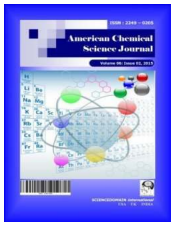


American Chemical Science Journal
11(4): 1-16, 2016, Article no.ACSJ.22639
ISSN: 2249-0205

SCIENCEDOMAIN *international*
www.sciencedomain.org

Structure-toxicity Relationships of Naphthylisoquinoline Derivatives as Antimalarial Agents Using Molecular Descriptors

Zakari Y. Ibrahim[1*], Adamu A. Uzairu[1] and Stephen E. Abechi[1]

[1]Department of Chemistry, Faculty of Science, Ahmadu Bello University, Zaria, Nigeria.

Authors' contributions

This work was carried out in collaboration between all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/ACSJ/2016/22639
Editor(s):
(1) Nagatoshi Nishiwaki, Kochi University of Technology, Japan.
(2) Francisco Marquez-Linares, Full Professor of Chemistry, Nanomaterials Research Group, School of Science and Technology, University of Turabo, USA.
Reviewers:
(1) Gyula Oros, Plant Protection Institute, Hungarian Academy of Sciences, Budapest, Hungary.
(2) Anonymous, University of Campinas, Campinas, Brazil.
(3) Anonymous, Meerut Institute of Engineering and Technology, Meerut, India.
Complete Peer review History: http://sciencedomain.org/review-history/12590

Original Research Article

Received 16th October 2015
Accepted 27th November 2015
Published 8th December 2015

ABSTRACT

The cytotoxicity data of 46 naphthylisoquinoline derivatives that will inhibit 50% of cell growth (pIC_{50}) were used to develop quantitative structure-activity relationships (QSAR). 433 molecular descriptors was obtained from DFT (B3LYP/6-311+G*) level of calculation for each molecule and used in multiple linear regression (MLR) analysis to generate 4 models, out of which the one with the highest statistical significance having correlation coefficient R = 0.791 and cross validated squared correlation coefficient Q^2 = 0.573 was selected as the best model. The QSAR model indicate that the MDE descriptors (MDEC-33) play an important role in the cytotoxicity of naphthylisoquinoline. The accuracy of the proposed MLR model was illustrated using the following evaluation techniques: cross-validation, Y-randomization and external validation on test set. The predictive ability of the model was found to be satisfactory and could be used for designing a similar group of anti-malarial drugs with lower cytotoxicity.

Keywords: QSAR; toxicity; naphthylisoquinoline; IC_{50}; MLR; MDE descriptors; DFT.

**Corresponding author: E-mail: ziyflux4doc@gmail.com;*

1. INTRODUCTION

Malaria represents a significant global health threat, with 40% of the world's population being at risk of contracting the disease. Malaria is a devastating disease in sub-Saharan Africa, where about 90% of cases and deaths occur, it is also a serious public health problem in certain regions of South East Asia and South America. About 630,000 (six-hundred and thirty thousand) people died from the disease in 2012 [1], with pregnant women and children under the age of five being the most vulnerable to infection [2]. Human malaria transmitted by female Anopheles mosquitoes is caused by four species of *Plasmodium*, which are, *P. falciparum*, *P. vivax*, *P. ovale* and *P. malariae*. Of these species, *P. falciparum* is responsible for the most severe form of malaria [3]. The endemic nature of malaria is rapidly getting worse partly due to the unavailability of effective drugs and partly due to development of resistance towards some traditional drugs such as chloroquine and pyrimethamine [4-8].

Although there are vast numbers of available antimalarial drugs, the control of this ancient infection is increasingly threaten by the emergence of drug-resistant strains of the malaria parasite, *Plasmodium* [9]. With drug resistance becoming a rapidly increasing problem and given the lack of suitable vaccines, the development of efficient, non-toxic, and inexpensive new drugs is an urgent task [10-12].

The naphthylisoquinoline alkaloids (NIQs) represent a class of natural products isolated from rare and difficult-to-cultivate tropical plants. They have manifold activities against various tropical diseases. Some C, C-linked naphthylisoquinolines, among them e.g. dioncophylline C and dioncopeltine A, show interesting *in vitro* and even *in vivo* activities against *P. falciparum* [13-15]. These and other promising bioactivities make the naphthylisoquinoline alkaloids suitable pharmaceutical lead structures for the synthesis of new potent agents [16,17].

The Structure– Property (Activity) Relationship (QSAR/QSPR) methods have been efficiently used for the study of toxicity mechanisms of various reactive chemicals. This is a powerful technique, which quantitatively relates variations in toxicity/activity to changes in molecular properties of the compounds, in terms of descriptors. In the area of computer – aided toxicity prediction, quantitative structure activity relationship (QSAR) have been seen as an attractive method for toxicity and fate assessment [18], which has been a problem for a very long time. The study of the quantitative relationship between toxicity/activity and molecular structure (QSTR/QSAR) is an important area of research in computational chemistry and has been widely used in the prediction of toxicity and other biological activities of organic compounds [19,20], thereby saving resources and expedite the process of the development of new molecules and drugs. Although various degree of research have been conducted on naphthylisoquinoline [21-24], research involving the use molecular descriptors to calculate molecular cytotoxicity of naphthylisoquinoline alkaloids have not been reported.

This research is aim at finding the accuracy of QSAR analysis in predicting the cytotoxicity of naphthylisoquinoline alkaloids, and also to investigate the descriptor(s) responsible for producing such toxicity. The result obtained will guide further structural optimization and predict the potency and physiochemical properties of clinical drug candidates.

2. MATERIALS AND METHODS

2.1 Methodology

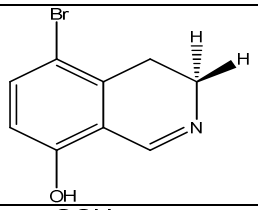
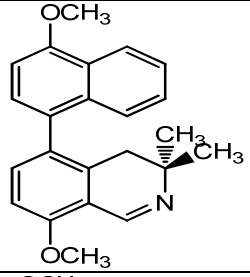
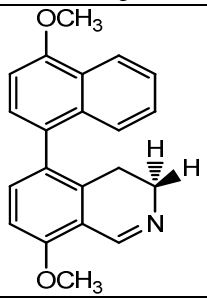
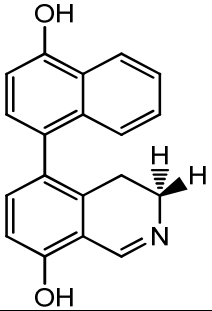
A data set of 44 compounds of naphthylisoquinoline with their cytotoxicity has been taken from published journals [22,23]. The Cytotoxicity value [IC_{50} (µM)] reported in the literature were converted to their molar units and then further to negative logarithm scale (pIC_{50}) and subsequently used as the dependent variable for the QSAR analysis. The structures and cytotoxic activity data of the compounds are listed in Table 1.

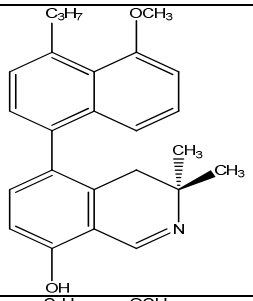
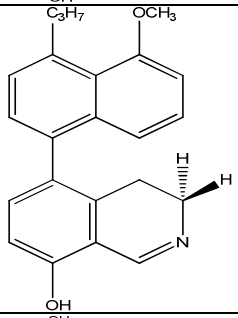
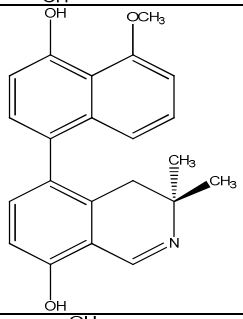
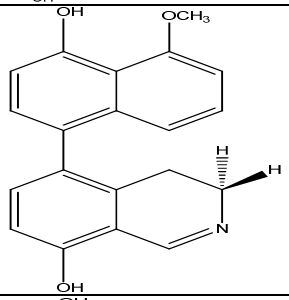
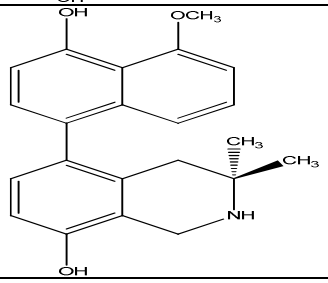
2.2 Analysis Procedures and Descriptor Generation

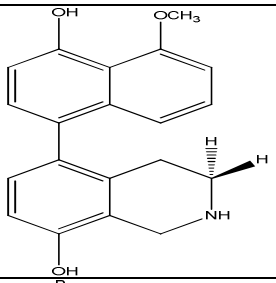
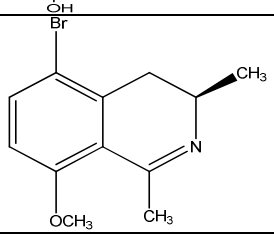
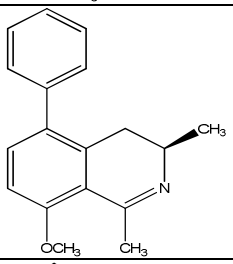
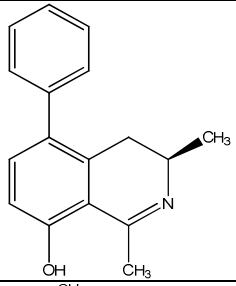
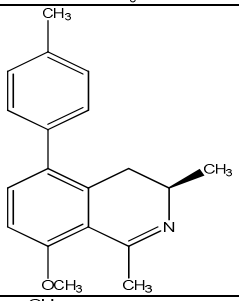
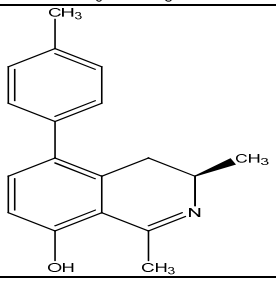
The compounds were sketched using the ChemBioDraw software. The sketched structures ware then transferred to Spartan 14 v.1.1.0 for the generation of 3D structures and geometric optimization. The geometries of the generated 3D were pre-optimized using molecular mechanics force field (MMFF) in the Spartan. The Geometry optimizations was performed through B3LYP/6-311G* at Spartan 14 v.1.1.0.

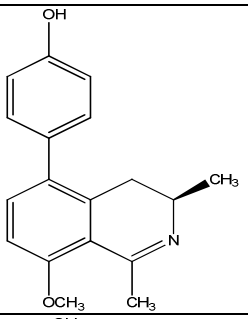
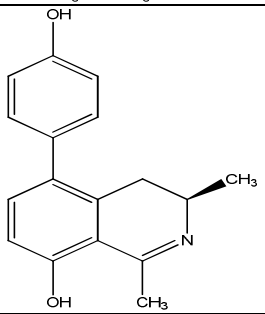
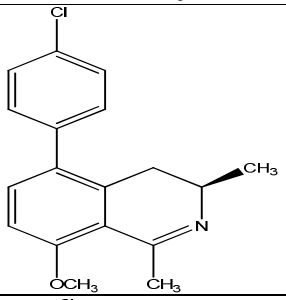
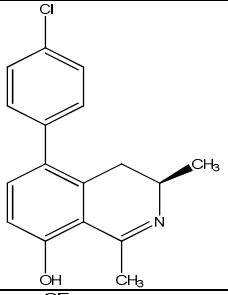
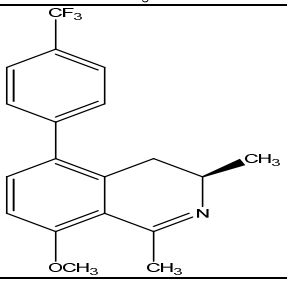
The minimum energy structures were used to obtain the electronic descriptors. Some descriptor values of all the molecules were calculated using PaDEL software while other descriptors were calculated from the ChemBio 3D Ultra software. Other chemical and physicochemical properties were determined by the chemical structure (lipophilicity, hydrophilicity descriptors, electronic descriptors, and energies of interaction). The preprocessing of the independent variables (descriptors) was done by removing invariable, which resulted in 433 descriptors in total, which was further treated by excluding descriptors that are highly inter-correlated to avoid data redundancy, bringing the descriptors numbers to 94 used for QSAR analysis and was considered as independent variables in this study. On the basis of Kennard-Stones algorithm, 34 compounds out of 44 were selected as the training set (for generation the models) and the remaining 10 were selected as the test set (for validation of the models). Compounds 20, 27, 29, 30, 31, 33, 35, 37, 39 and 40 as thus the test sets, while the remaining compounds are the training sets.

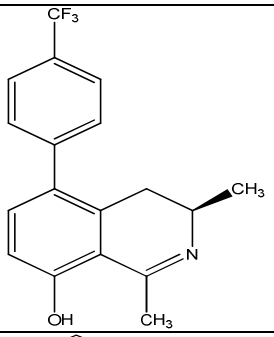
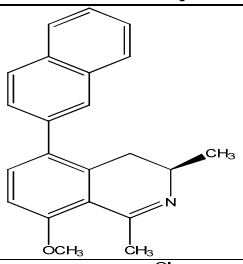
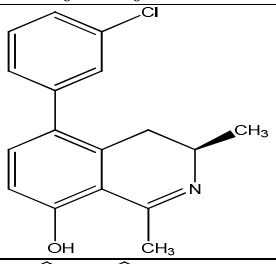
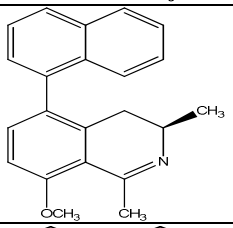
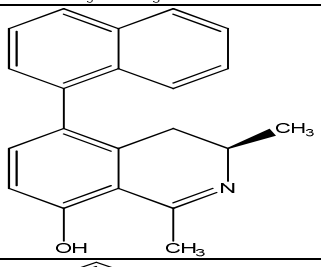
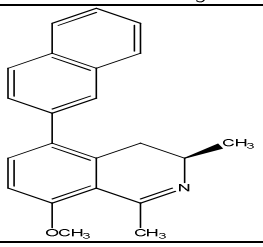
Table 1. Series of naphthylisoquinoline alkaloids with their cytotoxicity against L 6 cell line

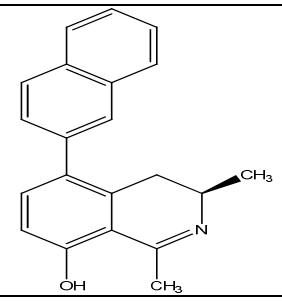
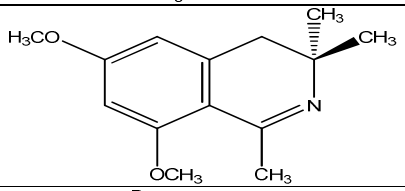
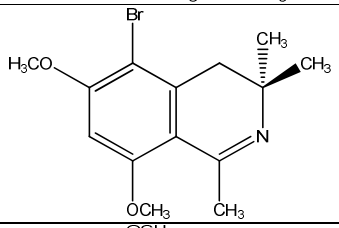
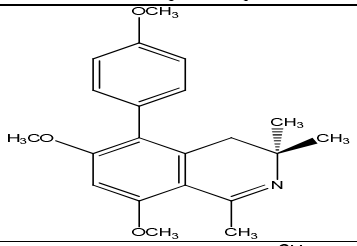
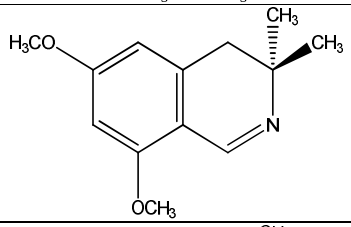
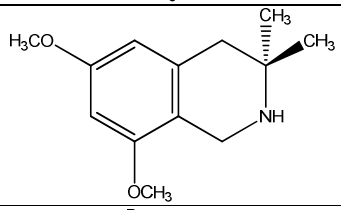
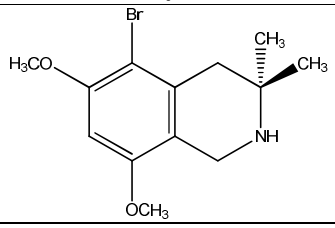
S/N	Structures	Cytotoxic activity IC_{50} (μM)	pIC_{50}
1		100	4.00
2		43.79	4.36
3		34.34	4.46
4		94.18	4.03

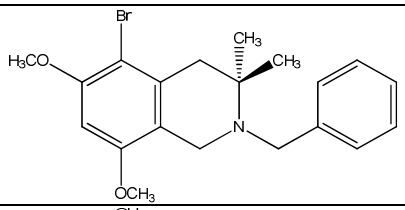
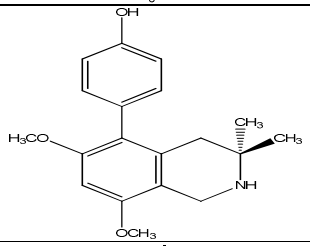
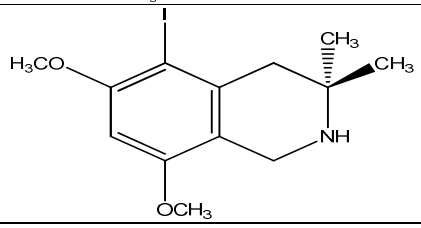
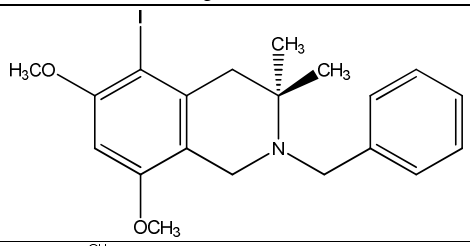
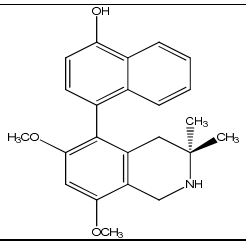
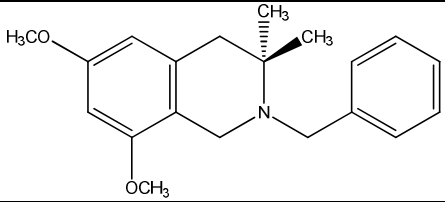
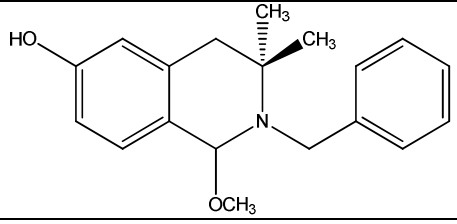
S/N	Structures	Cytotoxic activity IC_{50} (μM)	pIC_{50}
5		100	4.00
6		27.00	4.57
7		100	4.00
8		98.20	4.01
9		37.51	4.43

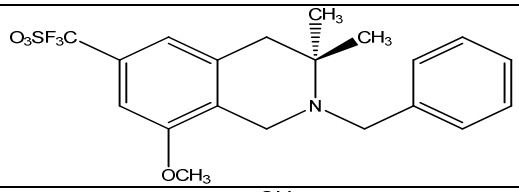
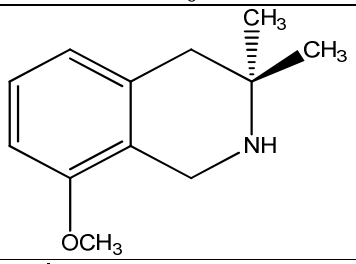
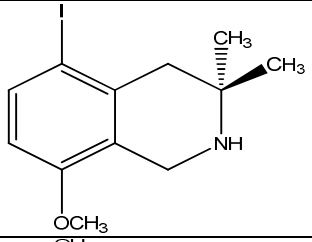
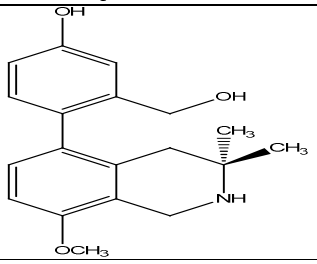
S/N	Structures	Cytotoxic activity IC_{50} (μM)	pIC_{50}
10		40.23	4.40
11		100	4.00
12		58.84	4.23
13		52.22	4.28
14		100	4.00
15		100	4.00

S/N	Structures	Cytotoxic activity IC_{50} (μM)	pIC_{50}
16		100	4.00
17		100	4.00
18		56.70	4.25
19		100	4.00
20		100	4.00

S/N	Structures	Cytotoxic activity IC_{50} (μM)	pIC_{50}
21		96.76	4.01
22		38.36	4.42
23		50.76	4.29
24		34.87	4.46
25		57.07	4.24
26		71.98	4.14

S/N	Structures	Cytotoxic activity IC_{50} (μM)	pIC_{50}
27		72.54	4.14
28		386.03	3.41
39		184.86	3.73
30		127.95	3.89
31		410.72	3.39
32		406.98	3.39
33		300.95	3.52

S/N	Structures	Cytotoxic activity IC_{50} (μM)	pIC_{50}
34		217.17	3.66
35		90.05	4.05
36		95.67	4.02
37		151.00	3.82
38		23.13	4.64
39		289.21	3.54
40		267.86	3.57

S/N	Structures	Cytotoxic activity IC_{50} (μM)	pIC_{50}
41		209.73	3.68
42		470.88	3.33
43		264.65	3.58
44		287.39	3.54

All the calculated descriptor values were considered as independent variable and cytotoxicity as dependent variable. BuildQSAR software was used to generate QSAR models by multiple linear regression analysis. Statistical measures used were: R^2_correlation-coefficient, Q^2- leave-one-out (LOO) cross-validation, F-test (Fischer's value) for statistical significance, s-standard deviation, R^2_{rand} - Y-randomization. The values of the calculated descriptors used for the multiple linear regression are showed in Table 2. The correlation matrix generated is showed in Table 3. In addition to low p-value or high F-statistics, a QSAR model is considered to be predictive if it satisfied the following conditions: $R^2 > 0.6$, $Q^2 > 0.6$ and $R^2_{pred} > 0.5$ [25,26]. From the predictive models selected, the model with largest values of R and R^2, smallest values of SEE, highest values of F and smallest values of PRESS and most especially highest value of R^2_m (overall) was choosing as the best model. The validation was done using the data from which the model was created (an internal method) and using a separate data set (an external method). In the internal validation, the following parameters were determined as seen in Table 5, least squares fit, R^2 (coefficient of determination) for the comparison between the predicted and experimental activities,. The leave-one-out (LOO) cross-validation, Q^2 used to evaluate the predictive power of the model and the Y-randomization, (R^2_{rand}) which ensures that the model is not due to a chance. However, a high Q^2 value does not necessarily give a suitable representation of the real predictive power of the model so, an external validation was performed in order to make more realistic validation of the predictive power of the models. The determined parameters as showed in Table 4 are R^2_{pred} which points to the external predictability of the model, $[(R^2 - R_o^2)/ R^2]$, $[(R^2 - R'^2_o)/ R^2]$ and R^2_m (overall), which includes prediction for both test set and training set (using LOO predictions) compounds.

Table 2. Descriptors used for generation of models

Compd No.	pIC50	BCUTp-1h	ETA_Eta_B	MDEC-33	XLogP	Weta3. unity	Weta1. polar	η
1	4.00	-1.0491	-1.5806	-1.3959	-1.7165	1.3643	0.5399	-0.0199
2	4.36	0.3215	0.0292	0.5551	0.6932	-0.6961	-0.1694	-0.9601
3	4.46	-0.0497	-1.8525	0.5551	0.0457	-0.0831	-0.0698	-0.8896
4	4.03	-0.1889	-1.3641	0.5551	-0.6079	0.5181	0.2531	-1.2187
5	4.00	0.4062	0.6030	1.2154	1.8171	1.3155	0.6031	-1.2657
6	4.57	0.2204	-1.2786	1.2154	1.1696	1.0703	0.4683	-1.7358
7	4.00	0.2634	0.8472	1.2154	-0.1457	-0.2102	-0.1137	-1.8533
8	4.01	-0.0350	-1.0345	1.2154	-0.7932	0.1441	-0.0740	-1.8533
9	4.43	0.2587	0.8472	1.2154	-0.2618	-0.0663	-0.1259	-0.9601
10	4.40	-0.0373	-1.0345	1.2154	-1.1546	0.3695	-0.0836	-0.9836
11	4.00	-0.3176	-0.5689	-0.3858	-0.6700	0.5375	-0.0771	0.0506
12	4.23	-0.2921	-1.0345	0.1831	0.6473	-0.5306	0.0845	0.1681
13	4.28	-0.4777	-0.7903	0.1831	0.3205	-0.5848	0.0743	0.2621
14	4.00	-0.2457	-0.3525	0.3413	1.0464	-0.5993	-0.1694	0.0741
15	4.00	-0.4313	-0.1083	0.3413	0.7196	-0.4901	-0.1374	0.1681
16	4.00	-0.2747	-0.3525	0.3413	-0.5092	-0.8112	-0.2575	-0.2080
17	4.00	-0.4604	-0.1083	0.3413	-0.8359	-0.7474	-0.2290	-0.1610
18	4.25	-0.2199	-0.3525	0.3413	0.1607	-0.9769	0.9299	0.1211
19	4.00	-0.4055	-0.1083	0.3413	-0.1661	-1.0236	-1.7910	0.2386
21	4.01	-0.4364	1.7461	0.3413	0.5720	-0.5763	-1.6542	0.3326
22	4.42	-0.0985	-0.8181	0.9212	1.8099	-0.1072	-1.7272	-0.6310
23	4.29	-0.3326	-0.1083	0.5019	-0.1661	-0.6973	-1.5539	0.3091
24	4.46	0.2439	-0.9263	1.2890	1.8099	0.1929	-1.4322	-0.4195
25	4.24	0.1510	-0.6821	1.2890	1.4832	-0.0831	0.2679	-0.2080
26	4.14	-0.0985	-0.8181	0.9212	1.8099	-0.1072	0.4540	-0.6310
28	3.41	-1.3308	0.4947	-0.9803	-1.0803	1.5652	0.6003	0.4032
32	3.39	-1.8477	-0.0791	-1.5120	-1.3806	2.2237	0.9094	2.1190
34	3.66	0.4092	0.5385	-0.8033	0.7410	-0.4475	-1.4252	1.4844
36	4.02	2.7641	0.3865	-0.8279	-0.3391	1.6128	-0.3562	0.2856
38	4.64	0.9195	0.6030	1.4195	0.0650	-0.9665	-0.0233	-1.0071
41	3.68	0.1039	4.0531	-1.4457	-0.5051	0.0936	0.8175	-1.1716
42	3.33	-2.1723	-0.5169	-1.8208	-1.4101	0.6935	0.2353	2.2835
43	3.58	2.1150	0.0569	-1.3959	-0.5835	0.6992	4.3539	0.4032
44	3.54	-0.0630	0.6030	-0.1779	-1.7837	-1.0235	-0.2280	1.2258

Table 3. Correlation matrix of physicochemical parameters used in model generation

	pIC50	BCUTp-1h	ETA_Eta_B	MDEC-33	XLogP	Weta3. unity	Weta1. polar	η
pIC50	1							
BCUTp-1h	0.31722	1						
ETA_Eta_B	-0.35127	0.17061	1					
MDEC-33	0.79096	0.22554	-0.25414	1				
XLogP	0.52444	0.27590	-0.08202	0.56454	1			
Weta3.unity	-0.34988	-0.05348	-0.10001	-0.41223	-0.22058	1		
Weta1.polar	-0.31851	0.19169	0.01322	-0.36138	-0.25725	0.39372	1	
η	-0.63759	-0.37003	0.03445	-0.69634	-0.32137	0.08917	-0.00572	1

Table 4. Golbraikh and Tropsha acceptable model criteria for selected models

Models	Q^2	R^2_{EXT}	$\lvert R_o^2 - R'^2_o \rvert$	$[(R^2 - R_o^2)/ R^2]$	K
1	0.5729	0.7865	0.0217	0.0056	0.9738
2	0.5251	0.8063	0.0336	0.0008	0.9749
3	0.5935	0.7044	0.0697	0.0053	0.9747
4	0.5762	0.6716	0.0182	0.0359	0.9695

Table 5. Statistical parameters of the 4 best scored models

Models	R^2	R^2_{adjs}	F	s	R^2_m (overall)	SPress	SDEP	n	k
1*	0.6256	0.6139	53.4729	0.2120	0.6339	1.6410	0.2197	34	1
2	0.6842	0.6526	21.6648	0.2011	0.5580	1.8246	0.2317	34	3
3	0.7098	0.6579	13.6937	0.1996	0.6236	1.5616	0.2143	34	5
4	0.7341	0.6751	12.4258	0.1945	0.5830	1.6284	0.2188	34	6

*NB. *= Best of the selected model*

Table 6. Comparison of observed toxicity with predicted toxicity of the selected models

Compd no.	Observed activity (μM)	Predicted cytotoxicity (μM)			
		*Model 1	Model 2	Model 3	Model 4
1	4.00	3.60	3.68	3.90	3.87
2	4.36	4.15	4.16	4.30	4.30
3	4.46	4.15	4.25	4.38	4.38
4	4.03	4.15	4.21	4.27	4.26
5	4.00	4.34	4.29	4.17	4.14
6	4.57	4.34	4.40	4.43	4.41
7	4.00	4.34	4.26	4.27	4.25
8	4.01	4.34	4.36	4.38	4.38
9	4.43	4.34	4.26	4.08	4.09
10	4.40	4.34	4.36	4.18	4.20
11	4.00	3.89	3.92	3.96	3.97
12	4.23	4.05	4.09	4.17	4.16
13	4.28	4.05	4.06	4.11	4.09
14	4.00	4.09	4.09	4.17	4.15
15	4.00	4.09	4.06	4.09	4.06
16	4.00	4.09	4.08	4.13	4.12
17	4.00	4.09	4.05	4.06	4.05
18	4.25	4.09	4.09	4.07	4.04
19	4.00	4.09	4.06	4.15	4.16
20^	4.00	4.09	3.96	3.96	3.93
21	4.01	4.09	3.93	3.95	3.92
22	4.42	4.25	4.27	4.45	4.44
23	4.29	4.13	4.10	4.09	4.11
24	4.46	4.36	4.39	4.38	4.40
25	4.24	4.36	4.37	4.22	4.21
26	4.14	4.25	4.27	4.33	4.29
27^	4.14	4.25	4.24	4.27	4.22
28	3.41	3.72	3.62	3.63	3.56
29^	3.73	3.93	3.91	3.79	3.83
30^	3.89	4.15	4.11	4.13	4.18
31^	3.39	3.57	3.49	3.66	3.56
32	3.39	3.57	3.49	3.28	3.22
33^	3.52	3.76	3.77	3.48	3.53
34	3.66	3.77	3.80	3.85	3.92
35^	4.05	3.96	3.94	3.88	3.92
36	4.02	3.76	3.98	3.77	3.99
37^	3.82	3.77	3.99	4.01	4.20
38	4.64	4.39	4.38	4.21	4.26
39^	3.54	3.59	3.59	3.70	3.68
40^	3.57	3.58	3.61	3.73	3.75
41	3.68	3.59	3.38	3.70	3.65
42	3.33	3.48	3.42	3.47	3.39
43	3.58	3.60	3.81	3.59	3.69
44	3.54	3.94	3.91	3.68	3.73

Numbers with ^ represent the test sets

Table 7. R^2 and Q^2 values after Y-randomization

Iteration	MLR	
	R^2	Q^2
1	0.015	0.000
2	0.006	0.000
3	0.124	0.015
4	0.189	0.096
5	0.128	0.020
6	0.000	0.000
7	0.003	0.000
8	0.017	0.000
9	0.131	0.009
10	0.001	0.000
11	0.023	0.000
12	0.038	0.000
13	0.011	0.000
14	0.009	0.000
15	0.016	0.000

The applicability domain (AD) for the best models was checked by the leverage approach to verify prediction reliability [27,28]. Not even a robust, significant, and validated QSAR model can be expected to reliably predict the modeled property for the entire universe of chemicals. In fact, only the predictions for chemicals falling within this domain can be considered reliable and not model extrapolations To visualize the applicability domain of a QSAR model, the Williams plot – the plot of standardized residuals versus leverage values for graphical detection of both response outliers (Y outliers) and structurally influential chemicals (X outliers) in a model – is used. Leverage values was calculated for both training compounds and test compounds. A leverage higher than the warning leverage *h** means that the compound predicted response can be extrapolated from the model, and thus, the predicted value must be used with great care. On the other hand, a standardized residual value greater than two indicates that the value of the dependent variable for the compound is significantly separated from the remainder training data, and hence, such predictions must be considered with much caution too. In this work, only predicted data for new compounds belonging to the applicability domain of the training set were considered reliable.

3. RESULTS AND DISCUSSION

3.1 QSAR Results

Selection descriptors that correlate to cytotoxic activity is an important step in QSAR modeling. In this research, a total number of 433 descriptors ranging from 0D- to 3D- classes of descriptors available in PaDEL have been computed and were subjected to MLR analysis. The QSAR models were constructed using a the BuildQSAR software - employed to systematically search for models with one or more variables which give rise to multiple linear regression (MLR) models – by setting the $R^2 > 0.65$. The correlation between the different physicochemical descriptors as independent variable and the pIC_{50} as dependent variable was determined.

Model 1.

pIC_{50} = + 3.9934 (± 0.0761) + 0.2810 (± 0.0785) MDEC-33
(n = 34; R = 0.791; s = 0.212; F = 53.473; p < 0.0001; Q^2 = 0.573; SPress = 0.226; SDEP = 0.223)

Model 2.

pIC_{50} = + 3.9971 (± 0.0730) + 0.0755 (± 0.0852) BCUTp-1h - 0.0655 (± 0.0702) ETA_Eta_B + 0.2464 (± 0.0802) MDEC-33
(n = 34; R = 0.827; s = 0.201; F = 21.665; p < 0.0001; Q^2 = 0.525; SPress = 0.247; SDEP = 0.235)

Model 3.

pIC_{50} = + 4.0053 (± 0.0722) - 0.1047 (± 0.0663) ETA_Eta_B + 0.0765 (± 0.0755) XLogP - 0.0859 (± 0.0917) Weta.unity - 0.0540 (± 0.0732) Weta1.polar - 0.1814 (± 0.0755) η
(n = 34; R = 0.842; s = 0.200; F = 13.694; p < 0.0001; Q^2 = 0.594; SPress = 0.236; SDEP = 0.218)

Model 4.

pIC_{50} = + 4.0121 (± 0.0711) + 0.0701 (± 0.0914) BCUTp-1h - 0.1157 (± 0.0663) ETA_Eta_B + 0.0611 (± 0.0764) XLogP - 0.0815 (± 0.0898) Weta3.unity - 0.0700 (± 0.0745) Weta1.polar - 0.1636 (± 0.0773) η
(n = 34; R = 0.857; s = 0.194; F = 12.426; p < 0.0001; Q^2 = 0.576; SPress = 0.246; SDEP = 0.222)

3.2 Discussion

Four (4) of the statistically significant models are given above and their statistical measures are listed in Table 5. The participated descriptors in these models are BCUTp-1h, MDEC-33,

ETA_Eta_B, XLogP, Weta3.unity, Weta1.polar and η. These descriptors represent, respectively, nlow highest polarizability weigheted BCUT, Molecular distance edge between all tertiary nitrogens, Branching index EtaB, partition coefficient, Directional WHIM, weighted by units weights, Directional WHIM, weighted by atomic polarizability and absolute hardness. While BCUT (Burden-CAS-University of Texas eigenvalues) descriptor, MDEC-33, Calculate Molecular Distance Edge (MDE) Descriptors for C, N and O, WHIM (Weighted Holistic Invariant Molecular descriptors) encodes information on the structural fragments, XLogP encodes prediction of the octanol/water partition coefficients of organic compounds and absolute hardness of the compounds. Extended topochemical descriptor. The descriptors, in all the models, have been scaled between the intervals 0 to 1[25] to ensure that a descriptor will not dominate simply because it has larger or smaller pre-scaled value compared to the other descriptors. In this way, the scaled descriptors would have equal potential to influence the QSAR models. The values of the coefficient of determinations as well as that of Q^2, all show that the model is predictive. The result of the Y-randomization performed as in Table 7 shows that the model did not occur by chance. The signs of the regression coefficients have indicated the direction of influence of explanatory variables in above models. The positive regression coefficient associated to a descriptor will augment the activity profile of a compound while the negative coefficient will cause detrimental effect to it. In all the generated models, the MDEC-33 descriptor is the most important descriptors. The positive regression coefficient of the descriptor ensures that decreasing the value of the descriptors would lead to lower cytotoxic prediction. In the Williams plot, it will be observe that all the compounds (Training + Test set) are all within the domain of the selected model.

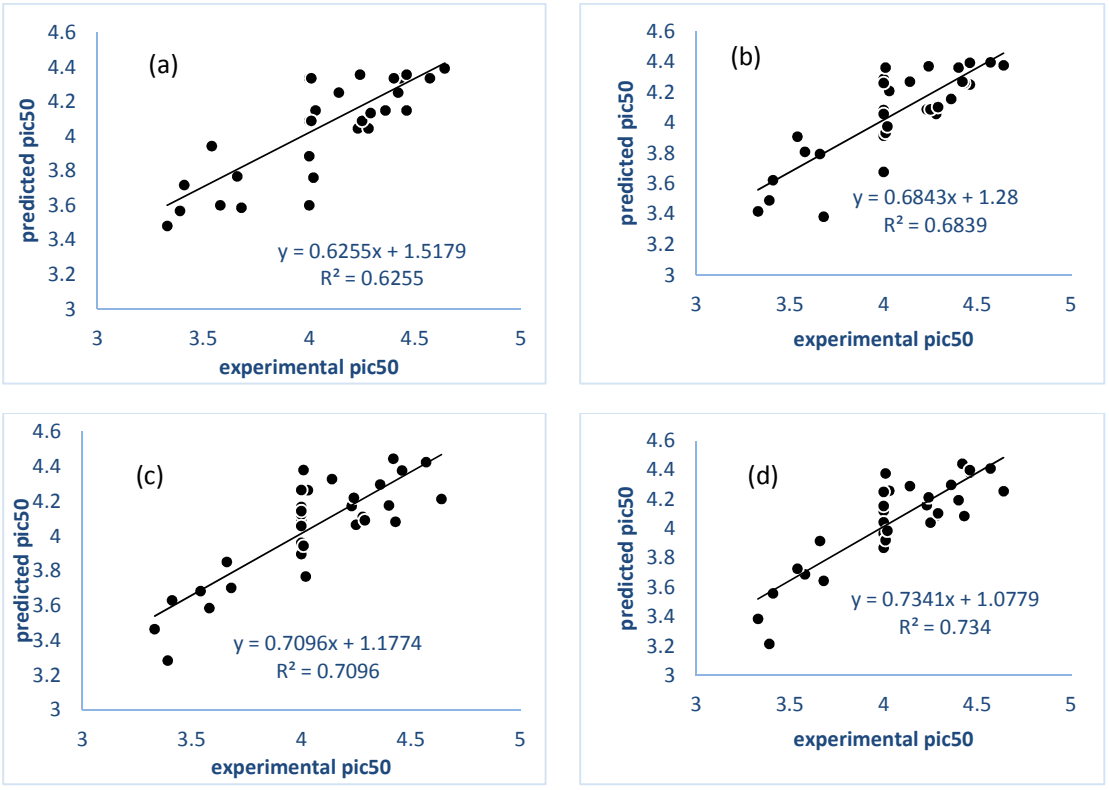


Fig. 1. (a) Correlation between the predicted pIC_{50} and the experimental pIC_{50} by eq. (1). (b) Correlation between the predicted pIC_{50} and the experimental pIC_{50} by eq. (2). (c) Correlation between the predicted pIC_{50} and the experimental pIC_{50} by eq. (3). (d) Correlation between the predicted pIC_{50} and the experimental pIC_{50} by eq. (4)

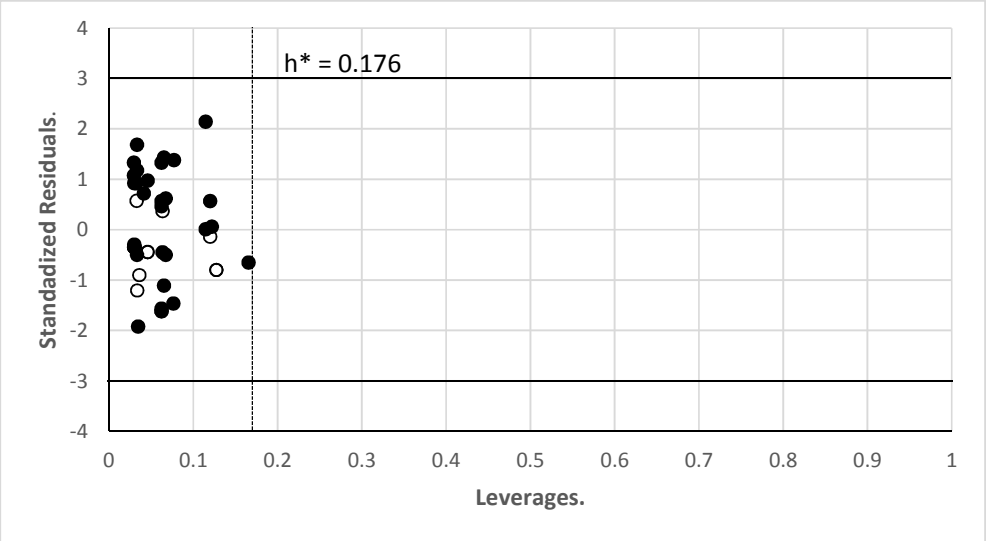


Fig. 2. Williams Plot for the training as well as the test sets. cut-off value: h* = 0.176
NB ● Training set ○ Test set

4. CONCLUSION

In this article, a QSAR study of 44 antimalarial drugs was performed based on the theoretical molecular descriptors calculated mostly using the PaDEL software. Models were generated, Out of which model 1 consisting of MDEC-33 was selected as the best model. Validations indicated that the QSAR model built was robust and satisfactory. And that the selected descriptor plays a vital role in predicting the cytotoxicity of naphthylisoquinolines. The descriptor can be considered for further designing of newer molecules with lower toxicity and better activity for the treatment of malarial.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. WHO: World Malaria Report 2013. Geneva: World Health Organization; 2013 Available:http://www.who.int/malaria/publications/world_malaria_report_2013/report/en/
(Date of access: 19.03.2014)
2. Shetty P. The numbers game. Nature. 2012;484:14–15.
3. Guerra CA, Howes RE, Patil AP, Gething PW, Van Boeckel TP, Temperley WH, Kabaria CW, Tatem AJ, Manh BH, Elyazar RE, Baird JK, Snow RW, Hay SI. The international limits and population at risk of *Plasmodium vivax* transmission in 2009. PLoS Negl Trop Dis. 2010;4(8):e774.
4. Nosten F, ter Kuile F, Chongsuphajaisiddhi T, Luxemburger C, Wester HK, Edstein M, Phaipum L, Thew KL, White NJ. Mefloquine-resistant falciparum malaria on the Thai-Burmese border. Lancet. 1982; 337:1140-1143.
5. Oduola AM, Milhous WK, Salako LA, Walker O. Desjardins RE. Reduced in-vitro susceptibility to mefloquine in West African isolates of *Plasmodium falciparum*. Lancet 1987;2:1304-5.
6. Carson PE, Flanagan CL, Ickes CE, Alving AS. Enzymatic Deficiency in Primaquine-Sensitive Erythrocytes. Science. 1956; 124:484-485.
7. Carson PE, Hohl R, Nora MV, Parkhurst GW, Ahmad T, Scalan S, Frischer H, Bull WHO. 1981;59:427.
8. Schmidt LH, Frank R, Genther CS, Rossan RN, Squires WI. Am J Trop Med Hyg. 1982;31:612.
9. Craft JC. Challenges facing drug development for malaria. Curr. Opin. Microbiol. 2008;11:428–433.
10. Hoet S, Opperdoes F, Brun R, Quetin-Leclercq J. Natural products active against African trypanosomes: A step towards new drugs. Nat Prod Rep. 2004;21:353-364.
11. Edwards G, Biagini GA. Resisting resistance: Dealing with the irrepressible

problem of malaria. Br J Clin Pharmacol. 2006;61:690–693.
12. Singh N. Drug resistance mechanisms in clinical isolates of *Leishmania donovani*. Indian J. Med. Res. 2006;123:411-422.
13. François G, Bringmann G, Phillipson JD, Aké Assi L, Dochez C, Rübenacker C, Schneider M, Wéry DC, Kirby GC. Phytochemistry. 1994;35:1461-1464.
14. François G, Timperman G, Eling W, Ake Assi L, Holenz J, Bringmann G. Antimicrob. Agents Chemother. 1997;41: 2533–2539.
15. Wiesner J, Ortmann R, Jomaa H, Schlitzer M. Angew. Chem. Int. Ed. 2003;42:5274-5293.
16. Bringmann G, Hinrichs J, Peters K, Peters EM. Synthesis of a chiral aryl-ferocenyl ligands by intramolecular coupling to a biaryl-related lactone. J. Org. Chem. 2001;66:629–632.
17. Hallock YF, Manfredi KP, Blunt KP, Cardellina II JH, Schäffer M, Gulden KP, Bringmann G, Lee AY, Clardy J, François G, Boyd MR. J. Org. Chem. 1994;59: 6349-6355.
18. Knight DJ, Breheny D. Alternatives to animal testing in the safety evaluation of products. Altern. Lab Anim. 2002;30(1):7-22.
19. Katritzky AR, Maran U, SLobanov V, krelson M. J. Chem. Inf. Comput. Sci. 2000;41:345-390.
20. Katritzky A, Pentrukhin R, Tatham R, Basak D, Benfenatim S, Karelaon E, Maran M. U. J. Chem. Inf. Comput. Sci. 2001;40:1–18.
21. Bringmann G, Rummey C. 3D QSAR investigations on antimalarial naphthylisoquinoline alkaloids by comparative molecular similarity indices analysis (CoMSIA), based on different alignment approach. J. Chem. Inf. Comput. Sci. 2003;43:304–316.
22. Bringmann G, Brun R, Kaiser M, Neumann S. Synthesis and antiprotozoal activities of simplified analogs of naphthylisoquinoline alkaloids. Eur J Med Chem. 2008;43:32-42.
23. Bringmann G, Bischof SK, Muller S, Gulder T, Winter C, Stich A, Moll H, Kaiser M, Brun R, Dreher J, Baumann F. QSAR guided synthesis of simplified antiplasmodial analogs of naphthylisoquinoline alkaloids. Eur J Med Chem. 2010;45:5370-5383.
24. Stiefl N, Bringmann G, Rummey C, Baumann K. Evaluation of extended parameter sets for the 3D-QSAR technique MaP: Implications for interpretability and model quality exemplified by antimalarially active naphthylisoquinoline alkaloids. J. Comput. Aided Mol. Des. 2003;17:347–365.
25. Golbraikh A, Tropsha A. Beware of q2! J Mol Graph Model. 2002;20:269-276.
26. Afantitis A, Melagraki G, Sarimveis H, Igglessi-Makorpoulon O, Kollias G. Eur J Med Chem. 2009;44:877.
27. Gramatica P, Corradi M, Consonni V. Modelling and prediction of soil sorption coefficients of non-ionic organic pesticides by molecular descriptors. Chemosphere. 2000;41:763–777.
28. Gramatica P. Minireview: Principles of QSAR models validation: Internal and external. QSAR Comb. Sci. 2007;26:694–701.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/12590