaria Liver Stage Development. *PLoS One* **7**, e47557 (2012).
726 32. M. Yuda, S. Iwanaga, S. Shigenobu, T. Kato, I. Kaneko, Transcription factor AP2-Sp and its target
727 genes in malarial sporozoites. *Mol. Microbiol.* **75**, 854–863 (2010).
728 33. M. Yuda, *et al.*, Identification of a transcription factor in the mosquito-invasive stage of malaria
729 parasites. *Mol. Microbiol.* **71**, 1402–1414 (2009).
730 34. M. Yuda, I. Kaneko, S. Iwanaga, Y. Murata, T. Kato, Female-specific gene regulation in malaria
731 parasites by an AP2-family transcription factor. *Mol. Microbiol.* **113**, 40–51 (2020).
732 35. E. Tintó-font, L. Michel-todó, T. J. Russell, N. Casas-, A heat-shock response regulated by the
733 PfAP2-HS transcription factor protects human malaria parasites from febrile temperatures. **6**,
734 1163–1174 (2021).
735 36. S. E. Lindner, E. K. De Silva, J. L. Keck, M. Llinás, Structural Determinants of DNA Binding by a P.
736 falciparum ApiAP2 Transcriptional Regulator. *J. Mol. Biol.* **395**, 558–567 (2010).
737 37. G. M. Morris, *et al.*, Automated docking using a Lamarckian genetic algorithm and an empirical
738 binding free energy function. *J. Comput. Chem.* **19**, 1639–1662 (1998).
739 38. F. J. Gamo, *et al.*, Thousands of chemical starting points for antimalarial lead identification.

740 *Nature* **465**, 305–310 (2010).
741 39. D. S. Wishart, *et al.*, DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic*
742 *Acids Res.* **46**, D1074–D1082 (2018).
743 40. M. D. Allen, K. Yamasaki, M. Ohme-Takagi, M. Tateno, M. Suzuki, A novel mode of DNA
744 recognition by a β-sheet revealed by the solution structure of the GCC-box binding domain in
745 complex with DNA. *EMBO J.* **17**, 5484–5496 (1998).
746 41. A. J. Geall, I. S. Blagbrough, A comparison of the properties of Sox-3 with Sry and two related
747 genes, Sox-1 and Sox-2. *J. Pharm. Biomed. Anal.* **22**, 849–859 (2000).
748 42. Y. Wang, H. Schellenberg, V. Walhorn, K. Toensing, D. Anselmetti, Binding Mechanism of
749 Fluorescent Dyes to DNA Characterized by Magnetic Tweezers. *Mater. Today Proc.* **4**, S218–S225
750 (2017).
751 43. J. Birnbaum, *et al.*, A genetic system to study Plasmodium falciparum protein function. *Nat.*
752 *Methods* **14**, 450–456 (2017).
753 44. B. F. C. Kafsack, H. J. Painter, M. Llinás, New Agilent platform DNA microarrays for transcriptome
754 analysis of Plasmodium falciparum and Plasmodium berghei for the malaria research community.
755 *Malar. J.* **11**, 187 (2012).
756 45. M. Llinás, Z. Bozdech, E. D. Wong, A. T. Adai, J. L. DeRisi, Comparative whole genome
757 transcriptome analysis of three Plasmodium falciparum strains. *Nucleic Acids Res.* **34**, 1166–1173
758 (2006).
759 46. G. K. Smyth, Linear models and empirical Bayes methods for assessing differential expression in
760 microarray experiments. *Stat. Appl. Ge- netics Mol. Biol.* **3**, Article 3 (2011).
761 47. A. Bahl, *et al.*, PlasmoDB: The Plasmodium genome resource. A database integrating
762 experimental and computational data. *Nucleic Acids Res.* **31**, 212–215 (2003).
763 48. D. P. Sangurdekar, The Rnits package for normalization and inference of differential expression in
764 time series microarray data . *Rvignette*, 1–13 (2014).
765 49. M. Yuda, I. Kaneko, Y. Murata, S. Iwanaga, T. Nishi, Targetome analysis of malaria sporozoite
766 transcription factor AP2-Sp reveals its role as a master regulator. *bioRxiv* (2022)
767 https:/doi.org/https://doi.org/10.1101/2022.02.23.481739.
768 50. P. R. Kensche, *et al.*, The nucleosome landscape of Plasmodium falciparum reveals chromatin
769 architecture and dynamics of regulatory sequences. *Nucleic Acids Res.* **44**, 2110–2124 (2015).
770 51. R. Bártfai, *et al.*, H2A.Z demarcates intergenic regions of the Plasmodium falciparum epigenome
771 that are dynamically marked by H3K9ac and H3K4me3. *PLoS Pathog.* **6**, e1001223 (2010).
772 52. G. A. Josling, *et al.*, A Plasmodium Falciparum Bromodomain Protein Regulates Invasion Gene
773 Expression. *Cell Host Microbe* **17**, 741–751 (2015).
774 53. L. Chappell, *et al.*, Refining the transcriptome of the human malaria parasite Plasmodium
775 falciparum using amplification-free RNA-seq. *BMC Genomics* **21**, 395 (2020).
776 54. C. G. Toenhake, *et al.*, Chromatin Accessibility-Based Characterization of the Gene Regulatory
777 Network Underlying Plasmodium falciparum Blood-Stage Development. *Cell Host Microbe* **23**,
778 557–569 (2018).
779 55. S. M. Ganesan, A. Falla, S. J. Goldfless, A. S. Nasamu, J. C. Niles, Synthetic RNA-protein modules
780 integrated with native translation mechanisms to control gene expression in malaria parasites.
781 *Nat. Commun.* **7**, 10727 (2016).
782 56. P. Prommana, *et al.*, Inducible Knockdown of Plasmodium Gene Expression Using the glmS
783 Ribozyme. *PLoS One* **8**, e73783 (2013).
784 57. D. G. Paton, *et al.*, Exposing Anopheles mosquitoes to antimalarials blocks Plasmodium parasite
785 transmission. *Nature* **567**, 239–243 (2019).
786 58. A. S. Nasamu, *et al.*, An integrated platform for genome engineering and gene expression
787 perturbation in Plasmodium falciparum. *Sci. Rep.* **11**, 342 (2021).

788 59. J. M. Murithi, *et al.*, The Plasmodium falciparum ABC transporter ABCI3 confers parasite strain-
789 dependent pleiotropic antimalarial drug resistance. *Cell Chem. Biol.*, 1–16 (2021).
790 60. C. S. Simon, *et al.*, An extended DNA-free intranuclear compartment organizes centrosome
791 microtubules in malaria parasites. *Life Sci. Alliance* **4**, e202101199 (2021).
792 61. Y. Wu, L. A. Kirkman, T. E. Wellems, Transformation of Plasmodium falciparum malaria parasites
793 by homologous integration of plasmids that confer resistance to pyrimethamine. *Proc. Natl.*
794 *Acad. Sci. U. S. A.* **93**, 1130–4 (1996).
795 62. R. Huey, G. M. Morris, Using AutoDock 4 with AutoDockTools : A Tutorial. 1–56.
796 63. L. Abrahams, M. Burke, w o T w o. 20–36 (2010).
797 64. A. W. Götz, *et al.*, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1.
798 generalized born. *J. Chem. Theory Comput.* **8**, 1542–1555 (2012).
799 65. J. M. Gaspar, Improved peak-calling with MACS2. *bioRxiv*, 1–16 (2018).
800 66. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features.
801 *Bioinformatics* **26**, 841–842 (2010).
802 67. F. Ramírez, F. Dündar, S. Diehl, B. A. Grüning, T. Manke, DeepTools: A flexible platform for
803 exploring deep-sequencing data. *Nucleic Acids Res.* **42**, 187–191 (2014).
804 68. S. C. Oehring, *et al.*, Organellar proteomics reveals hundreds of novel nuclear proteins in the
805 malaria parasite Plasmodium falciparum. *Genome Biol* **13**, R108 (2012).
806
807

808 ## Figure Legends

809

810 **Figure 1. Nine putative competitors of DNA binding by AP2-EXP were identified by an *in-***
811 ***silico* screen and tested *in vitro* and *in vivo***

812

813 A) Each compound that was prioritized based on the *in-silico* screen was assigned an identifier
814 (A-I) and tested for anti-*Plasmodium* activity in a 48-hour growth inhibition assay against
815 asexual blood stage *P. falciparum*.

816

817 B) Each putative AP2-EXP competitor was added to an EMSA containing the AP2-EXP AP2
818 domain. DNA binding competition will result in a loss of the shifted DNA probe. DMSO vehicle
819 was used as a control for normal DNA binding by AP2-EXP. 150 fmoles of DNA probe, 125ng of
820 AP2-EXP, and 300μM of each compound were used for each lane.

821

822 C) Chemical structures of AP2-EXP competing Compounds A, B, C, and I. Three of the four
823 compounds that compete DNA binding *in vitro* (Compounds A, B, and C) have a benzoxazole
824 core moiety, denoted by a red box.
825
826 **Figure 2. Analogues of Compound B compete AP2-EXP DNA binding with differing**
827 **efficacy**
828
829 A) Compounds B, B-1, B-2, B-3, and B-4 were added to an EMSA with AP2-EXP. DNA binding
830 competition will result in the loss of the shifted probe. 40 fmoles of DNA probe, 125ng of AP2-
831 EXP, and 300μM of each compound were used for each lane.
832
833 B) Compounds B, B-1 and B-4 were titrated at 25, 50, and 150μM into an EMSA with AP2-EXP
834 to determine whether there are differences in DNA binding competition. DMSO vehicle was
835 used as a control for normal DNA binding. 40 fmoles of DNA probe and 125ng of AP2-EXP were
836 used for each lane. The concentration increase for each compound is indicated by the triangle
837 from left to right.
838
839 C) Chemical structures of Compounds B, B-1, B-2, B-3, and B-4. Compounds that compete
840 DNA binding by AP2-EXP are highlighted in black. Each compound has a different substitution
841 pattern on the right-side benzene ring, denoted by a red box. The IC_{50} against asexual *P.*
842 *falciparum* of compounds B-1, B-2, B-3, and B-4 measured by Gamo *et al*(38) is indicated at the
843 bottom of each identifier. The IC_{50} of Compound B was determined in this study.
844
845 **Figure 3. AP2 Domain Competing Compounds B and C affect *P. falciparum* at in the mid**
846 **trophozoite stage, coinciding with the maximum expression of AP2-EXP**
847

848 A) AP2-EXP protein expression in the asexual blood stages ring (R), trophozoite (T), and
849 schizont (S), and Stage III gametocytes (Gam), was probed by western blot against AP2-EXP
850 endogenously tagged with GFP (AP2-EXP::GFP). The expected molecular weight of AP2-
851 EXP::GFP (147kDa) is indicated by an arrow. Wild type Pf3D7 protein was used as a negative
852 control. The full length western blot is provided in **Figure S10**.
853
854 B) Representative Giemsa-stained microscopy images of *P. falciparum* parasites spiked with
855 40μM Compounds B, C, or H. DMSO vehicle was used as a control for normal progression
856 through the asexual blood stage at 24 and 48 hours post spike-in.
857
858 C) Highly synchronous asexual blood stage *P. falciparum* parasites were cultured with 40μM
859 Compound C, which was washed out of the media at eight-hour intervals. IDC progression was
860 measured by counting ring stage parasites at 54 hpi. Each assay was done in triplicate. Error
861 bars represent standard deviation.
862
863 D) Highly synchronous asexual blood stage *P. falciparum* parasites were spiked with 40μM
864 Compound C at eight-hour intervals in the IDC. IDC progression was measured by counting ring
865 stage parasites at 54 hpi Each assay was done in triplicate. Error bars represent standard
866 deviation.
867
868 **Figure 4. Compound C disrupts the *P. falciparum* transcriptome specifically at 30hpi, with**
869 **bias towards AP2-EXP targets predicted by ChIP-seq**
870
871 A) *P. falciparum* parasites were spiked with 12μM (0.66xIC_{50}) Compound C or DMSO vehicle
872 control at 10 hpi. Total RNA was harvested at 6, 12, 18, 24, 30, 42, and 48 hpi for quantification

873 by DNA microarray. The full transcriptome Spearman Correlation between Compound C and
874 DMSO control spiked parasites was plotted as a correlogram. A * indicates *p-value* < 0.05.
875
876 B) Three replicates of AP2-EXP Chip-seq were collected at 30 hpi using two genetically tagged
877 parasite lines (AP2-EXP::GFP and AP2-EXP::HA). A no epitope control was collected using wild
878 type Pf3D7 parasites and anti-GFP antibodies. Log_2 enrichment of the immunoprecipitate (ChIP)
879 over Input DNA for one replicate of AP2-EXP::GFP, AP2-EXP::HA, and No Epitope Control is
880 plotted relative to the coding sequence start site of AP2-EXP target genes identified in at least
881 2/3 replicates. CATGA is the most overrepresented DNA motif within the peaks of AP2-EXP
882 occupancy conserved in at least 2/3 ChIP-seq experiments.
883
884 C) A volcano plot of the changes in abundance for AP2-EXP target genes at 24-30 hpi in the
885 presence of Compound C. 46/93 detected AP2-EXP target transcripts decrease in abundance
886 by log_2 fold change > 2.
887
888 D) GO-term analysis of transcripts that are both AP2-EXP targets and decreased in abundance
889 with respect to Compound C (Bonferroni *p-value* cutoff < 0.05).
890
891 **Figure 5. Compounds B and C are active against mosquito stage *Plasmodium berghei***
892 **parasites**
893
894 A) Midgut oocyst counts per mosquito for *P. berghei* infected mosquitoes on day 14 post
895 infection following compound injection on day 10 post infection. Compound identity is indicated
896 in the legend. Rep. 1 and Rep. 2 correspond to biological replicates one and two, respectively.
897

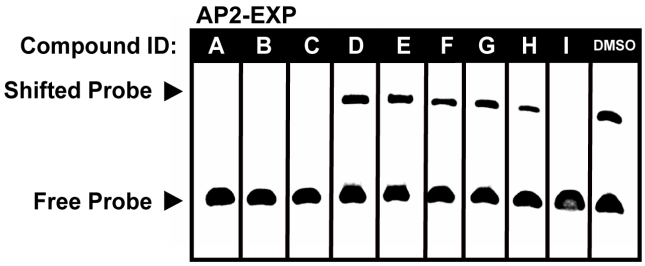
898 B) Midgut sporozoite counts per five infected mosquitoes on day 14 post infection following
899 compound injection on day 10 post infection. Compound identity is indicated in the legend. Rep.
900 1 and Rep. 2 correspond to biological replicates one and two, respectively.
901

Figure 1

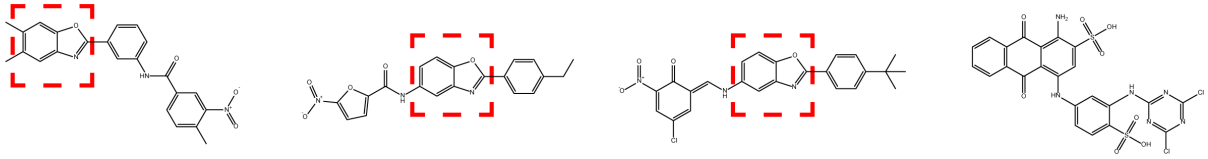
A

Letter Code	PubChem ID	IC_{50} (μM)
A	CID1365835	11
B	CID1365471	80
C	CID5750730	18
D	CID4541005	22
E	CID2343	170
F	CID4437	40
G	CID11390	34
H	CID6914273	40
I	CID25863	125

B



C



CID1365853
Compound A

CID1365471
Compound B

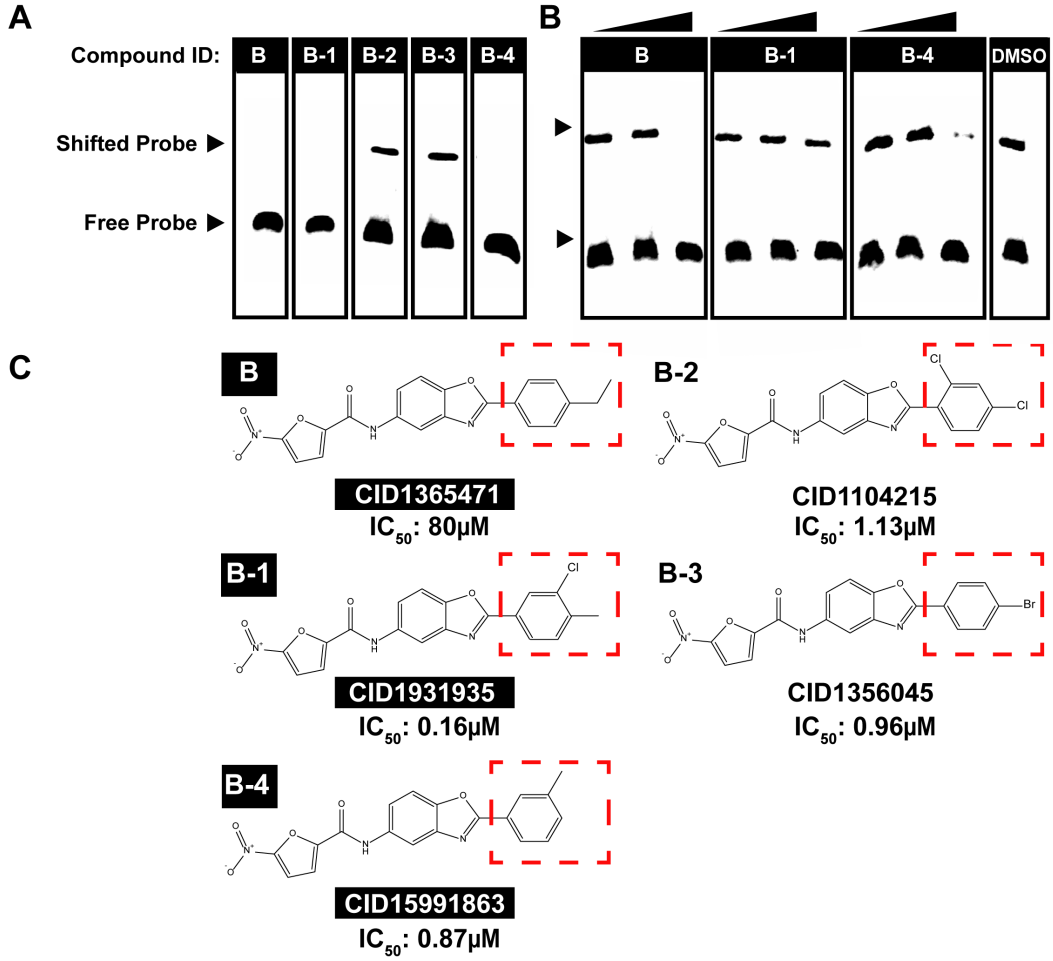
CID124220
Compound C

CID25863
Compound I

32 **Figure 2**

33

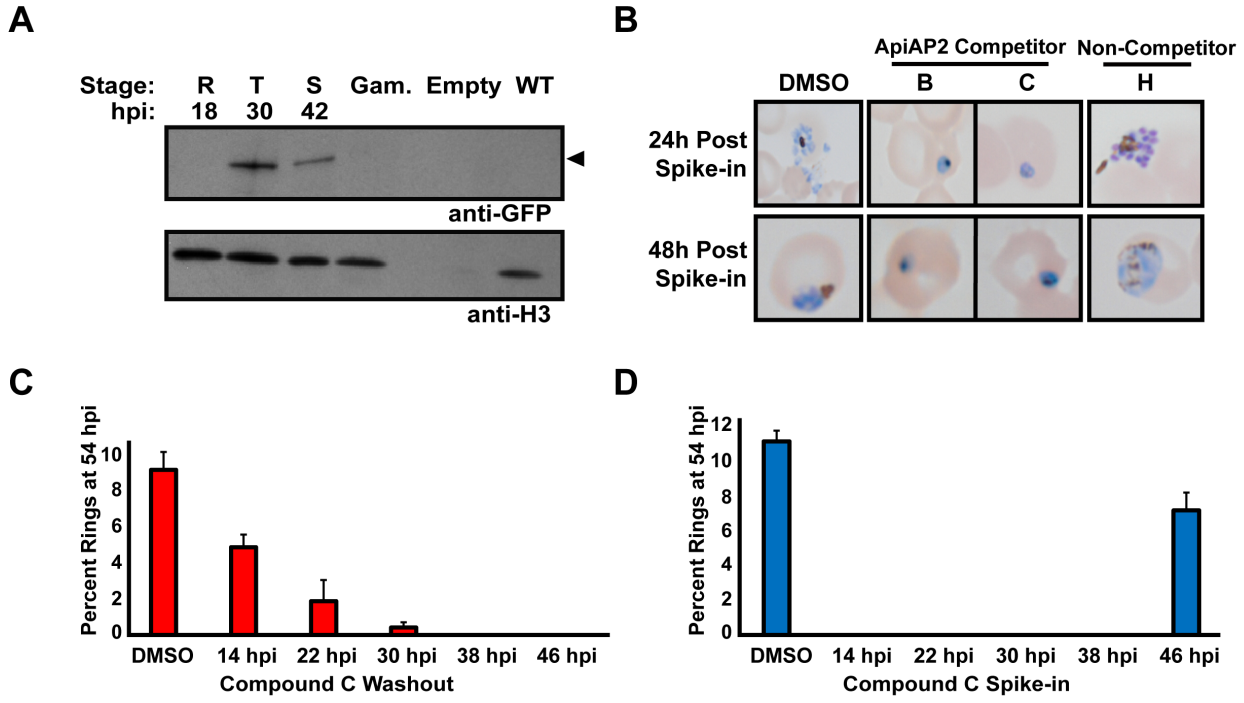
34



35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

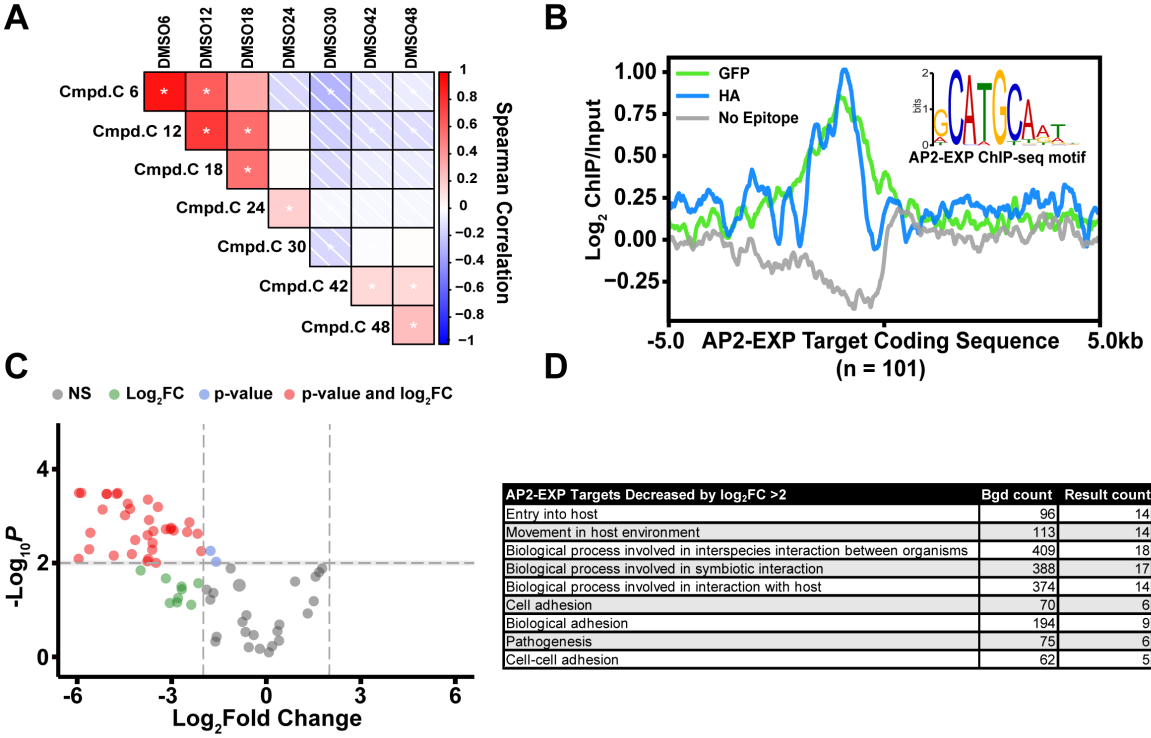
53 **Figure 3**

54



55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81

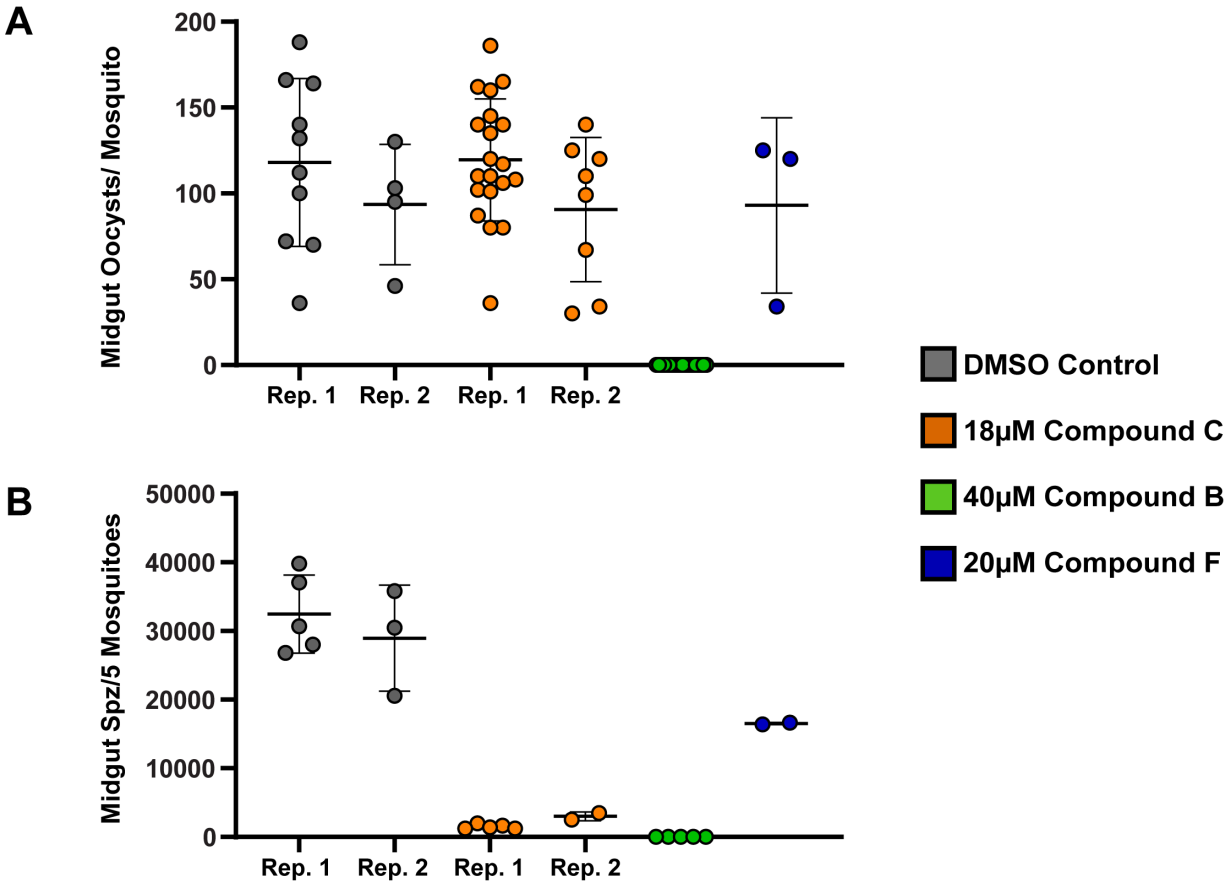
Figure 4



AP2-EXP Targets Decreased by log_2FC >2	Bgd count	Result count	Bonferroni
Entry into host	96	14	2.32E-12
Movement in host environment	113	14	2.43E-11
Biological process involved in interspecies interaction between organisms	409	18	2.03E-07
Biological process involved in symbiotic interaction	388	17	7.73E-07
Biological process involved in interaction with host	374	14	0.000201045
Cell adhesion	70	6	0.004098861
Biological adhesion	194	9	0.004708229
Pathogenesis	75	6	0.006101223
Cell-cell adhesion	62	5	0.028388607

109 **Figure 5**

110



111

SI Text

Extended Methods

In Silico Docking Screen

The Tres Cantos Antimalarial Set of 13533 small molecules was retrieved from the supplementary material of Gamo *et al*(1) and downloaded in Spatial Data File (SDF) format from BatchEntrez. 4603 small molecule structures were downloaded from DrugBank(2) version 2.5 in SDF format. Using the SDF library, a 3-dimensional MOL2 library was created with the use of BALLOON and the Merck Molecular Force Field. BALLOON was used to calculate up to 20 conformations using a genetic algorithm and to select the conformation with the lowest energy of conformation. The 3-dimensional MOL2 library was then prepared for AutoDock using the python script "ligand_prepare4.py" from AutoDockTools(3). The prepared ligands were then saved in PDBQT file format. The python script "ligand_prepare4.py" automatically defines the torsions of the molecule.

The homodimerized AP2-EXP AP2 domain was downloaded from the solved crystal structure (PDB ID 3IGM). AutoDockTools was used to remove water molecules, add hydrogens, and calculate charges. Residues Arg-88 and Asn-118 were found to be flexible using FlexPred and set as flexible residues for docking. The flexible residues were saved as flexible residues in a flexible residue file and the remaining molecule was saved as rigid residues in a rigid residue file as specified in the AutoDock manual(3). The rigid residues file was used to pre-calculate an energy grid of the macromolecule using autogrid4(3). A set of 14 grid map types S, Cl, F,A, Br, N, P, OA, SA, C, I, HD, Ca, NA were calculated for the macromolecule using AutoGrid(3).

From the prepared ligands a docking set was prepared for AutoDock with the macromolecule and pre-calculated energy grids. For each docking set a docking parameter file was prepared, with the Lamarckian Genetic Algorithm (LGA) specified as the search algorithm for AutoDock.

27 To allow for high throughput parallel processing PERL was used on top of AutoDock to manage
28 the dockings. The algorithm was initially set to run with a maximum of 25000 energy evaluations
29 and 20 repeats. Since the calculations are computationally intensive, this setting was used as
30 initial screening of the ligands, to identify a subset of ligands for a more thorough evaluation(3).
31 The 1000 best hits from the GSK compound evaluation and the DrugBank compound evaluation
32 were selected for a more thorough evaluation. The docking parameters were reconfigured to
33 250000 energy evaluations and 100 repeats and docking repeated. The docking results were
34 examined with PERL, to provide an automated approach to the interpretation of the docking
35 results. The docking results were evaluated using two different approaches. The top candidates
36 for competition of DNA were selected using a PERL script to filter based on having a geometric
37 center within 10 Å of the location of sense-strand DNA binding. Hits were then filtered based on
38 having a predicted free energy of interaction < -5 kJ/Mol.
39
40 **Electrophoretic Mobility Shift Assay**
41 EMSAs were run in DNA binding buffer (10mM Tris pH 7.5, 50mM KCl, 1mM DTT, 6mM MgCl2,
42 60ng/μL Poly DiDC, 65ng BSA). Recombinant proteins were titrated to empirically determine the
43 minimum mass required for DNA binding and this mass **(Fig S4)** was used for each gel-shift
44 unless otherwise specified. PAGE purified DNA probes with a 5' biotin ligated on the forward
45 DNA strand, along with an unlabeled complementary stand, were purchased from IDT **(Table**
46 **S1)**. DNA probes were double stranded by heating to 95°C, followed by stepwise cooling in
47 Annealing Buffer (10mMTtris-Cl pH7.5, 1mM EDTA, 10mM NaCl). Each recombinant protein
48 was incubated in DNA binding mixture plus competitor compound for 15 minutes prior to
49 addition of the cognate double stranded DNA oligonucleotide. Protein, competitor, and DNA
50 oligonucleotide were incubated together for an additional 5 minutes. The mixtures were
51 separated on a .5x TBE polyacrylamide gel, transferred to a nylon membrane (Amersham) at 50
52 Volts for 30 minutes, and probed using the Light Shift nucleic acid detection module

53 (Thermo) according to the manufacturer's protocol. All gels were imaged using a Bio Rad
54 chemiluminescence imager.
55
56 **Molecular Dynamics Simulations**
57 Five compounds were prepared for molecular dynamics (MD) simulation: B, B-1, B-2, B-3, and
58 B-4. In preparation for MD, parameters for the compounds were obtained using
59 Antechamber(4–6) with the Generalized Amber Force Field. Starting conformations for each
60 compound bound to AP2-EXP were based on predictions from docking. All complexes were
61 prepared using the tleap module of AmberTools(7) with the protein.ff14SB forcefield(8). Each
62 complex was solvated in an octahedral box of TIP3P water with a 10 -Å buffer around the
63 protein complex. Na^+ and Cl^- ions were added to neutralize the protein and achieve
64 physiological conditions. All MD minimizations and simulations were performed using Amber
65 with GPU acceleration(9, 10). First, complexes minimized with 5000 steps each of steepest
66 decent and conjugate gradient minimization with 500 kcal/mol·$Å^2$ restraints on all complex
67 atoms. Restraints were reduced to 100 kcal/mol·$Å^2$ and the minimization protocol was repeated.
68 Restraints were then retained only on the compound for a final minimization step. Following
69 minimization, all complexes were heated from 0 to 300 K using a 100-ps run with constant
70 volume periodic boundaries and 10 kcal/mol·$Å^2$ restraints on all protein and compound atoms.
71 To equilibrate complexes, 10 ns of MD was performed first with 10 kcal/mol·$Å^2$, then with 1
72 kcal/mol·$Å^2$ restraints on protein and compound atoms using the NPT ensemble. With
73 kcal/mol·$Å^2$ restraints retained on complexes, 500 ns production simulations were performed. A
74 2-fs timestep was used and all bonds between heavy atoms and hydrogens were fixed with the
75 SHAKE algorithm(11). A cut-off distance of 10 Å was used to evaluate long-range electrostatics
76 with Particle Mesh Ewald (PME) and for van der Waals forces. The 'strip' and 'trajout'
77 commands of the CPPTRAJ module(12) were used to remove solvent atoms and extract 50,000
78 evenly spaced frames from each simulation for analysis.

79

80 **Analysis of DNA Microarrays**

81 LIMMA(13) was used to normalize and extract signal intensities. Arrays were normalized using
82 robust splines normalization and within arrays parameters selected. Average expression values
83 were calculated per gene by averaging the $\log_2$ Cy5 (cDNA)/Cy3 (Reference Pool) signal
84 intensity across all probes. Correlation plots for the Compound C vs. DMSO control total
85 transcriptome and control genes were made using the $\log_2$ signal intensities for all detected
86 genes using the R package CorrPlot downloaded from: https://github.com/taiyun/corrplot. The
87 LIMMA eBayes function was used on average gene abundance data to determine changes in
88 transcript abundance between 24 and 30hpi that occur differentially in DMSO control vs.
89 Compound C dosed parasites. Volcano plots were made with LIMMA eBayes data using the
90 Enhanced Volcano package in R downloaded from:
91 https://github.com/kevinblighe/EnhancedVolcano. The Rnits(14) R package was used to model
92 differences in gene expression across the entire time course between the DMSO control and
93 Compound C spike in parasites. Rnits was run using the parameters: center genes, normalize
94 by intensity, and background filter probes. The Rnits model for differential expression was fit at
95 the gene level. All data from LIMMA, eBayes, and Rnits is provided in **Table S3**. Heat maps for
96 control genes and AP2-EXP target genes were made using Java Treeview(15) after means
97 centering and log transformation of transcript abundance values with Cluster3.0(16).

98

99 **ChIP-seq Data Analysis**

100 First, reads in each DNA library were trimmed to remove low quality base calls and Illumina
101 adaptor sequences using Trimmomatic(17) with a quality score cutoff of 20. FastQC(18) was
102 used to assess the quality of each DNA library following this step. Reads were then mapped to
103 the *Plasmodium falciparum* genome version 36 downloaded from PlasmoDB using BWA-
104 Mem(19). Multiply mapped reads were filtered out using Samtools(20). Filtered and mapped

105 Input and immunoprecipitate bam files were used for peak calling by MACS2(21) with the
106 following parameters: effective genome size 20000000, No Model, q 0.01. Peaks called in each
107 ChIP-seq replicate were overlapped using Bedtools(22) Intersect, and peaks that occur in a
108 minimum of 2/3 replicates were used for downstream analysis. For data visualization using IGV,
109 .bam files were converted to $\log_2$ Immunoprecipitate fraction/Input fraction bigwigs using
110 DeepTools(23) BamCompare. Bedtools(22) ClosestBed was used to correlate regions called as
111 peaks of occupancy by MACS2 to the closest gene. Results were filtered based on 1.5kb
112 proximity to the MACS2 peak and strandedness of the putative target gene. For comparison of
113 ChIP-seq library coverages to Transcription Start Sites and Start Codons, DeepTools
114 PlotHeatmap was used. Matrix files underlying PlotHeatmap were created using DeepTools
115 ComputeMatrix. A four-color plot of conserved AP2-EXP DNA binding sites was made using the
116 program cegr-tools four color plot downloaded from https://github.com/seqcode/cegr-
117 tools/tree/master/src/org/seqcode/cegrtools. DNA motifs enriched in AP2-EXP ChIP-seq peaks
118 were determined using the DREME Suite(24). DNA motifs were sorted from highest to lowest
119 level of conservation to the aggregate AP2-EXP DNA motif using FIMO(25).
120
121 For comparison of AP2-EXP peaks with existing post-translational histone marks, chromatin
122 post translational modification datasets(26, 27) were downloaded from
123 https://github.com/Daread/plasmodiumExprPrediction(28). Chromatin reader and nucleosome
124 occupancy datasets(29–31) were downloaded from NCBI. ChIP-seq data for BDP1 and HP1
125 was processed as described above. Nucleosome occupancy data was downloaded from NCBI
126 and processed as described above, with the exception that bam files were turned into bigwig
127 using BamtoBigWig with the --mnase parameter selected. DeepTools ComputeMatrix was used
128 to format the data. DeepTools PlotHeatmap was used to plot coverage of each dataset within
129 10kb of AP2-EXP peaks. For negative control coverage plots, Bedtools(22) ShuffleBed was

130 used to create random genomic intervals on the same chromosome and the same length as

131 AP2-EXP peaks of occupancy.

132

133 For comparison of AP2-EXP peaks identified by Shang *et al*(32) to this study, AP2-EXP peaks

134 of occupancy in the trophozoite stage were downloaded from

135 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE184658 . Peak summits from each

136 replicate were expanded to 200 nucleotide windows using Bedtools SlopBed(22). Conserved

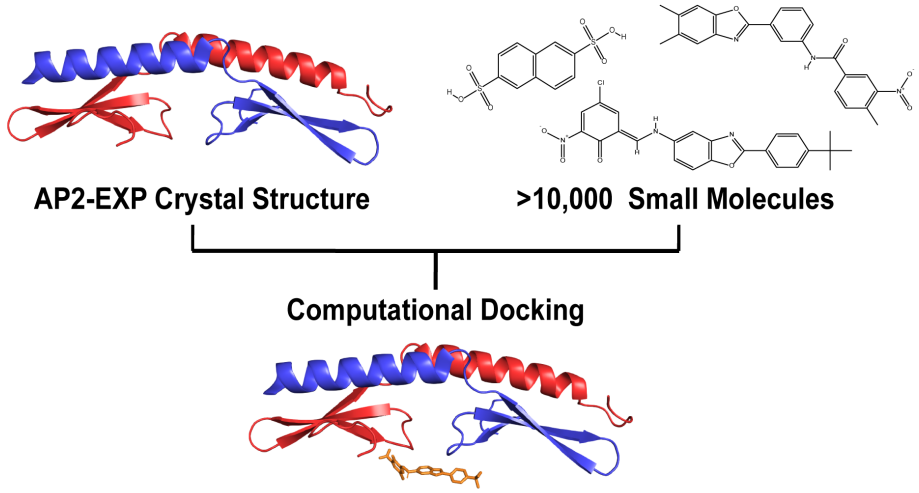
137 peaks in 2/2 replicates were identified using Bedtools intersect(22), then AP2-EXP peaks from

138 this study and Shang *et al* were overlapped using Bedtools Intersect(22).

139

140 **SI Figures**

A



B

Letter Code	Compound Used Originally for *in silico* Docking	Compound Name Used for Purchase	PubChem ID of Compound used in this study
A	GSK-like (TCMDC-124220)	CB5768506*	CID1365835
B	GSK (TCMDC-123924)	CB5842949	CID1365471
C	GSK-like (TCMDC-124220)	ChemDiv-8002-1285*	CID5750730
D	GSK-like (TCMDC-124220)	ChemDiv-8004-0752*	CID4541005
E	Drug Bank (DB00562)	Benzathiazide	CID2343
F	Drug Bank (DB04409)	Napthalene Trisulfonate	CID4437
G	Drug Bank (DB04640)	2,6 Napthalene Sulfonate	CID11390
H	Drug Bank (DB01219)	Dantrolene	CID6914273
I	Drug Bank (DB02633)	Procion Blue	CID25863

141

142 **Figure S1. Putative AP2-EXP competitors were identified using computational docking**

143 A) The crystal structure of AP2-EXP (PDB:3IGM)(33) was used as a template to

144 computationally dock thousands of small molecules *in silico* using AutoDock. Results were

filtered for compounds that dock within 10 Angstroms of DNA binding residues with a free energy less than -5kJ/mol. Compounds matching these criteria were sourced and used for further testing.

B) Seven compounds were identified as putative ApiAP2 competitors in an *in-silico* screen (Column 2). Five of these were available for direct purchase (Column 3). For the two remaining compounds, four alternate choices with a Tanimoto similarity score of .9 or greater were purchased (denoted by an asterisk in Column 3). The PubChem ID used to purchase each compound in this study is listed in Column 4.

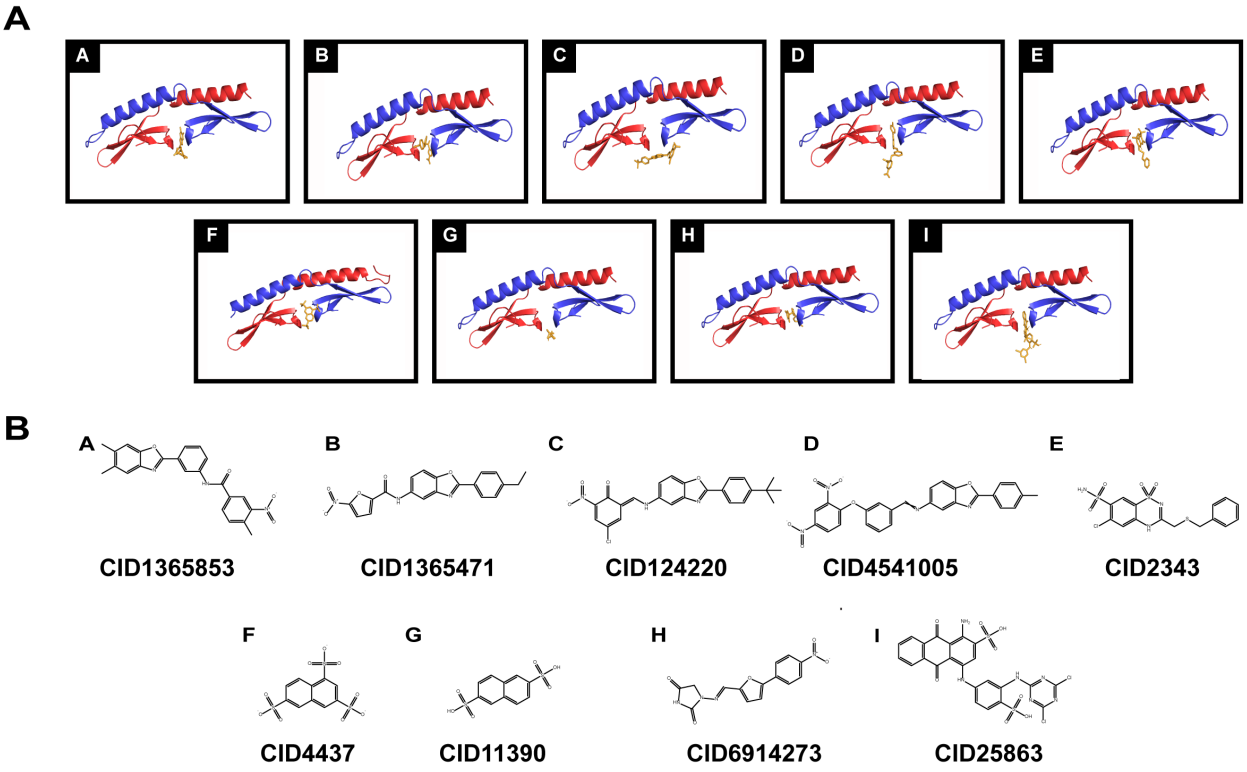


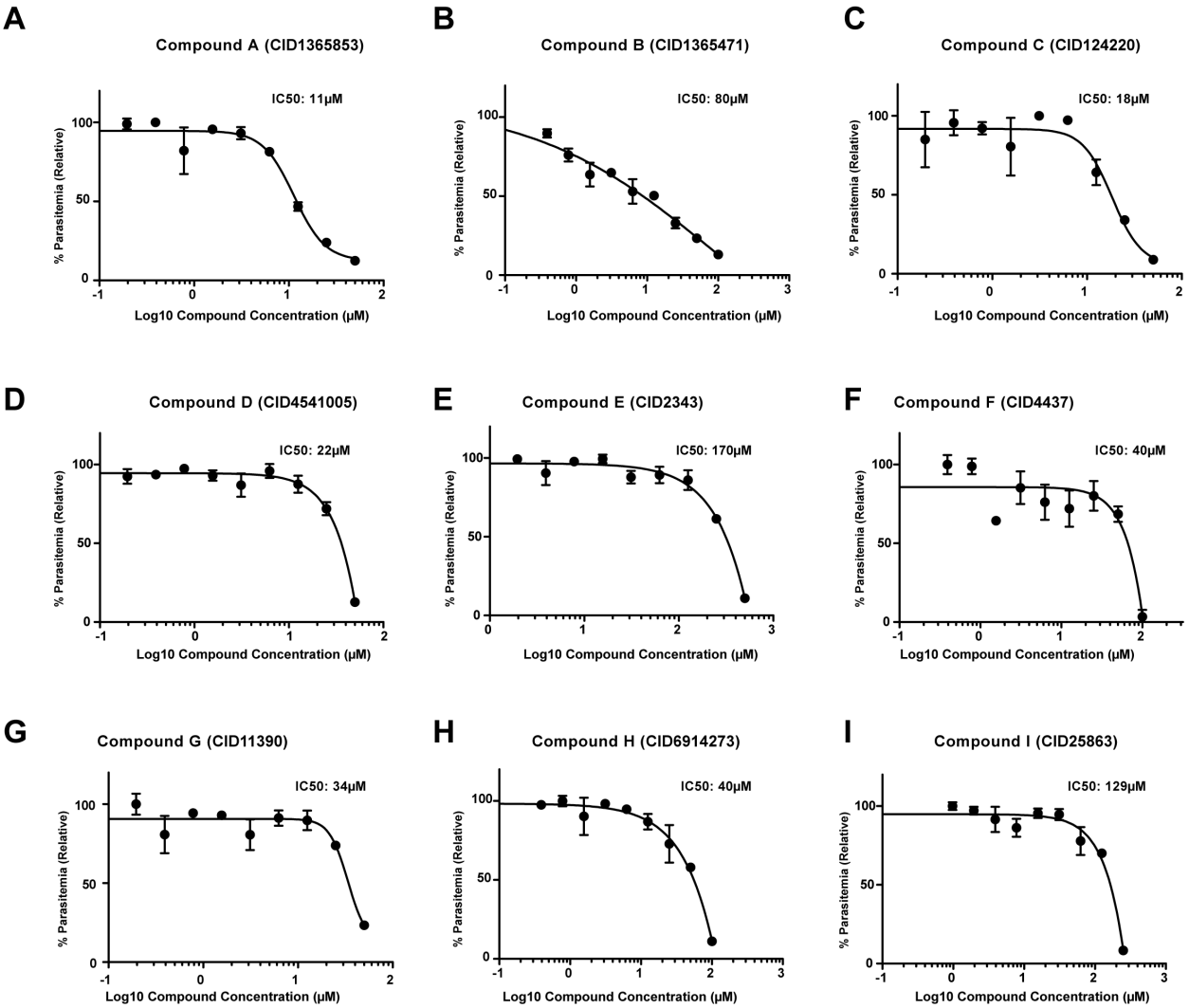
Figure S2. Docking conformations for Compounds A-I

A) The spatial conformation of each of the nine compounds (A-I) that dock within 10 Angstroms of the DNA binding pocket of AP2-EXP with a free energy less than -5kJ/mol is depicted above.

160 B) Chemical structures of each compound (A-I) corresponding to the molecular docking results

161 in panel A.

162



163

164 **Figure S3. IC50 assays for Compounds A-I**

165 A-I) 48-hour Sybr Green growth inhibition assays were conducted for each of the nine putative

166 ApiAP2 competitor compounds in order to determine IC50 values against asexual *P. falciparum*.

167 All growth assays were performed in triplicate. All compounds kill asexual stage *P. falciparum*

168 parasites in the micromolar (11-170μM) concentration range.

169

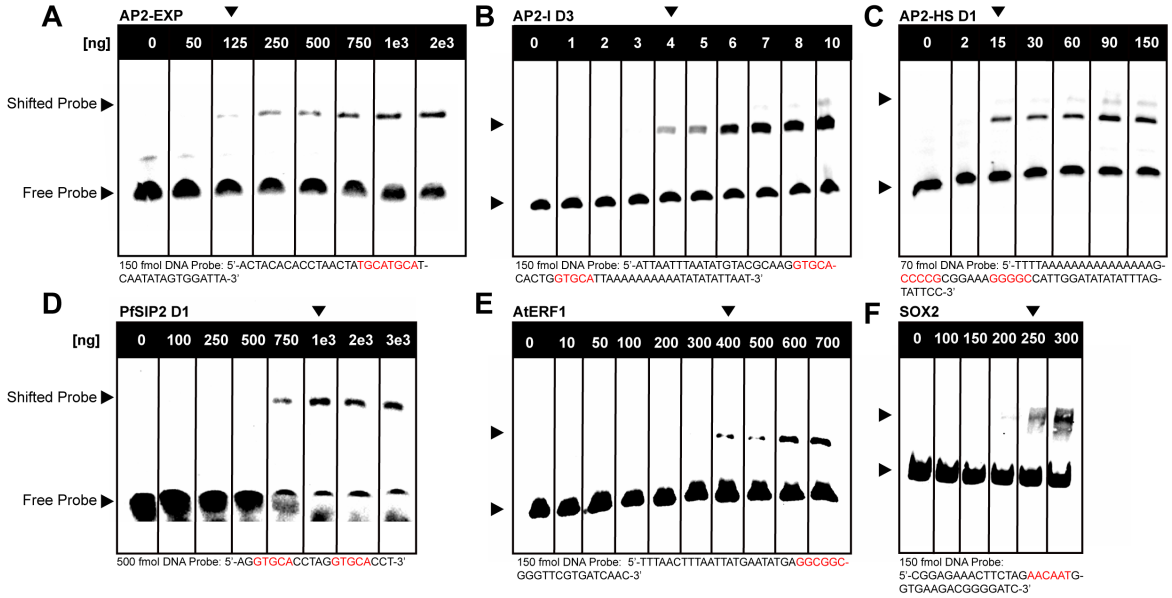


Figure S4. Titration of recombinant DNA binding domains to optimize competition electrophoretic mobility shift (EMSA) assays

A-F) DNA binding domains AP2-EXP, AP2-I D3, AP2-HS D1, PfSIP2 D1, AtERF1, and full length SOX2 were titrated against DNA oligos containing their respective binding motifs (highlighted in red) in an EMSA. Unless otherwise specified, the minimum mass of each recombinant DNA binding domain required to visualize DNA binding (denoted by an arrow) was used in competitive EMSAs with putative ApiAP2 competitor compounds.

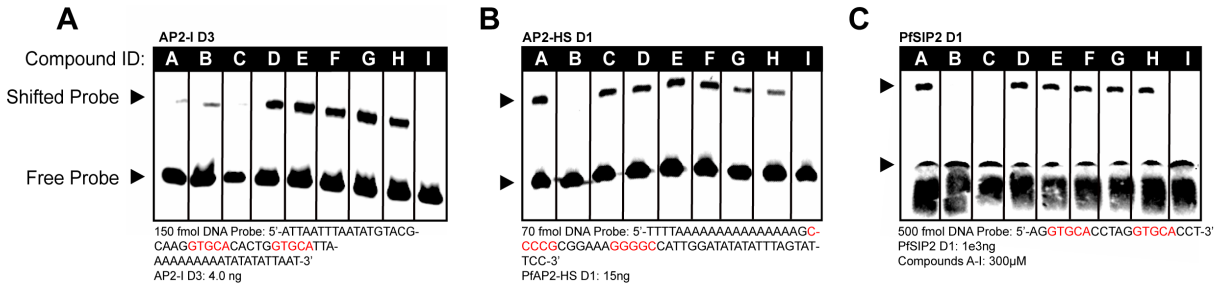
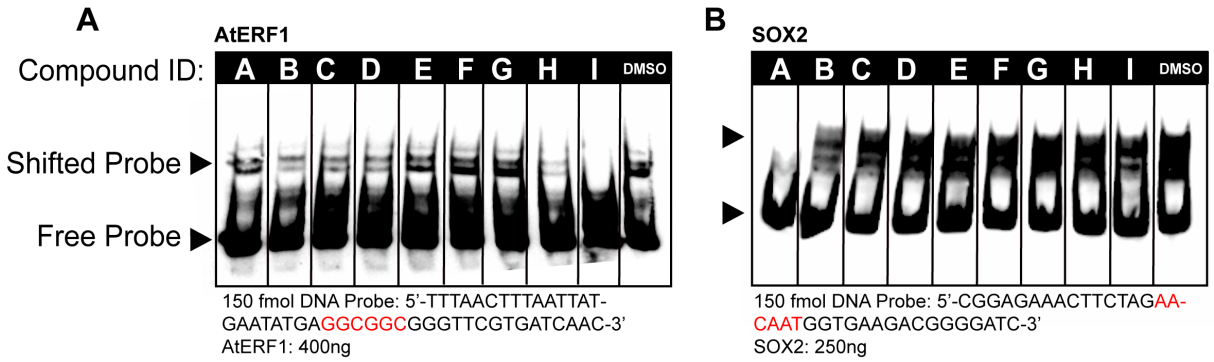


Figure S5. Putative ApiAP2 competitor compounds were tested against three asexual blood stage essential *P. falciparum* AP2 domains in addition to AP2-EXP in a competitive EMSA

183 AP2-I D3 (A) DNA binding activity is competed by Compounds A, B, C and I. AP2-HS D1 (B) is
184 competed by Compounds B and I, and PfSIP2 D1 (C) is competed by Compounds B, C, and I.
185 Compounds A, B, C and I all compete at least one AP2 domain in addition to AP2-
186 EXP. Cognate DNA motifs for each protein are highlighted in red. 300μM of each compound
187 was used per lane.
188



189

190 **Figure S6. Putative ApiAP2 competitor compounds were tested against off target DNA**
191 **binding domains in a competitive EMSA**
192 A) The plant encoded *Arabidopsis thaliana* AP2 domain from Ethylene Response Factor 1
193 (AtERF1) is competed by Compound I. The *Plasmodium* AP2 domain competitors Compound A,
194 B, and C do not compete AtERF1. The cognate AtERF1 DNA motif is highlighted in red. 300μM
195 of each compound was used per lane.
196
197 B) The human encoded High Mobility Group Box Domain transcription factor SOX2 is competed
198 by Compound A. Due to the lack of homology between SOX2 and AP2 domain proteins, this
199 result indicates that Compound A's DNA binding competition activity is not unique to the AP2
200 domain. The cognate SOX2 DNA motif is highlighted in red. 300μM of each compound was
201 used per lane.
202

Letter Code	AP2-EXP	AP2-I D3	PfSIP2 D1	AP2-HS D1	AtERF1	SOX2
A	+	+		-	-	+
B	+	+	+	+	-	-
C	+	+	+	-	-	-
D	-	-		-	-	-
E	-	-		-	-	-
F	-	-		-	-	-
G	-	-		-	-	-
H	-	-		-	-	-
I	+	+	+	+	+	-

Table S2. Summary of the Competitive EMSA results for each DNA binding domain tested

Compounds A, B, C, and I all can compete DNA binding by *Plasmodium* AP2 domains. Compounds A and I each compete one off-target protein domain.

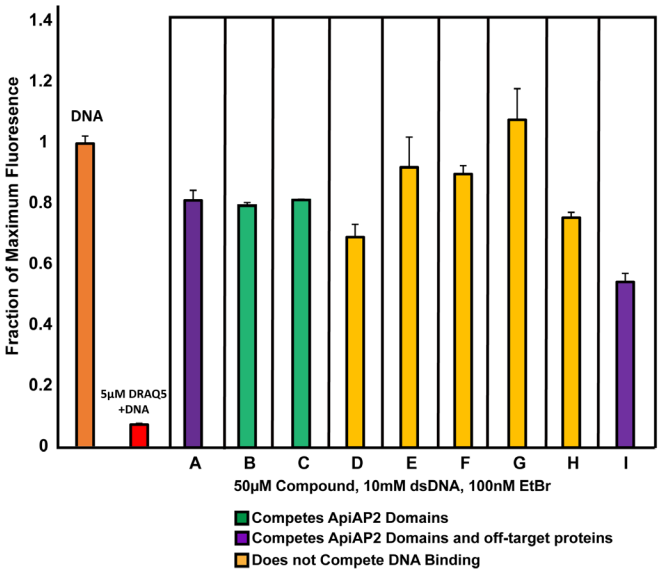


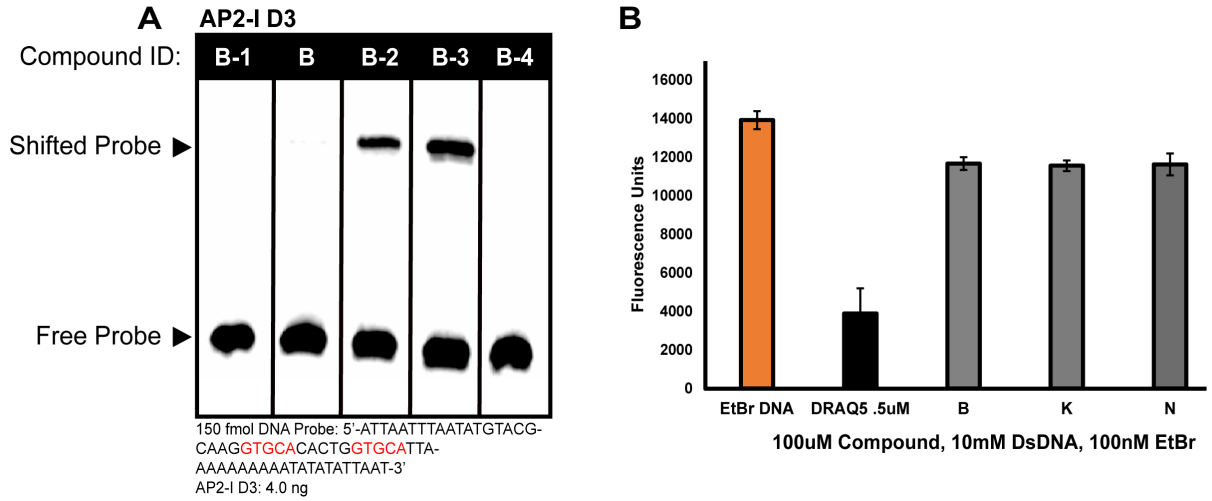
Figure S7. Compounds A-I were tested for DNA intercalation in an ethidium bromide exclusion assay

Each putative ApiAP2 competitor compound was added into a mixture containing double stranded DNA and ethidium bromide. The positive control DNA major groove intercalator DRAQ5 knocks down ethidium bromide fluorescence nearly completely relative to the DNA and ethidium bromide control. The legend indicates the cumulative result for each compound in

216 competitive EMSAs. Each assay was performed in triplicate. Error bars represent standard

217 deviation of the mean.

218



219

220 **Figure S8. Compound B analogues were tested against AP2-I D3 in a competitive EMSA**

221 **and checked for DNA intercalation ability**

222 A) Compounds B, B-1, B-2, B-3, and B-4 were added to an EMSA with AP2-I D3 to check

223 whether their DNA binding competition is consistent with AP2-EXP. The cognate AP2-I D3 DNA

224 motif is highlighted in red. 300μM of each compound was used per lane.

225

226 B) Compounds B, B-1, B-2, B-3, and B-4 were tested for DNA major groove intercalation in an

227 ethidium bromide exclusion assay. DRAQ5 was used as a positive control for intercalation.

228 Each assay was performed in triplicate. Error bars represent standard deviation of the mean.

229

230 **Movies S1-S5 Molecular dynamics simulations for Compound B and Compound B**

231 **analogues B-1, B-2, B-3, and B-4 interaction with AP2-EXP**

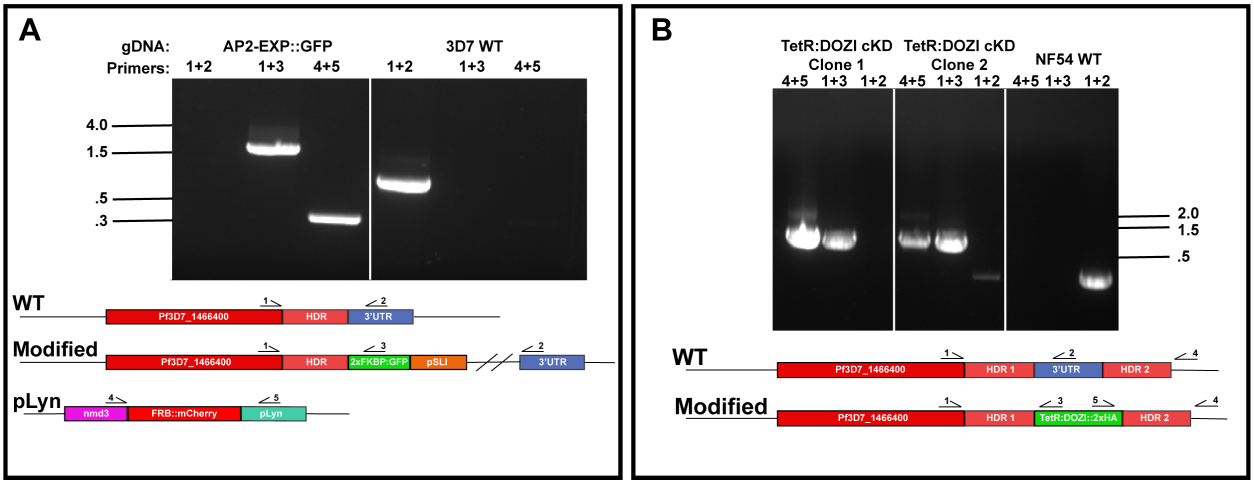
232 Compound B (Movie S1), Compound B-1 (Movie S2), Compound B-2 (Movie S3), Compound B-

233 3 (Movie S4), and Compound B-4 (Movie S5) interactions with AP2-EXP were each simulated

234 using molecular dynamics. Each molecule was initially started in the location predicted for

235 Compound B by molecular docking.

236



237

238 **Figure S9. Creation of endogenously tagged parasite lines AP2-EXP::GFP and AP2-**

239 **EXP::HA**

240 A) The Selection Linked Integration system was used to add 2xFKBP inducible mislocalization

241 protein and GFP to AP2-EXP by single homologous recombination. Successful integration was

242 confirmed by genotyping PCR. Genomic DNA from the wild type Pf3D7 parental line was used

243 as a control. In order to test the efficacy of the knock sideways system, the pLyn mislocalizer

244 plasmid was added to AP2-EXP::GFP and confirmed by PCR. DNA kb are indicated by the

245 marks to the left of the gel.

246

247 B) The PSN054 TetR:DOZI plasmid was used to add the TetR:DOZI mRNA repression module

248 and endogenous 2xHA tag to AP2-EXP by double homologous recombination. Correct

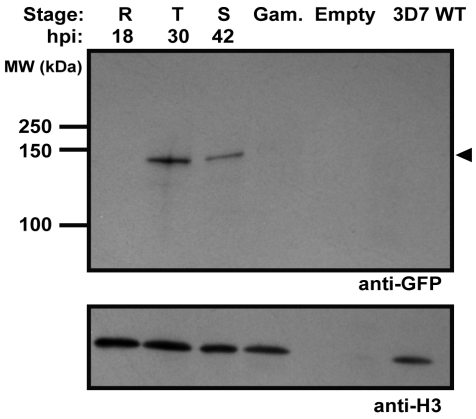
249 integration was confirmed by genotyping PCR. The parental NF54 parasite line was used as the

250 unedited control. Clonal populations one and two are indicated as C1 and C2, respectively.

251 Clone one had correct integration and complete absence of the wild type *ap2-exp* DNA locus

252 and was used for further experiments. DNA kb are indicated by the marks to the right of the gel.

253



254

255 **Figure S10. AP2-EXP protein expression in the AP2-EXP::GFP endogenously tagged**

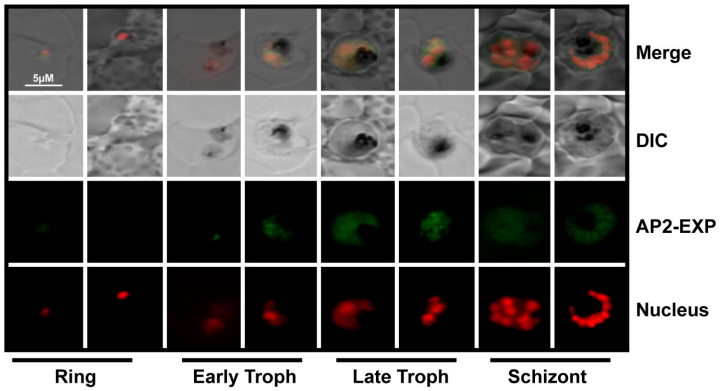
256 **parasite line (related to figure 3A)**

257 AP2-EXP expression was tracked throughout the IDC by harvesting protein from highly

258 synchronous asexual blood stage parasites followed by a western blot against the GFP tag.

259 Histone H3 was used as a loading control.

260



261

262 **Figure S11. Fluorescent microscopy tracking of AP2-EXP::GFP expression**

263 AP2-EXP::GFP expression was monitored in a highly synchronous parasite population by

264 fluorescent microscopy across the IDC. DRAQ5 was used as a nuclear stain for parasites.

265

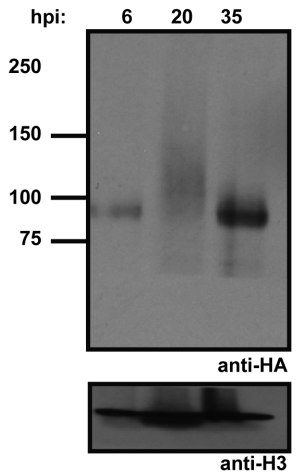


Figure S12. AP2-EXP protein expression in the AP2-EXP::HA endogenously tagged parasite line

AP2-EXP expression was tracked throughout the IDC by harvesting protein from highly synchronous asexual blood stage parasites followed by a western blot against the 2xHA tag. Histone H3 was used as a loading control.

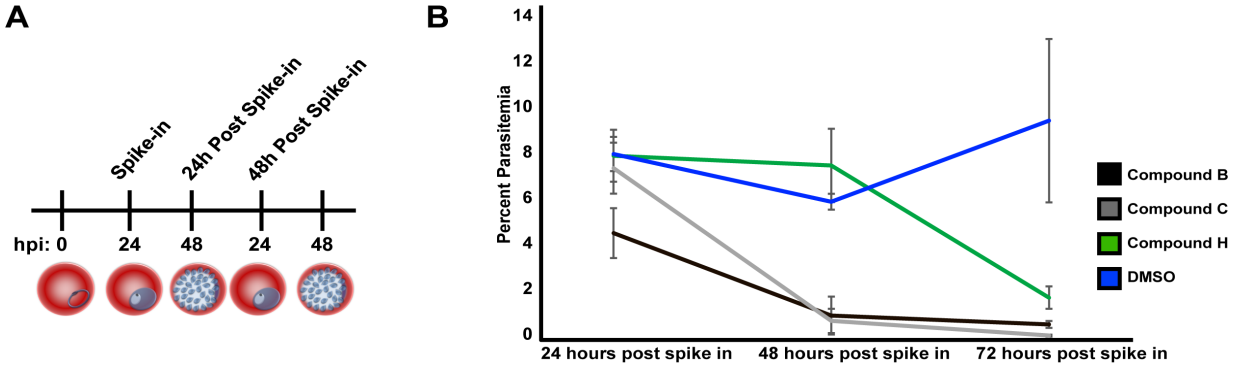


Figure S13. Parasite counts for 48-hour Compounds B, C, and H phenotyping time course, related to figure 3B

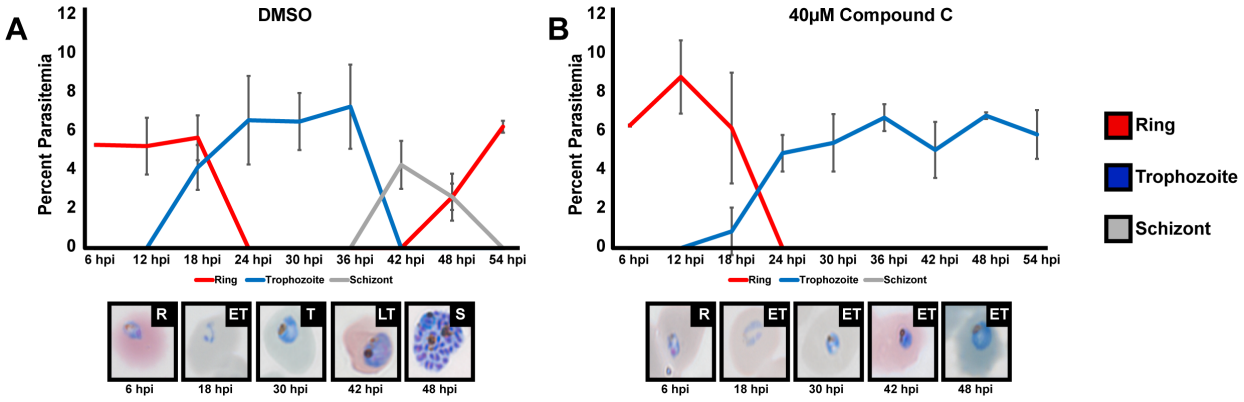
A) Schematic for AP2 domain competitor compound phenotyping time course.

B) 40μM of ApiAP2 competitor Compounds B or C, or non ApiAP2 competitor Compound H, was spiked into highly synchronous wild type Pf3D7 parasites at 24 hours post invasion. DMSO

280 vehicle was used as a reference for normal growth. All growth assays were performed in

281 triplicate. Error bars represent standard deviation of the mean.

282



283

284 **Figure S14. Compound C causes parasites to stop progressing in the early trophozoite**

285 **stage at 30 hpi**

286 Highly synchronous wild type Pf3D7 parasites were spiked with DMSO vehicle control (A) or

287 40μM Compound C (B) and monitored at 6-hour intervals throughout the IDC. Each growth

288 assay was performed in triplicate. Error bars represent standard deviation of the mean. Giemsa

289 stained parasite images are representative of each population at the specified time point. R, ET,

290 T, LT, and S indicate ring, early trophozoite, late trophozoite, or schizont morphologies,

291 respectively.

292

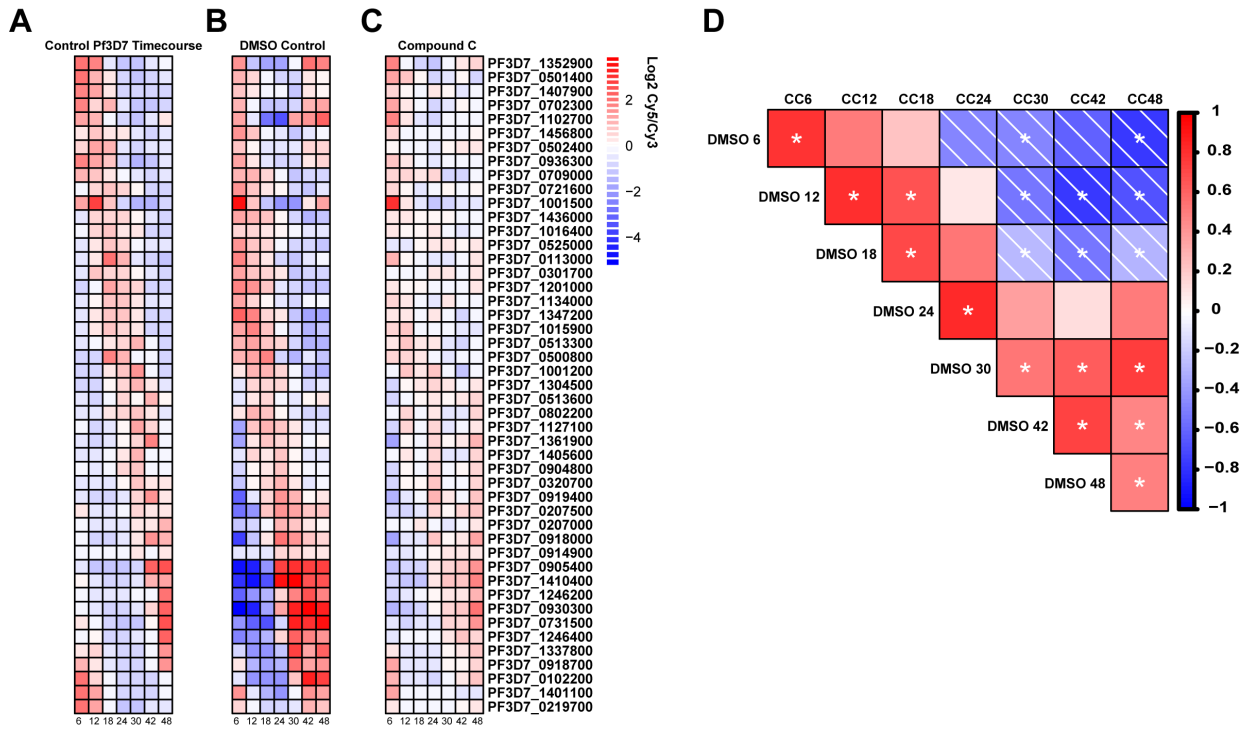


Figure S15. Quality control of DNA microarray data for DMSO vehicle control and Compound C parasites, related to figure 4

A) DNA microarray data from(34) for a set of highly periodic control genes expressed in the IDC.

B-C) The same set of highly periodic control genes as in panel A were plotted for the DMSO control and 12μM Compound C spiked parasites in order to compare parasite staging between the two experiments.

D) Correlogram depicting the Spearman Correlation value between control gene expression for DMSO (Panel B) and Compound C (Panel C) samples. A * indicates p value < .05.

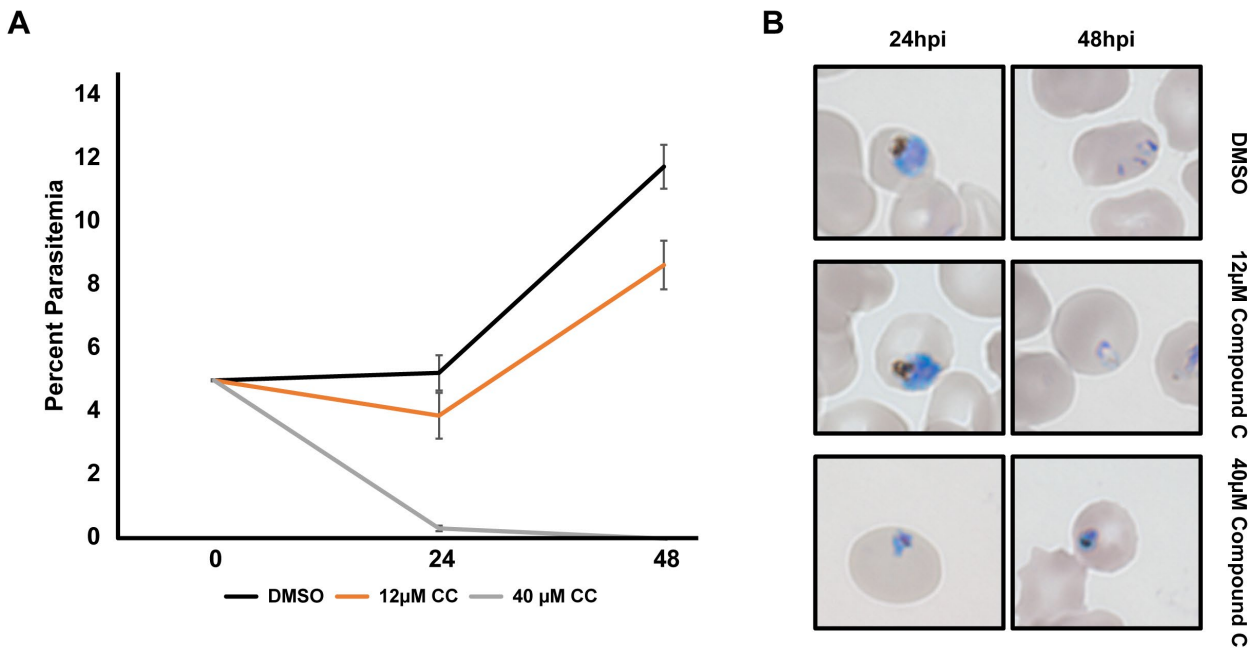


Figure S16. A 48-hour time course to determine the phenotype for 12μM Compound C, related to Figure 4A

A) Highly synchronous asexual blood stage Pf3D7 parasites were spiked with DMSO vehicle control, 12μM Compound C, or 40μM Compound C. 40μM Compound C was used as a positive control for complete parasite death at 48 hpi. Each growth assay was performed in triplicate. Error bars represent standard deviation of the mean.

B) Representative images of each parasite population (DMSO, 12μM Compound C, 40μM Compound C) at 24 and 48 hpi.

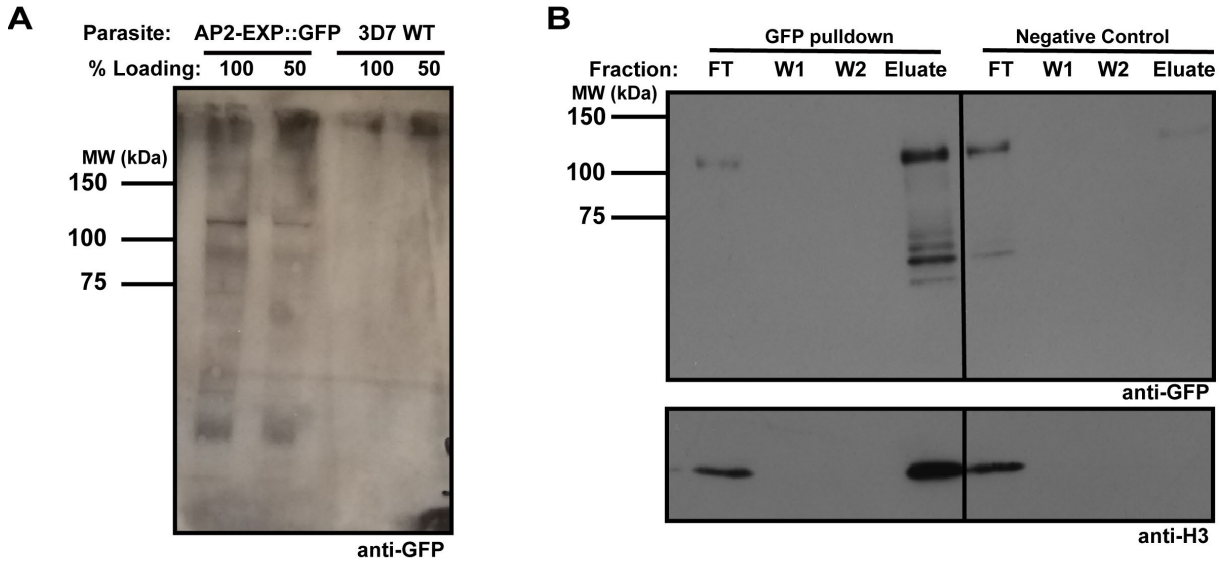


Figure S17. ChIP-seq protein quality control, related to figure 4B

A) Crosslinked nuclear material was blotted after sonication to ensure recovery of the full-length AP2-EXP protein during chromatin immunoprecipitation. Full length AP2-EXP is recovered, indicating that the protocol is suitable to analyze AP2-EXP DNA binding *in vivo*. Crosslinked nuclear material from the wildtype Pf3D7 parental parasite line was used as a negative control.

B) Anti-GFP beads were used to pull down GFP tagged AP2-EXP from AP2-EXP::GFP. Flowthrough (FT), Wash (W1 and W2) and Eluate fractions were saved and analyzed by western blot. The presence of AP2-EXP and Histone H3 in the Eluate lane indicates that AP2-EXP interacts with chromatin in the nucleus. The non-immune negative control beads do not enrich AP2-EXP or Histone H3 in the eluate.

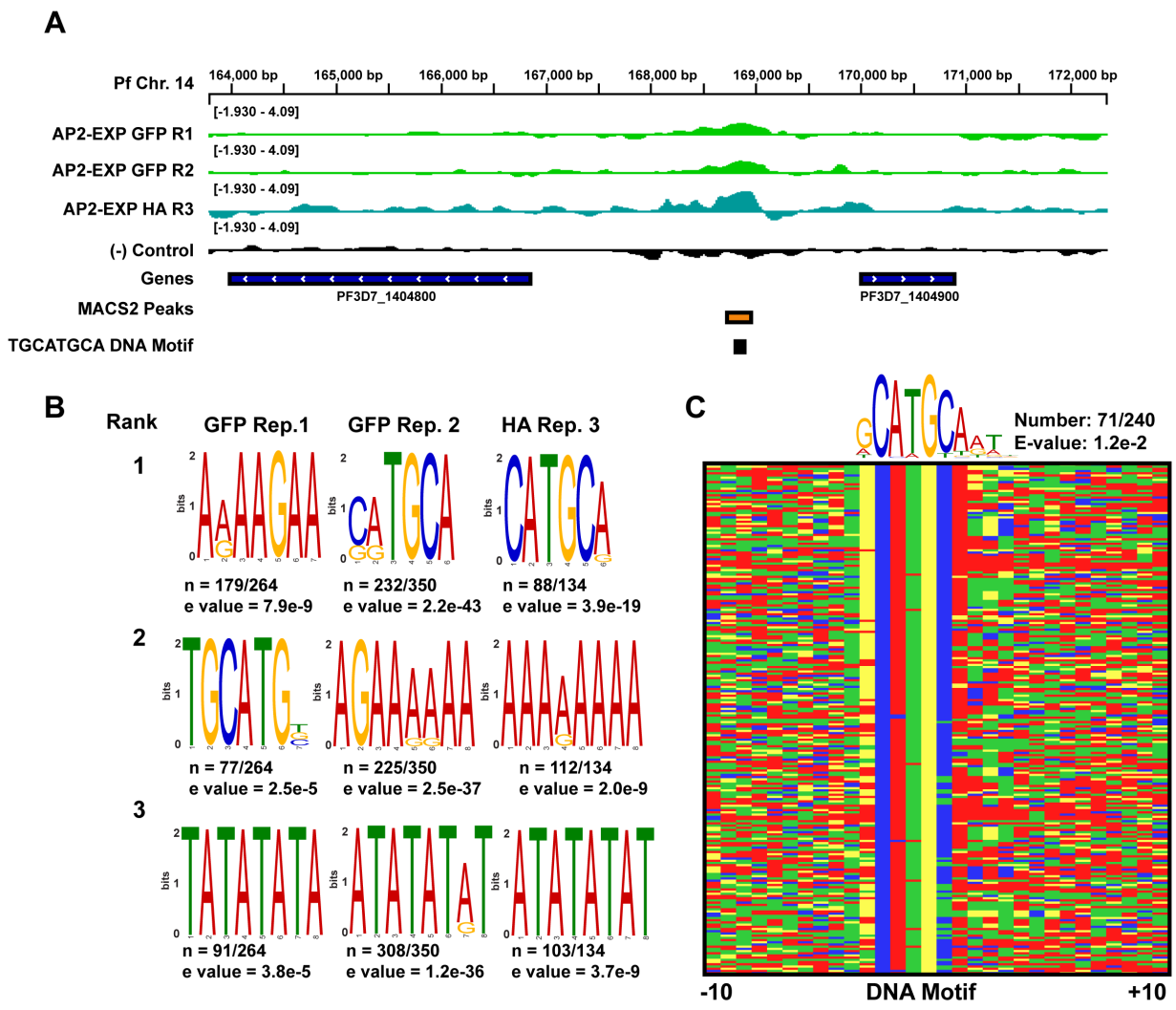


Figure S18. ChIP-seq extended data

A) Log2 immunoprecipitate/input ChIP-seq data from each replicate (2x AP2-EXP::GFP and 1xAP2-EXP::HA) of AP2-EXP ChIP-seq visualized by IGV at a representative DNA locus. The location of a conserved MACS2 called peak of occupancy and the TGCATGCA DNA motif is indicated by the bottom tracks. The (-) control lane is the coverage resulting from a no-epitope control ChIP-seq.

B) The top three ranked DNA motifs present within peaks of occupancy for each ChIP-seq replicate as determined by DREME(24). The core DNA motif CATGCA is overrepresented within each individual replicate.

C) The top overrepresented DNA motif within AP2-EXP peaks of occupancy conserved in 2/3 replicates of ChIP-seq as determined by DREME(24) plotted at the primary DNA sequence level. DNA sequences were sorted from highest to lowest degree of motif conservation using FIMO.(25)

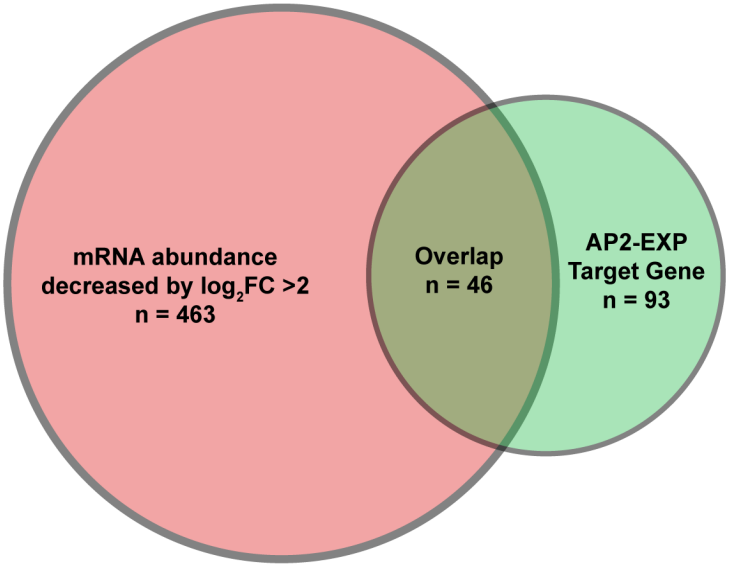


Figure S19. Comparison of AP2-EXP target genes with Compound C induced changes in transcript abundance

The total overlap between AP2-EXP gene targets detected in the Compound C RNA time course and global decrease in transcript abundance at 24-30 hpi by log_2 fold change >2.

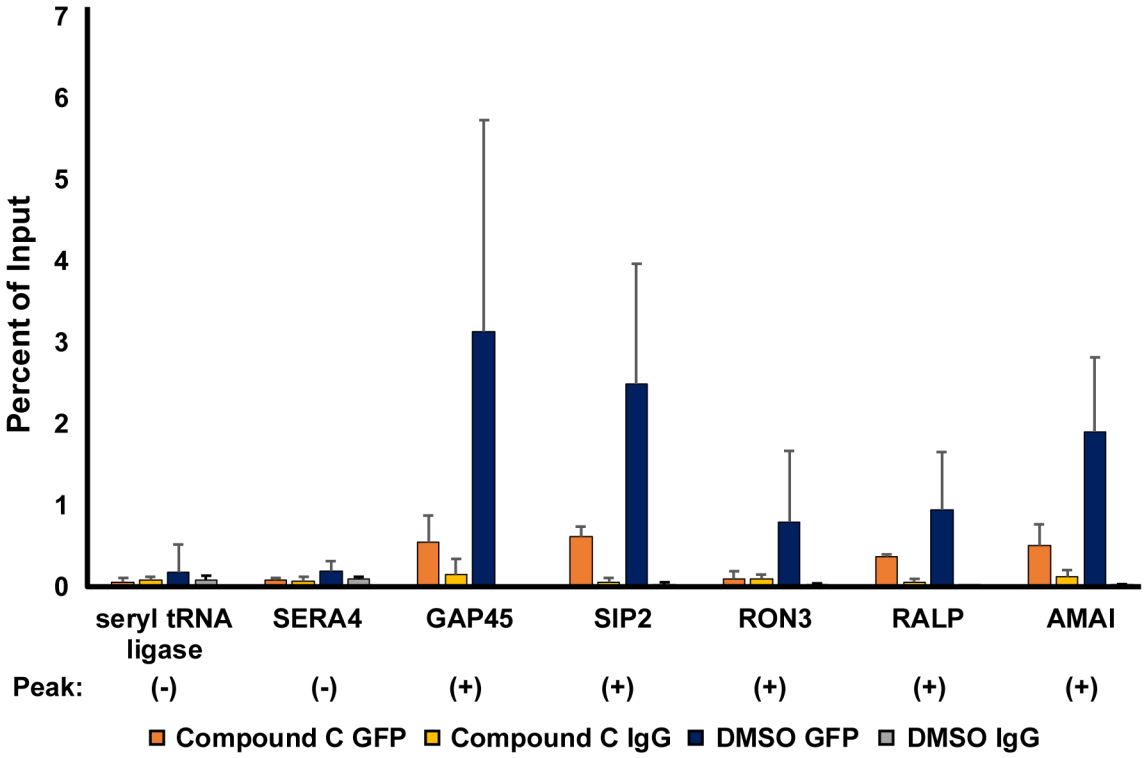


Figure S20. ChIP-Quantitative PCR to assess Compound C impact on AP2-EXP genomic occupancy

AP2-EXP::GFP parasites were spiked with 40μM Compound C or DMSO vehicle control at 30 hpi for two hours. ChIP samples were collected for each population using either anti-GFP or negative control IgG antibodies. The percent of input was determined by RT-qPCR. Each experiment was performed in triplicate. The presence or absence of an AP2-EXP peak of occupancy at each DNA locus based on ChIP-seq is indicated by a (+) or (-), respectively. Each assay was done in triplicate with the exception of AMAI, where n =2.

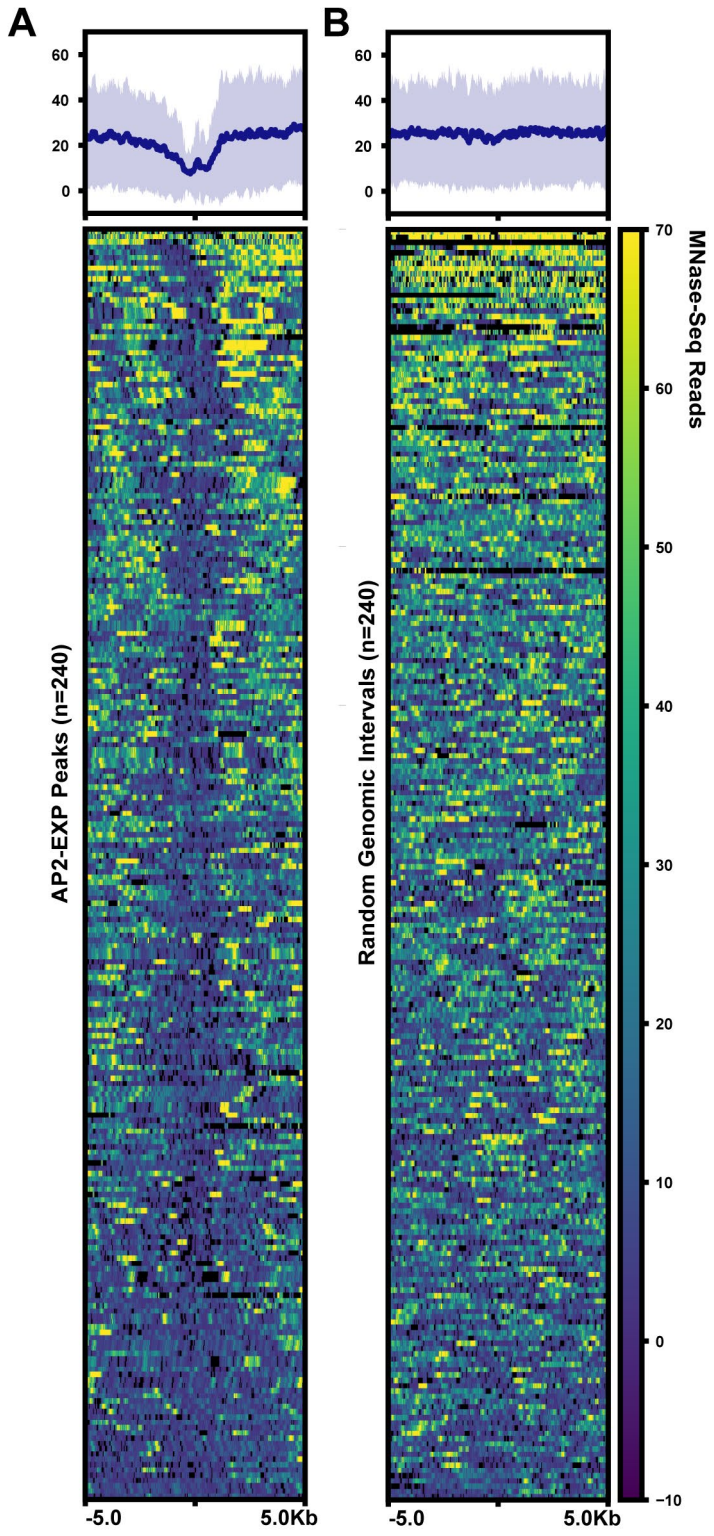
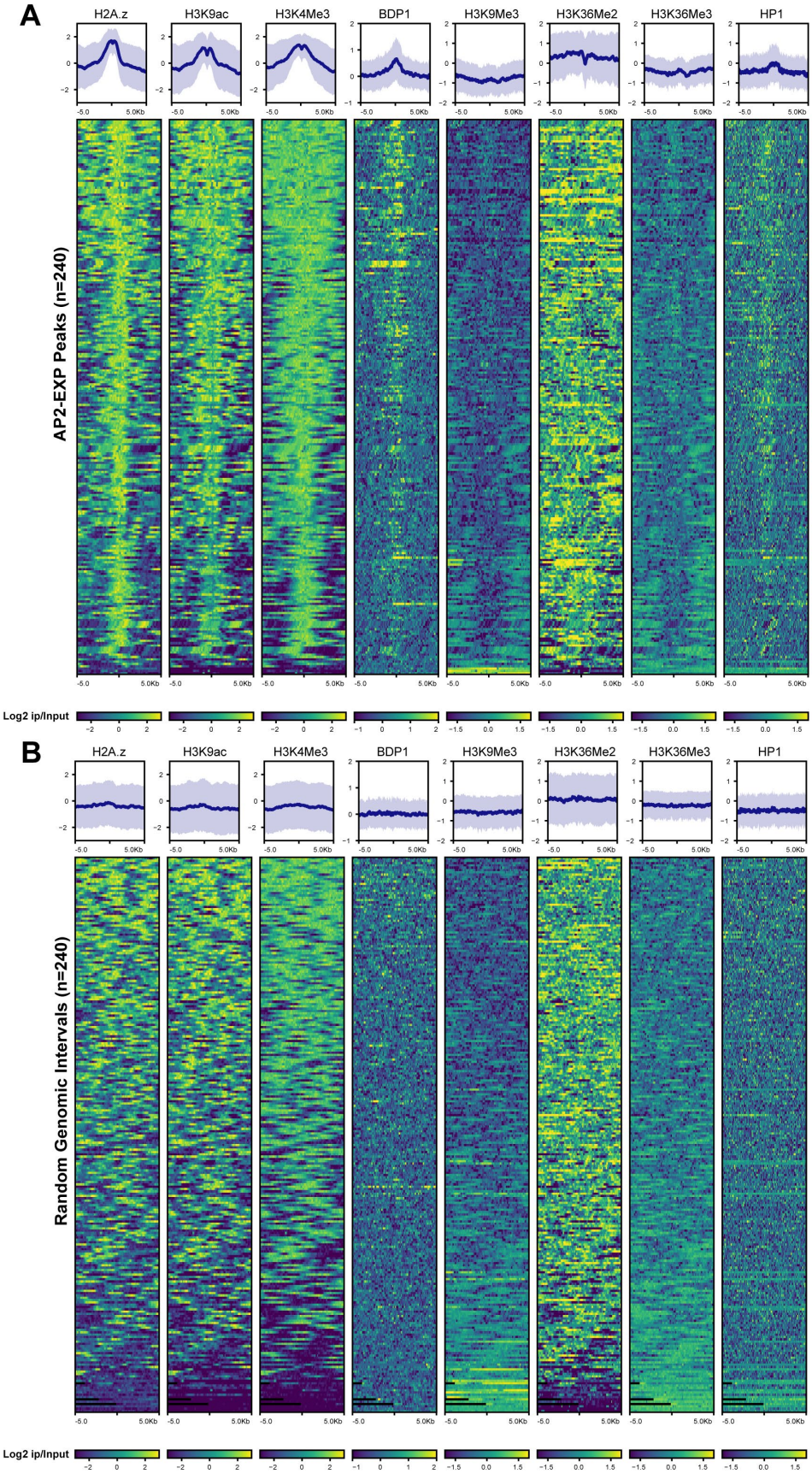


Figure S21. Nucleosome occupancy is depleted at AP2-EXP DNA binding sites

362 Mnase-seq data(30) was plotted against DNA binding sites conserved in 2/3 replicates of AP2-
363 EXP ChIP-seq (A) or random genomic intervals of equal length from the same chromosome as
364 the original peak (B).
365


A
H2A.z
H3K9ac
H3K4Me3
BDP1
H3K9Me3
H3K36Me2
H3K36Me3
HP1
AP2-EXP Peaks (n=240)
Log2 ip/Input
B
Random Genomic Intervals (n=240)
-5.0
5.0Kb

366

367 **Figure S22. Histone post translational modifications and chromatin reader occupancy at**

368 **AP2-EXP peaks**

369 The occupancy of histone variant H2A.z, histone modifications H3K9ac, H3K4me3(27),

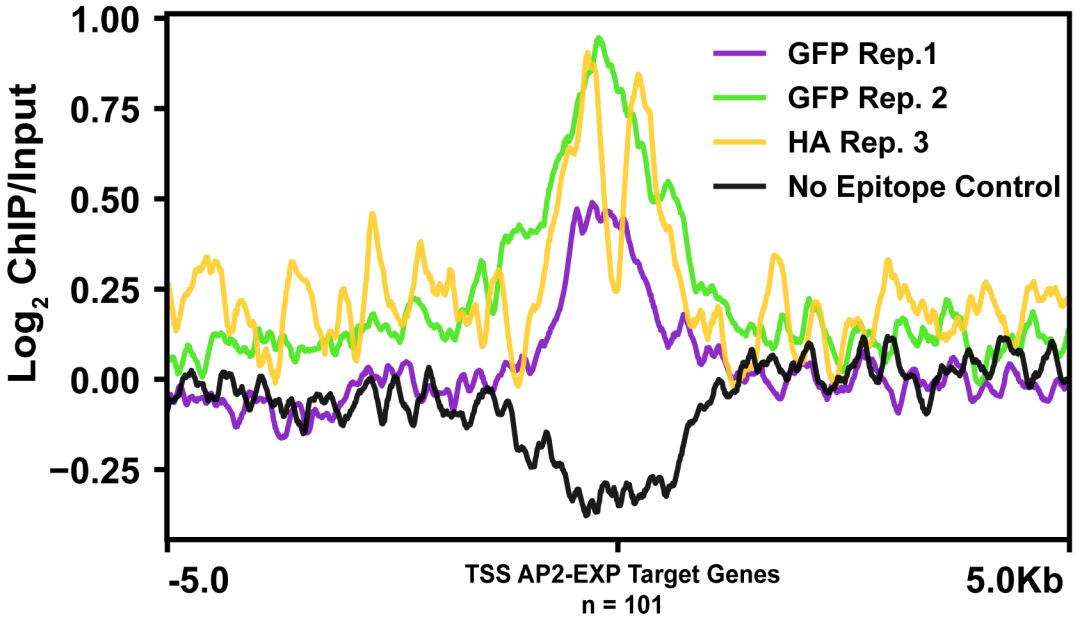
370 H3K9me3, H3K36me2/3(26), and chromatin readers BDP1(29) and HP1(31) were plotted

371 against AP2-EXP peaks of occupancy conserved in 2/3 ChIP-seq replicates (A) or random

372 genomic intervals of the same length (B), taken from the same chromosome on which the AP2-

373 EXP peak originally occurred.

374



375

376 **Figure S23. AP2-EXP DNA occupancy with respect to the Transcription Start Site (TSS) of**

377 **target genes**

378 Log_2 immunoprecipitate (ChIP)/Input ChIP-seq coverage for each replicate of AP2-EXP ChIP-

379 seq and the no epitope control was plotted against the TSS(35) of each target gene conserved

380 in 2/3 ChIP-seq replicates.

381

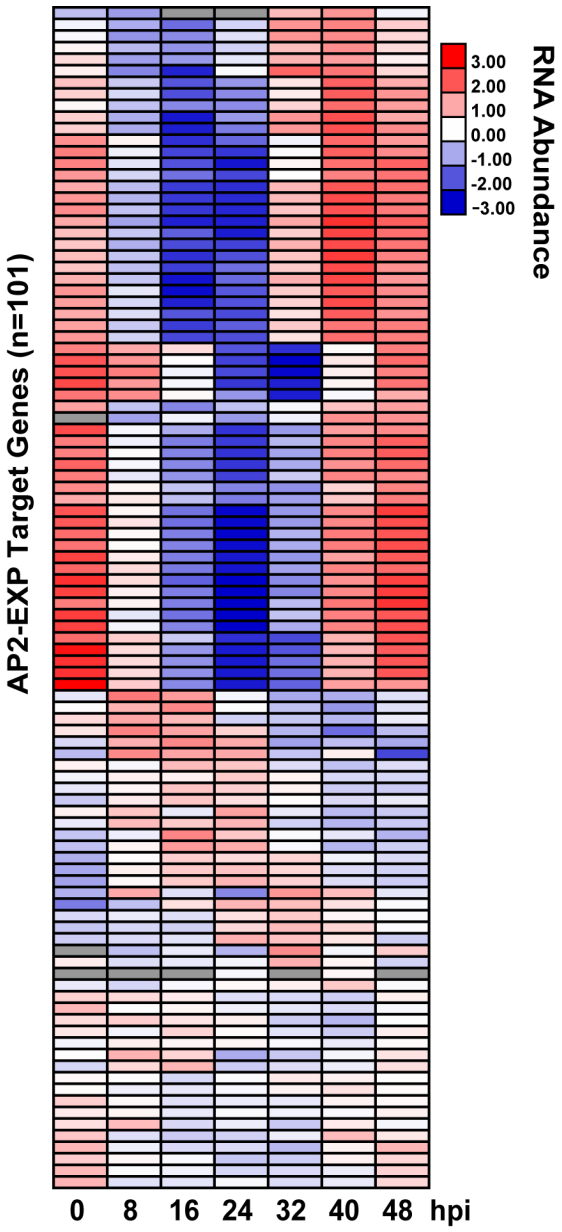


Figure S24. Normal Transcript Abundance of AP2-EXP Target Genes

AP2-EXP target genes were determined by ChIP-seq and their transcript abundance data during the 48-hour IDC was plotted using data from Chappell *et al*(35).

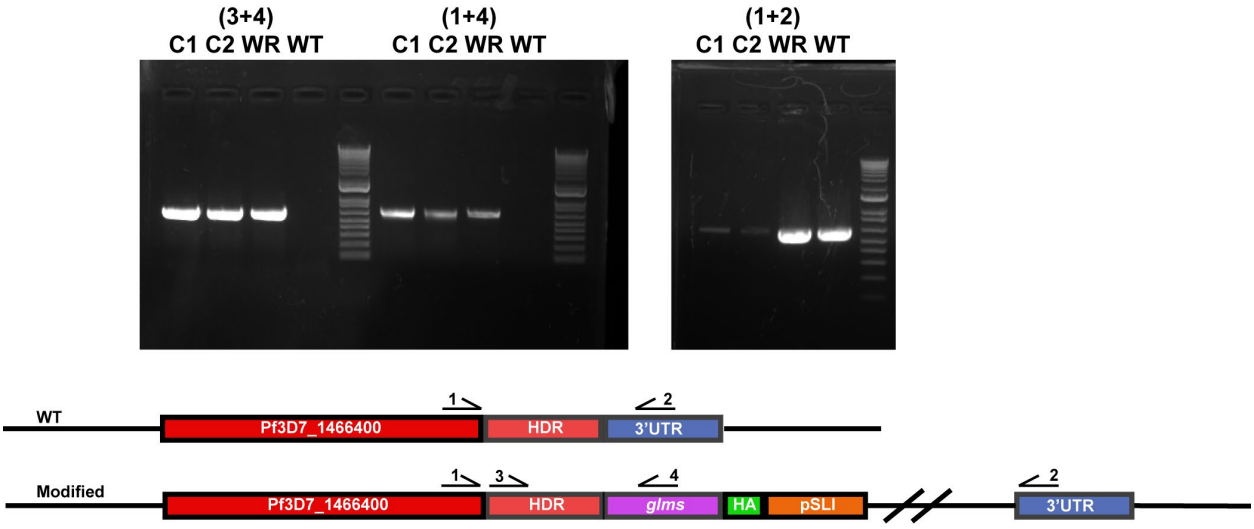


Figure S25. Creation of a *glms* ribozyme based knockdown line for AP2-EXP

Wildtype Pf3D7 *Plasmodium falciparum* parasites were transfected with pSLI:AP2-EXP*::glms*::HA to endogenously tag the AP2-EXP DNA locus with an inducible *glms* ribozyme and HA epitope tag. Successful integration to create AP2-EXP::*glms*::HA by single crossover homologous recombination was confirmed by genotyping PCR. C1 and C2 represent clonal populations selected for integration. WR represents a parasite population selected for the plasmid but not for integration. WT represents Pf3D7 wild type control parasite gDNA.

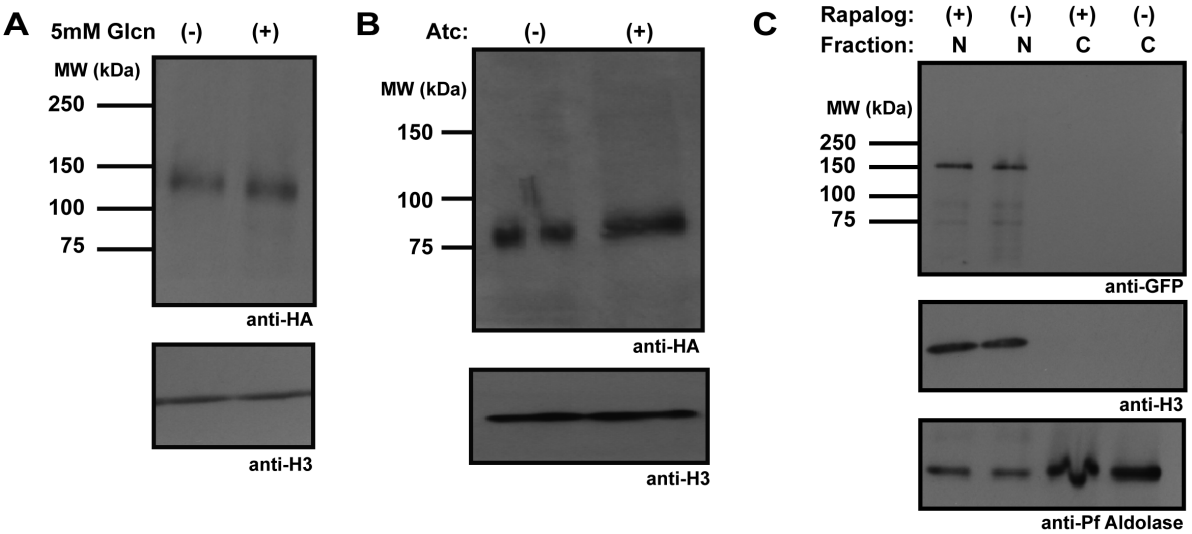


Figure S26. Western blot phenotyping of attempts to genetically knockdown AP2-EXP

398 A) To assess genetic knockdown of AP2-EXP by *glms* ribozyme tag in the parasite line AP2-
399 EXP::*glms*::HA, highly synchronous parasites were spiked with 5mM glucosamine or vehicle
400 control for 72 hours and AP2-EXP quantity was determined by anti-HA western blot. Histone H3
401 was used as a loading control. Glucosamine treatment did not impact the amount of AP2-EXP
402 protein present.
403
404 B) Genetic knockdown of AP2-EXP by the TetR:DOZI mRNA repression module was assessed
405 in the parasite line AP2-EXP::HA by washing anhydrotetracycline (aTc) out of the media for 72
406 hours. AP2-EXP quantity was determined by anti-HA western blot, and Histone H3 was used as
407 a loading control. Removal of aTc from the media did not impact the amount of AP2-EXP
408 protein present.
409
410 C) Genetic knockdown of AP2-EXP via protein mislocalization was assessed for the parasite
411 line AP2-EXP::GFP. 250nM rapalog was added to the media for 48 hours and AP2-EXP protein
412 localization was assessed by ant-GFP western blot. Histone H3 and Aldolase were used as
413 nuclear and cytosolic markers, respectively. The addition of rapalog did not cause any
414 detectable mislocalization of AP2-EXP from the nucleus to the cytosol. N indicates the nuclear
415 protein fraction, and C indicates the cytosolic fraction.
416

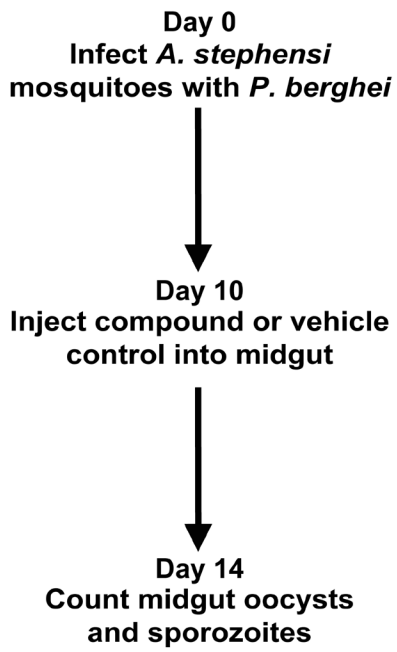


Figure S27. Mosquito stage *P. berghei* inhibition assay schematic

A. stephensi mosquitoes were infected with *Plasmodium berghei* parasites. On day 10 post infection, mosquito midguts were injected with the measured IC50 of Compounds B, C, F, or DMSO vehicle control. On day 14 post infection mosquitoes were dissected to count oocysts and midgut sporozoites. For Compound C and DMSO vehicle control, each experiment was performed in duplicate. Compounds B and F phenotyping were performed as a single experiment.

References

1. F. J. Gamo, *et al.*, Thousands of chemical starting points for antimalarial lead identification. *Nature* **465**, 305–310 (2010).
2. D. S. Wishart, *et al.*, DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**, D1074–D1082 (2018).
3. R. Huey, G. M. Morris, Using AutoDock 4 with AutoDockTools : A Tutorial. 1–56.
4. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general Amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
5. J. Wang, W. Wang, P. A. Kollman, D. A. Case, Automatic atom type and bond type perception in molecular

435 mechanical calculations. *J. Mol. Graph. Model.* **25**, 247–260 (2006).
436 6. P. A. K. and D. A. C. Wang, J., W. Wang, Antechamber: an accessory software package for molecular
437 mechanical calculations. *J. Am. Chem. Soc* **222**, U403 (2001).
438 7. D. Y. and P. K. Case, D., K. Belfon, I. Ben-Shalom, S. Brozell, D. Cerutti, T. Cheatham III, V. Cruzeiro, T.
439 Darden, R. Duke, G. Giambasu, M. Gilson, H. Gohlke, A. Goetz, R. Harris, S. Izadi, S. Izmailov, K.
440 Kasavajhala, A. Kovalenko, R. Krasny, T. Kurtzman, T. Lee, S. L, AMBER 2020 (2020).
441 8. A. Pérez, *et al.*, Refinement of the AMBER force field for nucleic acids: Improving the description of α/γ
442 conformers. *Biophys. J.* **92**, 3817–3829 (2007).
443 9. A. W. Götz, *et al.*, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1.
444 generalized born. *J. Chem. Theory Comput.* **8**, 1542–1555 (2012).
445 10. W. R. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Routine Microsecond Molecular Dynamics
446 Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh EwaldNo Title. *J Chem Theory Comput.*
447 **10**, 3878–3888 (2013).
448 11. J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, Numerical integration of the cartesian equations of motion of
449 a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
450 12. D. R. Roe, T. E. Cheatham, PTRAJ and CPPTRAJ: Software for processing and analysis of molecular
451 dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
452 13. G. K. Smyth, Linear models and empirical Bayes methods for assessing differential expression in microarray
453 experiments. *Stat. Appl. Ge- netics Mol. Biol.* **3**, Article 3 (2011).
454 14. D. P. Sangurdekar, The Rnits package for normalization and inference of differential expression in time series
455 microarray data . *Rvignette*, 1–13 (2014).
456 15. A. J. Saldanha, Java Treeview - Extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248
457 (2004).
458 16. M. J. L. de Hoon, S. Imoto, J. Nolan, S. Miyano, Open source clustering software. *Bioinformatics* **20**, 1453–
459 1454 (2004).
460 17. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data.
461 *Bioinformatics* **30**, 2114–2120 (2014).
462 18. S. Andrews, FastQC: A Quality Control Tool for High Throughput Sequence Data.
463 http://www.bioinformatics.babraham.ac.uk/projects/fastqc (2010).
464 19. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. **00**, 1–3 (2013).
465 20. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

466 21. J. M. Gaspar, Improved peak-calling with MACS2. *bioRxiv*, 1–16 (2018).
467 22. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*
468 **26**, 841–842 (2010).
469 23. F. Ramírez, F. Dündar, S. Diehl, B. A. Grüning, T. Manke, DeepTools: A flexible platform for exploring deep-
470 sequencing data. *Nucleic Acids Res.* **42**, 187–191 (2014).
471 24. T. L. Bailey, DREME: Motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653–1659
472 (2011).
473 25. C. E. Grant, T. L. Bailey, W. S. Noble, FIMO: Scanning for occurrences of a given motif. *Bioinformatics* **27**,
474 1017–1018 (2011).
475 26. L. Jiang, *et al.*, PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum.
476 *Nature* **499**, 223–227 (2013).
477 27. R. Bártfai, *et al.*, H2A.Z demarcates intergenic regions of the Plasmodium falciparum epigenome that are
478 dynamically marked by H3K9ac and H3K4me3. *PLoS Pathog.* **6**, e1001223 (2010).
479 28. D. F. Read, K. Cook, Y. Y. Lu, K. G. Le Roch, W. S. Noble, Predicting gene expression in the human malaria
480 parasite Plasmodium falciparum using histone modification, nucleosome positioning, and 3D localization
481 features. *PLoS Comput. Biol.* **15**, e1007329 (2019).
482 29. G. A. Josling, *et al.*, A Plasmodium Falciparum Bromodomain Protein Regulates Invasion Gene Expression.
483 *Cell Host Microbe* **17**, 741–751 (2015).
484 30. E. M. Bunnik, *et al.*, DNA-encoded nucleosome occupancy is associated with transcription levels in the
485 human malaria parasite Plasmodium falciparum. *BMC Genomics* **15**, 347 (2014).
486 31. S. A. Fraschka, *et al.*, Comparative Heterochromatin Profiling Reveals Conserved and Unique Epigenome
487 Signatures Linked to Adaptation and Development of Malaria Parasites. *Cell Host Microbe* **23**, 407–420
488 (2018).
489 32. X. Shang, *et al.*, Genome-wide landscape of ApiAP2 transcription factors reveals a heterochromatin-
490 associated regulatory network during Plasmodium falciparum blood-stage development. *Nucleic Acids Res.*
491 **50**, 3413–3431 (2022).
492 33. S. E. Lindner, E. K. De Silva, J. L. Keck, M. Llinás, Structural Determinants of DNA Binding by a P.
493 falciparum ApiAP2 Transcriptional Regulator. *J. Mol. Biol.* **395**, 558–567 (2010).
494 34. M. Llinás, Z. Bozdech, E. D. Wong, A. T. Adai, J. L. DeRisi, Comparative whole genome transcriptome
495 analysis of three Plasmodium falciparum strains. *Nucleic Acids Res.* **34**, 1166–1173 (2006).
496 35. L. Chappell, *et al.*, Refining the transcriptome of the human malaria parasite Plasmodium falciparum using

amplification-free RNA-seq. *BMC Genomics* **21**, 395 (2020).