Supporting Information

Probing the Open Global Health Chemical Diversity Library for multistage-active starting points for next-generation antimalarials

Authors: Matthew Abraham, Kerstin Gagaring, Marisa L. Martino, Manu Vanaerschot, David M. Plouffe, Jaeson Calla, Karla P. Godinez-Macias, Alan Y. Du, Melanie Wree, Yevgeniya Antonova-Koch, Korina Eribez, Madeline R. Luth, Sabine Ottilie, David A. Fidock, Case W. McNamara, Elizabeth A. Winzeler

*Correspondence to Elizabeth Winzeler (ewinzeler@ucsd.edu)

Table of Contents

Table S1. This file contains a description of the library, as well as results from cheminformatic as well as primary screening results (attached Excel file).

Table S2. GHCDL library characteristics

	mean	*min*	*max*
Molecular weight (Da)	320.7	184.2	492.2
H-bond donor	0.9		
H-bond acceptor	5.6		
logP	2.0	−3.27	– 6.35
TPSA (Å2)	64.2	16.13	160.88
FSP3	0.48	0	10
Cluster size	5.19	2	2131

Table S3. Known antimalarial compounds used for assessment of diversity (attached Excel file)

Table S4. ABS screening reconfirmation data (attached Excel file)

Table S5. Liver stage reconfirmation data (attached Excel file)

Table S6. Gametocyte reconfirmation data (attached Excel file)

Table S7. Genome sequencing statistics

Sample Name	*Mean Whole Genome Coverage*	*Number of Reads Aligned to Reference Genome*	*Percent Reads Mapping to Reference Genome*
Dd2-B2-Parent	79.8	24041607	92.9
DDD01061024--1-F5	94.3	32378286	92.3
DDD01061024--2-E8	74.3	25869385	92.4
DDD01061024--2-G6	80.3	26765318	89

All sequences have been deposited in the shortread sequence archive under accession PRJNA563182.

Table S8. Whole genome variant-calling results for lines evolved to resistance to DDD01061024.

Chromosome	*Position*	*Gene_Name*	*Gene_Descrip*	*Quality*	*Ref_ Base*	*Alt_ Base*	*Type*	*Effect*
Pf3D7_03_v3	63025			502.68	T	A	SNP	intergenic
Pf3D7_08_v3	1296769	PF3D7_0830500	tryptophan threonine-rich antigen TryThrA	6280.51	T	A	SNP	missense
Pf3D7_09_v3	39010			529.89	G	A	SNP	intergenic
Pf3D7_10_v3	1606365	PF3D7_1040200	stevor	1216.76	T	A	SNP	synonymous
Pf3D7_10_v3	1610720			1291.59	G	T	SNP	intergenic
Pf3D7_10_v3	1648826	PF3D7_1041300	erythrocyte membrane protein PfEMP1 VAR	552.72	G	T	SNP	missense
Pf3D7_10_v3	1648861	PF3D7_1041300	erythrocyte membrane protein PfEMP1 VAR	791.96	C	G	SNP	missense
Pf3D7_10_v3	1648866	PF3D7_1041300	erythrocyte membrane protein PfEMP1 VAR	854.96	T	G	SNP	synonymous
Pf3D7_10_v3	1648868	PF3D7_1041300	erythrocyte membrane protein PfEMP1 VAR	884.96	C	G	SNP	missense
Pf3D7_10_v3	1648873	PF3D7_1041300	erythrocyte membrane protein PfEMP1 VAR	905.96	T	G	SNP	missense
Pf3D7_10_v3	1648887	PF3D7_1041300	erythrocyte membrane protein PfEMP1 VAR	1205.87	T	C	SNP	synonymous
Pf3D7_10_v3	1648888	PF3D7_1041300	erythrocyte membrane protein PfEMP1 VAR	1508.97	T	C	SNP	missense
Pf3D7_11_v3	1981163			1201.88	A	T	SNP	intergenic
Pf3D7_13_v3	1296614	PF3D7_1330800	RNA-binding protein, putative	4628.52	G	T	SNP	missense
M76611	**4266**	**mal_mito_3**	**apocytochrome b cyb**	**320615.27**	**G**	**T,C**	**SNP**	**missense**
Pf3D7_03_v3	420818	PF3D7_0310000	50S ribosomal protein L9, apicoplast, putative	912.11	GA	G,GAA	INDEL	frameshift
Pf3D7_10_v3	1648856	PF3D7_1041300	erythrocyte membrane protein PfEMP1 VAR	986.92	T	TA	INDEL	frameshift
Pf3D7_10_v3	1648857	PF3D7_1041300	erythrocyte membrane protein PfEMP1 VAR	902.92	T	TGACAA	INDEL	frameshift
Pf3D7_10_v3	1648869	PF3D7_1041300	erythrocyte membrane protein PfEMP1 VAR	878.92	ACCT	A	INDEL	Disruptive inframe deletion
Pf3D7_13_v3	2159708			593.65	AT	A	INDEL	intergenic
Pf3D7_14_v3	1097185	PF3D7_1428000	conserved Plasmodium membrane protein unknown function	547.3	AT	A	INDEL	intron

Table S9. Genotyping results for clones resistant to DDD01061024.

Chromosome	*Position*	*Gene_Name*	*Fidock-Dd2B2-20170519-Parent*	*Fidock-Dd2B2-DDD01061024--1-F5*	*Fidock-Dd2B2-DDD01061024--2-E8*	*Fidock-Dd2B2-DDD01061024--2-G6*
Pf3D7_03_v3	63025		No mutation	Homozygous mutation	Mixed call	Mixed call
Pf3D7_08_v3	1296769	PF3D7_0830500	No mutation	Homozygous mutation	No mutation	No mutation
Pf3D7_09_v3	39010		No mutation	No mutation	Homozygous mutation	Homozygous mutation
Pf3D7_10_v3	1606365	PF3D7_1040200	No mutation	Homozygous mutation	Mixed call	Mixed call
Pf3D7_10_v3	1610720		No mutation	Mixed call	Mixed call	Homozygous mutation
Pf3D7_10_v3	1648826	PF3D7_1041300	No mutation	Homozygous mutation	Mixed call	Mixed call
Pf3D7_10_v3	1648861	PF3D7_1041300	No mutation	Homozygous mutation	No mutation	Mixed call
Pf3D7_10_v3	1648866	PF3D7_1041300	No mutation	Homozygous mutation	No mutation	Mixed call
Pf3D7_10_v3	1648868	PF3D7_1041300	No mutation	Homozygous mutation	No mutation	Mixed call
Pf3D7_10_v3	1648873	PF3D7_1041300	No mutation	Homozygous mutation	No mutation	Mixed call
Pf3D7_10_v3	1648887	PF3D7_1041300	No mutation	Homozygous mutation	No mutation	Mixed call
Pf3D7_10_v3	1648888	PF3D7_1041300	No mutation	Homozygous mutation	No mutation	Mixed call
Pf3D7_11_v3	1981163		No mutation	No mutation	No mutation	Homozygous mutation
Pf3D7_13_v3	1296614	PF3D7_1330800	No mutation	No mutation	No mutation	Homozygous mutation
M76611	**4266**	**mal_mito_3**	**No mutation**	**Homozygous mutation**	**Homozygous mutation**	**Homozygous mutation**
Pf3D7_03_v3	420818	PF3D7_0310000	No mutation	Mixed call	Mixed call	Mixed call
Pf3D7_10_v3	1648856	PF3D7_1041300	No mutation	Homozygous mutation	No mutation	Mixed call
Pf3D7_10_v3	1648857	PF3D7_1041300	No mutation	Homozygous mutation	No mutation	Mixed call
Pf3D7_10_v3	1648869	PF3D7_1041300	No mutation	Homozygous mutation	No mutation	Mixed call
Pf3D7_13_v3	2159708		No mutation	Mixed call	Mixed call	Mixed call
Pf3D7_14_v3	1097185	PF3D7_1428000	No mutation	Mixed call	Mixed call	Mixed call

The data show the positions that were mutated in Supplemental Table S8.

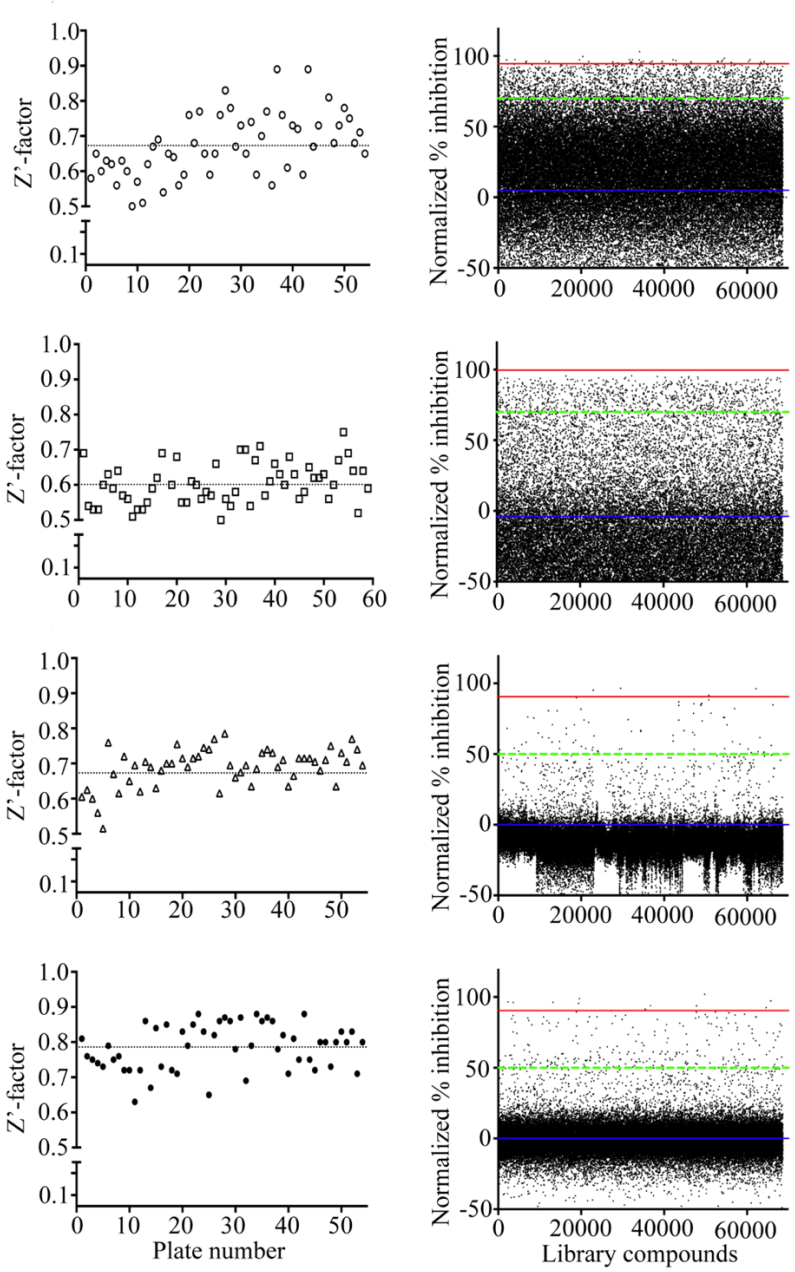

Figure S1. Primary screen assay statistics.
The Z'-factor was calculated per plate throughout the GHCDL HTS campaign. Screen averages (± SD) included a, 0.67±0.089 PbLuc (○), b, 0.60±0.062 ABS (•), c, 0.69±0.054 stage V GAMs (Δ), and d, 0.79±0.065 HepG2 cytotoxicity (●). Averages are depicted for each screen by the dotted line, and plates with a Z'-factor below 0.5 were repeated. Library-wide inhibition plots (column 2) show compound activity (black spots) with mean negative control (blue line), mean positive control (red line), and cutoff criteria for subsequent screening (green dashed line).

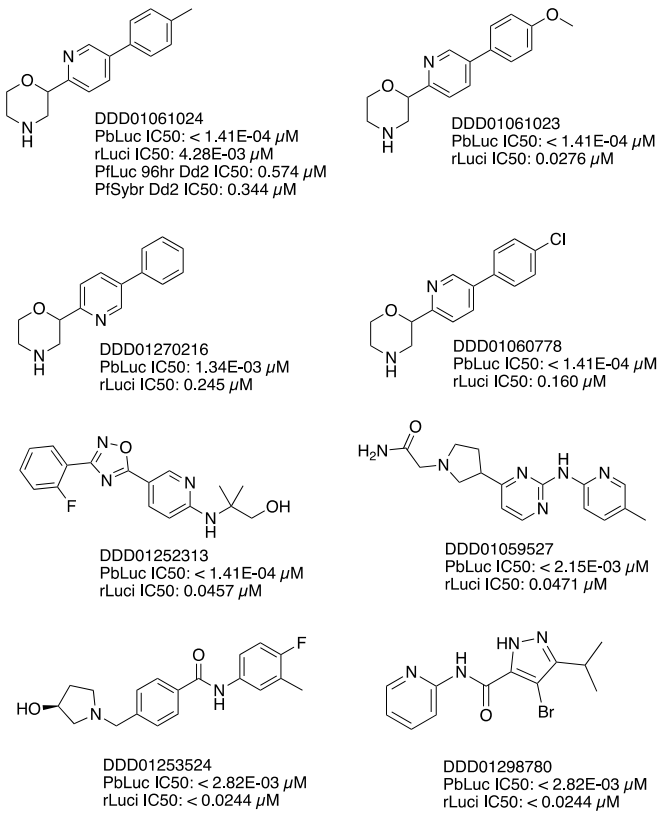

Figure S2. Recombinant luciferase inhibitors.
Compounds tested against firefly luciferase in dose response produced potent inhibition of overall bioluminescence. A cluster of four closely related analogs is heavily enriched in an otherwise diverse pool of molecules. All values are determined from biological duplicates and technical quadruplicates

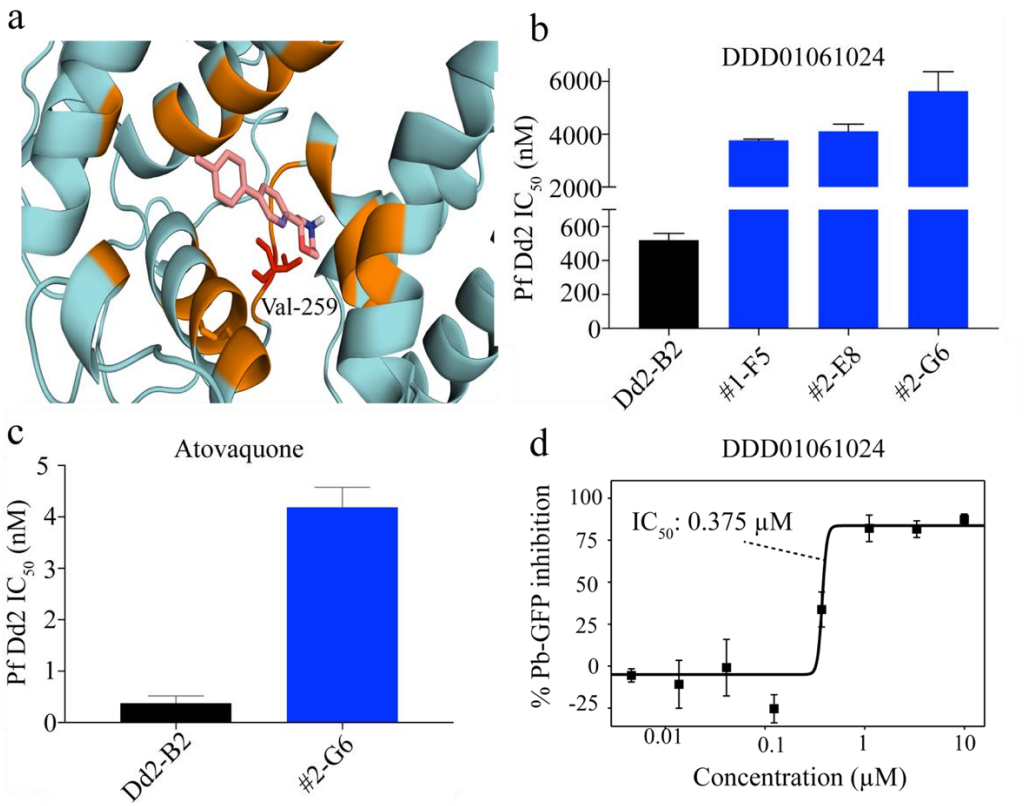

Figure S3. DDD01061024 is a cytochrome bc1 inhibitor.
The dual stage inhibitor, DDD01061024, is shown occupying the Q_0 site (residues outlined in orange) of a Pf Cyt Bc1 homology model (from PDB: 1BE3) (a). A local V259L substitution (red) in mutant clones may destabilize the interactions within this pocket, helping resistant clones tolerate higher doses of the inhibitor (b). DDD01061024-resistannt mutants show cross-resistance to atovaquone (c), another multistage active antimalarial which targets the Q_0 binding site. Dose dependent inhibition was determined using a metabolic endpoint (LDH) before normalizing to controls (n=2). Measuring the percentage of PbGFP infected hepatocytes using a high content imaging system validates the liver-stage potency of DDD01061024 in a fluorescence-based model (d). Infected hepatocytes were counted using luminescence-independent high content imaging in treated wells and normalized to controls over two replicates.

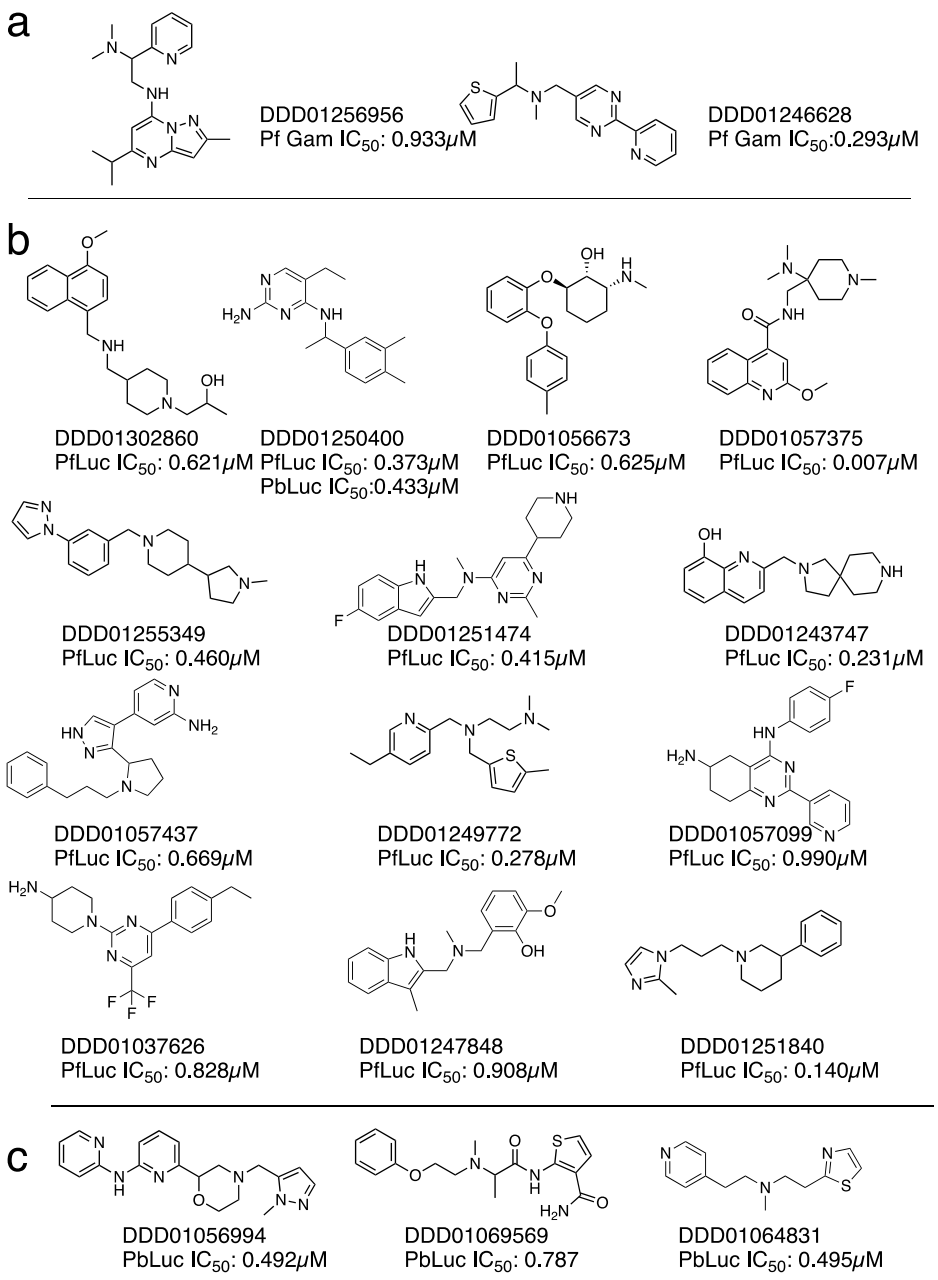

Figure S4. Potent singletons.
Structurally distinct scaffolds that showed no similarity to other hits using cutoff of 70% (DSC < 0.7). Submicromolar potency was reconfirmed for these unique scaffolds in their respective stage of action for a, stage V gametocytes; b, ABS, and c, liver stages. For Pfluc, *P. falciparum* Dd2 dose response values are shown for the 96hr time point. All values are determined from biological duplicates and technical quadruplicates