HAL open science

# Target Identification of an Antimalarial Oxaborole Identifies AN13762 as an Alternative Chemotype for Targeting CPSF3 in Apicomplexan Parasites

Valeria Bellini, Christopher Swale, Marie-Pierre Brenier-Pinchart, Tiffany Pezier, Sonia Georgeault, Fabrice Laurent, Mohamed-Ali Hakimi, Alexandre Bougdour

## To cite this version:

Valeria Bellini, Christopher Swale, Marie-Pierre Brenier-Pinchart, Tiffany Pezier, Sonia Georgeault, et al.. Target Identification of an Antimalarial Oxaborole Identifies AN13762 as an Alternative Chemotype for Targeting CPSF3 in Apicomplexan Parasites. iScience, 2020, 23 (12), 14 p. 10.1016/j.isci.2020.101871 . hal-03106392

**HAL Id: hal-03106392**

**https://hal.inrae.fr/hal-03106392v1**

Submitted on 15 Dec 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S2589004220310683
Manuscript_f44a5a60c4e6df5ea72bebc9bad5a12c

**Title**

# Target identification of an antimalarial oxaborole identifies AN13762 as an alternative chemotype for targeting CPSF3 in apicomplexan parasites.

**Authors**

Valeria Bellini[1], Christopher Swale[1], Marie-Pierre Brenier-Pinchart[1], Tiffany Pezier[2], Sonia Georgeault[3], Fabrice Laurent[2], Mohamed-Ali Hakimi[1,*,†], Alexandre Bougdour[1,4,*,†]

**Affiliations**

[1]Institute for Advanced Biosciences (IAB), Host-Pathogen Interactions and Immunity to Infection, INSERM U1209, CNRS UMR 5309, University Grenoble Alpes, 38000 Grenoble, France.

[2]INRAE, Université François Rabelais de Tours, Centre Val de Loire, UMR1282 ISP, Laboratoire Apicomplexes et Immunité Mucosale, 37380 Nouzilly, France.

[3]Plateforme des Microscopies, Université et CHRU de Tours, 37000 Tours, France.

**Author List Footnotes**

[4]Lead Contact.

[†]These authors contributed equally to this work.

**Contact Information**

*Corresponding authors. Email: alexandre.bougdour@inserm.fr (A.B.); mohamed-ali.hakimi@inserm.fr (M.-A.H.).



© 2020 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

**Summary**

Boron-containing compounds represent a promising class of molecules with proven efficacy against a wide range of pathogens, including apicomplexan parasites. Following lead optimization, the benzoxaborole AN13762 was identified as a preclinical candidate against the human malaria parasite, yet the molecular target remained uncertain. Here, we uncovered the parasiticidal mechanisms of AN13762, by combining forward genetics with transcriptome sequencing and computational mutation discovery, and using *Toxoplasma gondii* as a relevant model for Apicomplexa. AN13762 was shown to target *Tg*CPSF3, the catalytic subunit of the pre-mRNA cleavage and polyadenylation complex, as the anti-pan-apicomplexan benzoxaborole compound, AN3661. However, unique mutations within the *Tg*CPSF3 catalytic site conferring resistance to AN13762 do not confer cross-protection against AN3661, suggesting a divergent resistance mechanism. Finally, in agreement with the high sequence conservation of CPSF3 between *Toxoplasma* and *Cryptosporidium*, AN13762 shows oral efficacy in cryptosporidiosis mouse model, a disease for which new drug development is of high priority.

35 **Introduction**

36 The Apicomplexa phylum contains intracellular single-celled parasites several of which are
37 causative agents of animal and human diseases worldwide raising important public health problems
38 (De Rycker et al., 2018). The group comprises important human pathogens such as *Plasmodium*,
39 *Toxoplasma*, and *Cryptosporidium* responsible for malaria, toxoplasmosis and cryptosporidiosis,
40 respectively. For many of these diseases current treatments are suboptimal and there are few or no
41 alternative available for some. Indeed, the current standard of treatment for *Cryptosporidium*
42 infections, nitazoxanide, shows limited and immune-dependent effectiveness (Manjunatha et al.,
43 2016). Although the current medication against *Toxoplasma* is quite effective, it has adverse side
44 effects, particularly in immunocompromised patients, such as pyrimethamine-induced hematological
45 toxicity; and sulphonamide-induced skin rash, leukopenia and thrombocytopenia (Dunay et al.,
46 2018). In the case of malaria, emergence and spread of resistance to artemisinin-based combination
47 therapy, the primary form of treatment, poses a constantly growing threat (De Rycker et al., 2018).
48 Therefore, new classes of small-molecule drugs or drugs with novel modes of action are needed to
49 overcome these limitations.

50 In an effort to optimize efficacy of a novel class of boron-containing molecules against
51 malaria parasites, the lead candidate AN13762 was identified in a phenotype-based screening
52 (referred to as compound 46 in (Zhang et al., 2017)). Evaluation of pharmacokinetics showed that
53 AN13762 has improved potency and metabolic stability, is orally bioavailable and equally potent
54 across multidrug-resistant strains of *Plasmodium falciparum* (*P. falciparum*), demonstrating no cross-
55 resistance and a possible new mechanism of action. AN13762 has not exhibited either significant
56 toxicology or cytotoxicity liabilities at any dose tested and AN13762 was selected for preclinical
57 development by Medicines for Malaria Venture in 2017 (Zhang et al., 2017). Although previous
58 research on parental scaffold AN3661 identified Cleavage and Polyadenylation Specificity Factor 3

59 (CPSF3) as the direct target (Palencia et al., 2017; Sonoiki et al., 2017; Swale et al., 2019), a recent
60 study investigating the resistance mechanisms of AN13762 in *P. falciparum* identified multiple
61 components involved in prodrug activation or sumoylation and ubiquitination pathways along with
62 *Pf*CPSF3 suggesting that the latter was not the primary target (Sindhe et al., 2020).

63 The work described here was undertaken in order to shed light on the parasiticidal
64 mechanisms of AN13762 using *Toxoplasma gondii* (*T. gondii*) as a relevant representative of
65 apicomplexan parasites. Here, we present evidence that AN13762 is effective against both *T. gondii*
66 and *Cryptosporidium parvum* (*C. parvum*) *in vitro* at low micromolar concentrations, and *in vivo* in
67 mouse models of toxoplasmosis and cryptosporidiosis, respectively. Using a forward genetic
68 approach based on transcriptome sequencing, we identified its target as CPSF3, a common target of
69 several benzoxaboroles such as AN3661, a compound active against apicomplexan parasites
70 (Palencia et al., 2017; Sonoiki et al., 2017; Swale et al., 2019), or the trypanocidal compounds
71 AN11736 and acoziborole (Wall et al., 2018). Importantly, several point mutations found in *T. gondii*
72 CPSF3 conferring resistance against AN13762 were not effective against AN3661, suggesting a
73 divergent mode of resistance mechanisms between CPSF3 and benzoxaboroles. Hence this work
74 uncovers the molecular mechanism for the antiparasitic activity of a preclinical antimalarial
75 candidate AN13762 and extends the clinical spectrum of activity of this chemotype to other life-
76 threatening apicomplexan parasites.

77

78 **Results**

79 **AN13762 is active against *T. gondii in vitro* and *in vivo***

80 In order to assess the effectiveness of AN13762 against *T. gondii* parasites, growth of the
81 type I reference RH strain was monitored within Human Foreskin Fibroblasts (HFFs) treated with
82 AN13762, its parental scaffold AN3661 as a positive control (Figure 1A), pyrimethamine, the standard
83 of care for toxoplasmosis, or vehicle (DMSO). Efficient *in vitro* inhibition of *T. gondii* growth was
84 repeatedly confirmed, with the measured half maximum Effective Concentration ($EC_{50}$) of 2.1 µM,
85 which is almost 40 times higher than that of AN3661 (Figures 1B and 1C). Complete and sustained
86 inhibition of growth was observed at 10 µM AN13762 without any adverse effects for the host cells
87 (Figures 1D and S1).

88 When AN13762 was administered orally for 7 days to *T. gondii* infected mice, beginning on
89 the first day following intraperitoneal injection of parasites, 100% of the animals survived the lethal
90 infection by the highly virulent type I RH strain in contrast to untreated controls (Figures 1E and 1F).
91 Second lethal challenges to the mice that survived the first infection, confirmed that the initial 7-day
92 treatment with AN13762 resulted in a protective immune response to subsequent *T. gondii* infection
93 (Figures 1E and 1F), thus strengthening the biological and pharmacokinetic profile of AN13762 in
94 animal efficacy studies. Altogether, these results indicate that AN13762 is effective against *T. gondii*
95 both *in vitro* and *in vivo* allowing long-term cures in mouse model of acute toxoplasmosis with
96 comparable efficacy to current treatment.

97 **Selection of *T. gondii* parasites resistant to AN13762**

98 In an attempt to shed light on the mechanism of action of AN13762, we performed a forward
99 genetic screen combining chemical mutagenesis to isolate AN13762-resistant parasites and NGS
100 sequencing analysis to map mutations conferring drug resistance (Figure 2A). Central to our
101 approach, we reasoned that the gene(s) that would be mutated in more than one independently

102 mutagenized resistant clone might be relevant to the drug resistance mechanism and by this means
103 alleviating the notoriously difficult molecular mapping of point mutations induced by mutagens. For
104 this purpose, 7 independent ethyl methanesulphonate (EMS) mutagenesis experiments were
105 performed and the resulting mutagenized parasites were selected in the presence of 10 µM AN13762
106 (Figure 2B), which corresponds to approximately 5-fold the $EC_{50}$ value. Resistant parasites were
107 obtained from each of the 7-mutagenesis experiments, whereas none of the non-mutagenized
108 parasites survived the selection at 10 µM of AN13762, attesting once more to the parasiticidal
109 efficacy of this compound (Figure 2B). The resistant parasite lines were then cloned by limited
110 dilution and we selected a single clone from each mutagenesis experiment (named A1 to G1) for
111 transcriptome sequencing by RNA-Seq. All the resistant clones were able to grow and formed
112 plaques when grown in the presence of 10 µM AN13762 (Figures 2C-2E and S2). In parallel, the
113 parental strain was analyzed by RNA-Seq and used as a reference to identify EMS-induced mutations.
114 We use transcriptome sequencing as most drugs target expressed proteins, with levels of gene
115 expression and mutations being part of the sequencing results.

116 **Parasites resistant to AN13762 harbor mutations within *TgCPSF3***

117 To map the EMS-induced mutations that confer drug resistance, the Illumina sequencing
118 reads were aligned to the ~65-Mb *T. gondii* GT1 reference genome. The assembled sequences were
119 analyzed to identify single nucleotide variations (SNVs), small insertions or short deletions using the
120 parental strain as a reference (see *Transparent Methods*). By focusing on mutations present in coding
121 sequences, we identified a single gene, *CPSF3* (Cleavage and Polyadenylation Specific Factor 3,
122 *TGGT1_285200*), that harbored SNVs leading to amino acid substitutions in each of the 7 drug-
123 resistant lines that were not present in the parental strain (Figure 2F and Table 1). *CPSF3* encodes a
124 nuclear mRNA processing endonuclease that functions in pre-mRNA maturation (*6*), which has been
125 previously identified as the target of a several benzoxaborole compounds active against distantly
126 related pathogens (Lunde et al., 2019; Sonoiki et al., 2017; Swale et al., 2019; Wall et al., 2018),

127 including *T. gondii* (Palencia et al., 2017). Importantly, 4 different mutations were identified (G456S,
128 E545K, Y328H, and S519C; Figure 2G; Table 1), among which E545K conferring resistance against
129 AN3661 in *T. gondii* (Palencia et al., 2017). Mutations span from the metallo-β-lactamase domain to
130 the RNA-specificity domain of CPSF3 (Figure 2G). Therefore, these data suggest that mutations in
131 *CPSF3* were responsible for resistance against AN13762.

132 **Mutations within CPSF3 confer resistance to AN13762**

133 To confirm that the *CPSF3* mutations were sufficient to confer resistance to AN13762, we
134 reconstructed each of the mutation identified in AN13762-resistant parasites into the sensitive
135 parental wild-type strain using CRISPR/Cas9 system coupled to homology-directed repair for gene
136 editing in *T. gondii* (Figure 3A) (Palencia et al., 2017). Thus, RHΔ*ku80* parasites were co-transfected
137 with a vector expressing the Cas9 endonuclease and synthetic guide RNA (sgRNA), and the
138 corresponding homologous single-stranded donor oligonucleotides (ssODN) as repair template. After
139 selection with AN13762, emerging resistant parasites were cloned, and DNA sequencing established
140 that the mutations have been correctly inserted at *CPSF3* locus (Figures 3B and S3A). Transfections
141 with the Cas9 control vectors alone produced no surviving parasites. In the engineered parasites, we
142 observed that the *CPSF3* mutations E545K, G456S, S519C, and Y328H substantially decreased the
143 sensitivity against AN13762 when compared to wild-type parasites (Figures 3C-3E and S3B-3D). It is
144 noteworthy that Y328H mutation had a significant effect on parasite growth in the absence of drug
145 (Figure 3C, upper panel), suggesting that this mutation might affect basal activity of CPSF3 in
146 tachyzoites, which is in line with *CPSF3* being essential to parasite growth (Palencia et al., 2017; Sidik
147 et al., 2016). In addition, the *CPSF3* edited parasites harboring the mutations E545K or G456S were
148 also resistant to AN13762 treatment in mice (Figure 1F). Altogether, these data confirm the primary
149 role of *CPSF3* mutations in conferring resistance to AN13762 and indicate that AN13762, in a similar
150 fashion to AN3661, targets CPSF3 (Palencia et al., 2017; Swale et al., 2019).

151 **AN13762-resistant mutations G456S and S519C do not confer cross-resistance to AN3661**

152 We had previously found that mutations in CPSF3 were conferring resistance to another oxaborole
153 compound, AN3661 (Figure 2G, mutations Y328C, Y483N, and E545K; (Palencia et al., 2017)). To
154 examine whether the AN13762-resistant mutations in CPSF3 confer cross-resistance to AN3661, we
155 assayed AN3661 against reconstructed parasites harboring *CPSF3* mutations E545K, G456S, S519C
156 and Y328H. As expected, the most prevalent mutation E545K that was identified in the mutagenesis
157 experiments conducted against either AN13762 or AN3661, conferred resistance to both compounds
158 (Figures 3C-3E and S3B-3D). Note that the increase in resistance to AN3661 was more dramatic than
159 for AN13762 (~100- and ~3-fold increase in $EC_{50}$, respectively; Figure 3D). Very different results were
160 obtained for G456S and S519C mutations, which did not allow parasite growth when exposed to 5
161 µM AN3661 (Figure 3C). The $CPSF3^{G456S}$ mutation conferred the strongest resistance phenotype to
162 AN13762 with a ~42-fold increase in AN13762 $EC_{50}$ when compared to wild-type parasites, whereas
163 sensitivity to AN3661 remained unaffected (Figures 3D and 3E). Of note, the latter mutations were
164 not identified in the AN3661 screen, presumably reflecting their inability to protect against AN3661
165 at 5 µM. Conversely, the Y483N mutation identified in AN3661-resistant parasites conferred cross-
166 resistance to AN13762. Similarly, mutations affecting the Y328 residue of CPSF3 decreased sensitivity
167 to both compounds (Figures 3C and 3D). Altogether, these results further confirm the role of CPSF3
168 mutations in drug resistance and indicate a divergent mode of resistance between AN13762 and
169 AN3661.

170 **Molecular docking suggests a divergent resistance mechanism between oxaboroles**

171 Multiple sequence alignments show a high overall sequence conservation within the metallo-
172 β-lactamase (MBL), Beta-Casp and RNA specificity domain of CPSF3 within apicomplexan parasites
173 and humans (Figure 4A). One notable difference between the apicomplexan and human enzyme is
174 the presence of an extended loop or "apicomplexan specific insert" whose length varies from 20 to
175 59 residues. However, conservation of the generated resistant SNVs to AN13762 within *T. gondii*

176 CPSF3 coding sequence is absolute across species and appears close to the catalytic residues but are
177 never directly involved in the coordination of the catalytic zinc atoms. Next, we visualized the
178 resistance conferring mutations within the recently obtained structure of *Cryptosporidium* CPSF3
179 (*Ch*CPSF3) in co-crystal with AN3661 (pdb id 6Q55) (Swale et al., 2019). With the assumption that
180 AN13762 interacts with a comparable geometry as the AN3661 benzoxaborole group, notably
181 through the boron driven octahedral coordination of the two catalytic zinc ions, we placed the
182 AN13762 derivative in the same plane as AN3661 (Figure 4B). Through this modeling, we did not
183 generate any clashes with CPSF3, despite the much bigger size of AN13762 (13.4 Å in length against 7
184 Å for AN3661). When visualizing both AN3661 and AN13762 placement with regards to the
185 resistance-conferring mutations, two important features can be noted: First, most of the mutations
186 found (Y328C/H, E545K, S519C and Y483N) which rescue parasites from both compounds are not
187 directly observed in contact to the compound binding site. Instead, the mutated residues are
188 generally placed on loop regions lining the interfacial cavity between the RNA specificity domain and
189 Beta-Casp domain. These resistance-conferring mutations probably act indirectly on the compound
190 activity through either an allosteric mechanism preventing compound binding or by modifying RNA
191 recognition by CPSF3 as these loop regions are believed to regulate RNA access and recognition (Sun
192 et al., 2020). Second, the G456S mutation which exclusively rescues *T. gondii* parasites from
193 AN13762 is observed separated to the other resistance conferring mutations. Because of its close
194 proximity with the AN13762 pyrazine ring and methylazetidine (2.2 Å distance), the G456S mutant
195 probably introduces an important steric hindrance to AN13762 binding. AN3661, with a much shorter
196 organic extension does not come close enough for the mutation to have an effect on its binding and
197 activity. As a result, the G456S mutant remains sensitive to AN3661.

198 **AN13762 is active against *Cryptosporidium in vitro* and *in vivo***

199 The above data provide evidence that AN13762 targets CPSF3 enzyme. Given that it has been
200 shown that CPSF3 is a *bonafide* target for inhibiting *Cryptosporidium* development (Swale et al.,

201 2019), we assessed the anticryptosporidial activity of AN13762 *in vitro* and *in vivo*. The ability of
202 AN13762 to inhibit *C. parvum* INRAE Nluc fast growing strain in human ileocecal HCT-8 was assessed
203 with its parental scaffold AN3661 as a positive control. Although less potent than AN3661, an
204 efficient *in vitro* inhibition of *C. parvum* growth was repeatedly observed with AN13762 ($EC_{50}$ 13 ±9
205 µM) (Figures 5A-5B and S5). AN13762 presented no detectable toxicity for the host cells, even at 100
206 µM (Figure 5C). AN13762 activity was therefore assessed *in vivo* in a neonatal mouse model. Seven-
207 day-old neonates were orally treated with AN13762 mixed in carboxymethyl cellulose (CMC) 4 h
208 after *C. parvum* infection and daily until 3 days post-infection (dpi). Parasite load was assessed in the
209 intestine at 4 dpi by oocyst count and measuring Nluc activity representing transgenic expression by
210 the INRAE Nluc strain. Both methods revealed an impressive and significant inhibition of parasite
211 development as illustrated in Figure 5D and by scanning electron microscopy where only very scarce
212 parasites can occasionally be found on the intestinal villi of treated mice. Remarkably, the enzymatic
213 assay revealed a 4-log reduction in luminescence signals in treated mice, and oocysts were not
214 detected by coproscopic intestinal material examination, which is much less sensitive than the
215 former method. Altogether, these results indicate that AN13762 is effective against *C. parvum* both
216 *in vitro* and *in vivo* and provide an additional drug presumably acting by a different mode of action
217 than AN3661 to block CPSF3 activity.

218 **Discussion**

219 Whole cell phenotypic screening is an efficient approach in drug discovery that has led to the
220 identification of numerous antimicrobial lead compounds, although the targets and mode of action
221 remain unknown and challenging to determine. While clinical development remains possible without
222 this knowledge, lack of insight into the mechanism of action is one of the biggest obstacles for
223 further medicinal chemistry optimization or to predict and track drug resistance. Fortunately, a large
224 variety of target deconvolution technologies are currently available. The approach developed here

225 takes advantage of all the benefits of the EMS mutagenesis method, including its wide and mostly
226 unbiased coverage of the genome with virtually all types of mutations (Farrell et al., 2014). In this
227 work, by combining cost-effective RNA-Seq based variant calling, computational mutation discovery
228 and CRISPR/Cas9 genome editing, we identified CPSF3, the catalytic subunit of the pre-mRNA
229 cleavage and polyadenylation complex, as the target of AN13762 in *T. gondii* parasites.

230 In eukaryotes, CPSF3 is key to the 3' end processing of both polyadenylated and replication-
231 dependent histone precursor mRNAs (Shi and Manley, 2015). These distinct 3' ends are generated
232 co-transcriptionally by specialized 3' end processing machineries that recognize a conserved
233 hexanucleotide AAUAAA and a downstream G/U-rich sequence on the 3'end of nascent pre-mRNAs
234 destined for polyadenylation or cleave histone mRNA precursors few nucleotides downstream of a
235 highly conserved stem–loop structure (Marzluff et al., 2008). As a result, the majority of histone
236 genes are expressed as nonpolyadenylated transcripts that are poorly detected by poly-A purified
237 based RNA-Seq (Lyons et al., 2016; Zhao et al., 2018). Despite this technical bias, our transcriptomic
238 data indicate that histone mRNAs (e.g. H2Ba, H4, H2Ax, H2A1, and H2Bb; Table S2 and Figure S4)
239 were dramatically enriched in drug-resistant lines harboring CPSF3 mutations Y328H, E545K, and
240 S519C, but not in those mutants containing the CPSF3$^{G456S}$ allele (strains A1 and F1 in Table 1). This
241 suggests hypomorphic mutations of CPSF3 that retain sufficient activity to overcome lethality but
242 somehow favor histone pre-mRNA processing towards polyadenylation of transcripts that were
243 otherwise barely detected using our poly-A selected transcript experiment settings. Interestingly, the
244 mutations Y328H, E545K, and S519C are lining the channel accommodating the RNA substrate on
245 CPSF3 (Figure S6), whereas the G456S mutation that is observed distant from the other mutations
246 did not affect histone mRNAs accumulation. It is noteworthy that the Y328 mutations significantly
247 impacted the overall growth fitness (Figure 3C), suggesting a default in *Tg*CPSF3$^{Y328H/C}$ activity. As the
248 G456S mutation in *T. gondii* is equivalent to the G330S mutation found in the human CPSF3
249 counterpart conferring resistance against the anti-cancer agent JTE-607 (Ross et al., 2020), it is likely

250 that the mechanism of resistance is shared. Possibly, the G330S and G456S mutations can only be
251 effective for elongated molecules to clash with the compound thereby impeding binding without
252 affecting recognition of the substrate. Yet further studies are required to determine whether the
253 mutations in CPSF3 affect the access of the substrate to the catalytic site, complex assembly or its
254 conformational dynamics as shown recently by Sun *et al.* (Sun et al., 2020). Altogether, these results
255 underscore the advantage of using transcriptome sequencing to investigate mechanisms of drug
256 action and to provide functional insight into the molecular biology of the target protein.

257 In mammalian cells, CPSF3 is embedded in a large multisubunit complex including CPSF1,
258 CPSF2, CPSF4, CPSF7, cleavage stimulatory factor 1 (CSTF1), CSTF2, CSTF3, symplekin, and WDR33
259 (Dominski and Marzluff, 2007; Ryan, 2004). A quite similar complex was purified in *T. gondii* (Table
260 S2, Swale *et al.*, manuscript in preparation), and the identified subunits were all predicted to be
261 essential for tachyzoite growth i*n vitro* (Sidik et al., 2016). No mutations with significant enrichment
262 were found in the CPSF3 protein partners in the resistant strains, which is in agreement with our
263 docking model based on *Cryptosporidium hominis* (*C. hominis*) CPSF3 structural data where the
264 oxaboroles are enfolded within the CPSF3 scaffold, presumably precluding any interaction with other
265 components.

266 The mutations conferring resistance to AN13762 target *Tg*CPSF3 catalytic site, a gold
267 standard evidence for target confirmation of a bioactive small molecule. In the published structure of
268 AN3661 bound to *Ch*CPSF3, the oxaborole competes with the catalytic water molecules for zinc
269 atoms, hence blocking the phosphate bond cleavage of the pre-mRNA substrate (Swale et al., 2019).
270 Given the overall conservation of CPSF3 catalytic core in Apicomplexa and the high conservation of
271 the residues involved in drug resistance, it is likely that AN13762 binds to this site and disrupt the
272 pre-mRNA processing activity of *Tg*CPSF3 that is essential for parasite growth.

273 While it is clear that AN13762 targets CPSF3 in *T. gondii*, different results were observed in *P.*
274 *falciparum* where the mechanism of resistance is plural (Sindhe et al., 2020). In fact, while we were
275 investigating the mechanism of action of AN13762 in *T. gondii*, Sindhe and colleagues have shown
276 that *P. falciparum* resistance depends on the activity of Prodrug Activation and Resistance Esterase
277 (*Pf*PARE), an enzyme responsible for AN13762 processing, but also on enzymes involved in
278 ubiquitination and SUMOylation pathways or *Pf*CPSF3. The latter is responsible for the high-level of
279 resistance, thus suggesting that AN13762 or its refined derivative theoretically targets CPSF3 in
280 malaria parasites as well. Whether AN13762 is processed in *T. gondii* is not known. However, since
281 *Tg*CPSF3$^{G456S}$ selectivity towards AN13762 is based on steric hindrance over the methylazetidine
282 group which is cleaved off upon processing by the esterase, it seems unlikely that such a modification
283 occurs in *T. gondii*. Note that no mutations with significant enrichment were found in
284 *TGGT1_306330,* the closest homologue to *Pf*PARE in *T. gondii* (Table S2). Furthermore, as AN13762
285 processing is required for full antimalarial activity, it is tempting to speculate that the lack of
286 intracellular activation explains the decreased sensitivity observed in *T. gondii* and *Cryptosporidium*
287 ($EC_{50}$ values are in the µM range, Figures 1C and 5B) relative to *P. falciparum* ($EC_{50}$ values ranging
288 from 18 to 118 nM, (Sindhe et al., 2020)).

289 Based on the catalytic core sequence homology between *Tg*CPSF3 and *Cp*CPSF3, both
290 previously chemically validated targets for *Toxoplasma* and *Cryptosporidium* (Palencia et al., 2017;
291 Swale et al., 2019), we successfully laid the groundwork for pathogen hopping. In this respect,
292 AN13762 efficiently inhibits *C. parvum,* a species relevant to human health, *in vitro* and *in vivo* in
293 mouse model of infection. These results appear to be even more important for the treatment of
294 cryptosporidiosis, where druggable targets are scarce and there is a high demand for more efficient
295 therapies. However, further work will be needed to demonstrate that AN13762 acts as a direct
296 binder of the *Cp*CPSF3 and inhibits its mRNA processing activity, thereby restricting the growth of
297 parasites. The recent discovery of benzoxaborole-based chemistry has given rise to a series of

compounds with great potential against various infectious agents, including trypanosomatids and apicomplexan parasites by targeting different molecular targets (De Rycker et al., 2018). Remarkably, multiple compounds with known or suspected anti-CPSF3 activity across different organisms share a similar benzoxaborole scaffold that could be a prerequisite to CPSF3 binding (Begolo et al., 2018; Lunde et al., 2019; Palencia et al., 2017; Wall et al., 2018). Interestingly, the oxaborole such as Acoziborole can cross the blood brain barrier (Nare et al., 2010), offering a therapeutic option to eradicate persistent *Toxoplasma* cysts that are resistant to most, if not all, medications currently prescribed.

## Limitations of the Study

While our study is reasonably clear about AN13762 targeting CPSF3 in *Toxoplasma* and its activity against *Cryptosporidium* parasites, it remains possible that the mechanism of action in the latter is different and depends on prodrug activating enzyme(s) such as *Pf*PARE as described in *Plasmodium species*. Hopefully, recent advances in *Cryptosporidium* genetics will make it possible to carry out such investigations and genetically validate the *Cp*CPSF3 molecular target in this organism (Vinayak et al., 2020).

## Resource Availability

### Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Alexandre Bougdour (alexandre.bougdour@inserm.fr).

### Materials Availability

All unique materials generated in this study are available from the lead Contact upon request.

### Data and Code Availability

This study did not generate/analyze code.

322 The Illumina RNA-Seq dataset generated during this study is available at NCBI GEO:
323 GSE156685.

324 **Acknowledgments**

325 We acknowledge A. Palencia for initial fruitful discussions. We thank D. Cannella and C.
326 Corrao for technical assistance.

327 This research was supported by funds from the Agence Nationale pour la Recherche Project
328 ToxoP53 (grant no. ANR-19-CE15-0026) to A.B., Project HostQuest (grant no. ANR-18-CE15-0023),
329 and the European Research Council (ERC Consolidator grant no. 614880 Hosting TOXO) to M.-A.H.

330 **Author contributions**

331 F.L., M.-A.H., and A.B. conceptualized the research. A.B. supervised the research. V.B. designed and
332 conducted the *in vitro* studies performed in *T. gondii*. C.S. performed structural modellings. M.-P.B.P.
333 and V.B. designed and conducted the *in vivo* experiments with *T. gondii*. A.B. computed and analyzed
334 the RNA-seq data. F.L. supervised the work performed on *Cryptosporidium*. T.P. realized the *in vitro*
335 and *in vivo* studies performed with *Cryptosporidium*. S.G. performed the electron microscopy study.
336 V.B., C.S., and A.B. wrote the manuscript. Funding Acquisition, M.-A.H., and A.B. All the authors
337 contributed to the editing of the final version of manuscript, discussed and approved the results.

338 **Declaration of Interests**

339 The authors declare no competing interests.

340 **References**

341 Begolo, D., Vincent, I.M., Giordani, F., Pöhner, I., Witty, M.J., Rowan, T.G., Bengaly, Z., Gillingwater,
342 K., Freund, Y., Wade, R.C., Barrett, M.P., Clayton, C., 2018. The trypanocidal benzoxaborole
343 AN7973 inhibits trypanosome mRNA processing. PLoS Pathog 14, e1007315.

344 Casañal, A., Kumar, A., Hill, C.H., Easter, A.D., Emsley, P., Degliesposti, G., Gordiyenko, Y., Santhanam,
345 B., Wolf, J., Wiederhold, K., Dornan, G.L., Skehel, M., Robinson, C.V., Passmore, L.A., 2017.
346 Architecture of eukaryotic mRNA 3'-end processing machinery. Science 358, 1056–1059.

347 De Rycker, M., Baragaña, B., Duce, S.L., Gilbert, I.H., 2018. Challenges and recent progress in drug
348 discovery for tropical diseases. Nature 559, 498–506.

349 Dominski, Z., Marzluff, W.F., 2007. Formation of the 3' end of histone mRNA: Getting closer to the
350 end. Gene 396, 373–390.

351 Dunay, I.R., Gajurel, K., Dhakal, R., Liesenfeld, O., Montoya, J.G., 2018. Treatment of Toxoplasmosis:
352 Historical Perspective, Animal Models, and Current Clinical Practice. Clin Microbiol Rev 31.

353 Farrell, A., Coleman, B.I., Benenati, B., Brown, K.M., Blader, I.J., Marth, G.T., Gubbels, M.-J., 2014.
354 Whole genome profiling of spontaneous and chemically induced mutations in Toxoplasma
355 gondii. BMC genomics 15, 354.

356 Lunde, C.S., Stebbins, E.E., Jumani, R.S., Hasan, M.M., Miller, P., Barlow, J., Freund, Y.R., Berry, P.,
357 Stefanakis, R., Gut, J., Rosenthal, P.J., Love, M.S., McNamara, C.W., Easom, E., Plattner, J.J.,
358 Jacobs, R.T., Huston, C.D., 2019. Identification of a potent benzoxaborole drug candidate for
359 treating cryptosporidiosis. Nat Commun 10, 2816.

360 Lyons, S.M., Cunningham, C.H., Welch, J.D., Groh, B., Guo, A.Y., Wei, B., Whitfield, M.L., Xiong, Y.,
361 Marzluff, W.F., 2016. A subset of replication-dependent histone mRNAs are expressed as
362 polyadenylated RNAs in terminally differentiated tissues. Nucleic Acids Res 44, 9190–9205.

363 Manjunatha, U.H., Chao, A.T., Leong, F.J., Diagana, T.T., 2016. Cryptosporidiosis Drug Discovery:
364 Opportunities and Challenges. ACS Infect Dis 2, 530–537.

365 Marzluff, W.F., Wagner, E.J., Duronio, R.J., 2008. Metabolism and regulation of canonical histone
366 mRNAs: life without a poly(A) tail. Nat Rev Genet 9, 843–854.

367 Nare, B., Wring, S., Bacchi, C., Beaudet, B., Bowling, T., Brun, R., Chen, D., Ding, C., Freund, Y., Gaukel,
368 E., Hussain, A., Jarnagin, K., Jenks, M., Kaiser, M., Mercer, L., Mejia, E., Noe, A., Orr, M.,
369 Parham, R., Plattner, J., Randolph, R., Rattendi, D., Rewerts, C., Sligar, J., Yarlett, N., Don, R.,
370 Jacobs, R., 2010. Discovery of Novel Orally Bioavailable Oxaborole 6-Carboxamides That
371 Demonstrate Cure in a Murine Model of Late-Stage Central Nervous System African
372 Trypanosomiasis. AAC 54, 4379–4388.

373 Palencia, A., Bougdour, A., Brenier-Pinchart, M.-P., Touquet, B., Bertini, R.-L., Sensi, C., Gay, G.,
374 Vollaire, J., Josserand, V., Easom, E., Freund, Y.R., Pelloux, H., Rosenthal, P.J., Cusack, S.,
375 Hakimi, M.-A., 2017. Targeting Toxoplasma gondii CPSF3 as a new approach to control
376 toxoplasmosis. EMBO Mol Med 9, 385–394.

377 Ross, N.T., Lohmann, F., Carbonneau, S., Fazal, A., Weihofen, W.A., Gleim, S., Salcius, M., Sigoillot, F.,
378 Henault, M., Carl, S.H., Rodríguez-Molina, J.B., Miller, H.R., Brittain, S.M., Murphy, J.,
379 Zambrowski, M., Boynton, G., Wang, Y., Chen, A., Molind, G.J., Wilbertz, J.H., Artus-Revel,
380 C.G., Jia, M., Akinjiyan, F.A., Turner, J., Knehr, J., Carbone, W., Schuierer, S., Reece-Hoyes, J.S.,
381 Xie, K., Saran, C., Williams, E.T., Roma, G., Spencer, M., Jenkins, J., George, E.L., Thomas, J.R.,
382 Michaud, G., Schirle, M., Tallarico, J., Passmore, L.A., Chao, J.A., Beckwith, R.E.J., 2020.

383 CPSF3-dependent pre-mRNA processing as a druggable node in AML and Ewing's sarcoma.
384 Nat Chem Biol 16, 50–59.

385 Ryan, K., 2004. Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing
386 endonuclease. RNA 10, 565–573.

387 Shi, Y., Manley, J.L., 2015. The end of the message: multiple protein–RNA interactions define the
388 mRNA polyadenylation site. Genes Dev. 29, 889–897.

389 Sidik, S.M., Huet, D., Ganesan, S.M., Huynh, M.-H., Wang, T., Nasamu, A.S., Thiru, P., Saeij, J.P.J.,
390 Carruthers, V.B., Niles, J.C., Lourido, S., 2016. A Genome-wide CRISPR Screen in Toxoplasma
391 Identifies Essential Apicomplexan Genes. Cell 166, 1423-1435.e12.

392 Sindhe, K.M.V., Wu, W., Legac, J., Zhang, Y.-K., Easom, E.E., Cooper, R.A., Plattner, J.J., Freund, Y.R.,
393 DeRisi, J.L., Rosenthal, P.J., 2020. Plasmodium falciparum Resistance to a Lead Benzoxaborole
394 Due to Blocked Compound Activation and Altered Ubiquitination or Sumoylation. mBio 11.

395 Sonoiki, E., Ng, C.L., Lee, M.C.S., Guo, D., Zhang, Y.-K., Zhou, Y., Alley, M.R.K., Ahyong, V., Sanz, L.M.,
396 Lafuente-Monasterio, M.J., Dong, C., Schupp, P.G., Gut, J., Legac, J., Cooper, R.A., Gamo, F.-J.,
397 DeRisi, J., Freund, Y.R., Fidock, D.A., Rosenthal, P.J., 2017. A potent antimalarial
398 benzoxaborole targets a Plasmodium falciparum cleavage and polyadenylation specificity
399 factor homologue. Nat Commun 8, 14574.

400 Sun, Y., Zhang, Y., Aik, W.S., Yang, X.-C., Marzluff, W.F., Walz, T., Dominski, Z., Tong, L., 2020.
401 Structure of an active human histone pre-mRNA 3'-end processing machinery. Science 367,
402 700–703.

403 Swale, C., Bougdour, A., Gnahoui-David, A., Tottey, J., Georgeault, S., Laurent, F., Palencia, A., Hakimi,
404 M.-A., 2019. Metal-captured inhibition of pre-mRNA processing activity by CPSF3 controls
405 Cryptosporidium infection. Sci Transl Med 11.

406 Vinayak, S., Jumani, R.S., Miller, P., Hasan, M.M., McLeod, B.I., Tandel, J., Stebbins, E.E., Teixeira, J.E.,
407 Borrel, J., Gonse, A., Zhang, M., Yu, X., Wernimont, A., Walpole, C., Eckley, S., Love, M.S.,
408 McNamara, C.W., Sharma, M., Sharma, A., Scherer, C.A., Kato, N., Schreiber, S.L., Melillo, B.,
409 Striepen, B., Huston, C.D., Comer, E., 2020. Bicyclic azetidines kill the diarrheal pathogen
410 Cryptosporidium in mice by inhibiting parasite phenylalanyl-tRNA synthetase. Sci Transl Med
411 12.

412 Wall, R.J., Rico, E., Lukac, I., Zuccotto, F., Elg, S., Gilbert, I.H., Freund, Y., Alley, M.R.K., Field, M.C.,
413 Wyllie, S., Horn, D., 2018. Clinical and veterinary trypanocidal benzoxaboroles target CPSF3.
414 Proc Natl Acad Sci U S A 115, 9616–9621.

415 Zhang, Y.-K., Plattner, J.J., Easom, E.E., Jacobs, R.T., Guo, D., Freund, Y.R., Berry, P., Ciaravino, V.,
416 Erve, J.C.L., Rosenthal, P.J., Campo, B., Gamo, F.-J., Sanz, L.M., Cao, J., 2017. Benzoxaborole
417 Antimalarial Agents. Part 5. Lead Optimization of Novel Amide Pyrazinyloxy Benzoxaboroles
418 and Identification of a Preclinical Candidate. J. Med. Chem. 60, 5889–5908.

419 Zhao, S., Zhang, Y., Gamini, R., Zhang, B., von Schack, D., 2018. Evaluation of two main RNA-seq
420 approaches for gene quantification in clinical RNA sequencing: polyA+ selection versus rRNA
421 depletion. Sci Rep 8, 4781.

422 **Main figure titles and legends**

423 **Figure 1. Activity of AN13762 against *Toxoplasma gondii*.** (**A**) Chemical structures of benzoxaborole
424 leads AN13762 and AN3661. (**B**) Dose–response curves for inhibition of *T. gondii* growth *in vitro* in
425 response to increasing concentration of the indicated compounds. Confluent HFF monolayer were
426 infected with tachyzoites of *T. gondii* RH strain expressing the NanoLuc luciferase (RH Δ*ku80*
427 *UPRT::NLuc-P2A-EmGFP*). The *T. gondii* strains used in this study are listed in Table S1. Data are
428 presented as mean ± standard deviation (SD) of at least two independent biological assays, each with
429 3 technical replicates. Shaded error envelopes depict 95% confidence intervals. (**C**) $EC_{50}$ values of
430 each biological replicate were determined by non-linear regression analysis. $EC_{50}$ data are presented
431 as mean ±SD from at least 2 independent biological replicates, each with 3 technical replicates. (**D**)
432 HFF cells were infected with tachyzoites (RH Δ*ku80 UPRT::NLuc-P2A-EmGFP*) and incubated with 10
433 µM AN13762, 5 µM AN3661 or 0.1% DMSO as control. Cells were fixed 24 h post-infection and then
434 stained with antibodies against the *T. gondii* inner membrane complex protein GAP45 (magenta). The
435 cytosolic GFP is shown in green. Scale bars represent 10 µm. A complete data set can be found in
436 Figure S1. (**E**) Acute toxoplasmosis: timeline of mouse infections and treatments. Untreated mice
437 succumbed to infection, and thus a new group of healthy CBA/JRj mice was used for the second
438 challenge (n=3 in each group). (**F**) Survival curves of the CBA/JRj mice infected intraperitoneally (i.p.)
439 with $10^3$ tachyzoites of type I RH wild-type (WT) or the indicated *CPSF3* mutant strains (E545K or
440 G456S). During the first challenge, mice were treated orally with 40 mg/kg AN13762 or 200 mg/kg
441 sulphadiazine once daily beginning 1-day post-infection (n=6 for each condition; two independent
442 experiments, each with three mice per experimental group). Mice surviving the primary infection
443 were challenged a second time with the *T. gondii* WT strain ($2^{nd}$ challenge). A new group of naïve
444 mice was used as control (n=3 in each group).

445 **Figure 2. Strategy for AN13762 target deconvolution.** (**A**) Diagram of key steps of the forward
446 genetic screen to map mutations conferring drug resistance in *T. gondii* parasites. EMS mutagenized

447 population of *T. gondii* tachyzoites were selected in the presence of a lethal concentration of
448 AN13762 to isolate drug-resistant parasites. Analysis of the parental wild-type strain and multiple
449 resistant clones by variant-calling of sequencing reads generated by RNA-Seq to identify EMS-
450 induced mutations in coding sequences conferring drug-resistance. (**B**) Schematic of the strategy
451 used to obtain *T. gondii* resistant parasites to AN13762. From each mutagenesis experiment a single
452 clone (A1 to F1) was isolated and analyzed by RNA-Seq. (**C**) AN13762-resistant parasites form plaques
453 after 7 days of growth in the presence of 10 µM AN13762. Complete data set in shown in Figure S2A.
454 (**D**) Quantification of plaque sizes of wild-type parasites and resistant lines (A1 to G1) when cultured
455 in the presence of AN13762. n.d., not detected. Associated data are shown in Figure S2C. (**E**)
456 Fluorescence microscopy showing intracellular growth of *T. gondii* AN13762-resistant lines. HFF cells
457 were infected by the indicated *T. gondii* strains in the presence or absence of 10 µM AN13762. At 24
458 h post-infection, cells were fixed and stained with antibodies against GAP45 (magenta) and Hoechst
459 (blue) to detect IMC of parasites and nuclei, respectively. (**F**) Circos plot summarizing single
460 nucleotide variants (SNVs), insertions and deletions detected by transcriptomic analysis of the *T.*
461 *gondii* AN13762-resistant lines, grouped by chromosome (numbered in Roman numerals with size
462 intervals given outside). Each dot in the 7 innermost gray tracks corresponds to a scatter plot of the
463 mutations identified in the coding regions of the 7 drug-resistant strains, with each ring representing
464 one of the 7 drug-resistant lines (A1 to G1). In the second outermost track, lines depicting whole-
465 genome RNA-seq data of the *T. gondii* parental strain (RPKM values of genes are shown). Each bar in
466 the outermost track represents locations of selected archetypal essential genes. See Table S2 and
467 Figure S4 for transcriptomic analysis. (**G**) *Tg*CPSF3 domain architecture as predicted from PFAM
468 databases and crystal structures of *Cryptosporidium* CPSF3 (Swale et al., 2019). Positioning of
469 residues that were mutated in parasites resistant to AN13762 (Y328H, G465S, S519C, and E545K, in
470 red) or AN3661 (Y328C, Y483N, and E545K, in blue) are indicated.

471

472 **Figure 3. Validation of *T. gondii* CPSF3 as the benzoxaborole target.** (**A**) Schematic of the *CPSF3* gene
473 editing strategy in *T. gondii* parasites. Detailed view of *CPSF3* locus and CRISPR/Cas9-mediated
474 homology-directed repair with single-stranded oligo DNA nucleotides (ssODNs) carrying nucleotide
475 substitutions (orange letters). After homologous recombination (HR) events with ssODNs, CPSF3
476 recombinants were selected with AN13762. (**B**) Sanger sequencing validation of *CPSF3* editing.
477 Chromatograms of *CPSF3* DNA sequences from parental and engineered parasites are shown.
478 Nucleotide positions relative to the ATG start codon on genomic DNA are indicated. A complete data
479 set can be found in Figure S3A. (**C**) Effects of *CPSF3* mutations on *T. gondii* lytic cycle as determined
480 by plaque assay. Plaque sizes were measured for WT and the engineered CPSF3 mutant strains
481 (G456S, S519C, Y328H, E545K, Y483N, Y328C) after 7 days of growth in the absence or presence of 10
482 µM AN13762 or 5 µM AN3661. n.d., not detected. *P*-values corresponding to Kruskal−Wallis test with
483 Dunn's multiple comparisons with the wild-type (WT) strain are indicated. *ns*, not significative.

484 Associated data are shown in Figure S3B. (**D**) $EC_{50}$ values for Pyrimethamine (Pyr), AN13762, and
485 AN3661 were determined for WT and the engineered CPSF3 mutant strains (G456S, S519C, Y328H,
486 E545K, Y483N, Y328C). Data are mean from at least 2 independent biological replicates, each with 3
487 technical replicates. Associated dose–response curves are shown in Figure S3C. Mean $EC_{50}$ values ±SD
488 with fold changes (FC) in $EC_{50}$ relative to that of the WT parasites are indicated. (**E**) Fluorescence
489 microscopy showing intracellular growth of WT and the *CPSF3* edited parasites (G456S and E545K).
490 HFF cells were infected with tachyzoites of the indicated *T. gondii* strains expressing the *NLuc-P2A-*
491 *EmGFP* reporter gene and incubated with 10 µM AN13762, 5 µM AN3661 or 0.1% DMSO as control.
492 Cells were fixed 24 h post-infection and then stained with antibodies against the *T. gondii* inner
493 membrane complex protein GAP45 (magenta). The cytosolic GFP is shown in green. Scale bars
494 represent 10 µm. Complete data set in shown in Figure S3D.

495 **Figure 4. Docking studies for chemotypes AN13762 and AN3661.** (**A**) Multiple sequence alignment of
496 CPSF3 proteins from *T. gondii* (*Tg*), *C. hominis* (*Ch*), *P. falciparum* (*Pf*), and CPSF73 of *H. sapiens* (*Hs*).

497 The domain architecture is indicated as follows: blue, metallo-β-lactamase; green, β-CASP; orange,
498 RNA specificity domain; magenta, the insertion within the MBL domain of apicomplexan parasites.
499 Residues mutated in drug-resistant parasites are indicated by asterisks. The highly conserved
500 residues involved in the coordination the zinc atoms Zn1 or Zn2 are indicated in gray. Mutations
501 identified in parasites resistant to AN13762 or AN3661 are indicated in red and blue text,
502 respectively. (**B**) Schematic of *Cryptosporidium hominis* CPSF3/AN3661 co-crystal structure and
503 modelling with AN13762. CPSF3 is displayed in a cartoon fashion with the same domain color code as
504 in (A) surrounded by a light grey surface representation. Catalytic zinc atoms and coordinating
505 residues are shown in grey sticks while resistant mutations are shown in yellow and pink spheres. (**C**)
506 Zoom into the catalytic pocket with AN13762 manually placed colored in cyan. (**D**) Zoom into the
507 catalytic pocket binding AN3661 colored in dark purple.

508 **Figure 5. Efficacy against *C. parvum* in cell culture and neonatal mouse model.** (**A**) Comparative
509 inhibitory activity of AN13762 and AN3661 against *C. parvum* INRAE Nluc strain in human ileocecal
510 HCT-8 cells. The effect of both drugs in reducing parasitic load in epithelial cells was monitored by
511 the luminescence signal of transgenic Nluc parasites (each concentration point represents the
512 average of six measurements ± SD. Curves corresponding to AN13762- and AN3661- are in magenta
513 and blue respectively. Corresponding fluorescence microscopy images showing intracellular growth
514 of *C. parvum* parasites can be found in Figure S5. (**B**) Calculated $EC_{50}$ measurements are shown for
515 AN13762 and AN3661 (n=4 for each drug). (**C**) HCT-8 cell viability assay performed 48 h with
516 increasing concentration of AN13762. Percent viability compared to the untreated control is
517 displayed as a function of compound concentration in micromolar concentrations. Dotted line
518 represents 100% viability. (**D**) Schematic representation of the 4-day oral dosage of AN13762 [40
519 mg/kg] in CMC from day 0 (4 h post-infection) in 7-day-old neonatal mice previously infected with 5 ×
520 $10^5$ oocysts. The degree of infection was monitored by counting the oocysts in the small intestine of
521 the animals at 4 dpi (D.L., detection limit = $6.10^4$ oocysts/intestine) and by monitoring Nluc activity

522 on a small piece of ileum of each neonatal mouse (n = 14 animals per group). n.d., not detected.

523 Mann-Whitney test, ****$P < 0.0001$. Scanning electron microscopy (SEM) imaging of neonatal mice

524 ileum was performed at the end of the experiment on treated (AN13762) and mock treated (control)

525 animals.

526 **Main tables and legends**

| | | | | Variant calling | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Parental strain | Resistant mutants | | | | | | |
| Chr. | Gene | Annotation | Position | WT | A1 | B1 | C1 | D1 | E1 | F1 | G1 |
| V | *TGGT1_285200* | *CPSF3* | 2395898 | | | | E545K (GAG to AAG) | E545K (GAG to AAG) | | | E545K (GAG to AAG) |
| V | *TGGT1_285200* | *CPSF3* | 2395976 | | | | | | S519C (AGC to TGC) | | |
| V | *TGGT1_285200* | *CPSF3* | 2396165 | | G456S (GGC to AGC) | | | | | G456S (GGC to AGC) | |
| V | *TGGT1_285200* | *CPSF3* | 2396549 | | | Y328H (TAC to CAC) | | | | | |

527 **Table 1. Mutations found in candidate gene identified by RNA-Seq transcriptome analysis.** Amino

528 acid substitutions with the corresponding codons shown in parentheses are indicated for each

529 mutagenized *T. gondii* resistant mutant strain.

530 **Supplemental Excel table title and legends**

531 **Table S1. Strains and plasmids, primers and oligonucleotides, Related to Figures 1-3 and 5 and**

532 **Transparent Methods.**

533 **Table S2. RNA-Seq analysis of the EMS-induced drug-resistant lines of *T. gondii,* Related to Figure 2**

534 **and Transparent Methods.** RNA-Seq report, *T. gondii* transcripts (RPKM), Transcriptomic histones,

535 Transcriptomic CPSF complex, All variants, Coding region variants, Filtered variants in clones A1 to

536 G1, data used in Circos plot.

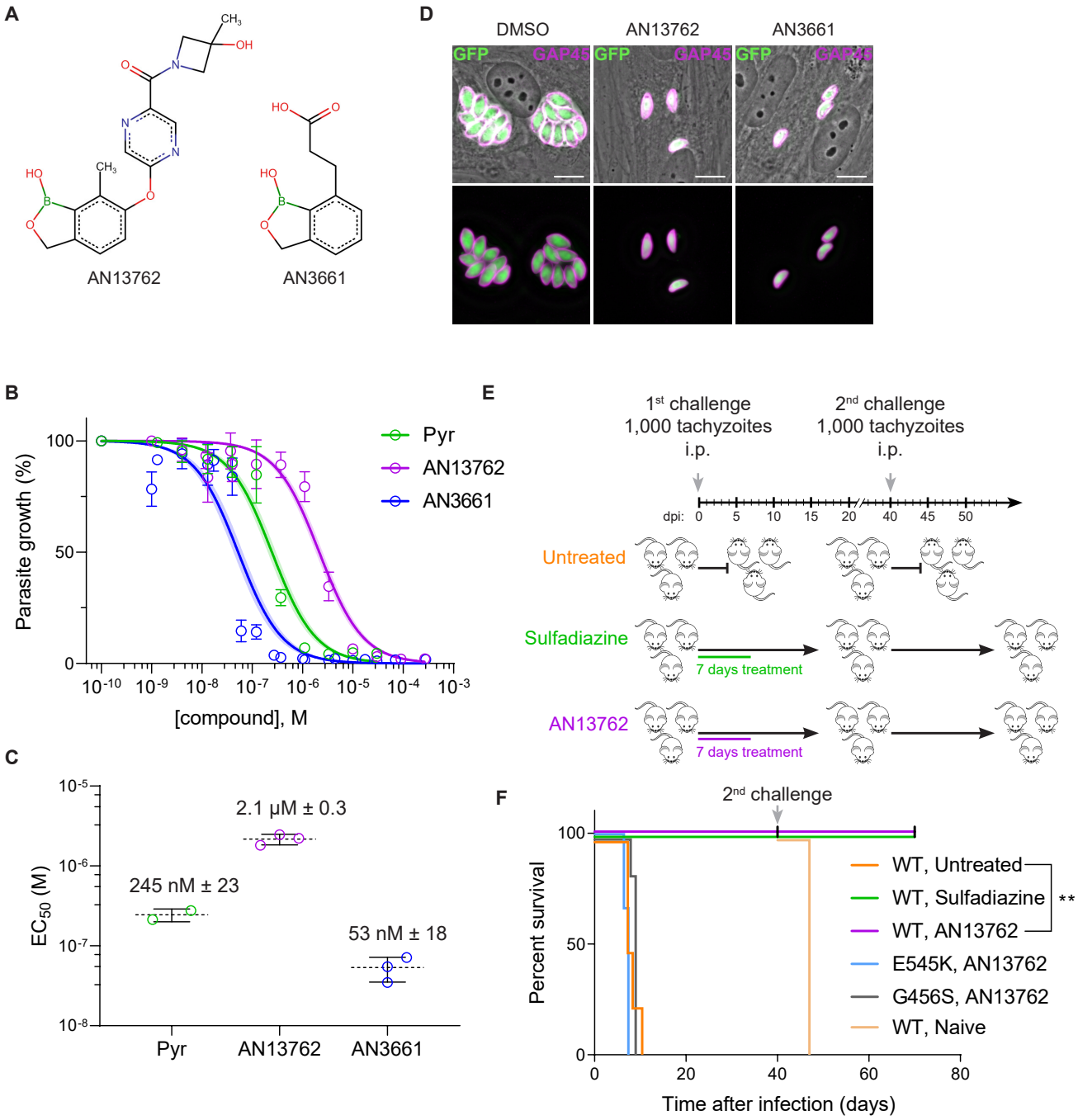
A
AN13762
AN3661
D
DMSO
AN13762
AN3661
GFP
GAP45
B
Parasite growth (%)
[compound], M
Pyr
AN13762
AN3661
C
EC50 (M)
2.1 μM ± 0.3
245 nM ± 23
53 nM ± 18
Pyr
AN13762
AN3661
E
1st challenge
1,000 tachyzoites
i.p.
2nd challenge
1,000 tachyzoites
i.p.
dpi: 0 5 10 15 20 40 45 50
Untreated
Sulfadiazine
7 days treatment
AN13762
7 days treatment
F
2nd challenge
Percent survival
Time after infection (days)
WT, Untreated
WT, Sulfadiazine
WT, AN13762
E545K, AN13762
G456S, AN13762
WT, Naive
**

**Figure 1**

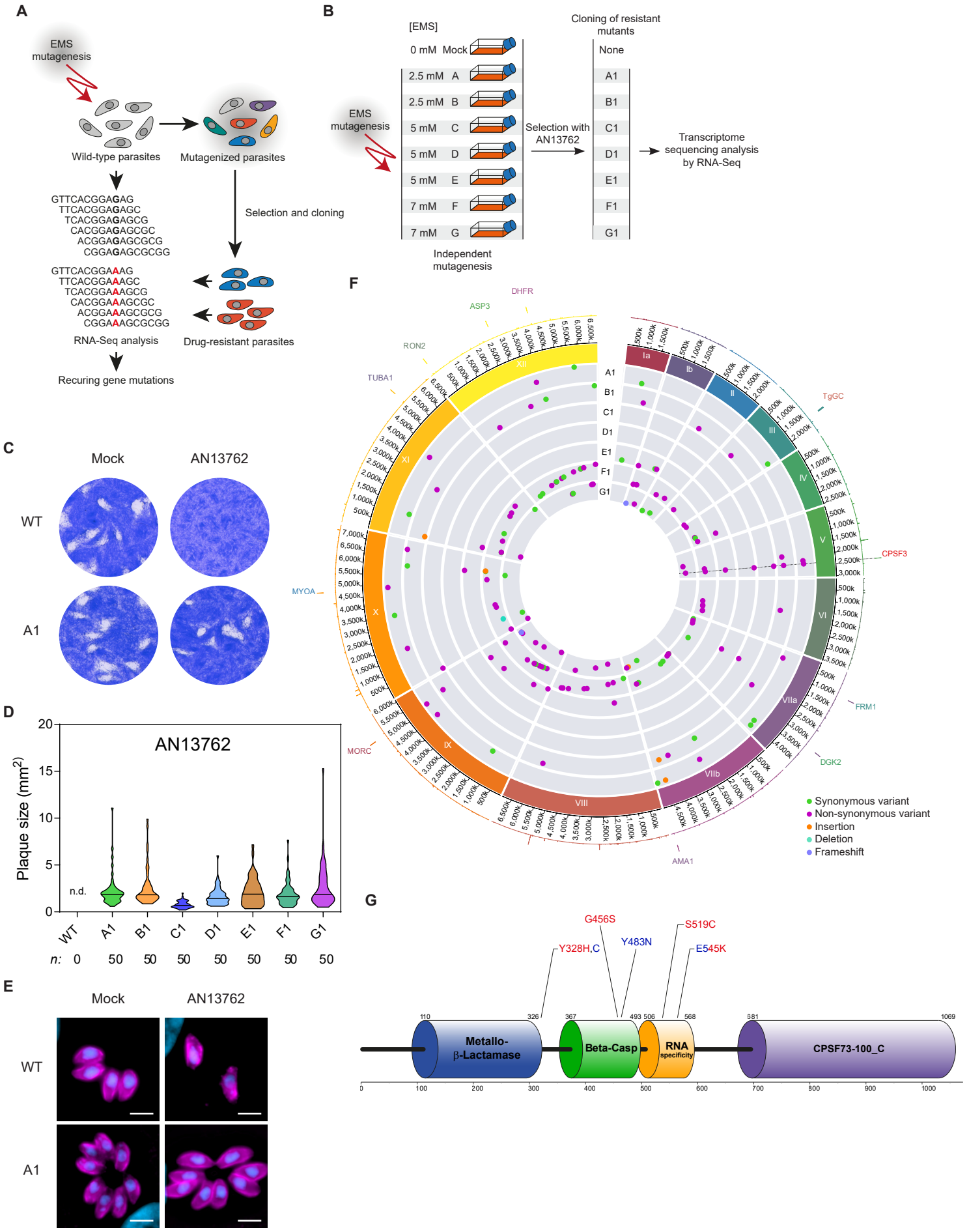
A
EMS mutagenesis
Wild-type parasites
Mutagenized parasites
GTTCACGGAGAG
TTCACGGAGAGC
TCACGGAGAGCG
CACGGAGAGCGC
ACGGAGAGCGCG
CGGAGAGCGCGG
GTTCACGGAAAG
TTCACGGAAAGC
TCACGGAAAGCG
CACGGAAAGCGC
ACGGAAAGCGCG
CGGAAAGCGCGG
Selection and cloning
RNA-Seq analysis
Drug-resistant parasites
Recuring gene mutations
B
[EMS]
0 mM Mock
2.5 mM A
2.5 mM B
5 mM C
5 mM D
5 mM E
7 mM F
7 mM G
Independent mutagenesis
EMS mutagenesis
Selection with AN13762
Cloning of resistant mutants
None
A1
B1
C1
D1
E1
F1
G1
Transcriptome sequencing analysis by RNA-Seq
C
Mock
AN13762
WT
A1
D
AN13762
Plaque size (mm2)
n.d.
WT A1 B1 C1 D1 E1 F1 G1
n: 0 50 50 50 50 50 50 50
E
Mock
AN13762
WT
A1
F
DHFR
ASP3
RON2
TUBA1
MYOA
MORC
AMA1
DGK2
FRM1
CPSF3
TgGC
Synonymous variant
Non-synonymous variant
Insertion
Deletion
Frameshift
G
G456S
S519C
Y328H,C
Y483N
E545K
Metallo-β-Lactamase
Beta-Casp
RNA specificity
CPSF73-100_C

**Figure 2**

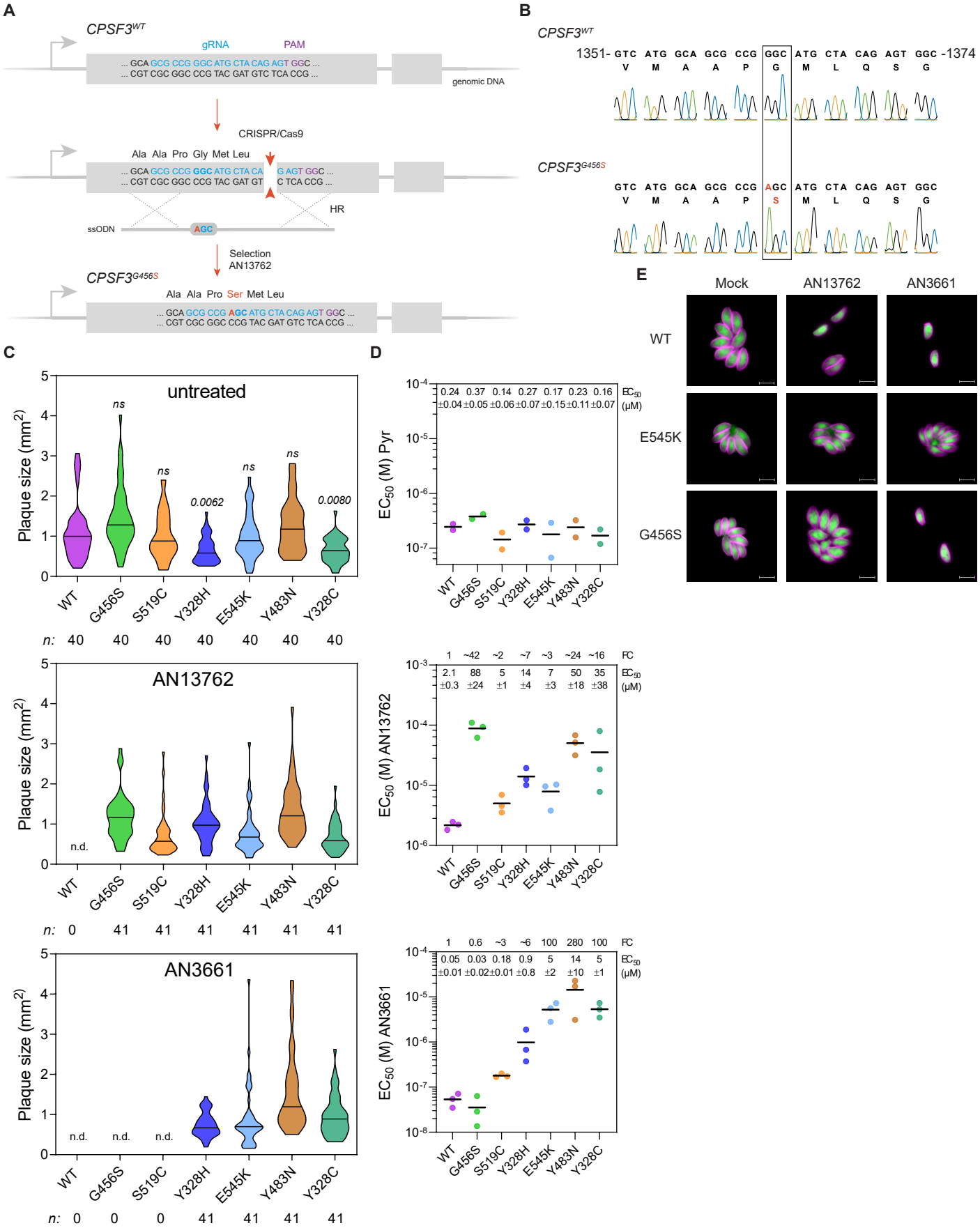
A
CPSF3WT
gRNA
PAM
genomic DNA
CRISPR/Cas9
Ala Ala Pro Gly Met Leu
HR
ssODN
AGC
Selection
AN13762
CPSF3G456S
Ala Ala Pro Ser Met Leu
B
CPSF3WT
1351- GTC ATG GCA GCG CCG GGC ATG CTA CAG AGT GGC -1374
V M A A P G M L Q S G
CPSF3G456S
GTC ATG GCA GCG CCG AGC ATG CTA CAG AGT GGC
V M A A P S M L Q S G
C
untreated
Plaque size (mm2)
ns
0.0062
0.0080
WT G456S S519C Y328H E545K Y483N Y328C
n: 40 40 40 40 40 40 40
AN13762
n.d.
n: 0 41 41 41 41 41 41
AN3661
n.d. n.d. n.d.
n: 0 0 0 41 41 41 41
D
EC50 (M) Pyr
0.24 0.37 0.14 0.27 0.17 0.23 0.16
±0.04 ±0.05 ±0.06 ±0.07 ±0.15 ±0.11 ±0.07
EC50 (μM)
EC50 (M) AN13762
1 ~42 ~2 ~7 ~3 ~24 ~16 FC
2.1 88 5 14 7 50 35
±0.3 ±24 ±1 ±4 ±3 ±18 ±38
EC50 (M) AN3661
1 0.6 ~3 ~6 100 280 100 FC
0.05 0.03 0.18 0.9 5 14 5
±0.01 ±0.02 ±0.01 ±0.8 ±2 ±10 ±1
E
Mock
AN13762
AN3661
WT
E545K
G456S

**Figure 3**

**A**

```
                                                  Metallo-β-Lactamase
                                                                                  Zn2
                    100                                 120                 140                                 160
TgCPSF3  SPLSAGVSGS PATVHIDASS KRRRLFLGED WVEITPLGAG CEVGRSCVIA RY-KGLTVMF DCGVHPAYSG LGALPIFDAV 157
ChCPSF3  -------EGE --TLYI---- -Q-------- -V----LGAG CEVGRSCVVV SF-KGRSVMF DCGIHPAFSG IGSLPVFDAI 71
PfCPSF3  --------G- ---------- ---------- ------GAS -EVGRSCVII ECDK-TSVML DCGIHPAFMG IGCLPIYDAY 51
HsCPSF3  ---------- -----I---- -R-------- -----PLGAG QEVGRSCIIL EF-KGRKIML DCGIHPGLEG MDALPYIDLI 59

                    180                                 200                 220                                 240
                 Zn1
TgCPSF3  DMTSVDVCLV THFHLDHCGA LPYLVTKTAF RGRVFMTEPT RVISKLVWL- -DYAR----- --MSA--FSQ ---GS-RD-N 221
ChCPSF3  DVSTIDLCLI THFHLDHSGA TPYFVSLTDF NGKVFMTEPT KAICKLVWQ- -DYAR----- --VNK--FSA ---GSIESEE 137
PfCPSF3  DISKVDLCLI THFHMDHSGA LPYLINKTRF KGRIFMTEAT KSICYLLWN- -DYARIEKYM NVVNKNKLSK NKKGG-EDEN 128
HsCPSF3  DPAEIDLLLI SHFHLDHCGA LPWFLQKTSF KGRTFMTHAT KAIYR--WLL SDYVK----- --VS------ ---------N 115

                 Apicomplexa  260  specific insertion                                                  Zn1
                                                  280                 300                                 320
TgCPSF3  Q---G--AAA AQAAAGS-Q- AE-KA----- ---G-G-A-F LYDEDDVDAT VRMVECLDFH QQV-EVGGIK VSCFGAGHVL 281
ChCPSF3  A-----PL-- ---S--S--- I--N------ ---------- LYTEKDIEKA INMTEIIDFR QQV-ELDGIR FSCYGAGHVL 183
PfCPSF3  GLNNGNMLLS NEYS--SDEN IDDNGDVYEN NDNGDGNSNV LYDENDIDKT MDLIETLNFH QNF-EFPNVK FTAYRAGHVI 205
HsCPSF3  -------I-- ------S--- AD-D------ ---------M LYTETDLEES MDKIETINFH E-VKEVAGIK FWCYHAGHVL 160

                    340                                 360  Y328H,C                380                                 400
                 Zn1-2                                   Zn2  *
TgCPSF3  GACMFLIEIG GVRMLYTGDF SRERDRHVPI AEVP---PVD VQLLICESTY GIHVHDDRQL RERRFLKAV- VDIV-NRGGK 356
ChCPSF3  GACMFLVEIG GVRILYTGDY SREDDRHVPR AEIP---PID VHVLICESTY GTRIHEPRID REKRFLGGV- QSIITRKG-K 258
PfCPSF3  GACMFLVEIN NIRFLYTGDY SREIDRHIPI AEIP---NID VHVLICEGTY GIKVHDDRKK REIRFLN-IL TSMIHNKG-K 280
HsCPSF3  GAAMFMIEIA GVKLLYTGDF SRQEDRHLMA AEIPNIKP-D I-LII-ESTY GTHIHEKREE REARFCNTV- HDIV-NRGGR 235

                    420                 Beta-Casp       440                 460                                 480
TgCPSF3  CLLPVFALGR AQELLLILEE YWTAHPEIRH -VPILFLSPL SSK--C-AVV FDAFVD-MCG EAVRSRALR- GENPFAFRFV 430
ChCPSF3  CLLPVFAIGR AQELLLILEE HWSRTPSI-Q NVPIIYASPM SIK--CMRV- FETYIN-QCG ESVR-RQADL GINPFQFNYI 332
PfCPSF3  VLLPVFALGR AQELLLILEE HWDKNKHL-Q NIPIFYISSM ATKSLC--I- YETFIN-LCG EFVK-KVVNE GKNPFNFKYV 354
HsCPSF3  GLIPVFALGR AQELLLILDE YWQNHPEL-H DIPIYYASSL AKK--CMAV- YQTYVNAM-N DKIR-KQIN- INNPFVFKHI 308

                    500          G456S                  520                 Y483N  540                                 560
                                  *                                            *
TgCPSF3  KNVKSV-EAA RVYIH-HD-G PAVVMAAPGM LQSGASREIF EAWAPDAKNG VILTGYSVKG TLADELKREP E---TI--Q- 501
ChCPSF3  KTVNSLNE-I KDIIY--NPG PCVVMAAPGM LQNGTSRDIF EIWAPDKRNG IILTGYAVRG TPAYELRKEP E---MI--Q- 403
PfCPSF3  KYAKSL-ESI SSYLYQDN-N PCVIMASPGM LQNGISKNIF NIIASDKKSG VILTGYTVKG TLADELKTEP E---FV--T- 426
HsCPSF3  SNLKSM-D-- ----HFDDIG PSVVMASPGM MQSGLSRELF ESWCTDKRNG VIIAGYCVEG TLAKHIMSEP EEITTMSGQK 381

                 S519C  580  Zn2                 RNA specificity  600  Zn2  E545K                620                                 640
                   *                                                          *
TgCPSF3  LPDRVLRRRC SFEMISFSAH SDYQQTQEFI GK-LKVPNVV LVHGERGEMR RLKEKLEEE- RPALSVFTPE ILQK--VSLQ 577
ChCPSF3  LGEKVIPMRA KFDQISFSAH SDFTQTQEFI NS-LKVPNVI LVHGERGECK KLKDKL-KEL SPSLAVFAPE ILQKVGLTFP 481
PfCPSF3  INDKVVKRKC RFEQISFSAH SDFNQTKTFI EK-LKCPNVV LVHGDKNELN RLKNKLIEE- KQYLSVFTPE LLQK--LSFH 502
HsCPSF3  LP---LKM-- SVDYISFSAH TDYQQTSEFI -RALKPPHVI LVHGEQNEMA RLKAALIRE- YE--D--NDE V-H---IEVH 446
```

**B**

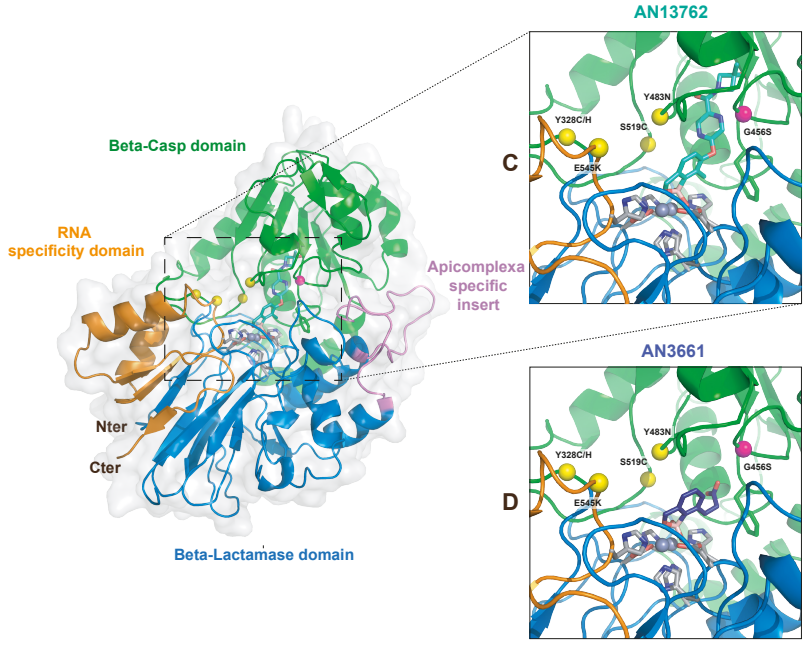


**Figure 4**

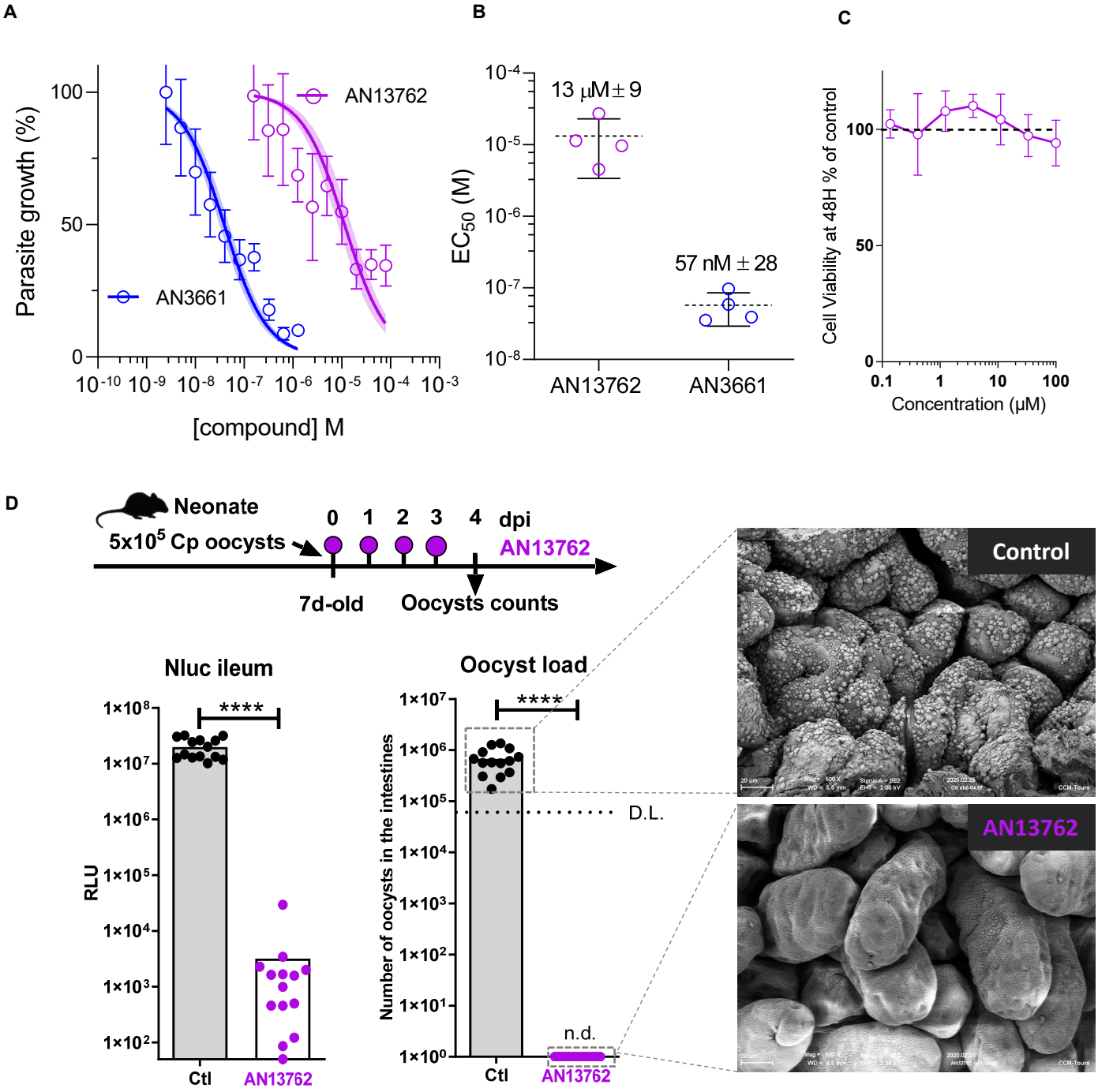
A
Parasite growth (%)
100
50
0
AN13762
AN3661
10-10 10-9 10-8 10-7 10-6 10-5 10-4 10-3
[compound] M
B
EC50 (M)
10-4
10-5
10-6
10-7
10-8
13 µM ± 9
57 nM ± 28
AN13762
AN3661
C
Cell Viability at 48H % of control
100
50
0
0.1
1
10
100
Concentration (µM)
D
Neonate
5x10^5 Cp oocysts
0 1 2 3 4 dpi
AN13762
7d-old
Oocysts counts
Nluc ileum
RLU
1×10^8
1×10^7
1×10^6
1×10^5
1×10^4
1×10^3
1×10^2
****
Ctl
AN13762
Oocyst load
Number of oocysts in the intestines
1×10^7
1×10^6
1×10^5
1×10^4
1×10^3
1×10^2
1×10^1
1×10^0
****
D.L.
n.d.
Ctl
AN13762
Control
AN13762

Figure 5