Journal of Molecular Graphics and Modelling 46 (2013) 105–124

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

2D, 3D-QSAR and molecular docking of 4(1*H*)-quinolones analogues with antimalarial activities

Thulie Paulinne Jiménez Villalobos [1], Ricardo Gaitán Ibarra [1], Joel José Montalvo Acosta [*,1]

Natural Products Group, Faculty of Pharmaceutical Science, Campus of Zaragocilla, University of Cartagena, Cartagena, Colombia

ARTICLE INFO

Article history:
Accepted 3 October 2013
Available online 14 October 2013

Keywords:
4(1*H*)-Quinolones
QSAR
Docking
Malaria
Cytochrome bc_1

ABSTRACT

Cytochrome bc_1 has become a major focus as a molecular target in malaria parasites, which are the most important vector-borne infectious disease in the world. The inhibition of cytochrome bc_1 blocks the mitochondrial respiratory chain and the consequent arrest of pyrimidine biosynthesis, which is essential for parasite development. The authors developed a theoretical study of two-dimensional, three-dimensional quantitative structure–activity relationships and a docking analysis of a series of 4(1*H*)-quinolones acting as cytochrome bc_1 inhibitors. The predictive ability of the quantitative structure–activity relationship models was assessed using internal (leave-one-out cross-validation) and external (test set with 8 compounds) validation. From the two-dimensional quantitative structure–activity relationship models, the authors emphasized the following descriptors: GCUT_SLOGP_0, SLogP_VSA_5, Kier molecular flexibility index, electrophilicity index, the partition coefficient and the charge of atom 5 of the quinolone ring as the most important to explain the antimalarial activity of the compounds studied. Three-dimensional quantitative structure–activity relationship models showed that the substituents R1 and R4 in 4(1*H*)-quinolones analogues are key modulators to enhance the antimalarial activity. The appropriate binding conformations and orientations of these compounds interacting with cytochrome bc_1 were also revealed by molecular docking. Based on the established models, 8 new compounds with highly predicted antimalarial activity have been theoretically designed and presented as a reference for synthesis and antimalarial evaluation.

1. Introduction

Around the world, malaria is the most significant parasitic disease of humans, and claims the lives of more children worldwide than any other infectious disease. In 2011, approximately 3.3 billion people were at risk of malaria; there were an estimated 219 million cases and an estimated 660.000 deaths [1]. Malaria is caused by protozoa of the genus *Plasmodium* and there are five species that infect humans (*P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale* and *P. knowlesi*) among which *P. falciparum* is the most lethal [2]. Although several attempts have been made to produce a vaccine, drugs are the only therapeutic alternative presently; however, the resistance to traditional therapies has increased morbidity and mortality from malaria, making the search for new antimalarial drugs that use novel molecular targets extremely urgent [3]. The mitochondrial electron transport chain has proved to be a valid chemotherapeutic target because of significant differences between plasmodial and analogous mammalian enzymes [4,5]. Atovaquone (**1**, Fig. 1), which is used in combination with proguanil to treat multidrug-resistant *P. falciparum* infections, is a potent inhibitor of the cytochrome bc_1, a protein subunit in the mitochondrial electron tranport chain [6,7]. Atovaquone selectively binds to the Qo site of cytochorome b, close to the site of interaction with the Rieske iron-sulfur protein (ISP). This inhibition of electron transport by atovaquone blocks respiration and produces a collapse of mitochondrial transmembrane potential [8–10].

Endochin (**2**, Fig. 1), a derivative of 4(1*H*)-quinolone, discovered in the 1940s, has prophylactic and therapeutic activity in avian malaria models, yet the lead languished because of inadequate preclinical models and a poor understanding of parasite biochemistry [11]. Currently, research on endochin and its derivatives have been reconsidered using modern biological models and powerful synthesis strategies in order to find new potential drugs containing the quinolone nucleus [12–14]. Mechanistic studies of 4(1*H*)-quinolone analogues have shown that the cytochrome bc_1 is the target for these compounds thus affecting the parasite mitochondrial respiration [15,16].

For many years, studies of quantitative structure–activity relationships (QSAR) have emerged as a powerful technique used in drug discovery [17]. This is a mathematical model of correlation

* Corresponding author. Tel.: +57 5 6698179; fax: +57 5 6698179.
E-mail addresses: montalvo161@gmail.com, jmontalvoa@unicartagena.edu.co (J.J. Montalvo Acosta).
[1] These authors contributed equally.

1093-3263/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2013.10.002

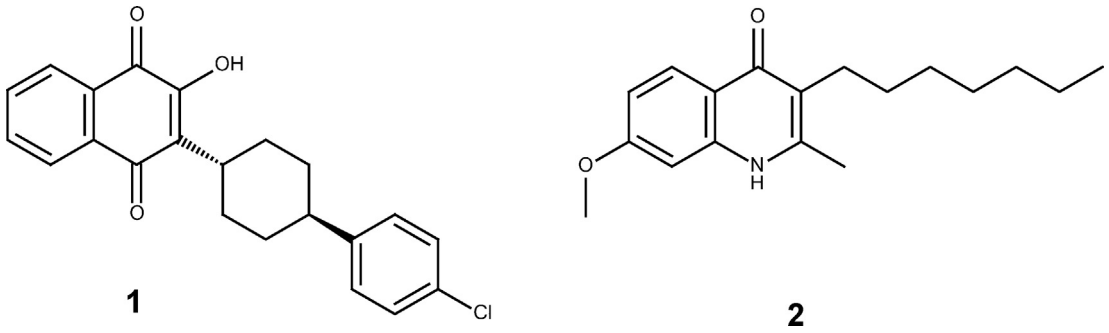


Fig. 1. Structures of atovaquone, **1**, and endochin, **2**.

statistically validated between the variation in chemical structure and biological activity profile of a series of compounds. QSAR is used primarily to correlate molecular descriptors with biological properties, but it can also be applied to predict activity values of non-synthesized compounds, structurally related to the molecules used to build the model. Molecular descriptors can be two-dimensional (2D, such as electronic and physicochemical properties) as well as three-dimensional (3D) [17,18]. Among the 3D-QSAR methods, the Comparative Molecular Field Analysis (CoMFA) and the Comparative Molecular Similarity Indices Analysis (CoMSIA) are extensively used in the current practice of rational drug design [19]. The 3D-QSAR models help to understand the non-bonding interaction characteristics between the drug molecule and the target because they are vivid and robust [20,21]. On the other hand, molecular docking studies provide the possible binding conformations of the ligands in a receptor. Combined with QSAR, they can offer more insight into understanding the detail of protein-inhibitor interactions and the factors affecting bioactivity, and thus providing information for designing new potential drugs [22–24].

In this study, a series of 4(1*H*)-quinolones with antimalarial activity reported recently in literature were chosen to perform a combined 2D-, 3D-QSAR and docking studies. The aim of this work focuses to establish reliable 2D/3D-QSAR models and determine the probably binding conformations for these compounds and provide a guideline for designing novel antimalarial drugs. Based on these models, new compounds with highly predicted antimalarial activity were theoretically designed and they are waiting for experimental verification.

2. Materials and methods

2.1. Data sets

A series of 48 4(1*H*)-quinolones synthesized by Cross et al. with antimalarial activity against the clinically relevant multidrug resistant malarial strain W2 (chloroquine and pyrimethamine resistant), were taken to perform this study [12]. The general structural formulae of the studied compounds are shown in Table 1. For the development of 2D, 3D-QSAR models the complete set of these molecules (48 compounds) were divided into a training set (40 compounds) to generate the models and a test set (8 compounds) to evaluate the predictive ability of the resulting models. The test compounds were selected manually in order to consider the structural diversity and a wide range of antimalarial activity. The *in vitro* antimalarial activities expressed as the Effective Concentration [EC_{50} (nM)] values were converted to the logarithmic scale pEC_{50} [$pEC_{50} = \text{Log}\,1/EC_{50}$ (M)] and used as dependent variables in the 2D- and 3D-QSAR analyses.

2.2. 2D-QSAR model

2.2.1. Calculations of 2D-QSAR descriptors

The geometry of the compounds were built with Marvin 5.12 software [25], and then optimized by the PM6 semi-empirical method in Mopac2009 [26] to find low-energy conformations for each compound. A set of quantum mechanical descriptors: Energy for highest occupied molecular orbital (E_{HOMO}), energy for lowest unoccupied molecular orbital (E_{LUMO}), dipole moment (μ), polarizability (α) and atomic charges on selected atoms (Q1, Q2, Q3, Q4 and Q5, see Table 1 for index identification) were obtained directly from the semi-empirical PM6 calculation. Another set of electronic descriptors such as chemical hardness (η), electronegativity (χ), electrophilicity index (ω) and chemical softness (S) were obtained based on the Koopmans theorem [27] from previously calculated electronic descriptors. Chemical hardness (η), electronegativity (χ) and electrophilicity index (ω) can be defined as follows:

$$\eta \approx -\frac{1}{2}(E_{HOMO} + E_{LUMO}) \approx \frac{1}{2}(I - A) \quad (1)$$

$$\chi \approx \frac{1}{2}(E_{HOMO} - E_{LUMO}) \approx \frac{1}{2}(I + A) \quad (2)$$

$$I \approx -E_{HOMO} \text{ and } A \approx -E_{LUMO} \quad (3)$$

$$\omega = \frac{\chi^2}{2\eta} \quad (4)$$

Where I and A are the ionization potential and electron affinity of the molecules, respectively. This study has also included the HOMO–LUMO energy gap and the partition coefficient (Log P) as quantum mechanical and thermodynamic descriptors, respectively. There are numerous applications of the HOMO–LUMO energy gap in establishing a correlation between the chemical structure and the biological activity. Log P values were calculated using the XLogP 2.0 software [28].

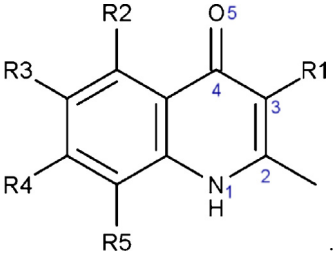
In addition to quantum mechanical descriptors, 185 2D descriptors for which neither energy minimization nor alignment is required were calculated for the built structures of the 4(1*H*)-quinolones, using QuaSAR module of the software Molecular Operating Environment (MOE version 2009.10) [29].

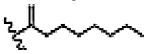
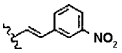
2.2.2. Statistical analysis

To select the predominant descriptors affecting antimalarial activity of the analogues of 4 (1*H*)-quinolones, a correlation analysis was performed by Minitab 14 statistical software [30], taking each descriptor as independent variables and pEC_{50} as dependent variable.

Initially, the descriptors set (185 2D + 15 QM descriptors) was reduced by eliminating out the descriptors with constant and near-constant values. After, a correlation matrix with the

Table 1
Structures and antimalarial activities values of a series of 4(1*H*)-quinolones.



Compounds	R1	R2	R3	R4	R5	EC_{50} W2 (nM)
1	Heptyl	H	H	OCH_3	H	8.59
2	Heptyl	H	H	H	H	93.2
3	Nonyl	H	H	H	H	24.1
4	Cyclohexyl	H	H	H	H	590.7
5	Phenyl	H	H	H	H	711
6	Br	H	H	H	H	453
7	1-nonenyl	H	H	H	H	12.9
8		H	H	H	H	39.6
9	2-Nonenyl	H	H	H	H	101.9
10		H	H	H	H	4599
11	Br	H	OCH_3	H	H	3000
12	Br	H	CH_3	H	H	288
13	Br	H	Br	H	H	1404
14	Br	H	Cl	H	H	845
15	Br	H	F	H	H	929
16		H	H	H	H	8288
17	Benzyl	H	H	OCH_3	H	733
18	Benzyl	OCH_3	H	H	H	6442
19	Phenyl	H	H	OCH_3	H	287
20	Phenyl	OCH_3	H	H	H	8154
21	Phenyl	H	H	H	OCH_3	2063
22	Benzyl	H	H	Cl	H	2168
23	Benzyl	Cl	H	H	H	630
24	Benzyl	H	Cl	H	H	2471
25	Phenyl	H	Cl	H	H	244
26	Phenyl	H	H	Cl	H	216
27	Phenyl	Cl	H	H	H	5064
28	Phenyl	H	Cl	Cl	H	569
29	Phenyl	Cl	Cl	H	H	659
30	Phenyl	Cl	H	Cl	H	2147
31	Phenyl	H	F	OCH_3	H	167
32	Phenyl	H	Cl	OCH_3	H	26.2
33	Phenyl	H	Br	OCH_3	H	79.4
34	Ethyl	H	F	OCH_3	H	368
35	Ethyl	H	Cl	OCH_3	H	48.2
36	Ethyl	H	Br	OCH_3	H	179
37	Heptyl	H	Cl	OCH_3	H	12.3
38	Nonyl	H	Cl	OCH_3	H	6.03
39	Cyclohexyl	H	Cl	OCH_3	H	237
40	Benzyl	H	Cl	OCH_3	H	93.7
41	Methyl	H	Cl	OCH_3	H	56.6
42	Isopropyl	H	Cl	OCH_3	H	112
43	Isobutyl	H	Cl	OCH_3	H	86.4
44	H	H	Cl	OCH_3	H	3175
45		H	Cl	OCH_3	H	2156
46		H	Cl	OCH_3	H	39.4
47	2-Heptyl	H	Cl	OCH_3	H	31.5
48		H	Cl	OCH_3	H	61.7

remaining descriptors and pEC_{50} was built; subsequently, those inter-correlated descriptors ($|r| > 0.9$) were removed in order to minimize redundant information. If multiple descriptors were inter-correlated, the descriptor with a higher correlation with pEC_{50} was kept.

The descriptors with a higher correlation with the pEC_{50} and lower inter-correlation (120 descriptors) were selected to carry out the stepwise multiple linear regression analysis to select the most important descriptors correlating with pEC_{50}. In the stepwise regression method, a multiple-term linear equation is built step-by-step where an initial model is recognized and then it is repeatedly altered by adding or removing a predictor variable according to the "stepping criteria". In this study, the authors used the "stepping criteria" alpha = 4 to enter and alpha = 4 to remove

(default values) in Minitab. The optimal 2D-QSAR equations using the best descriptors obtained from stepwise regression and pEC_{50} was built in BuildQSAR 2.1 software [31].

2.2.3. Validation of QSAR models

Next, the "leave-one-out" (LOO) cross-validation scheme was used to evaluate the predictive ability of the final 2D-QSAR equations. This step is necessary because a high value of the square of the correlation coefficient (r^2) indicates the best fit of the data but does not contain information about the ability to predict the dependent variable of no-include data in the training set. From the LOO cross-validation procedure, the square of cross-validation coefficient (q^2) is obtained, which is used as a criterion to evaluate both the robustness and the predictive ability of the models generated.

According to Tropsha, a QSAR model is considered predictive, if the following conditions are satisfied [32,33]:

$$q^2 > 0.5 \tag{5}$$

$$R^2 > 0.6 \tag{6}$$

$$\frac{R^2 - R_0^2}{R^2} < 0.1 \text{ or } \frac{R^2 - R'^2_0}{R^2} < 0.1 \tag{7}$$

$$0.85 < k < 1.15 \text{ or } 0.85 < k' < 1.15 \tag{8}$$

Mathematical definitions of R^2, R_0^2, R'^2_0, k and k' are based on regression for the test set of the observed activities against predicted activities and vice versa (regression of the predicted activities against observed activities). The definitions are presented clearly in [33] and are not repeated here for shortness.

Also in addition, according to Roy and Roy [34], the difference between values of $R^2{}_0$ and R'^2_0 must be studied and given importance. They suggested the following modified R^2 form:

$$R_m^2 = R^2\left(1 - \left|\sqrt{R^2 - R_0^2}\right|\right) \tag{9}$$

If R_m^2 value for a given model is >0.5, it indicates a good external predictability of the developed model. The model was further validated by applying the Y-randomization test to ensure the robustness of QSAR models and to avoid chance correlation.

Since one of the objectives of developing a QSAR model is for screening new compounds, it is necessary to define a domain of application [32,33]. Predictions for only those compounds that fall into this domain may be considered reliable. Extent of extrapolation [33] is one simple approach to define the applicability of the domain. It is based on the calculation of the leverage *hi* for each chemical, where the QSAR model is used to predict its activity:

$$hi = x_i^T(X^TX)^{-1}x_i \tag{10}$$

Where x_i is the descriptor-row vector of the query molecule and X is the $n \times k$ matrix containing the k descriptor values for each one of the n training molecules. A leverage value greater than $3k/n$ (leverage warning limit) is considered large and implies that the predicted response is the result of a substantial extrapolation of the model and may not be reliable.

2.3. 3D-QSAR models

2.3.1. Molecular alignment

Molecular alignment is one of the most important steps in 3D-QSAR methodologies, such as CoMFA/CoMSIA. For this study, an alignment by superimposing previously optimized ligands was developed using the atoms corresponding to Fig. 2a since the quinolone ring is common to all the compounds. The compound **45**, which is the most active, was used as a template for aligning the remaining training and test sets by using the Database Align option in SYBYL-X 1.3 [35]. The aligned compounds are displayed in Fig. 2b.

2.3.2. CoMFA/CoMSIA settings

CoMFA was performed using the QSAR option of SYBYL. The steric and electrostatic field energies were calculated using the Lennard–Jones and the Coulomb potentials, respectively, with a $1/r$ distance-dependent dielectric constant in all intersections of a regularly spaced grid. A sp^3 carbon atom with a radius of 1.53 Å and a charge of +1.0 was used as a probe to calculate the steric and electrostatic energies between the probe and the molecules using the Tripos force field. The truncation for both the steric and the electrostatic energies were set to 30 kcal mol^{-1} (default cut-off value).

Different combinations of parameters such as number of principal components, fields (steric and electrostatic) and grid spacing (1, 1.5 and 2 Å) were used to select the best CoMFA model.

In this CoMSIA study, a sp^3 carbon atom with a radius of 1.53 Å and a charge of +1.0 was used as the probe to calculate the CoMSIA similarity indices. The same grid spacing used in the best CoMFA model was employed in CoMSIA models. Gaussian-type distance dependence was used to measure the relative attenuation of the field position of each atom in the lattice. The similarity index $A_{F,k}$ for a molecule j with atoms i at a grid point q can be calculated as follows:

$$A_{F,k}^q(j) = -\Sigma_i(\omega_{\text{probe},k})(\omega_{ik})(e^{-\alpha r_{iq}^2}) \tag{11}$$

Where ω_{ik} is the actual value of the physicochemical property k of atom i; $\omega_{\text{probe},k}$ is the probe atom; r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the test molecule and the default value of 0.3 was used as the attenuation factor α, which showed the steepness of the Gaussian-type function.

As in CoMFA, the authors systemically altered the combination of fields and number of principal components to get the CoMSIA model with the best cross-validated q^2, the smallest errors and the largest F value.

2.3.3. Statistical analysis of CoMFA and CoMSIA models

In order to generate 3D-QSAR models, the partial least squares method (PLS) was used for correlating the variation in pEC_{50} values (dependent variable) with variations in the CoMFA/CoMSIA descriptors (independent variables) in the training set.

The PLS method in conjunction with the LOO cross-validation method was used to produce the square of cross-validation coefficient (q^2), the optimal number of principal components and the standard error of prediction (SEP). The PLS without cross-validation method was used to obtain the square of the conventional correlation coefficient (r^2), the standard error of estimate (SEE) and the value F (Fischer test). The best CoMFA and CoMSIA models were those with the highest values of r^2, q^2 and F values in addition to the lowest values of SEE and SEP. The results of the CoMFA and CoMSIA models were graphically represented by contour maps.

2.4. Molecular docking

Molecular docking studies were performed using GOLD software [36]. The ligands were docked within the Qo site of cytochrome bc_1 from *Saccharomyces cerevisiae* [37] (PDB code 3CX5). The three-dimensional structure of *P. falciparum* cytochrome bc_1 is not available, however, there is a high sequence identity between the *Saccharomyces cerevisiae* and plasmodial cytochrome bc_1 (~68%) within the Qo binding pocket [4,38]. The native ligand stigmatellin was removed, and the binding site was defined as all atoms within 7 Å of the crystallographic ligand. Hydrogens were added to amino acid residues and all

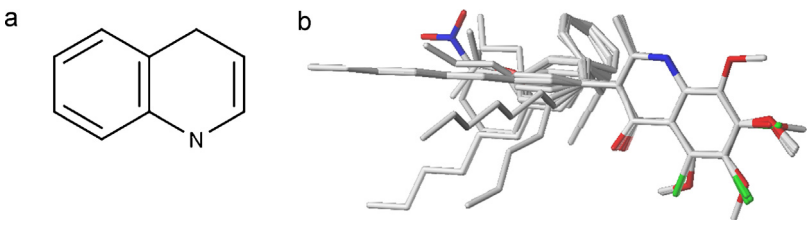


Fig. 2. (a) Common atoms of the ligands used in the alignment procedure. (b) Alignment of the training set using the compound **45** as a template.

crystallographic water molecules were removed, except for HOH7187, which has been described previously as key to the hydrogen bonding network [39]. HOH7187 was set in the "toggle" (On/Off) mode in GOLD (i.e. the decision about formation of interactions or displacement is made by the docking program). The docking with flexible ligands and rigid protein was performed using standard parameters except that the number of runs was 25. The scoring function used for ranking the docked ligands into the Qo site was GoldScore [40].

3. Results and discussion

3.1. 2D-QSAR models

Considering that the antimalarial activity of 4(1*H*)-analogues is caused by inhibiting the cytochrome bc_1 in the mitochondrial electron transport chain, it is reasonable to think that the electronic structure of these compounds is essential to explain their antimalarial activities. For this reason, in addition to 2D descriptors, 15 electronic (quantum chemical) descriptors were calculated and correlated with the antimalarial activity of the 4(1*H*)-quinolone derivatives in the training set.

To improve the 2D-QSAR correlations, the inter-correlated descriptors ($|r| > 0.9$) were eliminated, the stepwise multiple linear regression method was performed, and outliers identified as those molecules whose square of the prediction standard deviation were higher than 2 were excluded. In this study, two 2D-QSAR models were obtained which are discussed below (the input descriptors were normalized by centered on mean and scaled by variance):

Model 1

$$\mathrm{pEC_{50}} = 0.683\mathrm{Q5} + 0.553\mathrm{Log}\,P - 0.552\mathrm{G\_CUT\_SLOGP\_0} + 0.179\omega - 0.159\mathrm{Q2} + 6.723 \quad (12)$$

$$n = 36; r^2 = 0.796; R^2 = 0.768; s = 0.417; F = 23.427; p < 0.0001; q^2 = 0.665; \mathrm{SPRESS} = 0.535.$$

Where n is the number of compounds in the training set used to construct 2D-QSAR equation, r^2 and R^2 are the squares of the correlation coefficient of regression for the training and test sets, respectively; s is the regression standard deviation, F is the Fischer ratio value, p is the statistical confidence level, q^2 is the square of LOO cross-validated coefficient and S_{PRESS} is the standard deviation of sum of square of difference between predicted and observed values. The selected parameters in Eq. (12), as well as the deviations of regression and prediction are listed in Table 2. The plot of the predicted $\mathrm{pEC_{50}}$ values based on Eq. (12) versus experimental ones is shown in Fig. 3a.

Model 1 was built with 36 compounds after eliminating 4 compounds as outliers. Model 1 presented five descriptors correlated with $\mathrm{pEC_{50}}$ with high values of r^2, R^2, q^2 and F and low values of s and S_{PRESS}. In addition, this model passed all the tests for the predictive ability (Eqs. (5)–(9)) demonstrating that this 2D-QSAR model is very robust and predictive (Table 4), and it can offer some useful theoretical references for directing the molecular design of this kind of compound with high antimalarial activity.

The most important parameter in Model 1 is Q5 (coefficient of 0.683), that is, the atomic charge of the carbonyl oxygen in the quinolone ring. Q5 appears in the standardized equation with a positive sign which shows that the molecules with higher value of Q5 have greater value of pEC50. Compounds with the highest Q5 values (**7**, **43** and **45**) have the highest antimalarial activities (Table 2).

The second most important descriptor in Model 1 is the Log *P* (coefficient of 0.553). The positive sign of the Log *P* term in Eq. (12) indicates that the larger the Log *P* value, the higher the activity of the compound is. From Table 2, compounds **7**, **8**, **37** and **38**, which have the largest Log *P* values, also have the greatest antimalarial activities. These compounds possess large alkyl chains in R1; therefore, it is possible that there are hydrophobic interactions between the substituent R1 and the receptor.

The third most important descriptor in model 1 is GCUT_SLOGP_0. This descriptor has a very similar contribution to Log *P* in Model 1, however, it has a negative correlation with $\mathrm{pEC_{50}}$. The GCUT_SLOGP_0 descriptor used atomic contribution to Log *P* (using the Wildman and Crippen SlogP method [41]) instead of partial charge.

Another electronic parameter selected in Eq. (12) is ω, which has positive correlation to $\mathrm{pEC_{50}}$. Conceptually, the electrophilicity is a descriptor of reactivity that allows a quantitative classification of the global electrophilic nature of a molecule within a relative scale. The importance of this reactivity quantity has been recently demonstrated in understanding the toxicity of various pollutants in terms of their reactivity [42,43]. Table 2 presents the compound **10**, which has the smallest ω value (1.4339 eV), and has one of the lowest biological antimalarial activities ($\mathrm{pEC_{50}} = 5.337$) among this data set.

In addition, the electrostatic parameter Q2 exhibits negative correlation with $\mathrm{pEC_{50}}$ (negative sign of Q2 coefficient in Eq. (12)), so the larger Q2 value, the lower the activity of the compound is. Compounds with the highest Q2 values (**10**, **11** and **13**) have the lowest antimalarial activity (Table 2).

Model 2

$$\mathrm{pEC_{50}} = -1.042\mathrm{G\_CUT\_SLOGP\_0} - 0.702\mathrm{SLogP\_VSA5} + 0.580\mathrm{Q5} + 0.493\mathrm{KierFlex} - 0.332\mathrm{DM} + 6.659 \quad (13)$$

$$n = 38; r^2 = 0.784; R^2 = 0.768; s = 0.425; F = 23.252; p < 0.0001; q^2 = 0.674; \mathrm{SPRESS} = 0.523$$

Model 2 is developed for 38 compounds after eliminating 2 molecules as outliers. As in Model 1, Model 2 presented five descriptors correlated with $\mathrm{pEC_{50}}$ and good statistic consistency. Model 2 explained 76.8% of the variance in antimalarial activity (Table 3). The model was found to be predictive and robust, as evident from the higher q^2 and R^2. Also, the proposed model in Eq. (13) passed

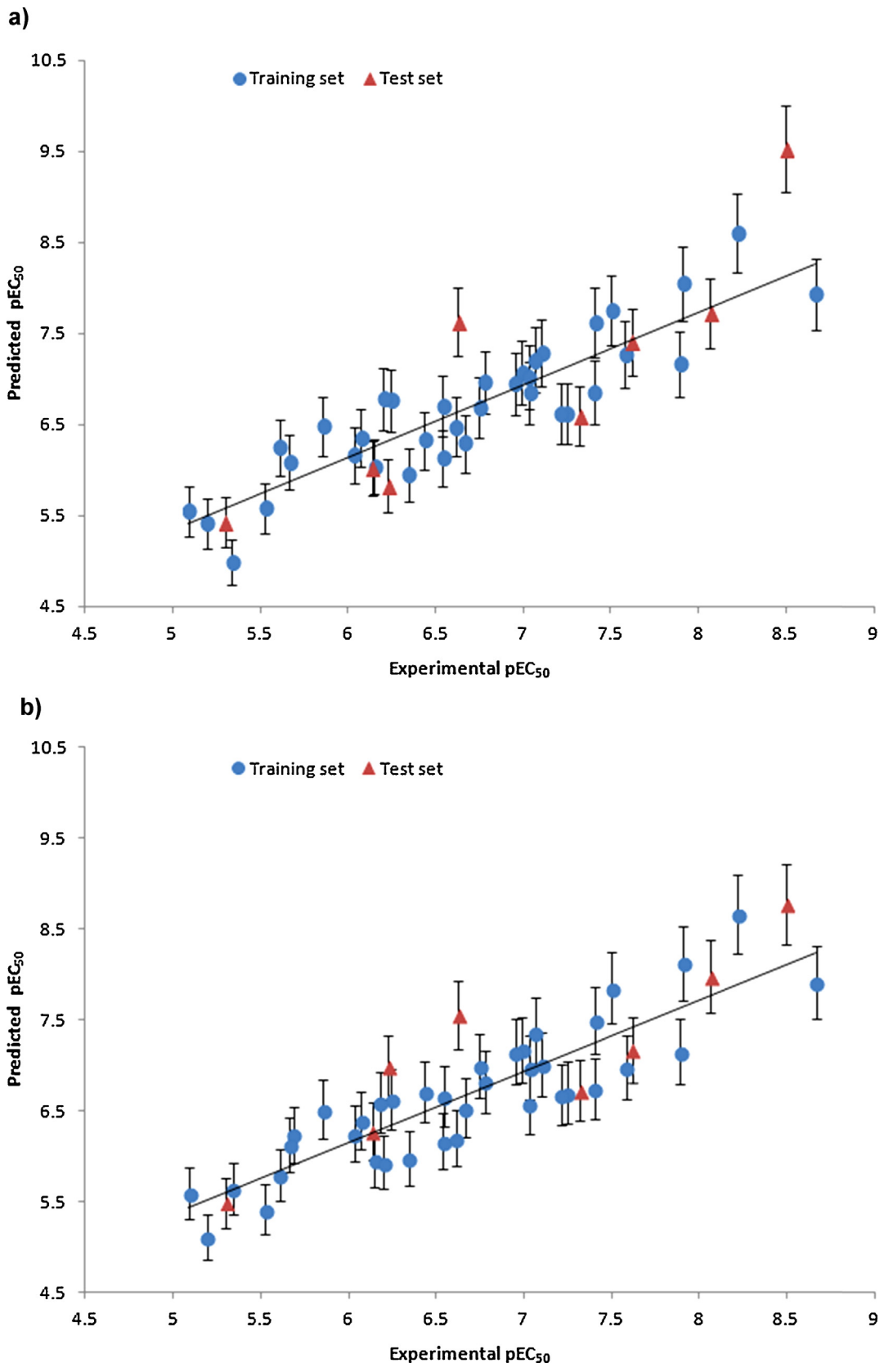


Fig. 3. Plot of predicted activities vs. experimental ones based on Eqs. (12) (a) and (13) (b).

Table 2
Selected descriptors, leverages (h) values and pEC_{50} predicted values from Model 1.

Compounds	GCUT_SLOGP_0	Q5 (u.a.)	ω (eV)	Log P	Q2 (u.a.)	pEC50 Exp.	pEC50 Calc.	Residual	h
2	−0.920466	−0.532	1.911	5.530	0.301	7.031	6.846	0.185	0.109
5	−0.920357	−0.515	1.871	3.940	0.344	6.148	6.032	0.116	0.075
6	−0.920269	−0.490	1.819	2.800	0.352	6.344	5.947	0.397	0.126
7	−0.920404	−0.529	1.836	6.560	0.300	7.889	7.167	0.722	0.145
8	−0.920402	−0.530	1.884	6.390	0.344	7.402	6.855	0.547	0.119
9	−0.920462	−0.537	1.906	6.350	0.299	6.992	7.068	−0.076	0.147
10	−0.920392	−0.533	1.434	4.310	0.338	5.337	4.981	0.356	0.481
11	−0.920257	−0.490	1.669	2.710	0.357	5.523	5.578	−0.055	0.115
12	−0.920284	−0.491	1.802	3.230	0.354	6.541	6.129	0.412	0.099
13	−0.920265	−0.478	1.679	3.600	0.352	5.853	6.487	−0.634	0.099
14	−0.920262	−0.478	1.661	3.420	0.353	6.073	6.352	−0.279	0.103
15	−0.920286	−0.477	1.648	2.960	0.356	6.032	6.171	−0.139	0.117
24	−0.920518	−0.524	1.708	4.380	0.315	5.607	6.244	−0.637	0.064
25	−0.920350	−0.502	1.706	4.560	0.344	6.613	6.473	0.140	0.048
28	−0.920325	−0.496	1.631	5.180	0.344	6.245	6.765	−0.520	0.089
18	−0.920767	−0.569	1.944	3.670	0.314	5.191	5.412	−0.221	0.226
22	−0.920499	−0.530	1.746	4.380	0.313	5.664	6.085	−0.421	0.079
23	−0.920487	−0.510	1.768	4.380	0.305	6.201	6.785	−0.584	0.034
19	−0.920746	−0.526	1.933	3.850	0.339	6.542	6.702	−0.160	0.085
20	−0.920613	−0.549	1.902	3.850	0.342	5.089	5.548	−0.459	0.110
26	−0.920331	−0.508	1.737	4.560	0.341	6.666	6.294	0.372	0.049
31	−0.920754	−0.514	1.804	4.020	0.340	6.777	6.963	−0.186	0.062
32	−0.920743	−0.506	1.719	4.480	0.339	7.582	7.270	0.312	0.081
33	−0.920744	−0.513	1.803	4.650	0.339	7.100	7.290	−0.190	0.067
34	−0.920842	−0.532	1.851	2.770	0.297	6.434	6.329	0.105	0.114
36	−0.920832	−0.531	1.863	3.400	0.296	6.747	6.689	0.058	0.078
37	−0.920847	−0.524	1.743	6.070	0.295	7.910	8.049	−0.139	0.108
38	−0.920847	−0.524	1.743	7.210	0.295	8.220	8.602	−0.382	0.200
40	−0.920903	−0.529	1.722	4.290	0.310	7.028	7.022	0.006	0.085
41	−0.920832	−0.522	1.737	2.760	0.268	7.247	6.620	0.627	0.139
42	−0.920843	−0.524	1.745	3.800	0.294	6.951	6.946	0.005	0.068
43	−0.920844	−0.524	1.745	4.360	0.295	7.063	7.207	−0.144	0.060
45	−0.920752	−0.503	1.734	5.240	0.318	8.666	7.932	0.734	0.097
46	−0.920785	−0.525	1.826	5.790	0.335	7.405	7.617	−0.212	0.090
47	−0.920787	−0.528	1.758	5.960	0.293	7.502	7.753	−0.251	0.083
48	−0.920800	−0.516	1.625	4.050	0.356	7.210	6.616	0.594	0.146
1[a]	−0.920851	−0.539	2.007	5.450	0.294	8.066	7.726	0.340	0.133
3[a]	−0.920467	−0.532	1.911	6.670	0.301	7.618	7.402	0.216	0.152
4[a]	−0.920502	−0.537	1.914	4.560	0.372	6.229	5.823	0.406	0.087
17[a]	−0.920907	−0.547	1.978	3.670	0.390	6.135	6.027	0.108	0.215
27[a]	−0.920319	−0.513	1.902	3.850	0.428	5.296	5.434	−0.138	0.201
35[a]	−0.920831	−0.525	1.747	3.230	0.296	7.317	6.596	0.721	0.086
39[a]	−0.920882	−0.529	1.843	5.100	0.303	6.625	7.630	−1.005	0.067
44[a]	−0.920663	−0.451	1.767	2.540	0.119	8.498	9.526	−1.028	0.717
16[b]	−0.920367	−0.505	1.820	4.700	0.325	5.082	6.803	−1.721	0.054
30[b]	−0.920294	−0.483	1.635	5.180	0.332	5.668	7.203	−1.535	0.107
21[b]	−0.920729	−0.513	1.799	3.850	0.348	5.686	6.798	−1.112	0.066
29[b]	−0.920314	−0.481	1.618	5.180	0.335	6.181	7.273	−1.092	0.118

[a] Compounds in the test set.
[b] Outliers.

the rest of the tests (Eqs. (5)–(9)) that were used for illustrating its predictive ability (Table 4). The selected parameters in Eq. (13), as well as the deviations of regression and prediction are listed in Table 3. The plot of the predicted pEC_{50} values based on Eq. (13) versus experimental ones is shown in Fig. 3b.

As in Model 1, the G_CUT_SLOGP_0 and Q5 descriptors were found in Model 2 and preserved the same sign. However, these descriptors presented different contributions to both models, thus G_CUT_SLOGP_0 is the most influential descriptor (coefficient of 1.042) in Model 2 while Q5 (coefficient of 0.580) is the third (in Model 1 was the opposite).

The second term in Model 2 was SLogP_VSA5 that is defined to be the sum of the v_i for over all atoms i (calculated with L_i between 0.15 and 0.20) and the atomic property p_i (Log P) for atom i as calculated in the SLogP descriptor (calculated with the Wildman and Crippen SLogP method [41]). The negative contribution of this descriptor reveals that the partition coefficient on the vdW surface of the molecules is unfavorable for the antimalarial activity.

KierFlex, molecular flexibility index, is given by (KierA1)(KierA2)/n. The Kier and Hall kappa molecular shape indices [44] compare the molecular graph with minimal and maximal molecular graphs. The positive contribution of molecular flexibility (KierFlex), a topological 2D parameter, indicates the influence of the molecular shape on antimalarial activity.

DM is the dipole moment. Dipole moments and other related physical quantities, such as multipole moments and polarizabilities, constitute another group of molecular descriptors which can be defined either in terms of classical physics or in terms of quantum mechanics. They encode information about the charge distribution in molecules [45]. The dipole moment (DM) is a vector quantity that encodes displacement with respect to the center of gravity of positive and negative charges in a molecule. This descriptor presented a small negative value in Eq. (13), therefore an increase in the charge distribution slightly decreases the antimalarial activity of the studied compounds.

Table 3
Selected descriptors, leverages (h) values and pEC50 predicted values from Model 2.

Compounds	GCUT_SLOGP_0	KierFlex	SlogP_VSA5	Dipole Moment (Debye)	Q5 (u.a.)	pEC50 Exp.	pEC50 Calc.	Residual	h
2	−0.920466	3.898	5.942	5.910	−0.532	7.031	6.970	0.061	0.095
5	−0.920357	1.786	2.757	6.411	−0.515	6.148	5.960	0.188	0.098
6	−0.920269	1.991	5.812	7.123	−0.490	6.344	5.973	0.371	0.077
7	−0.920404	4.592	5.942	5.726	−0.529	7.889	7.145	0.744	0.129
8	−0.920402	4.280	9.128	6.072	−0.530	7.402	6.738	0.664	0.099
9	−0.920462	4.592	5.942	5.762	−0.537	6.992	7.172	−0.180	0.134
10	−0.920392	2.758	5.942	7.156	−0.533	5.337	5.641	−0.304	0.110
11	−0.920257	2.670	41.196	5.718	−0.490	5.523	5.417	0.106	0.551
12	−0.920284	2.220	5.812	7.030	−0.491	6.541	6.161	0.380	0.068
13	−0.920265	3.216	8.867	7.709	−0.478	5.853	6.509	−0.656	0.117
14	−0.920262	2.765	5.812	7.759	−0.478	6.073	6.390	−0.317	0.101
15	−0.920286	2.157	5.812	7.777	−0.477	6.032	6.244	−0.212	0.103
24	−0.920518	2.740	24.811	7.075	−0.524	5.607	5.790	−0.183	0.059
25	−0.920350	2.357	2.757	7.373	−0.502	6.613	6.198	0.415	0.072
21	−0.920729	2.331	38.141	7.460	−0.513	5.686	6.236	−0.550	0.099
28	−0.920325	3.011	2.757	7.115	−0.496	6.245	6.623	−0.378	0.057
29	−0.920314	3.011	2.757	8.209	−0.481	6.181	6.593	−0.412	0.108
18	−0.920767	2.703	60.195	4.922	−0.569	5.191	5.109	0.082	0.318
22	−0.920499	2.740	24.811	5.495	−0.530	5.664	6.124	−0.460	0.066
23	−0.920487	2.740	24.811	7.340	−0.510	6.201	5.932	0.269	0.057
19	−0.920746	2.331	38.141	5.544	−0.526	6.542	6.664	−0.122	0.074
20	−0.920613	2.331	38.141	5.070	−0.549	5.089	5.587	−0.498	0.144
26	−0.920331	2.357	2.757	5.779	−0.508	6.666	6.531	0.135	0.085
31	−0.920754	2.492	38.141	6.298	−0.514	6.777	6.822	−0.045	0.065
32	−0.920743	2.961	38.141	6.875	−0.506	7.582	6.978	0.604	0.059
33	−0.920744	3.292	41.196	6.411	−0.513	7.100	7.011	0.089	0.041
34	−0.920842	2.413	41.326	5.874	−0.532	6.434	6.712	−0.278	0.100
36	−0.920832	3.388	44.381	5.958	−0.531	6.747	6.998	−0.251	0.049
37	−0.920847	5.616	41.326	6.608	−0.524	7.910	8.119	−0.209	0.128
38	−0.920847	6.825	41.326	6.624	−0.524	8.220	8.652	−0.432	0.257
40	−0.920903	3.374	60.195	6.506	−0.529	7.028	6.571	0.457	0.103
41	−0.920832	2.531	41.326	6.667	−0.522	7.247	6.698	0.549	0.093
42	−0.920843	3.456	41.326	6.601	−0.524	6.951	7.145	−0.194	0.061
43	−0.920844	3.959	41.326	6.602	−0.524	7.063	7.365	−0.302	0.058
45	−0.920752	4.378	41.326	6.104	−0.503	8.666	7.907	0.759	0.090
46	−0.920785	4.871	44.512	6.277	−0.525	7.405	7.490	−0.085	0.082
47	−0.920787	5.190	41.326	5.801	−0.528	7.502	7.841	−0.339	0.099
48	−0.920800	4.649	41.326	9.385	−0.516	7.210	6.673	0.537	0.294
1[a]	−0.920851	4.661	41.326	4.792	−0.539	8.066	7.969	0.097	0.118
3[a]	−0.920467	4.999	5.942	6.705	−0.532	7.618	7.170	0.448	0.135
4[a]	−0.920502	2.436	5.942	4.186	−0.537	6.229	6.980	−0.751	0.251
17[a]	−0.920907	2.703	60.195	5.158	−0.547	6.135	6.271	−0.136	0.134
27[a]	−0.920319	2.357	2.757	8.014	−0.513	5.296	5.486	−0.190	0.122
35[a]	−0.920831	2.979	41.326	6.936	−0.525	7.317	6.724	0.593	0.080
39[a]	−0.920882	3.812	41.326	6.064	−0.529	6.625	7.547	−0.922	0.073
44[a]	−0.920663	2.302	38.141	4.514	−0.451	8.498	8.761	−0.263	0.549
16[b]	−0.920367	2.847	5.942	6.152	−0.505	5.082	6.744	−1.662	0.056
30[b]	−0.920294	3.011	2.757	6.942	−0.483	5.668	6.910	−1.242	0.080

[a] Compounds in the test set.
[b] Outliers.

Both models were further validated by applying the Y-randomization. Several random shuffles of the Y vector were performed and the results are shown in Table 5. The low r^2 and q^2 values show that the good results in our original models were not based on chance correlation.

Table 4
Statistical parameters for evaluating the predictive abilities of different QSAR models in study.

Parameter	2D-QSAR		3D-QSAR	
	Model 1	Model 2	CoMFA	CoMSIA
q^2	0.665	0.674	0.635	0.537
R^2	0.768	0.768	0.817	0.776
R_0^2	0.765	0.721	0.478	0.301
R'^2_0	0.655	0.766	0.763	0.715
$(R^2 - R_0^2)/R^2$	0.004	0.061	0.415	0.612
$(R^2 - R'^2_0)/R^2$	0.148	0.002	0.066	0.078
k	1.008	1.016	0.972	0.970
k'	0.985	0.980	1.024	1.025
R_m^2	0.725	0.602	0.341	0.241

3.1.1. Defining model applicability domain

It needs to be emphasized that no matter how robust, significant and validated a QSAR/QSPR may be, it cannot be expected to reliably predict the modeled property for the entire universe of chemicals. Therefore, a very important step in QSAR model development is the definition of the applicability domain of classification or regression models, thus predictions for only those chemicals that fall in this domain may be considered reliable. In this study, the descriptor matrices for Model 1 and 2 were of the order 36×5 and 38×5, respectively and thus the warning leverage values ($h^* = 3k/n$) for Model 1 and 2 were 0.417 and 0.395, respectively.

To visualize the applicability of domain of the developed 2D-QSAR models, William plots were used. In the William plot, standardized residuals versus leverage values (h values in Tables 2 and 3) are plotted. This plot could be used for an immediate and simple graphical detection of both the response outliers, i.e., compounds with standardized residuals >2 standard deviation units and structurally influential compounds in a model, i.e., $h > h^*$. Compounds with $h > h^*$ hardly affect the goodness of fit of the

Table 5
r^2 and q^2 values after several Y-randomization tests for QSAR models in study.

Iteration	Model 1		Model 2		CoMFA		CoMSIA	
	r^2	q^2	r^2	q^2	r^2	q^2	r^2	q^2
1	0.076	−0.328	0.135	−0.159	0.135	−0.120	0.022	0.001
2	0.044	−0.327	0.034	−0.567	0.032	−1.123	0.039	0.068
3	0.120	−0.245	0.164	−0.340	0.003	0.002	0.000	−1.423
4	0.118	−0.251	0.183	−0.061	0.004	−0.033	0.020	−0.863
5	0.108	−0.197	0.128	−0.155	0.012	−0.050	0.009	−1.113
6	0.095	−0.333	0.184	−0.132	0.017	−0.962	0.155	0.015
7	0.112	−0.262	0.186	−0.066	0.034	0.021	0.050	0.076
8	0.250	−0.171	0.031	−0.387	0.000	−1.362	0.001	−1.323
9	0.157	−0.342	0.096	−0.202	0.006	−0.037	0.003	−0.963
10	0.162	−0.294	0.011	−0.335	0.077	−0.085	0.074	−0.057

developed model but these compounds may not be an outlier because of their low residuals. It must be noted that compounds with high value of leverage and good fitting in the developed model can stabilize the model. On the other hand, compounds with bad fitting in the developed model may be outliers. Thus, combination of leverage and the standardized residual could be used for assigning the applicability of domain.

The William plots for the developed models is shown in Fig. 4a and b. In Model 1 (Fig. 4a), the results indicate that out of the 48 compounds in study only two (**10** and **44**) have values greater than the warning leverage (dashed line), one of these influential compounds belong to the test set (**44**). Thus, 88% of the test compounds are within the applicability domain indicating that their predicted activity values are reliable. A total of 4 outliers were detected in Model 1, corresponding to the compounds **16**, **21**, **29** and **30**. Also, for the Model 2 (Fig. 4b), compound **44** had a leverage value higher than warning leverage (dashed line), in addition to compound **11**. In this Model, only two outliers were detected, corresponding to the compounds **16** and **21**.

3.2. 3D-QSAR models

3.2.1. Statistical results of CoMFA and CoMSIA Models

In addition to 2D-QSAR studies, 3D-QSAR CoMFA/CoMSIA models were also established to predict and interpret the antimalarial activities of the quinolones analogues. The CoMFA and CoMSIA models were developed using a training set of 40 quinolones analogues. The predictive ability of the best models was evaluated by using the same test set used in 2D-QSAR studies (a total of 8 compounds).

Different combinations of parameters (number of principal components, fields and grid spacing) were used for selection of the best CoMFA and CoMSIA models. A summary of all models

Table 6
Summary of CoMFA results.

No.	NC[a]	q^2[b]	SEP[c]	r^2[d]	SEE[e]	R^2[f]	*F*-Value	Grid spacing (Å)	Fields
1	4	0.466	0.679	0.935	0.236	–	126.619	2	S
2	5	0.474	0.682	0.948	0.214	–	125.064	2	S
3	6	0.541	0.647	0.957	0.197	0.521	123.880	2	S
4	7	0.578	0.630	0.961	0.191	0.548	112.855	2	S
5	8	0.584	0.635	0.967	0.180	0.583	112.502	2	S
6	9	0.592	0.640	0.971	0.172	0.578	110.206	2	S
7	3	0.392	0.725	0.782	0.427	–	430.413	2	E
8	4	0.422	0.705	0.889	0.309	–	69.846	2	E
9	5	0.422	0.739	0.906	0.289	–	65.462	2	E
10	2	0.445	0.699	0.848	0.351	–	103.532	2	S+E
11	3	0.469	0.666	0.917	0.264	–	131.747	2	S+E
12	4	0.486	0.665	0.952	0.202	–	174.828	2	S+E
13	5	0.498	0.667	0.961	0.187	–	166.276	2	S+E
14	6	0.498	0.678	0.967	0.173	–	162.352	2	S+E
15	7	0.508	0.680	0.971	0.166	0.562	150.865	2	S+E
16	3	0.512	0.682	0.884	0.311	0.566	91.589	1.5	S+E
17	4	0.514	0.646	0.918	0.266	0.581	98.009	1.5	S+E
18	5	0.585	0.606	0.936	0.238	0.717	99.333	1.5	S+E
19	6	0.610	0.597	0.948	0.219	0.775	99.295	1.5	S+E
20	7	0.622	0.596	0.956	0.195	0.782	108.275	1.5	S+E
21	8	0.631	0.599	0.963	0.189	0.784	101.153	1.5	S+E
22[g]	**9**	**0.635**	**0.405**	**0.970**	**0.173**	**0.817**	**108.238**	**1.5**	**S+E**
23	2	0.367	0.718	0.772	0.431	–	62.544	1	S+E
24	3	0.420	0.697	0.901	0.287	–	109.528	1	S+E
25	4	0.515	0.646	0.933	0.239	0.523	122.572	1	S+E
26	5	0.577	0.612	0.946	0.219	0.612	118.301	1	S+E
27	6	0.596	0.607	0.953	0.207	0.604	111.692	1	S+E
28	7	0.617	0.601	0.964	0.185	0.676	121.282	1	S+E

S = Steric; E = Electrostatic.
[a] Number of components.
[b] Cross-validated correlation coefficient after the leave-one-out procedure.
[c] Standard error of prediction.
[d] Non-cross-validated correlation coefficient for the training set.
[e] Standard error of estimate.
[f] Non-cross-validated correlation coefficient for the test set.
[g] The best selected model in bold.

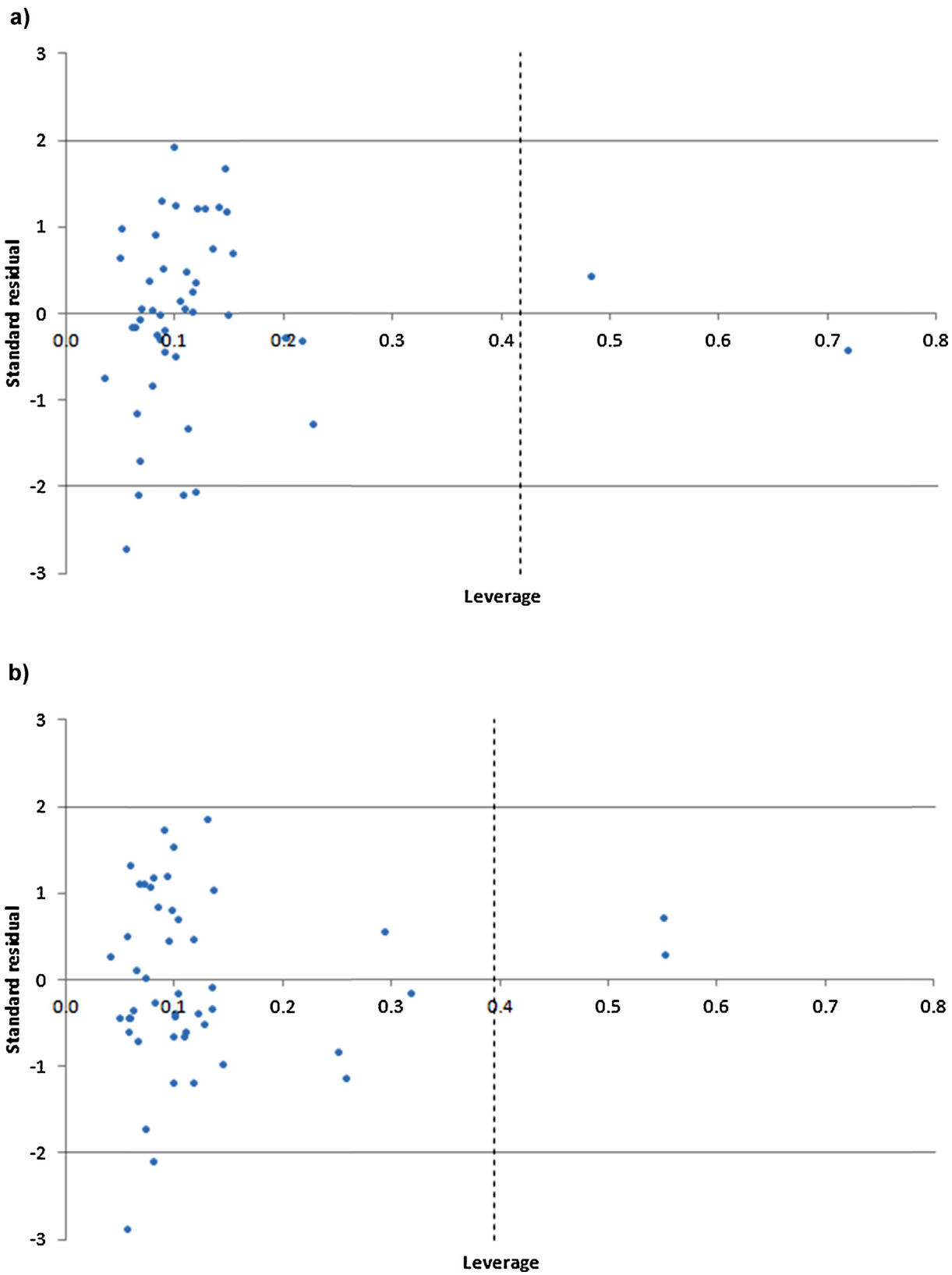


Fig. 4. William plot for the developed Models 1(a) and 2 (b).

Table 7
Summary of CoMSIA results.

No.	NC[a]	q^2[b]	SEP[c]	r^2[d]	SEE[e]	R^2[f]	*F*-Value	Grid spacing (Å)	Fields
1	2	0.446	0.695	0.601	0.570	–	27.885	1.5	H
2	4	0.450	0.688	0.760	0.442	–	58.614	1.5	E
3	3	0.437	0.701	0.806	0.403	–	49.783	1.5	S
4	3	−0.108	0.97	–	–	–	–	1.5	A
5	3	−0.060	0.92	–	–	–	–	1.5	D
6	4	0.510	0.649	0.840	0.371	0.521	46.016	1.5	S+H
7	3	0.506	0.694	0.810	0.399	0.502	51.036	1.5	D+S
8	4	0.437	0.669	0.880	0.346	–	64.304	1.5	D+S
9	5	0.510	0.746	0.891	0.310	0.485	55.737	1.5	S+H
10	3	0.465	0.676	0.811	0.422	–	32.023	1.5	S+H+E
11	4	0.480	0.669	0.880	0.361	–	52.112	1.5	S+H+E
12[g]	**5**	**0.537**	**0.443**	**0.921**	**0.264**	**0.775**	**79.755**	**1.5**	**S+H+E**
13	4	0.453	0.702	0.790	0.425	–	32.910	1.5	D+A+S
14	3	0.450	0.716	0.612	0.569	–	18.948	1.5	D+A+H
15	2	0.476	0.683	0.680	0.510	–	39.288	1.5	S+E+H+D+A
16	3	0.476	0.712	0.684	0.499	–	40.231	1.5	S+E+H+D+A

S = Steric; H = Hydrophobic; E = Electrostatic; D = Hydrogen bond donor; A = Hydrogen bond acceptor.
[a] Number of components.
[b] Cross-validated correlation coefficient after the leave-one-out procedure.
[c] Standard error of prediction.
[d] Non-cross-validated correlation coefficient for the training set.
[e] Standard error of estimate.
[f] Non-cross-validated correlation coefficient for the test set.
[g] The best selected model in bold.

generated in this study is in Tables 6 and 7. R^2 values were calculated only for those models with $q^2 > 0.5$ because to get reliable predictive models, the q^2 and R^2 values should be higher than 0.5 and 0.6, respectively. Statistical results for the CoMFA and CoMSIA final models are summarized in Table 8.

The final CoMFA model selected has a cross-validation coefficient (q^2) of 0.635 for nine components used and a standard error of prediction of 0.405. The non-cross-validated PLS analysis results in a conventional r^2 of 0.970 with an *F* value of 108.238 and an estimated standard error of 0.173. Also, we can observe that steric field descriptors explain 68.3% of the variance, while the electrostatic descriptors explain 31.7% of the variance.

The best CoMSIA model was developed using 3 fields: steric, electrostatic and hydrophobic. For the CoMSIA model, the q^2 value obtained was 0.537 with five principal components and a standard error of prediction of 0.443. The r^2 value was 0.921 with an *F* value of 79.755 and an estimated standard error of 0.264. The contributions to the variance of the steric, electrostatic and hydrophobic fields were 34%, 39.4% and 26.6%, respectively. Both CoMFA and CoMSIA models passed the tests proposed in Eqs. (5)–(8), showing that the predictive ability of these models was satisfactory (Table 4).

Table 8
Summary of the results obtained from the CoMFA and CoMSIA analyses.

Statistical parameters	CoMFA	CoMSIA
q^2[a]	0.635	0.537
r^2[b]	0.970	0.921
R^2[c]	0.817	0.775
ONC[d]	9	5
SEE[e]	0.173	0.264
SEP[f]	0.405	0.443
F-Value	108.238	79.755
Grid spacing (Å)	1.5	1.5
Contribution (%)		
Steric	68.3	34.0
Electrostatic	31.7	39.4
Hydrophobic		26.6

[a] Cross-validated correlation coefficient after the leave-one-out procedure.
[b] Non-cross-validated correlation coefficient for the training set.
[c] Non-cross-validated correlation coefficient for the test set.
[d] Optimum number of components.
[e] Standard error of estimate.
[f] Standard error of prediction.

Experimental and predicted activities from the best CoMFA and CoMSIA models for 48 compounds in this study are shown in Table 9. The correlation between predicted and experimental activities for both models is shown in Fig. 5a and b. As in 2D-QSAR, the final CoMFA and CoMSIA models were validated by applying the Y-randomization test (Table 5). From this test, low values of r^2 and q^2 were obtained, excluding the possibility of chance correlation or structural dependency of the training set.

3.2.2. Graphical Interpretation of CoMFA and CoMSIA models

To visualize the information contained in the 3D-QSAR models, maps of contours CoMFA and CoMSIA were generated, which can provide a major understanding of the biological activities of quinolone derivatives. The steric and electrostatic CoMFA contour fields are shown in Fig. 6. In the steric field CoMFA, the green contours represent regions where bulky groups enhance the biological activity, while bulky groups decrease activity in contact with yellow regions. Furthermore, in the electrostatic fields CoMFA, blue contours indicate regions where electropositive groups enhance the activity and red contours (not present) indicate the regions where electronegative groups increase the activity.

There are 3 yellow contours around the R1 substituent of the quinolone nucleus indicating that bulky groups in contact with these contours decrease the antimalarial activity. In Fig. 6b, compound **18**, a low activity compound, has a benzyl group oriented to the yellow contour. The same trend was found in other compounds with lower antimalarial activity, such as **20** (R1 = Phenyl), **22** and **24**. Moreover, the compounds **7**, **8**, **37**, **38**, **45** (see Fig. 6a) and **47**, which have long alkyl chains in R1, have high biological activity. This apparent contradiction which is due to the linear alkyl chains can be best accommodated in the steric "tunnel" formed by the three yellow contours owing to large conformational freedom of these chains (see Fig. 6a). Making a superposition of the steric contours obtained from CoMFA with the lipophilic potential surface of the Qo site in cytochrome bc_1 calculated by MOLCAD module in SYBYL (see Fig. 7a), the steric "tunnel" is composed by: I125, A126, F129, L130, I147, L165, F179 and F296. This cavity is deep and narrow; therefore, long linear alkyl substituents in R1 achieve better accommodation in the hydrophobic cavity than substituents with

Table 9
Experimental and predicted pEC_{50} and residuals ($pEC_{50Exp}-pEC_{50Pred}$) for the training and test sets using the best CoMFA and CoMSIA models.

Compounds	pEC_{50}	CoMFA		CoMSIA	
		Predicted	Residual	Predicted	Residual
2	7.031	6.785	0.246	7.030	0.001
5	6.148	6.494	0.346	6.156	0.008
6	6.344	6.209	0.135	6.038	0.306
7	7.889	7.914	0.025	7.348	0.541
8	7.402	7.452	0.050	7.488	0.086
9	6.992	7.011	0.019	7.218	0.226
10	5.337	5.296	0.041	5.335	0.002
11	5.523	5.735	0.212	5.859	0.336
12	6.541	6.190	0.351	6.149	0.392
13	5.853	6.040	0.187	6.233	0.380
14	6.073	6.054	0.019	6.155	0.082
15	6.032	6.002	0.030	6.013	0.019
16	5.082	5.150	0.068	5.221	0.139
24	5.607	5.648	0.041	5.944	0.337
25	6.613	6.385	0.228	6.263	0.350
21	5.686	5.701	0.015	5.793	0.107
30	5.668	5.975	0.307	5.793	0.125
28	6.245	6.188	0.057	6.186	0.059
29	6.181	6.416	0.235	6.239	0.058
18	5.191	5.221	0.030	5.222	0.031
22	5.664	5.654	0.010	5.919	0.255
23	6.201	5.953	0.248	5.809	0.392
19	6.542	6.521	0.021	6.598	0.056
20	5.089	5.143	0.054	5.127	0.038
26	6.666	6.399	0.267	6.244	0.422
31	6.777	6.715	0.062	6.582	0.195
32	7.582	7.401	0.181	7.038	0.544
33	7.100	7.057	0.043	6.842	0.258
34	6.434	6.449	0.015	6.654	0.220
36	6.747	6.859	0.112	6.915	0.168
37	7.910	8.144	0.234	8.232	0.322
38	8.220	8.204	0.016	8.218	0.002
40	7.028	7.168	0.140	7.056	0.028
41	7.247	7.258	0.011	7.293	0.046
42	6.951	6.940	0.011	6.992	0.041
43	7.063	7.040	0.023	7.197	0.134
45	8.666	8.679	0.013	8.403	0.263
46	7.405	7.287	0.118	7.565	0.160
47	7.502	7.537	0.035	7.836	0.334
48	7.210	7.168	0.042	7.241	0.031
1[a]	8.066	7.314	0.752	7.812	0.254
3[a]	7.618	6.846	0.772	7.016	0.602
4[a]	6.229	5.961	0.268	6.117	0.112
17[a]	6.165	6.303	0.138	6.616	0.451
27[a]	5.296	6.080	0.784	5.689	0.393
35[a]	7.317	7.181	0.136	7.113	0.204
39[a]	6.625	6.932	0.307	7.040	0.415
44[a]	8.498	8.082	0.416	7.254	1.244

[a] Compounds in the test set.

less conformational freedom (as benzyl and phenyl substituents) since it can cause repulsion within the cavity.

A green contour around the substituent R4 of quinolone ring shows that bulky groups in this position improves the biological activity; for example, compounds **37**, **38**, **45** (see Fig. 6a) and **47** have a methoxyl group in R4 and present higher biological activities than compounds **16**, **18** (see Fig. 6b), and **20** which have only hydrogen in position R4. In Fig. 7a, the substituent R4 is in an accessory pocket within the Qo site of the complex bc_1 where bulky groups in that position can increase the number of interactions with the cavity, and therefore, the biological activity.

Regarding electrostatic CoMFA maps, a large blue contour is found near the R4 substituent; therefore, positive charges in R4 increase the antimalarial activity. The most active compounds in study (**37**, **38** and **45**, see Fig. 6a) have methoxyl groups in this position which have partially positive charged hydrogen by electrodonor effect (inductive) from methoxyl oxygen to the aromatic ring of the quinolone. The compounds with hydrogen in position R4 have low antimalarial activity (as **18**, see Fig. 6b).

Fig. 7b shows the superposition of the electrostatics contours obtained from CoMFA with the electrostatic potential surface of the Qo site in cytochrome bc_1. The large blue contour is found in a cavity composed by the carbonyl groups of residues M139, S140, P179, C180, I269 and the side chain of Y279. This cavity rich in negative partial charges (negatively polarized oxygen) can interact favorably with partial positive charges (hydrogens) in the methoxyl group of the most active ligands.

Another blue contour was found near the yellow contours forming the steric "tunnel" around the R1 substituent (see Fig. 6a and b). The "tunnel" is mainly aromatic character (rich in phenylalanine residues), thus, positive charges in the R1 substituent of the ligands into the "tunnel" could lead cation–π interactions, and thus, enhance the antimalarial activity of the ligands.

3.2.3. Graphical Interpretation of CoMSIA model

Fig. 8a shows the distribution of the steric fields in the CoMSIA model. Compared with Fig. 6a, it is possible to find great

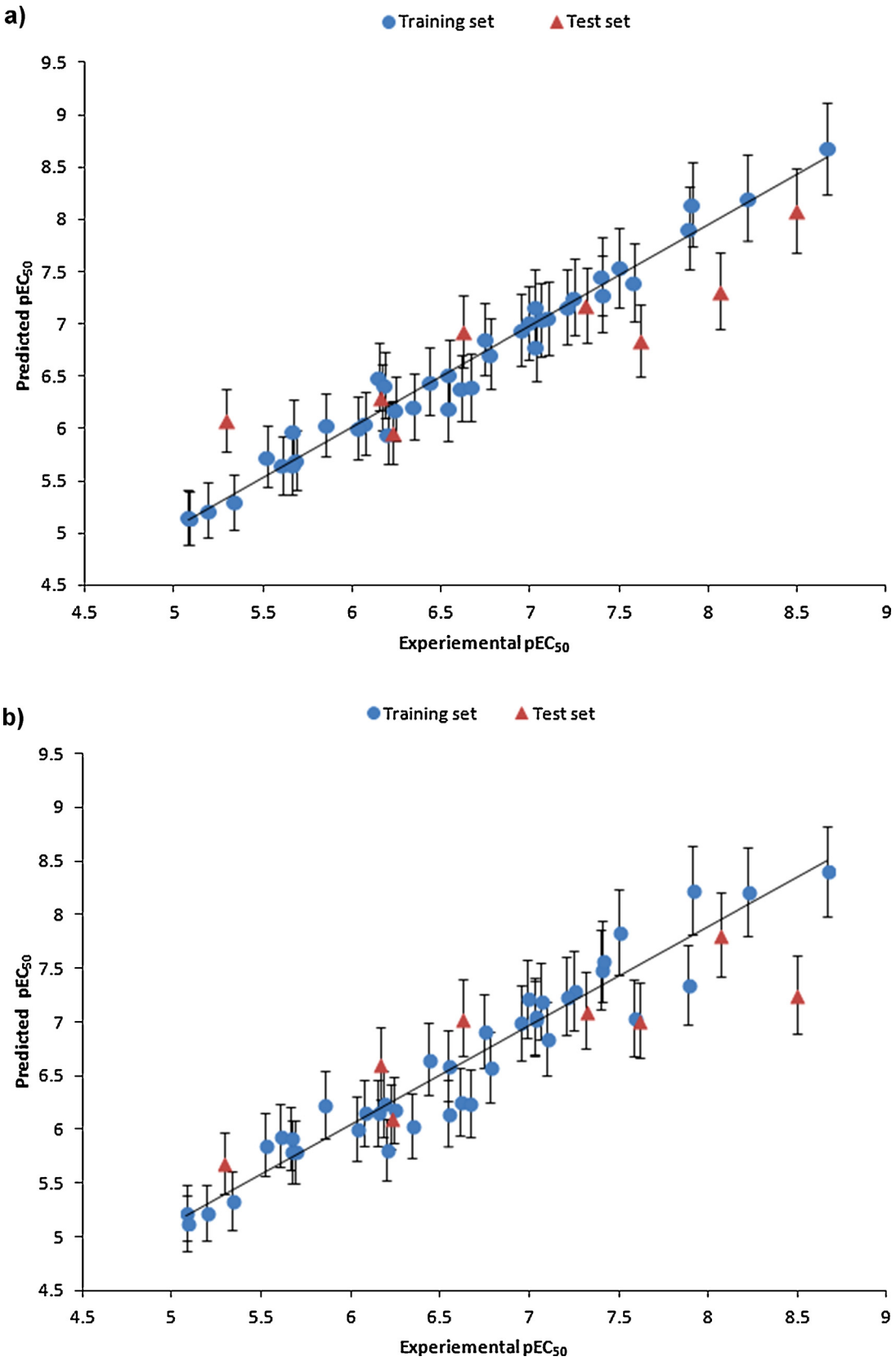


Fig. 5. Correlation between predicted and experimental activities of training (dots) and test (triangles) sets for the CoMFA (a) and CoMSIA (b) models.

similarity between the CoMSIA steric maps with those of CoMFA. However, some differences are set out in the CoMSIA steric fields. First, the green contour on the R4 substituent present in CoMFA fields has disappeared. Also, a new green contour appeared inside the yellow contours on the R1 substituent of the quinolone ring.

This sterically favorable contour confirms that compounds with bulky groups of linear alkyl chains in R1 position have high

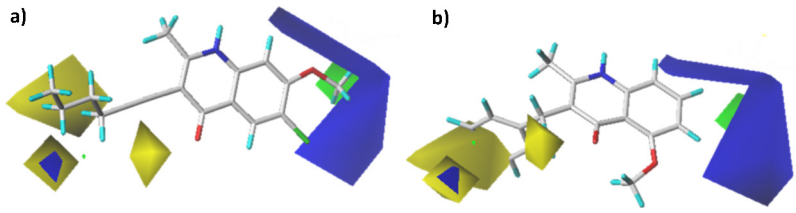


Fig. 6. Steric and electrostatic contour maps CoMFA using **45**(a) and **18** (b) as models.

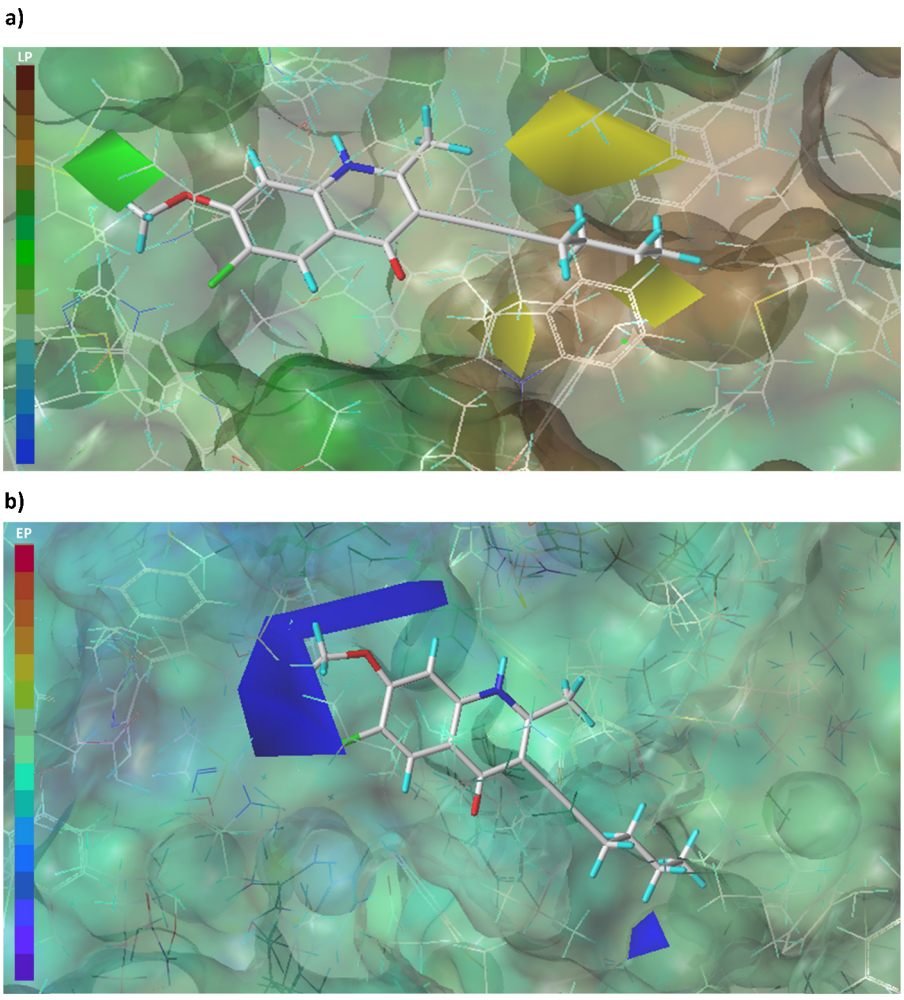


Fig. 7. Steric (a) and electrostatic (b) contour maps CoMFA within the binding site of cytochrome bc_1. Lipophilic surfaces and electrostatic potential of the enzyme was calculated using the Sybyl MOLCAD module.

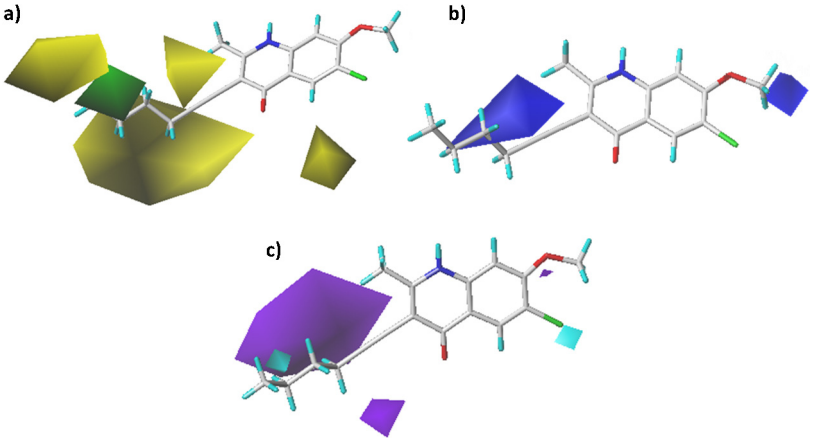


Fig. 8. Steric contour maps (a), electrostatic (b) and hydrophobic (c) using as a model CoMSIA **45** compound.

antimalarial potency due to a good interaction between R1 substituent and the steric "tunnel". Finally, a new yellow contour around the R2 substituent is shown in the CoMSIA model, whereby, bulky groups in this position will decrease the biological activity. This tendency is observed in the compounds with low potency such as: **18** (R2 = —OCH3), **20** (R2 = —OCH3) and **27** (R2 = —Cl), while compounds with high biological activity (**32**, **37**, **38** and **45**) have hydrogen at this position.

The CoMSIA electrostatic fields (Fig. 8b) maintain the same contour distributions as CoMFA as shown in Fig. 5a. However, from the quantitative point of view the CoMSIA model has a major contribution of the electrostatic field (39.4%) than CoMFA model (31.7%).

The hydrophobic contour map of the CoMSIA model is displayed in Fig. 8c. In the CoMSIA hydrophobic field, cyan contours represent favorable hydrophobic regions while purple contours represent favorable hydrophilic regions.

These contours groups have an arrangement similar to the steric contours CoMSIA previously analyzed where two purple polyhedrons around the R1 substituent (located in a similar position to the sterically unfavorable yellow field) on the quinolone ring indicate that hydrophobic groups in this region are unfavorable for antimalarial activity. In addition, a cyan contour (located in a similar position to the sterically favorable green polyhedron), favorable for hydrophobic groups inside the white contours, is beneficial for antimalarial activities (see Fig. 8c).

3.3. Molecular docking analysis

It is well known that 4(1*H*)-quinolones derivatives inhibit the plasmodial cytochrome bc_1. These compounds have structural similarity to other cytochrome bc_1 inhibitors such as atovaquone and stigmatellin; therefore, we expect quinolone derivatives to bind to the same binding site of these classical inhibitors. Because the molecular docking methodology can provide a greater understanding of the ligand–protein interactions, all derivatives of 4(1*H*)-quinolones in this study were docked in the Qo site of cytochrome bc_1.

Before docking, the reliability of the docking procedure was first validated using two approaches. In the first procedure, the co-crystallized stigmatellin was extracted from the complex and redocked to the binding site of cytochrome bc_1. In the second procedure, the initial 3D structure of stigmatellin was built from a 2D representation, then this conformation was optimized using the PM6 routine and docked to the binding site of cytochrome bc_1. This last approach follows the workflow of the docking experiment used to the 4(1*H*)-quinolone compounds in study. As a result (see Fig. 9), the root–mean–square deviation (RMSD) between the crystallographic conformation (green) and the redocked conformation of stigmatellin (red) was 0.75 Å and the RMSD between the crystallographic conformation and the PM6-optimized conformation of stigmatellin (blue) was 0.53 Å. Both procedures suggest an acceptable accuracy for the docking procedure, in special the second approach. These parameters of molecular docking can be extended to search for the binding conformations to Qo site for other cytochrome bc_1 inhibitors.

Fig. 10a–c shows the docking results of compounds **2**, **3** and **45**, respectively. They are the most potent antimalarial compounds in study and exhibit similar orientations to the crystalline ligand (see Fig. 10d). The three compounds have a long linear alkyl chain in R1, which interacts with a strongly hydrophobic "tunnel" formed by the amino acids I125, A126, F129, L130, I147, F151, L165, F179, F278 and F296. An important hydrogen bond is formed between HN of the quinolone ring and the side chain of the H181, which in turn is bonded to the 2Fe-2S cluster important for electron transport [10]. These results of molecular docking are in agreement with the CoMFA and CoMSIA results.

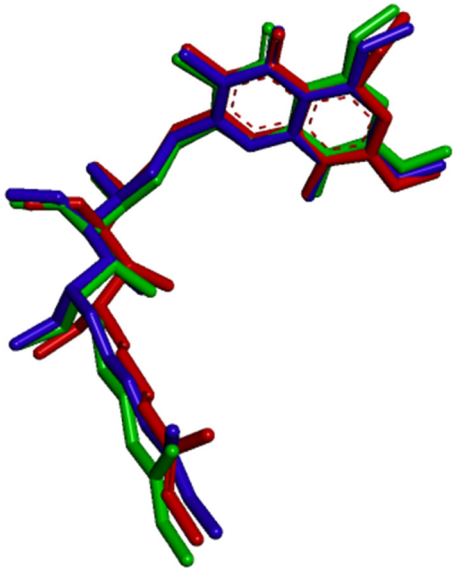


Fig. 9. Superimposition of the crystallographic (green), redocked (red) and PM6-optimized (blue) conformations of the stigmatellin ligand. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Fig. 11a and b displays the docking of compounds **13** and **18**, respectively. These compounds do not have a long alkyl chain in R1 and do not interact effectively with the hydrophobic "tunnel" which would explain their low antimalarial activity. The binding conformations achieved by molecular docking for these compounds significantly differ from the crystallographic conformation of stigmatellin. In addition, these binding conformations for **13** and **18** do not include hydrogen bonding interaction with H181. HOH7187 was consistently toggled off by GOLD for all 48 ligands in this study. Therefore, the presence of this water did not significantly influence the binding conformations of this kind of ligands in the Qo site of cytochrome bc_1.

3.4. Design of new 4(1H)-quinolones compounds

Based on the above discussion, we observed the R1 and R4 substituents play an important role in the antimalarial activity of the studied compounds. Then, in order to design new compounds with higher inhibitory activities of cytochrome bc_1, we propose maintaining a long alkyl chain in R1 and put positively charged groups on both R1 and R4. Bulky groups on R4 also potentiate the antimalarial activity. Compound **45** was used as template to initiate structural changes for new compounds because it is the most active compound of the 48 studied quinolones. The new compounds were designed based on information obtained from CoMFA and CoMSIA models, due to the ease implementing structural changes based on the results of these 3D-QSAR models.

The first modification to **45** was the introduction of bulky and positively charged groups (rings) and R4. For example, **TJ1** and **TJ3** compounds (see Table 10 and Fig. 12) were obtained by replacing hydrogen of the methoxyl group in R4 with the imidazole and piperidine rings, respectively. These compounds showed greater potency than the compound **45**. However, by putting a second positively charged group in the substituent R1, the biological activity was enhanced even more (see Table 10 and Fig. 12), as evidence by the **TJ2** and **TJ4** compounds. In order to enhance the ability to form hydrogen bonds of the new compounds designed, bioisosteric changes were developed in the quinolone nucleus (introduction of polar atoms in the quinolone scaffold).

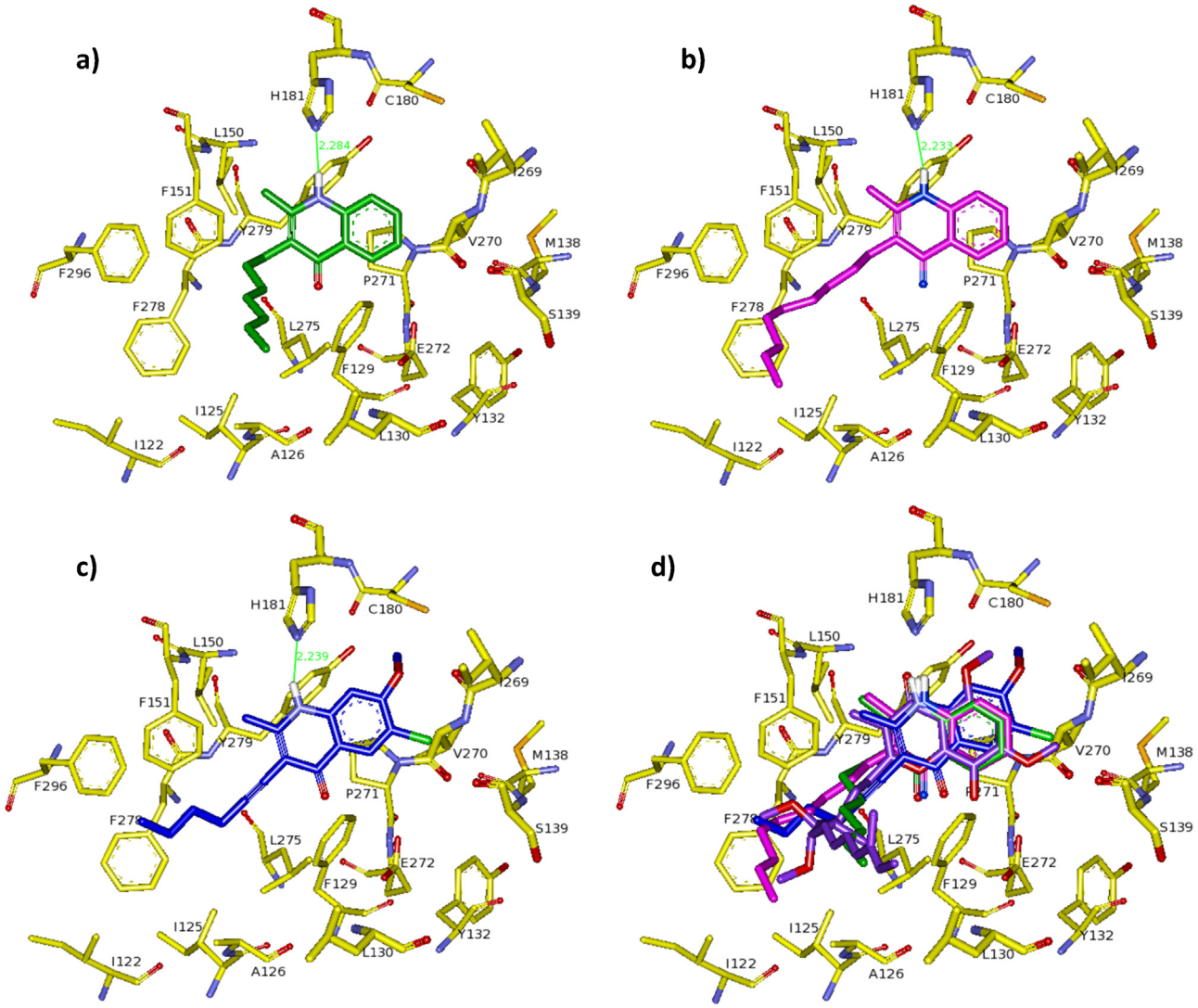


Fig. 10. Poses obtained from molecular docking for **2** (a), **3** (b) and **45** (c). Superimposition of **2** (green), **3** (pink), **45** (blue) and stigmatellin (purple) in the binding site of the cytochrome bc_1 (d). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

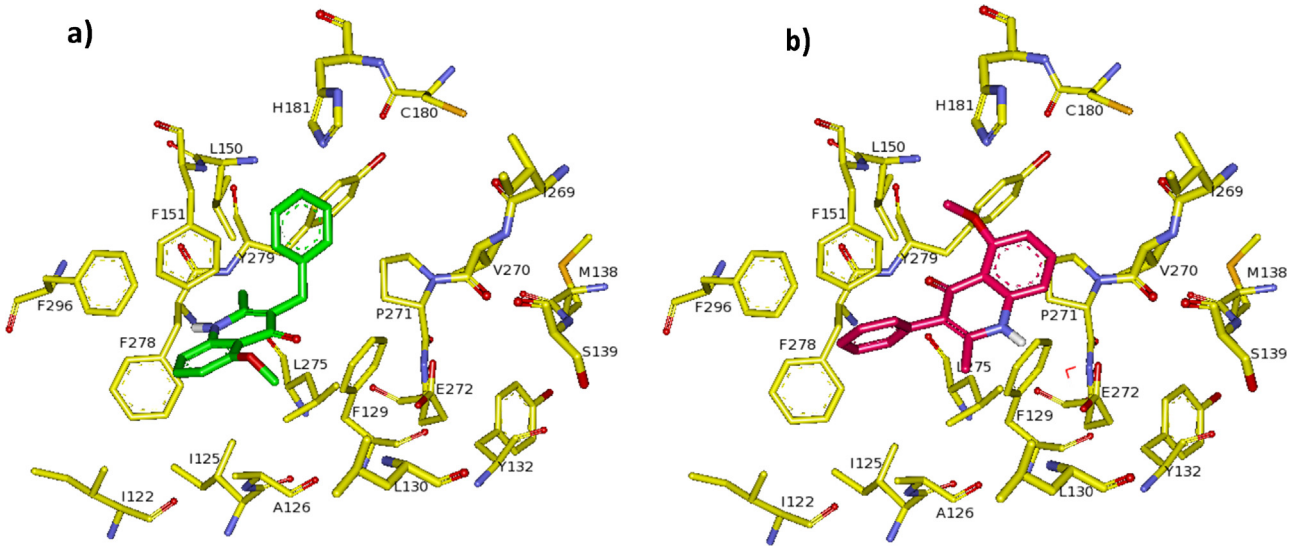


Fig. 11. Poses obtained from molecular docking for **18**(a) and **20** (b) in the binding site of the cytochrome bc_1.

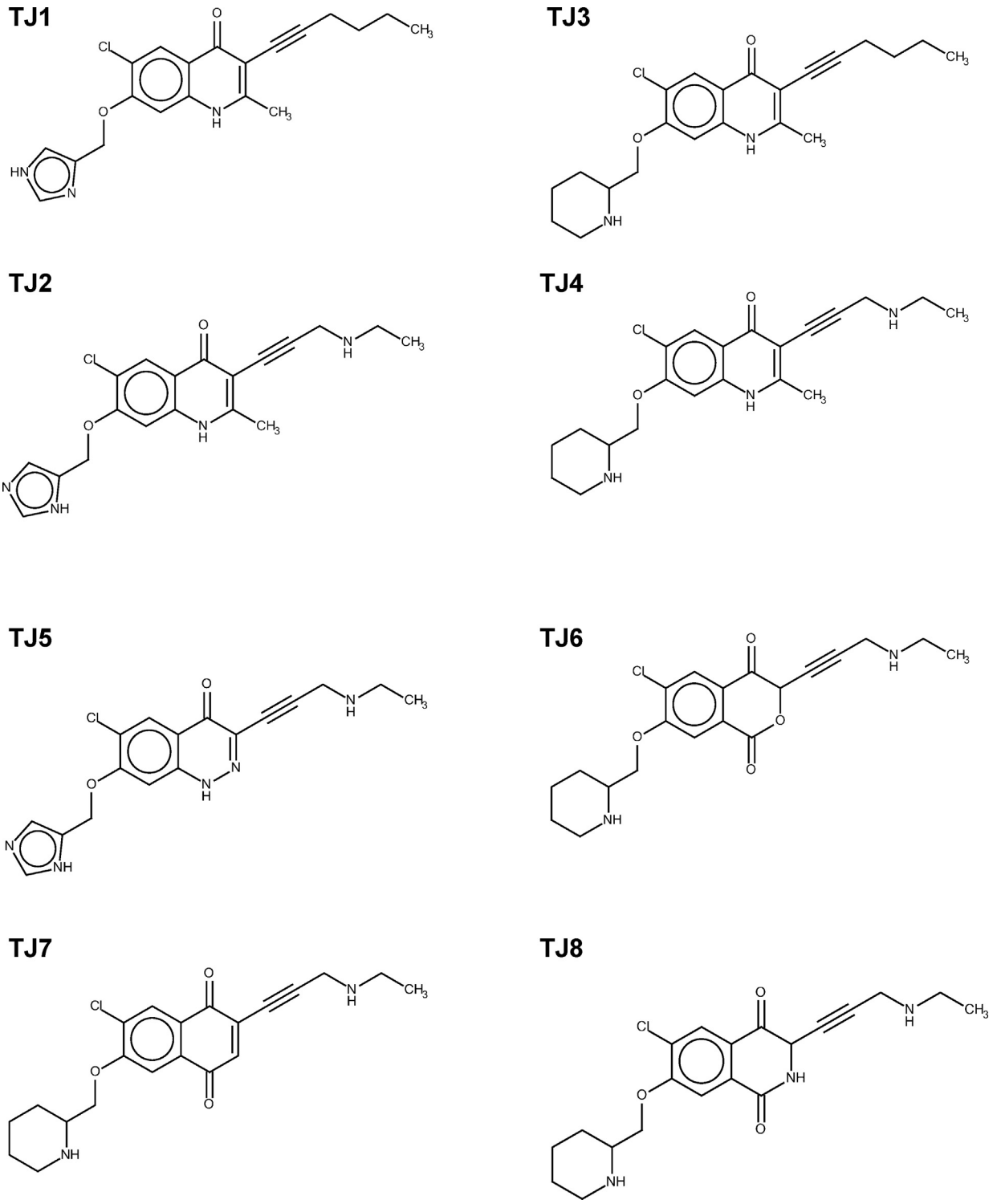


Fig. 12. Schematic structures of compounds designed **TJ1**–**TJ8**.

Initially, the carbon in position 2 with sp^2 hybridized, was replaced by nitrogen with sp^2 hybridized in **TJ2** to obtain **TJ5**. This compound had the highest predicted antimalarial activity of compounds designed for CoMFA and CoMSIA models (see Table 10 and Fig. 12).

Subsequently, the carbon 2 was replaced by oxygen and nitrogen with sp^3 hybridization in **TJ4** and the compounds **TJ6** and **TJ8** were obtained, respectively. Furthermore, the nitrogen of the quinolone ring was replaced by a carbonyl group in order to maintain the ability to form hydrogen bonding interactions in this region of the molecule. The 3D-QSAR models predict high antimalarial activity for these two new compounds.

Another transformation proposed was the conversion of the quinolone to quinone nucleus by replacing the amino group with a carbonyl group, subsequently, **TJ7** was obtained from **TJ4**. This compound also presents higher values of pEC_{50} predicted than **45** for both QSAR models.

Table 10
Predicted pEC_{50} and Rat IV LD_{50} values of the designed compounds.

Compounds	pEC_{50} predicted		Rat IV LD_{50} (mg/Kg)	Rat IV LD_{50} classification[a]
	CoMFA	CoMSIA		
TJ1	9.27202	9.28708	58.23	Class 4
TJ2	10.1993	8.83006	61.35	Class 4
TJ3	9.19638	9.13421	27.02	Class 3
TJ4	10.3334	8.70818	45.04	Class 4
TJ5	10.4458	10.85532	96.23	Class 4
TJ6	9.92591	9.59155	39.32	Class 3
TJ7	9.64801	9.5224	33.14	Class 3
TJ8	9.90446	9.60711	58.95	Class 4

[a] Acute Rodent Toxicity Classification of Chemicals by OECD Project.

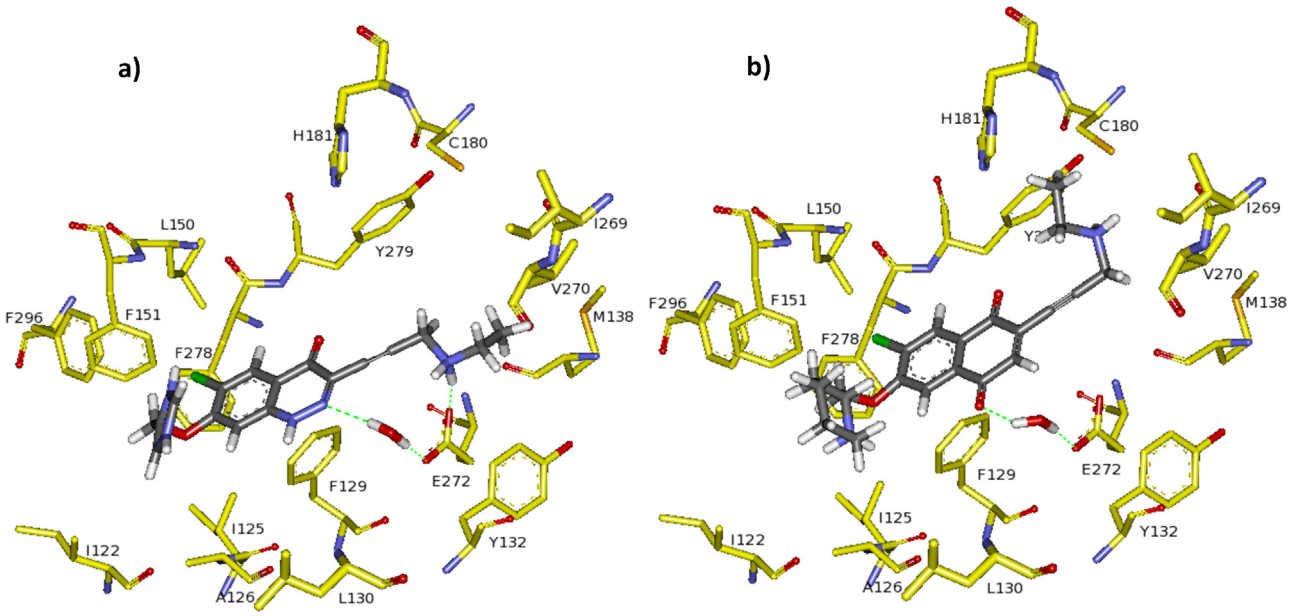


Fig. 13. Poses obtained by molecular docking of designed compounds **TJ5** (a) and **TJ7** (b) in the binding site of the cytochrome bc_1.

In addition to antimalarial activity predictions, we developed acute toxicological predictions for new compounds proposed in order to obtain a general outline about the toxicological profile that these compounds would have. Thus, *in silico* predictions of Lethal Doses 50 (LD_{50}) values for rats using intravenous (IV) administration were calculated using the free-server GUSAR-online [46]. Considering the acute rodent toxicity classification of chemicals by OECD project [46], our predictions (see Table 10) emphasize that designed compounds are considered as slightly toxic (class III) and practically non-toxic (class IV). Current antimalarial drugs such as atovaquone and proguanil hydrochloride present experimental rat IV LD_{50} (mg/kg) values of 40 and 36 [47], respectively, while the 8 new compounds present predicted rat IV LD_{50} values in the range of 27 and 97 (see Table 10). This show that the designed compounds (which exhibit functional groups very reactive, e.g. **TJ7** is a Michael acceptor) would have a toxicological profile similar to these drugs, which could be considered as safe.

Finally, the molecular docking procedure initially described was used to get the possible binding conformations of the compounds designed. Fig. 13a shows the best docked conformation predicted for **TJ5**. The imidazole group in R4 was found in the hydrophobic "tunnel" and the R1 substituent was situated in the cavity in which the substituent R4 was originally placed in the studied compounds. This exchange of orientation is favored because the imidazole group is both aromatic as positively charged favoring π–π interactions as well as cation–π interactions with phenylalanine residues in the hydrophobic "tunnel" (there are many phenylalanine residues in the hydrophobic "tunnel"). Replacing the C-2 with nitrogen in the aromatic ring generated a new hydrogen bond interaction with HOH7187 (toggled on) which also generates a hydrogen bond with the E272. This amino acid is important to electron transport in the cytochrome bc_1 [10]. A second hydrogen bond is formed between the E272 and the HN atom in R1.

The docked conformation for **TJ7** is displayed in Fig. 13b. **TJ7** also exhibits hydrogen bonding interaction with E272 through the intermediate HOH7187. This interaction was possible due to the bioisosteric change from amino to carbonyl group in the quinolone ring. The R1 and R4 substituents keep a similar accommodation in the binding site such as **TJ5**.

4. Conclusions

A theoretical study on a series of 4(1*H*)-quinolone acting as Qo site inhibitors at the cytochrome bc_1 was carried out using 2D-QSAR, CoMFA/CoMSIA methods and docking analysis. 2D-QSAR final models included electrostatic (Q2, Q5 and dipole moment), reactivity (electrophilicity index), physicochemical (Log*P*, GGUT_SLOGP_0 and SLogP_VSA_5) and topological (KierFlex) properties and showed a satisfactory statistical quality and predictive abilities as show the r^2, q^2 and R^2 values for both 2D-QSAR models. Meanwhile, the 3D-QSAR models generated indicating that the steric, electrostatic (in CoMFA) and hydrophobic (in CoMSIA) fields all have importance influences on the ligand–receptor interaction. The predictive ability of the CoMFA

(q^2 = 0.635 and R^2 = 0.817) and CoMSIA (q^2 = 0.537 and R^2 = 0.775) models was also as good as the 2D-QSAR models. Using molecular docking the appropriate binding conformations of these series of compounds interacting with cytochrome bc_1 was revealed.

The results obtained from the established models, 2D, 3D-QSAR, and docking analysis, show the following: (1) GCUT_SLOGP_0, Q5, Log *P*, SLogP_VSA_5, KierFlex, dipole moment, electrophilicity index and Q2 are the most important parameters to explain the antimalarial activity of the studied compounds. (2) Bulky groups in R4 position improve the biological activity, while long linear alkyl chains in the R1 chain further enhance the biological activity. (3) Positively charged groups in R1 and R4 increase the antimalarial activity. (4) The molecular docking highlights the hydrophobic interactions between the long alkyl chain linear in R1 and the hydrophobic "tunnel" formed by the amino acids I125, A126, F129, L130, I147, F151, L165, F179, F278 and F296. An important hydrogen bond interaction is formed between the HN quinolone ring and the side chain of H181, which is important for electron transport inside the cytochrome bc_1.

Based on the results obtained from the models, eight new compounds were designed by bioisosteric replacement on the **45** compound. From the set of designed compounds, **TJ5** is the most prominent antimalarial compound by presenting the highest pEC_{50} values predicted by the 3D-QSAR models and an attractive binding conformation predicted by molecular docking.

Acknowledgements

The authors wish to thank the program to support research groups, sponsored by the Vice-Rectory for research of the University of Cartagena from 2011 to 2012, Cartagena (Colombia). The authors also wish to thank Joseph Dunn at Writing Center at the University of Cartagena.

References

[1] WHO, World Malaria Report, 2012.
[2] M.C. Murray, M.E. Perkins, A.B. James, Chemotherapy of malaria Annual Reports in Medicinal Chemistry, vol. 31, Academic Press, Waltham, Massachusetts, 1996, pp. 141–150.
[3] A.V.S. Hill, Vaccines against malaria, Philosophical Transactions of the Royal Society B: Biological Sciences 366 (2011) 2806–2814.
[4] J.J. Kessl, K.H. Ha, A.K. Merritt, B.B. Lange, P. Hill, B. Meunier, et al., Cytochrome b mutations that modify the ubiquinol-binding pocket of the cytochrome bc1 complex and confer anti-malarial drug resistance in *Saccharomyces cerevisiae*, Journal of Biological Chemistry 280 (2005) 17142–17148.
[5] H.J. Painter, J.M. Morrisey, M.W. Mather, A.B. Vaidya, Specific role of mitochondrial electron transport in blood-stage *Plasmodium falciparum*, Nature 446 (2007) 88–91.
[6] S. Looareesuwan, J.D. Chulay, C.J. Canfield, D.B. Hutchinson, Malarone (atovaquone and proguanil hydrochloride): a review of its clinical development for treatment of malaria. Malarone Clinical Trials Study Group, American Journal of Tropical Medicine and Hygiene 60 (1999) 533–541.
[7] J.J. Kessl, B.B. Lange, T. Merbitz-Zahradnik, K. Zwicker, P. Hill, B. Meunier, et al., Molecular basis for atovaquone binding to the cytochrome bc_1 complex, Journal of Biological Chemistry 278 (2003) 31312–31318.
[8] N. Fisher, P.G. Bray, S.A. Ward, G.A. Biagini, The malaria parasite type II NADH: quinone oxidoreductase: an alternative enzyme for an alternative lifestyle, Trends in Parasitology 23 (2007) 305–310.
[9] M.W. Mather, E. Darrouzet, M. Valkova-Valchanova, J.W. Cooley, M.T. McIntosh, F. Daldal, et al., Uncovering the molecular mode of action of the antimalarial drug atovaquone using a bacterial system, Journal of Biological Chemistry 280 (2005) 27458–27465.
[10] T. Rodrigues, F. Lopes, R. Moreira, Inhibitors of the mitochondrial electron transport chain and de novo pyrimidine biosynthesis as antimalarials: the present status, Current Medicinal Chemistry 17 (2010) 929–956.
[11] W. Salzer, H. Timmler, H. Andersag, A new type of compounds active against avian malaria, Chemische Berichte 81 (1948) 12–19.
[12] R.M. Cross, A. Monastyrskyi, T.S. Mutka, J.N. Burrows, D.E. Kyle, R. Manetsch, Endochin optimization: structure activity and structure property relationship studies of 3-substituted 2-methyl-4(1H)-quinolones with antimalarial activity, Journal of Medicinal Chemistry 53 (2010) 7076–7094.
[13] R.M. Cross, N.K. Namelikonda, T.S. Mutka, L. Luong, D.E. Kyle, R. Manetsch, Synthesis, antimalarial activity, and structure–activity relationship of 7-(2-phenoxyethoxy)-4(1H)-quinolones, Journal of Medicinal Chemistry 54 (2011) 8321–8327.
[14] Y. Zhang, J.A. Clark, M.C. Connelly, F. Zhu, J. Min, W.A. Guiguemde, et al., Lead optimization of 3-carboxyl-4(1H)-quinolones to deliver orally bioavailable antimalarials, Journal of Medicinal Chemistry 55 (2012) 4205–4219.
[15] R. Cowley, S. Leung, N. Fisher, M. Al-Helal, N.G. Berry, A.S. Lawrenson, et al., The development of quinolone esters as novel antimalarial agents targeting the *Plasmodium falciparum* bc1 protein complex, MedChemComm 3 (2012) 39–44.
[16] R.W. Winter, J.X. Kelly, M.J. Smilkstein, R. Dodean, D. Hinrichs, M.K. Riscoe, Antimalarial quinolones: synthesis, potency, and mechanistic studies, Experimental Parasitology 118 (2008) 487–497.
[17] A.B. Richon, S.S. Young, An introduction to QSAR Methodology, Network Science Corporation, Saluda, NC, 1997.
[18] R. Todeschini, V. Consonni, P. Gramatica, Chemometrics in QSAR. Comprehensive Chemometrics, Chemical and Biochemical Data Analysis 4 (2009) 129–172.
[19] J. Verma, V.M. Khedkar, E.C. Coutinho, 3D-QSAR in drug design-a review, Current Topics in Medicinal Chemistry 10 (2010) 95–115.
[20] Y.-P. Li, X. Weng, F.-X. Ning, J.-B. Ou, J.-Q. Hou, H.-B. Luo, et al., 3D-QSAR studies of azaoxoisoaporphine, oxoaporphine, and oxoisoaporphine derivatives as anti-AChE and anti-AD agents by the CoMFA method, Journal of Molecular Graphics and Modelling 41 (2013) 61–67.
[21] A. Politi, S. Durdagi, P. Moutevelis-Minakakis, G. Kokotos, M.G. Papadopoulos, T. Mavromoustakos, Application of 3D QSAR CoMFA/CoMSIA and in silico docking studies on novel renin inhibitors against cardiovascular diseases, European Journal of Medicinal Chemistry 44 (2009) 3703–3711.
[22] S. Ekins, J. Mestres, B. Testa, In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling, British Journal of Pharmacology 152 (2007) 9–20.
[23] S. Jain, M. Ghate, K. Bhadoriya, S. Bari, A. Chaudhari, J. Borse, 2D, 3D-QSAR and docking studies of 1,2,3-thiadiazole thioacetanilides analogues as potent HIV-1 non-nucleoside reverse transcriptase inhibitors, Organic and Medicinal Chemistry Letters 2 (2012) 1–13.
[24] I. Halperin, B. Ma, H. Wolfson, R. Nussinov, Principles of docking: an overview of search algorithms and a guide to scoring functions, Proteins: Structure, Function, and Bioinformatics 47 (2002) 409–443.
[25] Marvin Sketch was used for drawing, displaying, and characterizing chemical structures, MarvinSketch 5.12.2, 2012, ChemAxon (http://www.chemaxon.com).
[26] MOPAC2009, J.J.P. Stewart, Stewart Computational Chemistry, 2008, Colorado Springs, CO, USA, http://OpenMOPAC.net
[27] T. Koopmans, Ordering of wave functions and eigenenergies to the individual electrons of an atom, Physica 1 (1933) 104–113.
[28] R. Wang, Y. Gao, L. Lai, Calculating partition coefficient by atom-additive method, Perspectives in Drug Discovery and Design 19 (2000) 47–66.
[29] Molecular Operating Environment (MOE), 2009.10; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2009.
[30] D.J. Wild, MINITAB Release 14, Journal of Chemical Information and Modeling 45 (2004) 212–220.
[31] D.B. De Oliveira, A.C. Gaudio, BuildQSAR: A New Computer Program for QSAR Analysis, Quantitative Structure–Activity Relationships 19 (2000) 599–601.
[32] A. Tropsha, Best practices for QSAR model development, validation, and exploitation, Molecular Informatics 29 (2010) 476–488.
[33] A. Tropsha, P. Gramatica, V.K. Gombar, The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, QSAR & Combinatorial Science 22 (2003) 69–77.
[34] P.P. Roy, K. Roy, On some aspects of variable selection for partial least squares regression models, QSAR & Combinatorial Science 27 (2008) 302–313.
[35] SYBYL-X 1.3, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.
[36] G. Jones, P. Willett, R.C. Glen, Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation, Journal of Molecular Biology 245 (1995) 43–53.
[37] C. Lange, C. Hunte, Crystal structure of the yeast cytochrome bc1 complex with its bound substrate cytochrome c, Proceedings of the National Academy of Sciences 99 (2002) 2800–2805.
[38] T. Rodrigues, R. Moreira, J. Gut, P.J. Rosenthal, P.M. ONeill, G.A. Biagini, et al., Identification of new antimalarial leads by use of virtual screening against cytochrome bc1, Bioorganic & Medicinal Chemistry 19 (2011) 6302–6308.
[39] G.A. Biagini, N. Fisher, A.E. Shone, M.A. Mubaraki, A. Srivastava, A. Hill, et al., Generation of quinolone antimalarials targeting the *Plasmodium falciparum* mitochondrial respiratory chain for the treatment and prophylaxis of malaria, Proceedings of the National Academy of Sciences 109 (2011) 8298–8303.
[40] G. Jones, P. Willett, R.C. Glen, A.R. Leach, R. Taylor, Development and validation of a genetic algorithm for flexible docking, Journal of Molecular Biology 267 (1997) 727–748.
[41] S.A. Wildman, G.M. Crippen, Prediction of physicochemical parameters by atomic contributions, Journal of Chemical Information and Computer Sciences 39 (1999) 868–873.
[42] R.G. Parr, L.S. Szentpaly, S. Liu, Electrophilicity index, Journal of the American Chemical Society 121 (1999) 1922–1924.
[43] R. Parthasarathi, V. Subramanian, D. Roy, P. Chattaraj, Electrophilicity index as a possible descriptor of biological activity, Bioorganic & Medicinal Chemistry 12 (2004) 5533–5543.

[44] The molecular connectivity chi indexes and kappa shape indexes in structure–property modeling, in: L.H. Hall, L.B. Kier, K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry, vol. 2, 2007, pp. 367–422.
[45] M. Karelson, V.S. Lobanov, A.R. Katritzky, Quantum-chemical descriptors in QSAR/QSPR studies, Chemical Reviews 96 (1996) 1027–1044.
[46] A. Lagunin, A. Zakharov, D. Filimonov, V. Poroikov, QSAR modelling of rat acute toxicity on the basis of PASS prediction, Molecular Informatics 30 (2011) 241–250.
[47] GlaxoSmithKline, Malarone: Atovaquone/Proguanil Hydrochloride, GlaxoSmithKline, United Kingdom, 2012 (product monograph).