COMPUTATIONAL MODELING OF ARTEMISININ AND ITS STRUCTURAL DERIVATIVES - INTERACTION, ACTIVITY AND MECHANISM

Synopsis

of

Ph. D Thesis

Submitted by

Mani Srivastava

Enrollment Number: 071755

Jaypee University of Information Technology
Waknaghat, P.O. Dumehar Bani, Kandaghat, Solan-173215, H.P.

November 2009

Introduction

Artemisinins (the plant extract from *Artemisia annua*) are the most potent antimalarials available, rapidly killing all asexual stages of *P. falciparum*. Several studies showed artemisinin to be an exceptional antimalarial agent with negligible toxicity and high efficacy against human malaria parasites, including those resistant to conventional antimalarials (1). Thus it has been the objective of numerous studies to prepare better and safer anti-malarial drugs. Artemisinin is structurally different from the previously known antimalarials. The compound is an unusually stable sesquiterpene lactone with an endoperoxide ring (empirical formula C15H22O5) (Figure 1). Presence of the endoperoxide moiety is the key to its antimalarial activity (2-5). The white needle crystals of artemisinin are hardly soluble in water or oil. However, since the peroxide bridge is stable under certain chemical reactions, several more soluble artemisinin derivatives, arteether, artemether, sodium artesunate, sodium artelinate and dihydroartemisinin (DHQ) have been synthesized for the treatment of malaria (Figure 1). Over the year there has been considerable increase in the number of artemisinin analogues. In a number of experimental studies during the past decade, a large number of analogues have been designed, synthesized and evaluated in order to prepare new, more potent and less toxic analogues with better therapeutic indices. These studies provided important insights into the development of the structure-activity relationship models between analogues.

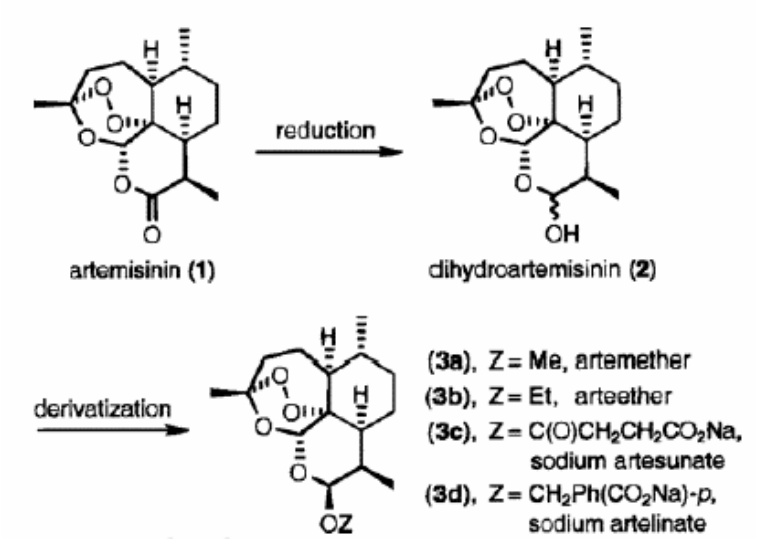


Figure 1.3. Artemisinin and its derivatives.

The mode of action of these antimalarials is not fully understood. Artemisinin act via mechanisms that are distinct from other antimalarial classes, including those that inhibit well defined targets such as enzymes of folate biosynthesis, the DOXP reductase pathway or the cytochrome electron transport system. The peroxide within the 1,2,4-trioxane system of artemisinin

is essential for antimalarial activity. Therefore, the peroxide structure becomes a focus for considerable chemical analysis aimed at trying to understand how artemisinin work. Two mechanism of action of artemisinin have been proposed. One the activated artemisinin form adducts with heme and leads to inhibition of heme polymerization. Secondly as shown recently, artemisinin, but not quinine or chloroquine, inhibit the sarco/endoplasmic reticulum Ca2+ ATPase (SERCA) orthologue (PfATP6) of *P. falciparum*. PfATP6 is essential for *P. falciparum* calcium homeostasis. However, the mechanism of interaction as well as the binding affinity of artemisinin with heme and PfATP6 has not yet known. In this regards the molecular modeling study could be very helpful to explore the mode of interaction. Docking, binding free energy and quantitative structure activity relationship (QSAR) are computational ways to explore the binding structure, binding affinity, interaction of ligand/receptor and development of activity model. In the work, several computational approaches were used to explore the binding of artemisinin and its structural derivatives in heme and PfATP6. Generating computational models of a sufficient pool of potential analogues of artemisinin, these virtual screening methods can be extended to facilitate the rational design of novel derivatives, guide the design of focused libraries based on the artemisinin skeleton and facilitate the search for related structures with similar biological activity from large databases.

Objective of the work

In the work, we tried to use computational methods to explore the binding structures, binding affinity and inhibition mechanism of artemisinin and its derivatives in their corresponding receptors. Two bio-systems were used in the work: heme polymerization and PfATP6 in *Plasmodium falciparum*. By studying these two systems, we want to produce more information for researcher to understand their biological function affected by artemisinin, how they affect both the receptors and what type of analogues will better inhibit the biosystems. Also we tried to develop approaches to calculate the biological activity of artemisinin analogues by ways of free energy of binding (FEB), binding affinity and quantitative structure-activity relationship (QSAR). These ways should be convenient approaches, which can be used to normal set of compounds to benefit ligand activity evaluations in a rational drug design.

Virtual library of artemisinin analogues: molecular building and preparation

An initial dataset of 144 artemisinin analogues were collected from published data (6-11) in which several different ring systems were represented. All of the analogues were either peroxides or trioxanes, which should act via similar mechanisms of action and were categorized into 10 classes.

Each of these compounds had associated in vitro bioactivity values (IC_{50} values reported in ng/ml) against the drug resistant malaria strain *P. falciparum* (W-2 clone). The log value of the relative activity (RA) of these compounds was used for analysis and was defined as:

$$\text{Log(RA)} = \log[(\text{artemisinin } IC_{50}/\text{analogue } IC_{50})(\text{analogue MW/artemisinin MW})] \quad (1)$$

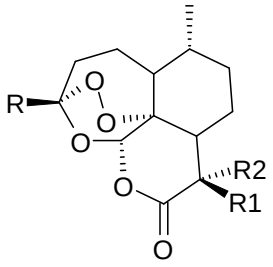
Molecular models of the artemisinin and its analogues (Table 1) were built using the Builder feature in Maestro (Schrodinger package) and energy minimized in a vacuum using Impact. Each structure was assigned an appropriate bond order using ligprep script shipped by Schrödinger and optimized initially by means of the OPLS2005 force field using default setting. Complete geometrical optimization of these structures was carried out with the HF/3-21G method using the Jaguar (Schrodinger Inc.). In order to check the reliability of the geometry obtained, we compared the structural parameters of the artemisinin 1,2,4-trioxane ring with theoretical (12) and experimental (13, 14) values from the literature. All calculations reproduced most of the structural parameters of the artemisinin 1,2,4-trioxane ring seen in X-ray structures. This applies especially to the bond length of the endoperoxide bridge which seems to be responsible for the antimalarial activity. All the analogues used in the study have been included in Table 2.

Table 1. Experimental and theoretical values of the 1,2,4-trioxane ring parameters in artemisinin (bond lengths in Å; bond angles and torsional angles in degrees).

	Theoretical			Experimental[c]	Experimental[d]
Parameters[a]	3-21G[b]	3-21G** [b]	6-31G[b]		
O1-O2	1.463	1.462	1.447	1.475(4)	1.469(2)
O2-C3	1.441	1.440	1.435	1.417(4)	1.416(3)
C3-O4	1.436	1.436	1.435	1.448(4)	1.445(2)
O4-C5	1.407	1.408	1.403	1.388(4)	1.379(2)
C5-C6	1.529	1.530	1.533	1.528(5)	1.523(2)
C6-O1	1.478	1.477	1.469	1.450(4)	1.461(2)
O1-O2-C3	106.9	107.070	108.800	107.600(2)	108.100(1)
O2-C3-O4	107.0	107.310	106.760	107.200(2)	106.600(2)
C3-O4-C5	115.6	115.700	117.300	113.500(3)	114.200(2)
O4-C5-C6	112.0	112.030	112.280	114.700(2)	114.500(2)
C5-C6-O1	111.1	111.589	110.910	111.100(2)	110.700(2)
C6-O1-O2	111.2	111.286	113.240	111.500(2)	111.200(2)
O1-O2-C3-O4	-74.9	-74.680	-71.840	-75.500(3)	-75.500(2)
O2-C3-O4-C5	31.8	32.150	33.390	36.300(4)	36.000(2)
C3-O4-C5-C6	29.4	28.400	25.320	24.800(4)	25.300(2)
O4-C5-C6-O1	-51.8	-50.769	-49.410	-50.800(4)	-51.300(2)
C5-C6-O1-O2	10.1	9.792	12.510	12.300(3)	12.700(2)
C6-O1-O2-C3	50.8	50.522	46.700	47.700	47.800(2)

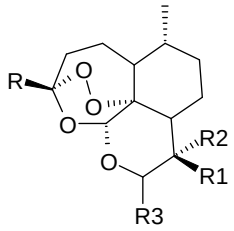
[a] Atoms are numbered according to Figure 1
[b] This work
[c] Values from Ref. (13) (experimental estimated standard deviations in brackets)
[d] Values from Ref. (14) (experimental estimated standard deviations in brackets)

Table 2. Artemisinin analogues with antimalarial activities against the drug-resistant malarial strain *P. falciparum* (W-2 clone) used in the work.



Compounds	R	R1	R2	log (RA)	IC_{50} (ng/ml)
1	CH_3	CH_3	H	1.00	0.040
2	C_4H_8Ph	H	H	0.45	0.194
3	CH_3	H	2-Z-Butenyl	-1.10	5.750
4	CH_3	H	H	0.79	0.065
5	CH_3	Allyl	H	0.34	0.550
6	CH_3	C_4H_9	H	0.17	0.311
7	C_4H_8Ph	C_4H_9	H	-0.32	1.310
8	$CH_2CH_2CO_2Et$	C_4H_9	H	1.36	0.025
9	C_4H_9	C_4H_9	H	-0.48	1.568
10	CH_3	C_2H_5	H	1.40	0.017
11	CH_3	C_6H_{13}	H	0.86	0.069
12	CH_3	i- C_6H_{13}	H	-0.04	0.547
13	CH_3	i-C_5H_{11}	H	0.07	0.408
14	C_3H_6(p-Cl-Ph)	H	H	0.10	0.457
15	C_4H_9	H	H	-0.74	2.416
16	$CH_2CH_2CO_2Et$	H	H	0.37	0.214
17	CH_3	C_3H_6(p-Cl-Ph)	H	1.37	0.025
18	CH_3	Br	CH_2Br	-1.64	27.24
19	CH_3	$=CH_2$	-	-0.89	3.083
20	CH_3	CH_2CH_3	-	-0.36	1.053
21	CH_3	-CH_2CH_2-	-	-0.94	3.632
22	CH_3	C_5H_{11}	H	1.02	0.046
23	CH_3	C_4H_8Ph	H	0.63	0.133
24	CH_3	C_4H_8Ph	H	0.12	0.400

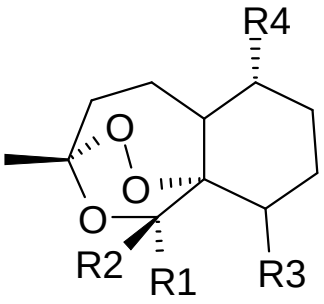
Table 2(continued). 10-Substituted artemisinin derivatives with antimalarial activities against the drug-resistant malarial strain *P. falciparum* (W-2 clone) used in the work.



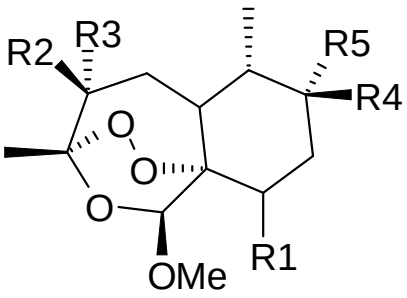
Compounds	R	R1	R2	R3	log (RA)	IC_{50} (ng/ml)
25	CH_3	CH_3	H	H	0.75	0.068
26	CH_3	CH_3	H	OH	0.55	0.114
27	CH_3	CH_3	H	OEt	0.34	0.202
28	CH_3	CH_3	H	OH	0.96	0.051
29	CH_3	H	Br	H	0.28	0.248
30	CH_3	CH_3	Br	NH-2-(1,3-thiazole)	0.66	0.134
31	CH_3	CH_3	Br	p-Cl-aniline	0.79	0.105
32	CH_3	CH_3	Br	aniline	0.18	0.397
33	CH_3	Br	CH_3	NH-2-pyridine	-0.09	0.768
34	CH_3	CH_3	Br	NH-2-pyridine	-0.77	3.667
35	CH_3	CH_3	H	α -OEt	0.32	0.212
36	CH_3	C_4H_9	H	H	1.32	0.021
37	CH_3	C_2H_5	H	H	0.67	0.086
38	CH_3	C_3H_7	H	OEt	-0.04	0.529
39	CH_3	H	H	OEt	0.43	0.157
40	CH_3	CH_3	H	C_3H_6OH	0.78	0.077
41	CH_3	CH_3	H	C_4H_9	0.06	0.400
42	CH_3	CH_3	H	OCH_2CO_2Et	0.52	0.158
43	CH_3	CH_3	H	$OC_2H_4CO_2Me$	0.10	0.433
44	CH_3	CH_3	H	$OC_3H_6CO_2Me$	-0.03	0.605
45	CH_3	CH_3	H	OCH_2 (4-PhCO_2Me)	-0.07	0.720
46	CH_3	CH_3	H	(R)-$OCH_2CH(CH_3)CO_2Me$	1.79	0.009
47	CH_3	CH_3	H	(S)-$OCH_2CH(CH_3)CO_2Me$	2.25	0.003
48	CH_3	CH_3	H	(R)- $OCH(CH_3)CH_2CO_2Me$	0.87	0.073
49	CH_3	CH_3	H	(S)-$OCH(CH_3)CH_2CO_2Me$	1.70	0.011
50	$CH_2CH_2CO_2Et$	H	H	H	0.70	0.096
51	C_4H_9	H	H	H	0.75	0.075
52	C_4H_8Ph	H	H	H	0.58	0.139
53	CH_3	-OCH_2-	-	OOH	-0.62	1.857
54	CH_3	-CH_2O-	-	OOH	-0.57	1.655
55	CH_3	$=CH_2$	-	OOH	-0.99	4.131
56	CH_3	C_5H_{11}	H	H	0.16	0.318
57	CH_3	C_3H_6Ph	H	H	1.40	0.021
58	CH_3	CH_3	H	OOt-C_4H_9	0.92	0.061
59	-	CH_3	OH	α-OH	-0.89	3.303
60	-	CH_3	H	CH_2CHF_2	0.11	0.366
61	-	CH_3	OH	OCH_2CF_3	0.33	0.243
62	-	CH_3	OH	OEt	-0.44	1.281

All R3 substituents are β except where noted.

Table 2(continued). Artemisinin derivatives lacking the D-ring with antimalarial activities against the drug-resistant malarial strain *P. falciparum* (W-2 clone) used in the work.



(63-74)



(75-79)

Compounds	R1	R2	R3	R4	log (RA)	IC_{50}(ng/ml)
63	O CH_3Ph	H	H	H	-0.09	0.530
64	O CH_3	H	$C_2H_4O_2CNEt$	H	-0.65	0.118
65	H	O CH_3	$C_2H_4OCH_3$	H	-0.39	0.996
66	H	O CH_3	$C_2H_4OCH2Ph$	H	0.75	0.091
67	H	O CH_3	C_2H_4O-allyl	H	0.40	0.184
68	H	O CH_3	$C_2H_4O_2Ph$	H	-0.59	2.086
69	H	O CH_3	$C_2H_4O_2C(4\text{-}PhCO_2Me)$	H	0.27	0.343
70	H	O CH_3	$C_2H_4O_2C(4\text{-}PhCO_2H)$	H	-0.81	3.856
71	-	O CH_3	-	-	1.70	0.398
72	H	O CH_3	$C_2H_4O_2C(4\text{-}PhCO_2C_2H_4NMe_2)$	H	0.25	2.790
73	H	O CH_3	$C_2H_4O_2CCH_2NCO_2\text{-}(t\text{-}C_2H_9)$	H	-0.04	0.670
74	H	O CH_3	$C_2H_4OCH_2$ (4-N-Me-pyridine)	H	-0.90	4.439

Compounds	R1	R2	R3	R4	R5	log (RA)	IC_{50}(ng/ml)
75	C_2H_4OH	H	CH_3	H	H	-1.80	26.849
76	C_2H_4OH	CH_3	H	H	H	0.23	0.251
77	C_2H_4OH	CH_3	CH_3	H	H	-1.80	28.102
78	$C_2H_4OCH_2Ph$	CH_3	CH_3	H	H	-1.80	36.157
79	$C_2H_4OCH_2$(4-py)	-	-	-	-	0.14	0.373

Table 2(continued). Miscellaneous Artemisinin derivatives with antimalarial activities against the drug-resistant malarial strain *P. falciparum* (W-2 clone) used in the work.

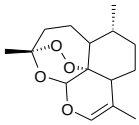
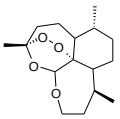
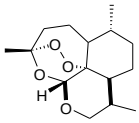
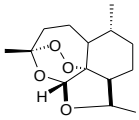
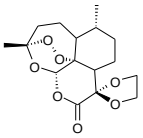
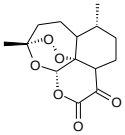
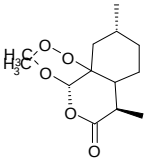
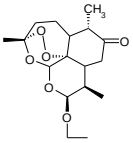
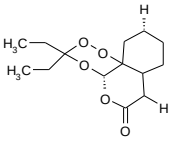
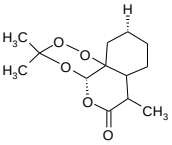
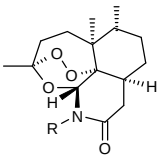
Compounds	structure	log (RA)	IC_{50} (ng/ml)
80		0.78	0.063
81		-1.20	6.340
82		-0.79	2.344
83		-0.64	1.573
84		-2.09	56.889
85		-2.49	123.612
86		-0.80	2.309
87		0.16	0.320
88		-0.60	1.525
89		-1.27	6.762
90		0.328	0.400

Table 2(continued). 9-Substituted Artemisinin derivatives with antimalarial activities against the drug-resistant malarial strain *P. falciparum* (W-2 clone) used in the work.

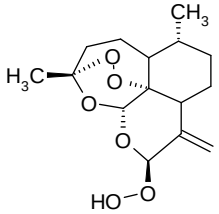
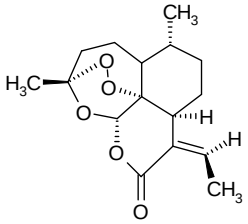
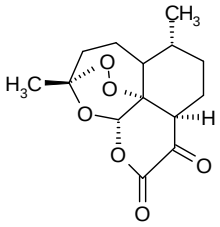
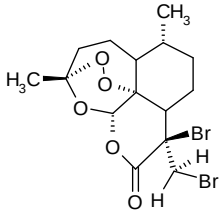
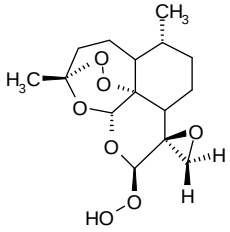
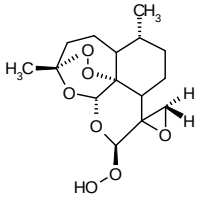
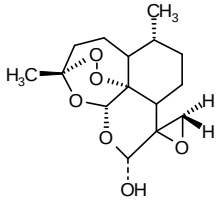
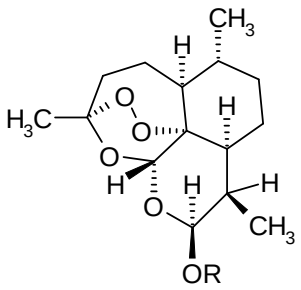
Compounds	Structure	log (RA)	IC_{50} (ng/ml)
91		-0.739	2.320
92		-0.197	0.657
93		-2.298	79.429
94		-1.487	19.143
95		-0.460	1.286
96		-0.409	1.143
97		-0.361	0.971

Table 2(continued). Dihydroartemisinin derivatives with antimalarial activities against the drug-resistant malarial strain *P. falciparum* (W-2 clone) used in the work.



Compounds	R	log (RA)	IC_{50} (ng/ml)
98	OR = H	0.487	0.123
99	(S)-$CH_2CH(CH_3)COOCH_3$	2.104	0.004
100	(S)-$CH(CH_3)CH_2COOCH_3$	0.599	0.137
101	1-adamantylmethyl	0.007	0.020
102	(S)-$CH_2CH(CH_3)COOH$	-0.658	0.603
103	(S)- $CH(CH_3)CH_2COOH$	-0.608	2.123
104	(R)-$CH(CH_3)CH_2COOH$	-0.383	2.380
105	OR= =O	-0.269	0.743
106	$CH_2PhCOOH$	0.176	0.394
107	(R)-$CH_2CH(CH_3)COOCH_3$	1.524	0.016
108	(R)-$CH_2CH(CH_3)COOH$	-0.463	1.520

Table 2(continued). Tricyclic 1.2.4 – trioxanes derivatives with antimalarial activities against the drug-resistant malarial strain *P. falciparum* (W-2 clone) used in the work.

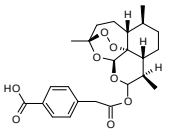
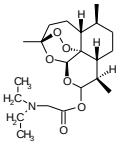
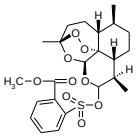
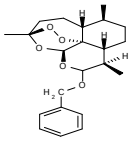
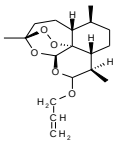
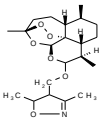
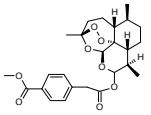
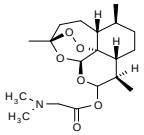
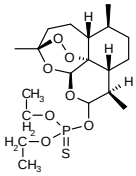
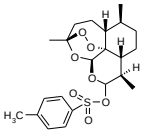
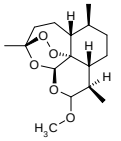
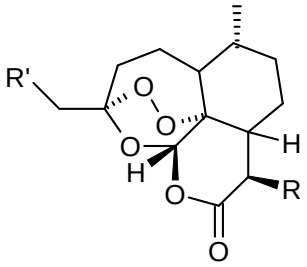
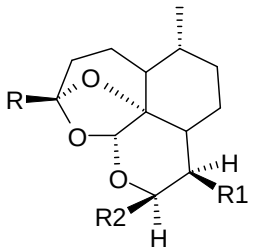
Compounds	structure	log (RA)	IC_{50} (ng/ml)
109		-0.475	1.886
110		0.995	0.057
111		-0.413	1.771
112		0.968	0.057
113		0.905	0.057
114		0.991	0.057
115		0.660	0.143
116		0.787	0.086
117		0.717	0.057
118		0.434	0.229
119		0.129	0.314

Table 2(continued). 3C- substituted artemisinin derivatives with antimalarial activities against the drug-resistant malarial strain *P. falciparum* (W-2 clone) used in the work.



Compounds	R1	R	log (RA)	IC_{50} (ng/ml)
120	CH_3	H	0.049	0.357
121	CH_3 CH_2	H	0.828	0.062
122	CH_3CH	H	-0.347	0.977
123	EtO_2CCH_2	H	0.365	0.216
124	$C_6H_5CH_2$	H	-2.000	50.78
125	p-$ClC_6H_4(CH_2)_2$	-	0.104	0.453
126	C_6H_5 $(CH_2)_3$	H	0.449	0.195
127	CH_3	CH_3 $(CH_2)_3$	0.410	0.187
128	$CH_3(CH_2)_2$	CH_3 $(CH_2)_3$	-0.481	1.573
129	$C_6H_5CH_2$	CH_3 $(CH_2)_3$	-2.000	58.72
130	p-$ClC_6H_4(CH_2)_2$	CH_3 $(CH_2)_3$	-0.276	1.239
131	$C_6H_5(CH_2)_3$	CH_3 $(CH_2)_3$	-0.319	1.306
132	EtO_2CCH_2	CH_3 $(CH_2)_3$	1.359	0.025

Table 2(continued). Deoxy-artemisinin derivatives with antimalarial activities against the drug-resistant malarial strain *P. falciparum* (W-2 clone) used in the work.



Compounds	R	R1	R2	log (RA)	IC_{50} (ng/ml)
133	CH_3	CH_3	OEt	-4	4198.58
134	CH_3	CH_3	OH	-4	3801.42
135	CH_3	C_4H_8Ph	-	-4	5248.23
136	CH_3	C_3H_7	-	-4	3971.63
137	CH_3	CH_3	-	-4	4567.37
138	CH_3	C_4H_9	H	-4	4170.21
139	$CH_2CH_2CO_2Et$	H	H	-4	4652.48
140	C_2H_4Ph	H	-	-4	3574.47
141	CH_2CH_3	H	-	-4	3574.47
142	CH_3	C_2H_4Ph	-	-4	4851.06
143	CH_3	C_3H_6Ph	-	-4	5049.64
144	CH_3	CH_3	-	-4	3773.05

Automated calculation of docking of artemisinin to heme

The high selectivity in the killing of parasites by artemisinin may be due to its interaction (15) with heme which accumulates in high quantities in parasitised red blood cells as a by-product

of hemoglobin lysis by the malaria parasite (16). The toxic heme (Fig. 2) is mostly detoxified by a specific mechanism of heme polymerization into non-toxic and insoluble polymer called as hemozoin which accumulates as a crystalline pallet in the cytosol of the erythrocytes (16). The heme polymerization is a target for some antimalarials, such as chloroquine, that inhibit this process (17). A recent study reported that artemisinin also inhibits heme polymerization (18). The chloroquine-resistant strain of *Plasmodium berghei* that lacks hemozoin, possibly because heme polymerization does not occur, is also resistant to artemisinin (19). This supports the view that inhibition of heme polymerization is the mode of action of artemisinin. It is very possible that artemisinin interacts with heme and hence inhibits the polymerization process.

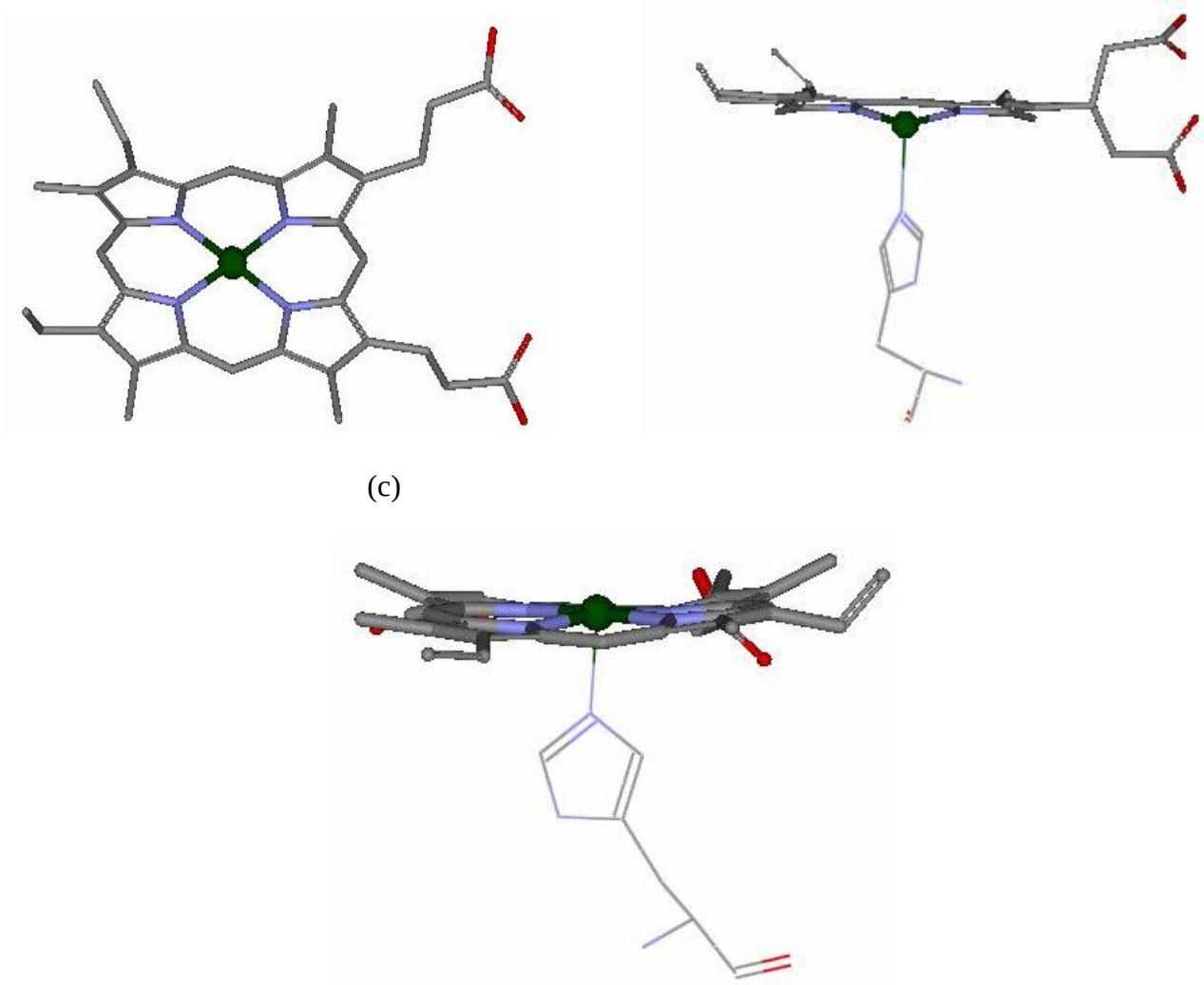
Understanding the nature of these interactions is very significant and theoretical calculations, in particular the molecular docking method, seem to be a proper tool for gaining such understanding. For the heme iron, accurate ab initio calculations were performed to obtain its atomic charge (and those of artemisinin) instead of using a crude approximation for the charge of iron, and specific parameters for iron were used in the docking calculations. The effects of different heme structures were also considered. Thus, three heme structures taken from the literature were studied. The knowledge obtained from this study has been used as a guide for series of docking experiments between heme and artemisinin derivatives. Also prediction models for predicting the antimalarial activity of artemesinin derivatives were developed based on binding interaction with heme as descriptor. We have used the molecular modeling techniques such as molecular docking and rescoring using Prime MMGB/SA.

Receptor preparation

The X-ray structure of three heme structures i.e., heme-pdb, heme-deoxy and heme-oxy were taken from the Protein Data Bank. These structures are all different owing to the source of heme and the oxidation state of iron. The first structure heme-pdb (PDB ID: 1CTJ) has planar heme structure with a strong positive charge on its central iron atom, which lies slightly above the porphyrin plane (Fig. 2a). In the process of hemoglobin degradation by the malaria parasite, the proximal ligand may possibly still be attached to the heme iron and, therefore, it is very possible that the histidine remains with the heme structure. As a result, the second and the third structures, heme-deoxy and heme-oxy, respectively, were obtained from the modifications of deoxy and oxy forms of hemoglobin which contain histidine as the proximal. Both deoxy and oxy forms of hemoglobin were taken from the Protein Data Bank (Id: 1A3N and 1HHO). In the heme-deoxy, the

histidine pulls the Fe atom to lie below the protoporphyrin plane and gives it a basin-like structure (Fig. 2b). In the oxy hemoglobin structure, there are six coordinations for heme iron, i.e. with four N atoms in the protoporphyrin ring, with the proximal ligand (histidine), and with O_2. Thus, for docking purposes, the O_2 coordination was deleted while maintaining the coordinates of the rest; this modified structure was taken as the receptor structure. As in heme-deoxy, the protoporphyrin plane has a basin-like structure, because of the attraction to the heme iron by histidine. Interaction with O_2 causes the Fe atom to be drawn up above the plane (Fig. 2c), however, and thus results in a structure which is markedly different from the heme-deoxy.

(a) (b)



(c)

Figure 2. The structures of three heme compounds: (a) heme-pdb, (b) heme-deoxy and (c) heme-oxy.

Atomic charge calculations

To investigate the effect of the atomic charge on docked configurations, atomic charges of both artemisinin and heme obtained at various levels of theory were used. For heme the HF/3-21G

and HF/6-311G** atomic charges were calculated. For artemisinin, atomic charge calculations were performed at HF/3-21G, HF/6-31G* and HF/6-311G**. All quantum chemical calculations were carried out using Jaguar (Schrodinger, Inc. 2000). It has been seen that all the 3 atomic charges (HF/3-21G, HF/6-31G* and HF/6-311G**) used for artemisinin gave similar results. Thus for the sake of saving CPU times, the HF/3-21G charges were chosen for atomic charge calculation and complete geometry optimization of all the artemisinin analogues used in the study.

Docking of the ligands

All the ligands were docked to the heme receptor using Glide 4.0. After ensuring that protein and ligands are in correct form for docking, the receptor-grid files were generated using grid-receptor generation program, using van der Waals scaling of the receptor at 0.4. The default size was used for the bounding and enclosing boxes. The grid box was generated at the centroid of the heme receptor. The ligands were docked initially using the "standard precision" method and further refined using "xtra precision" Glide algorithm. For the ligand docking stage, van der Waals scaling of the ligand was set at 0.5. Of the 50,000 poses that were sampled, 4,000 were taken through minimization (conjugate gradients 1,000) and the 30 structures having the lowest energy conformations were further evaluated for the favorable Glide docking score. A single best conformation for each ligand was considered for further analysis.

Rescoring using Prime/MM-GBSA

For each ligand, the pose with the lowest Glide score was rescored using Prime/MM-GBSA approach (20). This approach has been used to predict the free energy of binding for set of ligands to receptor. The docked poses were minimized using the local optimization feature in Prime and the energies of complex were calculated using the OPLS-AA force field and generalized-Born surface area (GBSA) continuum solvent model. The binding free energy (ΔG_{bind}) is then estimated using the equation:

$$\Delta G_{bind} = E_{R:L} - (E_R + E_L) + \Delta G_{solv} + \Delta G_{SA} \qquad (1)$$

where $E_{R:L}$ is energy of the complex, $E_R + E_L$ is sum of the energies of the ligand and unliganded receptor, the outcome of the use of OPLS-AA force field, ΔG_{solv} (ΔG_{SA}) is the difference between GBSA solvation energy (surface area energy) of complex and sum of the corresponding energies for the ligand and unliganded protein. Corrections for entropic changes were not applied in this type of free energy calculation.

Results and discussion

The docking calculations showed that the docking configurations depend on the heme-pdb atomic charges and especially the charge of Fe. All docking calculations agree that the heme iron binds with endoperoxide oxygens, where the O1–Fe distance is the shortest. Among these calculations, docking with HF/6-311G** charges yielded the shortest O1–Fe distance of 2.55 Å. This O1–Fe distance is markedly much shorter than those predicted using HF/3-21G (2.71 Å) charges. For the binding energy, the docking with HF/6-311G** charges gave the lowest energy. It is, therefore, reasonable to choose atomic charges from HF/6-311G** for heme in further docking calculations. To study the effect of atomic charges of artemisinin, the docking calculations using various charge schemes, i.e., HF/3-21G, HF/6-31G*, and HF/6-311G** for the artemisinin and HF/6-311G** charges for heme-pdb structure were performed. The dockings with ab initio charges (HF/3-21G, HF/6-31G*, and HF/6-311G**) gave similar results. Thus, for the sake of saving CPU times, the HF/3-21G charge was chosen for artemisinin.

To investigate the effect of the heme structure, three heme structures were selected as described in material and methods. The atomic charges were assigned as HF/6-311G** charges for all three heme molecules. For artemisinin compounds, the HF/3-21G optimized structure and atomic charges were used. The heme structure chosen does have an effect on the docking results. Although we could not observe agreement on O–Fe distances, all docking calculations with different heme structures (except heme-deoxy) suggested that artemisinin prefers to dock at endoperoxide oxygens (O1 and O2). Using heme-pdb for the heme structure, the docking results showed that artemisinin pointed its endoperoxide moiety toward the heme iron for the most occurring configuration (Fig. 3a). Owing to the planar structure of the heme-model, the repulsion between artemisinin and the protoporphyrin ring of heme prevents artemisinin from approaching the heme iron as closely as for heme-pdb. For the heme-deoxy, because of its basin-like structure (see Fig. 3b), the binding with the endoperoxide moiety of artemisinin is less favorable and a stronger O13–Fe attraction is resulted. This could be observed from the most occurring configuration, which has the shorter O13–Fe distance of 3.28 Å, compared with 5.85 and 5.63 Å for O1–Fe and O2–Fe (Fig. 3b). Still, this distance is longer than those obtained from the docking with other heme structures. For heme-oxy, the most occurring configuration has O1–Fe as the shortest heme–artemisinin distance (Fig. 3c). The O1–Fe and O2–Fe distances of 2.56 Å and 3.52 Å are comparable with those of heme-pdb. Note that hemeoxy and heme-pdb have similar structures. From the results from the three heme structures, it can be concluded that the structure of the heme

molecule has a significant effect on the docking configurations. The steric hindrance at the Fe position plays an important role in the binding.

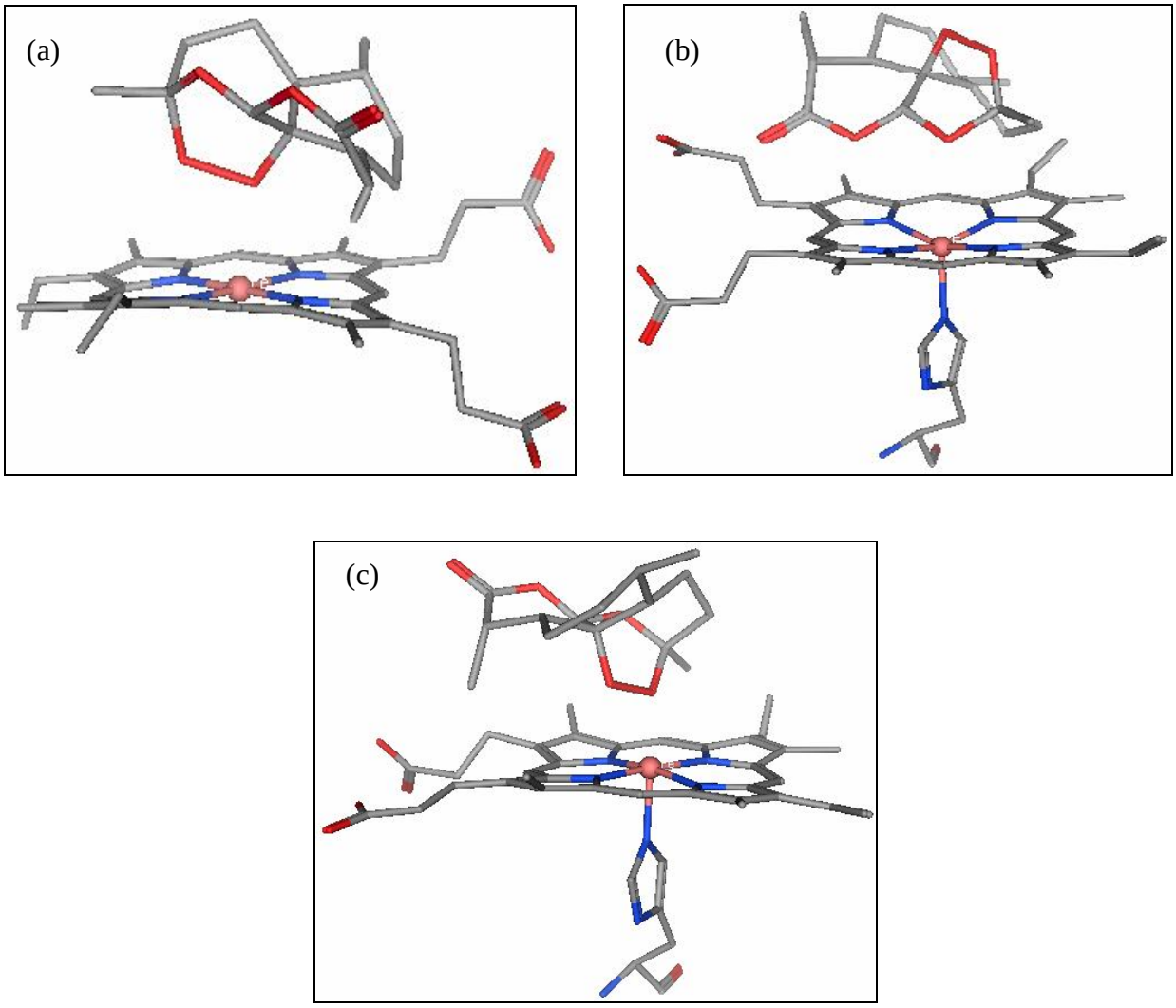


Figure 3. Docking configuration between artemisinin and (a) heme-pdb, (b) heme-deoxy and (c) heme-oxy structure.

Docking of artemisinin derivatives and development of prediction models

In this study heme-pdb was taken for docking of artemisinin derivatives. Docking methods has been applied here in this study to test whether the peroxide bridge performs an important role in binding to heme. For artemisinin (QHS) the atoms O1, O2, O13, O11 and for deoxy-artemisinin (DQHS) the interacting oxygen atoms; O2, O13, O11, respectively were being tested for interaction with heme-iron. The docking process was able to placed QHS derivatives at distances of 3.89 (± 1.19) Å, 4.46 (± 1.13) Å, 5.61 (±0.64) Å and 5.53 (± 0.63) Å with respect to O1, O2, O13 and O11 oxygen atoms respectively from heme-iron. DQHS derivatives have single oxygen instead of the peroxide bridge. The heme-iron could interact with DQHS in several ways. Form the docking results it has been seen that the heme-iron preferentially interact either at the side involving the

three nonperoxide oxygens O2, O13 and O14 or the peroxide derived oxygen O11. Thus the active antimalarial QHS clearly interacts with heme in a manner different from its inactive analogue DQHS. The orientation of QHS with respect to heme may be critical to formation of free radicals and thus for drug action. For the highly potent analogues the distances between O1-Fe, O2-Fe, O13-Fe and O11-Fe atom pairs between heme-iron and artemisinin were 2.59 (± 0.22) Å, 3.31 (± 0.36) Å, 5.49 (± 0.25) Å and 5.45 (± 0.24) Å respectively; the glide score obtained was -2.20 (± 0.11). A linear relationship between Glide score and optimized O-Fe distance was obtained with R^2 value of 0.659 (Fig. 4a). The optimized O-Fe distance was obtained by linear combination of O1-Fe, O2-Fe, O13-Fe and O11-Fe atom pairs between oxygen's of artemisinin and heme-iron. For the DQHS analogues (pIC_{50} < -3.0) which lack the peroxide bridge the Glide score was found to be -1.07 (± 0.09). The distances for O2-Fe, O13-Fe and O11-Fe atom pairs were 4.58 (± 0.59) Å, 5.34 (± 1.08) Å and 5.37 (± 0.74) Å, respectively. Further a linear relationship with R^2 value of 0.7702 was obtained between Glide score and the optimized O-Fe distance (Fig. 4c). The interaction of the artemisinin with heme is very much dependent upon the stereochemistry of artemisinin analogues, a mechanism that is controlled by steric hindrance. The analogues which approach the heme-iron as close as possible will have better interaction and thus the good glide score.

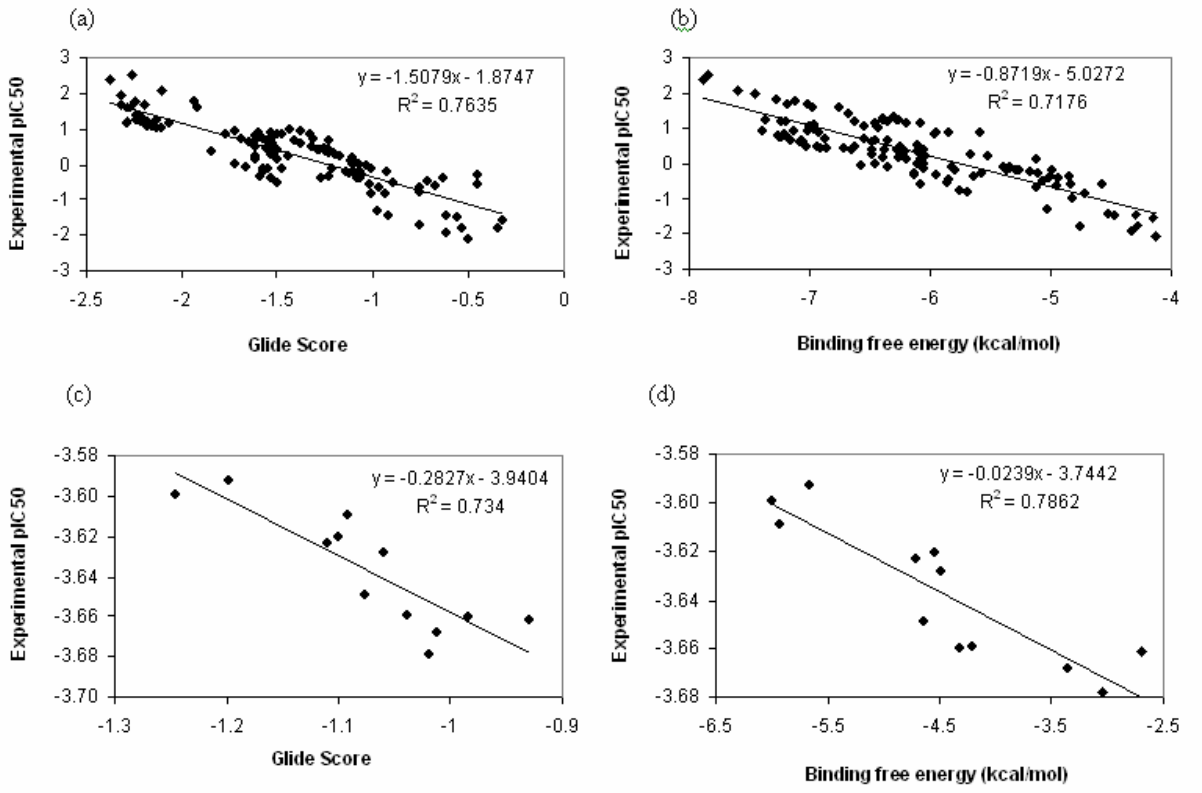


Figure 4(a-d). Models for predicting antimalarial activity (pIC_{50}) of the (a&b)Artemisinin (QHS) analogues and (c&d) Deoxy-artemisinin (DQHS) analogues based on Glide score and Binding free energy (ΔG_{bind}) as descriptor.

For each ligand in the virtual library, the pose with the lowest Glide score was rescored using Prime/MM-GBSA approach. Rescoring using Prime/MM-GBSA leads to minor changes of the ligand conformations within receptor site. These changes result from minimization of the ligand in receptor's environment and consequent stabilization of receptor:ligand complex. This approach is used to predict the binding free energy (ΔG_{bind}) for set of ligands to receptor. A linear relationship between linear combination of O-Fe distances and ΔG_{bind} energy was obtained (Fig. 4b & 4d) with R^2 value of 0.7073 and 0.7303 for the QHS and DQHS derivatives respectively.

A linear regression model for prediction of predicted pIC_{50} of antimalarial activity based on Glide score has been developed by considering analogues with known pIC_{50}.

$$pIC_{50} = - 1.88 (\pm 0.115) - 1.51 (\pm 0.074)* \text{G-score} \quad (2)$$
$$(N = 132, r^2 = 0.763, r^2_{cv} = 0.762, s = 0.428, F = 419.66)$$

Reasonably good agreement between predicted and experimental pIC_{50} are found (root mean square error = 0.36 and 0.01 for QHS and DQHS derivatives). Similar prediction model of predicted pIC_{50} of antimalarial activity has been developed by considering ΔG_{bind} energy as a descriptor. The equations 5 of the model and the corresponding statistics for QHS analogues are shown below:

$$pIC_{50} = -5.03 (\pm 0.298) – 0.872 (\pm 0.048)* \Delta G_{bind} \quad (3)$$
$$(N = 132, r^2 = 0.718, r^2_{cv} = 0.715, s = 0.471, F = 330.22)$$

The calculated pIC_{50} based on ΔG_{bind} energy descriptor was in good agreement with experimental pIC_{50} (root mean square error = 0.40 and 0.01 for QHS and DQHS derivatives) and suggested that the prediction model is robust and accurate.

A LIE-SGB continuum solvation model for binding affinity calculations of artemisinin analogues.

The structure-based linear interaction energy method implementing a surface generalized Born (SGB-LIE) (18) continuum model for solvation was used to build a binding affinity model for estimating the free energy of binding for a diverse set of artemisinin analogues. The analogues were docked into the heme as per the methods developed above. The docked complex corresponding to each analogue was transported to the LIAISON package for subsequent SGB-LIE calculations. The molecular dynamics (MD) technique has been used for LIE conformation space sampling. The system was initially heated to 300 K in 5 ps and then subjected to a MD simulation for 25 ps. A residue-based cutoff of 12 Å was set for the non-bonding interactions. The non-bonded pair list was updated every 10 fs. The time integration step of 1.0 fs and sampling LIE energies every 10 steps

was used. During the MD simulations, all the residues of the receptor beyond 12 Å from the bound ligand were frozen. Similarly, the average LIE energies for the ligand were obtained using the OPLS-2005 force field. The average LIE energy terms were used for building binding affinity model and free energy estimation for artemisinin analogues using the following equation.

$$\Delta G_{bind} = \alpha\,(\langle U^{b}_{ele}\rangle - \langle U^{f}_{ele}\rangle) + \beta\,(\langle U^{b}_{vdw}\rangle - \langle U^{f}_{vdw}\rangle) + \gamma\,(\langle U^{b}_{cav}\rangle - \langle U^{f}_{cav}\rangle) \qquad (4)$$

where bracketed terms represent the ensemble average of the energy terms such as van der Waals (*Uvdw*), electrostatic (*Uele*) and cavity (*Ucav*) energy. All the terms are evaluated for interaction between ligand, both in the free (f) and bound (b) state and its environment. The α, β and γ are LIE fitting parameters. The values obtained for the three fitting parameters, α, β and γ, are -0.141, -0.0926, and -1.07, respectively.

Results and discussion

A training set of 101 artemisinin analogues was used to build a binding affinity model for estimating the free energy of binding for 57 inhibitors (test set). The root mean square error (RMSE) between the experimental ΔG values and the values obtained by the SGB-LIE fit was 0.328 kcal/mol which is an indicator of the robustness of the fit. The quality of the fit can also be judged by the value of the squared correlation coefficient (r^2) which was 0.845 for the training set. The statistical significance of the SGB-LIE model is evaluated by the correlation coefficient r, standard error s, F-test value, leave-one-out cross-validation coefficient q^2 and predictive error sum of squares PRESS.

$$\Delta G = (-0.0271)\langle Uele\rangle + (-0.0902)\langle Uvdw\rangle + (-1.44)\langle Ucav\rangle \qquad (5)$$
$$(n = 102,\ r^2 = 0.845,\ S = 0.465,\ F = 234.1,\ q^2 = 0.844,\ PRESS = 21.38)$$

SGB-LIE model developed in this study is statistically (q^2 = 0.844, r^2 = 0.845, F = 234.1) best fitted and consequently used for prediction of antimalarial activities (pIC_{50}) of training and test sets of molecules. The free energy values for the test set were estimated based on optimized SGB-LIE parameters α, β and γ from the training set. There is a close match between the experimental and LIE free energy values of the ligands in the test set. Figure 5a & 5b graphically show the quality of fit for the training and test set. The predicted activity calculated from free energy of binding is satisfactory with small deviation compared with experimental activity of training and test sets of molecules with a root-mean-square error (RMSE) of 0.348 kcal/mol with respect to experimental data. Low levels of RMSE for the majority of analogues establish the structure-based LIE method as an efficient tool for generating more potent and specific inhibitors of tubulin by testing rationally

designed lead compounds based on podophyllotoxin derivatization. To test how sensitive the LIE method is to the underlying sampling techniques or in other words, how good the sampling techniques is in surfing the local conformation space, we also implemented LIE with the HMC sampling(21, 22). A Metropolis accept/reject criterion is checked every 5 steps of HMC's underlying MD simulation.

(a) (b)

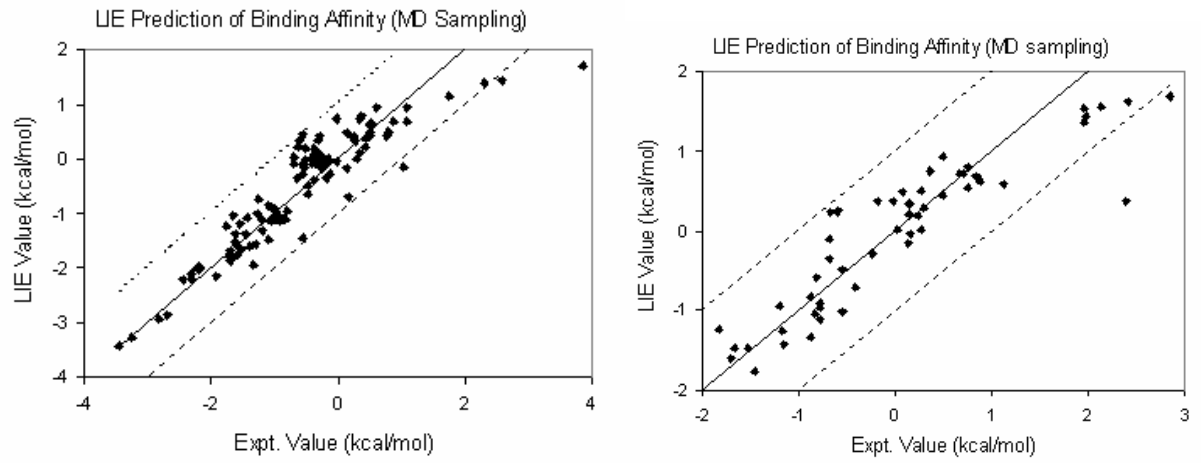


Figure 5. Free energy values estimated by the SGB-LIE method for (a) 76 podophyllotoxin analogues comprising the Training set plotted against corresponding experimental data; the RMS error is 0.48 kcal/mol and (b) for 36 podophyllotoxin analogues comprising the Test set plotted against corresponding experimental data; the RMS error is 0.56 kcal/mol.

The time step used in HMC's underlying MD is 3.0 fs with RESPA algorithm(23, 24). Using the same three-parameter model, the LIE predictions are shown in Figure 6.

(a) (b)

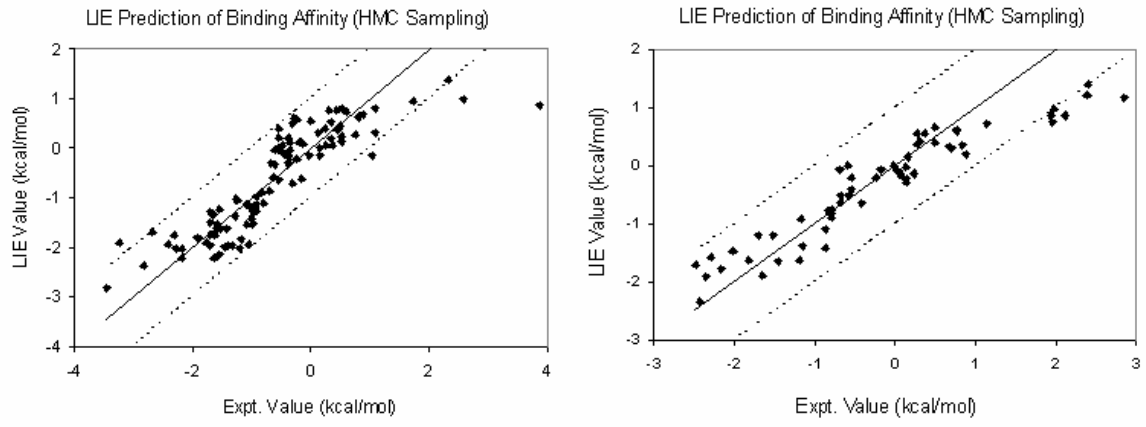


Figure 6. LIE binding energies for the Training set from HMC sampling. The overall RMS error is 0.415 kcal/mol for 101 ligands studied here. If LIE results agree perfectly with the experimental values, the data points (represented by diamonds) should be on the diagonal line. (a) Training set and (b) Test set.

The overall RMS error is 0.415 kcal/mol and the correlation coefficient is 0.758, which are comparable to those from MD sampling. The new parameters are found to be α = -0.0078, β = -0.0735 and γ = -1.20.

$$\Delta G = (-0.0078)\langle Uele\rangle + (-0.0735)\langle Uvdw\rangle + (-1.20)\langle Ucav\rangle \qquad (6)$$

(n = 103, r^2 = 0.758, S = 0.588, F = 133.6, P = 0.0001, q^2 = 0.757, PRESS = 36.51)

For the test set of 57 compounds the SGB-LIE model was able to predict their activity with an overall RMS error of 0.371. Figure 6 graphically shows the quality of fit between the SGB-LIE binding energy predictions vs the experimental values of the test set.

Computational screening of artemisinin analogues based on QSAR model

A set of 372 molecular descriptors were calculated using ADME Model Builder software package. We used a more systematic way of variable selection in order of missing value test → zero test → simple correlation test → genetic algorithm to obtain a set of good descriptors. QSAR equations were developed between the observed biological activity and the descriptors. The predictive capability of the equation (q^2) is determined using leave-one-out cross validation method. The relation for q^2 is as shown below.

$$q^2 = 1 - \frac{PRESS}{TOTAL} = 1 - \frac{\sum_{i=1}^{n}\left(y_{exp} - y_{pred}\right)^2}{\sum_{i=1}^{n}\left(y_{exp} - \overline{y}\right)^2}$$

where y_{pred}, $y_{\exp}$ and y are the predicted, experimental and mean values of activity, respectively. In order to evaluate the predictive ability of our QSAR model, we used the method described by Golbraikh et al., (25) and Roy et al., (26). The determination coefficient in prediction, q^2_{test} was calculated using the following equation (26).

$$q^2_{test} = 1 - \frac{\sum\left(Y_{pred_{Test}} - Y_{Test}\right)^2}{\sum\left(Y_{Test} - \overline{Y}_{Training}\right)^2}$$

Where $Y_{pred_{Test}}$ and Y_{Test} are the predicted value based on QSAR equation (model response) and experimental activity values, respectively, of the external test set compounds. $Y_{Training}$ is the mean activity value of the training set compounds. To further check the inter-correlation of descriptors variance inflation factor (VIF) analysis was performed. VIF value is calculated from $1/1\text{-}r^2$, where r^2 is the multiple correlation coefficient of one descriptor's effect regressed on the remaining molecular descriptors. If VIF value is larger than 10, information of descriptor can be hidden by correlation of descriptors (27, 28).

The 194 active compounds considered as potential of W-2 strain of *P. falciparum* inhibition were randomly divided into a training set of 156 compounds and a test set of 38 compounds. A set of 13 parameters selected based on systematic way of variable selection was used for development of QSAR equation. The best significant relationship between the molecular descriptors and antimalarial activity has been deduced to be:

pIC_{50} = 4.04 - 9.76 V7CH + 0.069 EMAX1 + 0.117 LOGP - 0.581 GEOM3 - 0.00014 STRA6 + 0.0035 STRA4 + 0.039 STRA2 - 1.09 L/B2 - 4.82 FVMN - 1.02 HOMO - 0.285 BOMX - 1.36 MOLC9 - 5.24 V6C (7)

(N = 141; r^2 = 0.845; s = 0.343; PRESS = 19.386; r^2_{adj} = 0.830; q^2_{cv} = 0.799; *F*-test = 53.40)

The QSAR model developed in this study is statistically (r^2 = 0.845, q^2_{cv} = 0.799, *F*-test = 53.40) best fitted and consequently used for prediction of antimalarial activity (pIC_{50}) of training and test sets of molecules. The r^2 and q^2_{cv} values of 0.845 and 0.799, respectively of the model corroborates with the criteria for a QSAR model to be highly predictive (Leban et. al., 1988) [27]. The inter-correlation of the descriptors used in the QSAR equation was very low (below 0.6). Satisfied with the robustness of the QSAR model developed using training set, we have applied the QSAR model to an external data set of artemisinin analogs comprising the test set. The overall root mean square error (RMSE) between the experimental and predicted pIC_{50} value was 0.325 which revealed good predictability. The squared correlation coefficient between experimental and predicted pIC_{50} values for the test set is also significant (r^2 = 0.871). The Figure 4a & 4b shows the quality of the fit. Since the value of q^2_{test} = 0. 876 and rm^2 = 0.788 were found to be in the acceptable range (28), thereby indicating the good external predictability of the QSAR model.

(a) (b)

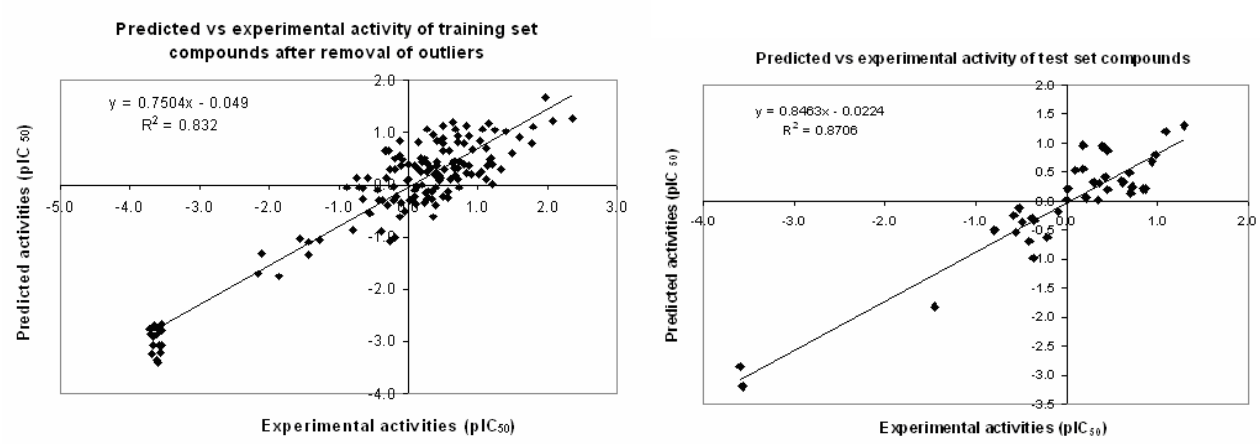


Figure 7. Relationship between predicted and experimental activities as per QSAR equation for the (a) training set and (b) test set.

The high predictive ability of the computational models developed in this study allows virtual screening of chemical databases or virtual libraries determined by either synthetic feasibility or commercial availability of starting materials to prioritize the synthesis of most promising candidates. Therefore, these models should facilitate the rational design of novel derivatives, guide the design of focused libraries based on the artemisinin skeleton and facilitate the search for related structures with similar biological activity from large databases.

Molecular interaction of artemisinin derivatives with PfATP6

An alternative hypothesis for the mode of action of artemisinins has been proposed, based on structural similarities between the sesquiterpene moieties of thapsigargin and in artemisinins. Thapsigargin (TG) (another plant product from *Thapsia garganica*) is an extremely potent inhibitor of Ca^{2+}-transporting ATPases (sarcoplasmic reticulum Ca^{2+}- transporting ATPases or SERCAs) from a wide variety of organisms. It was suggested that artemisinins may act in a similar way, but more specifically to inhibit the SERCA of malarial (PfATP6). Further the experimental studies revealed that artemisinins inhibit the sarcoendoplasmic reticulum Ca^{2+}-ATPase (SERCA) orthologue (PfATP6) of *P. falciparum* in Xenopus oocytes (Hawley *et. al.*, 1998) (10). PfATP6 is thought to be the real molecular target of artemisinins in spite of some disagreements to be resolved (29).

The amino acid sequence of PfATP6 is known (30) but the three-dimensional structure is not available. In this study, therefore we have constructed the 3D structure of PfATP6 by homology modeling using Prime 1.5. An identity of 43.5% with well-studied protein of SERCA (pdb id: 1iwo; resolution 1.3 Å) and contains the highly specific inhibitor thapsigargin (TG) provided a great strength for modeling the protein. Sequence alignment between them further improved by removing the mismatched sequence part (375-707) from the whole sequence and then constructed the three-diemnsional structure of PfATP6. Model evaluation was performed in PROCHECK v3.4.4 and VERIFY3D (31). The structures of SERCA and homology-modeled PfATP6 has been presented in Figure 8. The binding site of PfATP6 was identified with reference to the studies done on SERCA (pdb id: 1iwo) and thapsigargin as ligand. We carried out in silico studies to confirm these active sites, using SiteMap algorithm. The output from the SiteMap program showed coherent active sites for the target protein as reported in site directed mutagenesis study (32). All the 154 artemisinin ligands with known antimalarial activities (W2 clone) and thapsigargin (TG) were docked into the defined binding site using Glide-XP 4.0. The docking protocol was validated by extracting the crystal structure of TG from SERCA and redocking it. Thapsigargin in PfATP6 maintained the

same spatial coordinates as in SERCA. The RMSD value calculated out of ten accepted poses for docked TG and crystal structure of TG was found in between 0.37–1.57 Å (Fig. 9). The binding mode of both TG and artemisinin within the binding site is represented in Figure 9.

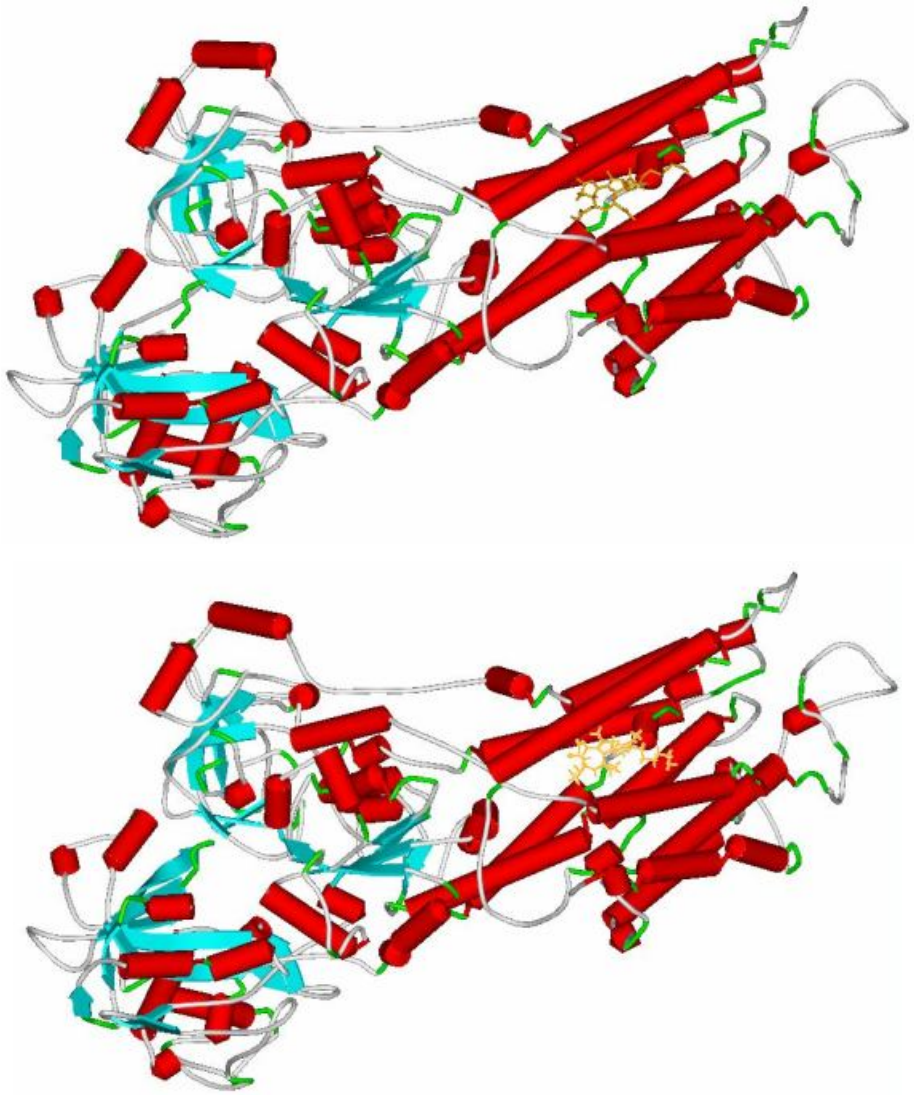


Figure 8 Schematic display of SERCA (above) and PfATP6 (below) generated using DS Visualizer for windows. Helices and sheets are represented as red cylinders and cyan arrows, respectively. The lignad TG (in brown stick) is included in the structure.

In this figure we can observe that both the molecule well fitted to the defined binding pocket. All the 154 artemisinin analogues were also found to be good binder with PfATP6. The binding modes of artemisinin and its derivatives showed hydrophobic interaction with PfATP6. As the Fe^{2+}-dependent activation and antimalarial activity of artemisinin do not depend on the heme binding, we can propose that the production of the carbon centered free radical (8) should not precede the binding to PfATP6. Therefore, artemisinin should be bound to PfATP6 before activiation by Fe^{2+} ion. For each ligand in the virtual library, the pose with the lowest Glide score was rescored using

Prime/MM-GBSA and EmBrace approaches. These approaches predict the binding free energy for set of ligands to receptor.

(a) (b)

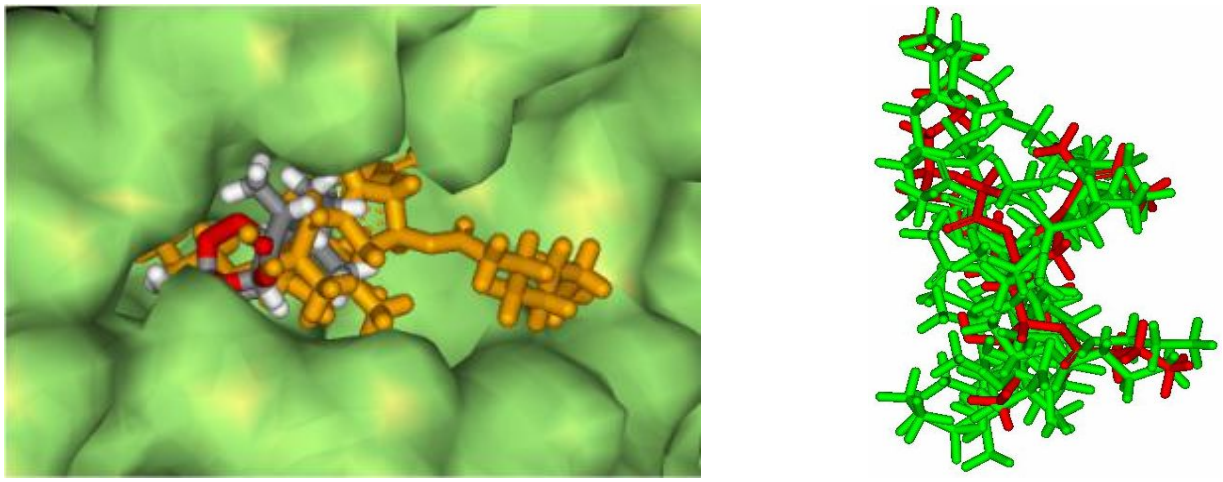


Figure 9. (a) Binding mode of artemisinin and TG within the binding site of PfATP6. TG: brown stick model and Artemisinin: green stick model. (b) Superposition of all docked configurations of TG on crystal structure (*red-stick*). RMSD (heavy atom) = 0.37- 1.57Å.

Building models for prediction of RA using Glide score and Prime/MM-GBSA energy

Both the docking scores and ΔG_{bind} well explained the activities of artemisinin derivatives. Figure 10 (a & b) shows good correlations of Glide XP score and ΔG_{bind} with relative antimalarial activities compared with artemisinin.

(a) (b)

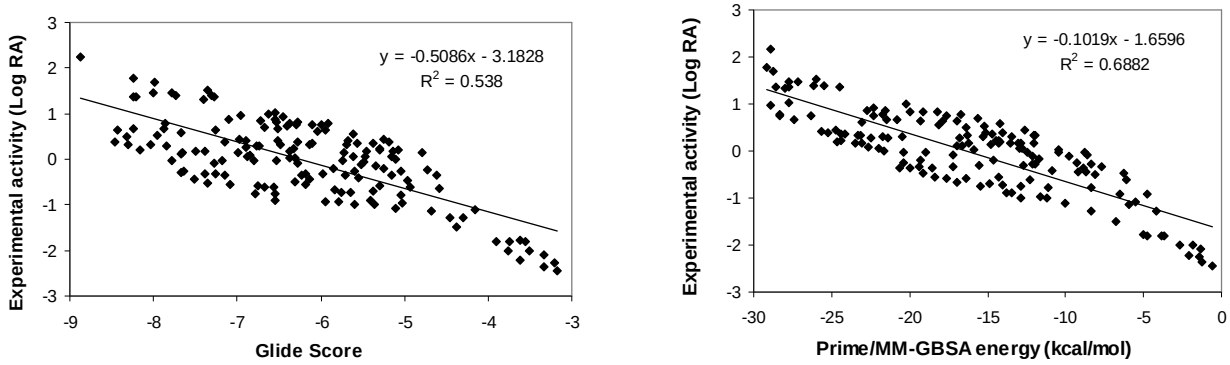


Figure 10. Models for predicting antimalarial activity (Log RA) of the artemisinin derivatives based on (a) Glide score and (b) Prime/MM-GBSA energy (ΔG_{bind}).

The equation (8) of the model and the corresponding statistics are shown below:

$$\text{Log RA} = -3.18\ (\pm 0.241) - 0.509\ (\pm 0.038) \times \text{G-score} \qquad (8)$$

($N = 154$; $r^2 = 0.538$; $s = 0.601$; $F = 178.22$; $r^2_{cv} = 0.525$; PRESS = 56.811)

The regression model developed in this study is statistically ($r^2_{cv} = 0.525$, $r^2 = 0.538$, $F = 178.22$) best fitted and consequently used for prediction of antimalarial activities (log RA) of the artemisinin

analogues. Rescoring using Prime/MM-GBSA leads to minor changes of the ligand conformations (due to energy minimization of the ligand in receptor's environment) and consequent stabilization of receptor and ligand complex. A linear regression model for prediction of predicted antimalarial activity (log RA) has been developed by considering analogues with known experimental activity. In this model we have taken ΔG_{bind} energy as a descriptor. The Equation (9) of the model and the corresponding statistics are shown below:

$$\text{Log RA} = -1.66\ (\pm 0.098) - 0.102\ (\pm 0.006) \times \Delta G_{bind} \qquad (9)$$
$$(N = 154;\ r^2 = 0.688;\ s = 0.495;\ F = 333.24;\ r^2_{cv} = 0.679;\ PRESS = 37.997)$$

The regression model developed based on ΔG_{bind} energy is statistically ($r^2_{cv} = 0.679$, $r^2 = 0.688$, F = 333.24) best fitted and consequently used for prediction of antimalarial activities (log RA) of the artemisinin analogues.

Linear optimization of energy parameters vs Activity

One docking structure with better Glide score from each molecule docking result was picked up as final docked structure in PfATP6 for further calculations. As the Glide program treats a receptor rigidly during docking simulation, an energy minimization was performed to the docked complex. A vdW energy and electrostatic energy between ligand and receptor were calculated for each minimized complex. Also a desolvation energy and solvent accessible surface area (SASA) change were calculated using EMBRACE (Schrodinger package). A scheme similar to Linear Response was used to develop a free energy of binding (FEB) relationship based on these energies, which can express the activity of these artemisinin derivatives. From the results of correlation factor analysis, it can be seen that ΔG_{vdW} has most significant correlation to the activity (log RA) and electrostatic energy (ΔGele) has less significant correlation to the activity. It indicates that the binding of these artemisinin derivatives to PfATP6 is almost hydrophobic. The correlation between predicted activity and actual activity is shown on Figure 11. The calculated activity has good correlation to the actual activity. The Equation (10) of the model and the corresponding statistics are shown below:

$$\text{Log RA} = -1.37 - 0.0035\ \text{SASA} - 0.0314\ \Delta G_{vdW} + 0.0029\ \Delta G_{ele} - 0.00398\ \Delta G_{solv} \qquad (10)$$
$$(N = 135;\ r^2 = 0.815;\ s = 0.291;\ F = 141.99;\ r^2_{cv} = 0.802;\ PRESS = 11.65)$$

The regression model developed based on linear response scheme is statistically ($r^2_{cv} = 0.802$, $r^2 = 0.815$, F = 141.99) best fitted and consequently used for prediction of antimalarial activities (log RA) of the artemisinin analogues. However, we may observe that model using linear response scheme is better for predicting antimalarial activity than model using Glide score and ΔG_{bind} descriptor as descriptors.

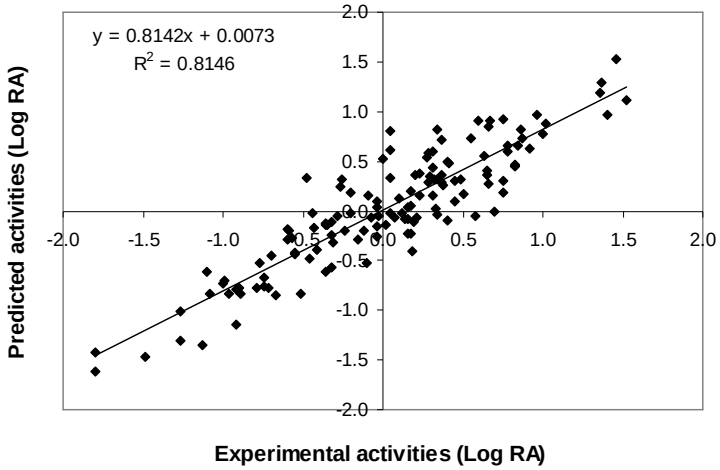


Figure 11. Models for predicting antimalarial activity (Log RA) of the artemisinin derivatives based on linear optimized energy parameters.

Conclusion

- Significant correlation co-efficient and low level of root mean square error, established the docking and prime MM-GB/SA based prediction model as an efficient tool for generating more potent and specific inhibitors of heme polymerization and PfATP6 in *Plasmodium falciparum* by testing rationally designed lead compounds based on artemisinin derivatives.
- The fast and accurate estimation of binding free energies based on LIE-SGB and linear optimized combination of energies parameters provides a means to screen the compound libraries for lead optimization and rational design. Hence, this could bring about the development of new and more effective drugs.
- Using a combination of topological, electro-topological-state indices, electronic and thermodynamic descriptors of chemical structures, we have built several robust QSAR models with high values of q^2 (for training sets) and predictive R^2 (for test set).
- Therefore, these models should facilitate the rational design of novel derivatives, guide the design of focused libraries based on the artemisinin skeleton and facilitate the search for related structures with similar biological activity from large databases.

List of Publications

- **M. Srivastava**; H. Singh; P. K. Naik, Application of the linear interaction energy method for rational design of artemisinin analogues as haeme polymerisation inhibitors journal of SAR and QSAR in Environmental Research, Volume 20,(2009) , pp: 327 – 355.

- **M. Srivastava**; H. Singh; P. K. Naik, Molecular modeling evaluation of the antimalarial activity of artemisinin analogues: molecular docking and rescoring using prime/MM-GBSA approach journal of current research in Biological Sciences (2009) (In Press).

- **M. Srivastava**; H. Singh; P. K. Naik, Quantitative structure-activity relationship (QSAR) of the artemisinin: the development of predictive in vitro antimalarial activity models Journal of Chemometrics (2009) (In Press).

- **M. Srivastava**; P. K. Naik, The Binding Modes and Binding Affinities of Artemisinin Derivatives with Plasmodium falciparum Ca^{2+}-ATPase (PfATP6). Communicated to Journal of Molecular Modeling (2009).

References

1. Li GQ, Guo XB, Fu LC, Jian HX, Wang XH (1994). Clinical trials of artemisinin and its derivatives in the treatment of malaria in China. Trans R Soc Trop Med Hyg 88 Suppl 1:S5-6.
2. Brossi A, Venugopalan B, Dominguez Gerpe L, Yeh HJ, Flippen-Anderson JL, Buchs P, Luo XD, Milhous W, Peters W (1988). Arteether, a new antimalarial drug: synthesis and antimalarial properties. J Med Chem 31:645-50.
3. Klayman DL (1985). Qinghaosu (artemisinin): an antimalarial drug from China. Science 228: 1049-55.
4. Lee IS, Hufford CD (1990). Metabolism of antimalarial sesquiterpene lactones. Pharmacol Ther 48: 345-55.
5. Luo XD, Shen CC (1987). The chemistry, pharmacology, and clinical applications of qinghaosu (artemisinin) and its derivatives. Med Res Rev 7: 29-52.
6. Acton N, Karle JM, Miller RE (1993). Synthesis and antimalarial activity of some 9-substituted artemisinin derivatives. J. Med. Chem. 36, 2552-2557.
7. Lin AJ, Margaret Lee, Daniel L, Klayman J (1989) Antimalarial Activity of New Water-Soluble Dihydroartemisinin Derivatives. 2. Stereospecificity of the Ether Side Chain. J. Med. Chem. 32: 1249-1252.
8. Posner GH, Oh CH, Gerena L, Wibur KM (1992) Extraordinarily Potent Antimalarial Compounds: New, Structurally Simple, Easily Synthesizes, Tricyclic 1,2,4-Trioxanes. J. Med. Chem. 35:2459-2467.
9. Avery MA, Bonk JD, Chong WKM, Mehrotra S, Miller R. et al. (1995). Structure-Activity Relationships of the Antimalarial Agent Artemisinin. 2. Effect of Heteroatom Substitution at O-11: Synthesis and Bioassay of N-Alkyl-11-aza-9-desmethylartemisinins. J. of Med. Chem., 38: 5038-5044.
10. Avery MA, Mehrotra S, Bonk JD, Vroman JA, Goins DK. et al. (1996). Structure-Activity Relationships of the Antimalarial Agent Artemisinin. 4. Effect of Substitution at C-3. J. of Med. Chem. 39: 2900-2906.
11. Pinheiro JC, Ferreira MMC, Romero OAS (2001) Antimalarial activity of dihydroartemisinin derivatives against P. falciparum resistant to mefloquine: a quantum chemical and multivariate study. Theochem-J. Mol. Struct., 572: 35-44.

12. Leban I, Golic L, Japelj M (1988). Crystal and molecular structure of qinghaosu: a redetermination. Acta. Pharm. Jugosl. 38: 71–77.
13. Lisgarten JN, Potter BS, Bantuzeko C, Palmer RA (1998). Structure, Absolute Configuration, and Conformation of the Antimalarial Compound, Artemisinin. J. Chem. Cryst. 28: 539–543.
14. Fersht AR (1984). Basis of biological specificity Trends Biochem. Sci., 9: 145.
15. Meshnick SR, Thomas A, Ranz A, Xu CM, Pan HZ *(*1991). Artemisinin (qinghaosu): the role of intracellular hemin in its mechanism of antimalarial action. *Mol Biochem Parasitol*. 49(2):181–189.
16. Goldberg DE, Slater AF, Cerami A, Henderson GB (1990). Hemoglobin degradation in the malaria parasite Plasmodium falciparum: an ordered process in a unique organelle. *Proc Natl Acad Sci U S A*. 87(8):2931–2935.
17. Slater AF (1993). Chloroquine: mechanism of drug action and resistance in Plasmodium falciparum. Pharmacol. Ther. 57: 203-235.
18. Pandey AV, Tekwani BL, Singh RL, Chauhan VS (1999). Artemisinin, an endoperoxide antimalarial, disrupts the hemoglobin catabolism and heme detoxification systems in malarial parasite. J.Biol. Chem 274:19383-19388.
19. Peters W, Li ZL, Robinson B L, Warhurst D C (1986). The chemotherapy of rodent malaria, XL. The action of artemisinin and related sesquiterpenes Ann. Trop. Med. Parasitol. 80: 483-489.
20. Lyne PD, Lamb ML, Saeh JC (2006) Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. J. Med. Chem., 49:4805-4808.
21. Duane S, Kennedy AD, Pendleton BJ, Roweth D (1987). Hybrid Monte Carlo. Phys. Lett. B: 195-216.
22. Zhou R, Berne BJ, (1997). Smart Walking: a new method for Boltzmann sampling of protein conformations. J. Chem. Phys. 107: 9185
23. Tuckerman M, Berne BJ, Martyna GJ, (1992) Reversible multiple time scale molecular dynamics. J. Chem. Phys., 97: 1990.
24. Zhou R, Berne BJ (1995). A new molecular dynamics method combing the reference system propagator algorithm with a fast multipole method for simulating proteins. J. Chem. Phys. 103: 9444.
25. Golbraikh A, Tropsha A (2002). Beware of q^2. J. Mol. Graph. Model. 20(4): 269-276
26. Roy PP, Roy K (2008). On some aspects of variable selection for partial least squares regression models. QSAR Comb. Sci. 27: 302-313.
27. Jaiswal M, Khadikar PV, Scozzafava A, Supuran CT (2004). Carbonic anhydrase inhibitors: the first QSAR study on inhibition of tumor-associated isoenzyme IX with aromatic and heterocyclic sulfonamides. Bioorg. Med. Chem. Lett.; 14: 3283-3290.
28. Shapiro S, Guggenheim B (1998). Inhibition of Oral Bacteria by Phenolic Compounds. Part 1. QSAR Analysis using Molecular Connectivity. Quant. Struct-Act. Relat. 17: 327-337.
29. Ellis DS et. al. (1985). The chemotherapy of rodent malaria, XXXIX. Ultrastructural changes following treatment with artemisinine of Plasmodium berghei infection in mice, with observations of the localization of [3H]-dihydroartemisinine in P. falciparum in vitro. Ann. Trop. Med. Parasitol. 79: 367–374.
30. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter

JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B (2002). Genome sequence of the human malaria parasite Plasmodium falciparum. Nature. 419: 498-511.
31. Eisenberg D, Luthy R, Bowie JU (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 277:396-404.
32. Uhlemann ACA, Cameron U, Eckstein-Ludwig J, Fischbarg P, Iserovich FA, Zuniga M, East A, Lee L, Brady RK, Haynes and Krishna S (2005). A single amino acid residue can determine the sensitivity of SERCAs to artemisinins. Nat. Struct. Mol. Biol. 12:628-629.