Prediction model for anti-malarial activities of hemozoin inhibitors using physicochemical properties

Running title: *"Prediction model for anti-malarial activities of hemozoin inhibitors"*

Farhana Mosaddeque[1,2,⊕], Shusaku Mizukami[3,⊕], Mohamed Gomaa Kamel[4,5,⊕], Awet Alem Teklemichael[1,2], Truong Van Dat[6], Satoshi Mizuta[7], Dinh Van Toan[6], Ali Mahmoud Ahmed[4,8], Nguyen Lam Vuong[4,6], Mohamed Tamer Elhady[4,9], Hoang Thi Nam Giang[4,10], Tran Ngoc Dang[4,6,11], Michiko Fukuda[1], Lam K. Huynh[12], Yoshimasa Tanaka[7], Timothy J. Egan[13], Osamu Kaneko[2,14,15], Nguyen Tien Huy[2,3,*], Kenji Hirayama[1,2,15,*]

[1]*Department of Immunogenetics, Institute of Tropical Medicine (NEKKEN), Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan*

[2]*Leading Program, Graduate School of Biomedical Sciences, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan*

[3]*Department of Clinical Product Development, Institute of Tropical Medicine (NEKKEN), Graduate School of Biomedical Sciences, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan*

[4]*Online Research Club (http://www.onlineresearchclub.org/)*

[5]*Faculty of Medicine, Minia University, Minia 61519, Egypt*

[6]*University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, 70000, Vietnam*

[7]*Center for Bioinformatics and Molecular Medicine, Graduate School of Biomedical Sciences, Nagasaki University, 1-14 Bunkyo, Nagasaki, Japan*

[8]*Faculty of Medicine, Al-Azhar University, Cairo 11651, Egypt*

[9]*Zagazig University Hospitals, Department of Pediatrics, Sharkia 44511, Egypt*

[10]*Faculty of Medicine and Pharmacy, The University of Da Nang, Da Nang, Vietnam*

[11]*Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba 305-8577, Japan*

[12]*International University, VNU-HCM, Ho Chi Minh City, Vietnam*

[13]*Department of Chemistry, University of Cape Town, Private Bag, Rondebosch, 7701, South Africa*

[14]*Department of Protozoology, Institute of Tropical Medicine (NEKKEN), Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan*

[15]*School of Tropical Medicine and Global Health, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan*

[⊕]*Authors equally contributed the work.*

**Correspondence to: Nguyen Tien Huy (E-Mail: tienhuy@nagasaki-u.ac.jp) or Kenji Hirayama (E-Mail: kenji@nagasaki-u.ac.jp).*

Emails:

FM: farhana.nu.10@gmail.com

SM: mizukami@nagasaki-u.ac.jp (ORCID: 0000-0002-7722-4884)

MGK: mohamedgomaa@s-mu.edu.eg (ORCID: 0000-0001-9848-2965)

AAT: awetekle@gmail.com

TVD: dattv@ump.edu.vn

SM: s-mizuta@nagasaki-u.ac.jp

DVT: mr.dvtoan@gmail.com

AMA: ali.m.ahmed@azhar.edu.eg

NLV: nguyenlamvuong@ump.edu.vn

MTE: mtielhd@gmail.com (ORCID: 0000-0003-1339-1138)

HTNG: htngiang@ud.edu.vn

TND: ngocdangytcc@gmail.com

MF: m-fuku@nagasaki-u.ac.jp

LKH: lamhuynh.us@gmail.com (ORCID: 0000-0002-3683-2787)

YT: ystanaka@nagasaki-u.ac.jp

TJE: Timothy.Egan@uct.ac.za

OK: okaneko@nagasaki-u.ac.jp

NTH: tienhuy@nagasaki-u.ac.jp (ORCID: 0000-0002-9543-9440)

KH: hiraken@nagasaki-u.ac.jp (ORCID: 0000-0001-9467-1777)

Keywords: Malaria – Anti-malarial – Hemozoin – Prediction – Drug.

Abstract

The rapid spread of strains of malaria parasites resistant to several drugs has threatened global malaria control. Hence, the aim of this study was to predict the anti-malarial activity of chemical compounds possessing anti-hemozoin formation activity as a new means of anti-malarial drug discovery. After the initial *in vitro* anti-hemozoin formation high-throughput screening (HTS) of 9,600 compounds, a total of 224 hit compounds were identified as hemozoin inhibitors. These 224 compounds were tested for *in vitro* erythrocytic anti-malarial activity at 10 μM using the chloroquine-mefloquine sensitive *Plasmodium falciparum* strain, 3D7A. Two independent experiments were conducted. The physicochemical properties of the active compounds were extracted from ChemSpider and SciFinder databases. We analyzed the extracted data using Bayesian model averaging (BMA). Our findings revealed that lower numbers of S atoms, lower values of log D pH 3, 4 and 5, and higher values of ACD log D pH 7.4 had a significant association with anti-malarial activity among compounds possessing anti-hemozoin formation activity. The BMA model revealed an accuracy of 91.23%. We report new prediction models containing the physicochemical properties that shed light on effective chemical groups for synthetic anti-malarial compounds and help *in silico* screening for novel anti-malarial drugs.

Introduction

Despite the recent advances in anti-malarial drug discovery and development, malaria remains one of the most serious medical burdens worldwide, resulting 212 million new cases with 429,000 deaths in 2015 (1). According to the different stages of the parasite life cycle, several therapies can affect its stages and exert their therapeutic roles. Quinine, amodiaquine, chloroquine, and mefloquine, which contain a quinoline scaffold, are fast-acting and highly effective blood schizontocidal therapies against malaria parasites, including *Plasmodium falciparum* (*PF*) (2).

Indeed, the resistance of malaria to chloroquine and other 4-aminoquinoline-based therapies, in addition to the antifolate combination sulfadoxine/pyrimethamine, has turned the light on artemisinin-based combinations to achieve higher response rates (3, 4). However, the rapidly spreading resistance of *PF* to artemisinin-based combinations has been reported causing a global challenge for malaria control (5, 6). Thus, it is important to discover new anti-malarial drugs, especially for malaria-endemic countries.

Recently, several new classes of anti-malarials have entered clinical studies, in patients with malaria, such as the fast-acting agents KAF156 (7), cipargamin (8), and artefenomel (9), whereas ferroquine remains the only long-acting novel anti-malarial in clinical development (10, 11). However, these drugs have not been approved yet and no vaccine to help in prevention, control, elimination, and eradication of malaria has been approved yet. Only one vaccine candidate, RTS, S/AS01, reached phase-III clinical trials with relatively low efficacy (12–14). Therefore, there is an urgent need for the discovery and development of novel anti-malarial chemotherapies for which there are no preexisting resistance mechanisms.

As of today, one of the most promising and ideal targets is interference with the parasite's heme detoxification pathway which is the target for some current anti-malarial drugs like quinine that is still efficacious against chloroquine-resistant *PF* (15–19). Recently, inhibition of the heme

detoxification pathway of the parasite has been highlighted as a target in several anti-malarial screening projects (20–22). This target is based on the inhibition of hemozoin which is a crystalline pigment produced by the malaria parasite as a result of the hemoglobin degradation process to protect it against the toxic heme produced as an end product of hemoglobin catabolism (23, 24).

Hemozoin formation is a protective physiochemical process that needs parasite protein (25–27) and/or food vacuole lipids or membranes (28, 29) for synthesis. Therefore, lipophilic detergents that mimic the intra-parasite condition like Nonidet P-40 and Tween 20 can be used as surrogate substances for high-throughput screening (HTS) of novel anti-malarials because they have the ability to promote the crystallization of heme (20, 30). This makes hemozoin inhibition suitable for research using HTS assays to build prediction models for novel anti-malarial drugs.

Recently, several studies used HTS and predicted models for β-hematin, synthetic hemozoin, inhibitors. Sandlin *et al* have screened 144,330 and produced 530 hits of which 171 were active against parasites and 73 had parasite IC_{50}s below 5 μM and 25 below 1 μM. (31). In addition, using physiochemical properties (22), we have recently developed an *in silico* model to predict drug-like compounds possessing anti-hemozoin activity.

As previously suggested, prediction models possess advantages for anti-malarial design because other approaches such as analog development based on existing agents or natural products mainly detect new anti-malarials by the chemical modifications of previously known compounds (32), however, new anti-malarial compounds can be discovered by the prediction equation based on a well-known metabolic target. Thus, the prediction models aid discovery of new chemical scaffolds. Moreover, specialized labware and expensive equipment are not required in these models, so millions of library compounds can be screened in-silico by using the prediction models. Also, the relationship between the compound's properties and anti-hemozoin activity is interpreted from the prediction models. Therefore, we continued the previous work by developing new prediction

models for novel anti-malarial activities of hemozoin inhibitors using the physiochemical properties of these small chemical compounds.

Methods

Ethics statement

Experiments which required human materials, red blood cell (RBC) and serum, were approved by the institutional ethical review board of Institute of Tropical Medicine (NEKKEN), Nagasaki University.

Materials

Chloroquine (Sigma-Aldrich Chemical Company, UK) was used to monitor the assay system, stored in 4°C freezer and mixed well before use. Dimethyl sulfoxide (DMSO; Wako Pure Chemicals Ltd., Osaka, Japan) was used as negative control. SYBR Green I (10,000 × stock conc.) was obtained from Lonza (Lonza, Rockland, ME, USA) and stored in -30 °C. Alamar Blue was purchased from Funakoshi (Funakoshi Co., Tokyo, Japan). Lysis buffer containing Tris (20 mM; pH 7.5), EDTA (10 mM), saponin (0.01%, wt/vol) and Triton X-100 (0.1%; vol/vol) was prepared in advance and kept at 4 °C. Human O^+ RBC and serum were obtained from the Japanese Red Cross Society (Reception Number: 28J0060). Compounds, which were identified positive from our previous study (22) were obtained from the Open Innovation Center for Drug Discovery (OCDD), University of Tokyo (Tokyo, Japan). All compounds received from the University of Tokyo at 10 µM concentration, however, the stock solutions were prepared in DMSO at 2 mM concentration. The final DMSO concentration used in the experiment was ≤ 0.5% which had no inhibitory effect on parasite culture.

Parasite culture

For this study, chloroquine-mefloquine sensitive *PF* strain 3D7A was used. *PF* strain (3D7A) was cultured and maintained as previously described (33) with slight modifications (34). In brief,

parasites were maintained in RPMI 1640 (Thermo Fisher Scientific, MA, USA) medium supplemented with 5% AB^+ human serum, 0.25% Albumax I (Thermo Fisher Scientific), 12.5 µg/ml gentamicin (Sigma-Aldrich) and human RBC (O^+) at a hematocrit level of 2% and 0.2 – 2% parasitemia in culture flasks at 37˚C under an atmosphere of 5% CO_2, 5% O_2 and 90% N_2.

Asexual anti-malarial assay

***In vitro* assay using chloroquine-mefloquine sensitive *PF* strain, 3D7A**

The erythrocytic anti-malarial activity of the compounds was measured by growth inhibition (%) of the parasites in the presence of compounds using SYBR Green I assay. In brief, 50 µL of *PF* strains were seeded into 96 well clear bottom black plate (Thermo Fisher Scientific) with 0.75% parasitemia and 2% hematocrit followed by addition of 50 µl of the compound at 10 µM final concentration. Wells, with the absence of parasite, are considered as a positive control and parasites in DMSO (≥ 5%) without any treatment considered as negative control. Plates were then incubated for 48 hours at 37°C with 5% CO_2, 5% O_2, and 90% N_2 using a closed Jar. After 48 hours of incubation at standard culture condition, 100 µl of lysis buffer containing SYBR Green I (1× final concentration) were directly added to the wells. Plates were then placed on a shaker with the gentle mix to incubate for 1 hour at room temperature in the dark. The fluorescence of each well then measured using a multi-plate reader (ARVO1430; PerkinElmer, MA, USA) in a fluorescence detection mode (Ex = 485 nm, Em = 515 nm) for a 0.1-second exposure. Two independent experiments were performed under similar condition. Subsequently, an *in vitro* dose-response assay was performed.

In vitro anti-malarial dose-response assay

Active compounds identified from the previous anti-malarial assay eventually goes for *in vitro* erythrocytic dose-response assay at ten different dilutions ranged between 0.5 nM and 10 µM to exclude any false positive compounds and measure the half maximal inhibitory concentrations

175 (IC_{50}). Each experiment was conducted twice using 96 well clear bottom black plate and the
176 fluorescence was measured as described above. Obtained data were analyzed to calculate the IC_{50}
177 value of each compound based on a typical sigmoid dose-response by using GraphPad Prism 6
178 software.

179 ***Anti-hemozoin dose-response assay***

180 The IC_{50} of potential hit compounds against hemozoin were further performed as previously
181 described (22). Briefly, hemin solution (10 mM heme in DMSO and 100 mM acetate buffer) were
182 incubated with detergent NP-40 and compounds (0 µM to 208 µM) for 250 minutes at 37˚C,
183 followed by addition of pyridine solution with 10 minutes shaking. Finally, the absorbance of each
184 compound was measured at 405/705 nm. The IC50 of each compound were calculated using
185 GraphPad Prism version 6 software.

186 ***Physicochemical properties of 224 hemozoin inhibitors***

187 Physicochemical properties of chemical compounds were obtained from SciFinder Scholar
188 Database (American Chemical Society, Washington, DC) or ChemSpider (www.chemspider.com)
189 as previously described (22).

190 ***Statistical analysis***

191 We conducted blinded experiments without knowledge of compound' names or structures. The
192 code was only opened after sending the anti-malarial and anti-hemozoin results to the Open
193 Innovation Center for Drug Discovery, University of Tokyo. Moreover, the statistician who
194 performed the analysis is not in the chemistry field and was not involved in the experiments.
195 Variables of physical properties and compounds with many missing data including pKa2, vapor
196 pressure (Torr), flash point (°C), enthalpy of vaporization (kJ/mol), and boiling point (°C) were
197 excluded. Finally, 206 compounds were included in our analysis.

198 The outcome variable was the anti-malarial activity of the chemical compounds (coded as 1 = active,
199 0 = inactive). The predictor variables were the physicochemical properties of our chemical
200 compounds. We performed uni-variable logistic regression (LR) to examine the association
201 between physicochemical properties and the aforementioned outcome. Subsequently, variables with
202 p-values *(P)* < 0.1 and/or with a prior significance, in our previous study (22) were included in the
203 multi-variable analyses using Bayesian model averaging (BMA) to find the independent predictors
204 of our outcome. We have multiplied index of refraction variable by 10 to have a smaller odds ratio
205 (OR).
206 *Development and validation of the prediction models*
207 We divided the original data randomly into training and testing sets with a ratio of 70:30, (nearly
208 157 (including 17 positive compounds) and 67 compounds (including 5 positive compounds),
209 respectively). The training data set was constructed to develop prediction models using the model.
210 Results are shown as a posterior probability (PP) and Bayesian information criteria (BIC) of each
211 model (35, 36) (Tables 2 and 3). The PP, P (B≠0), is the probability that a predictor variable has an
212 effect on the compound activity and a co-efficient is nonzero. Then, the best models were illustrated
213 visually by depicting the variables included in them (Figure 1).
214 *Comparison between the best models*
215 The sensitivity and specificity of the best models were calculated. The discriminatory powers of the
216 best prediction models obtained from BMA were compared according to the area under the curve
217 (AUC) from the receiver operating characteristic (ROC), and their accuracies (Figure 2). All
218 analyses were performed using RStudio software version 1.0.44 (https://www.rstudio.com/). $P <$
219 0.05 was considered statistically significant in all analyses. The data and R script can be obtained
220 from MGK or NTH.
221 **Results**
222 ***In vitro anti-malarial assay***

223 A total of 224 compounds having hemozoin inhibitory activity (22) were selected for *in vitro* anti-
224 malarial assay. Among those, 30 compounds having ≥ 45% growth inhibitory activity at 10 μM
225 concentration were further subjected to a dose-response assay to remove false positive compounds
226 from initial screening (Figure 1), resulting only 22 compounds with a clear sigmoid dose-response
227 curve to determine IC_{50} (Figure 2 and Table 1).
228 The interference of 22 compounds with SYBR Green l in the presence of malaria parasites was
229 measured by culturing *Plasmodium falciparum* 3D7A for 48 h and then adding 22 compounds
230 followed by addition of SYBR Green l with 1 h incubation in the dark. Subsequently, the
231 fluorescence was measured by using a multiplate reader (ARVO1430; PerkinElmer, MA) in the
232 fluorescence detection mode (excitation [Ex] at 485 nm and emission [Em] at 515 nm) for a 0.1-s
233 exposure. The results showed all fluorescence variations were less than 10%, indicating that there
234 was almost no interference of the compounds, artemisinin, and chloroquine with SYBR Green – l
235 fluorescence (Supplementary Figure). Furthermore, our results demonstrated that all 22 compounds
236 possessed anti-hemozoin with IC_{50} ranging from 4.58 to 198.1 μM (Table 1).
237 ***Development of prediction models for anti-malarial activities***
238 Number of C, and S atoms, H acceptors, H donor acceptor sum, log D pH 1, 2, 3, 4 and 5, mass
239 solubility pH 1 and 3 and polarizability had *P* value < 0.1 in the uni-variable LR, and log P, KOC
240 pH 5.5 and 7.4, log D pH 5.5 and 7.4, density, refractive index, rule number of five violations,
241 number of freely rotatable bonds, H bond donors, H, and O were significant from the previous study
242 (22), while number of N atoms and surface tension were significant from the current analysis and
243 the previous study, respectively (22) (Table 2).
244 The BMA analysis selected 37 models, from which the best 5 models had 5 variables, as the best
245 predictors of the chemical compound activity, including the number of S atoms, log D pH 3, 4 and
246 5, and ACD log D pH 7.4. (Figure 3, Tables 2 and 3). This reflected that the anti-malarial activity

was significantly associated with lower number of S atom, lower log D pH 3, 4, and 5, and higher ACD log D pH 7.4.

Validation of the prediction model

The BMA model, with a cut-off of 0.09, revealed a sensitivity, specificity, and maximum accuracy of 60%, 69.23%, and 91.23%, respectively.

Discussion

Our BMA prediction model identified several physicochemical properties from which we can design to develop novel drugs. We found that a lower number of S atoms and lower log D value at pH 3, 4 and 5 and a higher ACD log D value at pH 7.4 could serve as good predictors while developing novel anti-malarial drugs.

Ignatushchenko *et al.* have found that the replacement of the ring oxygen by sulfur to produce 4, 5-dihydroxythioxanthone, results in a decreased inhibitory activity on hemozoin formation and parasite growth when compared to the corresponding xanthone homolog (37).

A low log D, a low lipophilicity through high inhibitory activities against hemo crystallization, at pH 3, 4 and 5 were significantly associated with a better anti-malarial activity. It is well known that the anti-malarial drugs must enter to the parasite food vacuole where the pH is around 4 – 5.5. Noteworthy, log P is an indicator of lipophilicity and is a component of the Lipinski's rule of five, a rule of thumb, to predict the drug-likeness. However, log D is a better indicator at specific pH environment. Because the hemozoin formation occurs inside the acidic food vacuole, such compounds are expected to decrease its lipophilicity after the uptake into the parasite food vacuole and to be accumulated within it.

Log P is widely used, however, it fails to take into account the variation in the lipophilicity of a drug regarding the ionic states present at key biological pH values. Given that the majority of commercial pharmaceuticals contain an ionizable moiety, Bhal *et al.* that log D is a better descriptor for lipophilicity in the context of the rule of five. It gives more physiologically relevant results,

subsequently declining the potential false-negatives number incorrectly eliminated in screening. They also showed that the adapted rule of five using log D instead of log P gives a notable improvement in the pass rate for compounds that have the desired lipophilicity at a relevant physiological pH (38).

In 11 p-aminopyridine anti-malarials a quadratic association between log the lipid accumulation ratio (LAR), anti-log log D 7.4, and log resistance index was confirmed while log (LAR/vacuolar accumulation ratio (VAR)) versus log resistance index for 12 was linear. Both might be able to predict the utility of structural adjustments (39).

This confirms for chloroquine-sensitive parasites, that while charged drug concentration in food vacuole increases as vacuolar pH decreases, the concentration of uncharged drug (base) in vacuolar membrane and other lipid sites in the digestive vacuole has a constant value for each individual drug, because the proportion of drug partitioned into lipid decreases as log D (the log of [drug]lipid / [drug]water at equilibrium) decreases with a reduced pH.

Log D can predict gastrointestinal absorption and lipophilic properties because it is a pH-dependent function. Calculated values demonstrated the high gastrointestinal track absorption (pH 3–7) and lipophilic properties. It was observed a good correlation between the calculated distribution coefficients at pH 7.4 and pH 5.2 (log D 5.2 and 7.4), and the found inhibition percentages for the tested compounds (40).

In addition, the anti-malarial activity increases with increasing the ACD log D value at pH 7.4 is most probably because this favors the entry of the compound into the RBC, parasite and ultimately the digestive vacuole. On the other hand, a low ACD log D at a low pH (3 – 5) disfavors the egress of the compound out of the digestive vacuole. The consequence would be a greater accumulation in the digestive vacuole and hence higher concentrations at the site of action. Indeed, anti-log {ACD log D (7.4) – ACD log D (5)} is an experimental measure of the vacuolar pH trapping. It would make sense that this would only apply in the case of hemozoin inhibitors because these are the

compounds that act in the digestive vacuole. Since researchers prefer to use {ACD log D (7.4) - ACD log D (5)} for each compound. This might weaken the correlation and undermine this argument. Hence, we built a new model with that variable, {ACD log D (7.4) - ACD log D (5)} to investigate its effect, however, we found it insignificant.

Noteworthy, we have overcome our previous limitation in which we could not find any heavy atoms playing an important role in the refractive index which investigate the characteristics of the material (22). The presence and even the quantity of some heavy atoms and functional groups with a high refractive index, as sulfur (41) have an important role in increasing the molar refraction. It is noteworthy that out of the 224 included compounds, there were 87 compounds have at least 1 S atom with a mean of 1 atom and 131 compounds have no S atoms. Moreover, among negative and positive compounds there were 83 compounds and 4 compounds have at least one atom, respectively. Additionally, among our positive anti-malarial compounds, we have four compounds with quinolines which have no S atoms, hence this suggests a negative association between the S atoms and the anti-malarial activity. However, our results did not show an association between the refractive index and high anti-malarial activities, suggesting the refractive index is related only to anti-hemozoin but not in the anti-malarial activity. Although our models have not been validated in external samples, it can be the first clue for understanding the mechanism of action of anti-malarials (22). Moreover, future studies need to validate our results with a larger number of compounds, especially the positive ones.

Conclusion

Using physicochemical properties of the compounds, we reported new prediction models that could predict anti-malarial activity for chemical compounds possessing anti-hemozoin Our findings revealed that lower number of S atoms, lower values of log D pH 3, 4 and 5, and higher value of ACD log D pH 7.4 were independent predictors of a higher anti-malarial activity. This information can be used for *in silico* screening and modifying chemical structures to develop anti-malarials.

Acknowledgments

This study was conducted (in part) at the Joint Usage / Research Center on Tropical Disease, Institute of Tropical Medicine, Nagasaki University, Japan. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Moreover, we are grateful for all Online Research Club native speakers who revised the English language of this manuscript, especially Vincent Hou from McMaster University in Canada.

Conflict of interest

None.

Author contributions

FM, AAT, MF did the experiments under the supervision of ShuM, OK, NTH, and KH. FM, MGK, TVD, DVT, SaM, YT, AMA, HTNG, TND, LKH, and MTE extracted the data from ChemSpider and SciFinder. MGK and NLV performed the statistical analysis under the supervision of NTH and KH. FM, MGK, TVD, AMA, OK, TJE, and NTH interpreted the analysis wrote the manuscript. All authors revised and approved the final version.

References

1. World Health Organization. 2016. World Malaria Report 2015. World Health Organization.
2. Parhizgar A, Tahghighi A. 2017. Introducing new antimalarial analogues of chloroquine and amodiaquine: A narrative review. Iran J Med Sci.
3. Eastman R, Fidock D. 2009. Artemisinin-based combination therapies: a vital tool in efforts to eliminate malaria. Nat Rev Microbiol.
4. Plowe C. 2009. The evolution of drug-resistant malaria. Trans R Soc Trop.
5. Dondorp A, Ringwald P. 2013. Artemisinin resistance is a clear and present danger.

Trends Parasitol.

6. Dondorp A, Yeung S, White L. 2010. Artemisinin resistance: current status and scenarios for containment. Nat Rev.

7. White NJ, Duong TT, Uthaisin C, Nosten F, Phyo AP, Hanboonkunupakarn B, Pukrittayakamee S, Jittamala P, Chuthasmit K, Cheung MS, Feng Y, Li R, Magnusson B, Sultan M, Wieser D, Xun X, Zhao R, Diagana TT, Pertel P, Leong FJ. 2016. Antimalarial Activity of KAF156 in Falciparum and Vivax Malaria. N Engl J Med 375:1152–1160.

8. White NJ, Pukrittayakamee S, Phyo AP, Rueangweerayut R, Nosten F, Jittamala P, Jeeyapant A, Jain JP, Lefèvre G, Li R, Magnusson B, Diagana TT, Leong FJ. 2014. Spiroindolone KAE609 for Falciparum and Vivax Malaria. N Engl J Med 371:403–410.

9. Phyo A, Jittamala P, Nosten F. 2016. Antimalarial activity of artefenomel (OZ439), a novel synthetic antimalarial endoperoxide, in patients with Plasmodium falciparum and Plasmodium vivax. Lancet Infect.

10. Held J, Supan C, Salazar C, Tinto H. 2015. Ferroquine and artesunate in African adults and children with Plasmodium falciparum malaria: a phase 2, multicentre, randomised, double-blind, dose-ranging, non-. Lancet Infect.

11. McCarthy J, Rückle T, Djeriou E. 2016. A Phase II pilot trial to evaluate safety and efficacy of ferroquine against early Plasmodium falciparum in an induced blood-stage malaria infection study. Malaria.

12. Mahmoudi S, Keshavarz H. 2017. Efficacy of Phase 3 Trial of RTS, S/AS01 Malaria Vaccine in infants: a systematic review and meta-analysis. Hum Vaccin Immunother 00–00.

13. Rts S, Agnandji S, Lell B. 2012. A phase 3 trial of RTS, S/AS01 malaria vaccine in African infants. New Engl.

14. Targett G, Moorthy V. 2013. Malaria vaccine research and development: the role of the WHO MALVAC committee. Malaria.
15. Adams P, Berman P, Egan T, Marsh P. 1996. The iron environment in heme and heme-antimalarial complexes of pharmacological interest. J Inorg.
16. Dinio T, Gorka A, McGinniss A, Roepe P. 2012. Investigating the activity of quinine analogues versus chloroquine resistant Plasmodium falciparum. Bioorganic Med.
17. Egan T, Hunter R, Kaschula C. 2000. Structure− function relationships in aminoquinolines: effect of amino and chloro groups on quinoline− hematin complex formation, inhibition of β-hematin formation,. J Med.
18. Ridley RG, Dorn A, Vippagunta SR, Vennerstrom JL. 1997. Haematin (haem) polymerization and its inhibition by quinoline antimalarials. Ann Trop Med Parasitol 91:559–566.
19. Sullivan D. 2002. Theories on malarial pigment formation and quinoline action. Int J Parasitol.
20. Sandlin R, Carter M, Lee P. 2011. Use of the NP-40 detergent-mediated assay in discovery of inhibitors of β-hematin crystallization. Antimicrob agents.
21. Wicht KJ, Combrinck JM, Smith PJ, Egan TJ. 2015. Bayesian models trained with HTS data for predicting β-haematin inhibition and in vitro antimalarial activity. Bioorg Med Chem 23:5210–5217.
22. Huy N, Chi P, Nagai J, Dang T. 2017. High-Throughput Screening and Prediction Model Building for Novel Hemozoin Inhibitors Using Physicochemical Properties. Antimicrob agents.
23. Ihekwereme C, Esimone C. 2014. Hemozoin inhibition and control of clinical malaria. Adv.
24. Hänscheid T, Egan T, Grobusch M. 2007. Haemozoin: from melatonin pigment to drug

target, diagnostic tool, and immune modulator. Lancet Infect Dis.

25. Nakatani K, Ishikawa H, Aono S, Mizutani Y. 2014. Identification of essential histidine residues involved in heme binding and Hemozoin formation in heme detoxification protein from Plasmodium falciparum. Sci Rep.

26. Jr DS, Gluzman I, Goldberg D. 1996. Plasmodium hemozoin formation mediated by histidine-rich proteins. Sci YORK THEN.

27. Jani D, Nagarkatti R, Beatty W, Angel R. 2008. HDP—a novel heme detoxification protein from the malaria parasite. PLoS.

28. Dorn A, Stoffel R, Matile H, Bubendorf A, Ridley R. 1995. Malarial haemozoin/beta-haematin supports haem polymerization in the absence of protein. Nature.

29. Tripathi A, Gupta A, Garg S, Tekwani B. 2001. In vitro β-hematin formation assays with plasma of mice infected with Plasmodium yoelii and other parasite preparations: Comparative inhibition with quinoline and. Life Sci.

30. Huy N, Uyen D, Maeda A, Oida T. 2007. Simple colorimetric inhibition assay of heme crystallization for high-throughput screening of antimalarial compounds. Antimicrob agents.

31. Sandlin RD, Fong KY, Wicht KJ, Carrell HM, Egan TJ, Wright DW. 2014. Identification of β-hematin inhibitors in a high-throughput screening effort reveals scaffolds with in vitro antimalarial activity. Int J Parasitol Drugs drug Resist 4:316–25.

32. Rosenthal P. 2003. Antimalarial drug discovery: old and new approaches. J Exp Biol.

33. Trager W, Jensen J. 1976. Human malaria parasites in continuous culture. Science (80-).

34. 2011. PEXEL-independent trafficking of Plasmodium falciparum SURFIN4.2 to the parasite-infected red blood cell and Maurer's clefts. Parasitol Int 60:313–320.

35. Sandlin RD, Carter MD, Lee PJ, Auschwitz JM, Leed SE, Johnson JD, Wright DW. 2011. Use of the NP-40 detergent-mediated assay in discovery of inhibitors of beta-

hematin crystallization. Antimicrob Agents Chemother 55:3363–9.

36. Smilkstein M, Sriwilaijaroen N, Kelly J. 2004. Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening. Antimicrob agents.

37. 2007. Antimalarial drugs inhibiting hemozoin (β-hematin) formation: A mechanistic update. Life Sci 80:813–828.

38. Sanjivanjit K. Bhal, Karim Kassam, Ian G. Peirson and, Pearl* GM. 2007. The Rule of Five Revisited: Applying Log D in Place of Log P in Drug-Likeness Filters.

39. Warhurst DC, Craig JC, Raheem KS. 2016. Influence of LAR and VAR on Para-Aminopyridine Antimalarials Targetting Haematin in Chloroquine-Resistance. PLoS One 11:e0160091.

40. Rojas Ruiz FA, García-Sánchez RN, Estupiñan SV, Gómez-Barrio A, Torres Amado DF, Pérez-Solórzano BM, Nogal-Ruiz JJ, Martínez-Fernández AR, Kouznetsov V V. 2011. Synthesis and antimalarial activity of new heterocyclic hybrids based on chloroquine and thiazolidinone scaffolds. Bioorg Med Chem 19:4562–4573.

41. Liu J, Nakamura Y, Shibasaki Y, Ando S. 2007. High refractive index polyimides derived from 2, 7-bis (4-aminophenylenesulfanyl) thianthrene and aromatic dianhydrides. Macromolecules.

Figure legends

Figure 1. The workflow of this study. *First, 224 compounds that showed anti-hemozoin activity in the previous study* (22) *were underwent in vitro anti-malarial assay at 10 μM. Second, 30 compounds having ≥ 45% parasite inhibitory efficacy have underwent in vitro dose-response assay between 0.5 nM and 10 μM. Then, 22 compounds having* $IC_{50} \leq 10$ *μM were identified. Subsequently, the main structures and physicochemical properties of these compounds searched by*

SciFinder. Finally, the prediction models were generated by the traditional approach versus the Bayesian approach using their physical and chemical properties.

Figure 2. The initial in vitro anti-malarial screening of 224 compounds using chloroquine-mefloquine sensitive P. falciparum 3D7A strain. *The circle dots represent % of parasite inhibition of 224 compounds including 30 compounds having parasite inhibition ≥ 45%.*

Figure 3. Models selected by Bayesian model averaging. *Blue and red colors represent positive and negative variable estimates, respectively while uncolored variables were not included in the model. On the x-axis, models were listed in the order of the decline in the posterior probability.*

469 ***Table 1. Anti-malarial s and anti-hemozoin activitie of 22 positive compounds.***

Serial no.	Compound ID	*P. falciparum (3D7A)*	Anti-hemozoin
		IC_{50} (µM)	**IC_{50} (µM)**
1	UTDIF3A11	0.66 ± 0.10	78.1
2	UTDIF4A14	8.15 ± 2.60	34.67
3	UTDIF5A15	9.00 ± 1.40	14.01
4	UTDIF6C12	8	4.58
5	UTDIF7D16	10.50 ± 2.10	87.76
6	UTDIF8E4	10	38.54
7	UTDIF9E10	1.54 ± 0.07	53.44
8	UTDIF10E14	10.00 ± 1.40	41.18
9	UTDIF11F12	9	36.16
10	UTDIF12F15	8.95 ± 1.30	103.5
11	UTDIF13G5	9.26 ± 1.80	30.69
12	UTDIF14I6	7.00 ± 1.40	43.98
13	UTDIF15I15	1.01 ± 0.50	25.96
14	UTDIF16J16	9.28 ± 2.40	28.5
15	UTDIF17K7	4.80 ± 1.70	198.1
16	UTDIF18L5	3.06 ± 1.30	110
17	UTDIF19L16	6.80 ± 4.40	29.04
18	UTDIF20M7	8.00 ± 2.80	156
19	UTDIF21O9	0.56 ± 0.27	18.3
20	UTDIF22P6	0.06	42.98
21	UTDIF23P14	9.50 ± 0.70	24.72
22	UTDIF24G15	9	160

470

471

472 ***Table 2. Uni-variable logistic regression (LR) statistics of the physicochemical properties of the compounds.***

Predictors	**Uni-variable analysis**	
	OR (95% CI)	***P***
Number of C atoms	**1.11 (0.98 – 1.25)**	**0.092**
Number of H atoms*	1.07 (0.98 - 1.17)	0.111
Number of N atoms	**1.5 (1.15 - 1.95)**	**0.003**
Number of O atoms*	1.02 (0.77 - 1.35)	0.912
Number of S atoms	**0.45 (0.19 - 1.06)**	**0.067**
Number of F atoms	0 (0 - Inf)	0.992
Number of Cl atoms	1.45 (0.67 - 3.16)	0.348
Number of Br atoms	0 (0 - Inf)	0.991
Average mass atoms	1.01 (0.997 - 1.01)	0.214
Monoisotopic	1.01 (0.998 - 1.01)	0.192
Density*	0.89 (0.07 - 11.81)	0.929
Rotatable bonds*	1.03 (0.85 - 1.26)	0.753
H acceptors	**1.38 (1.08 - 1.77)**	**0.01**
H Donors*	1.11 (0.76 - 1.61)	0.605
H donor acceptor sum	**1.25 (1.03 - 1.52)**	**0.022**
KOC pH 1	1 (0.9999 - 1.0001)	0.647
KOC pH 2	1 (0.9999 - 1.0001)	0.568
KOC pH 3	1 (0.9998 - 1.0001)	0.499

KOC pH 4	1 (0.9998 - 1.0001)	0.457
KOC pH 5	1 (0.9998 - 1.0001)	0.469
KOC pH 6	1 (0.9999 - 1.0001)	0.562
KOC pH 7	1 (0.9999 - 1.0001)	0.538
KOC pH 8	1 (0.9999 - 1.0001)	0.65
KOC pH 9	1 (0.9999 - 1.0001)	0.603
KOC pH 10	1 (0.9998 - 1.0001)	0.526
LogD pH 1*	**0.73 (0.58 - 0.92)**	**0.007**
LogD pH 2	**0.74 (0.59 - 0.93)**	**0.009**
LogD pH 3	**0.73 (0.58 - 0.92)**	**0.008**
LogD pH 4	**0.72 (0.56 - 0.92)**	**0.008**
LogD pH 5	**0.76 (0.59 - 0.97)**	**0.027**
LogD pH 6	0.85 (0.66 - 1.09)	0.199
LogD pH 7	0.94 (0.74 - 1.2)	0.643
LogD pH 8	0.9995 (0.79 - 1.27)	0.997
LogD pH 9	1.04 (0.82 - 1.32)	0.756
LogD pH 10	1.09 (0.86 - 1.38)	0.481
Log P*	0.85 (0.62 - 1.17)	0.319
Mass intrinsic	0.97 (0.79 - 1.19)	0.762
Mass solubility pH 1	**1.002 (0.9999 - 1.003)**	**0.067**
Mass solubility pH 2	1.001 (0.9997 - 1.003)	0.106

Mass solubility pH 3	**1.002 (0.9998 - 1.003)**	**0.078**
Mass solubility pH 4	1.001 (0.9991 - 1.004)	0.24
Mass solubility pH 5	1.002 (0.9993 - 1.005)	0.136
Mass solubility pH 6	1.001 (0.998 - 1.004)	0.535
Mass solubility pH 7	0.997 (0.99 - 1.01)	0.616
Mass solubility pH 8	0.97 (0.89 - 1.06)	0.525
Mass solubility pH 9	0.84 (0.57 - 1.24)	0.376
Mass solubility pH 10	0.97 (0.91 - 1.03)	0.312
Mass solubility	0.47 (0.076 - 2.86)	0.409
Molar volume	1.01 (0.997 - 1.01)	0.218
Molecular weight	1 (0.997 - 1.01)	0.283
pKa1	1.05 (0.93 - 1.17)	0.444
Polar surface	1 (0.99 - 1.02)	0.733
Total score	1.22 (0.82 - 1.82)	0.334
Polarizability	**1.07 (0.996 - 1.14)**	**0.066**
Surface tension	1.02 (0.99 - 1.05)	0.256
Refractive index[ft]	1.36 (0.72 - 2.57)	0.348
Rule number of five violations	2 (0.67 - 5.9)	0.212
ACD logD pH 5.5*	0.87 (0.68 - 1.11)	0.256
ACD logD pH 7.4*	1.07 (0.83 - 1.36)	0.617

ACD BCF pH 5.5	1 (1 - 1)	0.771
ACD BCF pH 7.4	1 (0.9998 - 1.0001)	0.737
ACD KOC pH 5.5*	1 (1 - 1)	0.77
ACD KOC pH 7.4*	1 (0.9999 - 1.0001)	0.766

Abbreviations: OR = odds ratio; CI = confidence interval;

*Variables with prior significance in our previous study (22). Significant predictors are in bold and underscored.

[¶]Refractive index was multiplied by 10 to get a smaller OR.

477 ***Table 3. The best five parsimonious models selected by BMA.***

Model No.	Predictor	Coefficient of predictor	BIC of model	PP of model
1	Intercept	-2.21	-648.24	0.17
	Number of S atoms	-1.26		
	Log D pH 4	-0.86		
	ACD Log D pH 7.4	0.78		
2	Intercept	-2.76	-647.46	0.11
	Log D pH 4	-0.84		
	ACD Log D pH 7.4	0.81		
3	Intercept	-2.22	-646.94	0.09
	Number of S atoms	-1.3		
	Log D pH 3	-0.78		
	ACD Log D pH 7.4	0.68		
4	Intercept	-2.05	-646.02	0.06
	Number of S atoms	-1.22		
	Log D pH 5	-0.92		
	ACD Log D pH 7.4	0.85		
5	Intercept	-2.57	-645.61	0.05
	Log D pH 5	-0.92		
	ACD Log D pH 7.4	0.88		

478 Abbreviations: No. = number; BIC = Bayesian information criterion; PP = posterior probability. The cumulative PP = 0.468.

Figure 1

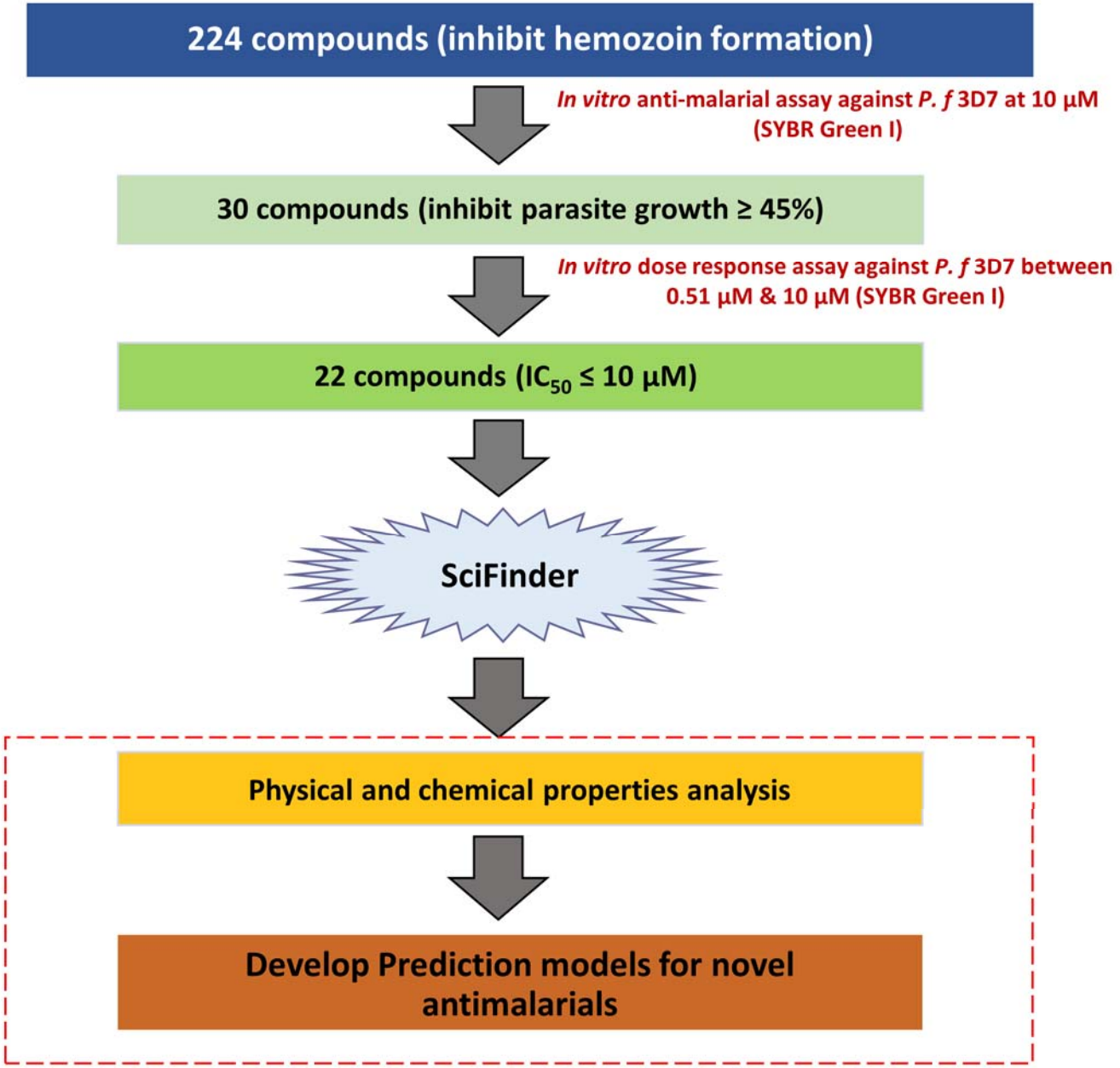

224 compounds (inhibit hemozoin formation)
In vitro anti-malarial assay against *P. f* 3D7 at 10 µM (SYBR Green I)
30 compounds (inhibit parasite growth ≥ 45%)
In vitro dose response assay against *P. f* 3D7 between 0.51 µM & 10 µM (SYBR Green I)
22 compounds (IC_{50} ≤ 10 µM)
SciFinder
Physical and chemical properties analysis
Develop Prediction models for novel antimalarials

Figure 2

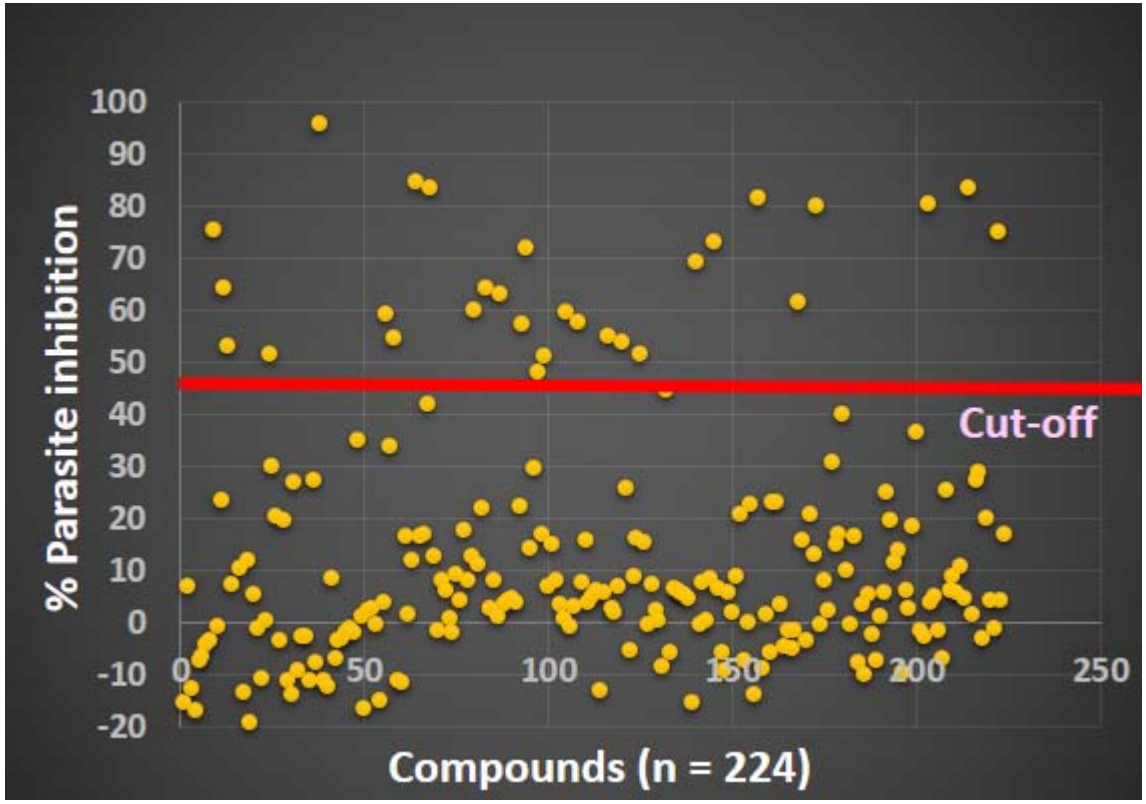
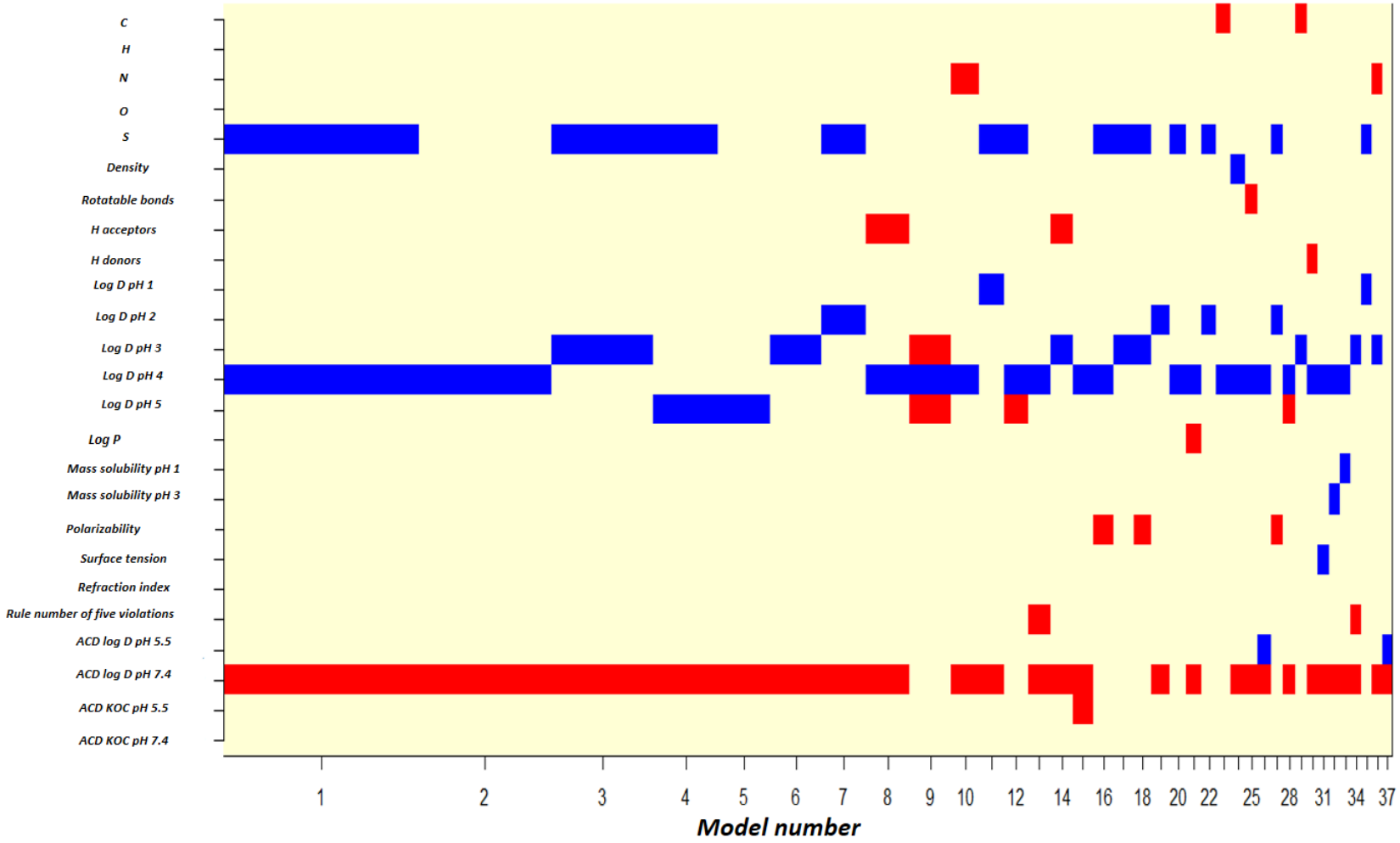

100
90
80
70
60
50
40
30
20
10
0
-10
-20
% Parasite inhibition
0
50
100
150
200
250
Compounds (n = 224)
Cut-off

Figure 3


C
H
N
O
S
Density
Rotatable bonds
H acceptors
H donors
Log D pH 1
Log D pH 2
Log D pH 3
Log D pH 4
Log D pH 5
Log P
Mass solubility pH 1
Mass solubility pH 3
Polarizability
Surface tension
Refraction index
Rule number of five violations
ACD log D pH 5.5
ACD log D pH 7.4
ACD KOC pH 5.5
ACD KOC pH 7.4
1 2 3 4 5 6 7 8 9 10 12 14 16 18 20 22 25 28 31 34 37
Model number

Supplementary Data

Method: *Plasmodium falciparum* sensitive strain, *3D7A* (3% parasitemia, 2% haematocrit) were seeded in 96-well black clear bottom plate. Subsequently, Chloroquine (10 µM), Artemisinin (10 µM), 22 compounds (10 µM) and DMSO (0.5%) were added onto the parasites (3D7A) in 96-well black clear bottom plate. Afterwards, lysis buffer containing SYBR Green – I was added, incubate for 1 hour in dark and measure the fluorescence using Perkin-Elmer plate reader.

Results: The % of fluorescence obtained from compounds & controls (positive – Artemisinin & Chloroquine, negative – 0.5% DMSO) were measured to check the interference of compounds with SYBR Green – I fluorescence. The negative control (0.5% DMSO) exhibited 100% fluorescence. Other 22 compounds and positive controls showed >90% fluorescence which indicates that there was almost no interference of the compounds with SYBR Green – I fluorescence.

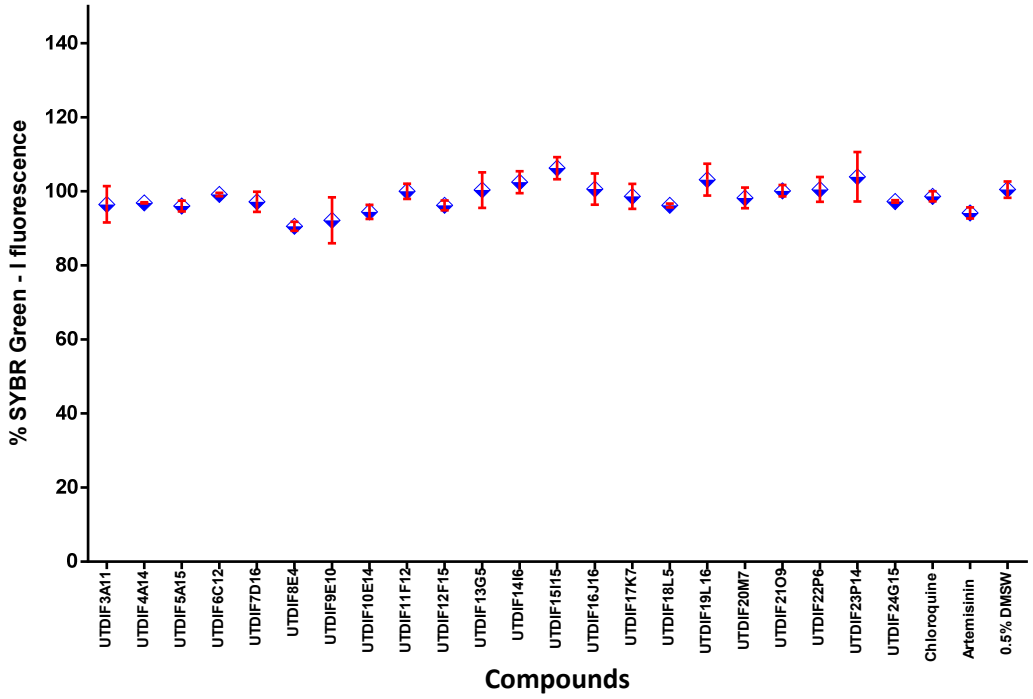


Fig A: The % of fluorescence of 22 compounds, Artemisinin& Chloroquine (positive control) and 0.5% DMSO (negative control) have shown. The % of fluorescence of all compounds are compared with the % of fluorescence of 0.5% DMSO (negative control.