Journal of Chronotherapy and Drug Delivery

ORIGINAL RESEARCH PAPER

QSAR Analysis of 7-Substituted 4-Aminoquinolines for Designing Potent Antimalarial Agents

Nitendra K. Sahu[1]* Vivek Chourasia[1], Sanjay J. Ingle[2], Kamlendra S. Bhadoriya[3]

[1]RKDF College of Pharmacy, SRK University, Bhopal 462026 (India)
[2]Vidyabharti College of Pharmacy, Amravati (India)
[3]School of Pharmacy, ITM University, Gwalior 474001 (India)

* For correspondence
Email: nksahupharma@gmail.com

Key words
QSAR, 7-substituted 4-aminoquinolines, antimalarial agents, MLR, PLSR, PCR

Abstract
The emergence and rapid spread of chloroquine resistant strains of Plasmodium falciparum has dramatically reduced the chemotherapeutic options. Towards this goal, the quantitative structure-activity relationship analyses of some synthesized 7-substituted 4-aminoquinolines were performed for their antiplasmodial activity against chloroquine-resistant parasites to find out the structural features responsible for the biological activity. The statistically significant best 2D QSAR model having correlation coefficient (r^2) = 0.8631 and cross validated squared correlation coefficient (q^2) = 0.8101 with external predictive ability (pred_r^2) = 0.6740 was developed by Partial Least Square Regression coupled with stepwise backward-forward method using Vlife MDS 3.5 software and showed that the parameters Average-ve potential, T_2_N_5, XcompDipole and QMDipoleX were highly correlated with antiplasmodial activity of 7-substituted 4-aminoquinolines. The all developed models are interpretable, with good statistical and predictive significance, and can be used for guiding ligand modification for the development of potential new antimalarial agents.

J Chronother Drug Deliv | Vol 7| Issue 2| 2016: 51-62

INTRODUCTION

Malaria is a devastating disease caused by four species of genus plasmodium which afflicts more than 40% of the world population, causing an estimated mortality of 1.5–2.7 million people annually.[1] During the past six decades, chloroquine (CQ) and other aminoquinolines have been the frontline antimalarial agents because of their therapeutic efficacy and lower cost. However, development of resistance has severely limited the choice of available antimalarial drugs,[2] which clearly highlights the urgent need of novel chemotherapeutic agents for the treatment of malaria.

A major initiative in this direction is to find enzyme targets that are critical to the disease process or essential for the survival of the parasite. Identification and design of novel chemical entities specifically affecting these targets could lead to better drugs for the treatment of malaria.[3] Among old and new drug targets of malaria, host heme molecule remains one of the most attractive target and 7-chloroquinoline compounds are very selective towards heme bindings.[4] So, rather than identifying new molecules for efficacy, modified 7-chloroquinolines having many advantages and efficiency are now in priority for antimalarial chemotherapy.

Quantitative structure activity relationship (QSAR) is one of the major tools in drug discovery to explore ligand–receptor/enzyme interactions, especially when the structural details of the target are not known or if there are multiple targets. The quantitative structure-activity relationship (QSAR) approach helps to correlate the specific biological activities or physical properties of a series of compounds with the measured or computed molecular properties of the compounds, in terms of descriptors.[5] QSAR methodologies save resources and expedite the process of the development of new molecules and drugs. There have been many QSAR researches related to design of anti-malarial drugs so far[6-10] but a systematic QSAR study is yet to be carried out for series of 7-substituted, 4-aminoquinolines.

The present study aimed to elucidate the structural features of 7-substituted, 4-aminoquinolines required for antimalarial activity and to obtain predictive QSAR models, which may guide the rational synthesis of novel plasmodium inhibitors.

EXPERIMENTAL

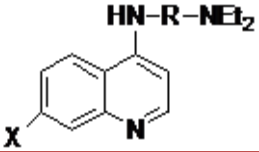
All molecular modeling studies were performed using the Molecular Design Suite (VLife MDS software package, version 3; supplied by VLife Sciences, Pune, India) on a Compaq PC with a Pentium IV processor and a Windows XP operating system. Structures of all the 40 compounds were sketched using the 2D draw application and converted to 3D structures. Energy minimization and geometry optimization were conducted using the Merck molecular force field (MMFF) method with the root mean square (RMS) gradient set to 0.01 kcal/mol Å and the iteration limit to 10,000.[11]

Data set

A number of 7-substituted, 4-aminoquinoline derivatives having antiplasmodial activity[12] were considered in the present QSAR study (Table 1). The biological activity values [IC_{50} (nM)] reported in literature were converted to their molar units and then further to negative logarithmic scale (pIC_{50}) and subsequently used as the dependent variable for the QSAR analysis.

The energy-minimized geometry was used for the calculation of the various 2D descriptors (Individual, Chi, Chiv, Path Count, Chi Chain, Chiv Chain, Chain Path Count, Cluster, Path Cluster, Kappa, Element Count, Estate numbers, Estate contributions, Semi Empirical, Hydophillic-Hydophobic and Polar Surface Area). The various alignment-independent (AI) descriptors were also calculated. For calculation of alignment, the independent descriptor was assigned the utmost three attributes. The first attribute was T to characterize the topology of the molecule. The second attribute was the atom type, and the third attribute was assigned to atoms taking part in the double or triple bond. The preprocessing of the independent variables (i.e., 2D descriptors) was done by removing invariable (constant column), which resulted in total 280 descriptors to be used for QSAR analysis.

Table 1. Structures of dataset used for QSAR analysis with corresponding observed antiplasmodial activities against CQ-resistant *P. falciparum* strain (Indochina I) of 7-Substituted 4-Aminoquinolines



S. No.	Compounds	X	-R-	IC_{50} (nM)	pIC_{50}
1	AQ-21	Cl	$(CH_2)_2$	5	8.301
2	AQ-13	Cl	$(CH_2)_3$	6	8.222
3	AQ-34	Cl	$CHMeCH_2$	10	8
4	AQ-26	Cl	$(CH_2)_4$	51	7.292
5	AQ-35	Cl	$(CH2)_5$	58	7.237
6	CQ	Cl	$CHMe(CH_2)_3$	95	7.022
7	AQ-33	Cl	$(CH_2)_6$	56	7.252
8	AQ-36	Cl	$(CH_2)_8$	41	7.387
9	AQ-41	Cl	$(CH_2)_{10}$	13	7.886
10	AQ-40	Cl	$(CH_2)_{12}$	11	7.959
11	AQ-44	I	CH_2CH_2	6	8.222
12	AQ-45	I	$CH_2CH_2CH_3$	4	8.398
13	AQ-46	I	CH_2MeCH_2	7	8.155
14	AQ-43	I	$CHMe(CH_2)_3$	35	7.456
15	AQ-57	I	$(CH_2)_{10}$	16	7.796
16	AQ-58	I	$(CH_2)_{12}$	20	7.699
17	AQ-22	Br	CH_2CH_2	7	8.155
18	AQ-37	Br	$CH_2CH_2CH_3$	6	8.222
19	AQ-47	Br	CH_2MeCH_2	12	7.921
20	AQ-24	Br	$CHMe(CH_2)_3$	90	7.046
21	AQ-56	Br	$(CH_2)_{10}$	17	7.77
22	AQ-55	Br	(CH2)12	18	7.745
23	AQ-25	F	CH_2CH_2	60	7.222
24	AQ-38	F	$CH_2CH_2CH_3$	120	6.921
25	AQ-42	F	CH_2MeCH_2	60	7.222
26	AQ-17	F	$CHMe(CH_2)_3$	500	6.301
27	AQ-61	F	$(CH_2)_{10}$	40	7.398
28	AQ-65	F	$(CH_2)_{12}$	25	7.602
29	AQ-39	CF3	CH_2CH_2	41	7.387
30	AQ-49	CF3	$CH_2CH_2CH_3$	18	7.745
31	AQ-59	CF3	CH_2MeCH_2	50	7.301
32	AQ-27	CF3	$CHMe(CH_2)_3$	45	7.347
33	AQ-54	CF3	$(CH_2)_{10}$	20	7.699
34	AQ-53	CF3	$(CH_2)_{12}$	25	7.602
35	AQ-23	OMe	CH_2CH_2	155	6.81
36	AQ-66	OMe	$CH_2CH_2CH_3$	900	6.046
37	AQ-67	OMe	CH_2MeCH_2	180	6.745
38	AQ-48	OMe	$CHMe(CH_2)_3$	3000	5.523
39	AQ-60	OMe	$(CH_2)_{10}$	179	6.77
40	AQ-68	OMe	$(CH_2)_{12}$	90	7.046

The sphere exclusion (SE) method[13-14] was adopted for division of training and test data set comprising of 30 and 10 molecules, respectively, with dissimilarity value of 2.0 where the dissimilarity value gives the sphere exclusion radius. The spherical exclusion method employs the following algorithm: (i) select a point

and include it in the training set; (ii) build a sphere with radius R with a center in this point; (iii) include all points within the sphere, except for the center, in the test set; (iv) discard all points in the sphere from the initial set; (v) if no points are left, stop, otherwise go to step (i). Ten compounds, namely, 24, 26, 36, 44, 47, 49, 53, 57, 60 and 66 were used as test set while the remaining molecules as the training set (Table 2). The unicolumn statistics of the training and test sets are reported in Table 3.

Table 2. Selected molecular descriptors used in 2D-QSAR analyses with values

Compound	Xcomp Dipole	Average-vePotential	QM DipoleX	T_2_N_5	Most-ve Potential	T_C_N_7
AQ-13	1.342	-0.022	0.826	2	-0.094	4
AQ-17	1.756	-0.028	-0.8	1	-0.095	5
AQ-21	1.941	-0.022	-1.08	3	-0.087	2
AQ-22	1.065	-0.022	0.635	3	-0.088	2
AQ-23	1.453	-0.03	-1.14	3	-0.11	3
AQ-24*	1.11	-0.022	1.259	1	-0.095	5
AQ-25	2.512	-0.027	-1.63	3	-0.087	2
AQ-26*	1.478	-0.023	-1.21	1	-0.092	5
AQ-27	2.019	-0.025	1.525	1	-0.099	5
AQ-33	0.568	-0.021	-0.48	1	-0.1	1
AQ-34	2.175	-0.025	-1.92	3	-0.092	2
AQ-35	1.037	-0.02	0.607	1	-0.098	4
AQ-36*	1.469	-0.02	-1.03	1	-0.098	3
AQ-37	1.272	-0.022	1.015	2	-0.097	4
AQ-38	2.1	-0.027	1.111	2	-0.092	4
AQ-39	2.133	-0.028	-1.67	3	-0.086	2
AQ-40	1.436	-0.018	-1.24	1	-0.096	3
AQ-41	1.421	-0.021	-1.18	1	-0.098	3
AQ-42	1.921	-0.026	-1.54	3	-0.088	2
AQ-43	1.851	-0.023	1.485	1	-0.092	5
AQ-44*	1.067	-0.024	-1.24	3	-0.086	2
AQ-45	1.226	-0.022	0.941	2	-0.097	4
AQ-46	0.98	-0.022	-1.24	3	-0.09	2
AQ-47*	1.062	-0.021	-1.29	3	-0.09	2
AQ-48	0.527	-0.03	0.473	1	-0.108	6
AQ-49*	2.308	-0.025	1.61	2	-0.092	4
AQ-53*	3.078	-0.024	2.009	1	-0.098	3
AQ-54	2.994	-0.025	1.94	1	-0.095	3
AQ-55	-0.43	-0.018	0.073	1	-0.098	3
AQ-56	0.298	-0.017	-0.92	1	-0.098	3
AQ-57*	0.116	-0.018	-0.33	1	-0.095	3
AQ-58	0.968	-0.017	0.644	1	-0.1	3
AQ-59	3.246	-0.028	-2.5	3	-0.082	2
AQ-60*	0.39	-0.024	0.237	1	-0.108	4
AQ-61	0.979	-0.021	0.586	1	-0.1	3
AQ-65	1.893	-0.023	1.515	1	-0.095	3
AQ-66*	0.922	-0.027	0.174	2	-0.104	5
AQ-67	0.948	-0.031	-0.96	3	-0.108	3
AQ-68	1.522	-0.022	-1.01	1	-0.107	4
CQ	1.837	-0.022	-1.3	1	-0.093	5

* Indicates the compounds considered in the test set

Table 3. Unicolumn statistics of the training and test sets for QSAR models

Data Set	Average	Max	Min	SD	Total
Training	7.4667	8.3979	5.5229	0.6347	223.9996
Test	7.3826	8.2218	6.0458	0.6350	73.8260

Max maximum, *Min* minimum, *SD* standard deviation

Statistical Computation

All the calculated descriptors were considered as independent variable and biological activity as dependent variable. VLife Molecular Design Suite (VLifeMDS software was used to generate QSAR models by Multiple Linear Regression (MLR), Partial Least Squares Regression (PLSR) and Principal Components Regression (PCR) method analysis.

In the selected 2D QSAR equations, the cross-correlation limit was set at 0.5, the number of variables at 10, and the term selection criteria at r^2. An F value was specified to evaluate the significance of a variable. The variance cutoff was set at 0, with autoscaling in which the number of random iterations was set at 10.

The developed QSAR models were evaluated using the following statistical measures: r^2 (the squared correlation coefficient), *F* test (Fischer's value) for statistical significance, q^2 (cross-validated correlation coefficient); pred_r^2, r^2 for external test set. The regression coefficient r^2 is a relative measure of fit by the regression equation. It represents the part of the variation in the observed data that is explained by the regression. However, a QSAR model is considered to be predictive, if the following conditions are satisfied: $r^2 > 0.6$, $q^2 > 0.6$ and pred_$r^2 > 0.5$[15]. The *F*-test reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the *F*-test indicate that the model is statistically significant. The low standard error of r^2 (r^2_se), q^2 (q^2_se) and pred_r^2 (pred_r^2se) shows absolute quality of fitness of the model.

Internal validation was carried out using 'leave-one-out' (q^2, LOO) method.[16] The cross-validated coefficient, q^2, was calculated using the following equation:

$$q^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{\text{mean}})^2}$$

Where y_i, and $\acute{y}_i$ are the actual and predicted activity of the ith molecule in the training set, respectively, and y_{mean} is the average activity of all molecules in the training set.

However, a high q^2 value does not necessarily give a suitable representation of the real predictive power of the model for antimalarial ligands. So, an external validation was also carried out in the present study. The external predictive power of the model was assessed by predicting pIC_{50} value of the 10 test set molecules, which were not included in the QSAR model development. The predictive ability of the selected model was also confirmed by pred_r^2.

$$\text{pred}\_r^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{\text{mean}})^2}$$

Where y_i, and $\acute{y}_i$ are the actual and predicted activity of the i^{th} molecule in the test set, respectively, and y_{mean} is the average activity of all molecules in the training set.

RESULTS AND DISCUSSION

The 2D QSAR study of 40 compounds (divided into 10 test and 30 training) for antiplasmodial activity through MLR, PLSR, and PCR analysis using VLife MDS resulted in the following statistically significant model, considering the term selection criterion as r^2, q^2 and pred_r^2. The training and test sets were selected by sphere exclusion method and the models were validated by both internal and external validation procedure. To ensure a fair comparison, the same training and test sets were used for each model's

development (Table 2). A Uni-Column statistics for training set and test set were generated to check correctness of selection criteria for training and test set molecules (Table 3).

The maximum and minimum value in training and set were compared in a way that:

1. The maximum value of pIC_{50} of test set should be less than or equal to maximum value of pIC_{50} of training set.
2. The minimum value of pIC_{50} of test set should be higher than or equal to minimum value of pIC_{50} of training set.

This observation showed that test set was interpolative and derived within the minimum–maximum range of training set. The mean and standard deviation of pIC_{50} values of sets of training and test provide insights to relative difference of mean and point density distribution of two sets. The mean of the test sets were higher than the train sets which indicates the presence of relatively more active molecules as compared to the inactive ones.

Interpretation of 2D-QSAR models

The best QSAR model was selected on the basis of value of statistical parameters like r^2 (square of correlation coefficient for training set of compounds), q^2 (cross-validated r^2), and pred_r^2 (r^2 for the test set of compounds). All 2D-QSAR models were validated and tested for its predictability using an external test set of ten compounds. The high value of LOO q^2 appears to be the necessary but not the sufficient condition for the model to have a high predictive power. So, the external validation is the only way to establish a reliable QSAR model. The results obtained for observed and predicted activity are presented in Table 4. In this section, the prediction performances of the methods proposed by (SW-PLSR, SW-MLR and SW-PCR) were evaluated. The regression equations obtained for the series of compounds are given below.

Model 1 (SW-PLSR)

pIC_{50} = 180.0310 (Average-vePotential) + 0.5004 (T_2_N_5) + 0.2723 (XcompDipole) + 0.1412 (QMDipoleX) + 10.4043

n = 30, r^2= 0.8631, q^2= 0.8101, *F*-test= 54.6582, r^2_se = 0.2480, q^2_se = 0.2921, pred_r^2 = 0.6740, pred_r^2se = 0.3661, ZScore R^2 = 7.66639 , ZScore Q^2 = 7.11067 , Best Rand R^2 = 0.54081, Best Rand Q^2 = 0.26864, Alpha Rand R^2 = 0.00000, Alpha Rand Q^2 = 0.00000, Z Score Pred R^2 = 2.76660, best Rand Pred R^2 = 0.65234, alpha Rand Pred R^2 = 0.01000

2D-QSAR Model 1 is among the best one obtained by the SW-PLS which shows positive contribution of Average-vePotential, T_2_N_5, XcompDipole and QMDipoleX and no negative correlation with any selected descriptors. Equation shows good squared correlation coefficient (r^2) of 0.8631, explains 86 % variance in biological activity, significant cross-validated correlation coefficient (q^2) of 0.8101. Another parameter for predictivity of test set compound is high pred_r^2 = 0.6740 and low pred_r^2se= 0.3661, which is showing good external predictive power of the model. It is apparent from model 1 that the descriptor Average-vePotential plays a pivotal role in determining activity. This descriptor signifies the average of the total -ve electrostatic potential on van der Waals surface area of the molecule. This is a positively contributing descriptor toward antiplasmodial activity as antimalarial and its contribution is approx 45.05%. The positive coefficient of Average-vePotential (45.05%) showed that increase in the values of this descriptor is beneficial for the antiplasmodial activity against malaria (like in compounds AQ-13, AQ-21, AQ-37, AQ-40, AQ-41, AQ-44, AQ-45, AQ-47, AQ-55, AQ-56, and AQ-57) and decrease in the values of this descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-17, AQ-23, AQ-38, AQ-48, AQ-66, and AQ-67).

The next important descriptor which influences the activity is T_2_N_5 that signifies the count of number of double bounded atoms (i.e. any double bonded atom, T_2) separated from Nitrogen atom by 5 bonds. This descriptor is also directly proportional to the activity (29.98%). The positive coefficient of T_2_N_5 (29.98%) showed that increase in the values of this descriptor is beneficial for the antiplasmodial activity

Table 4. Comparative observed and predicted activities (LOO) of 7-substituted 4-aminoquinolines by QSAR models

Compound	Observed Activity (pIC_{50})	Predicted Activity (pIC_{50})		
		MLR	PLSR	PCR
AQ-13	8.222	8.014	7.999	7.548
AQ-17	6.301	6.25	6.272	6.772
AQ-21	8.301	8.363	8.359	8.013
AQ-22	8.155	8.389	8.346	7.969
AQ-23	6.81	6.763	6.733	6.578
AQ-24	7.046	7.399	7.394	7.307
AQ-25	7.222	7.448	7.456	7.477
AQ-26	7.292	6.883	6.908	7.235
AQ-27	7.347	7.099	7.113	6.925
AQ-33	7.252	7.295	7.293	8.034
AQ-34	8	7.658	7.662	7.565
AQ-35	7.237	7.71	7.713	7.638
AQ-36	7.387	7.465	7.493	7.741
AQ-37	8.222	8.006	7.987	7.474
AQ-38	6.921	7.223	7.218	7.03
AQ-39	7.387	7.291	7.288	7.473
AQ-40	7.959	7.869	7.902	8.019
AQ-41	7.886	7.387	7.415	7.721
AQ-42	7.222	7.567	7.56	7.595
AQ-43	7.456	7.527	7.541	7.317
AQ-44	8.222	7.805	7.775	7.847
AQ-45	8.398	7.999	7.98	7.479
AQ-46	8.155	8.079	8.05	7.926
AQ-47	7.921	8.193	8.167	7.979
AQ-48	5.523	5.698	5.671	6.098
AQ-49	7.745	7.683	7.683	7.218
AQ-53	7.602	7.688	7.729	7.417
AQ-54	7.699	7.539	7.577	7.414
AQ-55	7.745	7.676	7.646	8.025
AQ-56	7.77	7.725	7.724	8.029
AQ-57	7.796	7.73	7.718	8.067
AQ-58	7.699	8.106	8.11	7.986
AQ-59	7.301	7.447	7.482	7.561
AQ-60	6.77	6.761	6.743	7.049
AQ-61	7.398	7.473	7.472	7.657
AQ-65	7.602	7.538	7.552	7.592
AQ-66	6.046	6.875	6.845	6.674
AQ-67	6.745	6.553	6.506	6.559
AQ-68	7.046	7.135	7.161	7.207
CQ	7.022	7.173	7.21	7.317

against malaria (like in compounds AQ-13, AQ-21, AQ-22, AQ-34, AQ-37, AQ-44, AQ-45, AQ-46, and AQ-47) and decrease in the values of this descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-17, AQ-38, AQ-48, AQ-60, and AQ-66). As a positive contributing descriptor, XcompDipole (this descriptor signifies the x component of the dipole moment (external coordinates)) is also a physicochemical descriptor influencing activity variation and is directly proportional to activity. The positive coefficient of XcompDipole (13.63%) showed that increase in the values of this descriptor is beneficial for the antiplasmodial activity against malaria (like in compounds AQ-13, AQ-21, AQ-22, AQ-34, AQ-37, AQ-40, AQ-41, AQ-44, AQ-45, and AQ-47) and decrease in the values of this

descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-48, AQ-60, AQ-66, and AQ-67). The last descriptor is QMDipoleX which is an individual descriptor that is induced dipole moment along X-axis and positively contributes to the biological activity (11.33%). The positive coefficient of QMDipoleX showed that increase in the values of this descriptor is beneficial for the antiplasmodial activity against malaria (like in compounds AQ-13, AQ-22, AQ-37, AQ-45, and AQ-55) and decrease in the values of this descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-17, AQ-23, and AQ-67). The descriptors selected for modeling inhibitory activity of the 7-substituted 4-aminoquinolines derivatives are summarized in Table 2. The correlation matrix between the physico-chemical parameters and the biological activity is presented in Table 5. The plots of observed activity *vs* predicted activity values of pIC_{50} and contribution chart are shown in Fig. 1(a) and Fig. 1(b).

Table 5. Correlation matrix for molecular (physicochemical and alignment independent) descriptors influencing the antimalarial activity against CQ-resistant *P. falciparum* strain (Indochina I) (2D-QSAR model 1)

Parameter	pIC_{50}	Xcomp Dipole	Average-vePotential	QMDipoleX	T_2_N_5
pIC_{50}	1				
XcompDipole	0.059	1			
Average-vePotential	0.598	-0.45	1		
QMDipoleX	0.018	0.019	0.113	1	
T_2_N_5	0.23	0.241	-0.448	-0.468	1

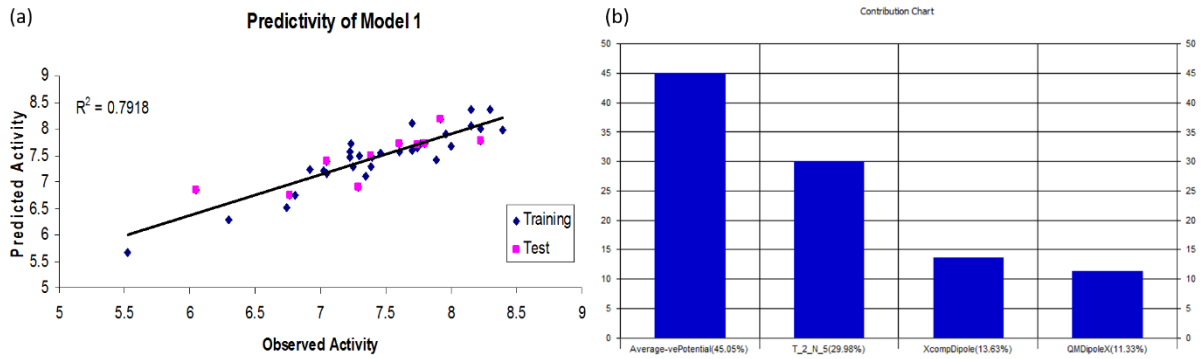


Fig 1. (a) Fitness plot of observed activity (pIC50) versus predicted activity (pIC50) for training set and test set compounds according to 2D-QSAR PLSR model 1
(b) Contribution of descriptors for biological activity developed using PLSR equation

Model 2 (SW-MLR)

pIC_{50} = 178.4470 (Average-vePotential) + 0.5219 (T_2_N_5) + 0.2444 (XcompDipole) + 0.1497 (QMDipoleX) + 10.3724

n = 30, r^2= 0.8645, q^2= 0.8142, *F*-test= 39.8863, r^2_se = 0.2516, q^2_se = 0.2947, pred_r^2 = 0.6626, pred_r^2se = 0.3724, ZScore R^2 = 8.05831, ZScore Q^2 = 7.40437, Best Rand R^2 = 0.38220, Best Rand Q^2 = 0.14349, Alpha Rand R^2 = 0.00000, Alpha Rand Q^2 = 0.00000, Z Score Pred R^2 = 2.79105, best Rand Pred R^2 = 0.41644, alpha Rand Pred R^2 = 0.01000

2D-QSAR Model 2 is the second best model obtained by the SW-MLR which shows positive contribution of all selected descriptors as Average-vePotential, T_2_N_5, XcompDipole and QMDipoleX. Equation shows good squared correlation coefficient (r^2) of 0.8645, explains 86% variance in biological activity, significant cross-validated correlation coefficient (q^2) of 0.8142. Another parameter for predictivity of test set compound is high pred_r^2 = 0.6626 and low pred_r^2se= 0.3724, which is showing good external

predictive power of the model. It is apparent from model 2 that the descriptor Average-vePotential plays a pivotal role in determining activity. This descriptor signifies the average of the total -ve electrostatic potential on van der Waals surface area of the molecule. This is a positively contributing descriptor toward antiplasmodial activity as antimalarial and its contribution is approx 44.58%. The positive coefficient of Average-vePotential (44.58%) showed that increase in the values of this descriptor is beneficial for the antiplasmodial activity against malaria (like in compounds AQ-21, AQ-13, AQ-44, AQ-45, AQ-37, AQ-41, AQ-40, AQ-57, AQ-47, AQ-56 and AQ-55) and decrease in the values of this descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-38, AQ-17, AQ-23, AQ-66, AQ-67 and AQ-48). The next important descriptor which influences the activity is T_2_N_5 that signifies the count of number of double bounded atoms (i.e. any double bonded atom, T_2) separated from Nitrogen atom by 5 bonds. This descriptor is also directly proportional to the activity (31.21%). The positive coefficient of T_2_N_5 (31.21%) showed that increase in the values of this descriptor is beneficial for the antiplasmodial activity against malaria (like in compounds AQ-21, AQ-34, AQ-44, AQ-46, AQ-22, and AQ-47) and decrease in the values of this descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-38, AQ-17, AQ-66, AQ-48 and AQ-60). As a positive contributing descriptor, XcompDipole (this descriptor signifies the x component of the dipole moment (external coordinates)) is also a physicochemical descriptor influencing activity variation and is directly proportional to activity. The positive coefficient of XcompDipole (12.22%) showed that increase in the values of this descriptor is beneficial for the antiplasmodial activity against malaria (like in compounds AQ-21, AQ-13, AQ-34, AQ-45, AQ-37, AQ-41, and AQ-40) and decrease in the values of this descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-66, AQ-67, AQ-48 and AQ-60). The last descriptor is QMDipoleX which is an individual descriptor that is induced dipole moment along X-axis and positively contributes to the biological activity (11.99%). The positive coefficient of QMDipoleX showed that increase in the values of this descriptor is beneficial for the antiplasmodial activity against malaria (like in compounds AQ-13, AQ-45, AQ-22, AQ-37 and AQ-55) and decrease in the values of this descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-17, AQ-23, and AQ-67). The descriptors selected for this modeling inhibitory activity of the 7-substituted 4-aminoquinolines derivatives are summarized in Table 2. The plots of observed activity *vs* predicted activity values of pIC_{50} and contribution chart for QSAR model 2 are shown in Fig. 2(a) and Fig. 2(b).

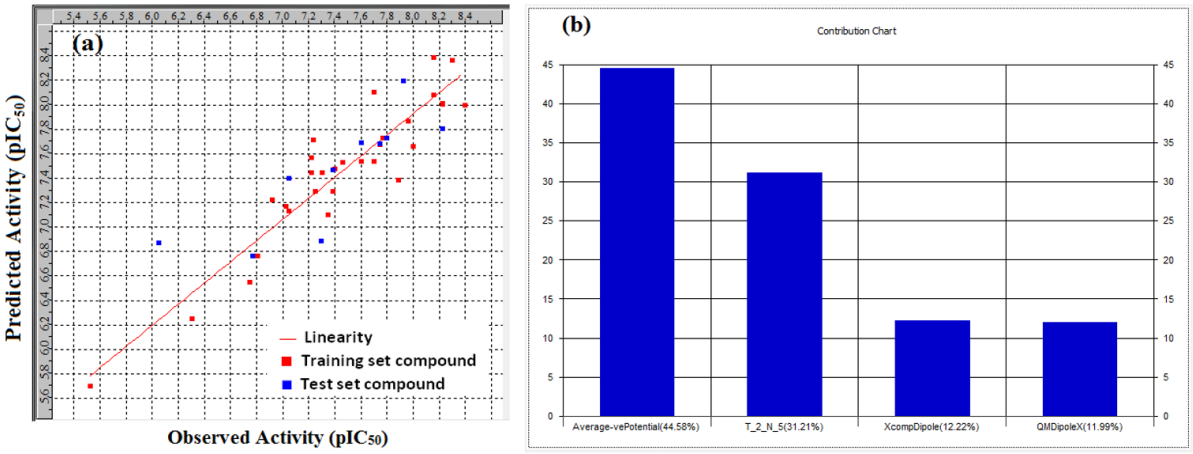


Fig 2. (a) Fitness plot of observed activity (pIC50) versus predicted activity (pIC50) for training set and test set compounds according to 2D-QSAR MLR model 2
(b) Contribution of descriptors for biological activity developed using MLR equation

To improve the external predictivity of the model, PCR analysis with the same data set was performed, which resulted in r^2 of 0.6027 and an internal predictive power of 52% (q^2 = 0.5203), with the good external predictivity of 67.6% (pred_r^2 = 0.6760).

Model 3 (SW-PCR)

pIC_{50} = 95.3804 (Average-vePotential) + 20.0116 (Most-vePotential) - 0.1705 (T_C_N_7) + 12.1643

n = 30, r^2= 0.6027, q^2= 0.5203, *F*-test= 20.4820, r^2_se = 0.4146, q^2_se = 0.4556, pred_r^2 = 0.6760, pred_r^2se = 0.3649, ZScore R^2 = 8.13422, ZScore Q^2 = 7.51310, Best Rand R^2 = 0.32855, Best Rand Q^2 = 0.18145, Alpha Rand R^2 = 0.00000, Alpha Rand Q^2 = 0.00000, Z Score Pred R^2 = 1.99180, best Rand Pred R^2 = 0.56429, alpha Rand Pred R^2 = 0.05000

2D-QSAR Model 3 is another model obtained by the SW-PCR which shows positive contribution of two descriptors as Average-vePotential, Most-vePotential and negative contribution of T_C_N_7. Equation shows good squared correlation coefficient (r^2) of 0.6027, explains 60 % variance in biological activity, significant cross-validated correlation coefficient (q^2) of 0.5203. Another parameter for predictivity of test set compound is high pred_r^2 = 0.6760 and low pred_r^2se = 0.3649, which is showing good external predictive power of the model. It is apparent from model 3 that the descriptor Average-vePotential plays a pivotal role in determining activity. This descriptor signifies the average of the total -ve electrostatic potential on van der Waals surface area of the molecule. This is a positively contributing descriptor toward antiplasmodial activity as antimalarial and its contribution is approx 51.89%. The positive coefficient of Average-vePotential (51.89%) showed that increase in the values of this descriptor is beneficial for the antiplasmodial activity against malaria (like in compounds AQ-21, AQ-13, AQ-44, AQ-45, AQ-37, AQ-41, AQ-40, AQ-57, AQ-47, AQ-56 and AQ-55) and decrease in the values of this descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-38, AQ-17, AQ-23, AQ-66, AQ-67 and AQ-48). The next important descriptor which influences the activity is Most-vePotential that signifies the highest value of -ve electrostatic potential on van der Waals surface area of the molecule. This descriptor is also directly proportional to the activity (19.43%). The positive coefficient of Most-vePotential (19.43%) showed that increase in the values of this descriptor is beneficial for the antiplasmodial activity against malaria (like in compounds AQ-21, AQ-13, AQ-34, AQ-44, AQ-46, AQ-22, AQ-57, and AQ-47) and decrease in the values of this descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-23, AQ-66, AQ-67, AQ-48 and AQ-60). The last important descriptor which influences the activity is T_C_N_7 that signifies the count of number of Carbon atoms (single double or triple bonded) separated from any other Nitrogen atom (single double or triple bonded) by 7 bonds in a molecule. This descriptor is indirectly proportional to the activity (-28.68%). The negative coefficient of T_C_N_7 (-28.68%) showed that decrease in the values of this descriptor is beneficial for the antiplasmodial activity against malaria (like in compounds AQ-21, AQ-34, AQ-44, AQ-46, AQ-22, AQ-41, AQ-40, AQ-57, AQ-47, AQ-56 and AQ-55) and increase in the values of this descriptor is detrimental for the antiplasmodial activity against malaria (like in compounds AQ-38, AQ-17, AQ-66, AQ-48 and AQ-60). The descriptors selected for this modeling inhibitory activity of the 7-substituted 4-aminoquinolines derivatives are summarized in Table 2. The plots of observed activity *vs* predicted activity values of pIC_{50} and contribution chart for QSAR model 3 are shown in Fig. 3(a) and Fig. 3(b).

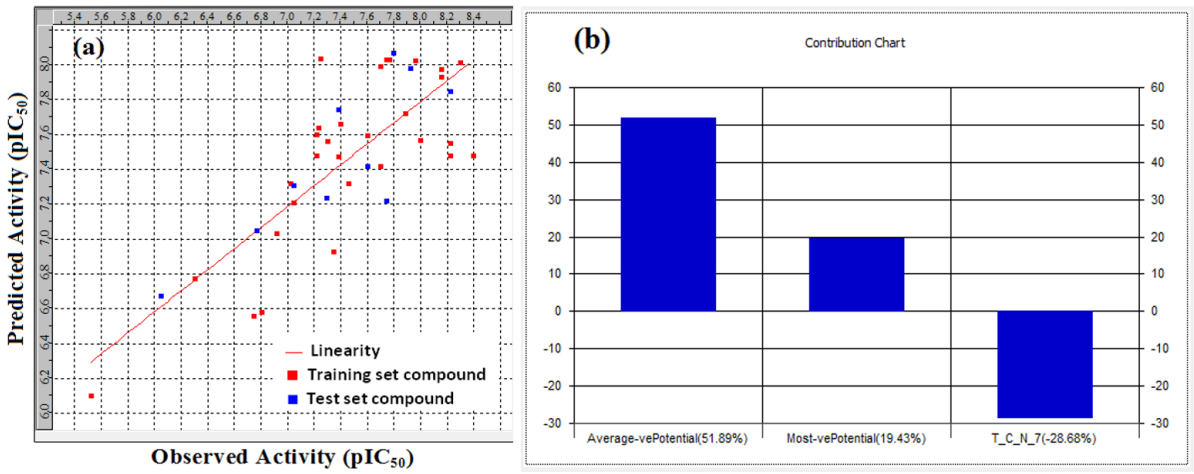


Fig 3. (a) Fitness plot of observed activity (pIC50) versus predicted activity (pIC50) for training set and test set compounds according to 2D-QSAR PCR model 3.
(b) Contribution of descriptors for biological activity developed using PCR equation.

CONCLUSION

The present study elucidates key structural requirements for the antiplasmodial activity against malaria utilizing 2DQSAR models. This work reveals how the antiplasmodial activities against malaria of various 7-substituted, 4-aminoquinoline derivatives may be treated statistically to uncover the molecular characteristics which are essential for high activity. The 2D-QSAR studies have been carried out on a series of 7-substituted, 4-aminoquinoline derivatives for the antiplasmodial activity against malaria. All the models are statistically significant. The 2D-QSAR model obtained by SW-PLSR method indicates that increase in Average-vePotential, T_2_N_5, XcompDipole and QMDipoleX of 7-substituted, 4-aminoquinoline derivatives leads to increases in potency of molecules (pIC50). A combination of the all above models is useful in understanding the structural requirements for design of novel, potent and selective antimalarial agents. Reliability of the models was confirmed by several statistical analyses and thus, the proposed models, due to the good predictive ability, offer a useful alternative for determining antiplasmodial activity against malaria of newly designed molecules. These efforts will guide synthetic medicinal chemists to design and synthesize new compounds with an increased biological activity in comparison to the reported compounds.

ACKNOWLEDGMENTS

The authors thank VLife Science Technologies Pvt. Ltd, Pune India for providing demo software. The authors also thank to Dean and Principal, RKDF College of Pharmacy, Bhopal, India for providing support.

DECLARATION OF INTEREST

It is hereby declared that this paper does not have any conflict of interest.

REFERENCES

1. Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI. The global distribution of clinical episodes of Plasmodium falciparum malaria. Nature 2005; 434: 214.
2. Ridley RG. Medical need, scientific opportunity and the drive for antimalarial drugs. Nature 2002; 415: 686-693.
3. Sahu NK, Sahu S. Kohli DV. Novel molecular targets for antimalarial drug development. Chem Biol Drug Des. 2008; 71: 287-297.
4. Vippagunta SR, Dorn A, Matile H, Bhattacharjee AK, Karle JM, Ellis WY, Ridely RG, Vennerstrom JL. Structural specificity of chloroquine-hematin binding related to inhibition of hematin polymerization and parasite growth. J Med Chem. 1999; 42: 4630- 4639.
5. Hansch C. Kurup A. Garg R. Gao H. Chem-bioinformatics and QSAR: a review of QSAR lacking positive hydrophobic terms. Chem. Rev. 2001, 101, 619-72.
6. Bhattacharjee AK, Geyer JA, Woodard CL, Kathcart AK, Nichols DA, Prigge ST, Li Z, Mott BT, Waters NC. A Three-Dimensional in Silico Pharmacophore Model for Inhibition of Plasmodium falciparum Cyclin-Dependent Kinases and Discovery of Different Classes of Novel Pfmrk Specific Inhibitors. J Med Chem. 2004; 47(22): 5418-5426.
7. Dascombe MJ, Drew MG, Morris H, Wilairat P, Auparakkitanon S, Moule WA, Alizadeh-Shekalgourabi S, Evans PG, Lloyd M, Dyas AM, Carr P, Ismail FM. Mapping Antimalarial Pharmacophores as a Useful Tool for the Rapid Discovery of Drugs Effective in Vivo: Design, Construction, Characterization, and Pharmacology of Metaquine. J Med Chem. 2005; 48(17): 5423-5436.
8. Katritzky AR, Kulshyn OV, Stoyanova-Slavova I, Dobchev DA, Kuanar M, Fara DC, Karelson M. 2006. Bioorg. Med. Chem. 14(7): 2333-2357.
9. Adane L, Bharatam PV. Modelling and informatics in the analysis of P. falciparum DHFR enzyme inhibitors. Curr Med Chem. 2008; 15(16): 1552-1569.
10. Deshpande S, Solomon VR, Katti SB, Prabhakar YS. Topological descriptors in modelling antimalarial activity: N1-(7-chloro-4-quinolyl)-1,4-bis(3-aminopropyl)piperazine as prototype. J. Enzym Inhib Med Chem. 2009; 24(1): 94-104.
11. Halgren TA. Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. J Comput Chem. 1996; 17: 553–586.
12. De D, Krogstad FM, Byers LD, Krogstad DJ. Structure-activity relationships for antiplasmodial activity among 7-substituted 4-aminoquinolines. J Med Chem. 1998; 41: 4918-4926.

13. Zheng W, Tropsha A. QSAR modeling using chirality descriptors derived from molecular topology. J Chem Inf Comput Sci. 2000; 40: 185-194.
14. Golbraikh A, Tropsha A. QSAR modeling using chirality descriptors derived from molecular topology. J. Chem. Inf. Comput. Sci. 2003; 43: 144 – 154.
15. Golbraikh A, Tropsha A. Beware of q2!. J Mol Graph Model. 2002; 20: 269-276.
16. Cramer RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA): Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc. 1988; 110: 5959-5967.

Received: July 10, 2016; Accepted: August 27, 2016