WORLD JOURNAL OF PHARMACEUTICAL RESEARCH
SJIF Impact Factor 5.045
Volume 4, Issue 1, 1673-1692. Research Article ISSN 2277– 7105

3D- QSAR OF N- SUBSTITUTED IMIDAZOLES AS ANTIFUNGAL AGENTS

Vithal M. Kulkarni*, Sandeep S. Pathare and Kakasaheb R. Mahadik

Department of Pharmaceutical Chemistry, Poona College of Pharmacy, Bharati Vidyapeeth University, Erandwane, Pune - 411038, Maharashtra, India.

Article Received on
12 Nov 2014,

Revised on 05 Dec 2014,
Accepted on 29 Dec 2014

***Correspondence for Author**
Dr. Vithal M. Kulkarni
Prof, Department of Pharmaceutical Chemistry, Poona College of Pharmacy, Bharati Vidyapeeth University, Erandwane, Pune - 411038, Maharashtra, India.

ABSTRACT

Azole antifungals are widely studied using various methods of drug design. A series of N-substituted imidazole derivatives was examined to determine the structural requirements for antifungal activity by three-dimensional quantitative structure-activity relationship (3D-QSAR) using comparative molecular field analysis (CoMFA). A training set of 50 compounds was used to establish the CoMFA model, which was validated by evaluation of a test set of 15 compounds. In this study, the superimposition of molecules was carried out by atom-based fit (rms), multi fit and field fit. The best QSAR model was obtained from rms fit with cross-validated r^2 = 0.725, conventional r^2 = 0.939 and predictive r^2 = 0.518. This series of compounds was also analyzed by genetic function approximation (GFA). The best model with cross-validated r^2 of 0.608 and predictive r^2 of 0.391 emphasized the importance of molecular shape analysis parameters. The models obtained from the present study may be useful for the development of new imidazole derivatives as potential antifungals.

KEYWORDS: QSAR, CoMFA, Antifungal activity, N-substituted Imidazole derivatives.

INTRODUCTION

The frequencies and type of life-threatening fungal infections have increased dramatically in immunocompromised patients.[1] The major opportunistic pathogen has been *Candida albicans.* The management of invasive fungal infection utilizes a variable multidisciplinary approach involving antifungals, appropriate surgery and immuno-correction. Currently available antifungal drugs have essentially three molecular targets: sterol 14α-demethylase

(azoles), ergosterol (polyenes) and β-1,3-glucan synthetase (echinocandins). Azoles antifungals interfere with cell membrane ergosterol synthesis via inhibition of cytochrome P450 14-α-sterol demethylase enzyme.[2,3] The natural substrate lanosterol is prevented through binding of the azole ring to the iron of porphyrin.[4] Limitations in their clinical applications include narrow spectrum (fluconazole), variable bioavailability (itraconazole), drug interactions, e.g. with cyclosporin, and emergence of drug resistance.[5] With the increase in the incidence of fungal infections and rise in azole resistance, there is urgent need for new potent antifungals. In the absence of crystal structure of *Candida albican* cytochrome P450 dependent 14-α-demethylase (CA-CYP51), the rational design of new antifungals has been carried out using various methods such as homology modeling,[6] pharmacophore modeling[7] and quantitative structure activity relationship studies.[8-9]

Recently, Di Santo et al[10] reported the synthesis and QSAR of a series of N-substituted derivatives of 1-[(aryl)(4-aryl-1H-pyrrol-3-yl)methyl]-1H-imidazole. In order to validate their pharmacophore models and gain further insight into the structure activity relationship, we applied a three dimensional quantitative structure activity relationship (3D QSAR) models for the same series using Comparative Molecular Field Analysis (CoMFA)[11] and genetic function approximation (GFA).[12]

CoMFA is one of the methods of rational drug design, which has been successfully applied in our laboratory for N-myristoylase inhibitors,[13] antihyperglycemic agents,[14] HIV-1 integrase inhibitors[15] and squalene epoxidase enzyme inhibitors.[16] The CoMFA approach has found wide application in drug design.

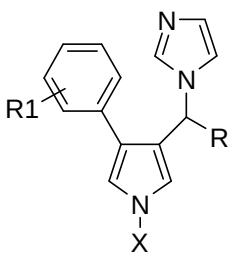
GFA, a genetic algorithm, generates a population of equations rather than one single equation, in correlating biological activity with physicochemical descriptors. GFA, as developed by Rogers,[17] involved combination of Friedman's multivariate adaptive regression splines (MARS) algorithm with Holland's genetic algorithm to evolve a population of equations that best fit the training set data. GFA models, which provide useful additional information such as relevance of a particular descriptor in the model and activity prediction, have been applied in the past to various therapeutic areas.[18-19]

Experimental Section

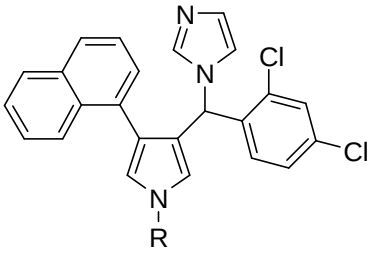
Biological data

Sixty-six molecules selected for the present study were taken from the published work by Di Santo, *et. Al.*[10] The structure of the compounds and their biological data are given in Table 1. The antimycotic activity against *C. albicans* was expressed as the minimum inhibitory concentration (MIC) in terms of (µmol/mL). In this QSAR study, the biological activity of each compound has been expressed as negative logarithm of MIC_{mean} (ratio of $MIC_{compd}/MIC_{bifonazole}$). Thus the data correlated linearly to the free energy change. Fluconazole was not considered either for generation or validation of QSAR models, as is doesn't have imidazole ring, which is common to all other structures. A training set of fifty molecules (Table 1) was used for generation of QSAR models. The training set molecules were selected in such a way that it contains information 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 fifteen molecules (indicated by *, Table 1) was used to access the predictive ability of the generated models. The test molecules represent range of biological activity similar to training set.

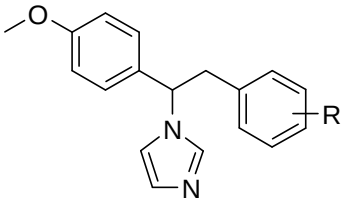
Table 1. Structures and activities of the molecules in the training and test sets.



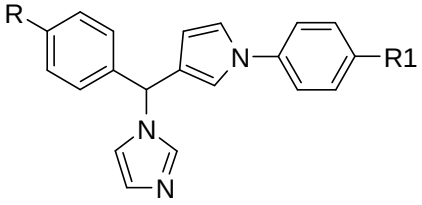
A



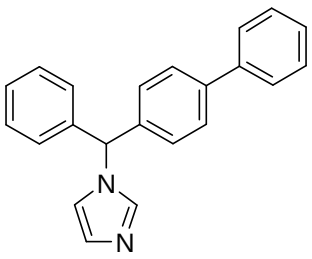
B



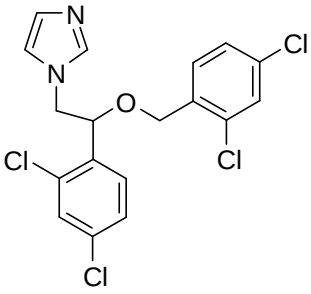
C



D

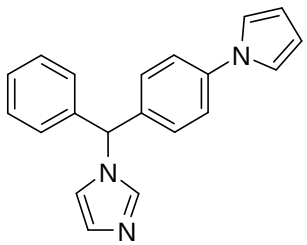


Bifonazole



Miconazole

Compd. No.	Scaffold	R	R_1	X	$MIC_{compd}/MIC_{bifonazole}$	
					a	b
1	A	Ph	4-Cl	H	0.660	0.180
2	A	Ph	4-Cl	CH_3	0.025	1.602
3	A	Ph	4-Cl	C_2H_5	0.110	0.959
4	A	Ph	4-Cl	C_3H_7	0.023	1.638
5	A	Ph	4-Cl	CH_2-c- C_3H_7	0.025	1.602
6	A	Ph	4-Cl	$CH=CH_2$	0.031	1.509
7	A	Ph	4-Cl	$CH2-CH=CH_2$	0.019	1.721
8	A	Ph	4-Cl	$CH2-CH=C(CH_3)_2$	0.043	1.367
9	A	Ph	4-Cl	$CH=CH-COOCH_3$	0.340	0.469
10	A	Ph	4-Cl	Ph	1.500	-0.176
11*	A	Ph	2,4-Cl_2	CH_3	0.11	0.959
12	A	Ph	2,4-Cl_2	C_2H_5	0.330	0.481
13*	A	2,4-Cl_2-Ph	4-Cl	H	0.920	0.036
14	A	2,4-Cl_2-Ph	4-Cl	CH_3	0.280	0.553
15	A	2,4-Cl_2-Ph	4-Cl	C_2H_5	0.310	0.509
16	A	2,4-Cl_2-Ph	4-Cl	C_3H_7	1.700	-0.230
17	A	2,4-Cl_2-Ph	4-Cl	$CH_2-CH=CH_2$	0.150	0.824
18*	A	2,4-Cl_2-Ph	4-Cl	$CH_2-CH\ (OCH_3)_2$	0.480	0.319
19	A	2,4-Cl_2-Ph	2-Cl	CH_3	0.360	0.444
20	A	2,4-Cl_2-Ph	2,4-Cl_2	CH_3	2.900	-0.462
21	A	2,4-Cl_2-Ph	2,4-Cl_2	C_2H_5	0.230	0.638
22	A	2,4-Cl_2-Ph	2,4-Cl_2	$CH_2-CH=CH_2$	0.120	0.921
23	A	4-CH_3-Ph	3,4-Cl_2	CH_3	0.190	0.721
24	A	4-CH3-Ph	3,4-Cl_2	C_2H_5	0.470	0.327
25	A	2,4-Cl_2-Ph	4-(1-pyrrolyl)	CH_3	0.620	0.207
26*	A	1-naphthyl	H	H	1.700	-0.230
27	A	2-naphthyl	H	H	3.800	-0.579
28*	A	2-naphthyl	H	CH_3	63.00	-1.799
29	A	1-naphthyl	H	CH_3	6.000	-0.778
30	A	2,4-Cl_2-Ph	4-Ph	H	4.300	-0.633
31	A	2,4-Cl_2-Ph	4-CF_3	H	2.200	-0.342
32	A	2,4-Cl_2-Ph	4-CN	H	4.300	-0.633
33	A	2,4-Cl_2-Ph	4-NO_2	H	4.500	-0.653
34*	A	2,4-Cl_2-Ph	4-NH_2	H	28.00	-1.447
35	A	2,4-Cl_2-Ph	4-(1-pyrrolyl)	H	1.300	-0.114
36*	A	2,4-Cl_2-Ph	4-OH	H	26.00	-1.415
37	A	2,4-Cl_2-Ph	4-SCH_3	H	28.00	-0.441
38*	B	H			0.210	0.678
39	B	CH_3			0.940	0.027
40	B	C_2H_5			1.100	-0.041
41*	C	2,4-Cl_2			1.000	0.00

42	C	4-NH_2			1.400	-0.146
43	C	3-(1-pyrrolyl)			45.00	-1.653
44	C	2-Cl			63.00	-1.799
45*	C	3-F			51.00	-1.708
46	C	2-NO_2			97.00	-1.87
47	C	2-(1-pyrrolyl)			7.800	-0.892
48	C	4-NO_2			17.00	-1.230
49*	C	3-Cl			44.00	-1.643
50	C	2-F			49.00	-1.690
51	C	4-F			23.00	-1.361
52*					6.800	-0.833
53	D	H	H		2.300	-0.362
54*	D	CH_3	H		1.200	-0.079
55	D	Cl	Cl		7.300	-0.863
56	D	F	Cl		2.800	-0.447
57	D	H	Cl		0.700	0.155
58*	D	CH_3	F		3.100	-0.491
59	D	Cl	F		2.700	-0.431
60	D	H	F		1.800	--0.255
61*	D	F	H		4.100	-0.613
62	D	CH_3	Cl		1.100	-0.041
63	D	F	F		1.400	-0.146
64	Bifonazole				1.000	0
65	Miconazole				0.140	-0.854

a = experimental $MIC_{compd}/MIC_{bifonazole}$ in terms of µmol/mL, b = -log ($MIC_{compd}/MIC_{bifonazole}$),

** indicates test set compounds*

Computational details

CoMFA study

All computational studies were performed using SYBYL 6.9.1[20] with standard Tripos force field.[21] The compounds were constructed from the fragments in the SYBYL database with standard bond lengths and bond angles. The chirality of asymmetric center was not specified as enantiomers were not specified. Geometry optimization was carried out using the standard Tripos forcefield with distance dependent-dielectric function and energy gradient of 0.001 kcal/mol Å. The initial conformations were obtained from systematic search. The lowest

energy conformers were selected and minimized by using Powell method till root-mean-square (rms) deviation 0.001 kcal/mol Å was achieved. Partial atomic charges required for calculation of the electrostatic interaction were computed by semiempirical molecular orbital method using AM1 in MOPAC program.

Alignment rules

The "alignment rule", i.e., 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. The most active molecule (07) was used as template for aligning the other molecules.

In the present study, we have superimposed molecules by three alignment rules: (1) Atom-based alignment, (2) Multifit alignment, (3) Field fit alignment (1) This was done by atom-based fitting of the atoms to the most active molecule, compound 07. The 11 heavy ring atoms (azole and phenyl ring attached to the asymmetric carbon atom) of the molecules were used for rms fitting, as shown in Fig. 1.

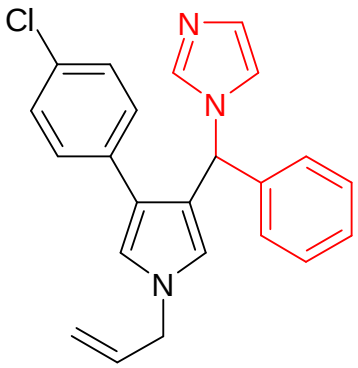


Fig. 1. Molecule 07 with atoms used for superimposition are marked in red.

(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 07. This involved energy calculations and fitting onto the template molecule by applying force (force constant 20 kcal/mol) and subsequent energy minimization.

(3) This was carried out using the SYBYL QSAR rigid body field fit command within SYBYL and using compound 07 as template molecule.

Generation of CoMFA fields

For each alignment, the steric and electrostatic potential fields for CoMFA were calculated at each lattice intersection of a regularly spaced grid of 2.0 Å in all 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 constant of 1.0 was used. A sp^3 carbon atom with van der Waals radius of 1.52 Å and + 1.0 charge was served as the probe atom to calculate steric and electrostatic fields. The steric and electrostatic contributions were truncated to ±30 kcal/mol, and the electrostatic contributions were ignored at lattice intersections with maximum steric interactions.

Partial Least Square (PLS) analysis

PLS[22-23] was used in conjugation with the cross-validation (leave-one-out) option to determine the optimum number of components. Final 3D-QSAR model without cross-validation was done using the optimal number of components. The results from cross-validation analysis were expressed as the cross-validated r^2 value (r^2_{cv}), which is defined as,

$$r^2_{cv} = 1 - \frac{PRESS}{\sum (Y - Y_{mean})^2}$$

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

The number of components that result in the highest r^2_{cv} and lowest standard error of predictions (SEP) were taken as the optimum. 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 leave one out (LOO)[24] method of cross-validation is rather obsolete and it generally gives high r^2 value. Final analysis was performed to calculate the r^2_{conv} with a number of cross-validation groups set to zero using the optimum number of components. To further assess the robustness and statistical confidence of the derived models, bootstrapping analysis (100 runs) was performed. The statistical results obtained for CoMFA analysis are shown in Table 2.

Table 2. Summary of CoMFA results.

	Alignments		
	1[a]	2[b]	3[c]
[d]r^2_{cv}	0.725	0.667	0.565
Components	6	4	7
SEP	0.516	0.556	0.657
r^2_{con}	0.939	0.913	0.921
SEE	0.244	0.284	0.280
F Value	109.534	118.251	70.066
Contrib. Steric	0.568	0.561	0.540
Electrostatic	0.432	0.439	0.460
r^2_{pred}	0.518	0.512	0.511
r^2_{BS}	0.964	0.945	0.943

[a]Alignment by RMS fit, [b]Alignment by Multi fit, [c]Alignment by Field fit, **[d]***cross-validated r^2 value was obtained from leave-one-out method.*

Leave half out (LHO)[25] method of cross-validation was performed for RMS fit analysis of CoMFA. 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^2_{cv} value was calculated by taking the mean of 100 runs. The r^2_{cv} 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. The statistical results obtained from LHO for CoMFA analysis are shown in Table 3.

Table 3. Results of analysis with cross validation for 2 and 5 groups and randomized biological activities for RMS fit

Sr. No	r^2cv for 2 groups[a]	r^2cv for 5 groups[b]	Randomized r^2 [c]
Mean	0.532	0.541	-0.185
SD	0.086	0.055	0.203
High	0.764	0.701	0.057
Low	0.439	0.504	-0.066

[a]Cross-validated r^2 for 2 groups with optimum number of components, average 25 runs

[b]Cross-validated r^2 for 5 groups with optimum number of components, average 25 runs

[c]Cross-validated r^2 with randomized biological activity, average of 25 runs

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^2_{cv} value for RMS fit of CoMFA. The process was repeated 100 times. Further cross validation with 5 groups was also carried out. The results are indicated in Table 3.

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

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

$$r^2_{pred} = \frac{SD - PRESS}{SD}$$

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.

Model building using GFA

All the models were developed using Cerius2 (version 4.10L) running on running Linux Red Hat Enterprise WS 3.0 on Intel Pentium IV 3.0 GHz processor.[26] Structures were constructed and partial charges were assigned using the charge equilibration method within Cerius2. Throughout the study Universal forcefield 1.02 was used.[27] The molecules were subsequently minimized using smart minimizer until root mean square deviation 0.001 kcal/mol Å was achieved and used in the study.[28]

Conformational sampling

The local minimized geometry was used as the initial structure for conformational analysis. Conformational ensembles were generated by random sampling using a rotation increment of 10^0 for all the torsional angles. In order to restrict the number of conformers being generated to maximum of 50, conformers with an energy threshold value of greater than 5 kcal/mol from the local minimized structure were rejected, thus selecting only energetically stable conformers.

Calculation of descriptors

Different physicochemical descriptors were calculated for each molecule in the study table using default settings within Cerius2. These descriptors included electronic, spatial, structural, thermodynamic and molecular shape analysis (MSA). The lowest energy conformer of compound 07 was taken as the reference for calculation of MSA. A complete list of descriptors used in the study is given in Table 4.

Table 4. Descriptors used in the GFA analysis.

Sr. No	Descriptor	Type	Description
1	DIFFV	MSA	Difference volume
2	COSV	MSA	Common overlap steric volume
3	Fo	MSA	Common overlap volume ratio
4	NCOSV	MSA	Non-common overlap steric volume
5	ShapeRMS	MSA	RMS to shape reference
6	SR Vol	MSA	Volume of shape reference compound
7	Vm	Spatial	Molecular volume
8	Area	Spatial	Molecular surface area
9	Density	Spatial	Molecular density
10	RadOfGyr	Spatial	Radius of gyration
11	PMI-mag	Spatial	Principal moment of inertia
12	Charge	Electronic	Sum of partial charges
13	Apol	Electronic	Sum of atomic polarizabilities
14	Dipole-mag	Electronic	Dipole moment
15	HOMO	Electronic	Highest occupied molecular orbital energy
16	LUMO	Electronic	Lowest unoccupied molecular orbital energy
17	Sr	Electronic	Super delocalizability
18	MW	Structural	Molecular weight
19	RotlBonds	Structural	Number of rotatable bonds
20	HbondAcc	Structural	Number of hydrogen bond accepters
21	HbondDon	Structural	Number of hydrogen bond donors
22	AlogP	Thermodyna	Logarithm of partition coefficient
23	Fh_2o	Thermodyna	Desolvation free energy for water
24	Foct	Thermodyna	Desolvation free energy for octanol
25	Hf	Thermodyna	Heat of formation
26	molRef	Thermodyna	Molar refractivity
27	Kier1	Topological	Kier index first order
28	Kier2	Topological	Kier index second order
29	Kier3	Topological	Kier index third order
30	Energy	Electronic	Energy

Generation of QSAR models

QSAR analysis establishes relationship between physicochemical properties and biological activity of the compounds studied. In the present study, QSAR model generation was performed by GFA technique using 25,000 crossovers, smoothness value of 1.00 and other

default settings. GFA was asked to consider, not more than four terms in the equation. The set of equations generated was evaluated on the following basis.

a. Lack of fit (LOF)
b. Variable terms in the equation
c. Predictivity of the equation (predictive r^2 value)

Cross-validated r^2 values (r^2_{cv}) were calculated using the cross-validation test option in statistical tools supported within Cerius2.

RESULTS

CoMFA

The CoMFA method was applied to derive a 3D-QSAR model for N-substituted derivatives of 1-[(aryl)(4-aryl-1H-pyrrol-3-yl)methyl]-1H-imidazole with antifungal activity. The negative logarithm of MIC_{mean} was used as biological activity in 3D-QSAR study (Table 1). Conformation of the molecules used in the study was obtained by systematic search and lowest energy conformer was selected and minimized using Powell method to rms 0.001 Kcal/mol Å.

Alignment of the molecules was carried out using three techniques, namely RMS fitting (atom-based), multifit (flexible fitting) and SYBYL QSAR rigid body field fit. The most active molecule (07) was used as the template molecule for alignment (Fig. 1). CoMFA models were generated using a training set of fifty molecules (Table 1), with column filtering value (σ min) 2.0. A training set of fifteen molecules (Table 1) was used to check the external predictivity of the models.

The atom-based alignment exhibited r^2_{cv} of 0.725 with six components, conventional r^2 (r^2_{conv}) of 0.939, predictive r^2 (r^2_{pred}) of 0.518, F value of 109.30. CoMFA models generated for multifit alignment showed r^2_{cv} of 0.667 with five components, r^2_{conv} of 0.913, r^2_{pred} of 0.512, F value of 118.25. Realignment of the molecules by field fit with respect to the fields of template molecule (molecule 07) yielded r^2_{cv} of 0.565 with six components, r^2_{conv} of 0.921, r^2_{pred} of 0.511, F value of 70.06. The external predictive ability, r^2_{pred} of the three CoMFA models are equally good, the model generated with atom-based alignment with good internal predictive ability (r^2_{cv} = 0.725) and small standard error of estimate (SEE = 0.244) was selected as the best model to explain SAR and to carry out further analysis. Results obtained

from the three different alignments are shown in Table 2. Observed and predicted biological activities of the training and test sets are plotted in Fig. 2 and Fig. 3 respectively.

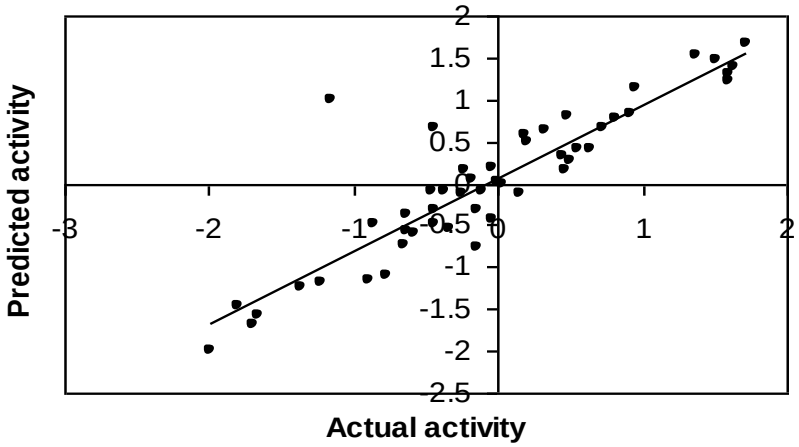


Fig. 2. Graph of observed activity versus predicted activities of training set molecules from multi fit alignment of CoMFA analysis, activity expressed as -log ($MIC_{compd}/MIC_{bifonazole}$).

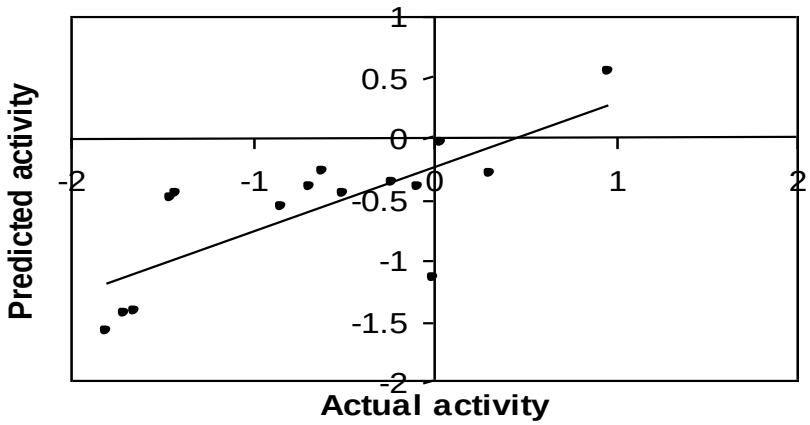


Fig. 3. Graph of observed activity versus predicted activities of test set molecules from multi fit alignment of CoMFA analysis, activity expressed as -log ($MIC_{compd}/MIC_{bifonazole}$).

To further assess the robustness and statistical confidence of the derived 3D QSAR model, bootstrapping analysis was performed and average result of 100 runs is 0.964 (r^2_{bs}). 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 0.764. Negative value of r^2_{cv} in randomized biological activity test revealed that the results were not based on chance

correlation. Results of bootstrapping analysis, LHO and randomization test are shown in Table 3.

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

GFA

Different QSAR equations were generated using the GFA algorithm in Cerius2 for a series of imidazole antifungals. A total of 50 compounds were used for QSAR model generation (Table 1). The predictive power of models was assessed by using test set of 15 compounds (Table 1), such that it represents various functional groups included in the training set and had an even distribution of biological activity.

For each of the compounds a conformational database was generated from local minimized structure by random sampling method. A total of 34 descriptors were calculated using the Cerius2 molecular modeling package. A list of descriptors is summarized in Table 4.

Different sets of equations were generated by altering chain length of the equations. The generated equations were evolved by repeating GA runs to check the stability of GFA models. The best model was selected on the basis of the values of r2 (square of the correlation coefficient for the training set compounds), r^2_{cv} (crossvalidated r^2), LOF (Friedman's Lack of Fit), r^2_{pred} (predictive r^2 for test set compounds). The best QSAR models are described in Table 5.

Table 5. QSAR equations generated using genetic function approximation for the training set of fifty molecules.

No	Equation	LOF	r^2	$^{b}r^2_{cv}$	^{c}BS r^2	F-test	$^{d}r^2_{pred}$
1	BA = 2.795 – 0.769 CHI-1 + 0.402 ShapeRMS + 0.904 RadGyr + 0.029 DIFFV +4.615 Fo	0.432	0.672	0.608	0.674	18.061	0.391
2	BA = 5.597 + 4.539 F0 – 0.709 CHI-1 + 0.031 DIFFV + 0.501 SapeRMS	0.422	0.647	0.536	0.648	20.605	0.354

The best model had six descriptors including a constant. The model exhibited good internal as well as external predictive ability. Observed and predicted biological activities of the test sets are plotted in Fig. 4.

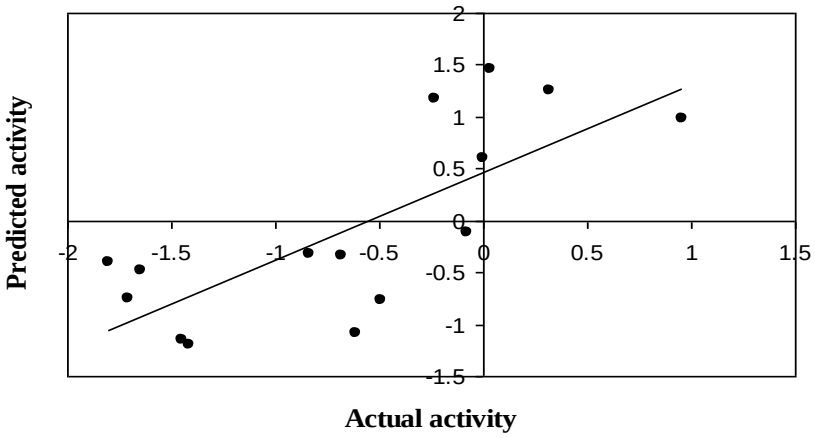


Fig. 4. Graph of observed activity versus predicted activities of test set molecules from GFA analysis, activity expressed as -log ($MIC_{compd}/MIC_{bifonazole}$).

DISCUSSION

Di Santo et al.[10] reported a quantitative pharmacophore model for antifungal azoles. The program Catalyst was applied to develop the pharmacophore model. The model constituted by the coordination feature, two hydrophobics, one ring aromatic and two excluded volumes. The most active molecules were found to match with all the pharmacophore. The authors suggested that the decrease in activities of the other compounds were due to the unmatching of pharmacophoric features. We have carried out 3D- QSAR to further validate the model.

CoMFA

RMS fit takes into account superimposition of the azole moieties as well as the superimposition of the phenyl rings common to all compounds. The CoMFA steric and electrostatic contour maps of RMS fit are shown in Fig. 5 and 6 respectively. The contour plots are to be considered as a representation of the lattice points, where differences in field values are strongly associated with difference in receptor binding affinity. It is likely that all the compounds studied exert same steric and/or electrostatic influence in certain area.

The maps were obtained using the PLS analysis STDEV* COEFF contouring by their contribution and displaying in transparent contour mode. Regions where steric bulk is favored are depicted in green while yellow indicates steric bulk is disfavored (Fig. 5), blue region depicts positive charge favored areas and red regions where negative charge favored areas (Fig. 6).

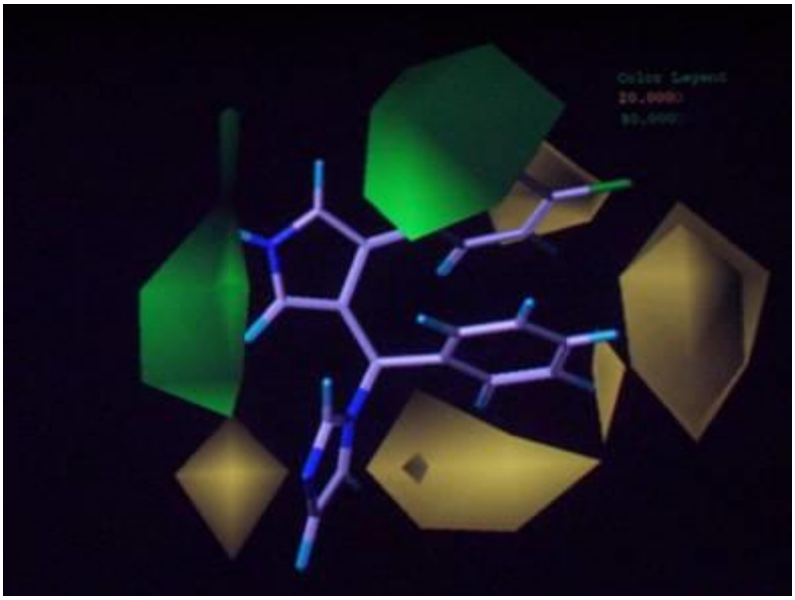


Fig. 5. CoMFA steric STDEV* contour plots from the atom based fit. Sterically favored areas are represented by green polyhedra. Sterically unfavorable areas are represented by yellow polyhedra. The active molecule 07 is shown in capped-sticks.

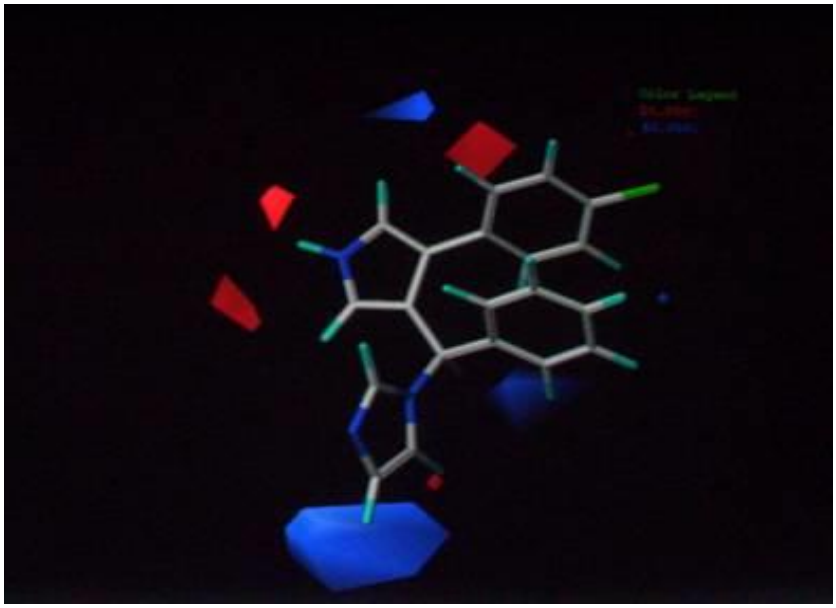


Fig. 6. CoMFA electrostatic STDEV*COEFF contour plots from atom based fit. Positive charge favored areas are represented by blue polyhedra. Negative charge favored areas are represented by red polyhedra. The active molecule 07 is shown in capped-sticks.

Fig. 5 depicts the steric contour plot using CoMFA. The sterically favorable green contour is found around N-substituent of pyrrole. This explains importance of steric interaction of the ligand with the receptor. N-substituted compounds are more active than unsubstituted

analogues. This green contour matched with hydrophobic groups in the pharmacophoric model developed by authors of the original paper. A large green contours found to surround the ortho position of C-4 phenyl ring at pyrrole ring of compound 07. But the para position shows sterically unfavorable yellow contours. This indicates there is a definite substituent requirement for steric interactions with the receptor site. This is further supported by analyzing compounds 27 to 41. Compound 27 to 30 lacking substituent at phenyl ring, were moderately active, while compounds 31 to 41 have bulky substituents at ortho or para positions were also less active. The excluded volume maps in pharmacophore model reported by Di Santo et al[10] also emphasize that this region could not contain any atoms or bonds. Yellow contours are seen near the carbon joining imidazole and pyrrole ring, which indicates that only one carbon atom is required as linkage. If it is replaced by any other linkage like a benzyl linkage the activity reduces. Yellow contours are also seen near second phenyl ring which indicate substitution at this position will reduce activity.

Fig. 6 displays the electrostatic contour plot using CoMFA. The electrostatically favorable red contours are found near N1 of pyrrole and aromatic ring of C-4 position of pyrrole ring. These electrostatically favorable contours are flanked with electrostatically unfavorable blue region. Red contour near N-1 of pyrrole ring shows the need for electron-rich atom for electrostatic interaction with the receptor to show good antifungal activity. Moderate activity of the compound 56 is may be due to improper spatial orientation of phenyl ring towards sterically favorable region. It is observed that compounds having aliphatic substituents are more active.

A red contour near the C-4 phenyl ring at pyrrole explains the need of an electron-withdrawing 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. Compounds 01 to 20, showed highest activity due to the presence of electron-withdrawing group than the unsubstituted aromatic compound 27-30. This explains the need of strong hydrophobic substituent with electron-withdrawing capacity. The electrostatic blue contours are observed near ortho position of second phenyl ring and imidazole moiety a blue electrostatic contour is observed. Both green and red contours are observed near pyrrole, which emphasizes need of a pyrrole ring substituted with hydrophobic as well as an electron rich group for better antifungal activity.

GFA

The antifungal activity of the series of N-substituted imidazole antifungals is thus a function of CHI-1 (Chi indices), ShapeRMS (RMS to shape preference), RadOfGyration (Radius of gyration), DIFFV (Difference volume) and Fo (Common overlap volume). All the descriptors indicate the importance of steric interaction of the ligand with the receptor for good antifungal activity.

Positive correlation of ShapeRMS, shape descriptor indicates there is a strict requirement in terms of molecular similarity with the reference molecule for volume to show antifungal activity. RadOfGyration, a spatial descriptor is positively correlated with the activity. It indicates the importance of spatial orientation of the compound for binding with the receptor. DIFFV and Fo, MSA parameters, indicate that there is a definite structural requirement for receptor binding. Overall steric interactions play a critical role in receptor binding.

In general, the CoMFA models are complimentary to the pharmacophore model developed by the authors of the original paper. It emphasizes importance of the N-substituent of pyrrole for antifungal activity. Overall steric interactions play a critical role in receptor binding.

CONCLUSIONS

The CoMFA method has been applied to a set of antifungal agents active against *C. albicans*. The compounds taken for the study belong to chemically diverse families of imidazole derivative. The CoMFA model explains the structure-activity relationships in 50 training set compounds. It also predicts accurately the biological activity of 15 test set compounds. RMS fit gives the best CoMFA model, which is further validated using various statistical methods. The CoMFA model gives as insight into binding mode of these imidazole antifungals. QSAR models generated using GFA also indicate that structural and spatial parameters are important for activity. This QSAR study emphasizes the importance of N-substituted pyrrole in interaction of these inhibitors with the enzyme.

ACKNOWLEDGMENT

The authors are grateful to Dr. Shivajirao S. Kadam, Vice Chancellor, Bharati Vidyapeeth University, Pune, Maharastra, India for providing the computational facility.

REFERENCES

1. Georgopapadakou, N. H.; Walsh, T. J. Antifungal agents: chemotherapeutic targets and immunologic strategies. Antimicrob. Agents. Chemother, 1996; 40: 279-291.
2. Sheehan, D. J.; Hitchcock, C. A.; Sibley, C. M. Current and emerging azole antifungal agents. Clin. Microbiol. Rev., 1999; 12: 40-79.
3. Asai, K.; Tsuchimori, N.; Okonogi, K.; Perfect, J. R.; Gotoh, O.; Yoshida, Y. Formation of Azole-Resistant Candida albicans by Mutation of Sterol 14-Demethylase P450. Antimicrob. Agents Chemother, 1999; 43: 1163-1169.
4. Odds, F.; Brown, A. J. P.; Gow, N. A. R. Antifungal agents: mechanisms of action. Trends Microbiol, 2003; 11: 272-279.
5. Lupetti, A.; Kelly, R.D., 'Molecular basis of resistance to azole antifungals", Trends In. Molecu. Med., 2002; 8: 76-81.
6. Xiao, L.; Madison, V.; Chau, A. S.; Loebenberg, D.; Palermo, R. E.; McNicholas, P. M. Three-dimensional models of wild-type and mutated forms of cytochrome P450 14alpha-sterol demethylases from Aspergillus fumigatus and Candida albicans provide insights into posaconazole binding. Antimicrob. Agents Chemother, 2004; 48: 568-574.
7. De Groot, M. J.; Ekins, S. Pharmacophore modeling of human cytochrome P450s. Adv. Drug Del. Rev, 2002; 54: 367-383.
8. Tafi, A.; Anastassopoulou, J.; Theophanides, T.; Botta, M.; Corelli, F.; Massa, S.; Artico, M.; Costi, R.; Di Santo, R.; Ragno, R. Molecular modeling of azole antifungal agents active against Candida albicans. 1. A comparative molecular field analysis study. J. Med. Chem, 1996; 39: 1227-1235.
9. Tafi, A.; Costi, R.; Botta, M.; Di Santo, R.; Corelli, F.; Massa, S.; Ciacci, A.; Manetti, F.; Artico, M. New derivatives of 1-[(aryl)[4-aryl-1H-pyrrol-3-yl]methyl]-1H-imidazole, synthesis, anti-candida activity, and quantitative structure-analysis relationship studies. J. Med. Chem, 2002; 45: 2720-2732.
10. Di Santo, R.; Tafi, A.; Costi, R.; Botta, M.; Artico, M.; Corelli, F.; Forte, M.; Caporuscio, F.; Angiolella, L.; Palamara. A. T., J. Antifungal agents. 11. N-substituted derivatives of 1-[(aryl)(4-aryl-1H-pyrrol-3-yl)methyl]-1H-imidazole: synthesis, anti-Candida activity, and QSAR studies. Med. Chem, 2005; 48: 5140-53.
11. Cramer, R. D.; Bunce, J. D.; Patterson, D. E. Crossvalidation, Bootstrapping, and Partial Least Squares Compared with Multiple Regression in Conventional QSAR Studies. Quant. Struct. Act. Relat, 1988; 7: 18-25.

12. Purushottamachar, P.; Kulkarni, V. M. 3D-QSAR of N-myristoyltransferase inhibiting antifungal agents by CoMFA and CoMSIA methods. Bioorg. Med. Chem, 2003; 11: 3487-3497.
13. Gokhale, V. M.; Kulkarni V. M. Comparative Molecular Field Analysis of Fungal Squalene Epoxidase Inhibitors. J. Med. Chem, 1999; 42: 5348-5358.
14. Murthy, V. S.; Kulkarni, V. M. 3D-QSAR CoMFA and CoMSIA on protein tyrosine phosphatase 1B inhibitors. Bioorg. Med. Chem, 2002; 10: 2267-2282.
15. Makhija, M. T.; Kulkarni, V. M. Molecular electrostatic potentials as input for the alignment of HIV-1 integrase inhibitors in 3D QSAR. J. Comput. Aided Mol. Des, 2001; 15: 961-968.
16. Modica M, Santagati M, Russo F, Parotti L, Gioia L D, Selvaggini C, Salmona M & Mennini T, J. [[(Arylpiperazinyl)alkyl]thio]thieno[2,3-d]pyrimidinone derivatives as high-affinity, selective 5-HT1A receptor ligands. Med Chem, 1997; 40: 574-85.
17. Rogers, D.; Hopfinger, A. J. Application of Genetic Function Approximation to Quantitative Structure-Activity Relationships and Structure-Property Relationship. J. Chem. Inf. Comput. Sci, 1994; 34: 854-866.
18. Kawakami, Y. Inoue, A. Kawai, T. Wakita, M. Sugimoto, H. and Hopfinger, A. The rationale for E2020 as a potent acetylcholinesterase inhibitor. Bioorg. Med. Chem, 1996; 4: 1429-46.
19. Kharkar, P. S. Desai, B. Varu, B. Loriya, R. Gaveria, H. Naliapara, Y. Shah, A. and Kulkarni, V. M. Three-dimensional quantitative structure-activity relationship of 1,4-dihydropyridines as antitubercular agents.J. Med. Chem, 2002; 45(22): 4858–4867.
20. SYBYL 6.9.1 molecular modeling software available from Tripose Associates Inc., 1699, St Hanley Road, St Louis, MO 63114-2913, USA.
21. Clark, M., Cramer, R. D., III; Van Opdenbosh, N. Validation of the general purpose tripos 5. 2 force field. J. Comput. Chem, 1989; 982-1012.
22. Wold S, Albano C, Dunn W J, Edlund U, Esbenson K, Gelad P, Hellberg S, Lindburg W & Sjostrom M, In Chemetrics; Int. Ed. Kowalski B Reidel, Dordrecht. The Netherlands, 1984; 17.
23. Dunn W J, Wolds U, Hellberg S & Gasteiger J, Multivariate Structure Activity Relations between data from a biological test and an esemble of structure descriptors: the PLS method. Quant Struct Act Relant Chem Bio, 1984; 3: 31.

24. Wold, S. Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models. Technometrics, 1978; 4: 397-405.
25. Ciubotariu, D.; Deretey, E.; Opera, T. I.; Sulea, T.; Simon, Z.; Kurunczi, L.; Chiriac, A. Multiconformational Minimal Steric Difference. Structure-Acetylcholinesterase Hydrolysis Rates Relations for Acetic Acid Esters. Quant. Struct. Act. Relat, 1993; 12: 367.
26. Computational results obtained using software program Cerius2 version 4.10L, Molecular Simulation Inc., 9685, Scration road, San Diego, CA. 92121, USA.
27. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. J. Am. Chem. Soc, 1992; 114: 10024-10026.
28. Rappe, A. K.; Goddard, W. A. Charge equilibration for molecular dynamics simulations. J. Phys. Chem, 1991; 95: 3358-3363.