Makara Journal of Science

Volume 28
Issue 3 *September*

Article 10

9-25-2024

Modeling of Benzimidazole Derivatives as Antimalarial Agents using QSAR Analysis

Ruslin Hadanu
Department of Chemistry Education, Faculty of Teacher Training and Education Science, Universitas Sembilanbelas November Kolaka, Sulawesi Tenggara 93517, Indonesia, ruslinhadanu@gmail.com

Marham Sitorus
Department of Chemistry, Faculty of Mathematics and Natural Sciences, Universitas Negeri Medan, Medan 20237, Indonesia

Makara Journal of Science, 28/3 (2024), 264-275
doi: 10.7454/mss.v28i3.2238

Modeling of Benzimidazole Derivatives as Antimalarial Agents using QSAR Analysis

Ruslin Hadanu[1*] and Marham Sitorus[2]

1. Department of Chemistry Education, Faculty of Teacher Training and Education Science, Universitas Sembilanbelas November Kolaka, Sulawesi Tenggara 93517, Indonesia.
2. Department of Chemistry, Faculty of Mathematics and Natural Sciences, Universitas Negeri Medan, Medan 20237, Indonesia.

[*]*E-mail: ruslinhadanu@gmail.com*

Received September 19, 2023 | Accepted August 12, 2024

Abstract

In this study, quantitative structure–activity relationship (QSAR) analysis was conducted on 20 homologous compounds of benzimidazole derivatives. The structures of the benzimidazole derivatives were optimized using the semiempirical Parameterized Model 3 method of HyperChem for Windows 8.0 to obtain 14 descriptors. Then, multiple linear regression (MLR) analysis was performed using the backward method. The results of the MLR analysis obtained four new QSAR equation models. Based on statistical criteria, model 1 was determined as the best QSAR equation model in predicting the theoretical IC_{50} values of the new benzimidazole derivatives. As many as 20 new compounds of benzimidazole derivatives were modeled, of which 13 new compounds (23, 24, 25, 26, 27, 28, 29, 30, 31, 37, 38, 39, and 40 compounds) were recommended for synthesis in the laboratory because these compounds of benzimidazole derivatives havethey theoretically had higher antimalarial activity than chloroquine.

Keywords: antimalarial, benzimidazole, HyperChem, modeling, QSAR

Introduction

Malaria is a major health problem in the world, and this can be seen in various new problems associated with it, including the resistance of *Plasmodium falciparum* to antimalarial drugs. In addition to resistance to the antimalarial drug chloroquine, there has also been resistance to artemisinine, thus causing a global threat in efforts to reduce morbidity and mortality and eradicate malaria [1]. Malaria is a life-threatening disease caused by a parasite transmitted to humans through the bite of an infected female *Anopheles* mosquito. In 2021, there were an estimated 247 million cases of malaria worldwide. The estimated number of deaths from malaria reached 619,000 in 2021. The World Health Organization Africa Region bears a very high global burden of malaria. In 2021, the region was home to 95% of malaria cases and 96% of malaria deaths. Children aged under 5 years account for about 80% of all malaria deaths in the region [2−4]. Malaria persists as a major problem in various countries in the world, such as in the Ebonyi State in Nigeria, which is still fighting a series of epidemics that have not subsided [2]. Moreover, four countries in Africa account for more than half of all deaths from malaria worldwide: Nigeria (31.3%), The Democratic Republic of the Congo (12.6%), The United Republic of Tanzania (4.1%), and Nigeria (3.9%) [4].

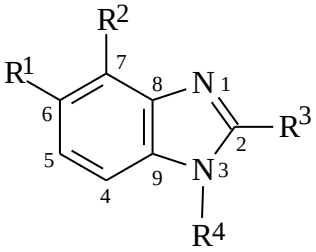
On the basis of the foregoing, the discovery and development of new, more potent antimalarial drugs is crucial for overcoming malaria parasite resistance to previously known antimalarial drugs. Researchers around the world are developing new antimalarial drugs, such as benzimidazole derivatives [5−7], imidazopyridazine compounds [8], artemisinin derivatives [9], pyrido[1,2-a] benzimidazole compounds [10], and natural product compounds [11], and pyrimido-[1]-2-a]benzimidazole derivatives [12]. In 2020, Mueller *et al.* [5] and his team synthesized several benzimidazole compounds by substituting various groups to obtain several benzimidazole derivatives with high potential activity as antimalarial drugs, but these did not exceed or equal the activity of the antimalarial drugs chloroquine and halofantrine or other drugs [5]. Accordingly, to obtain antimalarial drugs that are more potent or more active, new research on modeling benzimidazole derivatives must be conducted to obtain the most potential theoretical IC_{50} through the study of the quantitative relationship between the structures of benzimidazole derivatives and the IC_{50} obtained experimentally in pharmacology and parasitology laboratories.

Recently, other efforts in the development of new antimalarial drugs have often begun with the analysis of the quantitative structure–activity relationship (QSAR) of benzimidazole derivatives [13, 14], which are then synthesized in the laboratory, followed by the testing of antimalarial activity. However, further studies on the QSAR of benzimidazole derivatives [5] must be conducted using computational methods and QSAR analysis, which aims to obtain a QSAR equation model for designing new benzimidazole derivatives with higher antimalarial activity or equivalent to those of antimalarial drugs that have been on the market. Herein, QSAR analysis was used to look for a quantitative relationship between the empirical aspects of a compound and its theoretical biological activity. A mathematical relationship between activity descriptors was obtained, which could be used to predict new compounds with higher activity than those of previous compounds. To obtain the best QSAR equation model, the multiple linear regression (MLR) statistical method was used.

Experimental

Material and methods. The research materials for 20 benzimidazole derivative compounds, which are the results of the research of Mueller *et al*. [5], are presented in Table 1. The 20 benzimidazole derivatives in Table 1 were grouped into two: 15 training data compounds (compounds 2, 3, 4, 6, 7, 9, 11, 12, 13, 14, 15, 16, 18, 19, and 20) and five testing data compounds (compounds 1, 5, 8, 10, and 17).

Table 1. Chemical Structure and Activity Data of the Antimalarial Compounds of Benzimidazole Derivatives Obtained from Mueller *et al*. (2020) [5]



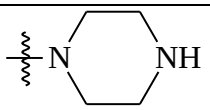
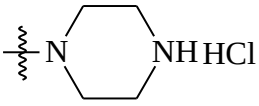
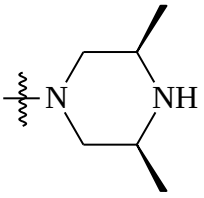
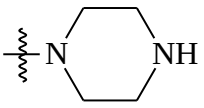
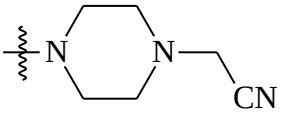
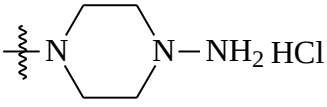
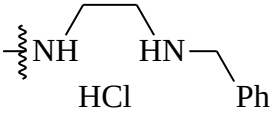
No.	R^1	R^2	R^3	R^4	IC_{50} (μM) Exp.
1	–H	–H		$-CH_2Ph$	0.810
2	–Cl	–H		$-CH_2Ph$	0.062
3	–Cl	–H		$-CH_2Ph$	0.216
4	–Cl	–H		$-CH_2Ph$	0.038
5	–Cl	–H		$-CH_2Ph$	0.306
6	–Cl	–H		$-CH_2Ph$	0.022
7	–Cl	–H		$-CH_2Ph$	0.086

Table 1. Continued

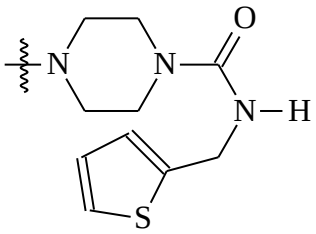
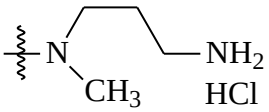
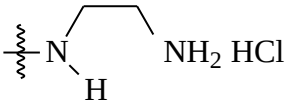
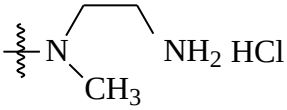
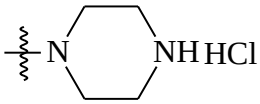
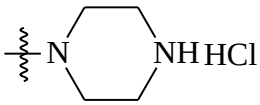
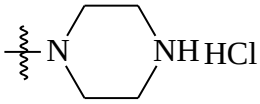
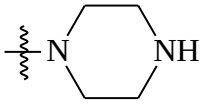
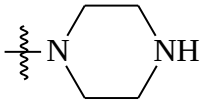
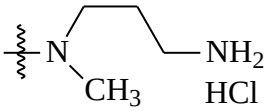
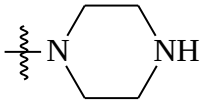
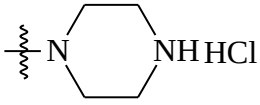
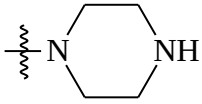
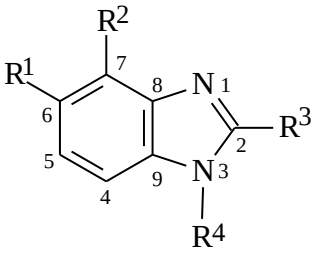
No.	R^1	R^2	R^3	R^4	IC_{50} (μM) Exp.
8	–Cl	–H		$–CH_2Ph–4–CH_3$	0.600
9	–Cl	–H		$–CH_2Ph$	0.348
10	–Cl	–H		$–CH_2Ph$	0.245
11	–Cl	–H		$–CH_2Ph$	0.233
12	$–CH_3$	–H		$–CH_2Ph$	0.098
13	–Br	–H		$–CH_2Ph$	0.042
14	$–CF_3$	–H		$–CH_2Ph$	0.310
15	–Br	–H		$–CH_2Ph$	0.028
16	$–CF_3$	–H		$–CH_2Ph$	0.138
17	–Br	–H		$–CH_2Ph$	0.382
18	–Cl	–H		–Ph	2.083`
19	–Cl	–Cl		$–CH_2Ph$	0.055
20	–Cl	–Br		$–CH_2Ph$	0.071

Table 2. Structures of Benzimidazole Derivatives from Modeling Results Based on Model 1 of the Best QSAR Equation



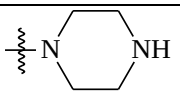
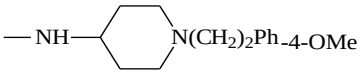
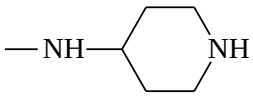
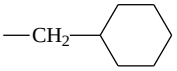
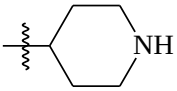
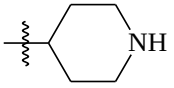
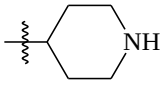
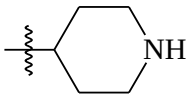
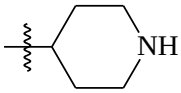
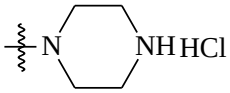
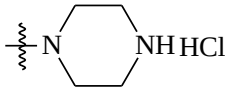
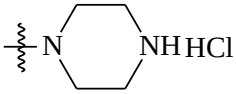
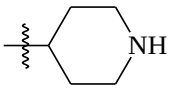
No.	R^1	R^2	R^3	R^4	IC_{50} (μM) Calculated
21	–CH_3	–H		–CH_2Ph	3.6×10^{-02}
22	–H	–H		–CH_2Ph-4-F	1.6×10^{-01}
23	–H	–H		–CH_2Ph-4-F	2.3×10^{-08}
24	–Cl	–H	–Cl	–CH_2Ph	2.3×10^{-16}
25	–Cl	–H	–Cl		1.4×10^{-16}
26	–Cl	–H	–Cl	–$(CH_2)_2$Ph	2.6×10^{-16}
27	–Cl	–Cl	–Cl	–Ph	6.0×10^{-08}
28	–Cl	–Cl	–Cl	–CH_2Ph	7.0×10^{-16}
29	–Cl	–Br	–Br	–CH_2Ph	1.0×10^{-20}
30	–Cl	–H		–CH_2Ph	9.0×10^{-04}
31	–Br	–H		–CH_2Ph	9.6×10^{-04}
32	–CF_3	–H		–CH_2Ph	8.4×10^{-03}
33	–CO_2Me	–H		–CH_2Ph	4.7×10^{00}
34	–SO_2Me	–H		–CH_2Ph	1.6×10^{06}
35	–Cl	–Ph		–CH_2Ph	6.8×10^{01}
36	–Ph	–H		–CH_2Ph	9.7×10^{-02}
37	–Ph–4–CF_3	–H		–CH_2Ph	5.7×10^{-07}
38	–Br	–Br	–Br	–CH_2Ph	1.1×10^{-19}
39	–Br	–Cl	–Br	–CH_2Ph	3.9×10^{-19}
40	–Cl	–Cl		–CH_2Ph	9.6×10^{-04}

Table 3. Dependent and Independent Variables used for QSAR Analysis of Benzimidazole Derivatives as Antimalarial Compounds

Comp.	Variables														log 1/IC_{50} Experiment
	qN1	qC2	qN3	qC4	qC5	qC6	qC7	qC8	qC9	μ	E_{LUMO}	E_{HOMO}	α (Å3)	log P	
Fitting Compounds															
2	−0.1493	−0.1277	0.2340	−0.0758	−0.0970	−0.1395	−0.0431	−0.0573	−0.1805	3.287	−0.3856	−8.8987	38.21	0.67	1.2811
3	−0.1411	−0.1195	0.225	−0.0782	−0.0982	−0.1404	−0.0411	−0.0566	−0.1798	4.608	−0.2418	−8.7407	40.14	1.46	0.6482
4	−0.1404	−0.1186	0.2250	−0.0783	−0.0982	−0.1405	−0.0410	−0.0566	−0.1796	4.708	−0.2441	−8.7527	36.47	0.64	1.3524
6	−0.1123	−0.1331	0.2355	−0.0825	−0.0937	−0.1437	−0.0342	−0.0643	−0.1795	6.225	−0.3036	−8.7960	39.56	0.29	1.6406
7	−0.1168	−0.1941	0.2633	−0.0741	−0.0924	−0.1400	−0.0386	−0.0661	−0.1765	5.078	−0.4877	−8.9380	46.81	1.50	1.0429
9	−0.1535	−0.1100	0.2153	−0.0773	−0.1018	−0.1379	−0.0460	−0.0498	−0.1816	1.733	−0.2728	−8.7999	38.98	0.65	0.4635
11	−0.1707	−0.1101	0.2217	−0.0748	−0.1028	−0.1361	−0.0496	−0.0467	−0.1829	5.168	−0.3258	−8.8351	37.15	0.60	0.5936
12	−0.1476	−0.1332	0.2299	−0.0871	−0.0928	−0.0907	−0.0413	−0.0623	−0.1830	4.784	−0.1371	−8.7675	38.12	1.04	1.0161
13	−0.1466	−0.1276	0.2337	−0.0816	−0.0788	−0.1249	−0.0196	−0.0636	−0.1759	3.169	−0.4385	−9.0621	38.91	0.94	1.2885
14	−0.1357	−0.1251	0.2312	−0.0919	−0.0477	−0.2024	0.0313	−0.0813	−0.1522	7.932	−0.5801	−9.1631	37.84	1.46	0.5487
15	−0.1374	−0.1191	0.2253	−0.0843	−0.0798	−0.1256	−0.0174	−0.0633	−0.1749	4.906	−0.2969	−8.8386	37.17	0.91	1.5933
16	−0.1313	−0.1146	0.2232	−0.0955	−0.0471	−0.2040	0.0254	−0.0784	−0.1506	6.871	−0.5060	−9.0145	36.11	1.43	0.8108
18	−0.1631	−0.1025	0.2522	−0.0639	−0.1080	−0.1311	−0.0533	−0.0391	−0.1842	4.760	−0.4694	−8.5808	34.64	0.09	−0.3294
19	−0.1388	−0.1276	0.2368	−0.0792	−0.0857	−0.1542	−0.0854	−0.0615	−0.1701	3.568	−0.5412	−8.7626	40.14	0.44	1.2459
20	−0.1310	−0.1162	0.2286	−0.0714	−0.0953	−0.1367	−0.0783	−0.0319	−0.1807	3.999	−0.3987	−8.8087	39.10	0.69	1.1514
Testing compounds															
1	−0.1232	0.2220	−0.0931	−0.1224	−0.0943	−0.1224	−0.0403	−0.0665	−0.1763	3.783	−0.0938	−8.6709	37.54	0.86	0.0309
5	−0.1360	0.2391	−0.0812	−0.1436	−0.0935	−0.1436	−0.0349	−0.0659	−0.1794	5.543	−0.3307	−8.8171	44.16	1.13	0.5195
8	−0.1245	0.2277	−0.0777	−0.1405	−0.0977	−0.1405	−0.0409	−0.0573	−0.1797	6.358	−0.6441	−8.7863	50.60	0.06	0.1152
10	−0.1307	0.2027	−0.0781	−0.1376	−0.0998	−0.1376	−0.0440	−0.0476	−0.1750	8.165	−0.3899	−8.8482	45.31	0.24	0.8383
17	−0.1108	0.2164	−0.0834	−0.1234	−0.0830	−0.1234	−0.0223	−0.0565	−0.1770	1.765	−0.3252	−8.8666	39.68	0.92	0.6642

Instrumentation. QSAR analysis was done using computer hardware (Dekstop-3I0S4MP; operating system, Windows Home Single Language 64-bit; system manufacturing, AsusTeK Computer Inc; processor, Intel(R) Core(TM) i7-7500 CPU@ 2.70GHz (4CPUs), ~2,9 GHz; memory, 8192 MB RAM; 500 GB HDD) and the HyperChem 8.0 software for Windows. HyperChem 8.0 for Windows was used to describe three-dimensional (3D) structures and optimize the geometry of the training data compounds (2, 3, 4, 6, 7, 9, 11, 12, 13, 14, 15, 16, 18, 19, and 20), testing data compounds (1, 5, 8, 10, and 17), and the structural modeling of benzimidazole derivative compounds (21–40), as presented in Table 2. MLR analysis was done using the SPSS 26.0 software for Windows to obtain the QSAR equation model.

General Procedure

Calculation of descriptors. The structures of the benzimidazole derivatives in Table 1 and the retrosynthesis design compounds in Table 2 were drawn in two dimensions (2D) and then converted into 3D using the HyperChem 8.0 software for Windows. The 3D structures of the benzimidazole derivatives were optimized using the "compute-geometry optimization" menu to obtain a more stable conformational (minimum energy) structure of the benzimidazole derivatives. These structures were optimized using the semiempirical Parameterized Model 3 (PM3) method and the Polak–Ribiere algorithm or the peer gradient method, with a root mean square value of 0.001 kcal/(Å·mol) in the ground state. The results of the optimization were recorded to obtain descriptor data using the "start log" menu. The variables were in the form of the total atomic charge (qN1, qC2, qN3, qC4, qC5, qC6, qC7, qC8, and qC9), dipole moment (μ), polarizability (α), and logP. Meanwhile, the E_{LUMO} and E_{HOMO} variables were obtained from the "compute," "vibration," and "orbital" menus. The data of all independent variables (qN1, qC2, qN3, qC4, qC5, qC7, qC8, qC9, μ, α, logP, E_{LUMO}, and E_{HOMO}) and dependent variables (log $1/IC_{50}$ experiments) are shown in Table 2, whereas the constants and coefficients of the independent variables from four QSAR equation models are presented in Table 4. One of the reasons and considerations in choosing the type of variable was adjusted to the research of Rajkhowa *et al.* [15], who conducted a QSAR analysis of artemisinin compounds as antimalarial drugs.

QSAR analysis using the MLR method. The log $1/IC_{50}$ values of the experimental results (Table 3) were used as dependent variables, whereas the net atomic charge (qN1, qC2, qN3, qC4, qC5, qC6, qC7, qC8, and qC9), μ, E_{LUMO}, E_{HOMO}, α, and logP were used as independent variables. The dependent and independent variables (Table 3) were analyzed using the MLR method to determine which descriptors affected the IC_{50} values of the benzimidazole derivatives. The results of the MLR analysis yielded four QSAR equation models. The statistical parameter values for these models (r, r^2, SE, F_{count}/F_{table}, and PRESS) are presented in Table 5. In the MLR analysis, the constant and coefficient values of each independent variable in the four QSAR equation models were used to determine the theoretical biological activity (IC_{50}) value of the benzimidazole derivatives and the best QSAR equation model.

Modeling new compounds of benzimidazole derivatives as antimalarials. The design or modeling phase of benzimidazole derivatives through QSAR analysis aims to find new compounds from benzimidazole derivatives as antimalarial drugs, where these compounds must have higher theoretical activity (low IC_{50} value) than the experimental IC_{50} values of the benzimidazoles derivatives presented in Table 1. The modeling method or design of new benzimidazole compounds was conducted by changing the atomic position and type of functional group in the main framework structure of the benzimidazole compound. The focus on changing the type and location of atoms or functional groups was conducted on the active central region based on disconnection and retrosynthesis studies while considering the feasibility of the synthesis pathway in an organic chemistry laboratory. It was expected that the placement of atoms or functional groups on the main framework of the benzimidazole derivatives would affect the antimalarial activity values of the new benzimidazole derivatives.

Results and Discussion

QSAR analysis. QSAR analysis is a computational method used to improve the efficiency of new drug development. Rational drug design, also known as in silico drug design, is conducted using computational methods such as QSAR analysis and information systems. One example of recent QSAR analysis research is the study of the relationship between the structure of benzamide derivatives and antituberculosis activity [16]. QSAR analysis is a valuable tool for researchers designing or developing drugs sustainably.

Many benzimidazole derivatives have been synthesized and tested for antimalarial activity (IC_{50}) by previous researchers, such as Mueller *et al*. [5]. To obtain a new benzimidazole derivative with high theoretical antimalarial activity, QSAR analysis is necessary. This research aims to develop a QSAR mathematical equation model to predict the antimalarial activity of benzimidazole derivatives. To design or model new benzimidazole derivative compounds using QSAR analysis, the following steps were taken: (1) examining the entire chemical structure of a series of benzimidazole derivative compounds with known antimalarial activity (IC_{50}) values [5] to determine the homologous basic framework; (2) optimizing the structure geometry; (3) determining the descriptor (independent variable); (4) descriptor calculation; (5) conduct-

ing bivariate analysis; (6) determining the QSAR equation models through MLR analysis; (7) selecting the best QSAR equation model; and (8) disconnection and retrosynthesis studies for modeling new compounds of benzimidazole derivatives.

One requirement for benzimidazole derivatives to be used as materials in this study is that the benzimidazole derivative series must have a basic structure that is homologous and must have an IC_{50} value from laboratory experiments. This is also a requirement to be met by fitting compounds, testing compounds, and modeling compounds through disconnection studies and retrosynthetic studies. The structures of the derivatives of the benzimidazole compounds as fitting compounds, testing compounds, and modeling compounds were optimized using the same method (PM3) to obtain the most stable structural form with the minimum or lowest energy. When the structural form of a benzimidazole compound was in the most stable state, variables (descriptors) were obtained for use in the QSAR analysis process using the multilinear regression method.

Before conducting a QSAR analysis, a bivariate analysis was first performed to determine the level of significant correlation between variables. Based on the results of the bivariate correlation analysis conducted using the SPSS 26.0 for Windows software, significant correlations were obtained between descriptors, which could be proven by the Pearson correlation r_{count} between descriptors: qN1–qC2 (−0.604), qN1–α (0.560), qN1–log1/IC_{50} (0.573), qC2–qN3 (−0.672), qC2–α (−0.888), qC4–qC5 (−0.840), qC4–qC7 (−0.840), qC4–qC7 (−0.761), qC4–qC8 (0.854), qC4–qC9 (−0.757), qC4–μ (−0.530), qC4–E_{HOMO} (0.645), qC4–log P (−0.572), qC5–qC6 (−0.753), qC5–qC7 (0.820), qC5–qC8 (−0.798), qC5–qC9 (0.963), qC5–μ (0.625), qC5–E_{LUMO} (−0.541), qC5–E_{HOMO} (−0.795), qC5–log P (0.596), qC6–qC7 (−0.609), qC6–qC8 (0.536), qC6–qC9 (−0.882), qC6–μ (−0.585), qC6–E_{LUMO} (0.688), qC6–E_{HOMO} (0.564), qC7–qC8 (−0.780), qC7–qC9 (0.757), qC7–μ (0.666), qC7–E_{HOMO} (−0.729), qC7–log P (0.609), qC8–qC9 (−0.763), qC8–μ (−0.566), qC8–E_{HOMO} (0.696), qC8–logP (−0.616), qC9–μ (0.635), qC9–E_{LUMO} (−0.640), qC9–E_{HOMO} (−0.715), qC9–logP (0.553), and E_{HOMO}–logP (−0.638). The bivariate correlation value is an absolute value, where a value of one indicates a very strong and positive correlation, whereas a value of −1 indicates a negative and very strong correlation. According to the results of the bivariate correlation analysis above, several independent variables and dependent variables have significant relationships, thus fulfilling the requirements for MLR analysis of the relationship between the serial structure of the benzimidazole compounds and biological activity (log 1/IC_{50}).

Based on the MLR analysis of the independent variables or descriptors (qN1, qC2, qN3, qC4, qC5, qC6, qC7, qC8, and qC9, μ, E_{LUMO}, E_{HOMO}, α, and logP) and the dependent variable (log 1/IC_{50}), four QSAR equation models were obtained, whose independent variable coefficients are presented in Table 4, whereas the statistical parameter criteria and PRESS validation test results are presented in Table 5. Based on the statistical parameter criteria and PRESS validation test results, model 1 of the QSAR equations was chosen as the best QSAR equation model as follows:

$$\text{Log } 1/IC_{50} = 26.892 + 49.892(qN1) - 55.394(qC2) - 128.558(qN3) - 20.978(qC4) + 10.162(qC6) - 6.361(qC7) - 55.991(qC8) - 207.272(qC9) + 0.076(\mu) - 14.914(E_{LUMO}) + 2.866(E_{HOMO}) - 0.460(\alpha) + 1.269(\text{log P}).$$

Table 4. Constants and Coefficients of Selected Independent Variables for Four QSAR Models

QSAR models	Coefficient of independent variables						
	N1	C2	N3	C4	C6	C7	C8
1	49.892	−55.394	−128.558	−20.978	10.162	−6.361	−55.991
2	49.343	−52.306	−124.637		12.350	−7.322	−58.882
3	50.775	−60.516	−146.789		16.583		−56.998
4	49.072	−57.328	−127.127		13.119		−49.738

QSAR Models	Coefficient of independent variables						*Constanta*
	C9	μ	E_{LUMO}	E_{HOMO}	$^{\alpha}$	log P	
1	−207.272	0.076	−14.914	2.866	−0.460	1.269	26.892
2	−181.020	0.080	−13.563	2.360	−0.439	1.145	28.006
3	−195.083	0.091	−15.964	3.812	−0.439	1.145	42.720
4	−167.770		−13.531	3.152	−0.426	0.982	38.177

Table 5. Four Selected Models and Their Statistical Parameters

QSAR Models	Variables	r	r^2	Adjusted R^2	SE	$F_{calc}/F_{table\ (0.05)}$	PRESS
1	qN1, qC2, qN3, qC4, qC6, qC7, qC8, qC9, μ, E_{LUMO}, E_{HOMO}, α, and log P	0.997	0.993	0.906	0.157	2.440	0.0045
2	qN1, qC2, qN3, qC6, qC7, qC8, qC9, μ, E_{LUMO}, E_{HOMO}, α, and log P	0.996	0.993	0.948	0.118	0.120	0.0067
3	qN1, qC2, qN3, qC6, qC8, qC9, μ, E_{LUMO}, E_{HOMO}, α, and log P	0.991	0.981	0.912	0.153	3.049	0.0347
4	qN1, qC2, qN3, qC6, qC8, qC9, E_{LUMO}, E_{HOMO}, α, and log P	0.981	0.963	0.871	0.185	2.834	0.1243

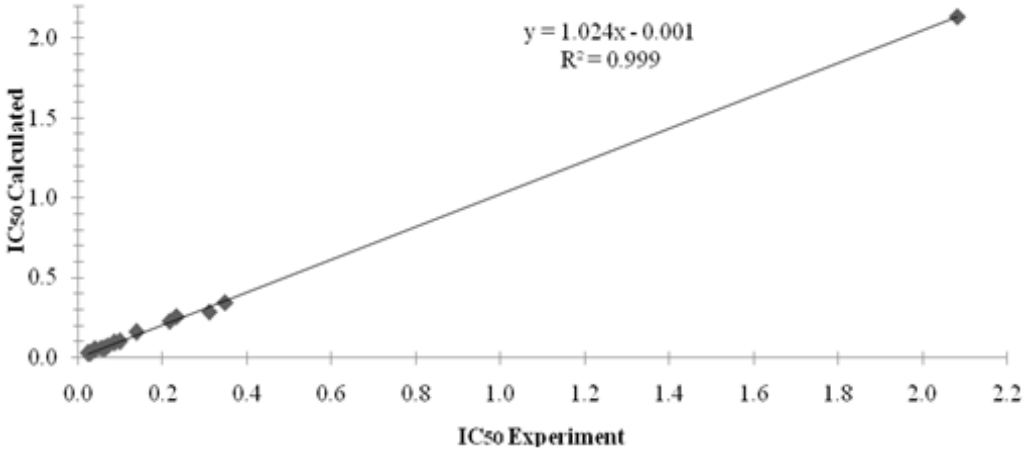


Figure 1. Relationship of the Calculated IC_{50} with the IC_{50} Experiments of Fitting Compounds

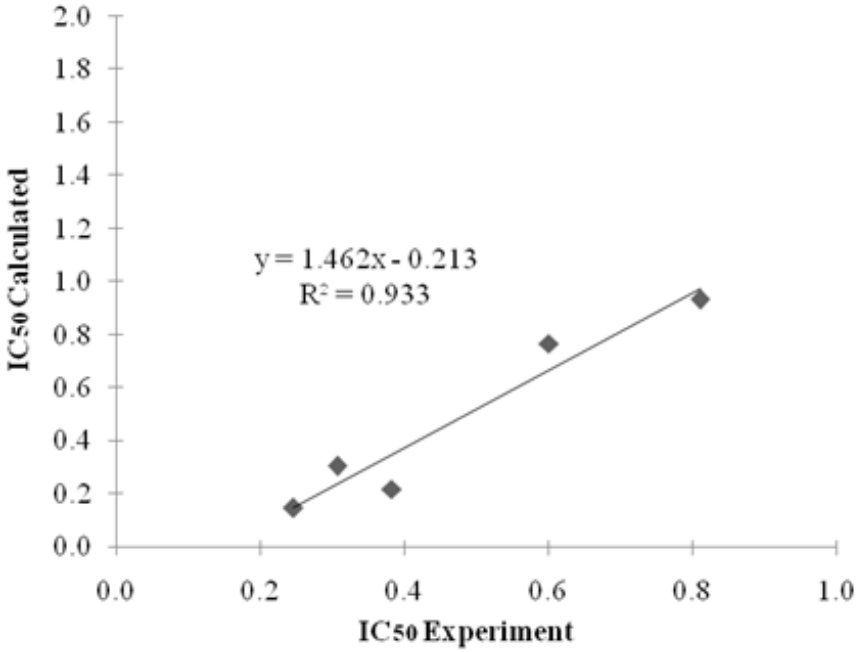


Figure 2. Relationship of the Calculated IC_{50} with the IC_{50} Experiments of Testing Compounds

Table 6. Log $1/IC_{50}$ Experiments and Log $1/IC_{50}$ Theoretical Values for Benzimidazole Derivatives as Testing Compounds from QSAR Model 1

Comp.	Log 1/IC_{50} Experiment	Log 1/IC_{50} Calculated
1.	0.0915	0.0309
5.	0.5143	0.5195
8.	0.2218	0.1152
10.	0.6108	0.8383
17.	0.4179	0.6642

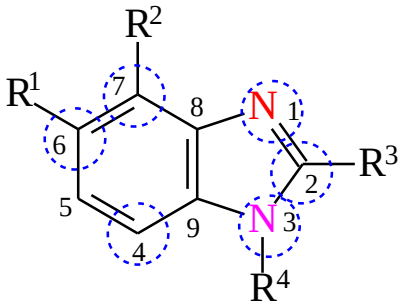


Figure 3. Active Site of the Benzimidazole Homologous Framework as an Antimalarial

Model 1 of the QSAR was chosen as the best QSAR equation because it had better statistical parameters and validation test results than the other models: n = 15, r = 0.997, r^2 = 0.993, adjusted r^2 = 0.906, SE = 0.157, F_{count}/F_{table} = 2.206, and PRESS = 0.0045. Other evidence regarding the validation of model 1 as the best equation could be seen in the graph of the linearity of the relationship between the predictive IC_{50} value and the experimental IC_{50} value (15 training data compounds), which had an r^2 value of 0.999 (Figure 1), and the graph of the linearity of the relationship between the predictive IC_{50} value and the experimental IC_{50} (five testing data compounds), which had an r^2 value of 0.933 (Figure 2). Based on the linearity graphs of the relationships between the predictive IC_{50} value and the experimental IC_{50} value for the internal test compound and the external test compound, model 1 was the best QSAR equation model because it had a very high level of significance.

The statistical parameter r^2 is an indicator of validity in selecting the best QSAR equation model, where an r^2 value close to 1 or −1 indicates that the independent variable strongly influences the dependent variable [17−18]. In addition to the r^2 parameter, small SE and PRESS values indicate a high level of confidence in the QSAR equation model, whereas a high F_{count}/F_{table} indicates the high validity of the QSAR equation model [19-23]. Based on these validation indicators, model 1 of the QSAR equations was the best model. Other evidence of the confidence level of model 1 as the best QSAR equation model is presented in Table 5.

In addition to statistical parameters, the validity of model 1, was tested again using the graph of the predicted IC_{50} relationship with the IC_{50} experiment (Figure 2) for the testing compounds (compounds 1, 5, 8, 10, and 17), which were not involved in the MLR analysis process. Model 1 was tested using the testing compounds to validate its ability to predict the theoretical IC_{50} of benzimidazole derivatives (Table 6). The validity of model 1 was presented as an r^2 value of 0.933 from the graph of the relationship between log 1/IC_{50} predictions and log 1/IC_{50} of the testing compound experiment (Figure 2), which showed that model 1 was the best QSAR equation model. The graph of the relationship between the log 1/IC_{50} predictions and log 1/IC_{50} experiments of the testing compound had an r^2 value of 0.933, which was close to one. This supported the validity of model 1 as the best QSAR equation in predicting the antimalarial activity of benzimidazole derivatives, both for the testing compounds (Table 6) and in the design of new benzimidazole derivative compounds through disconnection theory, retrosynthesis, and modeling (Table 6). Based on these statistical parameters, model 1 of the QSAR equation can predict well the theoretical IC_{50} value of the benzimidazole derivatives because the relationship between the theoretical log 1/IC_{50} and the experimental log 1/IC_{50} of the testing compound had a close and very significant relationship.

Design of new benzimidazole derivatives as antimalarial agents. The steps considered for designing and modeling benzimidazole derivatives are as follows: (1) conducting active site studies on the benzimidazole scaffold; (2) studying atoms or functional groups attached to the benzimidazole scaffold; (3) determining if the target compound (a new benzimidazole derivative) has functional groups that are common in antimalarials; (4) determining if raw materials, reaction reagents, and the synthesis pathway are available to ensure that the new benzimidazole derivative compounds can be synthesized in the laboratory; and (5) designing synthesis reactions based on the starting materials and reagents identified

through disconnection and retrosynthesis studies. We found that differences in atoms or functional groups attached to the basic framework of the target compound can cause differences in the net atomic charge, μ, E_{LUMO}, E_{HOMO}, α, and logP, as shown in Table 2. This also causes variations in antimalarial activity. This is consistent with the findings of Fitriani *et al*. [20] in their QSAR. The validity of this theory is supported by the research results in Tables 1, 3, and 6, which show that compounds with different structures have different biological activities (IC_{50} values), as reported previously [24-26]. Different benzimidazole derivatives have different structures even though they have the same parental structural framework (homologous). This is because of the differences in electronic properties, net atomic charge, μ, E_{LUMO}, E_{HOMO}, α, and logP, which result in different antimalarial activities.

The active sites for benzimidazole derivatives based on the best QSAR equation model (model 1 of the QSAR equation) are in the atoms N1, C2, N3, C4, C6, C7, C8, and C9, as shown in Figure 3. Based on the results of the active center analysis using model 1 of the QSAR equation on a set of benzimidazole derivatives, the independent variables that affected the theoretical IC_{50} value of the benzimidazole compounds were qN1, qC2, qN3, qC4, qC6, qC7, qC8, qC9, μ, E_{LUMO}, E_{HOMO}, α, and logP. The homologous benzimidazole scaffold contained atoms qN1, qC2, qN3, qC4, qC6, and qC7, so the design of new benzimidazole derivatives involved substituting atoms or functional groups on these atoms. This is why the theoretical IC_{50} values of the new benzimidazole derivative compounds differed, as shown in Table 2. The theoretical antimalarial activity (theoretical IC_{50} value) of each benzimidazole derivative was determined using model 1 of the best QSAR equation, which incorporated atomic charge and other descriptors. The aim was to identify the benzimidazole derivative with the smallest IC_{50} value, which could then be proposed as a new antimalarial compound to be synthesized in the laboratory.

A new benzimidazole derivative compound was designed through modeling, which had a smaller IC_{50} value (i.e., was more potent) than that of the antimalarial chloroquine (0.002 μM). This compound was recommended for synthesis in the laboratory [27]. The modeling of benzimidazole derivative compounds was conducted using disconnection, FGI, and retrosynthesis methods and obtained 13 compounds: 23 ($IC_{50} = 2.3 \times 10^{-08}$ μM), 24 ($IC_{50} = 2.3 \times 10^{-16}$ μM), 25 ($IC_{50} = 1.4 \times 10^{-16}$ μM), 26 ($IC_{50} = 2.6 \times 10^{-16}$ μM), 27 ($IC_{50} = 6.0 \times 10^{-08}$ μM), 28 ($IC_{50} = 7.0 \times 10^{-16}$ μM), 29 ($IC_{50} = 1.0 \times 10^{-20}$ μM), 30 ($IC_{50} = 9.0 \times 10^{-04}$ μM), 31 ($IC_{50} = 9.6 \times 10^{-04}$ μM), 37 ($IC_{50} = 5.7 \times 10^{-07}$ μM), 38 ($IC_{50} = 1.1 \times 10^{-19}$ μM), 39 ($IC_{50} = 3.9 \times 10^{-19}$ μM), and 40 ($IC_{50} = 9.6 \times 10^{-04}$ μM). Meanwhile, seven other benzimidazole derivative compounds had higher theoretical IC_{50} values (smaller potential) than chloroquine (0.002 μM) [28] and thus were not recommended for synthesis in the laboratory (Table 6): compounds 21 ($IC_{50} = 3.6 \times 10^{-02}$ μM), 22 ($IC_{50} = 1.6 \times 10^{-01}$ μM), 32 ($IC_{50} = 8.4 \times 10^{-03}$ μM), 33 ($IC_{50} = 4.7 \times 1^{00}$ μM), 34 ($IC_{50} = 1.6 \times 10^{06}$ μM), 35 ($IC_{50} = 6.8 \times 10^{01}$ μM), and 36 ($IC_{50} = 19.7 \times 10^{-02}$ μM). Benzimidazole derivatives were synthesized in the laboratory, prioritizing compounds with simple synthesis pathways as determined using disconnection, FGI, and retrosynthesis methods.

Conclusion

The QSAR analysis of 15 benzimidazole derivatives yielded the best QSAR equation model, which described the relationship between antimalarial activity and independent variables. The best QSAR equation model was model 1, presented as follows:

$\log 1/IC_{50} = 26.892 + 49.892(qN1) - 55.394(qC2) - 128.558(qN3) - 20.978(qC4) + 10.162(qC6) - 6.36(qC7) - 55.991(qC8) - 207.272(qC9) + 0.076(\mu) - 14.914(E_{LUMO}) + 2.866(E_{HOMO}) - 0.460(\alpha) + 1.269(\log P)$; $n = 15$, $r = 0.997$, $r^2 = 0.993$, adjusted $R^2 = 0.906$, SE = 0.157, $F_{count}/F_{table} = 2.206$, and PRESS = 0.0045. The active sites on the benzimidazole structure based on model 1 of the QSAR equation were best found in atoms N1, C2, N3, C4, C6, C7, C8, and C9. Therefore, N1, C2, N3, C4, C6, C7, C8, and C9, E_{LUMO}, E_{HOMO}, α, and log P were variables that affected the antimalarial activity (theoretical IC_{50} value) of the benzimidazole derivatives. A total of 20 new benzimidazole derivative compounds were modeled, including 13 benzimidazole derivative compounds with high (potential) antimalarial activity. These compounds were 23, 24, 25, 26, 27, 28, 29, 30, 31, 37, 38, 39, and 40. These compounds were recommended for synthesis in the laboratory.

Acknowledgment

This research is an independent research, but we would still like to thank the parties who helped from data processing and writing the article to the publication of this article.

Conflict of Interest

In the process of research, data collection, data processing, discussion of results, and the writing of this article, the author does not have a conflict of interest with anyone.

References

[1] Huang, F., Liu, H., Yan, H., Zhou, S., Xia, Z. 2021. Antimalarial drug resistance surveillance in China, 2016–2020. CCDC Weekly. 3(17): 366–371, https://doi.org/10.46234/ccdcw2021.099.

[2] Egwu, C.O., Aloke, C., Chukwu, J., Nwankwo, J.C. Irem, C., Nwagu, K.E., *et al*. 2023. Assessment of the antimalarial treatment failure in Ebonyi State,

Southeast Nigeria. J. Xenobiot. 13(1): 16–26, https://doi.org/10.3390/jox13010003.
[3] WHO. 2021. World Malaria Report 2021. Geneva. Switzerland. pp. 228.
[4] WHO. 2022. World Malaria Report 2022. South Africa. pp. 21.
[5] Mueller, R., Reddy, V., Nchinda, A.T., Mebrahtu, F., Taylor, D., Lawrence, N., *et al*. 2020. Lerisetron analogues with antimalarial properties: synthesis, structure–activity relationship studies, and biological assessment. ACS Omega. 5(12): 6967–6982, https://doi.org/10.1021/acsomega.0c00327.
[6] Singh, K., Okombo, J., Brunschwig, C., Ndubi, F., Barnard, L., Wilkinson, C., *et al*. 2017. Antimalarial pyrido[1,2-a]benzimidazoles: lead optimization, parasite life cycle stage profile, mechanistic evaluation, killing kinetics, and *in Vivo* oral efficacy in a mouse model. J. Med. Chem. 60(4): 1432–1448, https://doi.org/10.1021/acs.jmedchem.6b01641.
[7] Leshabane, M., Dziwornu, G.A., Coertzen, D., Reader, J., Moyo, P., Watt, M.V.D. *et al*. 2021. Benzimidazole derivatives are potent against multiple life cycle stages of *Plasmodium falciparum* malaria parasites. ACS Infect. Dis. 7(7): 1945–1955, https://dx.doi.org/10.1021/acsinfecdis.0c00910.
[8] Cheuka, P.M., Lawrence, N., Taylor, D., Wittlincd, S., Chibale. K. 2018. Antiplasmodial imidazopyridazines: Structure–activity relationship studies lead to the identification of analogues with improved solubility and hERG profiles. Med. Chem. Commun. 9: 1733–1745, https://doi.org/10.1039/c8md00382c.
[9] Pinheiro, L.C.S., Feitosa, L.M., Silveira, F.F.D., Boechat, N. 2018. Current antimalarial therapies and advances in the development of semi-synthetic artemisinin derivatives. An. Acad. Bras. Ciênc. 90(01 Suppl 2): 1251–1271, https://doi.org/10.1590/0001-3765201820170830.
[10] Okombo, J., Brunschwig, C., Singh, K., Dziwornu, G.A., Barnard, L., Njoroge, M., *et al*. 2019. Antimalarial pyrido[1,2-a]benzimidazole derivatives with mannich base side chains: Synthesis, pharmacological evaluation, and reactive metabolite trapping studies. ACS Infect. Dis. 5(3): 372–384, https://doi.org/10.1021/acsinfecdis.8b00279.
[11] Uzor, P.F., Prasasty, V.D., Agubata, C.O. 2020. Natural products as sources of antimalarial drugs. hind. Evid.-Based Compl. Alt. 2020(1): 9385125, https://doi.org/10.1155/2020/9385125.
[12] Nasr, I.S.A., Koko, W.S., Khan, T.A., Schobert, R., Biersack, B. 2023. Antiparasitic activity of fluorophenyl-substituted pyrimido[1,2-a]benzimidazoles. Biomedicines. 11(1): 219, https://doi.org/10.3390/biomedicines11010219.
[13] Mayoka, G., Njoroge, M., Okombo, J., Gibhard, L., Sanches-Vaz, M., Fontinha, D., *et al*. 2019. Structure-activity relationship studies and plasmodium life cycle profiling identifies pan-active n-aryl-3-trifluoromethyl pyrido[1,2- a]benzimidazoles which are efficacious in an *in vivo* mouse model of malaria. J. Med. Chem. 62(2): 1022–1035, https://doi.org/10.1021/acs.jmedchem.8b01769.
[14] Escala, N., Pineda, L.M., Ng, M.G., Coronado, L.M., Spadafora, C., del Olmo, E. 2023. Antiplasmodial activity, structure–activity relationship and studies on the action of novel benzimidazole derivatives. Sci. Rep. 13: 285, https://doi.org/10.1038/s41598-022-27351-z.
[15] Rajkhowa, S., Hussain, I., Hazarika, K.K., Sarmah, P., Deka R.C. 2013. Quantitative structure-activity relationships of the antimalarial agent artemisinin and some of its derivatives-a DFT approach. Comb. Chem. High T. Scr. 16(8): 590–602, https://doi.org/10.2174/1386207311316080002.
[16] Suardiani, L.A.K., Kesuma, D. 2022. Quantitative relationship of structure of compound N-(phenylcarbamothiol)-benzamide and its derivatives to anti-tuberculosis activity *in silico*. CERATA J. Pharm. Sci. 13(1): 31.
[17] Catalani, V., Botha, M., Corkery, J.M., Guirguis, A., Vento, A., Scherbaum, N. *et al*. 2021. The psychonauts' benzodiazepines; quantitative structure-activity relationship (QSAR) analysis and docking prediction of their biological activity. Pharmaceuticals. 14(8): 720, https://doi.org/10.3390/ph14080720.
[18] Zhu, L.L. 2018. QSAR study on the relative activity of thiazide diuretics in doping. IOP Conf. Ser. Earth Environ. Sci. 218: 012137, https://doi.org/10.1088/1755-1315/218/1/012137.
[19] Veerasamy, R., Rajak, H., Jain, A., Sivadasan, S., Varghese, C.P. Agrawal, R.K. 2011. Validation of QSAR models-strategies and importance. Int. J. Drug Design Discovery. 2(3): 511–519.
[20] Fitriani, I.N., Sholahuddin, M., Zikri, A.T. 2022. Analysis of quantitative structure-activity relationship (QSAR) of 1,8-napthalymide-4-aminoquinoline derivatives as antimalarial compounds. Walisongo J. Chem. 5(2): 194–201, https://doi.org/10.21580/wjc.v5i2.13412.
[21] Hadanu, R., Idris S., Sutapa, I.W. 2015. QSAR analysis of benzothiazole derivatives of antimalarial compounds based on AM1 semi-empirical method. Indones. J. Chem. 15(1): 86–92.
[22] Olasupo, S.B., Uzairu, A., Shallangwa, G., Uba, S. 2019. QSAR analysis and molecular docking simulation of norepinephrine transporter (NET) inhibitors as anti-psychotic therapeutic agents. Heliyon. 5(10): e02640, https://doi.org/10.1016/j.heliyon.2019.e02640.
[23] Edache, E.I., Uzairu, A., Mamza, P.A., Shallangwa, G.A. 2020. A comparative QSAR analysis, 3D-QSAR, molecular docking and molecular design of iminoguanidine-based inhibitors of HemO: A rational approach to antibacterial drug design. J. Drugs Pharm. Sci. 4(3): 21–36, https://doi.org/10.31248/JDPS2020.036.

[24] Hadanu, R., Adelin, L., Sutapa, I.W. 2018. QSAR studies of nitrobenzothiazole derivatives as antimalarial agents. Makara J. Sci. 22(1): 35–41, https://doi.org/10.7454/mss.v22i1.7620.
[25] Cherkasov, A., Muratov, E.N., Fourches, D., Varnek, A., Baskin, I.I., Cronin, M., *et al*. 2014. QSAR modeling: Where have you been? Where are you going to? J. Med. Chem. 57(12): 4977–5010, https://doi.org/10.1021/jm4004285.
[26] Zaki, M.E.A., Al-Hussain, S.A., Al-Mutairi, A.A., Masand, V.H., Samad, A., Jawarkar, R.D. 2020. Mechanistic analysis of chemically diverse bromodomain-4 inhibitors using balanced QSAR analysis and supported by X-ray resolved crystal structures. Pharmaceuticals. 15(6): 745, https://doi.org/10.3390/ph15060745.
[27] Souza, N.B.D., Carmo, A.M.L., da Silva, A.D., França, T.C.C., Krettli, A.U. 2014. Antiplasmodial activity of chloroquine analogs against chloroquine-resistant parasites, docking studies and mechanisms of drug action. Malaria J. 13: 469.