



Pergamon

3D-QSAR of *N*-Myristoyltransferase Inhibiting Antifungal Agents by CoMFA and CoMSIA Methods

P. Purushottamachar and Vithal M. Kulkarni*

Pharmaceutical Division, Institute of Chemical Technology, University of Mumbai, Matunga, Mumbai 400019, India

Received 4 December 2002; accepted 2 May 2003

Abstract—A series of benzofuran antifungals was examined to determine the structural requirements of *N*-myristoyltransferase (Nmt) enzyme inhibition by three-dimensional quantitative structure–activity relationship (3D-QSAR) using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) methods. Evaluation of 20 compounds (training set) served to establish the model, which was validated by evaluation of a set of 6 compounds (test set). The lowest energy conformer of the most active molecule obtained from systematic search was used as the template structure for the alignment. The best predictions were obtained with the CoMFA model from RMS fit, with $r_{cv}^2 = 0.828$, $r_{conv}^2 = 0.989$, $r_{pred}^2 = 0.754$ and with the CoMSIA model combining hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields with $r_{cv}^2 = 0.821$, $r_{conv}^2 = 0.978$ and $r_{pred}^2 = 0.747$. The models obtained from the present study can be useful for the development of new Nmt inhibitors as potential antifungals. The docking studies were also carried out wherein the active and inactive molecules were docked into the active site of the recently reported *Candida albicans* Nmt (CaNmt) crystal structure to analyze enzyme–inhibitor interactions. The results obtained from the present 3D-QSAR and docking studies were found complimentary.

© 2003 Elsevier Ltd. All rights reserved.

Introduction

Myristoyl CoA: protein *N*-myristoyltransferase (Nmt; EC 2.1.3.97) is a cytosolic monomeric enzyme which catalyzes the transfer of cellular fatty acid myristate (C14:0) from myristoyl-CoA to the N-terminal glycine amine of a variety of eukaryotic proteins.¹ These protein substrates include kinases, phosphatases, α -subunits of many heterotrimeric G-proteins and endothelial cell nitric oxide synthase. Structural and nonstructural proteins encoded by many viruses, including HIV-1, are also *N*-myristoyled.^{2,3} Different *N*-myristoyl proteins use myristate for various purposes, such as promotion of protein–protein and protein–lipid interactions.^{3,4} Genetic studies have established that Nmt is essential for the growth and survival of *Saccharomyces cerevisiae*. Nmt has also been proven to be essential for the viability of fungi, including medically important pathogenic fungi such as *Candida albicans* and *Cryptococcus neoformans*.^{5–7} Hence, Nmt is an attractive target for the design and development of novel antifungal agents

with a novel mode of action. It is also one of the best studied enzymes, and the wealth of acquired knowledge is useful for selectively targeting this enzyme to design inhibitors without disrupting the functions of host Nmt.^{8,9} Because of its novel mechanism of action, an Nmt inhibitor is expected to have advantage over azole antifungal agents in terms of activity against azole-resistant fungal strains and lack of the drug–drug interactions that are drawbacks of the azole antifungal agents.¹⁰

Although several peptidomimetic inhibitors of *C. albicans* Nmt (CaNmt) have been reported, their antifungal activity being only marginal.^{11–13} We have found novel nonpeptidic sulfonamide analogues¹⁴ as Nmt inhibitors and recently Tatsuo Ohtsuka et al. reported a series of novel benzofuran analogues as nonpeptidic Nmt inhibitors.^{15,16} During the initial SAR, it was found that substitution at C-2 and C-4 positions of benzofuran influence the observed antifungal activity. In view how modification at C-2 and C-4 with different functional groups will affect the antifungal activity the authors synthesized several compounds with wide spectrum of activities.

In our laboratory, we have been interested in finding novel antifungal agents against *C. albicans*. Various

*Corresponding author. Tel.: +91-22-2414-5616; fax: +91-22-2414-5614; e-mail: vithal@biogate.com

targets like the Cytochrome P450_{14 α} DM,^{17–19} dihydrofolate reductase²⁰ and squalene epoxidase²¹ have been used for rational design of antifungal agents. Considering the potential of Nmt as a new target against fungal infections, particularly *C. albicans*, we have generated pharmacophore model to derive the minimal structural requirements for Nmt inhibition.²² In order to gain further insight into the structure–activity relationship of the benzofuran antifungals,^{15,16} a three-dimensional quantitative structure–activity relationship (3D-QSAR) by both comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) methods were employed. Furthermore, we report docking studies wherein the most active and one of the inactive benzofuran analogues have been docked into the active site of *Ca*Nmt.

CoMFA is one of the methods of rational drug design, which has been successfully applied in our laboratory for antihyperglycemic agents,^{23,24} interleukin 1- β converting enzyme,²⁵ HIV-1 integrase inhibitors²⁶ and squalene epoxidase inhibitors.²¹ CoMFA approach has found wide applications in drug design. The CoMSIA technique is of particular interest since it includes a hydrophobic field, hydrogen bond donor and acceptor fields and is less alignment-sensitive than CoMFA.²⁷

Results

The CoMFA and CoMSIA methods were applied to derive a 3D-QSAR model for benzofuran analogues as Nmt inhibitors with antifungal activity. Antifungal activity was expressed as a negative logarithm of minimum inhibitory concentration (MIC) (μ M) against *C. albicans* was used (Table 1) for the study. Conformation of the molecules used in the study was obtained by systematic search and the lowest energy conformer was selected and minimized using Powell method until the root mean square (RMS) gradient of 0.001 kcal/mol Å was obtained.

The alignment of the molecules was carried out using three techniques, namely RMS fitting (atom-based), multfit (flexible fitting) and SYBYL QSAR rigid body field fit. The most active molecule **20** was used as the template for alignment by considering the heavy atoms of the benzofuran ring as shown in Figure 1. The superimposition of all the molecules is shown in Figure 2. CoMFA models were generated using 20 molecules (1–20, Table 1), with column filtering value (σ) of 2.0. The generated models were validated using a test set of six molecules (21–26, Table 1). All the results of cross-validation were analyzed by considering the fact that a value of cross-validated r^2 (r_{cv}^2) above 0.3 indicates the probability of getting correlation value by chance is less than 5%.²⁸

The atom-based alignment exhibited r_{cv}^2 of 0.828 with six components, conventional r^2 (r_{conv}^2) of 0.989 and predictive r^2 (r_{pred}^2) of 0.754. The steric and electrostatic contributions were 50:50.

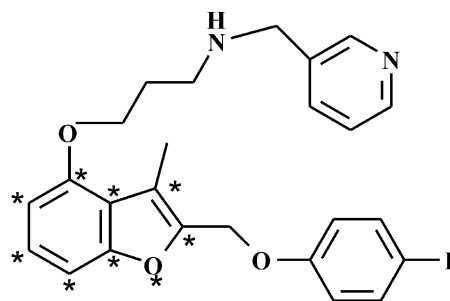


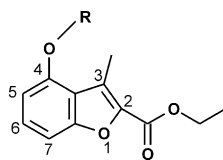

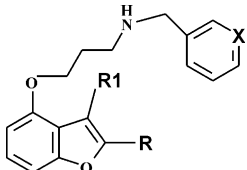
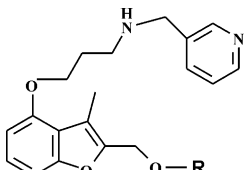
Figure 1. Molecule **20** with atoms used for superimposition are marked.

CoMFA models generated from multfit showed r_{cv}^2 of 0.813 with five components, r_{conv}^2 of 0.979 and r_{pred}^2 of 0.739. The steric and electrostatic contributions with this model were 50.5 and 49.5%, respectively. Realignment of the molecules by field fit with respect to the fields of template molecule (molecule **20**) yielded r_{cv}^2 of 0.692 with six components, r_{conv}^2 of 0.991 and r_{pred}^2 of 0.759. The steric and electrostatic contributions for this model were 38.7 and 61.3% respectively. The predictive ability and r_{cv}^2 of the three CoMFA models were equally good. The model generated with atom-based alignment with lower standard error of estimate (SEE = 0.192) and standard error of prediction (SEP = 0.204) was selected as the best model to explain the SAR and to carry out further analysis. The results obtained from the three different alignments are summarized in Table 2. Observed and predicted biological activities of the training and test sets are plotted in Figure 3. To further assess the robustness and statistical confidence of the derived 3D-QSAR model, bootstrapping analysis (100 runs) was performed and average result of 100 runs is 0.995 (r_{bs}^2).²⁹ To ascertain the true predictivity of the model a harder test using leave-half-out (LHO) method of cross-validation was performed 100 times and the mean r^2 is 0.766.³⁰ Negative value of r_{cv}^2 in randomized biological activity test revealed that the results were not based on chance correlation. Results of bootstrapping analysis, LHO and randomization test are indicated in Table 3.

The results of 3D-QSAR using CoMFA are represented as ‘coefficient contour’ maps. The contour maps obtained from the RMS fit model were used to explain the SAR of molecules in the present study.

CoMSIA was performed using steric, electrostatic, hydrogen bond donor (HB donor), hydrogen bond acceptor (HB acceptor) and hydrophobic fields. The atom-based alignment used for the CoMFA study served as alignment for CoMSIA. A total of 11 CoMSIA models were generated using combination of two, three, four and all fields. Though internal and external predictions of models obtained from two, four and all fields were equally good as that of model obtained from combination of three fields, the SEE and SEP were good in later case. The models obtained from combination of three fields, hydrophobic, HB donor and HB acceptor,

Table 1. Structures and activities of the molecules in the training (1–20) and test (21–26) sets

						
Compd	R	Antifungal activity ^a	CoMFA pred act ^b	CoMSIA pred act ^c		
1	O(CH ₂) ₂ NHC(CH ₃) ₃	−2.7708	−2.857	−2.754		
2	OCH ₂ CH ₂ (OH)NHCH(CH ₃) ₂	−2.5910	−2.666	−2.461		
3	OCH ₂ CH ₂ CH ₂ NHC(CH ₃) ₃	−2.4771	−2.491	−2.499		
4	OCH ₂ CH ₂ CH ₂ CH ₂ NHC(CH ₃) ₃	−2.2787	−2.093	−2.324		
5	OCH ₂ CH ₂ CH ₂ — 	−1.9294	−1.917	−1.877		
6	OCH ₂ CH ₂ CH ₂ CH ₂ CH ₂ NHC(CH ₃) ₃	−1.7242	−1.670	−1.789		
21	OCH ₂ CH ₂ CH ₂ NHCH(CH ₃) ₂	−2.6720	−2.466	−2.323		
22	OCH ₂ CH ₂ (OH)NHC(CH ₃) ₃	−2.3222	−2.234	−2.229		
						
Compd	R	R1	X	Antifungal activity ^a	CoMFA pred act ^b	CoMSIA pred act ^c
7	CH ₂ CH ₂ Ph	CH ₃	N	−1.1139	−1.223	−1.005
8	CONHPh	CH ₃	N	−1.0791	−1.043	−1.216
9	CH ₂ SPh	CH ₃	N	−1.0413	−1.008	−0.957
10	COOC ₂ H ₅	Cyclopropane	N	−1.0413	−1.140	−0.958
11	COOC ₂ H ₅	CH ₃	CH	−1.0000	−0.981	−1.048
12	COOC ₂ H ₅	CH ₃	N	−0.2041	−0.196	−0.280
23	COOC ₂ H ₅	CH ₂ CH ₃	N	−1.0791	−0.129	−0.535
						
Compd	R	Antifungal activity ^a	CoMFA pred act ^b	CoMSIA pred act ^c		
13	2-Cyanophenyl	−0.0791	0.344	−0.346		
14	Phenyl	0.4317	0.847	1.072		
15	3-Fluorophenyl	0.9586	0.952	1.157		
16	2,3-Difluorophenyl	1.0969	0.971	1.324		
17	2-Fluorophenyl	1.3979	0.760	1.213		
18	2,3,4-Fluorophenyl	1.4559	1.655	1.445		
19	2,4-Fluorophenyl	1.5228	1.390	1.313		
20	4-Fluorophenyl	1.6777	1.576	1.203		
24	4-Chlorophenyl	0.0362	1.371	1.543		
25	4-Cyanophenyl	0.7958	1.595	0.665		
26	2-Fluoro-4-bromophenyl	1.1307	1.103	1.881		

^aAntifungal activity is expressed as log (1/MIC) against *C. albicans* CY 1002.^bCoMFA predicted biological activity.^cCoMSIA predicted biological activity.

showed higher r_{cv}^2 (0.821) and a reasonable r_{pred}^2 (0.747). The model, characterized by a lower SEE (0.271) and a SEP (0.777), was selected as the best model to generate contour maps and explain the SAR. The results of CoMSIA are summarized in Table 4 and the observed versus predicted biological activities of the training and test sets are plotted in Figure 4.

Compound **24** was over predicted in both models, making it an outlier. However, retention of compound **24** in the generation of the model does not affect the statistical significance; hence, no attempts were made to perform PLS analysis without it.

The automated docking study of the benzofuran antifungals into the active site of the CaNmt revealed the

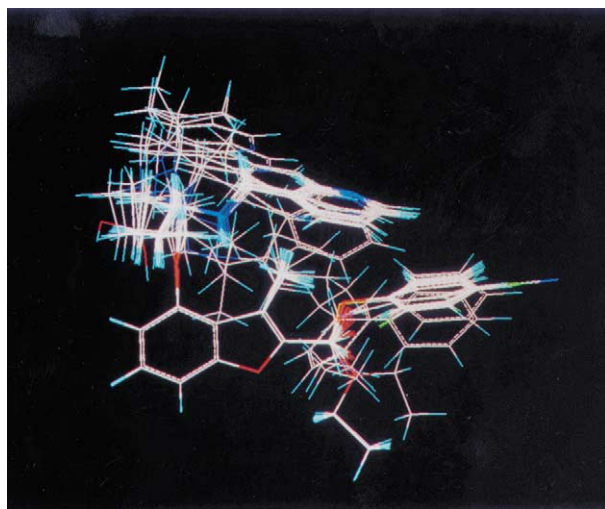


Figure 2. Superimposition of molecules (training and test set) using RMS fitting.

Table 2. Summary of CoMFA results

	Alignments		
	I ^a	II ^b	III ^c
r_{cv}^{2d}	0.828	0.813	0.692
Components	6	5	6
SEP	0.204	0.354	0.327
r_{con}^2	0.989	0.979	0.991
SEE	0.192	0.256	0.174
<i>F</i> value	195.773	130.773	239.771
Contrib. steric	50	50.5	38.7
electrostatic	50	49.5	61.3
r_{pred}^2	0.754	0.739	0.759

^aAlignment by RMS fit.

^bAlignment by multi fit.

^cAlignment by field fit.

^dCross-validated r^2 value was obtained from leave-one-out method.

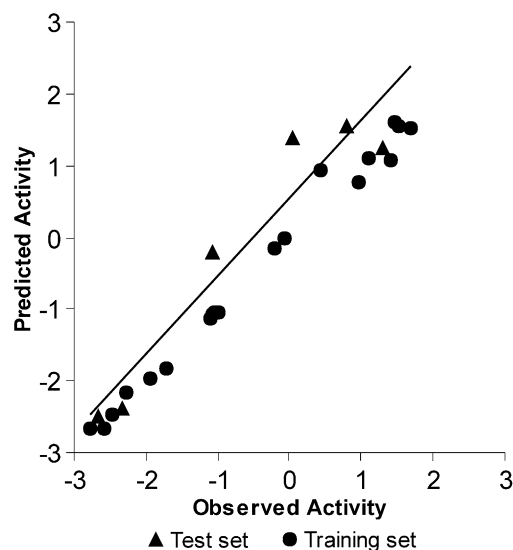


Figure 3. Graph of observed activity versus predicted activities of training and test set molecules from atom-based alignment of CoMFA analysis, activity expressed as log (1/MIC, μ M).

Table 3. Results of analysis with bootstrapping, leave-half-out cross-validation and randomized biological activities

	r_{bs}^{2a}		r_{cv}^{2b}		r^{2c}	
	CoMFA ^d	CoMSIA ^e	CoMFA ^d	CoMSIA ^e	CoMFA ^d	CoMSIA ^e
Mean	0.995	0.986	0.766	0.705	−0.134	−0.231
SD	0.004	0.009	0.011	0.024	0.139	0.286
High	—	—	0.877	0.848	0.169	0.246
Low	—	—	0.600	0.539	−0.741	−0.893

^aFrom 100 bootstrapping runs.

^bCross-validated r^2 by leave-half-out with optimum number of components, average 100 runs.

^cCross-validated r^2 with randomized biological activity, average of 100.

^dCoMFA model generated by RMS fit.

^eCoMSIA analysis by combined hydrophobic, donor and acceptor fields.

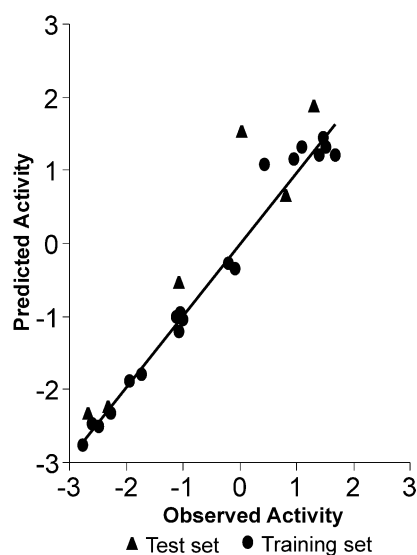


Figure 4. Graph of observed activity versus predicted activities of training and test set molecules from CoMSIA analysis with hydrophobic, donor and acceptor fields, activity expressed as log (1/MIC, μ M).

similarity in the binding mode to that observed in the crystal structure.³¹ In the active site, the pocket for the inhibitor binding site is hydrophobic in nature as it is lined by residues like Tyr107, Phe115, Phe117, Tyr119, Phe176, Tyr225 and so on, and also residues like Leu451 and Asn392 govern the electrostatic interactions. The interactions of compounds **20** (IC_{50} = 0.0052 μ M, MIC 0.021 μ M) and **3** (IC_{50} = 1.7 μ M, MIC = 300 μ M) are shown in Figures 9 and 10, respectively.

Compound **20** was observed to bind in the enzyme active site in an extended conformation. The aryl group at the C-2 position was located in the hydrophobic pocket formed by Phe115, Phe240 and Phe339. The benzofuran oxygen interacted with His227 through the hydrogen bond. The amide NH_2 of Asn392 and ether oxygen, though within 2.5 Å, could not form a H-bond, as the angle is not favorable. The secondary amine group of alkyl chain formed a H-bond with Leu451. The pyridine ring of the inhibitor interacted with

Table 4. Summary of CoMSIA results

	r^2_{cv} ^b	Components	SEP	r^2_{conv}	SEE	<i>F</i> value	<i>C</i> ^a					r^2_{pred}
							S	E	H	D	A	
S+E	0.613	1	0.969	0.744	0.788	52.370	24.2	75.8	—	—	—	0.875
D+A	0.757	6	0.905	0.697	0.322	63.811	—	—	—	25.3	74.7	0.859
S+D+A	0.766	6	0.888	0.969	0.324	67.233	14.8	—	—	19.2	66	0.873
E+D+A	0.693	2	0.889	0.892	0.527	70.157	—	33.2	—	20.8	45.9	0.815
S+E+D	0.590	1	0.998	0.727	0.814	48.039	21.8	39.2	—	39	—	0.849
S+E+A	0.710	2	0.864	0.892	0.528	69.917	14.5	36.2	—	—	49.3	0.845
S+E+H	0.644	8	1.190	0.990	0.204	130.370	13.1	50.7	36.1	—	—	0.615
H+D+A	0.821	6	0.777	0.978	0.271	97.102	—	—	38.8	15.5	51.8	0.747
S+E+D+A	0.693	2	0.888	0.893	0.526	70.592	11.1	30.3	—	17.8	40.7	0.816
S+H+D+A	0.814	6	0.791	0.971	0.298	95.377	9.2	—	28.8	13	49	0.714
ALL	0.700	1	0.879	0.902	0.503	77.810	9	26.4	16	14.4	34.3	0.819

S, steric; E, electronic; H, hydrophobic; D, hydrogen bond donor; A, hydrogen bond acceptor.

^a*C*, contribution.

^bCross-validated r^2 values obtained from leave-one-out method.

Tyr107 and Tyr119 through π – π stacking interactions and through H-bonding with Tyr119.

The compound **3** exhibited a similar docking mode into the enzyme active site. Since it lacked the pyridine ring, the hydrophobic and H-bonding interactions of the pyridine ring with enzyme residues were absent. The C-2 alkyl substituent of the inhibitor was located in the hydrophobic cavity formed by Phe240, Phe115 and Phe339.

Discussion

The CoMFA steric and electrostatic contour maps of the most active molecule **20** are shown in Figures 5 and 6, respectively. The contour plots are to be considered as a representation of the lattice points, where difference in field values is strongly associated with difference in receptor binding affinity. The absence of lattice points does not indicate that a given substructure element has no influence on the biological activity. It is likely that all the compounds studied exert the same steric and/or electrostatic influence in a certain area. Though CoMFA contour maps cannot be used as receptor maps, still they generate many useful interpretations.

Figure 5 depicts the CoMFA steric contour plot. The lipophilic fragments of the molecules are surrounded by the sterically favored green contours, which are flanked with a small unfavorable yellow region, suggesting that there is a definite requirement of a substructure with appropriate shape to exhibit biological activity. Large and small green contours are found surrounding the terminal phenyl ring at the C-2 position and pyridyl group attached to side chain at the C-4 position, which explains the importance of steric interactions of the ligand with the receptor. This is further supported by analyzing compounds **12** and **3–6**. Compound **12**, lacking hydrophobic functionality at the C-2 position, was moderately active, while compounds **3–6** lacking both hydrophobic functionality at C-2 and C-4 positions were inactive.

Figure 6 displays the electrostatic contour plot using CoMFA. The negative charge favorable red contours

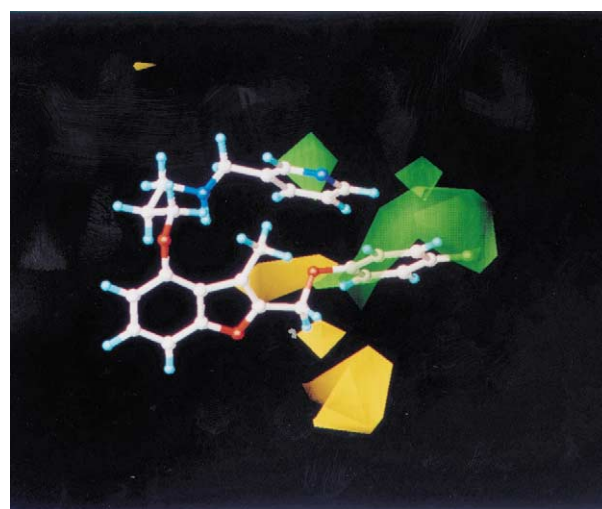


Figure 5. CoMFA steric STDEV* contour plots from the RMS fit. Sterically favored areas are represented by green polyhedra. Sterically disfavored areas are represented by yellow polyhedra. The active molecule **20** is shown in ball-and-sticks.

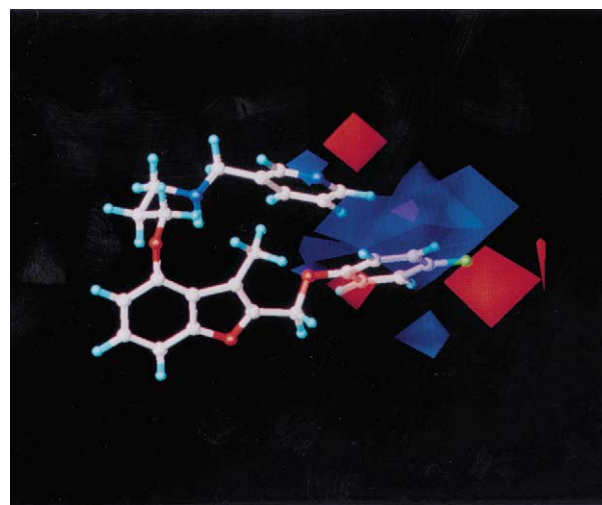


Figure 6. CoMFA electrostatic STDEV*COEFF contour plots from RMS fit. Positive charge favored areas are represented by blue polyhedra. Negative charge favored areas are represented by red polyhedra. The active molecule **20** is shown in ball-and-sticks.

are found near the ether linkage, aromatic ring of the C-2 position and near the pyridyl group at the C-4 position of benzofuran ring. These negative charge favorable contours are flanked with positive charge favorable blue region.

Red contour near ether linkage at the C-2 position of benzofuran ring shows the need for electron-rich atom for electrostatic interaction with the receptor to show good antifungal activity. Inactivity of the amide and alkyl spacer containing compounds **7** and **8**, respectively, are due to the lack of electron-rich atom to interact with the receptor through hydrogen bonding. Moderate activity of the compound **9** is due to difference in hydrogen acceptor capacity of sulphur than oxygen atom. Analysis of molecules **7**, **8** and **9** clearly indicates that ether linkage is optimum for potent antifungal activity.

A red contour near the phenyl ring explains the need of an electron-rich (electronegative) group to show potent antifungal activity. An aromatic ring having electron-withdrawing group is known to interact with another aromatic ring more strongly than does an unsubstituted aromatic ring.³² Compounds **20**, **25** and **26** showed highest activity due to the presence of electron-withdrawing group on aromatic ring at C-2 position than the unsubstituted aromatic compound **14**. This explains the need of strong hydrophobic substituent with electron-withdrawing group.

Both the green and red contours were found around the end of the C-4 side chain. It indicates the need of both hydrophobic and electron-rich substituent to show potent antifungal activity. In compounds **1**, **3**, **4** and **6**, the *t*-butyl group is oriented away from sterically favorable region (Fig. 7), which may be detrimental to their activity. Even though benzyl group of compound **11** having a proper spatial orientation towards sterically favorable region was moderately active because it lacks the electron-rich functional group. High activity of

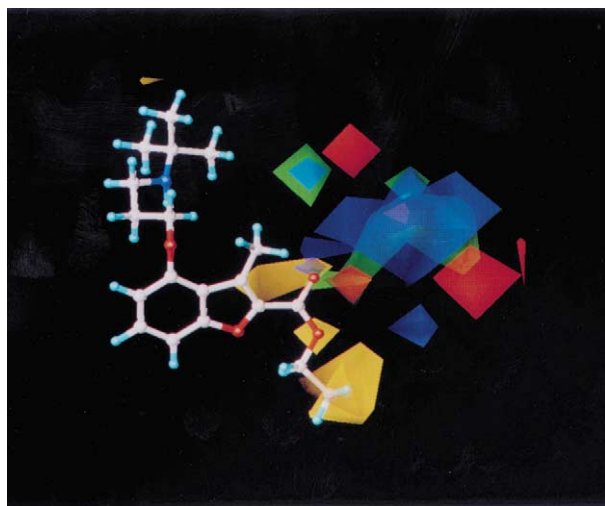


Figure 7. CoMFA steric and electrostatic STDEV*COEFF contour plots from RMS fit. The inactive molecule **3** is shown in ball-and-sticks.

compound **12** was mainly due to the proper orientation and the availability of pyridyl ring nitrogen for favorable steric and electrostatic interactions, respectively.

As combination of hydrophobic, HB donor and HB acceptor fields gave good statistical results (Table 4), this model was used to analyze CoMSIA 3D-plots. In Figure 8, yellow, white, cyan, purple, magenta and red contours found around different functional groups of most active compound **20** indicate the need of properly oriented groups to show potent antifungal activity. Hydrophobically favorable yellow region were found around the aromatic ring at the C-2 position of benzofuran ring system of most active molecule. Hydrophobically disfavorable white region adjacent to favorable yellow region indicates that for molecule to exhibit good antifungal activity, they should possess substructures with appropriate size and shape to occupy the hydrophobically favorable region. This is further supported by the fact that the same functional group in CoMFA analysis was surrounded by sterically favorable region. This clearly indicates the possibility of π - π stacking interaction with the amino acid residues in the active site. Molecules **1–6** and **21–23**, which lack this functional group at the C-4 position of benzofuran ring system, are inactive, where molecules having this functional group with electron-withdrawing substituents (compounds **15–20**, **25**, and **26**) are highly active than the unsubstituted aromatic compound **14**.

HB acceptor region indicated by magenta is flanked with HB acceptor disfavorable red region. This is mainly found around the 2-methyl pyridine group of the side chain at C-4 position of benzofuran ring system. Molecules which lack this HB acceptor pyridyl ring nitrogen (compounds **1–6**, **11** and **21–23**) are inactive.

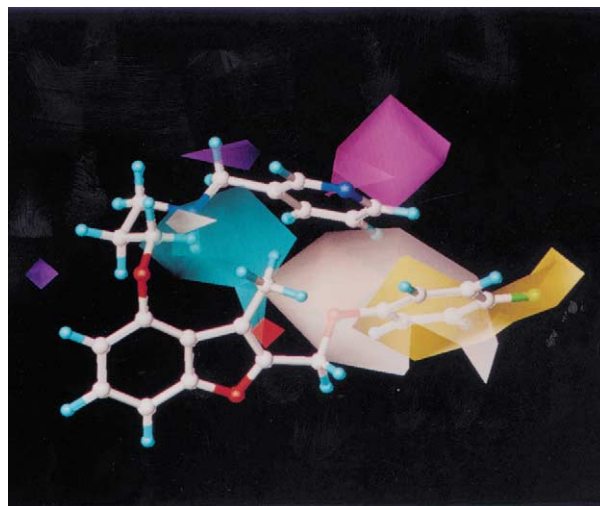


Figure 8. CoMSIA hydrophobic, HB donor and HB acceptor fields. Yellow indicates regions where hydrophobic substituents enhance activity; white indicates hydrophobic substituents reduces activity; cyan indicates regions where hydrogen bond donor functional groups enhance activity; purple indicates hydrogen bond donor functional group reduces activity, where magenta indicates regions where hydrogen bond acceptor functional group enhances activity; red indicates hydrogen bond acceptor substituents reduces activity. The active molecule **20** shown in ball-and-sticks.

The molecules **10** and **12** are moderately active even though they possess this functional group. This is due to lack of hydrophobic functional group at the C-2 position. This clearly indicates the need of both hydrophobic and HB acceptor functional group at C-4 and hydrophobic functional group with electron-rich substituent at the C-2 position of benzofuran ring, respectively, to show potent antifungal activity.

HB donor favorable cyan colour region was found along with small purple disfavorable region. This was observed around the secondary amine functional group at the C-4 position of benzofuran ring system. This indicates that the HB donor functional group may increase the activity. By observing the structures of all molecules, it clearly indicates the need of other two functional groups to exhibit potent antifungal activity.

Docking studies revealed that substituent at the C-2 position of benzofuran ring is located in a hydrophobic cavity formed by Phe115, Phe240 and Phe339. This is clearly evident from the presence of sterically favorable CoMFA contour (green) and hydrophobically favorable in CoMSIA contours (yellow) around the substituent of the C-2 position. It indicates that, for exhibiting antifungal activity, molecules should have aromatic substituent rather than aliphatic substituent at the C-2 position to interact with hydrophobic aromatic amino acid residues. In the CoMFA study, a negative charge favorable red contour near aromatic substituent at the C-2 position indicates the requirement of electron-rich substituent on the aromatic ring system for better π - π stacking interactions.

Similarly, a negative charge favorable red contour near ether linkage in CoMFA and the location of ether linkage near a hydrophilic amino acid like Asn392 indicates the necessity of ether linkage for potent antifungal activity. Whereas in molecule **3**, the ester oxygen is also located near Asn392 but its poor activity is due to the weak hydrogen bond forming capacity of ester oxygen and lack of aromatic substituent for π - π stacking interaction with active site residues.

The docking studies also revealed that the pyridyl ring at the C-4 position of molecule **20** (Fig. 9) was surrounded by hydrophobic amino acids Tyr107, Tyr119 and Phe176, justifying the presence of a sterically favorable green contour around pyridyl ring in CoMFA contours. A negative charge favorable red contour in CoMFA and HB acceptor favorable magenta contour in CoMSIA around the pyridyl ring are complimentary to the H-bond formation between the pyridyl ring nitrogen and OH of Tyr119 in the active site. Whereas, in molecule **3** (Fig. 10), the presence of the *t*-butyl group diminishes the additional hydrogen bonding interaction necessary for formation of a stronger complex between the ligand and the enzyme for better activity.

In CoMSIA study, the HB donor favorable cyan contour found around secondary amine of the C-4 substituent indicates the proper orientation of the secondary amine along with other requirements is

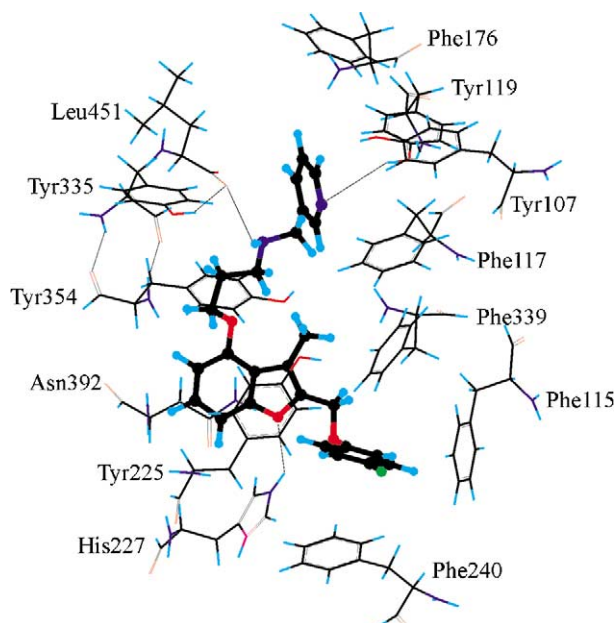


Figure 9. Stereoview of the binding mode of compound **20** (ball-and-sticks) in the active site of CaNmt. The active site residues in the binding pocket are shown in color by element.

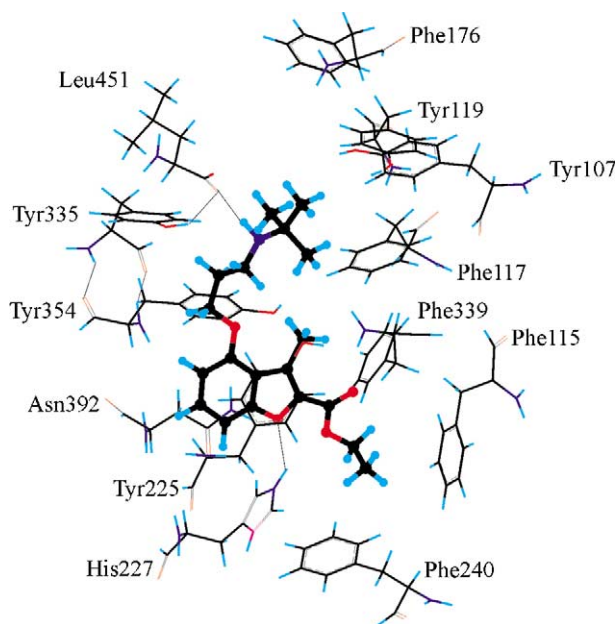


Figure 10. Stereoview of the binding mode of compound **3** (ball-and-sticks) in the active site of CaNmt. The active site residues in the binding pocket are shown in color by element.

essential for potent antifungal activity. This amine functionality is located near the Leu451 amino acid residue in the active site.

Conclusions

Comparative molecular field analysis and comparative molecular similarity indices analysis studies of *N*-myristoyltransferase inhibitors as antifungals were aimed at deriving structural requirements for antifungal activity. A data set of 26 molecules was divided into a training

set (20 molecules) and a test set (six molecules). The negative logarithm of inhibitory concentration (MIC) expressed in μM against fungi *C. albicans* was used as biological activity. Analysis of this data set resulted in CoMFA models with good correlative and predictive properties by the RMS fit method. Both the steric and electrostatic parameters significantly contribute towards the activity. The CoMSIA model generated by combination of hydrophobic, HB donor and HB acceptors fields showed good correlative and predictive properties. Predictive power of both derived models was assessed by the test set molecules. To check the robustness and statistical confidence of the derived models, bootstrapping analysis was performed. To further ascertain the true predictivity of the model, harder tests using LHO method of cross-validation and a cross-validation by randomization were performed. Derived CoMFA and CoMSIA models were able to predict the activities to a better extent and good predictions were obtained. The results were found to be consistent with the previously reported SAR studies.

The docking analysis of the benzofuran antifungals in CaNmt enzyme using an automated docking method gave a qualitative picture of ligand and macromolecule interactions which are not only consistent with the one in the reported crystal structure but are also complementary with the 3D-QSAR studies.

The statistical results obtained from the present study are moderately predictive. The structural information derived from the 3D-QSAR and docking studies were currently being used in the design of potent antifungal agents.

Methods

Biological data

Twenty-six molecules selected for the present study were taken from the published work by Tatsuo et al.^{15,16} These molecules have been reported as antifungals against *C. albicans* CY1002 through Nmt enzyme inhibition. The negative logarithm of inhibitory concentration (MIC) expressed in μM were used as biological activity in the 3D-QSAR study, thus correlating the data linear to the free energy change. A training set of 20 molecules (**1–20**, Table 1) was used for generation of QSAR models. The training set molecules were selected in such a way that there was no redundancy in information content in terms of both structural features and activity ranges. The most active compounds were included so that they provide critical information on pharmacophore requirements. Several moderately active and inactive compounds were also included to spread the activity ranges. A test set of six molecules (**21–26**, Table 1) was used to assess the predictive ability of the generated models. Selection of the test set molecule was made by considering the fact that test molecules represent a range of biological activity similar to the training set. The mean of biological activity of training and test sets were -0.539 and -0.685 , respectively. Thus, the test set is the true representative of the training set.

Computational details

3D-QSAR. All CoMFA and CoMSIA studies were performed on Silicon Graphics Indy R5000 workstation. Structural manipulations were performed using Sybyl 6.6³³ with standard Tripos force field.³⁴ Molecules were constructed using standard geometries and bond lengths. The initial conformations were obtained from systematic search. The lowest energy conformers were selected and minimized by using Powell method until root-mean-square (rms) deviation $0.001 \text{ Kcal/mol \AA}$ was achieved. Partial atomic charges required for calculation of the electrostatic interaction were computed by the semi-empirical molecular orbital method using the MOPAC program. The charges were computed using the AM1 method.

Alignment rule. The ‘alignment rule’, that is the positioning of a molecular model with the fixed lattice, is by far the most important input variable in CoMFA, since the relative interaction energies depend strongly on relative molecular positions. In the present study, we have superimposed molecules by three alignment rules.

Alignment 1. This was done by RMS fitting of the atoms to the most active molecule, compound **20**, by considering the heavy atoms of benzofuran ring as shown in Figure 1. The superimposition of all the molecules is shown in Figure 2.

Alignment 2. In this case, alignment of the molecules was carried out by flexible fitting (multifit) of atoms, of the molecules to the template molecule, compound **20**. This involved energy calculations and fitting onto the template molecule by applying force (force constant 20 Kcal/mol) and subsequent energy minimization.

Alignment 3. This was carried out using the SYBYL QSAR rigid body field fit command within SYBYL and using compound **20** as template molecule. Field fit adjusts the geometry of the molecules such that its steric and electrostatic fields match the fields of the template molecule.

Generation of CoMFA and CoMSIA fields

The steric and electrostatic potential fields for CoMFA were calculated at each lattice intersection of a regularly spaced grid of 2.0 \AA . The lattice was defined automatically, and is extended 4 \AA units past van der Waals volume of all the molecules in *X*, *Y* and *Z* directions. The van der Waals potential (Lennard–Jones, 6–12) and columbic term, which represent, respectively, steric and electrostatic fields, were calculated using the Tripos force field. A distance-dependent dielectric expression $\epsilon = \epsilon_0 \text{ Rij}$ with $\epsilon = 1.0$ was used. An sp^3 carbon atom with van der Waals radius of 1.52 \AA and $+1.0$ charge was served as the probe atom to calculate steric and electrostatic fields. The steric and electrostatic contributions were truncated to $\pm 30 \text{ kcal/mol}$, and the

electrostatic contributions were ignored at lattice intersections with maximum steric interactions.

The five CoMSIA similarity index fields available within Sybyl (steric, electrostatic, hydrophobic, HB donor and HB acceptor) were calculated at lattice points using a common probe atom of 1 Å radius, as well as the charge, hydrophobicity and hydrogen bond properties of 1 and an attenuation factor of 0.3.

Partial least square (PLS) analysis

PLS^{35,36} was used in conjugation with the cross-validation option to determine the optimum number of components, which were then used in deriving the final 3D-QSAR model without cross-validation. The results from cross-validation analysis were expressed as the cross-validated r^2 value (r_{cv}^2). Which is defined as,

$$r_{cv}^2 = 1 - \frac{\text{PRESS}}{\Sigma(Y - Y_{\text{mean}})^2}$$

where $\text{PRESS} = \Sigma (Y - Y_{\text{pred}})^2$.

The number of components that result in the highest r_{cv}^2 and lowest standard error of predictions (SEP) were taken as the optimum. Cross-validation was performed using leave-one-out (LOO) method^{29,37} in which one compound is removed from the data set and its activity is predicted using the model derived from the rest of the data set. LOO cross-validation was carried out with the number of components set equal to 10 and equal weights were assigned to steric and electrostatic fields using CoMFA_STD scaling option. To speed up the analysis and reduce the noise, a minimum filter value ' σ ' of 2.0 Kcal/mol was used. The LOO method of cross-validation is rather obsolete and it generally gives a high r^2 value. To ascertain the true predictivity of the model a harder test using LHO³⁰ method of cross-validation was performed for best models of CoMFA and CoMSIA methods. In this case, the data set is randomly divided into two groups, and the activity of the compounds from one group is predicted using the model from the other group. The process of group cross-validation was performed 100 times. The final r_{cv}^2 value was calculated by taking the mean of 100 runs. The r_{cv}^2 obtained from LOO and LHO were compared for each PLS analysis. In each case the optimum number of components was found to be the same as that obtained by the LOO cross-validation procedure.

To further assess the robustness and statistical confidence of the derived models, bootstrapping analysis (100 runs) was performed.²⁹ The statistical results obtained from bootstrapping analysis for both CoMFA and CoMSIA analysis are shown in Table 3. To check the probability of chance correlation, PLS analysis was performed by randomization of the biological activity. This was done by randomly changing biological activity data and performing PLS analysis to calculate the r_{cv}^2 value for both CoMFA and CoMSIA models. The

process was repeated 100 times. The results are indicated in Table 3.

Predictive r^2 value (r_{pred}^2)

To validate the derived CoMFA models, biological activities of the test set molecules were predicted using models derived from training set.

Predictive r^2 value was calculated using formula

$$r_{\text{pred}}^2 = \frac{\text{SD} - \text{PRESS}}{\text{SD}}$$

where SD is the sum of squared deviation between the biological activities of the test set molecule and the mean activity of the training set molecules and PRESS is the sum of squared deviations between the actual and the predicted activities of the test molecules.

Docking

All docking studies described herein were performed on a 750 MHz Intel Pentium III processor running Windows 98 using Molecular Operating Environment (MOE) 2002. 03 molecular modeling software.³⁸

For the current study we have used recently reported crystal structure of CaNmt enzyme (PDB: 1IYL). The reported crystal structure is a tetramer and only C fraction of the enzyme along with its inhibitor CaNmt-5 was isolated for further studies. Initially hydrogens were added to the enzyme and it was subjected to a well-defined minimization protocol. MOE-Dock was used to search for favorable binding configurations between a small, flexible ligand and a rigid macromolecular target. Searching was conducted within a specified 3-D docking box, using simulated annealing as the search protocol and MMFF94 forcefield^{39–41} The search protocol tries to optimize both purely spatial contacts as well as electrostatic interactions. MOE-Dock performs a series of independent docking runs and writes the resulting conformations and their energies to a molecular database file. During the calculations, the ligand molecule takes on conformations from the search trajectory.

The search protocol explores the ligand configuration space by generating random changes to the ligand's coordinates. Each such change to the ligand's conformation is called a *move*. The method has an acceptance (rejection) test which is applied to each move and which drives the search along a trajectory towards an optimal solution. Simulated annealing is a global optimization technique based on the Monte Carlo method. It explores various states of a configuration space by generating small random changes in the current state and then accepting or rejecting each new state according to the Metropolis criterion.⁴² According to this criterion, moves that decrease the energy of the system are always accepted, while moves that increase the energy of the system are accepted according to probability p :

$$p = e^{\frac{-\Delta u}{kt}} \quad \text{where } \Delta u = u_1 - u_0$$

where u_0 is the energy of the current state, u_1 is the energy of the new state, t is the 'temperature' of the simulation, and k is Boltzmann's constant. A simulated annealing run consists of a sequence of Monte Carlo cycles, each cycle consisting of a number of moves, or steps. The temperature is held constant during each cycle, and is systematically reduced from one cycle to the next.

The energy of a configuration is the sum of the electrostatic and dispersive interaction energy between the ligand and the target as well as the intra-molecular energy of the ligand due to its conformation. To calculate the interaction energies between the ligand and the target, MOE-Dock can use either the built-in potential function or *grid-based potential fields*. In the grid-based method, the interaction energy is calculated using electrostatic and van der Waals potential fields that have been sampled on a grid overlaying the docking box. The fields are interpolated at the atom positions by tri-linear interpolation. The van der Waals parameters are taken from the currently active forcefield, and the electrostatic field is calculated in a Coulombic manner using the current dielectric. The grid-based method calculates the potential energy grids only once, at the beginning of the docking procedure. Thus, subsequent energy calculations are considerably faster than if the built-in potential energy function were being used. However, when using grids, the solvation energy correction is not calculated.

For the grid-based docking of inhibitors, simulated annealing was used as the search protocol with a total of 25 runs, six cycles per run and 8000 steps per cycle and initial temperature of 1000 K. The best docking orientation was selected on the basis of H-bonding interactions, binding energy, hydrophobic interactions and conformational energy difference. The resulting complex was subjected to energy minimization using the following parameters: a distance-dependent dielectric constant, nonbonded cutoff of 12 Å, MMFF94 force field, and 200 steps of steepest descent followed by 200 steps of conjugate gradient and truncated Newton minimization until an energy gradient tolerance of 0.01 kcal/mol Å was satisfied.

Acknowledgements

The authors thank FDC Ltd, Mumbai. P.P. thanks Mr. Anand V. Raichurkar and Mr. Prashant S. Kharkar for the helpful discussion.

References and Notes

- Johnson, D. R.; Bhatnagar, R. S.; Knoll, L. J.; Gordon, J. I. *Annu. Rev. Biochem.* **1994**, *64*, 869.
- Boutin, J. A. *Cell Signal.* **1997**, *9*, 15.
- Bhatnagar, R. S.; Gordon, J. I. *Trends Cell. Biol.* **1997**, *7*, 14.
- McLaughlin, S.; Aderem, A. *Trends Biochem. Sci.* **1995**, *20*, 272.
- Duronio, R. J.; Towler, D. A.; Heuckeroth, R. O.; Gordon, J. I. *Science* **1989**, *16*, 796.
- Weinberg, R. A.; McWherter, C. A.; Freeman, S. K.; Wood, D. C.; Gordon, J. I.; Lee, S. C. *Mol. Microbiol.* **1995**, *16*, 241.
- Ledge, J. K.; Jackson-Machelski, E.; Toffaletti, D. L.; Perfect, J. R.; Gordon, J. I. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12008.
- Stkorski, J. A.; Devdas, B.; Zupec, M. E.; Freeman, S. K.; Brown, D. L.; Lu, H.; Nagarajan, S.; Mehta, P. P.; Wadw, A. C.; Kishore, N. S.; Bryant, M. L.; Getman, D. P.; McWherter, C. A.; Gerdon, J. I. *Biopolymers* **1997**, *43*, 43.
- White, T. C.; Marr, K. A.; Bowden, P. A. *Clin. Microbiol. Rev.* **1998**, *11*, 382.
- Benedetti, M. S.; Bani, M. *Drug. Metab. Rev.* **1991**, *31*, 665.
- Nagarajan, S. R.; Gordon, J. I.; Sikorski, J. A. *J. Med. Chem.* **1997**, *40*, 1422.
- Devdas, B.; Gordon, J. I.; Sikorski, J. A. *J. Med. Chem.* **1997**, *40*, 2609.
- Devdas, B.; Gordon, J. I.; Sikorski, J. A. *J. Med. Chem.* **1998**, *41*, 996.
- Karki, R. G.; Kulkarni, V. M. *Indian Drugs* **2001**, *38*, 406.
- Miyako, M.; Kenichi, K.; Hirosato, E.; Yoshihiko, I.; Shinji, T.; Satoshi, S.; Toshihiko, F.; Kiyoaki, S.; Yasuhiko, S.; Yuko, A.; Tatsuo, O.; Nobuo, S. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1833.
- Hirosato, E.; Miyako, M.; Pingi, L.; Kenichi, K.; Kenji, M.; Satoshi, S.; Michiko, H.; Toshihiko, F.; Kiyoaki, S.; Hidetoshi, S.; Yasuhiko, S.; Yuko, A.; Tatsuo, O.; Nobuo, S. *Bioorg. Med. Chem. Lett.* **2002**, *10*, 607.
- Tanaji, T. T.; Santosh, S. K.; Kulkarni, V. M. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 958.
- Tanaji, T. T.; Hariprasad, V.; Kulkarni, V. M. *Drug Des. Dis.* **1998**, *15*, 181.
- Tanaji, T. T.; Kulkarni, V. M. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 204.
- Gokhale, V. M.; Kulkarni, V. M. *J. Comput. Aided Mol. Des.* **2000**, *14*, 495.
- Gokhale, V. M.; Kulkarni, V. M. *J. Med. Chem.* **1999**, *42*, 5348.
- Karki, R. G.; Kulkarni, V. M. *Eur. J. Med. Chem.* **2001**, *36*, 147.
- Santosh, S. K.; Lalji, K.; Gediya; Kulkarni, V. M. *Bioorg. Med. Chem.* **1999**, *7*, 1475.
- Murthy, S. V.; Kulkarni, V. M. *Bioorg. Med. Chem.* **2002**, *10*, 2267.
- Kulkarni, S. S.; Kulkarni, V. M. *J. Med. Chem.* **1999**, *42*, 373.
- Mahindra, M. T.; Kulkarni, V. M. *J. Comput. Aided Mol. Des.* **2002**, *16*, 181.
- Klebe, G.; Abraham, U. *J. Comput. Aided. Mol. Des.* **1999**, *13*, 1.
- Kaminsky, J. J.; Doweyko, A. M. *J. Med. Chem.* **1997**, *40*, 427.
- Cramer, R. D., III; Bunce, J. D.; Patterson, D. E. *Quant. Struct. Act. Relat.* **1988**, *7*, 18.
- Ciubotariu, D.; Deretey, E.; Opera, T. I.; Sulea, T.; Simon, Z.; Kurunczi, L.; Chiriac, A. *Quant. Struct. Act. Relat.* **1993**, *12*, 367.
- Satoshi, S.; Miyako, M.; Kiyoaki, S.; Takaaki, A. F.; Kenji, M.; Yasuhiko, S.; Hirosato, E.; Kenichi, K.; Yuko, A.; Nobuo, S.; Allan, D'A.; Fritz, K. W.; David, W. B.; Tatsuo, O. *Chem. Biol.* **2002**, *9*, 1119.
- Hunter, C. A.; Sanders, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 5525.
- SYBYL 6.6 molecular modeling software is available from

Tripose Associates Inc., 1699, St Hanley Road, St Louis, MO 63114-2913, USA.

34. Clark, M.; Cramer, R. D., III; Van Opdenbosh, N. *J. Comput. Chem.* **1989**, 982.

35. Wold, S.; Albano, C.; Dunn, W. J.; Edlund, U.; Esbenson, K.; Geladi, P.; Hellberg, S.; Lindburg, W.; Sjostrom, M. In *Chemometrics*; Kowalski, B., Int. Ed.; Reidel: Dordrecht, The Netherlands, 1984; p 17.

36. Geladi, P. *J. Chemom.* **1998**, 2, 231.

37. Wold, S. *Technometrics* **1978**, 4, 397.

38. MOE software is available from Chemical Computing Group Inc., 1010 Sherbrooke Street West, Suite 910, Montreal, Canada H3A 2R7.

39. Halgren, T. A. *J. Am. Chem. Soc.* **1990**, 112, 4710.

40. Halgren, T. A. *J. Comp. Chem.* **1996**, 17, 490.

41. Halgren, T. A. *J. Comp. Chem.* **1999**, 20, 730.

42. Metropolis, N.; Resenbluta, A. W.; Resenbluta, M. N.; Teller, A. H. *J. Chem. Phys.* **1953**, 21, 1087.