

3D-QSAR analysis of antimalarial farnesyltransferase inhibitors based on a 2,5-diaminobenzophenone scaffold

Aihua Xie,^a Prasanna Sivaprakasam^a and Robert J. Doerksen^{a,b,*}

^aDepartment of Medicinal Chemistry, School of Pharmacy, University of Mississippi, MS 38677-1848, USA

^bResearch Institute of Pharmaceutical Sciences, School of Pharmacy, University of Mississippi, MS 38677-1848, USA

Received 7 June 2006; revised 20 June 2006; accepted 21 June 2006

Available online 11 July 2006

Abstract—With annual death tolls in the millions and emerging resistance to existing drugs, novel therapies are needed against malaria. Wiesner et al. recently developed a novel class of antimalarials derived from farnesyltransferase inhibitors based on a 2,5-diaminobenzophenone scaffold. The compounds displayed a wide range of activity, including submicromolar, against the multi-drug resistant *Plasmodium falciparum* strain Dd2. In order to investigate quantitatively the local physicochemical properties involved in the interaction between drug and biotarget, we used the 3D-QSAR methods CoMFA and CoMSIA to study some of the series, including the screened lead compound 2,5-bis-(acylamino)benzophenone, 28 cinnamic acid derivatives, 29 *N*-(3-benzoyl-4-tolylacetylaminophenyl)-3-(5-aryl-2-furyl)acrylic acid amides, and 34 *N*-(4-substituted-amino-3-benzoylphenyl)-[5-(4-nitrophenyl)-2-furyl]acrylic acid amides. We found that steric, electrostatic, and hydrophobic properties of substituent groups play key roles in the bioactivity of the series of compounds, while hydrogen bonding interactions show no obvious impact. We built several highly predictive 3D-QSAR models, including a CoMSIA one composed of steric, electrostatic, and hydrophobic fields, with $r^2 = 0.94$, $q^2 = 0.63$, and $r^2_{\text{pred}} = 0.63$. The results provide insight for optimization of this class of antimalarials for better activity and may prove helpful for further lead optimization.

© 2006 Elsevier Ltd. All rights reserved.

1. Introduction

Malaria is one of the most widespread and deadly infectious diseases in the world, leading to more than 300 million cases and the death of over 1 million people every year. Nearly 50% of the world population, particularly Africans, Latin Americans, and South-east Asians, are at high risk of catching malaria (World Health Organization, World Malaria Report, 2005).^{1,2} Out of the four species of the genus *Plasmodium* which cause human malaria, *Plasmodium falciparum*³ is the most vicious and lethal one. Due to the continuing development of resistance of *P. falciparum* to conventional antimalarial drugs,^{4–6} efforts to find new agents against multi-drug resistant *Plasmodium* strains are long-term and vital tasks for medicinal researchers.^{4,7,8}

Post-translational modification of proteins by prenylation is an important mechanism of cellular regulation. One of the key enzymes for this mechanism is the heterodimeric zinc-containing protein farnesyltransferase.^{9,10} Inhibitors of this enzyme are under development as anti-cancer agents.^{11–13} Several groups have suggested that farnesyltransferase could be an effective target for drug development against malaria and other parasitic diseases such as trypanosomatid infections, and have developed inhibitors with significant in vitro and in some cases in vivo activity against *P. falciparum*.^{5,8,14–32} We focus here on a series of particularly interesting compounds by Wiesner and Schlitzer et al.^{8,17–25} based on a 2,5-diaminobenzophenone scaffold. The lead compound was first developed as a farnesyltransferase inhibitor, and then found to be active in vitro against *P. falciparum*.¹⁸ During the development of the compounds as antimalarials with unreported mode of action^{18–24} they were also reported by Schlitzer and co-workers. in a related series of papers as yeast farnesyltransferase inhibitors (cf., e.g., Ref. 33). Members of this novel class of antimalarials were only later confirmed as killing *P. falciparum* by inhibiting farnesyl-

Keywords: Malaria; Farnesyltransferase inhibitors; CoMFA; CoMSIA.

* Corresponding author. Tel.: +1 662 915 5880; fax: +1 662 915 5638; e-mail: rjd@olemiss.edu

transferase, and then also shown to be active in vivo against *P. falciparum*.²⁵ The compounds displayed a considerable IC₅₀ (50% inhibitory concentration) range against the multi-drug resistant *P. falciparum* strain Dd2, including some submicromolar activity. Now, as pointed out by Wiesner/Schlitzer,^{25–28,32} lead optimization is still needed to obtain compounds with the same or improved in vitro and especially in vivo antimalarial activity coupled with suitable solubility and membrane penetrability, plus low toxicity or side effects (especially since the target patient group includes a high proportion of children and pregnant women).

We decided to investigate further the structure–activity relationship of this series of compounds using quantitative models in order to gain understanding and as a step toward developing a rational strategy for further lead optimization. We report here on several highly significant and predictive 3D-QSAR models (quantitative structure–activity relationships considering explicitly the 3D-conformation of the ligands) based on many of the reported compounds and their activities,^{18–24} using the CoMFA (comparative molecular field analysis)³⁴ and CoMSIA (comparative molecular similarity indices analysis)^{35,36} methods. QSAR techniques have been successfully applied to study the structural features of other classes of antimalarials,^{6,37–40} of other farnesyltransferase inhibitors,^{41–46} and, in particular, of antimalarial farnesyltransferase inhibitors.⁴⁷ CoMFA calculates steric and electrostatic properties in the space surrounding each of the aligned molecules in a data set according to Lennard–Jones and Coulomb potentials, respectively.³⁴ The CoMSIA approach calculates similarity indices around the molecules, with similarity expressed in terms of five different physicochemical properties: steric occupancy, electrostatic field, local hydrophobicity, and hydrogen-bond donor and acceptor properties.³⁵ Böhm's study has shown that CoMSIA possesses better predictive power and greater robustness compared to CoMFA in some cases.³⁶ Each of the methods leads not only to a QSAR but also to a set of plots visualizing regions around the ligands in which chemical modifications are suggested to improve or worsen the activity. An additional strength of CoMSIA is that five different contour maps can be generated and visualized, compared to two for CoMFA.

For 3D-QSAR studies, the assumption is that all the compounds adopt the same binding mode with the same target,⁴⁸ so usually compounds which target a protein whose X-ray structure has been determined would be expected to yield a robust model based on binding site constrained alignment.^{49,50} Although the X-ray crystal structures for rat⁵¹ and human⁵² farnesyltransferase with inhibitors and a substrate analogue have been determined, allowing the identification of the active site residues, the structure for the *P. falciparum* homologue of the enzyme remains unsolved.⁷ However, for compounds whose target has not been verified or for which no X-ray crystal structure is available, an alignment based on a carefully selected template can also yield enlightening models.^{37,39}

2. Data set and computational methods

Molecular modeling and CoMFA and CoMSIA were performed using SYBYL 7.0.⁵³

2.1. Data set

Table 1 shows the structures of the 92 compounds used in this study and their observed activities (pIC₅₀). The reported series of compounds includes the original screened lead compound 2,5-bis-acylaminobenzophenone,¹⁸ 28 cinnamic acid derivatives,^{19,20} 29 *N*-(3-benzoyl-4-tolylacetylaminophenyl)-3-(5-aryl-2-furyl)acrylic acid amides,^{21,24} and 34 *N*-(4-substituted-amino-3-benzoylphenyl)-[5-(4-nitrophenyl)-2-furyl]acrylic acid amides.^{22,23} The compound selection and experimental activities (IC₅₀) against the multi-drug resistant *P. falciparum* strain Dd2 all originated from the same laboratory.^{18–24} Figure 1 shows the distribution of pIC₅₀ for the whole data set, ranging from 4.37 to 7.43 with a variance of 0.47. The compounds fit roughly into three compound clusters. We divided the data set, choosing 74 compounds to be in the QSAR training set and 18 compounds in the test set. The compound number and pIC₅₀ values of the test set are given in bold in Table 1. Care was taken to ensure a uniform distribution of structurally different compounds from each cluster with a wide range in pIC₅₀ for both training and test sets.

2.2. Identification of bioactive conformation

The selection of the putative bioactive conformation proceeded as follows: the three most active compounds (**18**, **56**, and **71**) were extracted separately from the three above-mentioned compound clusters and minimized. For each compound, conformation searching was done using simulated annealing and energy minimizations using the MMFF94 force field. For comparison, the final conformations of the three compounds were rigidly aligned, by fitting all common atoms, in SYBYL, and showed root-mean-squared deviation (rmsd) for the common atoms of 0.0276 Å between **18** and **71**, 0.0053 Å between **56** and **71**, and 0.0301 Å between **18** and **56**. This shows that the three compounds had very similar conformations at the energy minima resulting from simulated annealing. Because **71** is from the largest compound cluster, and its activity is only marginally lower than that of the most active compound **56**, we chose **71** as the alignment template.

2.3. Alignment

Once the putative binding conformation was determined (for **71**), the geometry of each of the other 91 structures was minimized, adjusted manually to approximate the template conformation, and then minimized again to obtain the final geometry used for 3D-QSAR. For each compound, the minimum corresponding to the putative binding conformation was of lower energy than the originally minimized conformer. Finally, the minimized conformations were aligned to **71**, using eight reference atoms in the backbone of the 2,5-diaminobenzophenone

Table 1. Numbering, structure, and experimental activity^a of the data set

Compound	Structure	pIC ₅₀
1^b		5.57
Compound	R	pIC ₅₀
2	–H	5.24
3	–NO ₂	5.19
4	–CHO	4.40
5	–COOCH ₃	6.00
6	–CF ₃	5.24
7	–Cl	5.26
8	–Br	5.49
9	–NH ₂	5.26
10	–CH=C(CN) ₂	4.37
11	–CH ₃	5.85
12	–O–CH ₃	5.89
13	–CH ₂ –CH ₃	5.92
14	–CH(CH ₃) ₂	5.92
15	–C(CH ₃) ₃	5.52
16	–O–CH ₂ –CH ₃	6.07
17	–O–(CH ₂) ₃ –CH ₃	5.96
18	–O–(CH ₂) ₂ –CH ₃	6.47
Compound	R	pIC ₅₀
19		5.05
20		5.47
21		5.60
22		5.89
23		5.62
24		6.46

Table 1 (continued)

Compound	R	pIC ₅₀
25		6.51
26		5.55
27		6.92
28		4.62
29		4.64
Compound	R	pIC ₅₀
30		6.38
31		6.00
32		6.70
33		6.92
34		7.06
35		7.07
36		6.52
37		6.89
38		6.52
39		6.12
40		6.68

(continued on next page)

Table 1 (continued)

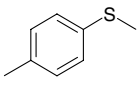
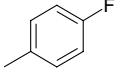
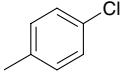
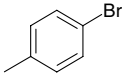
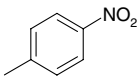
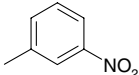
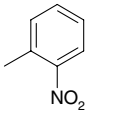
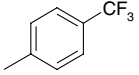
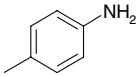
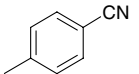
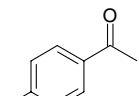
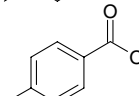
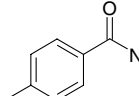
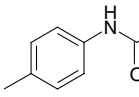
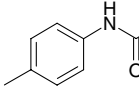
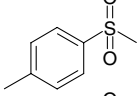
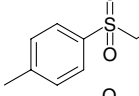
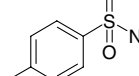
Compound	R	pIC ₅₀
41		7.08
42		6.49
43		6.84
44		6.90
45		7.12
46		6.23
47		6.17
48		7.11
49		6.59
50		6.66
51		7.17
52		6.77
53		6.25
54		6.55
55		6.25
56		7.43
57		7.22
58		6.70

Table 1 (continued)

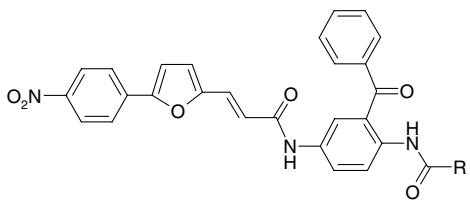


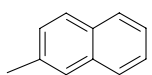
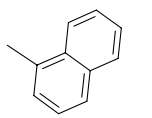

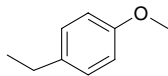






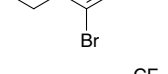

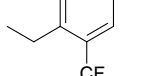
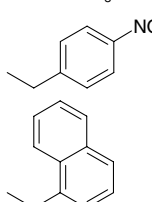
Compound	Structure	pIC ₅₀
		
Compound	R	pIC ₅₀
59		5.52
60		5.60
61		5.60
62		6.11
63		6.57
64		6.49
65		6.82
66		6.19
67		6.64
68		7.19
69		7.15
70		6.00
71		7.33
72		6.00
73		5.85
74		6.60

Table 1 (continued)

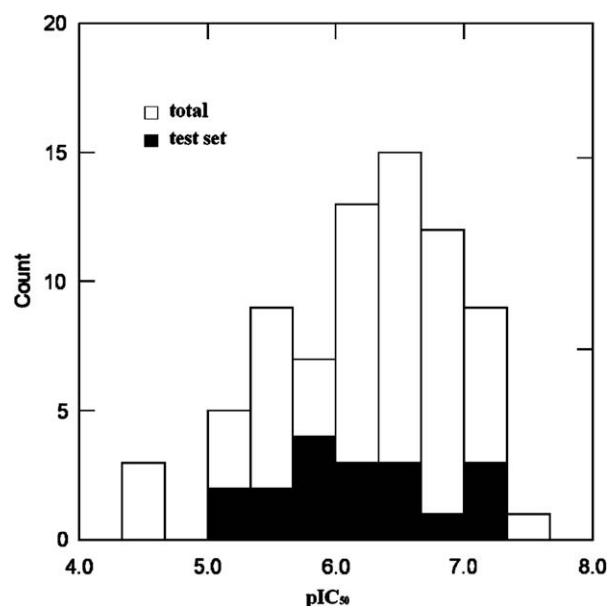
Compound	R	pIC ₅₀
75		6.68
76		6.00
77		5.26
78		6.51
79		5.48
80		5.89
81		6.36
82		7.21
83		6.20
84		5.96
85		6.36
86		6.89
87		6.77
88		5.85
89		6.05
90		5.92
91		6.38

Table 1 (continued)

Compound	R	pIC ₅₀
92		6.24

^a pIC₅₀ against the multi-drug resistant *Plasmodium falciparum* strain Dd2.

^b The lead compound.

Figure 1. Distribution of pIC₅₀ for the 2,5-diaminobenzophenones.

scaffold (Fig. 2). The structural diversity of the aligned compounds is shown in Figure 2.

3. Results and discussion

By thorough study we found that steric, electrostatic and hydrophobic features are key factors that impact the bioactivity of this series of compounds, whereas the hydrogen-bond donor and acceptor fields have no significant correlation with antimalarial activity.

3.1. 3D-QSAR models

Table 2 summarizes the full investigation based on the training set of 74 compounds. Grubbs testing⁵⁴ yielded no outliers in the training set for any of the QSAR models. For the CoMSIA models, statistical parameters showed significance for steric, electrostatic, and hydrophobic features, but hydrogen-bond donor or acceptor fields made insignificant contributions to the models and hence such features are predicted not to be an influence on the activity of the compounds.

Reliable predictions only come from statistically valid QSAR models. There are several statistical parameters, such as r^2 , q^2 , r^2_{pred} , and F , which can be used to evaluate the robustness of a QSAR model. Consider first which models are best based on high q^2 . Model 4, the CoMSIA

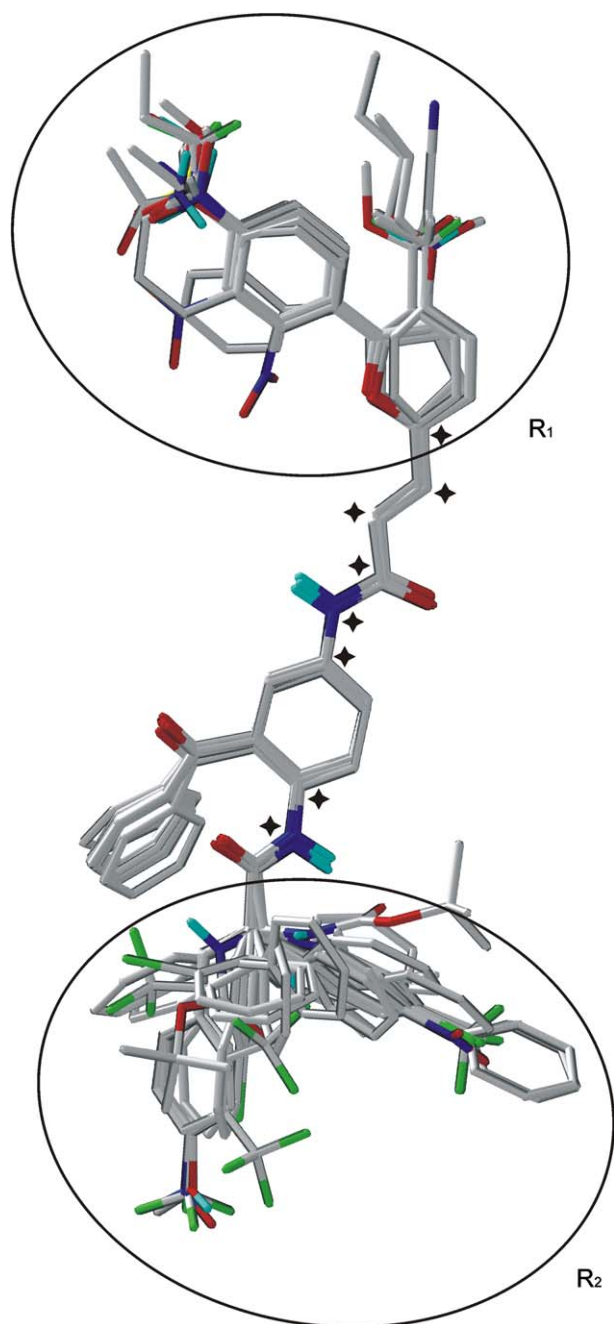


Figure 2. Alignment of the 74 compounds in the training set with template compound **71**. The regions of structural diversity, R_1 and R_2 , are highlighted. The eight reference atoms used for alignment with template compound **71** are starred.

model which contains only a hydrophobic field as descriptor, is the most interpretive single field for the SAR of this series of compounds, with $q^2 = 0.556$ and $r^2 = 0.867$. Thus, it appears that hydrophobicity is the most significant individual factor that impacts the activity. For the models that were derived from both steric and electrostatic features, the CoMFA model (Model 1) is better than the CoMSIA model (Model 7). The best model is CoMSIA Model 8, which includes steric, electrostatic, and hydrophobic fields, with $q^2 = 0.630$ and $r^2 = 0.943$.

Next consider which models are best based on high r^2_{pred} . Tropsha⁵⁵ has emphasized that high q^2 (>0.5) is a necessary but not sufficient condition for a predictive QSAR model. Once models with low r^2 and q^2 are ruled out, the best model should be chosen based on high r^2_{pred} .⁵⁵ For each model, Tables 2 and 3 list r^2_{pred} and the residuals for each compound in the test set. Nearly all the models performed well in the prediction of the activity of the test compounds. In almost all cases, the predicted pIC_{50} values are close to the observed values, with residuals no larger than 1 log unit. From the r^2_{pred} of the CoMSIA models we can easily tell that the main factors that impact the activity of the test set are steric and hydrophobic fields. The r^2_{pred} of Model 2 (only a steric field) and of Model 4 (only a hydrophobic field) are much higher than that of Model 3 (only an electrostatic field). In fact, Model 4 has the highest r^2_{pred} of all the models we considered. For the models containing only steric and electrostatic fields, r^2_{pred} for CoMFA (Model 1) is much higher than that for CoMSIA (Model 7), also suggesting that Model 1 is better than Model 7. Considering the flexibility of all the molecules and the lack of target information, the predictive qualities of the QSAR models are satisfying. Overall, based on r^2 , q^2 , and r^2_{pred} , as well as on the highest value on the Fisher F -test, the best model is Model 8. This CoMSIA model, which includes steric, electrostatic, and hydrophobic fields, is the one we use for graphical analysis below, along with CoMFA Model 1. Either of these could be selected as the final screening tool for lead optimization.

3.2. Graphical interpretation of the results

The contribution maps obtained by CoMFA (Model 1) and CoMSIA (Model 8) show how 3D-QSAR methods can identify features important for the interaction between ligands and the target protein. They allow identification of those positions that require a particular physicochemical property to enhance the bioactivity of a ligand.

In order to understand the CoMFA and CoMSIA maps it is important to consider two factors regarding the alignment, shown in Figure 2. First, the R_1 substituents adopt a different spatial orientation in the cinnamic acid derivatives than in the other two compound clusters. In Figure 2, they project straight vertically, whereas the R_1 substituents of the acrylic acid amides are pointed toward the upper left. A second factor is that R_2 substituents also show two different orientations, in this case based on the length of the carbon chain (linking chain) which connects the substituent group to the scaffold. Compounds with a one-carbon linking chain, as found in **21–28** and **63–67**, have R_2 substituents pointing away from the benzoyl group, while compounds with a two-carbon linking chain, which is more flexible, as found in **78–88**, adopt an orientation with the R_2 group parallel to the benzoyl group. Therefore in the CoMFA and CoMSIA contour maps below we positioned both the template molecule **71** and representative compound **29**, which is a cinnamic acid with a two-carbon linker, in order to depict these two orientation factors.

Table 2. Statistical parameters of the CoMFA and CoMSIA models for the training set of 74 2,5-diaminobenzophenones

	CoMFA				CoMSIA			
	1	2	3	4	5	6	7	8
q^2	0.565	0.474	0.453	0.556	0.086	−0.025	0.496	0.630
S_{PRESS}	0.480	0.540	0.538	0.492	—	—	0.513	0.449
r^2	0.888	0.812	0.732	0.867	—	—	0.786	0.943
SEE	0.244	0.323	0.371	0.270	—	—	0.334	0.177
F	107.6	35.0	63.7	61.29	—	—	63.39	155.2
N	5	8	3	7	—	—	4	7
Field ^a	Contribution							
S	0.567	1.00	—	—	—	—	0.382	0.192
E	0.433	—	1.00	—	—	—	0.618	0.422
H	—	—	—	1.00	—	—	—	0.386
A	—	—	—	—	1.00	—	—	—
D	—	—	—	—	—	1.00	—	—
r^2_{pred}	0.512	0.533	0.138	0.722	—	—	0.348	0.628

^a S , steric; E , electrostatic; H , hydrophobic; D , hydrogen-bond donor; A , hydrogen-bond acceptor.

Table 3. Residuals of test set prediction for CoMFA and CoMSIA models

Compound	pIC_{50}^a	Residuals from different models					
		1	2	3	4	7	8
2	5.24	0.33	0.02	0.25	0.04	0.21	0.15
9	5.26	0.41	0.09	0.64	0.24	0.61	0.36
20	5.47	0.16	0.07	0.19	0.18	0.12	0.05
22	5.89	0.04	0.05	0.10	0.41	0.04	0.57
30	6.38	0.22	0.24	0.44	0.41	0.19	0.42
36	6.52	0.53	0.80	0.60	0.44	0.47	0.59
41	7.08	−0.40	−0.34	0.73	−0.28	0.68	−0.37
46	6.23	0.26	0.15	−0.14	0.33	0.13	0.20
55	6.25	0.28	0.37	0.42	0.39	0.39	0.39
60	5.60	0.01	0.08	0.42	0.17	0.32	0.24
64	6.49	0.28	−0.10	0.19	−0.23	0.04	−0.45
68	7.19	−0.43	−0.57	0.83	−0.22	0.64	−0.26
70	6.00	0.00	0.00	0.35	−0.07	0.38	−0.17
80	5.89	0.50	0.44	0.15	−0.03	0.28	0.13
82	7.21	−0.87	−1.03	−1.20	−0.69	1.12	−0.57
86	6.89	−0.55	−0.59	0.77	−0.34	0.74	−0.29
88	5.85	0.50	0.39	0.35	0.05	0.26	0.36
90	5.92	0.68	0.08	0.82	0.36	0.62	0.47
Variance ^b		0.18	0.17	0.32	0.10	0.24	0.14

^a Experimental pIC_{50} against the multi-drug resistant *P. falciparum* strain Dd2.

^b RSS/n , for n compounds.

3.2.1. Steric field. Figure 3 shows the steric field contours from CoMFA (a) and CoMSIA (b). Green represents where more bulky groups are favorable for improving bioactivity, while yellow means in those positions bulky groups are unfavorable for improving activity. The two contour maps are basically consistent with each other except for differences in polyhedron volume. The steric contours obtained through CoMFA Model 1 give more localized and detailed polyhedra than those from CoMSIA Model 7, as is obvious from Figure 3.

R_1 substitutions. The CoMSIA map shows that bulky substituents anywhere on the terminal phenyl ring will improve activity, while the CoMFA map shows that only in the *para* and *meta* positions will bulky groups be beneficial, while the *ortho* position does not tolerate bulky substituents. Also, the *para* position needs bulky, but not too bulky or long chain, groups, as indicated by a small yellow CoMFA contour. This interpretation

matches with the structural features exhibited by **30–58**. By contrast, comparing **30** with the less active **46** and **47** indicates that nitro group substitution at the *ortho* or *meta* positions, respectively, was not favorable for activity improvement. Though it is true that the nitro group is bulkier than H, this behavior is best attributed to the electrostatic and hydrophilic properties of the nitro group (see below). The small yellow block in the upper right side of the maps indicates that cinnamic acid derivatives are less active relative to the clusters with a furyl ring, since the bulk associated with their six-membered ring is indicated as detrimental to activity.

R_2 substitutions. Among the compounds containing a two-carbon linker, substituents at the *ortho* position of the terminal phenyl ring will lead to an activity decrement (e.g., **84** with *ortho* CF_3 was less active than the corresponding unsubstituted derivative, **78**). For the one-carbon linker compounds, **66**, **70**, and **72** with *ortho*

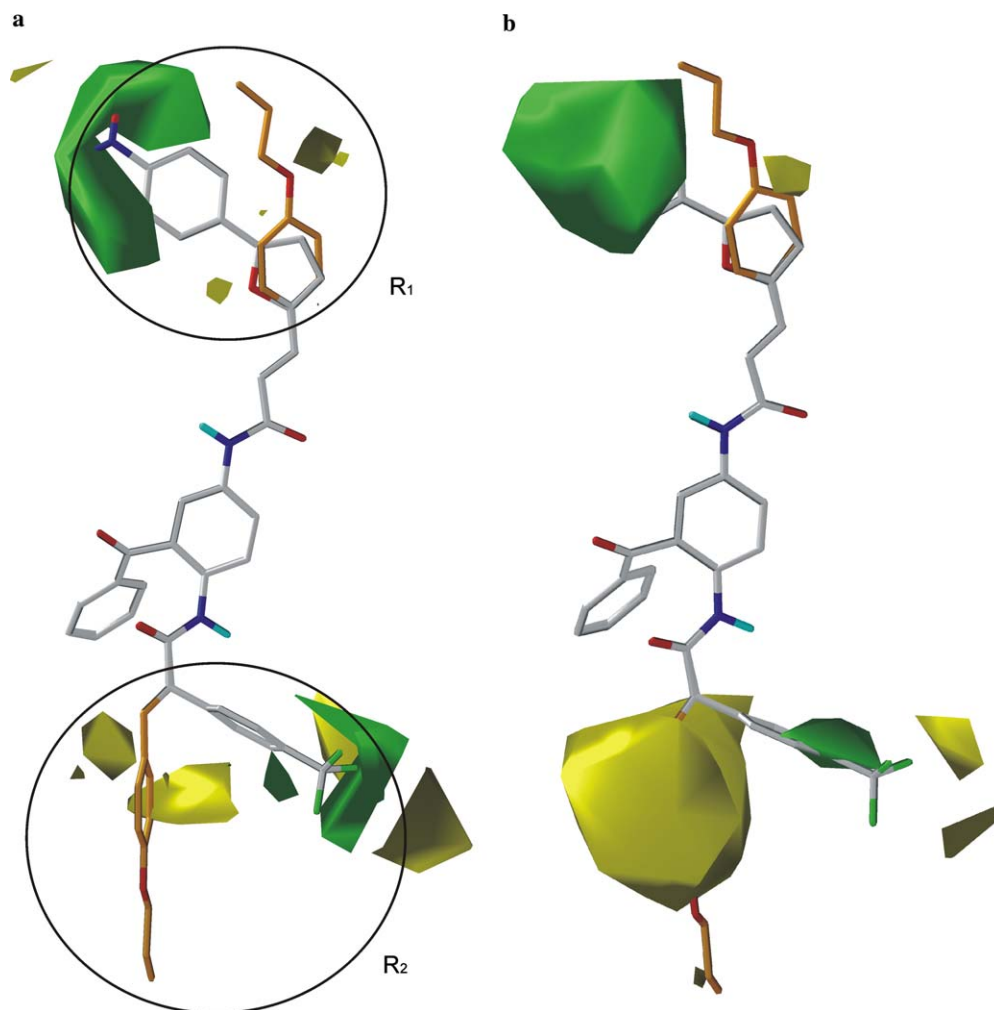


Figure 3. Steric field contour map (a, CoMFA; b, CoMSIA) around the template compound, **71**, as well as **29**, which are overlapped and both shown in stick representation, with carbons colored orange in the end groups of **29**. Green (yellow) regions show where bulky groups should increase (decrease) the antimalarial activity.

methyl, bromo, and trifluoromethyl groups, respectively, had reduced activity. Both these effects correspond to the small yellow block near the middle of the R_2 circle in the CoMFA map (Fig. 3a) and are included in the large yellow polyhedron in the CoMSIA map (Fig. 3b). There is a small green block (in both the CoMFA and CoMSIA maps) indicating enhanced activity for bulky substituents at the *meta* position; **65**, with a *meta* methyl substituent, was an example of this. For the *para* position, also, bulky groups improved the activity, though the yellow blocks which are further away indicate that long chains will not be beneficial. To illustrate this, compounds with short bulky groups like **68**, **69**, and **71** showed improved activity, while ones with a longer chain substituent (such as **76**) showed decreased activity. The big yellow block (CoMSIA) or two smaller yellow blocks (CoMFA) at the lower left of the Figure 3 images indicate that the position occupied by big R_2 groups in structures such as **61**, **62**, and **77** (which have no linker carbon chain) is sterically disfavored. Compounds **64** and **73** do not match our conclusions regarding steric effects of the data set, perhaps because their relative activities were determined more by other physicochemical properties (see below).

3.2.2. Electrostatic field. The electrostatic contour maps of CoMFA (a) and CoMSIA (b) are shown in Figure 4. Blue means an electronegative group will favor the activity, while red represents regions where electronegative groups will reduce the activity of the molecule.

R_1 substitutions. An electronegative group at the *para* position of the terminal phenyl of the arylfurylacryloyl group will increase the activity. For example, **40**, **42–45**, **48**, **50–52**, and **56–58** all showed improved activities compared to the unsubstituted **30**. In the CoMFA map, there is a blue block close to the *para* position of the phenyl ring, and a red block just a little further away. This means *para* groups should have electron withdrawing atoms attached close to the phenyl, and electron donating groups further away from those atoms in order to improve activity. This interpretation can be verified by comparing the increasing trend of activity of **30**, **40**, **37**, **41**, and **51**. Wiesner et al.²⁴ investigated the role of hydrogen-bond acceptor properties at the *para* position of **45** (which itself has a nitro group there), and found that some of them (**48**: trifluoromethyl; **51**: acetyl) were equipotent and some (**56**: methylsulfonyl) had improved antimalarial activity. Our study

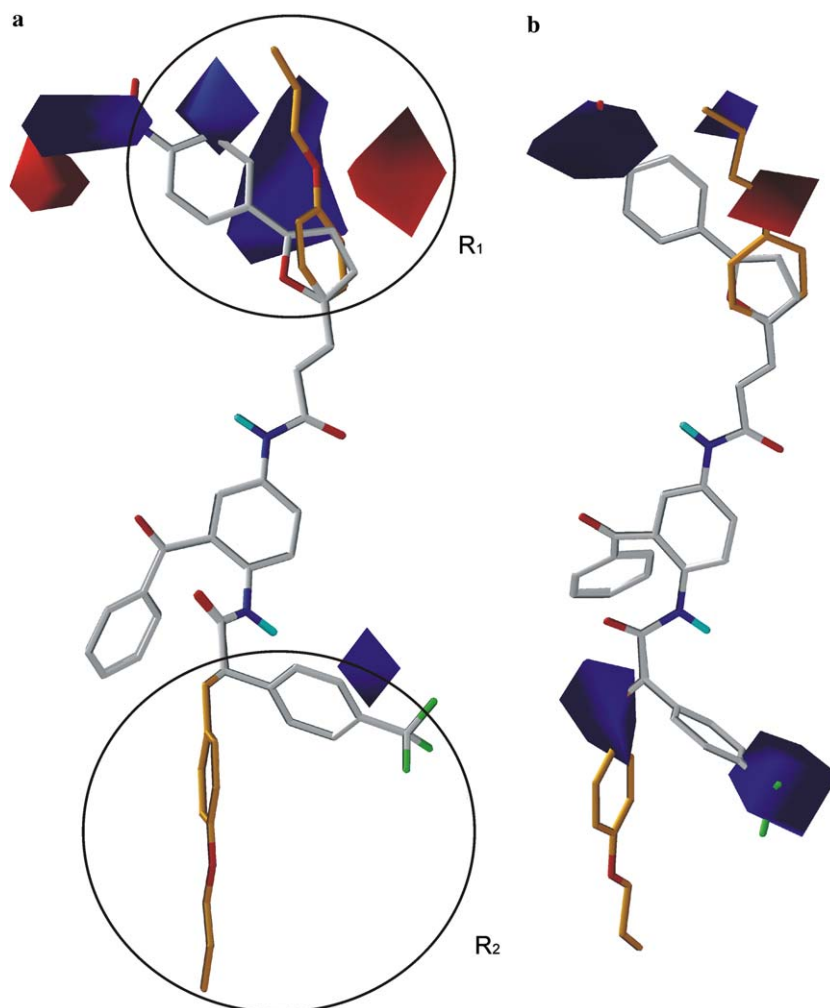


Figure 4. Electrostatic field contour map (a, CoMFA; b, CoMSIA) around the template compound, **71**, as well as **29**, which are overlapped and both shown in stick representation, with carbons colored orange in the end groups of **29**. Blue (red) regions show where electronegative groups should increase (decrease) the antimalarial activity.

supplements this observation by indicating that this should be mainly due to the electrostatic property of the hydrogen-bond acceptor groups at the *para* position. For cinnamic acid derivatives, at the *para* position of the terminal ring, the substituent needs an appropriate combination of electron withdrawing and donating atoms, as shown clearly in the CoMSIA figure with a red block close by and a blue block further away. This is illustrated by **16–18**.

***R*₂ substitutions.** An electron withdrawing group at the phenyl's *para* position (**24**, **25**, **27**, **67**, **68**, **69**, and **71**) increased the activity. However, **64**, with an electron donating methoxy group at the *para* position, decreased the activity in spite of its favorable steric profile. **73** was an outlier for both steric and electrostatic rules. It will be discussed in the hydrophobic rule, below. The blue block at the lower left corner of the CoMSIA map shows that an electron donating group at this position (such as in **89**, **91**, and **92**) will decrease the activity.

3.2.3. Hydrophobic field. Figure 5 shows the hydrophobic contour from CoMSIA Model 8. Magenta means

that a hydrophobic group will increase antimalarial activity, while cyan means that a hydrophobic group will decrease the activity.

***R*₁ substitutions.** Magenta patches in Figure 5 indicate that a hydrophobic group at the *ortho* or *meta* position of the terminal phenyl of the arylfurylacryloyl group will increase the activity. For example, **46** and **47** showed that a polar nitro substituent at the *meta* or *ortho* position, respectively, worsened the activity. The *para* position needs a hydrophobic group with a hydrophilic element close to the phenyl ring (such as in **56** and **57**) for increased activity. The big cyan block covering the furyl ring (of template **71**) or phenyl ring (of **29**) presumably comes from the interpretation of cinnamic acid derivatives (such as **31**, **32**, **46**, and **47**) as follows: since most of their *R*₁ groups are hydrophobic, and most of those compounds were relatively inactive, thus the contour map shows that here hydrophilic groups should increase activity.

***R*₂ substitutions.** The *para* position of the phenyl ring needs a small hydrophobic group (**71**: CF₃; **69**: Br; **45**: CH₃) to display good activity (pIC₅₀ = 7.33, 7.15, and

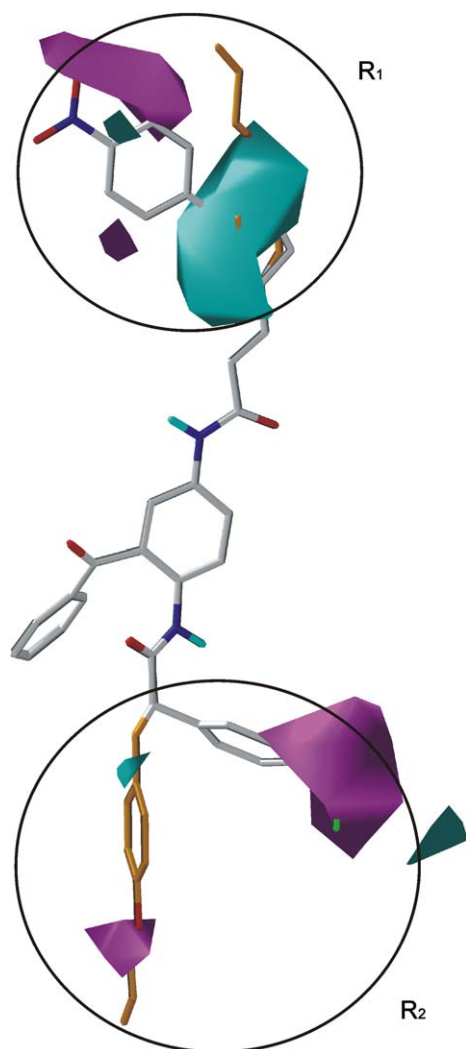


Figure 5. CoMSIA hydrophobic field contour map around the template compound, **71**, as well as **29**, which are overlapped and both shown in stick representation, with carbons colored orange in the end groups of **29**. Magenta (cyan) regions show where hydrophobic groups should increase (decrease) the antimalarial activity.

7.12, respectively). In fact **71**, which contains CF_3 , the most hydrophobic substituent among those three, was one of the most active compounds. The activity of **73**, with a polar nitro group on the *para* position, was relatively low. On the other hand, the cyan blocks at the lower right corner of the template structure indicate that too large a hydrophobic group will decrease the antimalarial activity. For example, **76** and **28** which each contain a *para* phenyl ring were found to have relatively low activity. For compounds containing a two-carbon linker (**29**, **78–92**), the small cyan patch at the lower left indicates the activity trend of **77** and **89–92**, which are all branched sidechains, showing that branching at this position lowered the activity. The magenta block at bottom showing the phenyl *para* position of those ligands also welcomes hydrophobic groups. For example, compounds with hydrophobic *para* groups (**81**, **82**, **85**, **86**, and **87**) had better or nearly equipotent activity compared to unsubstituted **78**, while **79**, **80**, and **88** had *para* hydrophilic groups and lower activity.

3.3. Discussion of related work

Gupta et al.⁴⁷ published a 3D-QSAR model for 16 of the Wiesner et al.^{18,19} diaminophenones, using thermodynamic, electronic, and spatial descriptors. Their best model had $r^2 = 0.69$; $q^2 = 0.43$, $F = 9.47$, $r^2_{\text{pred}} = 0.61$. Note that their best model is for many fewer ligands and is rather less reliable than our work, according to the statistical measures r^2 , q^2 , and F . Their r^2_{pred} is relatively good for their small, 4-compound test set.

González et al.⁴⁵ used multiple linear regression analysis and nonlinear neural networks for 3D-QSAR using farnesyltransferase (FT) inhibition data, including for 22 of the molecules in our work. They used radial distribution function descriptors for their analysis. They attempted to fit models to two classes of inhibitors, thiol and nonthiol, together or for the individual classes. Note that it is thought that the two classes have a different binding mode. While it is advantageous to obtain a very accurate model by such an approach which could then be used for virtual screening, the resulting model is not easy to interpret in terms of which chemical groups and interactions are important for ligand binding, particularly since it is applied to diverse inhibitors which most likely bind in different modes. Any comparison between our results and theirs should be qualified by the consideration that their models were fit to FT inhibition data, whereas ours were fit to antimalarial activity. The ability to interpret QSAR data in terms of individual, localized interactions and the role of chemical groups is a strength of the CoMFA and CoMSIA methods used in our work. Their interpretive conclusions instead focus on overall qualities of FT inhibition. One of their conclusions is similar to ours, that steric and hydrophobic interactions are crucial for FT inhibition. Contrary to their other main conclusion, that electronegativity or polarizability effects are unimportant for FT inhibition, we have discussed above that there are some specific contributions of electrostatics to improved activity at least for the nonthiol inhibitors in our work, based on the sort of detailed analysis possible with CoMFA and CoMSIA, which goes beyond the whole-molecule topological descriptors they used.⁴⁵

Wiesner/Schlichter et al. have gained considerable knowledge about the active site through homology modeling of *P. falciparum* farnesyltransferase based on the X-ray structures of farnesyltransferase of rat and of human, plus docking of farnesyltransferase inhibitors into that model's active site.²⁶ The putative binding conformation adopted in our study is quite similar to their docked conformation.²⁶ Similarly, Glenn et al.³⁰ have also reported a homology model developed based on rat farnesyltransferase, containing a deep hydrophobic pocket for the active site. Though our 3D-QSAR contour plots cannot be compared directly to the active site of the protein as depicted in the reported homology models,^{26,30} our observations about the hydrophobic and steric profiles of the 2,5-diaminobenzophenone farnesyltransferase inhibitors confirm the spacious hydrophobic volume of the active site of those homology models.

4. Conclusion

For the 2,5-diaminobenzophenone scaffold based series of novel antimalarials, though there is a high degree of flexibility and no available X-ray crystal structure, we have been able to obtain 3D-QSAR models with good statistical significance and good predictive ability using CoMFA and CoMSIA, based on a putative bioactive conformation obtained from simulated annealing searching. Molecular surface property (steric, electrostatic, and hydrophobic) mapping of these models highlighted the features that affected bioactivity and provided explicit indications for lead compound optimization. In this study, the CoMFA model with steric and electrostatic fields (Model 1) performed better than the corresponding CoMSIA model (Model 7). However, the best CoMSIA model, Model 8 composed of steric, electrostatic, and hydrophobic fields, was significantly more accurate than other models, with $r^2 = 0.94$, $q^2 = 0.63$, and $r^2_{\text{pred}} = 0.63$. In this study, the hydrophobic property is the most significant factor that correlated with activity. The strong impact of hydrophobicity on antimalarial activity—as exemplified by Model 4, containing only the hydrophobic field, with $r^2 = 0.87$, $q^2 = 0.56$, and $r^2_{\text{pred}} = 0.72$, the highest r^2_{pred} of any single-field CoMSIA model—gives a clue for further investigation of antimalarials. We also have found that the activity of other series of antimalarials is significantly correlated to their relative hydrophobicity. We will continue studying this property of antimalarials to see if it can be developed into a criterion of antimalarial drug candidate screening.

5. Computational details

5.1. Conformational search

Conformation searching was done using simulated annealing⁵³ for **18**, **56**, and **71**. Each annealing cycle was commenced from an initially assigned temperature of 1000 K (with initial velocities taken from a Boltzmann distribution), which was maintained for 1 ps to allow for exploration of conformational space, and then followed by exponential quenching in 2 ps down to 200 K. The timestep used was 1 fs. Five hundred such cycles generated 500 low-temperature conformations which were then energy ranked. The 50 lowest energy conformers were further minimized, at which point the lowest energy conformation was selected as the putative binding conformation. For minimization, the Powell conjugate-gradient algorithm was used with a gradient convergence criterion of $0.01 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$. Simulated annealing and energy minimization both used the MMFF94 force field, with partial atomic charges also taken from MMFF94 and distance-dependent dielectric constant = 1 to represent vacuum.

5.2. 3D-QSAR methodology

For CoMFA and CoMSIA, a 3D cubic lattice needs to be created to encompass the aligned molecules.^{34,35}

According to experience and the literature, a 2.0 Å grid spacing is most appropriate for partial least squares (PLS) analysis in order to extract sufficient field information while avoiding excessive noise. Other CoMFA/CoMSIA details are given as [Supplementary data](#).

The CoMFA or CoMSIA descriptors were used as independent variables, and pIC_{50} values were the dependent variables in PLS analysis to derive QSAR models using the standard implementation in the SYBYL package.⁵³ To evaluate the quality of the QSAR models, we used three standard squared correlation coefficients, r^2 , q^2 , and r^2_{pred} . Following Leach and Gillet's description,⁵⁶ each coefficient can be calculated using the form $1 - (RSS/TSS)$, where TSS (total sum of squares) is the sum of squared deviations of the experimental activities from their mean and RSS (residual sum of squares) is the sum of squared deviations of the calculated activities from their experimental counterparts. For r^2 and q^2 , the sum is over all the compounds in the training set, whereas for r^2_{pred} , the sum is rather over all of the test set compounds. For r^2 , the calculated values are those determined by the QSAR model as originally generated, whereas for q^2 , the calculated values are those predicted through a leave-one-out procedure.⁵⁷ Here the SAMPLS (sample-distance PLS) algorithm⁵³ was used to determine the optimal number of components (N , given in [Table 2](#)) in reference to the predictive power, which is measured by how much q^2 increases as N is increased. Generally, robust statistical models should have high r^2 , q^2 , and r^2_{pred} . The optimal number of components was then used to derive the final QSAR model. The predictive power was tested by calculating r^2_{pred} for all models with q^2 more than 0.5. The variance was calculated as RSS/n , for n compounds. For each model we used the Grubbs test⁵⁴ to check for outliers.

For each CoMFA or CoMSIA model, figures depicting coefficient contour maps for particular fields—in which the coefficients are scaled by the standard deviation in order to be on a % scale—give insight into regions around the aligned ligands in which functional group modifications will lead to improved or worsened activity.³⁴ Here we illustrate such properties by depicting regions having scaled coefficients either greater than 80% (favored) or less than 20% (disfavored).

Acknowledgments

Funding from CDC Grant U01/CI000211-02, from NSF EPS-0556308, and from University of Mississippi, including from its Faculty Research Program, as well as computing facilities of Laboratory for Applied Drug Design and Synthesis and MCSR, are greatly appreciated. This investigation was conducted in a facility constructed with support from research facilities improvement program C06 RR-14503-01 from the NIH National Center for Research Resources. Thanks to Pankaj R. Daga for technical assistance and helpful suggestions.

Supplementary data

Supplementary data associated with this article including calculation details and predicted activities can be found, in the online version, at [doi:10.1016/j.bmc.2006.06.041](https://doi.org/10.1016/j.bmc.2006.06.041).

References and notes

1. Snow, R. W.; Guerra, C. A.; Noor, A. M.; Myint, H. Y.; Hay, S. I. *Nature* **2005**, *434*, 214.
2. Bell, D. R.; Jorgensen, P.; Christophel, E. M.; Palmer, K. L. *Nature* **2005**, *437*, E3.
3. LaCount, D. J.; Vignali, M.; Chettier, R.; Phansalkar, A.; Bell, R.; Hesselberth, J. R.; Schoenfeld, L. W.; Ota, I.; Sahasrabudhe, S.; Kurschner, C.; Fields, S.; Hughes, R. E. *Nature* **2005**, *438*, 103.
4. Baird, J. K. N. *Engl. J. Med.* **2005**, 352, 1565.
5. Eastman, R. T.; White, J.; Hucke, O.; Bauer, K.; Yokoyama, K.; Nallan, L.; Chakrabarti, D.; Verlinde, C. L.; Gelb, M. H.; Rathod, P. K.; Van Voorhis, W. C. *J. Biol. Chem.* **2005**, *280*, 13554.
6. Fidock, D. A.; Rosenthal, P. J.; Croft, S. L.; Brun, R.; Nwaka, S. *Nat. Rev. Drug Discovery* **2004**, *3*, 509.
7. Brady, R. L.; Cameron, A. *Curr. Drug Targets* **2004**, *5*, 137.
8. Wiesner, J.; Ortmann, R.; Jomaa, H.; Schlitzer, M. *Angew. Chem., Int. Ed.* **2003**, *42*, 5274.
9. Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *J. Biol. Inorg. Chem.* **2005**, *10*, 3.
10. Casey, P. J.; Solski, P. A.; Der, C. J.; Buss, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 8323.
11. Appels, M.; Beijnen, J. H.; Schellens, J. H. *Oncologist* **2005**, *10*, 565.
12. Pan, J.; Yeung, S. C. *Cancer Res.* **2005**, *65*, 9109.
13. Gelb, M. H.; Van Voorhis, W. C.; Buckner, F. S.; Yokoyama, K.; Eastman, R.; Carpenter, E. P.; Panethymitaki, C.; Brown, K. A.; Smith, D. F. *Mol. Biochem. Parasitol.* **2003**, *126*, 155.
14. Buckner, F. S.; Eastman, R. T.; Yokoyama, K.; Gelb, M. H.; Van Voorhis, W. C. *Curr. Opin. Investig. Drugs* **2005**, *6*, 791.
15. Ohkanda, J.; Lockman, J. W.; Yokoyama, K.; Gelb, M. H.; Croft, S. L.; Kendrick, H.; Harrell, M. I.; Feagin, J. E.; Blaskovich, M. A.; Sebt, S. M.; Hamilton, A. D. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 761.
16. Chakrabarti, D.; Azam, T.; DelVecchio, C.; Qiu, L.; Park, Y. L.; Allen, C. M. *Mol. Biochem. Parasitol.* **1998**, *94*, 175.
17. Böhm, M.; Mitsch, A.; Wissner, P.; Sattler, I.; Schlitzer, M. *J. Med. Chem.* **2001**, *44*, 3117.
18. Wiesner, J.; Wißner, P.; Dahse, H.-M.; Jomaa, H.; Schlitzer, M. *Bioorg. Med. Chem.* **2001**, *9*, 785.
19. Wiesner, J.; Mitsch, A.; Wißner, P.; Jomaa, H.; Schlitzer, M. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 423.
20. Wiesner, J.; Kettler, K.; Jomaa, H.; Schlitzer, M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 543.
21. Wiesner, J.; Mitsch, A.; Wißner, P.; Kramer, O.; Jomaa, H.; Schlitzer, M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2681.
22. Wiesner, J.; Kettler, K.; Sakowski, J.; Ortmann, R.; Jomaa, H.; Schlitzer, M. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 361.
23. Wiesner, J.; Fucik, K.; Kettler, K.; Sakowski, J.; Ortmann, R.; Jomaa, H.; Schlitzer, M. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1539.
24. Wiesner, J.; Mitsch, A.; Jomaa, H.; Schlitzer, M. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2159.
25. Wiesner, J.; Kettler, K.; Sakowski, J.; Ortmann, R.; Katzin, A. M.; Kimura, E. A.; Silber, K.; Klebe, G.; Jomaa, H.; Schlitzer, M. *Angew. Chem., Int. Ed.* **2004**, *43*, 251.
26. Kettler, K.; Wiesner, J.; Silber, K.; Haebel, P.; Ortmann, R.; Sattler, I.; Dahse, H.-M.; Jomaa, H.; Klebe, G.; Schlitzer, M. *Eur. J. Med. Chem.* **2005**, *40*, 93.
27. Kettler, K.; Wiesner, J.; Fucik, K.; Sakowski, J.; Ortmann, R.; Dahse, H.-M.; Jomaa, H.; Schlitzer, M. *Pharmazie* **2005**, *60*, 677.
28. Schlitzer, M. *Curr. Med. Chem. Anti-Infect. Agents* **2005**, *4*, 277.
29. Nallan, L.; Bauer, K. D.; Bendale, P.; Rivas, K.; Yokoyama, K.; Horney, C. P.; Rao Pendyala, P.; Floyd, D.; Lombardo, L. J.; Williams, D. K.; Hamilton, A. D.; Sebt, S.; Windsor, W. T.; Weber, P. C.; Buckner, F. S.; Chakrabarti, D.; Gelb, M. H.; Van Voorhis, W. C. *J. Med. Chem.* **2005**, *48*, 3704.
30. Glenn, M. P.; Chang, S. Y.; Hucke, O.; Verlinde, C. L.; Rivas, K.; Horney, C.; Yokoyama, K.; Buckner, F. S.; Pendyala, P. R.; Chakrabarti, D.; Gelb, M.; Van Voorhis, W. C.; Sebt, S. M.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **2005**, *44*, 4903.
31. Ryckebusch, A.; Gilleron, P.; Millet, R.; Houssin, R.; Lemoine, A.; Pommery, N.; Grellier, P.; Sergheraert, C.; Henichart, J.-P. *Chem. Pharm. Bull.* **2005**, *53*, 1324.
32. Kettler, K.; Wiesner, J.; Ortmann, R.; Dahse, H.-M.; Jomaa, H.; Schlitzer, M. *Pharmazie* **2006**, *61*, 63.
33. Kettler, K.; Sakowski, J.; Silber, K.; Sattler, I.; Klebe, G.; Schlitzer, M. *Bioorg. Med. Chem.* **2003**, *11*, 1521.
34. Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
35. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130.
36. Böhm, M.; Stürzebecher, J.; Klebe, G. *J. Med. Chem.* **1999**, *42*, 458.
37. Avery, M. A.; Muraleedharan, K. M.; Desai, P. V.; Bandyopadhyaya, A. K.; Furtado, M. M.; Tekwani, B. L. *J. Med. Chem.* **2003**, *46*, 4244.
38. Parenti, M. D.; Pacchioni, S.; Ferrari, A. M.; Rastelli, G. *J. Med. Chem.* **2004**, *47*, 4258.
39. Xue, C. X.; Cui, S. Y.; Liu, M. C.; Hu, Z. D.; Fan, B. T. *Eur. J. Med. Chem.* **2004**, *39*, 745.
40. Marrero-Ponce, Y.; Iyarreta-Veitia, M.; Montero-Torres, A.; Romero-Zaldivar, C.; Brandt, C. A.; Avila, P. E.; Kirchgatter, K.; Machado, Y. *J. Chem. Inf. Model.* **2005**, *45*, 1082.
41. Sung, N.-D.; Cho, Y.-K.; Kwon, B.-M.; Hyun, K. H.; Kim, C. K. *Arch. Pharmacol. Res.* **2004**, *27*, 1001.
42. Wan, S.; Yi, X.; Guo, Z. *Yaoxue Xuebao* **2001**, *36*, 423.
43. Horvath, D. In *QSPR/QSAR Studies by Molecular Descriptors*; Diudea, M. V., Ed.; Nova Science Publishers: Huntington, NY, 2001; p 389.
44. Polley, M. J.; Winkler, D. A.; Burden, F. R. *J. Med. Chem.* **2004**, *47*, 6230.
45. González, M. P.; Caballero, J.; Tundidor-Camba, A.; Helguera, A. M.; Fernández, M. *Bioorg. Med. Chem.* **2006**, *14*, 200.
46. Puntambekar, D.; Giridhar, R.; Yadav, M. R. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1821.
47. Gupta, A. K.; Soni, L. K.; Hanumantharao, P.; Sambasivarao, S. V.; Arockia Babu, M.; Kaskhedikar, S. G. *Asian J. Chem.* **2004**, *16*, 67.
48. Kubinyi, H. In *3D QSAR in Drug Design*; Kubinyi, H., Folkers, G., Martin, Y. C., Eds.; Kluwer Academic Publisher: Dordrecht, 1998; Vol. 2, p 1.
49. Nair, A. C.; Jayatilke, P.; Wang, X.; Miertus, S.; Welsh, W. J. *J. Med. Chem.* **2002**, *45*, 973.
50. Liao, C.; Xie, A.; Zhou, J.; Shi, L.; Li, Z.; Lu, X.-P. *J. Mol. Model.* **2004**, *10*, 165.

51. Park, H. W.; Boduluri, S. R.; Moomaw, J. F.; Casey, P. J.; Beese, L. S. *Science* **1997**, 275, 1800.
52. Long, S. B.; Hancock, P. J.; Kral, A. M.; Hellings, H. W.; Beese, L. S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, 98, 12948.
53. SYBYL 7.0. Tripos, Inc., St. Louis, MO, USA, **2005**.
54. Grubbs, F. *Technometrics* **1969**, 11, 1.
55. Tropsha, A. In *Chemoinformatics in Drug Discovery*; Oprea, T. I., Ed.; Wiley: Weinheim, 2005; p 437.
56. Leach, A. R.; Gillet, V. J. In *An Introduction to Chemoinformatics*; Kluwer Academic Publishers.: Dordrecht, 2003.
57. Hawkins, D. M.; Basak, S. C.; Mills, D. J. *Chem. Inf. Comput. Sci.* **2003**, 43, 579.