



Molecular Docking and 3-D-QSAR Studies on the Possible Antimalarial Mechanism of Artemisinin Analogues

Feng Cheng, Jianhua Shen,* Xiaomin Luo, Weiliang Zhu, Jiande Gu, Ruyun Ji, Hualiang Jiang* and Kaixian Chen

Center for Drug Discovery & Design and State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences, 294 Taiyuan Road, Shanghai 200031, PR China

Received 17 December 2001; accepted 6 May 2002

Abstract—Artemisinin (Qinghaosu) is a natural constituent found in *Artemisia annua* L., which is an effective drug against chloroquine-resistant *Plasmodium falciparum* strains and cerebral malaria. The antimalarial activities of artemisinin and its analogues appear to be mediated by the interactions of the drugs with hemin. In order to understand the antimalarial mechanism and the relationship between the physicochemical properties and the antimalarial activities of artemisinin analogues, we performed molecular docking simulations to probe the interactions of these analogues with hemin, and then performed three-dimensional quantitative structure–activity relationship (3-D-QSAR) studies on the basis of the docking models employing comparative molecular force fields analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). Molecular docking simulations generated probable ‘bioactive’ conformations of artemisinin analogues and provided a new insight into the antimalarial mechanism. The subsequent partial least squares (PLS) analysis indicates that the calculated binding energies correlate well with the experimental activity values. The CoMFA and CoMSIA models based on the bioactive conformations proved to have good predictive ability and in turn match well with the docking result, which further testified the reliability of the docking model. Combining these results, that is molecular docking and 3-D-QSAR, together, the binding model and activity of new synthesized artemisinin derivatives were well explained. © 2002 Elsevier Science Ltd. All rights reserved.

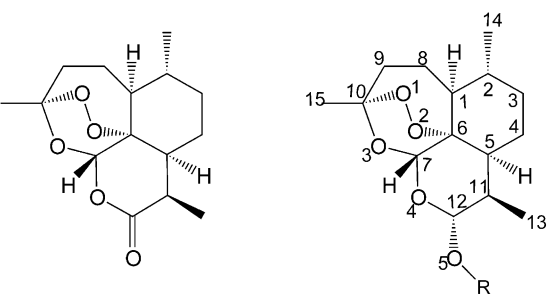
Introduction

Artemisinin (Qinghaosu, QHS, Fig. 1) was first isolated in 1971 from *Artemisia annua* L., an ancient Chinese herbal remedy. It is an effective antimalarial agent against chloroquine-resistant *Plasmodium falciparum* strains and cerebral malaria.¹ However its poor solubility restricts its employment. In order to find more potent antimalarial derivatives, our institute has been engaged in a long-term project on artemisinin chemistry including the structural modification, reactions, synthesis, and biological assay since 1978. Hypothesizing that introduction of groups to the 12-position of the dihydroartemisinin can improve their solubility, a series of C12 ether and ester analogues have been synthesized.² Their structures and biological activities, designated SD₉₀, are shown in Figure 1. In order to get new insights into the antimalarial mechanism and structure–

activity relationship of these analogues, molecular docking and 3-D-QSAR studies have been performed in the present paper.

The structure of artemisinin was identified to be an endoperoxide-containing sesquiterpene lactone.¹ The presence of endoperoxide bridge is essential for its antimalarial activity.³ Recent studies on the mode of action of artemisinin showed that free hemin could be the targeting molecule of artemisinin in biological systems.⁴ The hemin is therefore chosen as the receptor of these analogues in this paper. We first determined the probable binding (bioactive) conformations of these analogues with hemin using docking approaches and then performed both traditional and 3-D-QSAR analyses (CoMFA⁵ and CoMSIA⁶). Our results show that the docking method provides reliable bioactive conformations. The 3-D-QSAR models based on the bioactive conformations could be used to explain the difference of biological activity and to verify the docking result. Our models are proved to have good predictive ability and are consistent with the experiment data, and thereby may be helpful in designing new artemisinins analogues.

*Corresponding authors. Tel.: +86-21-6431-1833x222; fax: +86-21-6437-0269; e-mail: jhshen@iris3.simm.ac.cn (J. Shen); hljiang@mail.shenc.ac.cn (H. Jiang).



Artemisinin C16 connects with O5

Compound.	R	SD ₉₀ (mg/kg/day)	-log C ^a
Training Set			
DHQHS1	-H (α)	3.65	4.89
DHQHS2	-H (β)	3.65	4.89
SM229	-CH ₃ (α)	1.16	5.41
SM224	-CH ₃ (β)	1.02	5.47
SM227	-C ₂ H ₅ (β)	1.95	5.21
SM220	-n-C ₃ H ₇ (β)	1.70	5.28
SM245	-CH(CH ₃) ₂ (β)	2.24	5.16
SM247	-CH ₂ CH ₂ CH(CH ₃) ₂ (β)	5.65	4.8
SM249	-CH ₂ CH ₂ OCH ₃ (β)	4.10	4.92
SM277	-CH ₂ CH ₂ OCH ₂ CH ₂ CH ₃ (β)	3.71	5.02
SM280	-CH ₂ CH ₂ COCH ₃ (α)	2.28	5.18
SM105	-COCH ₃ (α)	1.20	5.44
SM108	-COC ₂ H ₅ (α)	0.66	5.71
SM241	-COC ₃ H ₇ (α)	0.65	5.74
SM223	-COOC ₂ H ₅ (α)	0.63	5.75
SM242	-COOC ₃ H ₇ (α)	0.50	5.87
SM273	-CO(p-CH ₃)C ₆ H ₄ (α)	1.73	5.37
SM233	-COCH=CHC ₆ H ₅ (α)	0.74	5.75
SM270	2-Oxacuculophety (β)	4.70	4.89
SM232	-COC ₆ H ₅ (β)	3.42	5.04
Test Set			
SM242a	-COOC ₃ H ₇ (β)	1.32	5.45
SM276	-COOC ₆ H ₅ (α)	0.63	5.81
SM374	-COCH ₂ C ₆ H ₄ (α)	0.95	5.62

^aC=SD₉₀/(molecular weight) × 1000(mol)

Figure 1. Structure formulas and antimalarial activities of artemisinin analogues.

Materials and Methods

Models

The 3-D structures of 23 artemisinin analogues were constructed using standard geometric parameters of molecular modeling software package SYBYL, release 6.5 (Tripos Inc., St. Louis, MO, USA). Atomic charges were calculated using the Gasteiger–Hückel protocol. This method is a combination of two charge computational approaches: the Gasteiger–Marsili method to calculate the σ component of the atomic charge and the Hückel method to calculate the π component of the atomic charge. With full minimization by the tripos force field, the optimized structures of these analogues were obtained.

In parasite metabolism, heme is left free after digestion of hemoglobin by the parasite.⁷ Therefore, in our study, we chose free heme as the targeting receptor of artemisinin instead of heme complexing with hemoglobin. The structure of heme is shown in Figure 2.

Docking calculations

The most negative charged region of artemisinin and its analogues locates around the peroxide moiety⁸ and the most positive charged region of heme locates around the Fe²⁺ ion. Chemical intuition indicates that these two parts may interact when artemisinins bind to heme. It is in agreement with the experiments related to antimalarial action of artemisinin, which revealed that a series of O- and C-radicals were produced through the electron transfer from Fe²⁺ to the peroxide bond when artemisinin reacts with heme.⁹ According to this clue, in the present study the peroxide bond of the ligand was placed above the Fe atom and the entire molecule of the ligand parallels to the plane of the porphyrin ring of heme. The initial distance between the middle point of the peroxide bond (M) and Fe atom was set 3 Å.

Genetic algorithm-based flexible docking method, FlexiDock in SYBYL 6.5 was employed to search the possible binding conformations of both ligands and heme. During the docking calculation the torsion angle of O1–M–

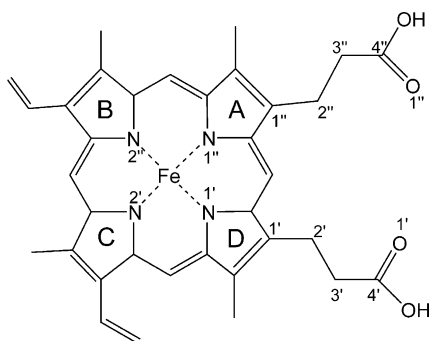


Figure 2. Structure of hemin.

Fe–N1' was scanned from 0 to 359° in an increment of 10° to find the most stable configuration of the complex. At each step of torsion scanning, flexible docking was performed setting all the rotational bonds in both artemisinins and hemin as variables. In this way, the possible global energy-minimum configuration of each ligand-hemin complex was obtained. Finally, the configurations were then optimized using the functionality of DOCK module in SYBYL 6.5. Thus the minimum energy conformations were obtained and final binding energies were calculated.

Because SYBYL 6.5 has not completely parameterized for Fe ion, the atomic parameters of Fe²⁺ had to be assigned by approximation. Shukla et al.¹⁰ has proven that Fe²⁺ in docking routine of SYBYL can be replaced by other similar atoms during molecular docking and it has little influence on the structure of hemin. Accordingly, the Fe²⁺ of hemin was substituted by Ca²⁺ in the docking calculation. FlexiDock and DOCK calculations were performed using the default parameters of SYBYL except that extra size was set to 5 Å and the maximum lattice points to 2,000,000.

QSAR analyses

Structure alignment. In the earlier 3-D-QSAR studies on the artemisinin analogues, molecules have been geometrically optimized using molecule mechanics, semi-empirical quantum chemical methods (CNDO and AM1 methods) or ab initio molecular orbital methods.¹¹ However, due to lack of consideration of the binding between these analogues and their receptor, the local low-energy conformations rather than the binding conformations of the artemisinins were used in the QSAR studies. Considering that CoMFA and CoMSIA is based on the assumption that drug-receptor interactions are responsible for biological activities, The binding conformations with hemin are therefore more suitable to 3-D-QSAR analyses.

The optimized complexes of these compounds with hemin derived from FlexiDock and DOCK were aligned using fit atom protocol in SYBYL (fit atoms: Fe, N1', N1'', N2' and N2'' of hemin, Fig. 2). The binding conformations of the artemisinins and their alignment were extracted from the complexes alignment for the 3-D-QSAR analyses.

CoMFA. For the CoMFA calculation, steric and electrostatic interactions were calculated using Tripos force field with a distance-dependent dielectric constant. The cutoff was set to 30 kcal/mol. All models were investigated using full cross-validated Partial Least Squares (PLS)¹² (leave one out) method with CoMFA standard options for scaling of variables. Minimum-sigma (column filtering) was set to 2.0 kcal/mol to improve the signal-to-noise ratio by omitting those lattice points whose energy variation is below this threshold.

CoMSIA. Three physicochemical properties, steric, electrostatic and hydrophobic similarity index fields, have been evaluated and the standard parameters implemented in SYBYL 6.5 were used.

All molecular modeling and QSAR-related calculations were carried out on a SGI O2 workstation.

Results and Discussion

Docking models and binding energy

Global energy-minimum configurations (Fig. 3) of the artemisinin analogue–hemin complexes indicate that the endoperoxide bridge exactly points towards the Fe ion, which is in agreement with the hypothesis derived from the experiments.¹³ Some important structural parameters of the ligand–hemin complexes are listed in Table 1. The optimized M·····Fe distance ranges from 2.6 to 2.7 Å, and the torsion angle of O1–M–Fe–N1' changes from 114 to 124°. The two carboxyl ethyl groups in hemin are perpendicular to the porphyrin ring, their carboxyl groups interacting with the hydrogens attached to C1 and C9, respectively: the two hydrogens attached to C9 weakly hydrogen bond with the carbonyl group of one carboxyl ethyl (C=O1), the C9·····O1 distance is about 3.1 Å; the hydrogen attached to C1 forms a weak hydrogen bond with the carbonyl of another carboxyl ethyl (C=O1'), C1·····O1' distance ranges from

Table 1. The geometry parameters for the complexes of the artemisinin analogues and hemin

Complex	H-A	M-Fe	C9-O1	C1-O1'	O1-M-Fe-N1'
DHQHS1	2.459	2.651	3.220	3.449	122.1
DHQHS2	2.426	2.639	3.29	3.414	119.1
SM229	2.417	2.680	3.107	3.532	119.8
SM224	2.389	2.692	3.376	3.518	117.0
SM227	2.492	2.623	3.087	3.349	123.6
SM220	2.427	2.688	3.087	3.451	122.1
SM245	2.423	2.641	3.241	3.503	117.9
SM247	2.460	2.695	3.063	3.455	123.7
SM249	2.460	2.695	3.063	3.455	123.7
SM277	2.509	2.630	3.094	3.343	124.6
SM280	2.438	2.652	3.087	3.365	118.0
SM105	2.423	2.693	3.102	3.469	120.1
SM108	2.426	2.701	3.120	3.462	119.9
SM241	2.460	2.695	3.063	3.455	123.7
SM223	2.406	2.695	3.270	3.672	114.9
SM242	2.415	2.709	3.127	3.634	116.2
SM273	2.423	2.689	3.105	3.515	118.4
SM233	2.419	2.693	3.117	3.556	117.6
SM232	2.411	2.666	3.262	3.416	119.6
SM270	2.436	2.655	3.102	3.503	119.1

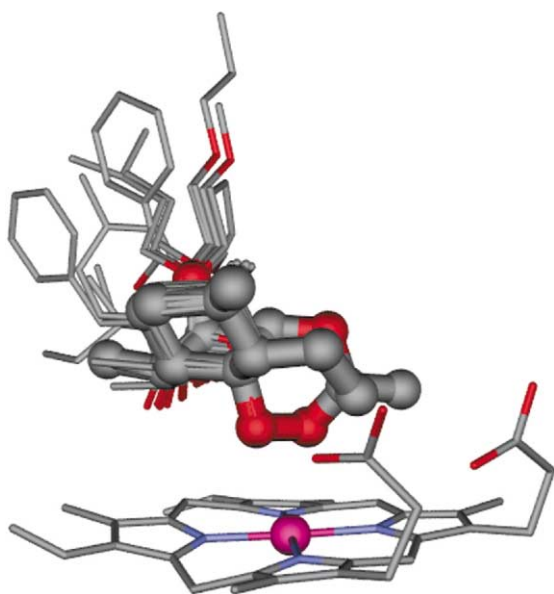


Figure 3. The final docking configurations of the 20 artemisinin analogues in training set. The common structure of these analogues and Fe ion are indicated by ball and stick model. Hydrogens of the molecules are omitted for clarity.

3.3 to 3.6 Å. These three major interactions determine the orientation of the rigid rings of artemisinins and the conformations of the two carboxylethyls in heme of the ligand–heme complexes. This result differs from that of Shukla et al.¹⁰ They did not consider the flexibility of the side chains of heme during the docking simulation; and therefore the aforementioned two important interactions between the two carbonyls and artemisinins, which are essential to control the configurations of the complexes, were overlooked. Because the flexibility of the ether and ester side chains of artemisinin were also considered in the docking calculation, the best binding conformations of these side chains were recognized by FlexiDock. In addition, the methyl group interacts perfectly with one of imidazole rings in heme through cation- π -like interaction,¹⁴ and one hydrogen atom exactly points to the center of the imidazole ring, which is beneficial to the dispersion interaction between the methyl and the imidazole. The interaction distance (from H to the center of imidazole ring) is about 2.4 Å.

Besides these important contacts between artemisinin framework and heme, ligand-specific differences are also noted in C12-substituents interactions with the receptor. It can be seen from Figure 3 that the carbonyl group of the ester α -analogues exactly interacts with the positive electrostatic field formed by the methylene end of the vinyl chain of ring B and the methylidyne group between rings B and C. As to β -analogues and ether α -analogues, the interactions disappear because of the long distances and lack of the carbonyl group respectively. The additional electrostatic interaction may be the reason why the ester α -analogues are more potent than β -analogues and ether α -analogues.

To further evaluate the reliability of the binding configurations predicted by FlexiDock and DOCK methods,

a classical QSAR were then performed to explore whether the antimalarial activities of these compounds could be correlated to the binding energy with heme. Employing partial least squares (PLS) analysis, we calculated regression equations for the antimalarial activities, $-\log C$ s, using E_{electro} (electrostatic interaction energy), E_{steric} (steric interaction) and E_{total} (total interaction energy) as descriptor variables, respectively (Table 2 and eqs 1–3).

$$-\log C = 2.339 - 0.082 \times E_{\text{electro}} \\ n = 20, s = 0.186, r = 0.850, F = 47.060 \quad (1)$$

$$-\log C = 4.194 - 0.070 \times E_{\text{steric}} \\ n = 20, s = 0.344, r = 0.253, F = 1.235 \quad (2)$$

$$-\log C = 1.411 - 0.075 \times E_{\text{total}} \\ n = 20, s = 0.187, r = 0.848, F = 46.262 \quad (3)$$

where s is the standard error, r is the correlation coefficient and F is the testing factor of the reliability.

PLS analysis indicated that a good correlation was found between the total interaction energy, E_{total} , and the antimalarial activity (eq 3). This finding is in agreement with the experiment of Paitayatat et al.,⁷ who determined the dissociation constants (K_d) for the interaction of a series of artemisinin derivatives with heme and found that the dissociation constants show a correlation with the antimalarial activity of the derivatives. The best correlation was found using the electrostatic interaction energy, E_{electro} , as the single descriptor variable (eq 1). It is shown that electrostatic effects variable is enough to fully describe the active data. This conclusion correlates well with the interaction model derived from the configurations of ligand–heme complexes, in which the dominant interactions are electrostatic interactions, including the weak hydrogen bonds and cation- π -like interaction.

Table 2. Calculated binding energy of the artemisinin analogues with the heme

Compd	$-\log C$	$E_{\text{electrostatic}}$ (kcal/mol)	E_{steric} (kcal/mol)	E_{binding} (kcal/mol)
DHQHS1	4.89	−32.640	−16.326	−48.966
DHQHS2	4.89	−32.720	−16.560	−49.262
SM229	5.41	−35.376	−16.873	−52.249
SM224	5.47	−36.438	−14.244	−50.674
SM227	5.21	−31.548	−16.518	−48.066
SM220	5.28	−35.844	−13.419	−48.061
SM245	5.16	−31.713	−16.348	−50.522
SM247	4.8	−34.950	−14.536	−49.486
SM249	4.92	−37.044	−14.705	−51.748
SM277	5.02	−32.644	−15.782	−48.426
SM280	5.18	−33.518	−14.398	−47.916
SM105	5.44	−37.659	−16.511	−54.211
SM108	5.71	−40.113	−16.337	−56.450
SM241	5.74	−40.458	−17.251	−57.709
SM223	5.75	−40.914	−15.201	−56.115
SM242	5.87	−42.167	−16.511	−58.679
SM273	5.37	−36.804	−18.366	−55.170
SM233	5.75	−41.367	−15.035	−56.403
SM232	5.04	−32.556	−14.147	−46.703
SM270	4.89	−31.888	−15.139	−47.027

It was reported that the antimalarial mechanism of artemisinin might involve two steps: artemisinin firstly forms an adduct with hemin, and then is catalyzed to generate free radical intermediates to kill the parasite.⁴ Our finding strengthens the role of hemin in antimalarial mechanism of artemisinin and its analogues. A good correlation between antimalarial activity and binding energy of artemisinin derivatives with hemin prove that the noncovalent binding or interaction for artemisinin and its derivatives with the hemin probably play an important role in the parasite killing. The complexes of these analogues and hemin are formed mainly by the electrostatic force between the peroxy group of the analogues and Fe ion in hemin. This structural character implies that this binding mode is beneficial to the electron transfer from peroxy group to Fe ion, which may lead to the rupture of the peroxy bond and the formation of free radical.

3-D-QSAR analyses

CoMFA. PLS analysis gave a correlation with a cross-validated r_{cv}^2 of 0.595, with the optimum number of components is 3. The non-cross-validated PLS analysis was repeated with the optimum number of components, as determined by the cross-validated analysis, to give r^2 of 0.973. These values indicate a good conventional statistical correlation and the CoMFA model has a good predictive ability. The steric field descriptors (1210 variables) explain 66.6% of the variance, while the electrostatic descriptors explain 33.4%. The non-cross-validated calculations are shown in Table 3.

Table 3. Experimental and calculated activities of CoMFA and CoMSIA

Compound	EA ^a	CoMFA		CoMSIA	
		PA ^b	δ ^c	PA	δ
Training set					
DHQHS1	4.89	4.92	−0.03	4.93	−0.04
DHQHS2	4.89	4.99	−0.10	5.10	−0.21
SM229	5.41	5.30	0.11	5.21	0.20
SM224	5.47	5.43	0.04	5.43	0.04
SM227	5.21	5.19	0.02	5.16	0.05
SM220	5.28	5.23	0.05	5.23	0.05
SM245	5.16	5.21	−0.05	5.18	−0.02
SM247	4.80	4.80	0.00	4.94	−0.14
SM249	4.92	4.96	−0.04	4.91	0.01
SM277	5.02	5.00	0.02	4.88	0.04
SM280	5.18	5.23	−0.05	5.29	−0.11
SM105	5.44	5.52	−0.08	5.51	−0.07
SM108	5.71	5.72	−0.01	5.67	0.04
SM241	5.74	5.74	0.00	5.73	0.01
SM223	5.75	5.76	−0.01	5.71	0.04
SM242	5.87	5.93	−0.06	5.91	−0.04
SM273	5.37	5.29	0.08	5.38	−0.01
SM233	5.75	5.72	0.03	5.72	0.03
SM232	5.04	4.95	0.09	5.09	−0.05
SM270	4.89	4.92	−0.03	4.80	0.09
Test set					
SM242a	5.45	5.41	0.04	5.32	0.13
SM276	5.81	5.70	0.11	5.82	−0.01
SM374	5.62	5.57	0.05	5.67	−0.05

^aExperimental activities.

^bCalculated.

^cResidual values.

CoMFA coefficient contour maps. The QSAR produced by CoMFA, with its hundreds or thousands of terms, was usually represented as a 3-D ‘coefficient contour’. The CoMFA steric and electrostatic fields for the analysis based on alignments of binding conformations are presented as contour plots in Figure 4. To aid in visualization, all 20 analogues of artemisinin in the training set are displayed in the maps.

The CoMFA steric contour map indicates that areas in which molecular steric bulk might have a favorable (green) or unfavorable (yellow) effect on the activity of an analogue. As shown in Figure 4, the green contour surrounding the C12-substituted group of α -analogues reflects the increase in activity found by introducing steric bulk in this region. In contrast, the two yellow contours at the end of the C12-substituted group of β -analogues indicate that steric occupancy with bulk groups in this region will decrease activity. This model illustrates why the α -analogues are, in general, more active than β -analogues.

The CoMFA electrostatic map indicates red contours in regions where high electron density might play a favorable role in activity, and blue contours in areas in which the negative charge is predicted to decrease activity. With the artemisinin series, there is only a major blue contour located around C16 of the β -analogues. This indicates that more positive charge is favored in this region. This blue contour indicates that the ester analogues are more active than ether analogues because the charge of the C16 of the former analogues is more positive than that of the latter.⁹ This is also consistent with the experiment data.

CoMSIA. Using steric, electrostatic, hydrophobicity as descriptors, a CoMSIA model with a r_{cv}^2 value of 0.554 for three components and a conventional r^2 of 0.923

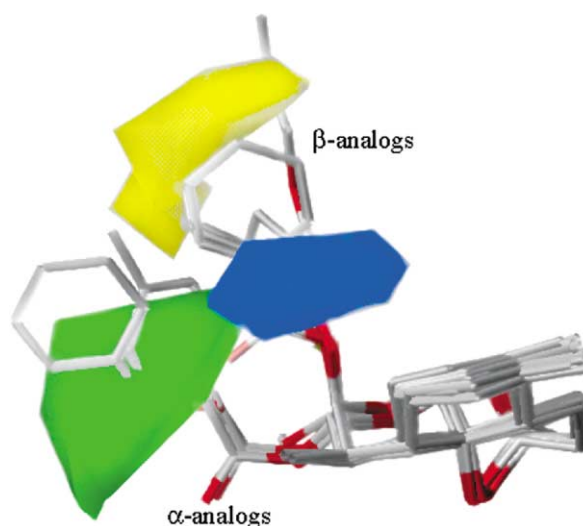


Figure 4. Contour maps from the final CoMFA analysis in combination with all artemisinin analogues. Hydrogens are omitted for clarity. Green contours refer to sterically favored regions; yellow contours indicate disfavored areas. Blue contour refer to regions where negatively charged substituents are disfavored; red contours indicate regions where negatively charged substituents are favored.

was obtained. The CoMSIA-predicted activities for these analogues are shown in Table 3. The CoMSIA model exhibited a good cross-validated correlation, indicating that it is also highly predictive. The steric field descriptors explain 25.7% of the variance, while the proportion of electrostatic descriptor remains in the same range with 30.8%. Here the additional hydrophobic field explains the remaining 43.5% of variance. The total field distribution of steric field and hydrophobic field is 69.2%, which is approximately equal to the steric distribution in CoMFA (66.6%). Thus the CoMFA steric field contribution can be seen as a combination of pure steric and hydrophobic effects.

CoMSIA coefficient contour maps. The inspection of CoMSIA steric coefficient contour maps shows a high correspondence to the CoMFA results (Fig. 5a). The steric field distribution is nearly the same as that of CoMFA (Fig. 4). In the electrostatic contour map, the blue polyhedral has moved a little below the C16, which still indicates that increasing the positive charge in the C16 will benefit to the antimalarial activity. In addition, a large red contour surrounding the oxygen atoms near the C16 of the α -analogues appears in CoMSIA map (Fig. 5b). This indicates the negative interactions, for α -analogues, are favored in this region, which is also in agreement with the interaction model derived from docking.

One of the advantages of CoMSIA is that hydrophobic contributions, which can not be completely treated using Lennard–Jones and Coulombic fields, are evaluated using a hydrophobic similarity index field. In the hydrophobic coefficient contour map, regions, where enhanced hydrophilicity and hydrophobic are favorable, are indicated by white and yellow, respectively. Seen from Figure 5c, the end of C12-substituent is surrounded by white polyhedra, which means that more hydrophilic substituent in this region is beneficial to the activity. Yellow polyhedra centralizes on the atoms C16, C12 and O5; around these regions hydrophobic substituent is favored. This is also in consistent with the CoMFA steric fields, for there are overlaps between the CoMFA steric fields (Fig. 4) and the CoMSIA hydrophobic fields (Fig. 5c),

and CoMFA static fields reflect about 40% hydrophobic fields as mentioned above.

The hydrophobic contour map is in consistent with the results derived by the pharmacological experiment. Ye et al. had synthesized some artemisinin analogues and showed that enhancing the lipophilicity of the artemisinin in the regions near C12 may increase the binding affinity of the membrane of parasite and accordingly increase their antimalarial activity.¹⁵ Wu et al.⁸ determined the logPs of these artemisinin analogues by HPLC; the activity ($-\log C$) was assumed to follow a parabolic distribution with respect to logP, and the optimized value of logP ranges from 2.6 to 2.9. Adding hydrophobic groups into C12 resulted in increasing the activity, however, the large and long hydrophobic groups (with more hydrophobicity) may decrease the activity. This is in agreement with the hydrophobic field distribution, because large and long hydrophobic groups may reach the white polyhedra. We can, therefore, conclude that CoMSIA hydrophobic field reasonably reflects the hydrophobicity–activity relationship of the artemisinin analogues.

Validation of the strategy for finding binding conformations and 3-D-QSAR models. An important aspect of a QSAR model is the ability to quantitatively calculate the relative activities of the other compounds not included in the training set. SM242a, SM276 and SM374 (Table 1) were excluded from the training set to serve as test compounds to evaluate the predictive ability of the present CoMFA and CoMSIA models. These particular compounds were selected because they represent compound diversity. SM242a is an β -analogue while SM276 and SM374 are α -analogues, of which the former is ester analogue and the latter is ether one. The results for the test-set compounds are summarized in Table 3, in which the observed antimalarial activities are listed along with the corresponding activities predicted by our CoMFA and CoMSIA models. The predicted values agreed remarkably well with the experimental results, the largest standard errors of the predictive values of both CoMFA and CoMSIA are 0.13, which shows our model is able to accurately reproduce the results of the test

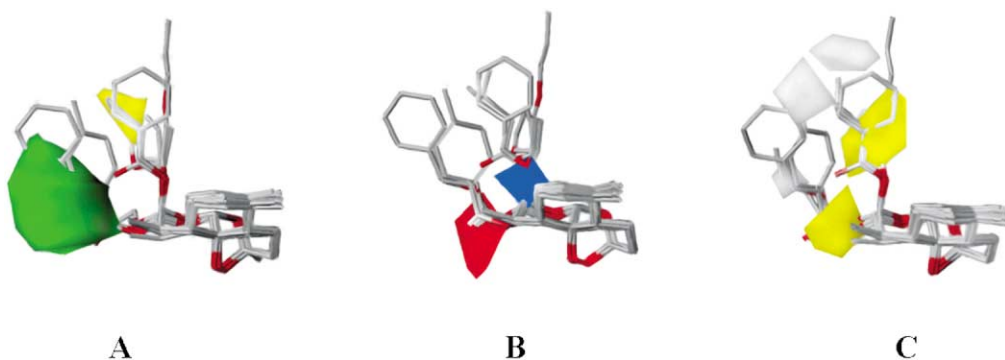


Figure 5. Contour maps from the final CoMSIA analysis in combination with all analogues. Hydrogens are omitted for clarity. (A) Steric coefficient contour map. Green contours refer to sterically favored regions; yellow contours indicate disfavored areas. (B) Electrostatic coefficient contour map. Blue contours refer to regions where negatively charged substituents are disfavored; red contours indicate regions where negatively charged substituents are favored. (C) Hydrophobic coefficient contour map. Yellow contours refer to regions where hydrophobic substituents are favored; white contours indicate regions where hydrophilic substituents are favored.

compounds with the same experimental condition as training compounds.

The antimalarial activities of the 20 compounds in the training set spread only one order of magnitude. Another set of 32 dihydroartemisinin analogues¹¹ (**1**–**32**) were therefore chosen to validate our strategy for finding binding conformations and 3-D-QSAR models. Their structures and activities are listed in Figure 6 and Table 4.

The yellow contours in the CoMFA and CoMSIA steric contour maps derived from compounds of the training set indicate that steric occupancy with bulk groups at the end of the C12-substituted groups of β -analogues may decrease activity. All of 32 analogues are β -analogues. It can be therefore concluded that most of the analogues may be less power than dihydroartemisinin (**11**). This is highly consistent with the actual activities (Table 4).

As shown in Table 4, some analogues such as **5**–**8**, **13**–**15**, **30** and **32** exhibit weak antimalarial activity. Compounds **5**–**8** possess negative charges at the end of the

C12-substituted groups. The C16 atoms of **13**, **14** and **15** encode in the aromatic rings, which bring negative charges. A major blue contour locates in the CoMFA and CoMSIA electrostatic contour maps show that more negative charge is not favored around C16 of the β -analogues. Thus it can explain low activities of these compounds.

Compounds **29**–**32** are all C11 hydroxyl-substituted, however, **29** and **31** are more active than **30** and **32**. Inspection of the molecular docking structures (Fig. 3), we found that different from compounds **29** and **31**, the C11 methyl group of **30** and **32** may point directly to the hemin plane. The steric hindrance weakens the binding of the two compounds with hemin, which leads that compounds **30** and **32** are special inactive. These results all indicate that the experiment data can be well explained and both the 3-D-models of artemisinins complexing with hemin and the CoMFA and CoMSIA models on the basis of binding conformations are reliable.

After the qualitative explanation on the experiment data, molecular docking and 3-D-QSAR analyses were

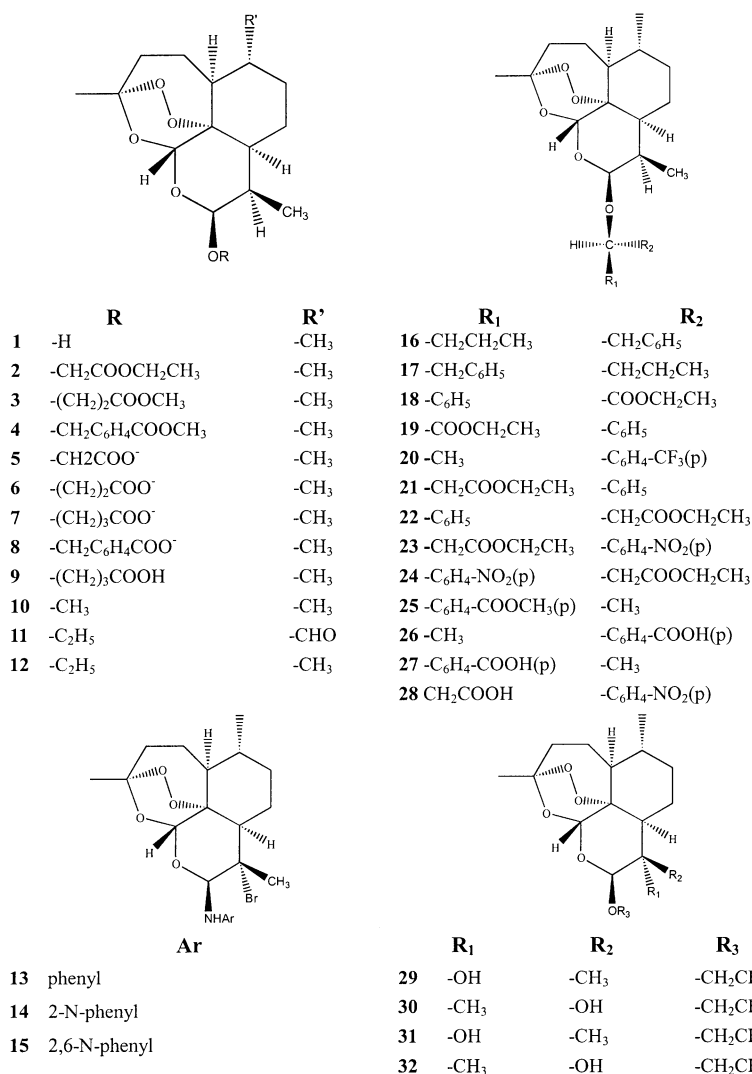


Figure 6. Structures of the 32 dihydroartemisinin analogues collected by Tonmunphean et al.¹¹ Compound **11** is DHQHS2 in the training set, it was involved in this testing set using its relative activity.

Table 4. Biological activities of the 32 dihydroartemisinin analogues and their binding energies with hemin derived from docking calculation

No.	Log(RA) ^a	E _{binding} (kcal/mol)
1	0.854	−49.262
2	0.689	−50.781
3	0.202	−49.234
4	0.580	−58.014
5	−1.264	−46.100
6	−1.463	−45.629
7	−1.411	−48.384
8	0.226	−52.075
9	−0.786	−49.998
10	0.423	−50.674
11	0.079	−49.873
12	0.146	−48.066
13	−0.786	−46.342
14	−1.139	−46.272
15	−1.666	−46.728
16	−0.122	−53.049
17	0.375	−51.803
18	0.904	−52.237
19	0.655	−55.397
20	0.199	−53.733
21	0.667	−49.613
22	0.700	−51.305
23	0.612	−49.672
24	0.971	−53.419
25	0.522	−52.617
26	−0.399	−47.725
27	−0.105	−50.842
28	−0.094	−49.592
29	0.255	−47.306
30	−0.824	−35.654
31	−0.347	−47.198
32	−1.097	−36.352

^aThe activities were measured against one strains of *Plasmodium falciparum*, D-6 clone. RA: the relative activity, which was calculated by the equation as follows: $\log(\text{RA}) = \log[(\text{artemisinin IC}_{50}/\text{analogue IC}_{50})(\text{analogue } M_r/\text{artemisinin } M_r)]$.

then performed on these 32 compounds using the same process as for the training compounds. The binding energies calculated by the molecular docking method are listed in Table 4. PLS analysis indicates that an acceptable correlation exists between the total binding energies, E_{total} s, and the activities of 30 analogues (eq 4) with the exception of compounds **30** and **32**. The deviation of these two compounds may be caused by the steric hindrance of the methyl with the side chains of hemin as mentioned above. It shows the binding energy between artemisinin and hemin is indeed associated with biological significant.

$$-\log C = 2.339 - 0.082 \times E_{\text{total}} \quad (4)$$

$$n = 30, s = 0.580, r = 0.670$$

The subsequent analysis gives $r_{\text{cv}}^2 = 0.654$, $n = 5$, $r^2 = 0.945$, $F = 169.25$, $s = 0.174$ for CoMFA model and $r_{\text{cv}}^2 = 0.570$, $n = 5$, $r^2 = 0.945$, $F = 90.12$, $s = 0.199$ for CoMSIA. The non-cross-validated calculations are shown in Table 5. These parameters all show good conventional statistical correlations. These good statistical results indicate a further application for our binding conformation recognition approach as an appropriate alignment

Table 5. Experimental and predictive activities of CoMFA and CoMSIA for the 32 dihydroartemisinin analogues

No.	EA ^a	CoMFA		CoMSIA	
		PA ^b	δ^c	PA	δ
1	0.85	0.72	0.13	0.36	0.49
2	0.69	0.78	−0.09	1.01	0.32
3	0.20	0.16	0.04	0.26	−0.06
4	0.58	0.49	0.09	0.54	0.04
5	−1.26	−1.27	0.01	−1.25	−0.01
6	−1.46	−1.43	−0.03	−1.51	0.05
7	−1.41	−1.37	−0.04	−1.45	0.04
8	0.23	0.16	0.07	0.47	−0.24
9	−0.79	−0.71	−0.08	−0.66	−0.13
10	0.42	0.33	0.09	0.29	0.13
11	0.08	0.17	−0.09	−0.09	0.17
12	0.15	0.13	0.02	0.15	0.00
13	−0.79	−1.17	0.38	−1.00	0.21
14	−1.14	−1.28	0.14	−1.25	0.11
15	−1.67	−1.39	−0.27	−1.33	−0.34
16	−0.12	−0.01	−0.11	−0.08	−0.04
17	0.38	0.49	−0.11	0.39	−0.01
18	0.90	1.03	−0.13	0.79	0.11
19	0.65	0.68	−0.03	0.64	0.01
20	0.20	0.31	−0.11	0.33	−0.13
21	0.67	0.80	−0.13	0.65	0.02
22	0.70	0.74	−0.04	0.63	0.07
23	0.61	0.64	−0.02	0.27	0.34
24	0.97	0.76	0.21	1.02	−0.05
25	0.52	0.26	0.26	0.24	0.28
26	−0.40	−0.51	0.11	−0.37	−0.03
27	−0.11	−0.05	−0.06	0.12	−0.22
28	−0.09	0.09	−0.18	0.15	−0.24
29	0.25	−0.16	0.41	0.38	−0.13
30	−0.82	−0.81	−0.01	−0.70	−0.12
31	−0.35	−0.25	−0.10	−0.31	−0.04
32	−1.10	−0.79	−0.31	−1.15	0.05

^aExperimental activities.

^bPrediction.

^cResidual values.

engine for CoMFA and CoMSIA calculation on other artemisinin analogues.

Conclusions

In this study, Molecular docking and 3-D-QSAR were successfully combined. The binding conformations of a series of artemisinin analogues were firstly determined by using docking procedure. Based on the docking model, consistent and highly predictive 3-D-QSAR models were derived, which are consistence with the experimental results and could be mapped back to the original the ligand–receptor complex models. This leads to a better understanding of antimalarial mechanism and thereby provides guidelines for new artemisinin analogues design.

Acknowledgements

We gratefully acknowledge financial support from the National Natural Science Foundation of China (Grant 29725203), the ‘863’ High-Tech Program of China (Grant 863–103–04–01), the State Key Program of Basic Research of China (Grant 1998051115).

References and Notes

1. Hien, T. T.; White, N. J. *Lancet* **1993**, 341, 603.
2. Li, Y.; Yu, P. L.; Chen, Y. X.; Li, L. Q.; Gai, Y. Z.; Wang, D. S. *Ke Xue Tong Bao* **1979**, 24, 667.
3. Jiang, H. L.; Chen, K. X.; Tang, Y.; Chen, J. Z.; Wang, Q. M.; Ji, R. Y.; Zhuang, Q. K. *Ind. J. Chem.* **1997**, 36B, 154.
4. Meshnick, S. R. *Ann. Trop. Med. Parasitol.* **1996**, 90, 367.
5. Marshall, G. R.; Cramer, R. D. d. *Trends Pharmacol. Sci.* **1988**, 9, 285.
6. Klebe, G.; Abraham, U. *J. Comput. Aided Mol. Des.* **1999**, 13, 1.
7. Paitayatat, S.; Tarnchompoo, B.; Thebtaranonth, Y.; Yuthavong, Y. *J. Med. Chem.* **1997**, 40, 633.
8. Wu, J. A.; Ji, R. Y. *Chung Kuo Yao Li Hsueh Pao* **1982**, 3, 55.
9. Wu, W. M.; Wu, Y. K.; Wu, Y. L.; Yao, Z. J.; Zhou, C. M.; Li, Y.; Shan, F. *J. Am. Chem. Soc.* **1998**, 120, 3316.
10. Shukla, K. L.; Gund, T. M.; Meshnick, S. R. *J. Mol. Graph.* **1995**, 13, 215.
11. Tonmunphean, S.; Kokpol, S.; Parasuk, V.; Wolschann, P.; Winger, R. H.; Liedl, K. R.; Rode, B. M. *J. Comput. Aided Mol. Des.* **1998**, 12, 397.
12. Bush, B. L.; Nachbar, R. B. *J. Comput. Aided Mol. Des.* **1993**, 7, 587.
13. Chen, H. Y.; Chen, Y.; Zhu, S. M.; Bian, N. S.; Shan, F.; Li, Y. *Talanta* **1999**, 48, 143.
14. Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, 97, 1303.
15. Ye, B.; Wu, Y. L.; Li, G. F.; Jiao, X. Q. *Yao Hsueh Hsueh Pao* **1991**, 26, 228.