Parihar and Nandi *SpringerPlus (2015)4:819*
DOI 10.1186/s40064-015-1593-3

SpringerPlus
a SpringerOpen Journal

RESEARCH Open Access

In-silico combinatorial design and pharmacophore modeling of potent antimalarial 4-anilinoquinolines utilizing QSAR and computed descriptors

Neha Parihar and Sisir Nandi*

*Correspondence: sisir.iicb@gmail.com
Division of Pharmaceutical Chemistry, Global Institute of Pharmaceutical Education and Research, Affiliated to Uttarakhand Technical University, Kashipur 244713, India

Abstract

There are very few studies for combinatorial library design and high throughput screening of 4-anilinoquinoline antimalarial compounds having activities against parasitic strain of *P. falciparum*. Therefore, an attempt has been made in the present paper to design potent lead compounds in this congener utilizing quantitative structure activity relationship utilizing theoretical molecular descriptors. QSAR models for a series of 4-anilinoquinolines considering various theoretical molecular descriptors including topological, constitutional, geometrical, functional group and atom-centered fragments has been carried out by stepwise forward–backward variable selections assimilating multiple linear regression (MLR) methods showing the topological indices contribute maximum impact on parasitic *P. falciparum* strain. A combinatorial library of 2160 compounds has been generated and finally, 16 compounds were screened through high throughput screening as promising 4-anilinoquinoline antimalarial hits based on their predicted activities utilizing topological descriptor based validated QSAR model. Highly predicted active compounds were then undergone for pharmacophore modeling to predict mode of binding and to optimize leads having greater affinity towards malarial *P. falciparum* parasitic strain.

Keywords: 4-Anilinoquinolines, Combinatorial library generation, Virtual screening, QSAR, Pharmacophore, Topological indices

Background

Malaria is an Anopheles mosquito borne parasitic disease triggered by four species of genus plasmodium including *P. falciparum*, *P. vivax*, *P. ovale*, and *P. malariae*. Amongst these, *P. falciparum* is the most dangerous species because it can penetrate into deeper tissues and infect red blood corpuscles leading to its breakdown and rupture, forming sticky lump like mass structure in the blood capillary which may ground circulatory arrest such as cerebral attack causing death of the individual (Tham and Kennedy 2015). As per the updated reports approximately 3.4 billion cases of malaria occur every year and about 1.3 million deaths occurred in the year of 2013 worldwide (Seder 2014). Brutal death of more than 1 million people globally cries to develop new antimalarial chemotherapeutics. One of the promising antimalarial chemotherapeutics is 4-anilinoquinoline

derivatives including amodiaquine and piperaquine which act as blood schizontoside and haemazoin inhibitors. Due to drug resistance and lack of knowledge of exact mechanism of action of these series of compounds, it is really urgent to design and develop new congeneric leads utilizing structure activity-property relationship studies. Although the structure–property-activity relationships were developed since long years back (Crum-Brown and Fraser 1968), but now it is a multidisciplinary area of molecular design and are widely used for the prediction of properties, activities and/or toxicities of new chemicals by developing quantitative relationship between molecular activity or property (such as partition coefficient (log *P*), boiling point, melting point, acid and base constant, chromatographic retention index, toxicity, or reactivity) and computed structural properties such as constitutional, electrostatic, geometrical, topological, or quantum chemical molecular characteristics (Basak et al. 1997; Pompe and Novic 1999; Randic 1975; Roy et al. 2015a, b, c). Therefore in the present paper QSAR modelling has been carried out for antimalarial 4-anilinoquinolines based on the computed structure–property-activity correlations.

In this connection, a series of N^1,N^1-diethyl-*N*2-(4-quinolinyl)-1,2-ethanediamine derivatives having various groups substituted at the 7-position on the quinoline nucleus have been synthesized by Kaschula et al. (2002) who tested in vitro antimalarial activity of the same compounds against chloroquine sensitive D10 strain of *P. falciparum* showing that an electron attracting group at the 7th position bears with lower p*K*a of both the quinoline nitrogen atom and the tertiary anilino nitrogen in the alkyl side chain. O'Neill et al. (2003) synthesized a new series of amodiaquine analogues by interchanging hydroxyl at the 3′ position and the 4′ Mannich side-chain function of anilino moiety of quinoline which can produce non-toxic metabolite. Hwang et al. (2011) synthesized many 4-anilinoquinoline compounds introducing diaryl, ether, biaryl, and alkylaryl groups to the basic nucleus and tested their antimalarial activity against the chloroquine-sensitive strain 3D7 and the chloroquine-resistant K1 strain as well as for cytotoxicity against mammalian cell lines. In vitro screening and in vivo pharmacokinetic estimation of virtual libraries of newly designed chloroquine scaffold based 4-anilinonoquinolines showed highly potent antimalarial activity in mice found out two lead compounds utilizing ADMET predictions (Ray et al. 2010). Solomon et al. (2005, 2007) synthesized new 4-anilinoquinoline derivatives with evaluating its in vitro activity against chloroquine sensitive strain of *P. falciparum* strain and chloroquine resistant N-67 strain of *P. yoelii* in vivo whereas the same group generated another new series of 4-anilinoquinoline analogs which can form complex with hematin to act as hemazoin inhibitors showing affinity towards heme polymerization target.

To predict the biochemical mechanisms of 4-anilinoquinolines, quantitative structure–property-activity relationship studies were being executed recently by many researchers. Gupta et al. carried out QSAR on antimalarial activity and cytotoxicity of 4-anilinoquinoline using structural descriptors and identified that the antimalarial activity are being correlated with topological, 2D autocorrelation and functional group descriptors while cytotoxicity is being correlated with atom centered descriptors. This model suggests that the analogues with aromatic primary amines, aliphatic secondary amines are responsible for antimalarial activity and aromatic ethers, CH_2R_2 and CH_3X contributed to cytotoxicity. With another work the author developed topological

descriptor based QSAR model using electrons enrich species in aniline substituent showing better structure activity correlations quantitatively (Gupta et al. 2005; Gupta and Prabhakar 2006). Descriptors based QSAR modeling has been performed by many other authors and co-workers which are cited here (Masand et al. 2014; Sahu et al. 2014; Deshpande et al. 2009).

QSARs utilizing topological structural indices have been carried out but there is hardly any studies based on in silico virtual screening of combinatorial compounds and pharmacophore modeling of 4-anilinoquinoline compounds. One of the important techniques to focus mode of binding of the ligand is pharmacophore generation when the crystal structure of target is unknown. Delarue et al. synthesized a number of 4-anilinoquinolines having two proton accepting side chains and in vitro antimalarial activity has been evaluated on *P. falciparum* FcB1R strain whereas toxicity of the same compounds have been studied using MRC-5 cells and macrophages respectively (Delarue et al. 2001). A number of experimental and theoretical studies for the design of potent 4-anilinoquinolines have been performing. But experimental design of a single molecule involves a series of reactions and processes from the starting material of synthesis, structure elucidation and biological assays for activity studies. This total process consumes long years, enormous manpower, monetary issues and a number of animal sacrifices. So theoretical modeling utilizing QSAR based on topological indices computed solely from the structures of these compounds was carried out a lot. But there is scarcely any in silico design of 4-anilinoquinoline derivatives using pharmacophore modeling and virtual screening. In the present article, an attempt has been made to design thousands of combinatorial compounds at a time considering 4-anilinoquinoline scaffold with potential antimalarial activities. Such compounds are screened for potency and selectivity utilizing high-throughput screening techniques. HTS is based on lead optimization which incorporates Lipinski rule of five, QSAR and pharmacophore modeling. Such advances lead to greater understanding of new entity design having higher affinity towards the target.

Methodology

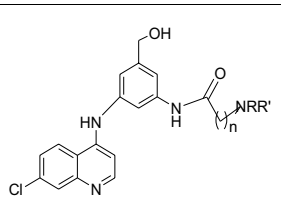
Experimental methods

Data base

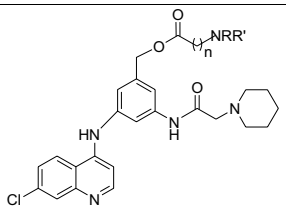
A number of 4-anilinoquinolines having antimalarial activities against *P. falciparum* have been synthesized by Delarue et al. (2001). Different protons accepting side chains were substituted at 3′ and 5′ positions of the amino moiety to produce potent compounds which are tested for in vitro antimalarial activities against the chloroquine resistant *P. falciparum* FcB1R strain. Table 1 contains structure and antimalarial activities in terms of IC_{50} of 62 congeneric 4-anilinoquinoline derivatives. These IC_{50} values were converted into their negative logarithms (pIC_{50}) which are taken into consideration in the present calculation as dependent variables whereas computed descriptors calculated by using optimized 3D-structure of 4-anilinoquinoline compounds are considered as independent variables for statistical multivariate regression modeling.

Molecular optimization is carried out by minimization of molecular surface energy. For this purpose, 2D structures of 4-anilinoquinolines, drawn in ChemDraw software, were converted into 3D modules incorporated into Chem3D Ultra (Mills 2006). The 3D

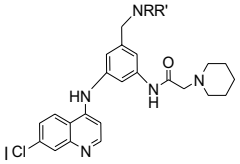
Table 1 Biological activity data



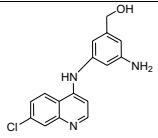
(2-19)



(20-32)



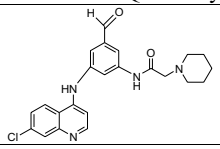
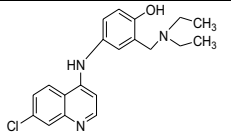
(3 3-61)

Comp. No.	Substituents		IC_{50} (μM)	pIC_{50} (μM)
	n	NRR′		
1	-		0.2789	0.554
2	1	4-methyl piperidine	0.0777	1.110
3*	2	4-methyl piperidine	0.1355	0.868
4	4	4-methyl piperidine	0.4650	0.332
5	5	4-methyl piperidine	0.3045	0.516
6	7	4-methyl piperidine	0.3085	0.511
7	11	4-methyl piperidine	0.0860	1.065
8*	1	piperidine	0.0445	1.352
9	1	N-methyl piperazine	0.344	0.463
10	1	4-hydroxy piperidine	1.050	–0.021
11	1	pyrrolidine	0.0440	1.356
12	1	thiazolidine	0.5880	0.231
13*	1	$NHC(CH_3)_3$	0.0509	1.293
14	1	NEt_2	0.0416	1.381
15	1	$NHCH_2C_6H_5$	0.1927	0.715
16	1	$NHCH_2C_6H_4Cl$(para)	0.855	0.068
17*	1	1,2,3,4-tetra hydroisoquinoline	1.450	–0.161
18	1	3-amino pyrazole	0.8065	0.093
19	0	H	0.3598	0.444
20	1	piperidine	0.0540	1.268
21	2	piperidine	0.0783	1.106
22*	4	piperidine	0.0155	1.810
23	5	piperidine	0.0274	1.562
24	7	piperidine	0.1650	0.782
25	4	pyrrolidine	0.0230	1.638
26	4	morpholine	0.0141	1.851
27	4	N-methyl piperazine	0.0651	1.186
28	4	NEt_2	0.1411	0.850

structures were energetically optimized using Merck Molecular Force Field with a value of 0.01 as Dielectric Constant.

Input MDL mol files of fully optimized molecules were then browsed into DRAGON software (Todeschini and Consonni et al. 2006, 2009) for computation theoretical structural descriptors. A total number of 1664 structural invariance including topological, 3D and geometrical, constitutional and molecular property, functional group and atom centered fragments have been calculated. The descriptors with same or almost near values or perfectly inter-correlated were reduced from the descriptor data to improve the degree of freedom. Thus, after reduction, a total number of 1367 different descriptors

Table 1 continued

29	4	Br	0.1511	0.821
30	0	Phenyl	0.0755	1.122
31	0	Quinol-4yl	0.0256	1.592
32			0.2817	0.550
33	-	$NHCH_2CH_2NMe_2$	0.0155	1.810
34	-	$NHCH_2CH_2CH_2NMe_2$	0.134	0.873
35*	-	$NMeCH_2CH_2NMe_2$	0.0050	2.301
36	-	$NHCH_2CH_2$pyrrolidine	0.0102	1.991
37	-	$NHCH_2CH_2$pyrrolidine	0.0066	2.180
38	-	$NHCH_2CH_2$pyrrolidine	0.0079	2.046
39	-	NEt_2	0.0082	2.086
40	-	NHtBu	0.0037	2.432
41	-	piperidine	0.0102	1.991
42	-	pyrrolidine	0.0069	2.161
43*	-	N-methyl piperazine	0.0075	2.125
44	-	morpholine	0.0095	2.022
45	-	4-hydroxy piperidine	0.0065	2.187
46*	-	N-(2-hydroxy ethyl) piperazine	0.0115	1.939
47*	-	N-phenyl piperazine	0.0055	2.260
48*	-	N-benzyl piperazine	0.0130	1.886
49	-	N-(diphenyl methyl) piperazine	0.0125	1.903
50	-	N-(4-chloro benzyl) piperazine	0.0115	1.939
51	-	N-(4-methoxy benzyl) piperazine	0.0110	1.959
52	-	N-(4-nitro benzyl) piperazine	0.0132	1.879
53	-	N-(4-diethyl amino benzyl) piperazine	0.0119	1.924
54	-	N-(4-cyano benzyl) piperazine	0.0133	1.876
55	-	N (piperonyl) piperazine	0.0069	2.161
56	-	NHC_6H_5	0.0212	1.674
57*	-	$NHCH_2C_6H_5$	0.0085	2.071
58	-	$NHCH(C_6H_5)_2$	0.0091	2.041
59	-	$NHCH_2C_6H_4Cl$(para)	0.0046	2.337
60	-	$NHCH_2C_6H_4OMe$(para)	0.0044	2.356
61	-	$NHCH_2C_6H_4CF_3$(para)	0.0055	2.260
62*	 (Amodiaquine)		0.0074	2.131

*Test compounds

were selected for further quantitative-structure activity relationship modeling. Descriptor classes along with their names and standard symbols as calculated by the DRAGON software are given in Additional file 1: Table S1 (Batra et al. 2015).

Statistical data analysis

The descriptor data has been analyzed by multiple linear regression (MLR) method. MLR can generate QSAR by correlating a set of computed structural invariance to compound's antimalarial response endpoints. In the present data set sum of descriptors greatly beats the number of compounds. MLR may be applied when the numbers of descriptors are more or lower than the number of compounds (Batra et al. 2015;

Katritzky et al. 2001; Tropsha et al. 2003; Draper and Smith 1998; http://www.minitab.com).

Since a large number of descriptor data have been calculated, so selection of variables is one of the decisive footsteps in QSAR modeling to predict the significant descriptors responsible for producing significant biological activities. If the association between the parameter(s) selected and activity is strong, then activity predictions will be possible. If there is only weak association, knowing the value of the parameter(s) will not help in predicting activity. Thus, for a given study, parameters should be selected which are relevant to the activity for the series of molecules under investigation and these parameters should have values which are obtained in a consistent manner. There are a number of methods for descriptor selection which includes genetic algorithm (de Campos and de Melo 2014; Broadhurst et al. 1997), stimulated annealing (Kirkpatrick et al. 1983), stepwise forward–backward selection (Hoskuldsson 1988; Nandi and Bagchi 2014), etc. Of them stepwise forward–backward feature selection is mostly user-friendly incorporated in Minitab software (http://www.minitab.com) which can select significant variables at 5 % level used in the present study for the generation of a number of QSAR models utilizing different sets of computed molecular descriptors including topological, 3D and geometrical, constitutional and molecular property, functional group and atom centered fragments, respectively.

F statistic value of 4.0 has been selected in the present calculation for inclusion and exclusion of the variables. Four different QSAR models have been formulated which were statistically validated by incorporating test and training sets approaches. The division of the total data set into training and test sets was performed at a random basis. Compounds with asterisk mark in Table 1 were selected as test set compounds. The quality of training model is denoted by R^2 (R is the square root of multiple R-square for regression) and Q^2 (cross-validation R^2) respectively.

R^2 and Q^2 of a model are calculated by

$$R^2 = 1 - \left[\sum (Y_{obs} - Y_{calc})^2 \Big/ \sum \left(Y_{obs} - \bar{Y}\right)^2\right] \quad \text{and}$$

$$Q^2 = 1 - \left[\sum \left(Y_{obs} - Y_{pred}\right)^2 \Big/ \sum \left(Y_{obs} - \bar{Y}\right)^2\right]$$

where Y_{obs}, Y_{calc} and Y_{pred} denote observed, calculated and predicted activity values, respectively, and $\bar{Y}$ indicates mean activity value of training molecules. Q^2 denotes predictive statistics which should be greater than 0.5. The validated QSAR's can identify the most significant contribution of the descriptor data modeled. Such most reliable validated model can be used to predict the highly active congeneric compounds which may be real or virtual, generated by combinatorial library design.

Combinatorial library generation

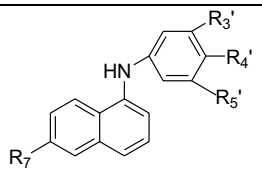
Increase in the drug development cost and big pressure of discovering new molecules, pharmaceutical and biotech companies are crying to design new entity paying least money with increased profitability and productivity. The concept of combinatorial chemistry has been at the forefront of new molecule and drug discovery since 1990 but not quite as powerful a tool currently like a reliable pharmacophore model or

structure-based methods (Ecker and Crooke 1995; Janda 1994; Davies 1996). However, one of the most efficient tools for design of millions of compounds paying least time and cost is computer aided combinatorial library generation. Incorporation of combinatorial library design and high throughput virtual screening including QSAR and pharmacophore modeling or structures based design as a major tools for lead optimization methods applied in the chemo-bioinformatics has dramatically altered the character of new lead discovery research paying least time and cost. Using traditional methods of synthesis, a medicinal chemist can produce limited number of compounds within certain time span. Early SAR studies are based on the use of physical properties and physicochemical substituent constants for the prediction of other more complex physicochemical, bio medicinal, and toxicological properties. Such property–property correlations are useful only when such properties for all compounds are available whereas on application of computer aided combinatorial library design, one can generate millions of compounds within a few time. Most of these compounds have no physicochemical data. Hence, there is a need to develop QSAR models using non-empirical parameters utilizing computed molecular descriptors for the screening of promising lead compound. Once high throughput screening started to make an impact the demand for optimized lead to test experimentally increased dramatically and the researchers began to develop new lead and scaffold more efficiently. The aim of this approach is to screen few potent lead like candidate structures which could be proposed for further synthesis, structure elucidation and biological activity testing using synthetic experiments (Lowe 1995; Terrett et al. 1995; Gallop et al. 1994; Nandi and Bagchi 2011a, b).

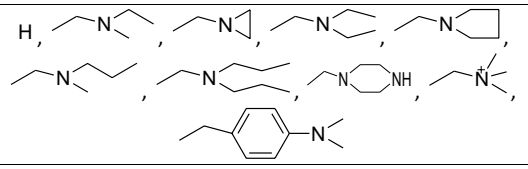
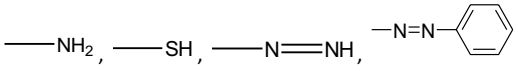
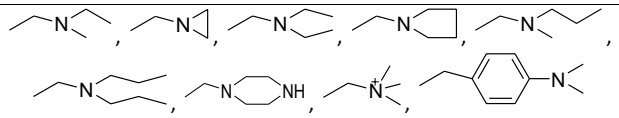
Generation of combinatorial chemical libraries is based on the designing of a scaffold which is a common substructure of the congeneric series. A number of different aliphatic and aromatic substituents are introduced at the specified substitution points of the common nucleus to produce large virtual libraries. In the present article, a total number of 2160 compounds have been generated by introducing different substituents at points of diversity including R_3', R_4', R_5' and R_7, respectively associated to the parent 4-anilinoquinoline nucleus. The following Table 2 represents different possible substituents and the scaffold nucleus structures to develop combinatorial library.

Virtual compounds were then screened by the application of high throughput screening techniques comprising of validated training QSAR, pharmacophore generation (Marshall et al. 1979; Beusen et al. 1999; Golender et al. 1993; Vilar and Koehlar 2000) and Lipinski's 'rule of five' (Lipinski et al. 1997), respectively. The biological activities of the virtual compounds were predicted using the validated training QSAR model based on topological indices. Although this type of activity prediction is the conventional way for predicting active ligands, the method is not beyond contest as we do not have experimental measurement how so far the predicted activity is accurate. Therefore a comparative study between the observed activity of the known amodiaquine lead and predicted activity of the highly active virtual compounds and mode of binding prediction through pharmacophore modeling has been carried out for the highly predicted active congeneric compounds as well as active known leads (such as amodiaquine).

Table 2 Scaffold and possible substituents attached to develop the virtual library



(4-anilinoquinolines)

Substitution points	Possible substituents
R_3'	
R_4'	
R_5'	
R_7	Cl, F, Br, —C≡CH, CCl_3 CF_3

Development of pharmacophore model

A pharmacophore consists of three-dimensional structural topographies for a given series of diverse molecules by ensuring the interaction of molecules with the biological target triggering the biological activity. It provides an estimate of common molecular interaction capabilities of a group of bioactive compounds for its target receptor structure. It does not represent any molecule or a functional group (Leach et al. 2009). All the active molecules sharing maximum number of common features are identified within the conformational flexible active binding region of space (Shoichet 2004; Mason et al. 2001). Therefore 3D pharmacophore assumes the mode of binding of structurally diverse molecules towards the biological target in a possibly common binding mode. These features are denoted as hydrogen bond donor, hydrogen bond acceptor, hydrophobicity of the moiety, aromatic rings, positive ionization properties (cation), negative ionization properties (anion), respectively (Langer and Krovar 2003; Koes and Camacho 2011). However, the concept is very insightful for understanding the molecular recognition aspects of a target receptor shared by a set of bioactive compounds. Pharmacophore modeling methods are not only useful for virtual screening and identification of new hits from databases but also useful for providing insights to de novo design of novel compounds and for understanding the complementary requirements for binding to the active sites of unknown candidate structures as well. Since pharmacophore transcends the chemical structural class and captures only the features responsible for activity, use of pharmacophore has the advantage for identification of potentially new biologically active compounds or chemical scaffolds as novel leads. Therefore in the present

study an attempt has been made to focus on the 3D structural features based pharmacophore generation of 4-anilinoquinolines which are active against *P. falciparum* FcB1R using Portable InteLigandScout software (version 2.02) (Wobler and Langer 2005). Being a fully automated and convenient software tool, Ligand Scout is widely running on all operating systems with works being successfully published (Schuster and Langer 2005). In the present study common binding mode of the congeneric active ligands was analyzed by the development of pharmacophore model considering amodiaquine lead compound using default Ligand Scout settings. In addition to this pharmacophore model has been used as a predictive tool to optimize top 16 highly predicted active combinatorial compounds by generating its individual pharmacophore and compare with amodiaquine pharmacophore to correlate the mode of binding. Very interesting comparative predictive results were found which have been discussed in the next section.

Results and discussion

QSAR modeling

Earlier publications stated that topological indices can produce maximum impact on antimalarial activity of these congeneric compounds (Gupta 2015; Gupta and Prabhakar 2006; Masand et al. 2014; Sahu et al. 2014; Deshpande et al. 2009). Therefore in the present work a number of QSAR models were generated utilizing topological, functional group and atom centered fragments, constitutional and molecular property descriptors, respectively. Impact of different types of descriptors on antimalarial activities is focused in terms of R^2 and its validation is done by calculating cross-validated R2 (R^2_{cv}) while treating the data set using MLR coupled with stepwise forward–backward selection methods. Outcomes were given in the following Table 3.

The above MLR models described that topological indices can produce highest influences in terms of R^2 and R^2_{cv} calculated as 0.870 and 0.810 followed by functional group and atom centered fragments, constitutional and molecular property, 3D and geometrical indices, respectively, which can contribute moderate impact on the inhibition of *P. falciparum* parasitic strain. Therefore, in the next attempt, topological indices have been selected to develop a number of QSAR models which are validated statistically by incorporating training and test sets concept as well as external validations. External validations are carried out by calculating predicted R^2 and r^2_m respectively. Topological descriptor based best training QSAR model, along with its quality and interpretation of modeled parameters are explained in the following Table 4. Predicted R^2 and r^2_m are calculated by the following formula.

$$R^2_{\text{Pred}} = 1 - \frac{\sum \left(Y_{\text{pred(Test)}} - Y_{\text{(Test)}}\right)^2}{\sum \left(Y_{(Test)} - \bar{Y}_{training}\right)^2}$$

Table 3 Impact of descriptors on biological activity

Descriptor class	R^2	R^2_{cv}
Topological	0.870	0.810
Functional group + atom centered fragments	0.812	0.744
Constitutional + molecular property	0.784	0.644
3d + geometrical	0.713	0.634

Table 4 Topological indices based training QSAR model and interpretation of the modeled descriptors

Validated training QSAR model	
$pIC50 = (-10.234) + (-9.3) * BIC5 + (-1.83) * GATS7e + (-3.19) * MATS7e + (4.81) * BEHe4 + (-1.21) * EEig12r + (-0.065) * DP12 + (1.89) * BELm7 + (3.1) * PCR$	
$N = 50, \quad R^2 = 0.870, \quad Q^2 = 0.810, \quad Pred\_R2 = 0.737, \quad r_m^2 = 0.659,$ Average $r_m^2\left(\overline{r_m^2}\right) = 0.682, \quad$ Delta $r_m^2\left(\Delta r_m^2\right) = 0.04, \quad SEE = 0.282$	
Modeled descriptors	**Interpretation**
BIC5	Bond information content index (neighborhood symmetry of 5-order)
GATS7e	Geary autocorrelation—lag 7/weighted by atomic Sanderson electronegativities. 2D autocorrelations
MATS7e	Moran autocorrelation—lag 7/weighted by atomic Sanderson electronegativities
BEHe4	Highest eigen value of number 4 of burden matrix/weighted by atomic Sanderson electronegativities
EEig12r	Eigenvalue 12 from edge adj. matrix weighted by resonance integrals
DP12	Molecular profile number 12
BELm7	Lowest eigenvalue number 7 of Burden matrix/weighted by atomic masses
PCR	ratio of multiple path count over path count

where, $Y_{pred(test)}$ and $Y_{(test)}$ indicate predicted and observed activity values respectively of the test set compounds and $\bar{Y}_{training}$ indicates mean of observed activity values of the training set. For a predictive QSAR model, the value of R^2_{pred} should be more than 0.5 (Nandi and Bagchi 2011a, b).

Further, external predictability of the generated QSAR models was scrutinized by calculating modified r^2 (r_m^2), average modified r^2 ($\overline{r_m^2}$) and delta modified r^2 (Δr_m^2), respectively which are given as

$$r_m^2 = r^2\left(1 - \left|\sqrt{r^2 - r_o^2}\right|\right)$$

where, r^2 and r_o^2 are squared correlation coefficient between the observed (Y axis) and predicted (X axis) activity values of the test set with and without intercept, respectively. r_m^2 value must be greater than 0.5 to have a significant model (Roy and Roy 2008, 2009; Roy et al. 2013). Change of the axes gives the value of r'^2_0 and r'^2_m is calculated by the following formula which depends on the value of r'^2_0.

$$r'^2_m = r^2 \times \left(1 - \sqrt{r^2 - r'^2_0}\right)$$

where, r^2 and r'^2_0 are squared correlation coefficient between the observed (X axis) and predicted (Y axis) activity values of the test set with and without intercept, respectively. Therefore, average r_m^2 and delta r_m^2 are now calculated by

$$\text{Average } r_m^2\left(\overline{r_m^2}\right) = \left(r_m^2 + r'^2_m\right)/2 \quad \text{and} \quad \text{delta } r_m^2\left(\Delta r_m^2\right) = \left|r_m^2 - r'^2_m\right|$$

It is noticeable that an acceptable QSAR model should give the value of "Average r_m^{2}" > 0.5 and "Delta r_m^{2}" should be <0.2, respectively. Values of modified r^2 (r_m^2), average r_m^2($\overline{r_m^2}$) and delta r_m^2(Δr_m^2) have been efficiently computed by web free software link of http://aptsoftware.co.in/ rmsquare/and http://203.200.173.43:8080/rmsquare/, respectively (Roy et al. 2013).

The selected model can explain and predict 87 and 81 % of variances of the antimalarial activity of the deliberated compounds. This model can also produce 73.7 % external predictability and r_m^2 value of 0.659 whereas values of average r_m^2 of 0.682 and delta r_m^2 of 0.04 extend more efficient evidence of external predictability of the generated QSAR model. Further the above QSAR model is confirmed its external predictability by predicting the response activities of the test molecules, as specified in Table 5.

From the Table 5 it is obvious that the predicted responses of all the test compounds are in good treaty with their corresponding observed responses and ideal fit is attained produced by plotting a graph (Fig. 1) by correlating observed activity versus predicted activity of the test set compounds. The squared correlation coefficient is calculated as 0.771.

Once the QSAR model formulated and validated properly, its utility is to predict the biological responses of the compounds which are generated by combinatorial deign and experimentally non-investigated.

Table 5 Predicted activity for the test set molecules

Compound number	Observed activity	Predicted activity
3	0.8681	0.754
8	1.3516	0.535
13	1.2933	1.362
17	−0.1614	0.153
22	1.8097	1.234
35	2.301	1.606
43	2.1249	1.909
46	1.9393	2.031
47	2.2596	2.051
48	1.886	2.013
57	2.0706	1.982
62	2.130	2.023

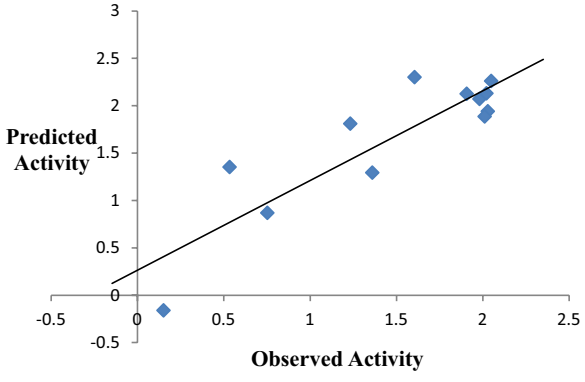


Fig. 1 Observed versus predicted activities of the test molecules

Combinatorial library generation and virtual screening

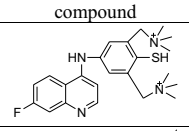
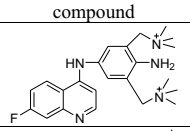
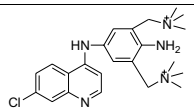
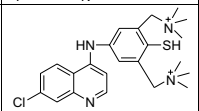
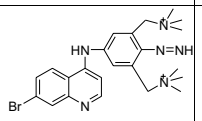
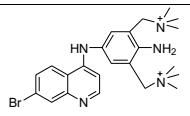
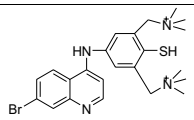
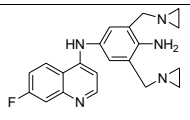
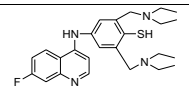
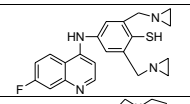
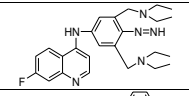
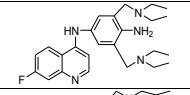
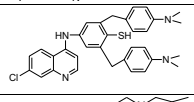
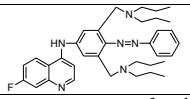
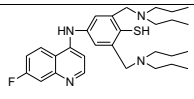
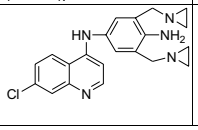
In the present study a total number of 2160 compounds have been generated by introducing a number of 10, 4, 9 and 6 different substituents at various substitution points including R_3', R_4', R_5' and R_7', respectively connected to the common template of 4-anilinoquinoline. The rationale behind this group selection is to undergo literature survey to find out the active lead in these congeners including amodiaquine and isoquine respectively (O'Neill et al. 2003). Let us consider the different functional groups associated to the scaffold of these lead compounds and modify these substituents based on the developed pharmacophore model for active lead. The special feature of the new family reported here are the presence of basic side chain at both the 3′ and 5′ positions and therefore the impossibility of nucleophillic addition of proteins even in the case of a metabolic hydroxylation at the hindered 4′-site (Delarue et al. 2001). R_4' position should be substituted by bio-isosteres of –OH group whereas electron drawing moiety is favorable at R_7 position of the common substructure.

All the optimized virtual compounds were screened by predicting biological activities in terms of pIC_{50} utilizing best topological indices based training QSAR model described in Table 4. As per the prediction, a number of top 16 highly predicted active compounds (Table 6) were reported as hits for further lead optimization process. It was shown that predicted activity of all these 16 highly active virtual hits much greater than the AQ lead. It was also calculated that these highly predicted active virtual compounds match with the property ranges as prescribed by Lipinski's 'Rule of Five' which include following properties such as number of hydrogen bond acceptor, number of hydrogen bond donor, XlogP, molecular weight and rotatable bond count respectively.

As none of these virtual compounds are experimentally tested, it is very vital to test whether these compounds are within the chemical applicability domain (AD) of the developed model, especially in view of that all 16 hit molecules have biological activity values much higher than those of the training set compounds (i.e., the hit compounds are outside the activity domain of the training molecules). The applicability domain of a training QSAR model determines its acceptance by the regulatory bodies such as Organization for Economic Cooperation and Development (OECD) for its applications to predict new molecules. The OECD Principle 3 defines 'a defined domain of applicability' for the developed QSAR model. The Setubal Workshop report (Jaworska et al. 2005) presented the following regulation for the AD assessment: "The applicability domain of a (Q)SAR is the physico-chemical, structural, or biological space, knowledge or information on which the training set of the model has been developed, and for which it is applicable to make predictions for new compounds. The applicability domain of a (Q)SAR should be described in terms of the most relevant parameters, i.e., usually those that are descriptors of the model. Ideally the (Q)SAR should only be used to make predictions within that domain by interpolation not extrapolation" (OECD 2007).

In the present study, applicability domain of the training model as well as top 16 virtual 4-anilino quinolone hits were calculated by using "AD using Standardization approach" which is a free ware tool (Roy et al. 2015a, b, c) to find out whether query compounds are located outside the applicability domain of the built QSAR model and it also detects outliers present in the training set compounds. The results depicted that training molecules

Table 6 Top 16 highly predicted active compounds along with their predicted biological activity

Comp ID	Structure of the highly predicted active combinatorial compound	Predicted Activity	Leverage (h)	Comp ID	Structure of the highly predicted active combinatorial compound	Predicted Activity	Leverage (h)
659		4.943	0.142	649		4.919	0.133
289		4.681	0.116	299		4.656	0.119
1029		4.543	0.137	1009		4.527	0.156
1019		4.458	0.163	403		4.129	0.134
454		4.100	0.139	413		4.091	0.137
464		4.089	0.141	444		4.089	0.139
340		4.057	0.139	597		4.032	0.294
577		4.006	0.156	43		3.929	0.134

1 and 19 were detected as outlier whereas all the predicted hits are situated within the zone of AD.

Further, to cross check the accuracy of this model validation, leverage value (h) and warning leverage (h*) for each of screened hits were calculated. The leverage value (h) of a compound in the original variable space which measures its influence on the model may be defined as

$$h_i = x_i^T \left(X^T X\right)^{-1} x; \quad (i = 1, 2, \ldots, n)$$

where, x_i is the descriptor row-vector of the i-th compound, x_i^T is the transpose of x_i, X is the descriptor matrix, X^T is the transpose of X $(X^T X)^{-1}$ is the inverse of matrix $X^T X$.

The warning leverage (h*) may be calculated by $h^* = 3k/n$

where, n is the number of training compounds and k is the number of model parameters (Hong et al. 2009; Hemmateenejad and Yazdani 2009; Nandi et al. 2011). The leverage value of all hit compounds was mentioned in Table 6. The calculated warning leverage

value is of 0.480. A leverage (h) greater than warning leverage (h*) means that the predicted response is the result of substantial extrapolation of the model and therefore may not be reliable. For the present investigation, it was observed that the leverages of all hit compounds are lower than h* which are pretty acceptable.

Further lead optimization through pharmacophore modeling

Finally, 16 compounds were predicted as promising 4-anilinoquinoline hits active against chloroquine resistant *P. falciparum* FcB1R strain. As the crystal structure of the *P. falciparum* target is unknown, therefore, the predicted hits were subjected to pharmacophore generation to investigate the mode of interaction with the receptor target. Fully optimized 3D structure of AQ was considered as a reference for pharmacophore generation because AQ is an established potential lead-like drug in this in 4-anilinoquinoline congeneric series. To focus the inhibitor's crucial features required for binding with malarial *P. falciparum* FcB1R strain, a comparative study between the amodiaquine (lead) and 4-anilinoquinoline (highly active virtual compounds) pharmacophore has been studied. The pharmacophore model generated by us for amodiaquine (lead) is given in Fig. 2.

The above model predicted five features including three aromatic points (blue circle), six hydrophobicity (yellow ball), two HBAs (orange color), two HBDs (green arrow and lawn green ball) and one positive ionization (blue star). Quinoline nucleus itself should be aromatic and hydrophobic. N1 of the quinoline is a hydrogen bond acceptor whereas 4-amino group is hydrogen bond donor. 4-anilino benzene contributes aromaticity. R_3' may interact with the target by creating hydrophobicity and positive ionization. R_4' can produce hydrogen bond interaction. Electron withdrawing moiety is favorable at R_7 position of the quinoline nucleus which is also responsible for producing hydrophobic interaction.

The detailed comparative pharmacophoric 3D features for AQ (lead) and top 16 predicted active congeners have been given in the following Table 7.

For AQ lead and rest of the highly predicted active compounds considered by us after screening (Table 7), it is seen that the aromatic quinoline ring should interact with hydrophobic residues. N1 of the quinoline, 4-amino and R_4' of the aniline moiety may produce hydrogen bond interactions. 4-anilino benzene shows aromaticity. R_7 position of the quinolone interacts with hydrophobic residues whereas R_3' and R_5' must be

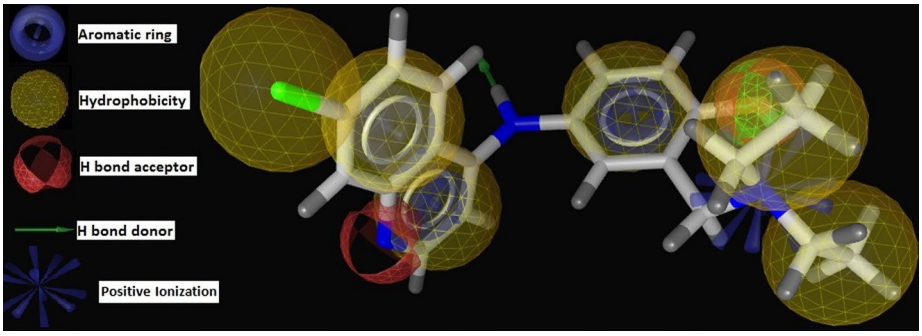


Fig. 2 Pharmacophore of amodiaquine (AQ)

Table 7 Comparative pharmacophoric 3D features for AQ (lead) and top 16 predicted active congeners

Compound ID	Pharmacophoric 3D features predicted by our models							
	Quinoline	N1	4-amino	4-anilino benzene	R_7	R'_3	R'_4	R'_5
AQ (Lead)	H; AR	HBA	HBD	AR	H	PI; H	HBA; HBD	-
659	H; AR	HBA	HBD	AR	H; HBA	PI	HBA;HBD	PI
649	H; AR	HBA	HBD	AR	H; HBA	PI	HBD	PI
289	H; AR	HBA	HBD	AR	H	PI	HBD	PI
299	H; AR	HBA	HBD	AR	H	PI	HBA;HBD	PI
1029	H; AR	HBA	HBD	AR	H	PI	HBA;HBD	PI
1009	H; AR	HBA	HBD	AR	H	PI	HBD	PI
1019	H; AR	HBA	HBD	AR	H	PI	HBA;HBD	PI
403	H; AR	HBA	HBD	AR	H; HBA	PI	HBD	PI
454	H; AR	HBA	HBD	AR	H; HBA	PI; H; H	HBA;HBD	PI; H; H
413	H; AR	HBA	HBD	AR	H; HBA	PI	HBA;HBD	PI
464	H; AR	HBA	HBD	AR	H; HBA	PI; H; H	HBA;HBD	PI; H; H
444	H; AR	HBA	HBD	AR	H; HBA	PI; H; H	HBD	PI; H; H
340	H; AR	HBA	HBD	AR	H	AR; H; H	HBA;HBD	AR; H; H
597	H; AR	HBA	HBD	AR	H; HBA	PI; H; H	HBA; H; AR	PI; H; H
577	H; AR	HBA	HBD	AR	H; HBA	PI; H; H	HBA;HBD	PI; H; H
43	H; AR	HBA	HBD	AR	H	PI	HBD	PI

H hydrophobicity, *AR* aromaticity, *HBD* hydrogen bond donor, *HBA* hydrogen bond acceptor, *PI* positive ionization

substituted by the groups which contribute positive ionizations responsible for ligand receptor interactions. Therefore this is an important attempt for prediction of biochemical mechanisms of the top virtual hits generated by combinatorial library design. Although experimental validation of the screened hits are necessary utilizing in vitro and in vivo analyses, however, an integration of pharmacophore modeling, virtual screening, structure-based methods, molecular biology and combinatorial chemistry together can provide a better basis for more efficient drug discovery and design reducing both costs and time.

New focus on compound's mechanism of action

When the pharmacophore models developed for screened highly active combinatorial hits (Figs. 3, 4) were compared with the AQ lead, a significant comparable performance was noted in terms of mode of interactions with the target protein. As per the prediction, compound ID 659 and 649 were predicted as highest active hits with predicted activities (pIC50) are 4.943 and 4.919 μM respectively. The pharmacophoric interaction patterns of selected active hits were shown in Figs. 3 and 4. Some more active virtual hits were predicted as ID 454, 464, 444, 597 and 577 of those modes of binding and predicted activity are likely with compound ID 659 and 649.

The predicted activities of these two compounds and mode of bindings are almost similar, there is a sharp change in pharmacophore when compared with AQ lead and the other two compounds. From the pharmacophore models (Figs. 3, 4) it is clear that R_3' and R_5' must be substituted by the basic groups containing tertiary amino moiety responsible for producing positive ionization (PI). More PI in these regions can increase affinity of the ligand towards negative environment of the acidic parasitic digestive

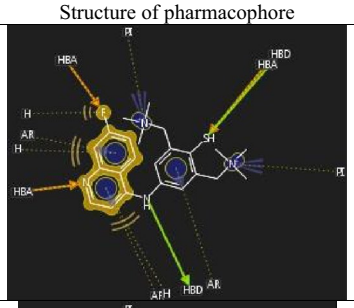
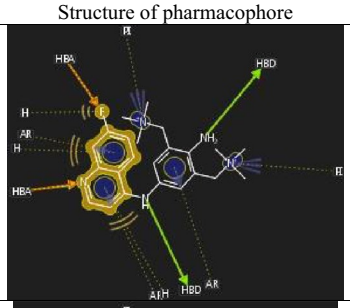
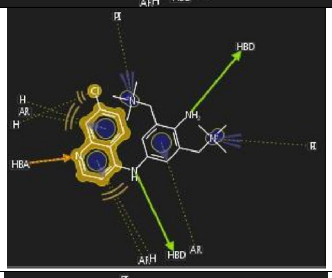
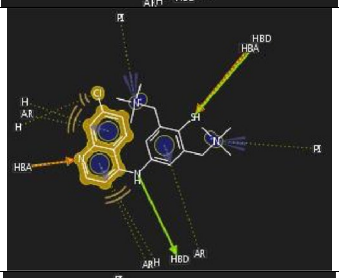
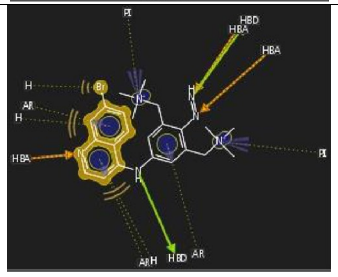
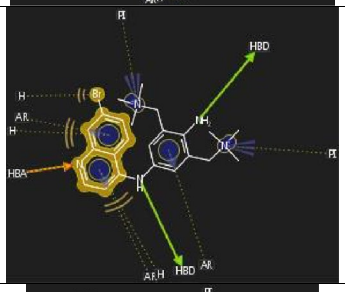
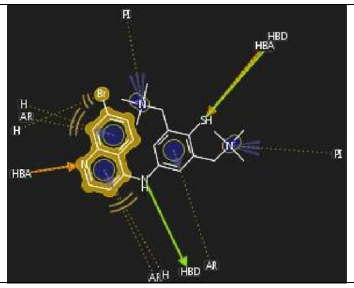
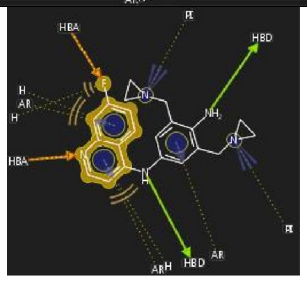
ID	Structure of pharmacophore	ID	Structure of pharmacophore
659		649	
289		299	
1029		1009	
1019		403	

Fig. 3 Pharmacophore models of selected highly active virtual hits

vacuoles. R_7 electron withdrawing moiety of quinoline may produce hydrogen bond interaction with the target protein. These added pharmacophoric features in compare with AQ (lead) pharmacophore may enhance the antimalarial activity of these compounds. The special feature of the new family already reported is the presence of a basic side chain at both the R_3' and R_5' positions and therefore the impossibility of nucleophillic addition of proteins even in the case of a metabolic hydroxylation at the hindered R_4' site (Delarue et al. 2001). Metabolic hydroxylation of R_4' substituent may produce toxic metabolite (O'Neill et al. 2003). To generate least toxic and highly active compounds an attempt has been made in the current study for the designing of 4-anilinoquinoline compounds by substituting thiol, diazo, phenyl diazo and amino groups instead of hydroxyl group.

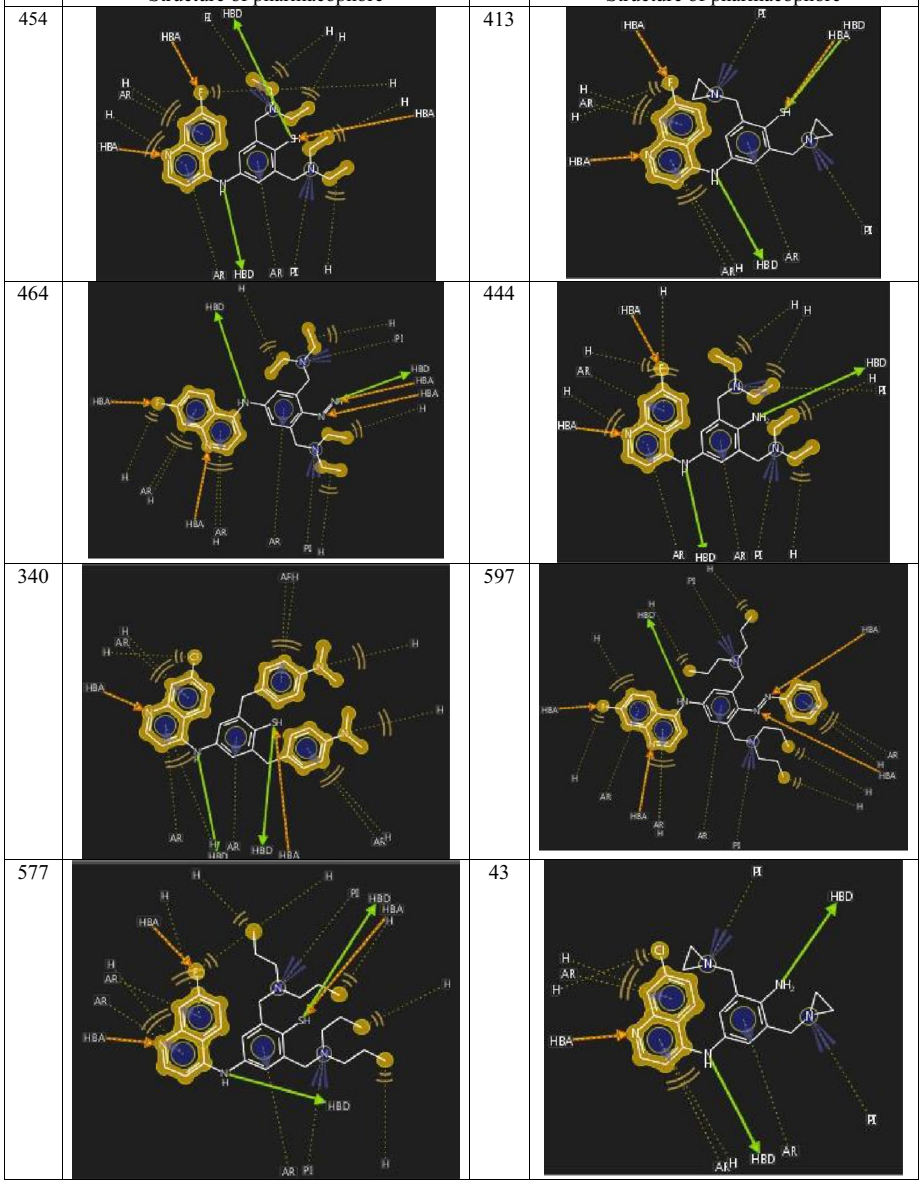


Fig. 4 Pharmacophore models of selected highly active virtual hits

Conclusion and future direction

The advance research, so far yet, focused that 4-anilinoquinolines are basic in nature. They can deposit into the acidic digestive vacuoles of the plasmodium and interact with the heme and interfere with the parasitic DNA sequestration (Valderramos and Fidock 2006). From the present study of pharmacophore modeling, it has been found that tertiary amino group (basic) associated with R_3' and R_5' positions impart positive ionizations. Parasitic nucleic acid bases such as quinine and uracil may undergo nucleophillic attack. The tert-N-group of the compound may interact with this guanine and uracil bases via positive ionization and thus breaks the DNA chain length of the malarial parasite. R_3' and R_5' substituents such as dialkylaminoalkyl moiety may also impart

hydrophobicity and cause hydrophobic interaction with the hydrophobic amino acid residues such as histidine of the parasitic proteins. Heme is bound with histidine and lipid to undergo DNA sequestration. Thus positive charge ionization and hydrophobicity are responsible to inhibit DNA sequestration. Therefore the virtual compounds ID including 454, 464, 444, 597 and 577 are predicted as highly active hits in this congeneric series. Compounds ID 659 and 649 are predicted as highest top two active lead like compounds because R_7 electron withdrawing moiety of these compounds may contribute an additional hydrogen bond interaction which is decisive for producing antimalarial activity. In comparison to AQ lead, other predicted active hits ID including 289, 299, 1029, 1009, 1019, 340 and 43 bear almost same mode of pharmacophoric interaction patterns with an additional feature of PI at R_5' position. Therefore, these predicted active virtual compounds may be recommended for further synthesis and testing as potent agents against *P. falciparum* FcB1R strain. Studies in this direction may help to design new congeneric active leads with least toxicity. Further, synthesis, testing for activity and toxicity study may be carried out in near future to model potent antimalarial compounds in this series.

Additional file

Additional file 1: Table S1. Computed theoretical molecular descriptors used in the study.

Authors' contributions

NP carried out this work. SN initiated and supervised this project. Both authors read and approved the final manuscript.

Acknowledgements

SN is sincerely thankful to National Institute of Chemistry, Slovenia for availing DRAGON and Ligand Scout softwares used in the present work. Neha shows deep sense of gratitude to Dr. Sisir Nandi for the excellent guidance and motivation. Neha deeply acknowledges her beloved sister Anshika for constructive inspiration in this work.

Competing interests

The authors declare that they have no competing interests.

Received: 20 September 2015 Accepted: 4 December 2015
Published online: 29 December 2015

References

Basak SC, Mills D, Gute BD (1997) Predicting bioactivity and toxicity of chemicals from mathematical descriptors: a chemicalcum—biochemical approach, in Advances in Quantum Chemistry. Elsevier, Academic Press

Batra A, Nandi S, Bagchi MC (2015) QSAR and pharmacophore modeling of indole-based C-3 pyridone compounds as HCV NS5B polymerase inhibitors utilizing computed molecular descriptors. Med Chem Res 24:2432–2440

Beusen DD, Marshall GR, Guner O (1999) Pharmacophore definition using the active analog approach: In: pharmacophore perception, development and use in drug design. International University Line La Jolla, pp 21–45

Broadhurst D, Goodacre R, Jone Rowland JJ, Kell BD (1997) Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry. Anal Chim Acta 348:71–86

Crum-Brown A, Fraser TR (1968) On the connection between chemical constitution and physiological action. Part 1. On the physiological action of the ammonium bases, derived from Strychia, Brucia, Thebaia, Codeia, Morphia and Nicotia. Trans R Soc Edinburgh 25:151–203

Davies K (1996) Using pharmacophore diversity to select molecules to test from commercial catalogs, including DIVERSet and HTS Chemicals. In: Chaiken IM, Janda KD (eds) Molecular diversity and combinatorial chemistry: libraries and drug discovery, Chapter 27. American Chemical Society, Washington DC, pp 309–316

de Campos LJ, de Melo EB (2014) Modeling structure-activity relationships of prodiginines with antimalarial activity using GA/MLR and OPS/PLS. J Mol Graph Model 54:19–31

Delarue S, Girault S, Maes L, Fontaine D MA, Labaeı¨d M, Grellier P, Sergheraert C (2001) Synthesis and in vitro and in vivo antimalarial activity of new 4-anilinoquinolines. J Med Chem 44:2827–2833

Deshpande S, Solomon VR, Katti BS, Prabhakar SY (2009) Topological descriptors in modelling antimalarial activity: N^1-(7-chloro-4-quinolyl)-1, 4-bis (3-aminopropyl) piperazine as prototype. J Enzyme Inhib Med Chem 24:94–104

Draper NR, Smith H (1998) Applied regression analysis, 3rd edn. Wiley, New York
Ecker DR, Crooke ST (1995) Combinatorial drug discovery: which methods will produce the greatest value? Biotechnology 13:351–360
Gallop MA, Barrett RW, Dower WJ, Fodor SPA, Gordon AM (1994) Applications of combinatorial technologies to drug discovery 1. J Med Chem 37:1233–1251
Golender VE, Vorpagel ER, Kubinyi H (1993) Computer-assisted pharmacophore identification in 3D QSAR in drug design: theory methods and applications. ESCOM, Leiden, pp 137–149
Gupta KM (2015) CP-MLR/PLS directed QSAR studies on the antimalarial activity and cytotoxicity of substituted 4-aminoquinolines. Med Chem Research 22:3497–3509
Gupta KM, Prabhakar SY (2006) Topological descriptors in modeling the antimalarial activity of 4-(3′, 5′-disubstituted anilino) quinolines. J Chem Inf Model 46:93–102
Hemmateenejad B, Yazdani M (2009) QSPR models for half-wave reduction potential of steroids: a comparative study between feature selection and feature extraction from subsets of or entire set of descriptors. Anal Chim Acta 634:27–35
Hong Q, JingWen C, Ying W, Bin W, XueHua L, Fei L, YaNan W (2009) Development and assessment of quantitative structure-activity relationship models for bioconcentration factors of organic pollutants. Chin Sci Bull 54:628–634
Hoskuldsson A (1988) PLS regression methods. J Chemometrics 2:211–228
Hwang YJ, Kawasuji T, Takashi JD, Clark AJ, Connelly CM, Zhu FG, Sigal SM, Wilson BE, DeRisi LJ, Guy RK (2011) Synthesis and evaluation of 7-substituted 4-aminoquinoline analogues for antimalarial activity. J Med Chem 54:7084–7093
Janda KD (1994) Tagged versus untagged libraries: methods for the generation and screening of combinatorial chemical libraries. Proc Acad Sci USA 91:10779–10785
Jaworska J, Nikolova-Jeliazkova N, Aldenberg T (2005) QSAR applicability domain estimation by projection of the training set descriptor space: a review. ATLA Altern Lab Anim 33:445
Kaschula HC, Timothy EJ, Hunter R, Basilico N, Parapini S, Taramelli D, Pasini E, Monti D (2002) Structure-activity relationships in 4-aminoquinoline antiplasmodials: the role of the group at the 7-Position. J Med Chem 4:3531–3539
Katritzky AR, Petrukhin R, Tatham D, Basak S, Benfenati E, Karelson M, Maran U (2001) Interpretation of quantitative structure-property relationships. J Chem Inf Comput Sci 41:679–685
Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. Am Assoc Adv Sci 220:671–680
Koes DR, Camacho CJ (2011) Pharmer: efficient and exact pharmacophore search. J Chem Inf Model 51:1307–1314
Langer T, Krovat EM (2003) Chemical feature-based pharmacophores and virtual library screening for discovery of new leads. Curr Opin Drug Discov Dev 6:370–376
Leach AR, Gillet VJ, Lewis RA, Taylor R (2009) Three dimensional pharmacophore methods in drug discovery. J Med Chem 53:539–558
Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv 23:3–25
Lowe G (1995) Combinatorial chemistry. ChemSoc Rev 24:309–382
Marshall GR, Barry CD, Bosshard HE, Dammkoehler RA, Dunn DA, Olson EC, Christoffersen RE (1979) The conformation parameter in drug design: the active analog approach in computer–assisted drug design. American Chemical Society, Washington, pp 205–226
Masand HV, Toropov AA, Toropova PA, Mahajan TD (2014) QSAR models for antimalarial activity of 4-aminoquinolines. Curr Comput Aided Drug Des 10:75–82
Mason JS, Good AC, Martin EJ (2001) 3D pharmacophores in drug discovery. Curr Pharm Des 7:567–597
Mills N (2006) ChemDraw Ultra 10.0. J Am Chem Soc 128:13649–13650
Minitab® Statistical Software (2010) Minitab. http://www.minitab.com
Nandi S, Bagchi MC (2011a) In silico design of potent EGFR kinase inhibitors using combinatorial libraries'. Mol Simul 37:196–209
Nandi S, Bagchi MC (2011b) Activity Prediction of Some Nontested Anticancer Compounds Using GA-Based PLS Regression Models. Chem Biol Drug Des 78:587–595
Nandi S, Bagchi MC (2014) QSAR modeling of 4-anilinofuro [2,3-b]quinolines: an approach to anticancer drug design. Med Chem Res 23:1672–1682
OECD (2007) Guidance document on the validation of (quantitative) structure–activity relationships (Q)SARs Models, ENV/JM/MONO(2007)2
O'Neill PM, Mukhtar A, Stocks AP, Randle EL, Hindley S, Ward AS, Storr CR, Bickley FJ, O'Neill IA, Maggs LJ, Hughes HR, Winstanley AP, Bray GP, Park BK (2003) Isoquine and related amodiaquine analogues: a new generation of improved 4-aminoquinoline antimalarials. J Med Chem 46:4933–4945
Pompe M, Novič M (1999) Prediction of gas-chromatographic retention indices using topological descriptors. J Chem Inf Comput Sci 39:59–67
Randic M (1975) On characterization of molecular branching. J Am Chem Soc 79:6609–6615
Ray S, Madrid BP, Catz P, LeValley ES, Furniss JM, Rausch LL, Guy RK, DeRisi LJ, Iyer VL, Green EC, Mirsalis CJ (2010) Development of a new generation of 4-aminoquinoline antimalarial compounds using predictive pharmacokinetic and toxicology models. J Med Chem 53:3685–3695
Roy K, Chakraborty P, Mitra I, Ojha PK, Kar S, Das RN (2013) Some case studies on application of "r^{2}_{m}" metrics for judging quality of quantitative structure–activity relationship predictions: emphasis on scaling of response data. J Comput Chem 34:1071–1082
Roy K, Kar S, Das RN (2015a) Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment, 1st edn. Academic Press, USA
Roy K, Kar S, Das RN (2015b) A Primer on QSAR/QSPR modeling: fundamental concepts (SpringerBriefs in Molecular Science). Springer, New York
Roy K, Kar S, Ambure P (2015c) On a simple approach for determining applicability domain of QSAR models. Chemom Intell Lab Sys 145:22–29

Roy PP, Roy K (2008) On some aspects of variable selection for partial least squares regression models. QSAR Comb Sci 27:302–313
Roy PP, Roy K (2009) Comparative chemometric modeling of cytochrome 3A4 inhibitory activity of structurally diverse compounds using stepwise MLR, FA-MLR, PLS, GFA, G/PLS and ANN techniques. Eur J Med Chem 44:2913–2922
Sahu KN, Sharma CM, Mourya V, Kohli DV (2014) QSAR studies of some side chain modified 7-chloro-4-aminoquinolines as antimalarial agents. Arab J Chem 7:701–707
Schuster D, Langer T (2005) The identification of ligand features essential for PXR activation by pharmacophore modeling. J Chem Info Model 45:431–439
Seder R (2014) Public health: the malaria wars. Nature 514:166
Shoichet BK (2004) Virtual screening of chemical libraries. Nature 432:862–865
Solomon VR, Haq W, Srivastava K, Puri KS, Katti BS (2007) Synthesis andantimalarial activity of side chain modified 4-aminoquinoline derivatives. J Med Chem 50:394–398
Solomon VR, Puri KS, Srivastava K, Katti BS (2005) Design and synthesis of new antimalarial agents from 4-aminoquinoline. Bioorganic Med Chem 13:2157–2165
Terrett NK, Gardner M, Gordon DW, Kobylecki RJ, Steele J (1995) Combinatorial synthesis: the design of compound libraries and their application to drug discovery. Tetrahedron 51:8135–8173
Tham WH, Kennedy AT (2015) Malaria: a master lock for deadly parasites. Nature 522:158–159
Todeschini R, Consonni V (2009) Molecular descriptors for chemoinformatics, revised and enlarged edition, 2nd edn. Wiley, Weinheim
Todeschini R, Consonni V (2006) Dragon software (version 5.4-2006). Milano, Italy
Tropsha A, Gramatica P, Gombar VJ (2003) The importance of being Earnest: validation is the absolute essential for successful application and interpretation of QSPR models. QSAR Comb Sci 22:69–77
Valderramos SG, Fidock DA (2006) Transporters involved in resistance to antimalarial drugs. Trends Pharmacol Sci 27:594–601
Villar HO, Koehlar RT (2000) Comments on the design of chemical libraries for screening. Mol Divers 5:13–24
Wolber G, Langer T (2005) LigandScout: 3D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J Chem Inf Model 45:160–169