Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Medicinal Chemistry 39 (2004) 745–753
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
www.elsevier.com/locate/ejmech

Original article

3D QSAR studies on antimalarial alkoxylated and hydroxylated chalcones by CoMFA and CoMSIA

C.X. Xue [a], S.Y. Cui [a], M.C. Liu [a], Z.D. Hu [a,*], B.T. Fan [b]

[a] *Department of Chemistry, Lanzhou University, Lanzhou 730000, P.R. China*
[b] *Université Paris 7-Denis Diderot, ITODYS 1, rue Guy de la Brosse, 75005 Paris, France*

Received 7 October 2003; received in revised form 19 May 2004; accepted 27 May 2004

Available online 11 August 2004

Abstract

The 3D QSAR analyses of antimalarial alkoxylated and hydroxylated chalcones were first conducted by Comparative molecular field analysis (CoMFA) and Comparative similarity indices analysis (CoMSIA) to determine the factors required for the activity of these compounds. Satisfactory results were obtained after performing a leave-one-out (LOO) cross-validation study with cross-validation q^2 and conventional r^2 values of 0.740 and 0.972 by the CoMFA model, 0.714 and 0.976 by the CoMSIA model, respectively. The results provided the tools for predicting the affinity of related compounds, and for guiding the design and synthesis of novel and more potent antimalarial agents.

Keywords: CoMFA; CoMSIA; Malaria; Chalcone; QSAR

1. Introduction

Despite years of continual effort, malaria is still one of the most deadly diseases affecting third-world countries, claiming more than one million lives annually, most of whom are children [1]. A major thrust to combat the malaria is the identification of new targets that are critical to the disease process or essential for the survival of the parasite. Recently, there has been strong interest in the potential antimalarial activity of chalcones [2–5]. Chemically, chalcone is 1,3-diphenyl-2-propen-1-one (see Fig. 1). Depending on the substitution pattern on the two aromatic rings, a wide range of pharmacological activities have been identified for various chalcones. These include antileishmanial [6], anti-inflammatory [7], antimitotic [8], and modulation of P-glycoprotein-mediated multidrug resistance [9].

The antimalarial activity of chalcones was first noted when licochalcone A, a natural product isolated from Chinese liquorice roots, was reported to exhibit potent in vivo and in vitro antimalarial activity [2]. Subsequently, a synthetic analogue, 2,4-dimethoxy-4′butoxychalcone, was reported to have outstanding antimalarial activity [3]. Antimalarial chalcones are widely thought to act against malarial cysteine protease [4], an enzyme used by the parasite for the degradation of host hemoglobin for its nutritional purposes. Chalcones are very readily synthesized, and various substitution patterns can be attempted on the two aromatic rings to give a large number of potential analogues. A sound understanding of the structural requirements for antimalarial activity in the chalcones is important in guiding and optimizing drug design efforts. However, a review of the literature has shown that no comprehensive structure-activity relationships of chalcones have been reported.

Quantitative structure activity relationship (QSAR) based on the three dimensional (3D) structures of the ligands alone involves two methods: Comparative molecular field analysis (CoMFA) [10] and Comparative similarity indices analysis (CoMSIA) [11]. Recently, more advanced techniques have

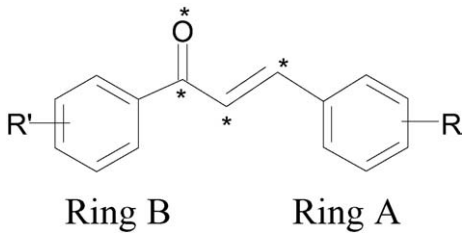

Fig. 1. General structure of chalcones. Stars indicate the atoms selected as the fitting centers.

* Corresponding author. Tel.: +86-931-891-2578; fax: +86-931-891-2582.
E-mail address: huzd@lzu.edu.cn (Z.D. Hu).

0223-5234/$ - see front matter © 2004 Elsevier SAS. All rights reserved.
doi:10.1016/j.ejmech.2004.05.009

attempted to model the receptor environment from the perspective of the ligand structure. QSAR studies incorporate 3D information for the ligands and provide a more detailed analysis of the ligand–receptor interactions.

In the present study, the 3D QSAR studies of 63 ring B alkoxylated and 30 ring B hydroxylated chalcones by CoMFA and CoMSIA were performed. Thus, the resulting CoMFA and CoMSIA studies will not only illustrate the conformation or spatial orientation of antimalarial chalcones but also provide useful indicators for further design of new drug candidates for malaria.

2. Materials and methods

2.1. Data set for analysis

A data set of 93 compounds was taken from the literature [12]. The structures of chalcones used in this study were listed in Table 1 and their biological activity IC_{50} values μM (for inhibition of [^{3}H] hypoxanthine uptake into *P. falciparum* (K1) in the presence of drug), which were expressed as pIC_{50} (the –log IC_{50}) were given in Tables 2 and 3. The general structure of chalcones was shown in Fig. 1. Of the compounds listed in Table 1, there are 38 compounds, which are not listed in Chemical Abstracts databases (1967 to date). The pIC_{50} was used as a dependent variable in the QSAR study. The whole data set was randomly divided into two subsets: the training set (2, 3, 4, 6...) and the test set (1, 5, 9...) (69 and 24 points, respectively). The training set of chalcones was used for 3D QSAR analysis. In addition, 24 compounds, selected from the various structures of chalcones, were kept to test the actual prediction of the model.

2.2. Computer modeling

All molecule modeling techniques and 3D QSAR studies described herein were performed on SGI INDIG O_2 workstations using SYBYL 6.9 molecular modeling software [13]. The use of a reasonably low energy conformation in the alignment is a useful starting point for statistical comparisons of flexible structures within the CoMFA and CoMSIA models. In this study, each structure of 93 compounds was first minimized using molecular mechanics with the MMFF94 force field with a 0.05 kcal/mol energy gradient convergence criterion. Charges were calculated by the MMFF94 method at the beginning and Gasteiger–Huckel was considered for the further calculations.

2.3. Alignment rules

Structural alignment is perhaps the most subjective, yet critical, step in CoMFA study, inasmuch as experience shows that the resulting 3D QSAR model is often sensitive to the particular alignment scheme [14]. Here one thing should be pointed out: in contrast to CoMFA, CoMSIA is not sensitive to changes in orientation of the superimposed molecules in the lattice [15]. The selected template molecule is typically one of the following [16,17]: (1) the most active compound; (2) the lead and/or commercial compound; (3) the compound containing the greatest number of functional groups. Generally, the low energy conformation of the most active compound is set as a reference.

The conformational search was performed using the multisearch routine in SYBYL. Compound 7 (most active compound) was chosen as the template molecule, on which other molecules were aligned.

The following are three alignments of superimposition [18,19]: (1) rms fitting, which means all of the molecules are aligned by minimizing the rms distance between atom pairs belonging respectively to the fitting molecule and to the template molecule; the atoms used for alignment were marked with asterisk (*) in Fig. 1. Fig. 2 showed the aligned molecules (including the test set). (2) Flexible fitting (multifit). In this case, alignment of the molecules is carried out by flexible fitting (multifit) of atoms of the molecules to the template molecule. This involves energy calculation and fitting onto the template molecule by applying force (force constant = 20 kcal/mol Å) and subsequent energy minimization. The aligned molecules were shown in Fig. 3. (3) It is carried out by using the SYBYL QSAR rigid body field fit command within SYBYL. Field fit adjusts the geometry of the molecule so that its steric and electrostatic fields match the fields of the template molecule.

In the present study, only the first and second alignment rules, which are often used for molecular superimposition were adopted. The molecular superimposition was nearly very similar in the two alignments.

2.4. CoMFA analysis

Following alignment, the molecules are placed one by one into a 3D cubic lattice with 2 Å grid. The steric (van der Waals) and electrostatic (Coulombic with 1/r dielectric) fields are calculated at each grid point using an sp^3-hybridized carbon probe with a +1.0 charge. A 30 kcal/mol energy cutoff was applied, which means the steric and electrostatic energies greater than 30 kcal/mol are truncated to that value, thus, can avoid infinity of energy values inside molecule.

As for modeling methods, partial least squares (PLS) linear regression was used to correlate the activities (pIC_{50}) with the CoMFA values that contain the magnitude of either the steric or electrostatic potentials.

Firstly, to avoid overfitted 3D QSAR, the optimum number of components *(N)* used in the model derivation is chosen from the analysis with the highest cross-validated correlation coefficient (q^2), which is defined as follows:

$$q^2 = 1 - \frac{\sum(Y - Y_{\text{pred}})^2}{\sum(Y - Y_{\text{mean}})^2}$$

The cross-validated q^2 quantifies the predictive ability of the model. It was determined by a leave-one-out (LOO)

Table 1
Structure of chalcones

No.	R′	R	No.	R′	R
1	2′,3′,4′-trimethoxy	2,4-dichloro	48	2′,4′-dimethoxy	4-ethyl
2		4-dimethylamino	49		3-quinolinyl
3		4-trifluoromethyl	50		4-quinolinyl
4		2,4-dimethoxy	51		4-methyl
5		4-methyl	52		4-methoxy
6		4-ethyl	53		4-dimethylamino
7		3-quinolinyl	54		4-fluoro
8		4-quinolinyl	55		4-chloro
9		4-methoxy	56		4-bromo
10		4-fluoro	57		2-chloro-4-fluoro
11		4-phenyl	58		3,4-dichloro
12		2,4-difluoro	59		4-nitro
13		4-nitro	60		1-naphthalenyl
14		3,4-dichloro	61	4′-methoxy	4-hydroxy
15		4-chloro	62		2,4-difloro
16		2-chloro	63		4-methoxy
17		3-chloro	64		3-quinolinyl
18		H	65		4-quinolinyl
19	4′-butoxy	2,4-dimethoxy	66		4-fluoro
20	2′,4′-dimethoxy	2,4-dichloro	67		2,4-dichloro
21		4-trifluoromethyl	68		4-trifluoromethyl
22		2,4-difluoro	69		2,4-dimethoxy
23		2,4-dimethoxy	70		4-methyl
24		4-nitro	71	2′,4′-dihydroxy	2-naphthalenyl
25		4-dimethylamino	72		4-pyridinyl
26		4-cyano	73		4-quinolinyl
27		H	74		4-chloro
28	4′-ethoxy	2,4-difluoro	75	4′-hydroxy	1-naphthalenyl
29		4-methoxy	76		3-quinolinyl
30		3-quinolinyl	77		2-pyridinyl
31		4-quinolinyl	78		4-quinolinyl
32		4-fluoro	79		2-naphthalenyl
33		2,4-dichloro	80		4-chloro
34		4-trifluoromethyl	81		2-chloro
35		2,4-dimethoxy	82		3-chloro
36		4-methyl	83		4-methoxy
37		4-nitro	84		4-methyl
38		4-dimethylamino	85		3-methyl
39		4-cyano	86		4-butyl
40		H	87		4-trifluoromethyl
41	2′,4′-dihydroxy	2,4-dichloro	88		4-nitro
42		3-quinolinyl	89		4-fluoro
43		2,4-difluoro	90		3,4-dichloro
44		2,4-dimethoxy	91		4-dimethylamino
45		1-naphthalenyl	92		2,4-difluoro
46		4-trifluoromethyl	93		H
47		2-pyridinyl			

procedure of cross-validation in which each compound is successively removed from the model derivation and its pIC_{50} value can be predicted using the model built from the remaining compounds. To speed up the analysis and reduce noise, column filtering was set at 2.0 kcal/mol so that only those steric and electrostatic energies with values greater than 2.0 kcal/mol are considered in the PLS analysis. In above stage, the predictive quality of the ''best'' correlation model is determined. Then, the optimal number of components *(N)* is employed to do no validation PLS analysis to get the final model parameters such as correlation coefficient (r^2), standard error (SE) and F value. At the same time, the CoMFA color contour maps are derived for the steric and electrostatic fields. The quality of the final CoMFA model is measured by two statistical parameters: r^2 and q^2. The value of q^2, which indicates the predictive capacity of the model,

Table 2
Experimental activities and predictive activities by CoMFA and CoMSIA (training set)

No. [a]	pIC_{50}	CoMFA	Error	CoMSIA	Error	No. [a]	pIC_{50}	CoMFA	Error	CoMSIA	Error
2	4.74	4.68	–0.06	4.70	–0.04	48[b]	5.62	5.48	–0.14	4.53	0.11
3 [b]	5.52	5.19	–0.33	5.29	–0.23	50[b]	4.57	4.63	0.06	4.59	0.02
4 [b]	4.78	4.79	0.01	4.78	0	51	4.03	4.11	0.08	3.94	–0.09
6 [b]	4.78	4.90	0.12	4.86	0.08	52	3.89	3.81	–0.08	3.82	–0.07
7 [b]	5.70	5.68	–0.02	5.70	0	54	3.49	3.39	–0.10	3.41	–0.08
8 [b]	4.22	4.18	–0.04	4.16	–0.06	55	3.47	3.37	–0.10	3.40	–0.07
10 [b]	5.02	4.97	–0.05	4.78	–0.24	56	3.27	3.36	0.09	3.43	0.16
11 [b]	4.58	4.76	0.18	4.64	0.06	58[b]	3.53	3.58	0.05	3.57	0.04
12 [b]	4.73	4.64	–0.09	4.78	0.05	59	3.38	3.30	–0.08	3.41	0.03
14 [b]	4.84	4.88	0.04	4.88	0.04	60	3.49	3.52	0.03	3.51	0.02
15 [b]	4.84	4.96	0.12	4.78	–0.06	62[b]	4.57	4.66	0.09	4.66	0.09
16 [b]	4.38	4.47	0.09	4.60	0.22	63	4.66	4.85	0.19	4.55	–0.11
18	4.80	4.78	–0.02	4.88	0.08	64[b]	5.32	5.23	–0.09	5.37	0.05
19	3.97	4.07	0.10	3.97	0	66	4.84	4.87	0.03	4.90	0.06
20	4.73	4.72	–0.01	4.68	–0.05	67	4.80	4.64	–0.16	4.73	–0.07
22	5.21	5.27	0.06	5.30	0.09	68[b]	4.72	4.64	–0.08	4.77	0.05
23	5.68	5.65	–0.03	5.67	–0.01	70	4.15	4.22	0.07	4.19	0.04
24	4.00	4.02	0.02	3.98	–0.02	71[b]	4.70	4.77	0.07	4.66	–0.04
26	4.02	3.99	–0.03	3.95	–0.07	72	3.92	3.93	0.01	3.83	–0.09
27	4.26	4.15	–0.11	4.44	0.18	74	4.91	4.98	0.07	4.91	0
28 [b]	4.55	4.54	–0.01	4.46	–0.09	75[b]	4.40	4.33	–0.07	4.32	–0.08
30 [b]	4.60	4.55	–0.05	4.64	0.04	76[b]	4.39	4.34	–0.05	4.39	0
31 [b]	4.00	3.97	–0.03	4.02	0.02	78	4.29	4.42	0.13	4.32	0.03
32 [b]	4.62	4.72	0.10	4.62	0	79[b]	4.56	4.59	0.03	4.49	–0.07
34 [b]	4.62	4.62	0	4.67	0.05	80	4.42	4.47	0.05	4.45	0.03
35	4.52	4.39	–0.13	4.54	0.02	82	4.48	4.63	0.15	4.44	–0.04
36 [b]	4.42	4.45	0.03	4.53	0.11	83	4.49	4.51	0.02	4.47	–0.02
38	4.52	4.56	0.04	4.51	–0.01	84	4.60	4.58	–0.02	4.67	0.07
39 [b]	3.27	3.30	0.03	3.28	0.01	86[b]	4.68	4.65	–0.03	4.67	–0.01
40	4.37	4.33	–0.04	4.40	0.03	87	4.52	4.58	0.06	4.56	0.04
42 [b]	4.79	4.78	–0.01	4.77	–0.02	88	4.69	4.58	–0.11	4.71	0.02
43	4.80	4.86	0.06	4.83	0.03	90	4.74	4.74	0	4.73	–0.01
44	4.25	4.33	0.08	4.45	0.20	91	4.75	4.78	0.03	4.72	–0.03
46	4.58	4.45	–0.13	4.46	–0.12	92	4.61	4.63	0.02	4.59	–0.02
47	4.71	4.65	–0.06	4.54	0.02						

* Test set.

[a] The number is the serial number of chalcones in Table 1;
[b] Compounds not listed in Chemical Abstract database (1967 to present);

should be greater than 0.40; and the value of r^2, which shows the self-consistency of the model, should be greater than 0.90.

2.5. CoMSIA analysis

Similar to CoMFA, a data table is constructed from similarity indices calculated at the intersections of a regularly spaced lattice (2 Å grid) in CoMSIA. The similarity indices between the compounds and the probe atom are calculated according to:

$$A^q_{F,k}(j) = \sum_i w_{\text{probe},k} w_{ik} e^{-\alpha r^2_{kq}}$$

where A is the similarity index at grid point q, summed over all atoms i of the molecule j under investigation; $w_{\text{probe},k}$ is the probe atom with radius 1 Å, charge +1, hydrophobicity +1, hydrogen bond donating +1, hydrogen bond accepting +1; w_{ik} is the actual value of the physicochemical property k of atom i; r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the test molecule; α is the attenuation factor, and the default value of α is 0.3 [20].

Larger values of α will result in a steeper Gaussian function and an increasing attenuation of the distance-dependent effects of molecular similarity. On the other hand, reducing α to smaller values means that a probe placed at a particular lattice point detects molecular similarity in its neighborhood more globally.

In this study, five physicochemical properties were evaluated: steric, electrostatic, hydrophobic, and hydrogen-bond donor or acceptor properties. These fields are selected to cover the major contributions to ligand binding. Using all the five CoMSIA descriptors for the explanatory variables, a LOO run and a no validation PLS analysis were performed.

Table 3
Experimental activities and predictive activities by CoMFA and CoMSIA (test set)

No. [a]	PIC_{50}	CoMFA	Error	CoMSIA	Error
1	5.27	4.78	–0.49	4.92	–0.35
5 [b]	4.59	5.01	0.42	4.90	0.31
9 [b]	4.60	5.33	0.73	4.88	0.28
13 [b]	4.65	5.23	0.58	4.35	–0.30
17 [b]	4.61	4.86	0.25	4.81	0.20
21	5.23	5.17	–0.06	5.54	0.31
25	4.15	4.41	0.26	4.64	0.49
29	4.48	4.32	–0.16	4.31	–0.17
33	4.02	4.32	0.30	4.18	0.16
37	4.41	4.48	0.07	4.56	0.15
41	4.16	4.19	0.03	3.97	–0.19
45	4.61	4.43	–0.18	4.04	–0.57
49 [b]	5.66	5.56	–0.10	5.59	–0.07
53	4.26	3.88	–0.38	4.20	–0.06
57 [b]	3.22	3.59	0.37	3.25	0.03
61	5.15	4.43	–0.72	5.06	–0.09
65 [b]	4.37	4.43	0.06	4.53	0.16
69	5.19	4.78	–0.41	4.74	–0.45
73 [b]	4.03	4.32	0.29	4.50	0.47
77	4.79	4.70	–0.09	4.49	–0.30
81	4.21	4.06	–0.15	4.41	0.20
85	4.59	4.50	–0.09	4.47	–0.12
89	4.66	4.61	–0.05	4.56	–0.10
93	4.53	4.52	–0.01	4.56	0.03

[a] The number is the serial number of chalcones in Table 1;
[b] Compounds not listed in Chemical Abstract database (1967 to present).

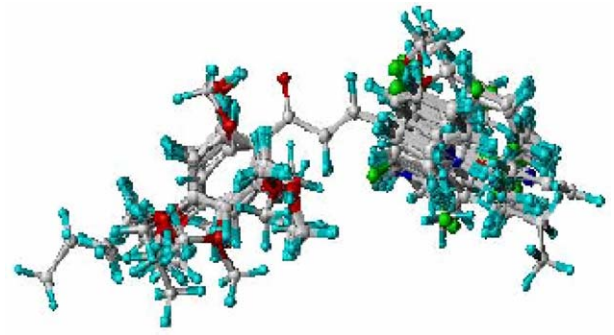

Fig. 2. 3D-view of all aligned compounds by alignment (1): rms fitting.

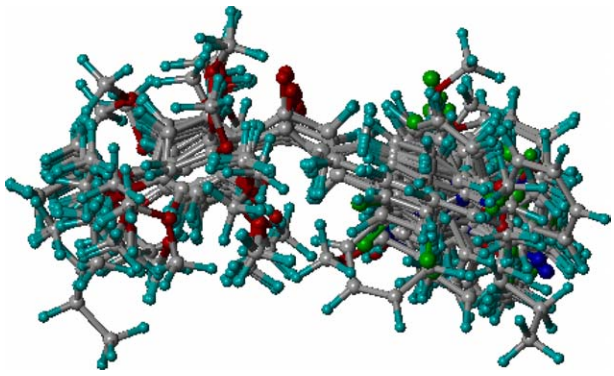

Fig. 3. 3D-view of all aligned compounds by alignment (2): flexible fitting.

3. Results and discussion

3.1. Results of the CoMFA analysis

The results from the CoMFA studies were summarized in Table 4. The alignment (1) (CoMFA (1)) with MMFF94 charges showed a cross-validated $q^2 = 0.740$ with six components, alignment (2) (CoMFA (2)) showed 0.720 with seven components. The model with alignment (1) was selected for further study judged by the cross-validated correlation. Gasteiger–Huckel charge was calculated and this model showed a reduced $q^2 = 0.709$ with six components (CoMFA (3)). So, the following discussions will only refer to the model generated by CoMFA (1). A non-cross-validated $r^2 = 0.972$ with $F = 357.227$ was also observed with this model. In this analysis almost equal contributions were observed from steric (45.6%) and electrostatic (54.4%) fields. The predicted activities using the 3D QSAR model from CoMFA (1), and the residue values of the training set and the test set were given in Tables 2 and 3, respectively. Figs. 4 and 5 showed the plot of observed versus predicted antimalarial activity for the training set and the test set, respectively.

The CoMFA steric and electrostatic fields based on PLS analysis were represented as 3D contour plots in Figs. 6 and 7, using compound 7 as reference structure. The steric contour map shows a green region surrounding the C-2′, 4′ position of the ring B, indicating that a bulky substituent is preferred in the position to produce higher antimalarial activity. This conclusion is well consistent with the experiments

Table 4
Summary of CoMFA model

Model	q^2	N	r^2	SE	F	Static	Elec
CoMFA(1)	0.740	6	0.972	0.093	357.227	0.456	0.544
CoMFA(2)	0.720	7	0.965	0.098	336.254	0.447	0.553
CoMFA(3)	0.709	6	0.958	0.107	319.847	0.451	0.549

that hydroxylated chalcones are generally less active than their alkoxylated counterparts (Table 1). Among the alkoxylated chalcones, 12 compounds (1, 3, 7, 10, 21, 22, 23, 48, 49, 61, 64, 69) have pIC_{50} values over 5.00 and the most active compound is 1-(2′, 3′, 4′-trimethoxyphenyl)-3-(3-quinolinyl)-2-propen-1-one (compound 7), with a pIC_{50} of 5.70. There is another green region near the C-3,4 position of the ring A implying that a relatively bulky substituent is beneficial to the antimalarial activity. The association of the 3-quinolinyl ring with good activity is an interesting recurring feature among all alkoxylated chalcones. The methoxychalcone with a substituted A ring (compound 27) has poor antimalarial activity (pIC_{50} = 4.26). The yellow region near C-4′ position of ring B indicated that a less bulky substituent is preferred for higher activity and this can explain well why 4′-butoxy-2,4-dimethoxy-chalcone and 2′,4′-dimethoxy-4-butoxy-chalcone (compound 19 and 52) has low activity (pIC_{50} = 3.97 and 3.89, respectively).

There is a major blue region around C-3,4 of ring A and C-2′ of ring B in the electrostatic contour map suggesting that the substituting which can increase the positive charge on the ring systems would result in a higher activity. Indeed, it is found that an electron-withdrawing group like a chlorine or fluorin atom of the ring A system increases the antimalarial activity, such as compounds 1, 3, 10, 21, 22 ($pIC_{50} > 5.00$). In addition, two red polyhedrals focus on the C-3′,4′ of ring B, which shows that where high electron density might play a favorable role in activity. This can also well explain why the hydroxylated chalcones are generally less active than their alkoxylated counterparts. The 3′ and 4′ substituted analogues 23, 48, 61, 69 ($pIC_{50} > 5.00$) have more potent activity.

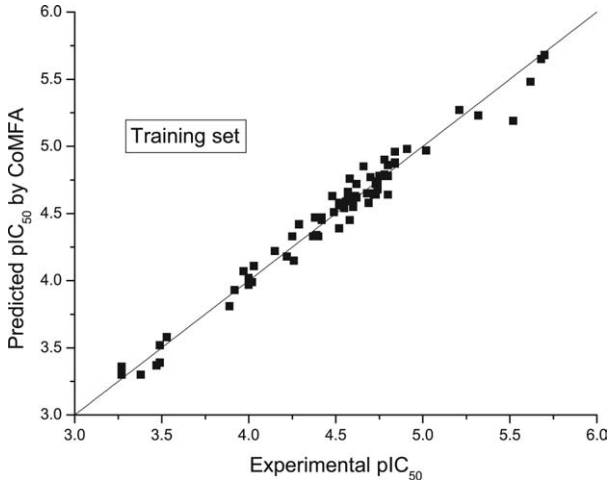

Fig. 4. Predictive vs. experimental pIC_{50} values derived from the CoMFA model of the training set.

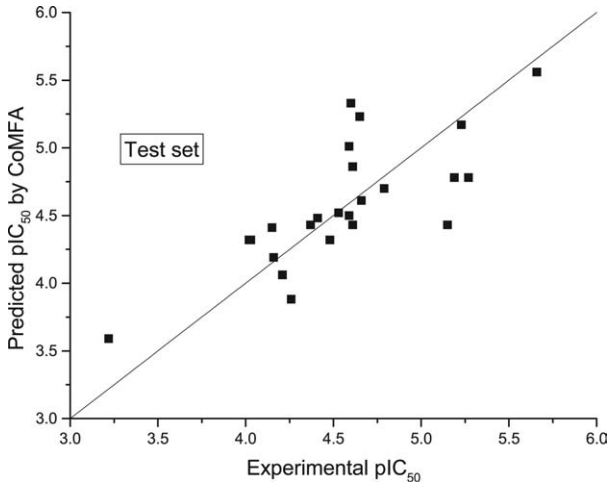

Fig. 5. Predictive vs. experimental pIC_{50} values derived from the CoMFA model of the test set.

3.2. Results of the CoMSIA analysis

The results of CoMSIA analysis were presented in Table 3. All analyses were based on a grid spacing of 2.0 Å,

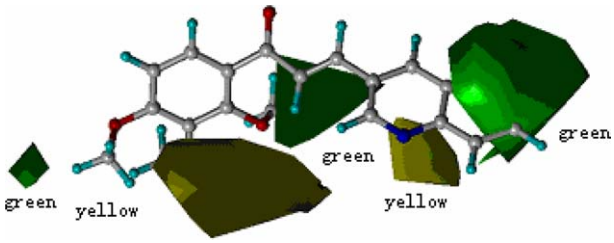

Fig. 6. The steric contour maps of CoMFA model (st.dev. × coeff.). The favorable steric areas (contribution levels 80%) with more bulk are indicated by green isopleths, where as the disfavorable steric areas (contribution levels 20%) are shown by yellow isopleths.

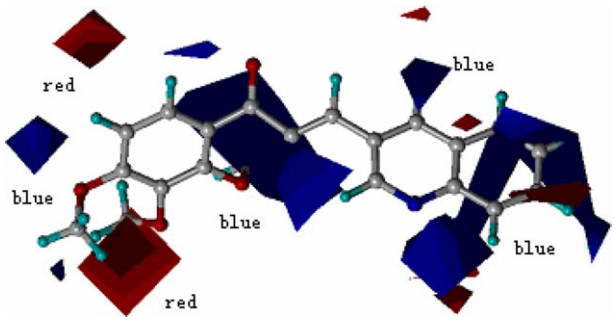

Fig. 7. The electrostatic contour maps of CoMFA model (st.dev. × coeff.). The favorable electrostatic areas (contribution levels 80%) with positive charges are indicated by blue isopleths, whereas the favorable electrostatic areas (contribution levels 20%) with negative charges are show by red isopleths.

the column filtering was set at 2.0 kcal/mol. Here, different α was set. Compared with CoMSIA (5) ($q^2 = 0.709$, $N = 10$), the CoMSIA (4) gave an improved cross-validated correlation coefficient ($q^2 = 0.704$) with a lower number of optimal components ($N = 8$), so, CoMSIA (4) was selected for further study. This model gave a non-cross-validated r^2 of 0.972 with $F = 357.227$ and SE = 0.088. The calculated activities and the residue values predicted by the CoMSIA (4) model for the training set and the test set were given in Tables 2 and 3, respectively. Figs. 8 and 9 showed the plot of observed vs. calculated antimalarial activity for the training set and the test set, respectively.

The CoMSIA approach presently considered five different property fields. The purpose of using five different fields is not to increase the significance and predictive power of the 3D QSAR models but to partition the various properties into spatial locations where they play a decisive role in determining biological activity. At this point, the major advantage of CoMSIA is the better ability to visualize and interpret the obtained correlations in terms of field contributions. Strictly speaking, the plots representing isocontours of the obtained coefficients from PLS indicate those lattice points where a particular property significantly contributes and thus explain the variation in affinity data. An excellent insight into the relationships between structure and activity for the different physicochemical properties of the considered structures were given by CoMSIA contour maps. The isocontour plots of CoMSIA (Figs. 10–14) were field contributions of the steric, electrostatic, hydrophobic, hydrogen-bond donor and acceptor properties, respectively. And the molecule in all the maps was the most active compound 7.

In the steric contour maps (Fig. 10), areas contoured by green indicate regions favorable for steric occupancy, while areas contoured by yellow indicate the opposite. The green

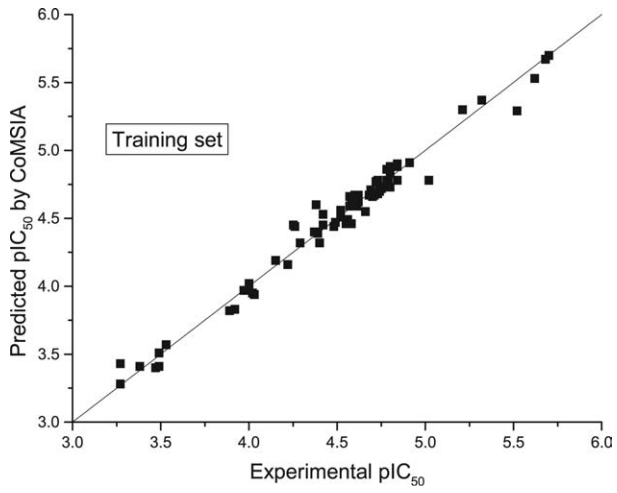

Fig. 8. Predictive vs. experimental pIC_{50} values derived from the CoMSIA model of the training set.

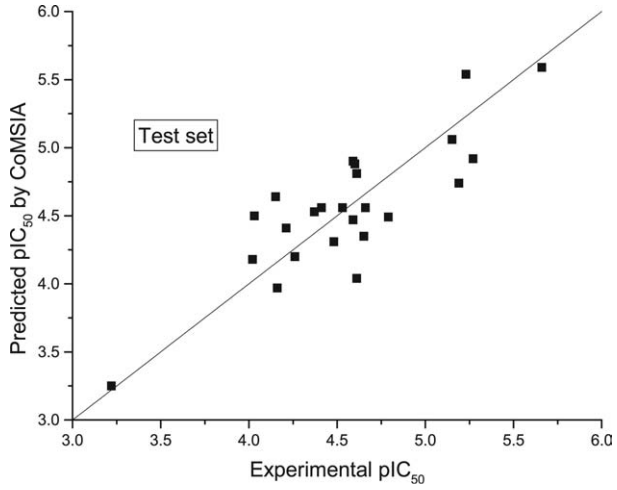

Fig. 9. Predictive vs. experimental pIC_{50} values derived from the CoMSIA model of the test set.

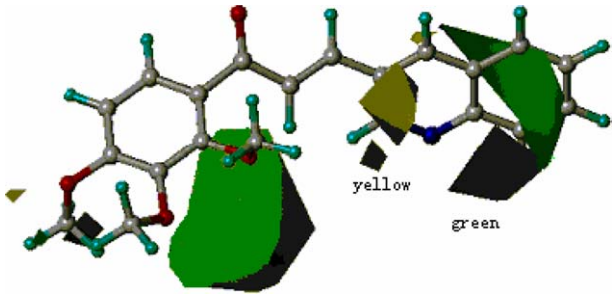

Fig. 10. The steric contour maps of CoMSIA model (st.dev. × coeff.). The favorable steric areas (contribution levels 80%) with more bulk are indicated by green isopleths, where as the disfavorable steric areas (contribution levels 20%) are shown by yellow isopleths.

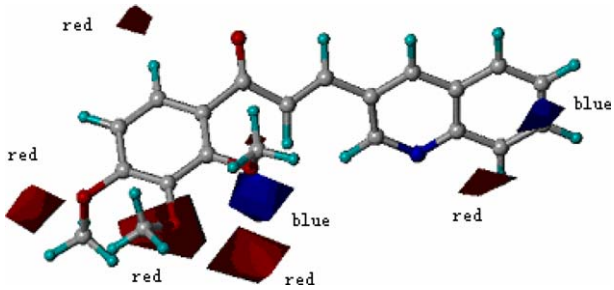

Fig. 11. The electrostatic contour maps of CoMSIA model (st.dev. × coeff.). The favorable electrostatic areas (contribution levels 80%) with positive charges are indicated by blue isopleths, whereas the favorable electrostatic areas (contribution levels 20%) with negative charges are shown by red isopleths.

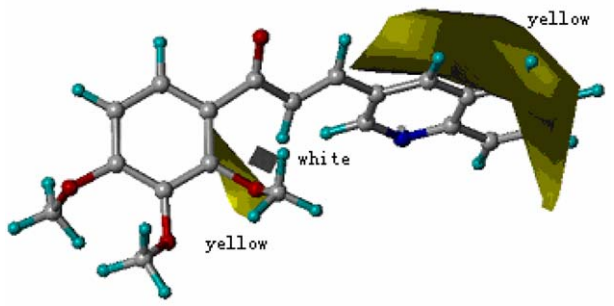

Fig. 12. The hydrophobic contour maps of CoMSIA model (st.dev. × coeff.). The favorable hydrophobic areas (contribution levels 80%) indicated by yellow isopleths, whereas the disfavorable hydrophobic areas (contribution levels 20%) are shown by white isopleths.

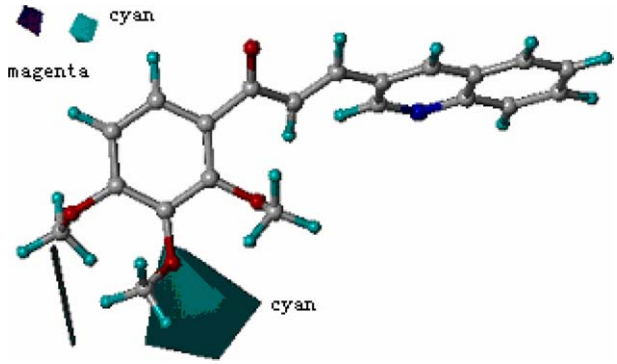

Fig. 13. The hydrogen-bond donor contour maps of CoMSIA model (st.dev. × coeff.). Cyan isopleths contour maps (contribution levels 80%) beyond the ligands where an hydrogen-bond donor group in the ligand will be favorable for biological activity, while purple isopleths (contribution levels 20%) represents hydrogen-bond accepter in the ligands unfavorable for bioactivity.

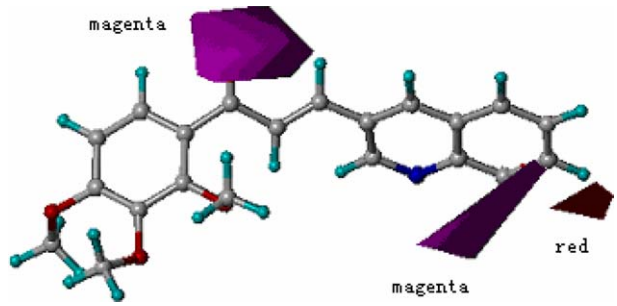

Fig. 14. The hydrogen-bond acceptor contour maps of CoMSIA model (st.dev. × coeff.). Magenta isopleths contour maps (contribution levels 80%) beyond the ligands where an hydrogen-bond acceptor group in the ligand will be favorable for biological activity, while red isopleths (contribution levels 20%) represents hydrogen-bond donor in the ligands unfavourable for bioactivity.

polyhedra of Fig. 10 represent a preferred occupancy of the pocket of the acceptor. Thus groups of increasing steric bulk in this region will enhance binding affinity and then increase the activity. An alkoxylated B ring would be more advantageously placed in the green zones and this was indeed observed when the steric contours around the 2′, 4′ dihydroxy chalcone (compound 44, pIC_{50} = 4.25) and its 2′, 4′-dimethoxy analogue (compound 23, pIC_{50} = 5.68) were compared. The distribution of steric contours around ring B in this model would explain why alkoxylated chalcones are generally better antimalarials than their hydroxylated counterparts. This coincided with the contours of CoMFA (Fig. 6).

In Fig. 11, the electrostatic properties were summarized. Areas where negatively charged groups enhance activity are contoured by red polyhedra and where positively charged groups increase activity are surrounded by blue isopleths. Red areas were observed surrounding the C-3′, 4′ position of the ring B, which means negative charge in these regions corresponds to high activity. An electron-rich substituent would be more favorably placed in the red zones and this was indeed observed when dimethoxy chalcone (compounds 23 and 48 with pIC_{50} values of 5.68 and 5.62) and its dihydroxy analogue (compounds 44 and 42 with pIC_{50} values of 4.25 and 4.79) were compared. The distribution of electrostatic contours in this model coincided with the contours of CoMFA (Fig. 7).

Hydrophobicity is one of the most important properties related to biomolecular interactions. In CoMSIA study, hydrophobic similarity index fields were constructed too and hydrophobic contour maps were shown in Fig. 12. In general, orange polyhedra indicate that hydrophobic substituents are good for increasing the potency, while hydrophilic substituents are beneficial to the activity at the regions of white contours. There is an orange area near the C-3, 4 position of the ring A, which means favorable for hydrophobic. The same region has been indicated in the steric maps (Figs. 6 and 10) in green to favor a more bulk group. This correlation can be explained by analyzing the trends in biological data for some of the compounds, using compounds 7, 49 and 64, which have pIC_{50} values over 5.00 as examples. Thus, the hydrophobic and bulky group of quinoline ring would be suitably placed in ring A.

The graphical interpretation of the field contributions of the hydrogen-bond property was shown in Fig. 13 (hydrogen-bond donor field) and Fig. 14 (hydrogen-bond acceptor field). In principle, they should highlight areas beyond the ligands where putative hydrogen partners in the acceptor could form hydrogen-bonds that influence binding affinity significantly. Cyan and magenta contours in these maps indicate that groups possessing the property will be favorable for binding, whereas purple and red contours show that this property should be absent in this area. In the donor field (Fig. 13), the cyan polyhedra around the compound represent areas where the presence of a hydrogen-bond donor would increase affinity. A magenta isopleth in the acceptor field (Fig. 14) surrounds C-3, 4 position of ring A, indicating this area to be favorable for hydrogen-bond acceptors. This correlation can be explained by analyzing the trends in biological data for some of the compounds, using compounds 47, 77 and 88, which have pIC_{50} values greater than or equal to 4.69 as examples.

By analyzing the results in Table 5 obtained by the CoMFA and the CoMSIA models, respectively, it can be seen that the relative error for compound 9, 13, and 61 was 15.9%, 12.5%, and –14.0% predicted by CoMFA model, and 6.1%, –6.5%, and –1.7%, by CoMSIA model, respectively, showing that these compounds were predicted better by the CoMSIA model than by CoMFA. This might be well due to the addition of the hydrogen-bond field compared with CoMFA, which is the benefit of the CoMSIA analysis over the CoMFA. The three compounds mentioned above have a function group of hydrogen-bond acceptor at C-4 position of ring A, but in the CoMFA model, no hydrogen-bond field was considered like in the CoMSIA model, leading to the poor predicted results of these compounds. Indeed, Fig. 14 proved again the existence of the field of the hydrogen-bond acceptor. On the contrary, some compounds results was quite surprising for the activity was predicted more poorly by the CoMSIA model than by the CoMFA model, particularly for compounds 45.

Table 5
Summary of CoMSIA model

Model	*A*	q^2	*N*	r^2	*SE*	*F*	Static	Elec [a]	Hydro [b]	H-d [c]	H-a [d]
CoM-SIA(1)	0.2	0.380	6	0.819	0.233	57.181	0.126	0.449	0.243	0.055	0.127
CoM-SIA(2)	0.3	0.559	10	0.893	0.161	155.850	0.142	0.398	0.249	0.058	0.153
CoM-SIA(3)	0.4	0.681	10	0.954	0.120	182.951	0.163	0.322	0.282	0.057	0.176
CoM-SIA(4)	0.5	0.714	8	0.976	0.088	237.032	0.160	0.332	0.268	0.084	0.155
CoM-SIA(5)	0.6	0.709	10	0.969	0.101	222.425	0.165	0.310	0.282	0.056	0.186

[a] Relative contribution of electrostatic fields.
[b] Relative contribution of hydrophobic fields.
[c] Relative contribution of hydrogen-bond donor fields.
[d] Relative contribution of hydrogen-bond acceptor fields.

The level-dependent contouring of CoMFA field contributions highlights those regions in space where the aligned molecules would favorably or unfavorably interact with a possible environment. In contrast, the CoMSIA field contributions denote those areas within the region occupied by the ligands that 'favor' or 'dislike' the presence of a group with a particular physicochemical property. Thus, contour maps of CoMSIA are easier to interpret by enabling partitioning of variance into the different field types. In summary, the CoMFA and CoMSIA models for antimalarial activity indicated a preference for a sterically large (alkoxylated) ring B and a hydrophobic ring A.

4. Conclusion

In this study, the 3D QSAR analyses, CoMFA and CoMSIA were applied to predict the antimalarial activity of a set of chalcones. The QSAR models gave a good statistical results in terms of q^2 and r^2 values. The CoMFA model provided the most significant correlation of steric and electrostatic fields with the biological activities. The effects of the steric, electrostatic, hydrophobic, and hydrogen-bond donor and acceptor fields around the aligned molecules on their activities were clarified by analyzing the CoMSIA contour maps. The information obtained in this study provides the tools for predicting the affinity of related chalcones compounds, and for guiding further structural modifying and synthesizing new potent antimalarial agents.

Acknowledgments

This work was supported by the Ministry of Sciences and Technology of China and the Ministry of Foreign Affairs of France (AFCRST PRA SI 02-02).

References

[1] M. Wahlgren, M.T. Bejarano, Nature 400 (1999) 506–507.
[2] M. Chen, T.G. Theander, B.S. Christensen, L. Hviid, L. Zhai, A. Kharazmi, Antimicrob., Agents Chemother. 38 (1994) 1470–1475.
[3] M. Chen, B.S. Christensen, L. Zhai, M.H. Rasmussen, T.G. Theander, S. Frokjaer, S. Sterffansen, J. Davidsen, A. Kharazmi, J. Infect. Dis. 176 (1997) 1327–1333.
[4] R. Li, G.L. Kenyon, F.E. Cohen, X. Chen, B. Gong, J.N. Dominguez, E. Davidson, G. Kurzban, R.E. Miller, E.O. Nuzum, P.J. Rosenthal, J.H. Mekerrow, J. Med. Chem. 38 (1995) 5031–5037.
[5] V.J. Ram, A.S. Saxena, S. Srivastava, S. Chandra, Bioorg. Med. Chem. Lett. 10 (2000) 2159–2161.
[6] S.F. Nielsen, S.B. Christensen, G. Cruciani, A. Kharazmi, T. Liljefors, J. Med. Chem. 41 (1998) 4819–4832.
[7] F. Herencia, M.L. Ferrandia, A. Ubeda, J.N. Dominguez, J.E. Charris, G.M. Lobo, M.J. Alcaraz, Bioorg. Med. Chem. Lett. 8 (1998) 1169–1174.
[8] S. Ducki, R. Forrest, J.A. Hadfield, A. Kendall, N.J. Lawrence, A.T. Mcgown, D. Rennison, Bioorg. Med. Chem. Lett. 8 (1998) 1051–1056.
[9] F. Bois, A. Boumendjel, A. Mariotte, G. Conseil, A. Di Petro, Bioorg. Med. Chem. 7 (1999) 2691–2695.
[10] R.D. Cramer III, D.E. Patterson, J.D. Vunce, J. Am. Chem. Soc. 110 (1988) 5959–5967.
[11] G. Klebe, U. Abraham, T. Mietzner, J. Med. Chem. 37 (1994) 4130–4146.
[12] M. Liu, P. Wilairat, M.L. Go, J. Med. Chem. 44 (2001) 4443–4452.
[13] , Tripos Associates, St. Louis, MO, USA, 1999;http://www.tripos.com/ available from.
[14] X.J. Zou, L.H. Lai, G.Y. Jin, Z.X. Zhang, J. Agric. Food Chem. 50 (2002) 3757–3760.
[15] M. Bohm, J. Sturzebecher, G. Klebe, J. Med. Chem. 42 (1999) 458–477.
[16] A. Agarwal, E.W. Taylor, J. Comput. Chem. 14 (1993) 237–245.
[17] N. Baurin, E. Vangrevelinghe, L. Morin-Allory, J.Y. Mérour, P. Renard, M. Payard, G. Guillaumet, C. Marot, J. Med. Chem. 43 (2000) 1109–1122.
[18] E.R. Collantes, X. Li, J. Agric. Food Chem. 47 (1999) 5245–5251.
[19] A.K. Debnath, J. Med. Chem. 42 (1999) 249–259.
[20] G. Klebe, U. Abraham, J. Comput. Aid. Mol. Des. 13 (1999) 1–10.