Three-Dimensional Quantitative Structure−Activity Relationship of 1,4-Dihydropyridines As Antitubercular Agents

Prashant S. Kharkar,[†] Bhavik Desai,[‡] Harsukh Gaveria,[‡] Bharat Varu,[‡] Rajesh Loriya,[‡] Yogesh Naliapara,[‡] Anamik Shah,[‡] and Vithal M. Kulkarni*,[†]

Pharmaceutical Division, Institute of Chemical Technology, University of Mumbai, Matunga, Mumbai 400 019, India, and Department of Chemistry, Saurashtra University, Rajkot 360 005, India

Three-dimensional quantitative structure−activity relationship (3D QSAR) methods, comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA), were applied on a series of 1,4-dihydropyridines possessing antitubercular activity. The study was performed using 33 compounds, in which 22 molecules were used for the derivation of the 3D QSAR models (training set) and 11 molecules were used to evaluate the predictive ability of the derived models (test set). Superimpositions were performed using three alignment rules: atom-based fitting, SYBYL QSAR rigid body field fit of the steric and electrostatic fields of the molecules, and flexible fitting (multifit). Both methods were analyzed in terms of their predictive abilities and produced comparable results with high internal as well as external predictivities. Steric and electrostatic fields of the inhibitors were found to be relevant descriptors for SAR. Use of lowest unoccupied molecular orbital energies or ClogP as additional descriptors in the QSAR table did not improve the significance of the 3D QSAR models. Both CoMFA and CoMSIA models based on multifit alignment showed better correlative and predictive properties than other models. A QSAR study using genetic function approximation was also performed for the same set of molecules using different types of physicochemical descriptors to deal with cell-based activity data. The QSAR models revealed the importance of spatial properties and conformational flexibility of side chains for antitubercular activity. Inclusion of fractional polar solvent accessible surface area as a descriptor in the model generation resulted in models with significant internal and external predictivities for the same test set molecules, which may support the possible mode of action of these compounds.

Introduction

Mycobacterium tuberculosis, a human pathogen causing tuberculosis (TB), claims more human lives than any other bacterial pathogens.[1−3] About one-third of the world population is infected with *M. tuberculosis*, 10% of which will develop the disease at some point in their lives.[4] The current treatment of active TB is basically a four drug regimen comprising isoniazid (INH), rifampin, pyrazinamide, and ethambutol for a period of at least 6 months. The failure of patients to complete the therapy has led to the emergence of multidrug resistant TB (MDRTB). Moreover, the pandemic of human immunodeficiency virus (HIV), which dramatically increases susceptibility to develop active TB, has exacerbated the situation. The growing number of cases of MDRTB has become such a public health threat that the World Heath Organization (WHO) has declared TB a global public health emergency.[5] There is an urgent need for new chemotherapeutic agents to combat the emergence of the resistance and strategies, which can effectively shorten the duration of chemotherapy.

INH, a well-known antitubercular drug, is believed to kill mycobacteria by inhibiting the biosynthesis of mycolic acids−critical components of the mycobacterial cell wall. The catalase and peroxidase activities are thought to participate in the drug sensitivity mechanism by converting INH in vivo into its biologically active form, which then acts on its intracellular target.[6] In analogy to INH, pyridines substituted with alkylated tetrazoles (designed as lipophilic precursors of isosteres of isonicotinic acid) have been reported to possess antitubercular activity against the $H_{37}Rv$ strain of *M. tuberculosis*.[7] These compounds, after penetration of the mycobacterial cell wall, could be biotransformed by esterases or peroxidase-catalases. They are more active than the unmodified polar isosteres of isonicotinic acid, which may be due to better penetration of these agents into the cell wall of the mycobacteria.

On the basis of these observations, some 4-substituted phenyl-2,6-dimethyl-3,5-bis-N-(substituted phenyl)carbamoyl-1,4-dihydropyridines were synthesized and tested against *M. tuberculosis* $H_{37}Rv$.[8]

To further explore the structural requirements of 1,4-dihydropyridines for the antitubercular activity, two methods of three-dimensional quantitative structure−activity relationship (3D QSAR), comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA), were performed. The CoMFA method of 3D QSAR was introduced in 1988 by Cramer,[9] in which an assumption is made that the interaction between an inhibitor and its molecular target is primarily noncovalent in nature and shape-dependent. Therefore, a QSAR may be derived by

* To whom correspondence should be addressed. Tel: +91-22-4145616. Fax: +91-22-4145614. E-mail: vithal@biogate.com.
† University of Mumbai.
‡ Saurashtra University.

sampling the steric and electrostatic fields surrounding a set of ligands and correlating the differences in these fields to biological activity. CoMFA calculates steric fields using Lennard–Jones potential and electrostatic fields using a Coulombic potential. This approach has been widely accepted[10,11] and exceptionally valuable, although it is not without problems. In particular, both of the potential functions are very steep near the van der Waals surface of the molecule, causing rapid changes in surface descriptions and requiring the use of cutoff values so that calculations are not done inside the molecular surface. In addition, a scaling factor is applied to the steric field, so that both fields can be used in the same PLS (partial least squares) analysis. Finally, changes in orientation of the superimposed molecule set, relative to the calculation grid, can cause significant changes in CoMFA results, again probably due to strict cutoff values. The alignment rule, i.e., molecular conformation and orientation, is one of the most sensitive input areas for CoMFA studies. Several researchers to date have used traditional atom-based alignment methods. Recently, Kulkarni et al. have proposed a new alignment method based on molecular electrostatic potential (MEP).[12] Also, alignment based on the pharmacophore models has been used as an input for CoMFA studies.[13]

The CoMSIA method of 3D QSAR analysis was introduced by Klebe[14] in 1994, in which using a common probe atom, similarity indices are calculated at regularly spaced grid points for the prealigned molecules. For the distance dependence between the probe atom and the molecule atoms, a Guassian function is used. Because of the different shape of the Guassian function, the similarity indices can be calculated at all grid points, both inside and outside the molecular surface. Also, CoMSIA is not sensitive to changes in the orientation of the superimposed molecules in the lattice, and the correlation results obtained by CoMSIA can be graphically interpreted in terms of the field contribution maps allowing physicochemical properties relevant for binding to be easily mapped back onto molecular structures.[15,16]

Although the CoMSIA approach presently considers five different property fields (steric, electrostatic, hydrophobic, and hydrogen bond donor and acceptor fields) to focus on different physicochemical properties, the fields are highly interdependent on each other. Hence, to avoid duplication of fields and for direct comparison with CoMFA, only steric and electrostatic fields have been considered. Furthermore, the hydrophobic field has been treated separately and in combination with the steric and electrostatic fields. Because the hydrogen bond donor and acceptor properties are implicitly considered in electrostatic interactions, they have not been considered here.

Biological activity of these compounds is based upon cell-based assays. In the present study, we have derived QSAR models for these compounds using different molecular descriptors to make better predictions for the cell-based activity data. We have used genetic function approximation (GFA) technique to generate different QSAR models from various descriptors. The GFA technique was used since it generates a population of equations rather than one single equation for correlation between biological activity and physicochemical proper-

ties. GFA, developed by D. Rogers, involved a 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.[17–24]

This is done as follows: (i) An initial population of equations is generated by random choice of descriptors. The fitness of each equation is scored by lack-of-fit (LOF) measure

$$\text{LOF} = \text{LSE}/\{1 - [(c + d \times p)/m]\}^2$$

where LSE is the least squares error, c is the number of basis functions in the models, d is the smoothing parameter that controls the number of terms in the equation, and p is the number of features contained in all terms of the models. (ii) Pairs from the population of the equation are chosen at random and "crossovers" are performed and progeny equations are generated. (iii) The fitness of each progeny equation is assessed by the LOF measure. (iv) If the fitness of the new progeny equation is better, then it is preserved.

A distinct feature of GFA is that it produces a population of models. GFA models provide some useful additional information such as relevance of a particular descriptor in the model and activity prediction. The combination of robust statistical technique GFA coupled with the use of different types of molecular descriptors would result in better prediction of biological activity for antitubercular agents. In this paper, we present the CoMFA, CoMSIA, and QSAR using GFA of the antitubercular 1,4-dihydropyridines.

Materials and Methods

Biological Data. A data set of 33 molecules belonging to the 1,4-dihydropyridine class, synthesized in our laboratory and evaluated for antitubercular activity, was used for this study.[8] The antitubercular activity was determined using the modified BACTEC 460 system[25] against *M. tuberculosis* H_{37}-Rv ATCC 27294. The activity was expressed as percent inhibition (p), which is defined as

$$p = [1 - (\text{GI of test sample/GI of control sample})] \times 100$$

The biological activity data, originally reported as "response at a standard concentration" form, were converted for QSAR purposes to "concentration needed to produce a standard response" form by using eq 1, the "logit transformation".[26]

$$\text{activity} = \log \{\% \text{ activity of standard/(100\%} -$$
$$\% \text{ activity of standard)}\} \quad (1)$$

It is imperative to evaluate the predictivity of the 3D QSAR models generated. The molecules were divided into training set and test set. Selection of the training set and the test set molecules was done by considering the fact that test set molecules represent a range of biological activity similar to that of the training set. Thus, the test set is the true representative of the training set. The structures of the training and test set molecules are given in Tables 1 and 3, respectively.

CoMFA and CoMSIA Analyses. All of the molecular modeling studies, CoMFA, and CoMSIA reported herein were performed on a Silicon Graphics INDY R5000 workstation using SYBYL 6.6 molecular modeling software from Tripos, Inc., St. Louis, MO.[27] All of the compounds were built from fragments in the SYBYL database. Each structure was fully geometry-optimized using the standard Tripos force field[28] with a distance-dependent dielectric function until a root mean square (rms) deviation of 0.001 kcal/mol Å was achieved. The

Table 1. Structures of Molecules Belonging to the Training Set (Ref 8)

compd	R	R′	R2	R3	R4	R5
1	H	Cl	H	NO_2	H	H
2	H	Cl	OH	H	OCH_3	H
3	H	Cl	H	OCH_3	OCH_3	OCH_3
4	H	NO_2	Cl	H	Cl	H
5	H	NO_2	H	H	$N(CH_3)_2$	H
6	H	NO_2	H	H	Cl	H
7	Cl	H	H	H	SCH_3	H
8	Cl	H	Cl	H	H	H
9	Cl	H	H	H	Cl	H
10	Cl	H	Cl	H	Cl	H
11	Cl	H	SH	H	H	H
12	OCH_3	H	H	H	SCH_3	H
13	OCH_3	H	Cl	H	Cl	H
14	OCH_3	H	H	NO_2	H	H
15	OCH_3	H	Cl	H	H	H
16	H	H	H	NO_2	H	H
17	CH_3	CH_3	H	H	SCH_3	H
18	CH_3	CH_3	H	H	H	H
19	CH_3	CH_3	H	H	OCH_3	H
20	H	NO_2	Cl	H	H	H
21	H	NO_2	Cl	H	Cl	H
22	H	NO_2	H	H	H	H

Table 2. Observed vs Calculated Activity Values for Molecules in the Training Set

compd	observed activity	predicted activity		
		CoMFA[a]	CoMSIA[a]	GFA[b]
1	−0.3888	−0.0242	−0.1413	0.0393
2	0.6297	0.8282	0.5897	0.4289
3	1.1949	1.2392	1.0726	1.1815
4	1.3802	1.2136	1.0076	1.1643
5	1.5096	1.3292	1.7435	1.4642
6	0.2126	0.3863	0.5946	0.9305
7	0.5496	0.3780	0.1999	0.3405
8	−0.6585	−0.9159	−0.7518	−0.8080
9	−0.6886	−0.4326	−0.5456	−0.6190
10	−0.9079	−0.5567	−0.5669	−0.3647
11	−0.9542	−0.9428	−0.7899	−0.3478
12	1.1233	1.2046	1.1426	0.8690
13	1.0047	0.4953	0.5593	0.5121
14	0.2498	0.2424	0.6680	1.1209
15	−0.2311	0.2553	0.3570	−0.1396
16	0.4775	0.0902	−0.1135	0.2652
17	0.1580	−0.0439	−0.0769	−0.1259
18	−0.8653	−1.0646	−1.0399	−1.1758
19	−0.1580	−0.0991	−0.1511	−0.2345
20	1.5096	1.5595	1.4332	1.5970
21	1.6901	1.8011	1.6006	1.3101
22	1.1949	1.0887	1.2400	0.6256

[a] Results using alignment III. [b] Predicted biological activity calculated using eq 6 in Table 9.

partial atomic charges required for the electrostatic interaction were computed by a semiempirical molecular orbital method using the MOPAC 6.0 program. The charges were computed using the PM3 Hamiltonian (keywords: 1SCF, MMOK).[29,30]

The conformational search was performed using systematic search protocol. The rotatable bonds in all molecules were searched from 0 to 360° in 10° increments. The minimum energy conformations thus obtained were minimized using the Tripos force field and subsequently used in the analysis.

Table 3. Structures of Molecules in the Test Set (Ref 8)

compd	R	R′	R″	R2	R3	R4
23	H	H	NO_2	H	H	SCH_3
24	H	NO_2	H	H	H	SCH_3
25	H	NO_2	H	Cl	H	H
26	OCH_3	H	H	H	H	$N(CH_3)_2$
27	H	H	NO_2	OH	H	H
28	H	Cl	H	H	H	OH
29	H	H	H	H	H	SCH_3
30	CH_3	CH_3	H	H	NO_2	H
31	Cl	H	H	H	H	OCH_3
32	CH_3	H	CH_3	OH	H	H
33	Cl	H	H	OH	H	H

Figure 1. Atom-based alignment.

Compound **21** (most active compound) was chosen as the template molecule, on which other molecules were aligned.

Alignment Rule. In the present study, three different alignment rules were adopted.

Alignment I. In this, heavy atoms of the 1,4-dihydropyridine nucleus were used for rms fitting onto the corresponding atoms of the template structure (compound **21**). The atoms used for alignment are marked with an asterisk (*) in Figure 1. The alignment of molecules using rms fitting is shown in Figure 8a.

Alignment II. This was carried out by using the SYBYL QSAR rigid body field fit command within SYBYL and using compound **21** as the template structure. Field fit adjusts the geometry of the molecule such that its steric and electrostatic fields match the fields of the template molecule. The field fit alignment of the molecules is shown in Figure 8b.

Alignment III. In this case, alignment of the molecules was carried out by flexible fitting (multifit) of atoms of the molecules to the template structure. This involved energy calculations and fitting onto the template molecule by applying force (force constant = 20 kcal/mol Å) and subsequent energy minimization. The heavy atoms of the 1,4-dihydropyridine nucleus were considered for this alignment (Figure 1). The aligned molecules are shown in Figure 8c.

CoMFA steric and electrostatic interaction fields were calculated at each lattice intersection point of a regularly spaced grid of 2.0 Å. The grid pattern, generated automatically by the SYBYL/CoMFA routine, extended 4.0 Å units in *X, Y,* and *Z* directions beyond the dimensions of each molecule. The steric term, which represents van der Waals (Lennard−Jones, 6−12) interactions, and the Coulombic term, which represents the electrostatic interactions, were calculated using the standard Tripos force field. A distance-dependent dielectric expression $\epsilon = \epsilon_0 R_{ij}$ with $\epsilon_0 = 1.00$ was adopted. An sp3 carbon atom with a van der Waals radius of 1.52 Å and +1.0 charge was used as the probe to calculate the steric and electrostatic fields. Values of the steric and electrostatic fields were truncated at 30 kcal/mol. The electrostatic fields were ignored at the lattice points with maximal steric interactions.

Table 4. Observed vs Predicted Activity Values for Molecules in the Test Set

		predicted activity		
compd	observed activity	CoMFA[a]	CoMSIA[a]	GFA[b]
23	1.6901	2.1747	2.2078	1.8045
24	1.5096	1.2541	1.3834	1.4635
25	1.2787	0.8733	0.8437	0.7450
26	1.1233	1.3892	1.4809	0.8370
27	1.0606	1.2600	1.4481	1.4077
28	0.5754	−0.1829	−0.3523	−0.6351
29	0.3474	0.7904	0.2145	−0.4453
30	0.0521	−0.7104	−0.4597	0.2869
31	−0.3679	−0.1816	0.0485	0.1573
32	−0.4319	−0.8590	−1.0642	−0.8799
33	−1.1949	−0.5632	−0.6004	−0.3561

[a] Results using alignment III. [b] Predicted biological activity calculated using eq 6 in Table 9.

CoMSIA calculates similarity indices at the intersections of a surrounding lattice. The similarity index A_F [31] or a molecule j with I atoms at the grid point q is determined as follows:

$$A_{F,k}{}^q(j) = -\sum_i \omega_{\text{probe},k}\, \omega_{ik}\, e^{-\alpha r_{iq}{}^2}$$

where ω_{ik} is the actual value of the physicochemical property k of atom i; $\omega_{\text{probe},k}$ is the probe atom with charge +1, radius 1 Å, and hydrophobicity = 1; and r_{iq} is the mutual distance between the probe atom at grid point q and atom i of the molecule. Three physicochemical properties k (steric, electrostatic, and hydrophobic) were evaluated, using a common probe atom with 1 Å radius, charge, and hydrophobicity. A Guassian type distance dependence was considered between the grid point q and each atom i of the molecule. The value of the so-called attenuation factor α was set to 0.3. A lattice of 2 Å grid spacing was generated automatically.

PLS Analysis. To obtain a 3D QSAR, PLS[32] fitting was used. The PLS method has been applied successfully in numerous QSAR studies aiming to rationalize those structural features affecting biological activity. PLS regression seeks a relationship between Y and X, where vector Y is the response or dependent variable and X represents the descriptor data.

PLS analyses were performed following the CoMFA standard implementation in SYBYL. The different descriptor blocks have been scaled to each other using the CoMFA standard scaling option. To check statistical significance of the models, cross-validations were done by means of the leave-one-out (LOO) procedure. The results from the cross-validation analysis were expressed as the cross-validated r^2 (r_{cv}^2). It is defined as $r_{cv}^2 = 1 - \text{PRESS}/\sum(Y - Y_{\text{mean}})^2$, where PRESS = $\sum(Y - Y_{\text{pred}})^2$.

The optimum number of components was determined by selecting the smallest s_{press} value. s_{press} is the root mean predictive error sum of squares. It is an expected uncertainty in the prediction for an individual compound based on the data available from other compounds in the set.

$$s_{\text{press}} = (\text{PRESS}/(n - c - 1))^{1/2}$$

where n = number of rows and c = number of components. Usually the smallest s_{press} value corresponds to the highest r_{cv}^2 value. The optimal number of components was subsequently used to derive the final QSAR models. For all conventional analyses (no cross-validation), the "minimum σ" standard deviation threshold was set to 2.0 kcal/mol. The r_{cv}^2, s_{press}, r_{conv}^2, and SE values were computed as defined in SYBYL. SE is the standard error of estimate. It is a measure of the target property uncertainty still unexplained after the QSAR has been derived.

In Table 5, $\text{Pr}^2 = 0$ means the probability of obtaining the observed F ratio value by chance alone, if the target and the explanatory variables themselves are truly uncorrelated. When

Table 5. Summary of Results from CoMFA and CoMSIA Analyses

	alignment I		alignment II		alignment III	
	CoMFA	CoMSIA	CoMFA	CoMSIA	CoMFA	CoMSIA
r_{cv}^2	0.533	0.558	0.547	0.509	0.563	0.562
s_{press}	0.687	0.649	0.657	0.684	0.646	0.646
r_{conv}^2	0.930	0.887	0.956	0.890	0.919	0.884
SE	0.267	0.327	0.204	0.324	0.278	0.330
components	5	4	4	4	4	4
F values	42.326	33.525	92.649	34.390	48.649	32.320
$\text{P}r^2 = 0$	0.00	0.00	0.00	0.00	0.00	0.00
			fraction			
steric	0.583	0.332	0.550	0.330	0.598	0.328
electrostatic	0.417	0.668	0.450	0.670	0.402	0.672
r_{pred}^2	0.716	0.700	0.686	0.660	0.707	0.676
r_{bs}^2 [a]	0.960	0.913	0.948	0.936	0.947	0.919
SD	0.022	0.051	0.021	0.036	0.031	0.038

[a] Results from 100 runs of bootstrapped analysis.

Table 6. Results of Analysis with Randomized Biological Activities and Cross-Validation Using Five Groups

	r_{vb}^2 [a]		r_{cv}^2 [b]	
	CoMFA[c]	CoMSIA[c]	CoMFA[c]	CoMSIA[c]
mean	−0.149	−0.292	0.497	0.430
SD	0.260	0.177	0.202	0.076
high	0.619	−0.010	0.706	0.572
low	−0.650	−0.623	0.132	0.305

[a] Cross-validated r^2 with randomized biological activity average of 100 runs. [b] Cross-validated r^2 using five groups with optimum number of components average of 25 runs. [c] Results of analysis with alignment III.

$\text{Pr}^2 = 0$ is zero, then results are not by chance and are significant. Additionally, to perform an even more rigorous statistical test, several runs of cross-validations using five groups were done in which each target property value is predicted by a model based on about 4/5 or 80% of the available data. To further assess the robustness and statistical confidence of the derived models, bootstrapping analysis (100 runs) was performed.

A common test to check the consistency of the models is to scramble the biological data and repeat the model derivation process, allowing detection of possible chance correlations. After our data set was randomized in several distinct ways, in all cases, we only observed very low or negative r_{cv}^2 values in the PLS analyses (Table 6).

QSAR and GFA. All of the molecular modeling studies were done using Cerius2 (version 3.5) molecular modeling software[33] running on Silicon Graphics O2 R5000 workstation. All of the molecules were imported within Cerius2. Partial charges were assigned using charge equilibration method.[34] All of the molecules were minimized until root mean square deviation 0.001 kcal/mol Å was achieved and used in the study.

Calculation of Molecular Descriptors. Various types of molecular descriptors were calculated for each molecule in the study table using the default settings within Cerius2. A total of 31 descriptors categorized as (i) molecular shape analysis (MSA), (ii) spatial, (iii) electronic, (iv) structural, and (v) thermodynamic were calculated. A complete list of descriptors used in the study is given in Table 8.

MSA Descriptors.[35] MSA descriptors were calculated using the MSA module within Cerius2. The lowest energy conformer of compound **21**, the most active compound, was taken as reference for the calculation of the MSA descriptors.

The QSAR analysis builds models of biological activity using physicochemical properties of a series of compounds. The basic assumption is that the variations of biological activity within a series can be correlated with changes in measured or computed molecular features of the molecules. To fit the equations and thereby generate QSAR models, we have used the recently developed GFA method.

Table 7. Results of CoMSIA Analysis with Hydrophobic, Steric, and Electrostatic Fields Alone and in Combination

	hydrophobic[a]	steric[a]	electrostatic[a]	CoMSIA[a,b]
r^2_{cv}	0.424	0.287	0.489	0.603
s_{press}	0.847	0.947	0.719	0.678
r^2_{conv}	0.908	0.844	0.862	0.939
SE	0.339	0.410	0.374	0.266
components	8	6	5	7
r^2_{pred}	−0.347	0.798	0.905	0.578

[a] Results of analysis with alignment III. [b] CoMSIA results using steric, electrostatic, and hydrophobic fields taken together.

A training set of 22 compounds, which was used for the CoMFA and CoMSIA studies, was used for the QSAR study. GFA was performed using 100 000 crossovers, a smoothness value of 1.0, and other default settings within Cerius2.

A population of 100 equations was generated in each run by using a set of 31 descriptors and linear basis functions. The set of equations generated for each chemical class was evaluated on the following basis: (i) LOF measure, (ii) variable terms in the equation, and (iii) 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.

Predictive r^2 Value. The predictive r^2 was based only on the molecules not included in the training set and is defined as

$$r^2_{pred} = (SD - PRESS)/SD$$

where SD is the sum of the squared deviations between the biological activity of the molecules in the test set and the mean biological activity of the training set molecules and PRESS is the sum of the squared deviations between predicted and actual activity values for every molecule in the test set. Like r^2_{cv}, the predictive r^2 can assume a negative value reflecting a complete lack of predictive ability of the training set for the molecules included in the test set.[36]

Results and Discussion

CoMFA and CoMSIA Analyses. The results of CoMFA and CoMSIA studies are summarized in Table 5. All of the analyses reveal comparable cross-validated r^2 values. CoMFA analysis based on alignment I yielded a correlation with an r^2_{cv} of 0.533 (five principal components) and a conventional r^2 of 0.930. This model displayed a good external predictivity with r^2_{pred} of 0.716. CoMSIA studies using alignment I yielded an r^2_{cv} of 0.558 using four components, and r^2_{conv} was 0.887. This model also exhibited a good predictive ability with r^2_{pred} of 0.7.

Realignment of the compounds using rigid body field fit (alignment II) improved the significance of the CoMFA model with r^2_{cv} of 0.547 (4 pc) and r^2_{conv} of 0.956. The model also revealed good external predictivity ($r^2_{pred} = 0.686$). CoMSIA analysis using alignment II led to a slightly decreased internal predictivity ($r^2_{cv} = 0.509$ with 4 pc). The conventional r^2 for this model was 0.890. The model obtained showed external predictivity with an r^2_{pred} of 0.660.

CoMFA analysis based on flexible fitting (alignment III) produced an internally predictive ($r^2_{cv} = 0.563$ with 4 pc) and statistically significant ($r^2_{conv} = 0.919$) model. A high bootstrapped (100 runs) r^2 of 0.947 ± 0.031 adds a high confidence limit to this analysis. The model exhibited good external predictivity with an r^2_{pred} of 0.707 for compounds in the test set. CoMSIA analysis with alignment III also resulted in a model with best internal predictivity ($r^2_{cv} = 0.562$ with 4 pc) and statistical significance ($r^2_{conv} = 0.884$). The bootstrapped analysis (100 runs) resulted in an r^2_{bs} of 0.919 with a standard deviation of 0.038. The model showed a high predictive r^2 of 0.676.

Table 8. Descriptors Used in the Present Study

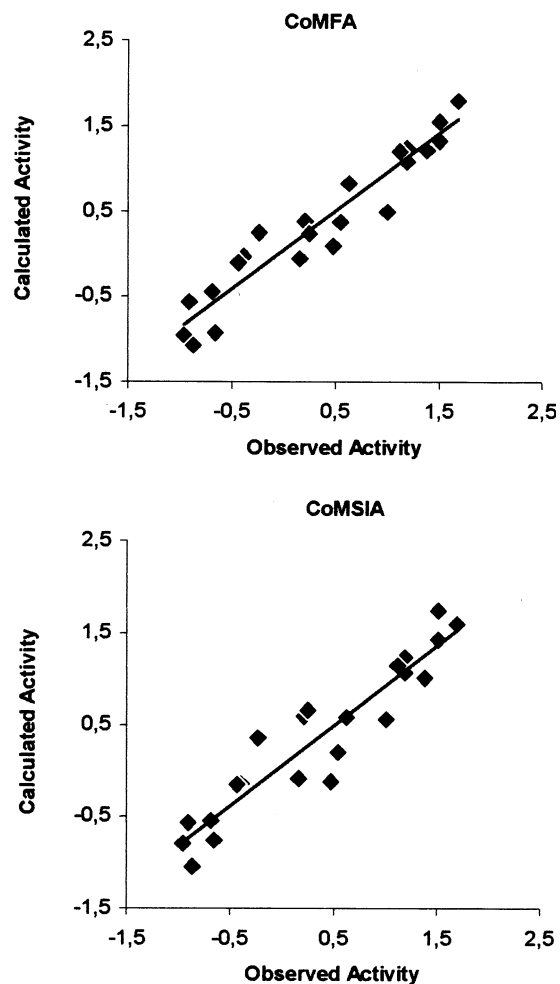| no.	descriptor	type	description
1	DIFFV	MSA	difference volume
2	COSV	MSA	common overlap steric volume
3	Fo	MSA	common overlap volume ratio
4	NCOSV	MSA	noncommon overlap steric volume
5	ShapeRMS	MSA	RMS to shape ref
6	SRVol	MSA	volume of shape ref compd
7	Vm	spatial	molecular volume
8	Area	spatial	molecular surface area
9	Density	spatial	molecular density
10	RadOfGyr	spatial	radius of gyration
11	PMI_mag	spatial	principal moment of inertia
12	PMI_X	spatial	principal moment of inertia X component
13	PMI_Y	spatial	principal moment of inertia Y component
14	PMI_Z	spatial	principal moment of inertia Z component
15	Apol	electronic	sum of atomic polarizabilities
16	Dipole_mag	electronic	dipole moment
17	Dipole_X	electronic	dipole moment, X component
18	Dipole_Y	electronic	dipole moment, Y component
19	Dipole_Z	electronic	dipole moment, Z component
20	HOMO	electronic	highest occupied molecular orbital energy
21	LUMO	electronic	LUMO energy
22	Sr	electronic	superdelocalizability
23	MW	structural	molecular weight
24	RotlBonds	structural	no. of rotatable bonds
25	HbondAcc	structural	no. of hydrogen bond acceptors
26	HbondDon	structural	no. of hydrogen bond donors
27	AlogP	thermodynamic	logarithm of partition coefficient
28	Fh2o	thermodynamic	desolvation free energy for water
29	Foct	thermodynamic	desolvation free energy for octanol
30	Hf	thermodynamic	heat of formation
31	MolRef	thermodynamic	molar refractivity

Figure 2. Calculated vs observed activity from CoMFA and CoMSIA analyses of the training set. Both results are from alignment III.

The LOO cross-validation method might lead to high r_{cv}^2 values, which do not necessarily reflect a general predictiveness of the models. Therefore, cross-validation using five groups was performed. In this method, a model based on about 80% of the available data predicts each target property value. The random formation of the cross-validation groups may have an effect on the results. Hence, cross-validation was performed 25 times for all of the analyses. The results of cross-validation using five groups and randomized activities are reported in Table 6. The mean r_{cv}^2 values were slightly lower as compared to the values obtained in the LOO method. In no case were r_{cv}^2 values negative. The results obtained suggest that there is a good internal consistency in the underlying data set.

The real test for the model predictiveness is to predict the activity of compounds, which were not used in the model generation. To check the external predictivity of the model, we used the test set, which was comprised of 11 compounds. Both the CoMFA and the CoMSIA models exhibited a good predictiveness on these compounds. The observed and calculated activity values for training and test set molecules are given in Tables 2 and 4, respectively. The plots of calculated vs observed activity values for training set molecules and predicted vs observed activity values for the test set molecules are shown in Figures 2 and 3, respectively.
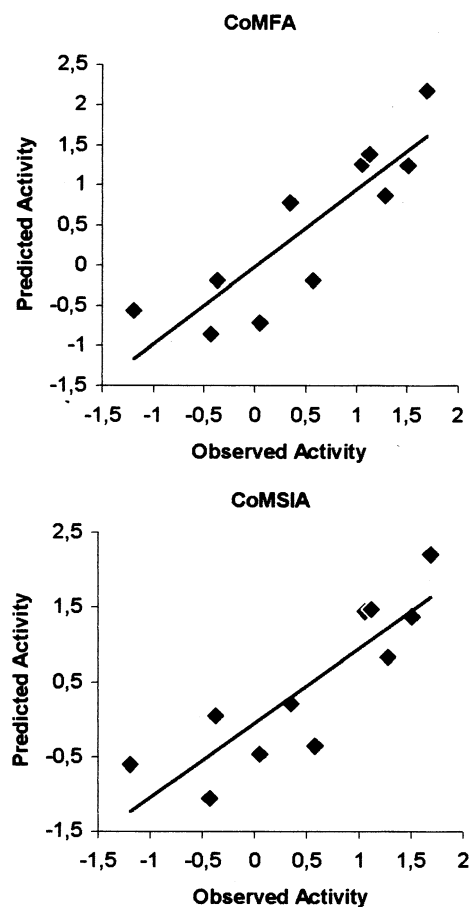
Figure 3. Predicted vs observed activity from CoMFA and CoMSIA analyses of the test set. Both results are from alignment III.

The compounds used in the present study may act as precursors and after penetration into the mycobacterial cell wall may lead to the 3,5-dicarboxylate anions by enzymatic hydrolysis. These anions may then interact with some receptor site to inhibit important biochemical function(s) crucial for the survival of mycobacteria. To consider such interactions in CoMFA and CoMSIA, lowest unoccupied molecular orbital (LUMO) energies (eV) of the compounds were used as a third descriptor. The analysis of the QSAR table shows that LUMO decreased the statistical significance of both the CoMFA and the CoMSIA models (data not shown). Inclusion of ClogP,[37] a physicochemical parameter for lipophilicity, did not improve the significance of the CoMFA model (data not shown).

To further explore whether hydrophobic interactions contribute significantly to the antitubercular activity, hydrophobic fields were computed and correlated alone as well as in combination with the steric and electrostatic fields with the activity. The results of this analysis are summarized in Table 7. Alone with the hydrophobic fields, the model showed a significant r_{cv}^2 of 0.424, but the poor r_{pred}^2 value of -0.342 was obtained. Along with the steric and electrostatic fields, r_{cv}^2 improved slightly but the external predictivity of this model decreased significantly. It is apparent that the steric and electrostatic fields alone are adequate in both CoMFA and CoMSIA to characterize the variance in biological activity of these antitubercular compounds.
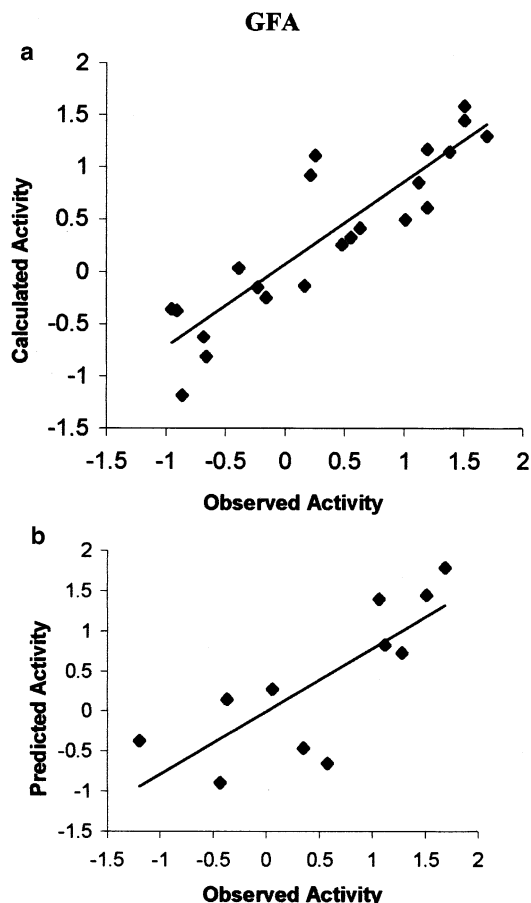
Figure 4. Plots of (a) calculated vs observed activity for the training set and (b) predicted vs observed activity for the test set molecules.

The correlations obtained were not based on chance correlations as can be seen from the cross-validated r^2 values (Table 6). All of these analyses were repeated 100 times with LOO cross-validation each time after randomly interchanging the biological activities (random permutation testing) between the compounds. The means of the 100 runs are reported in Table 6. In all analyses, the mean r^2_{cv} values were negative. This indicates greater than 99.9% likelihood that the relationship using the correct assignment of the biological activities did not arise by chance.

Graphical Interpretation of the Results. To visualize the information content of the derived 3D QSAR model, CoMFA contour maps were generated by interpolating the products between the 3D QSAR coefficients and their associated standard deviations. Figure 5 shows the stereoview of the CoMFA contour map from the analysis based on alignment III using compound **21** (most potent compound of the series) as a reference structure. The green and yellow polyhedra describe regions of space around the molecules where increases in steric bulk, respectively, enhance or diminish antitubercular activity. The red and blue polyhedra describe regions where high electron density (i.e., negative charge or polarity) within the substrate structure enhances or diminishes activity, respectively.

Besides greater robustness and better predictive power, CoMSIA provides significantly improved contour diagrams. They allow the correlation results to be mapped back onto the molecular structures. The coef-
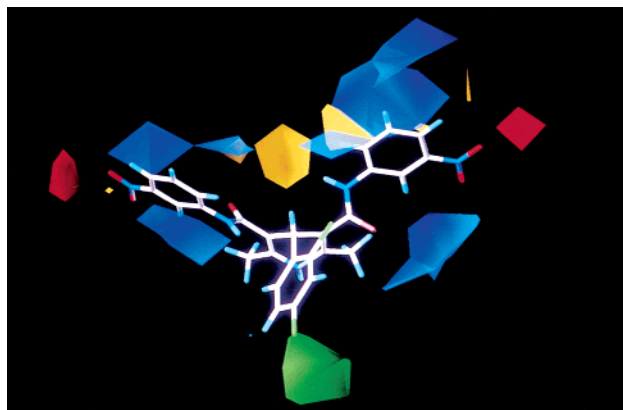
Figure 5. CoMFA STDEV*COEFF contour plot from analysis based on alignment III. Sterically favored areas (contribution level of 80%) are represented by green polyhedra. Sterically disfavored areas (contribution level of 20%) are represented by yellow polyhedra. Negative charge-favored areas (contribution level of 80%) are represented by blue polyhedra. Negative charge-disfavored areas (contribution level of 20%) are represented by red polyhedra.

ficient contour maps using the field type "STDEV* COEFF" were generated. The steric contribution contour plots from the analysis based on alignment III are plotted in Figure 6. In these figures, the green contours represent regions of high steric tolerance (80% contribution), while the yellow contours represent regions of low steric bulk tolerance (20% contribution). The yellow region in the interior of the molecule arises from the fact that CoMSIA fields are calculated inside as well as outside the molecular surface. The regions of molecules inside its van der Waals radius are not involved in drug−receptor interaction. The presence of sterically unfavored yellow contours near R has a negative effect on the activity as shown by compound **11** (Figure 6). The contour map also shows a large green contour in the vicinity of R_4. The occupation of this area by a bulky group will have a positive effect on the biological activity as represented by compound **21** (Figure 6), which has more activity than compound **22** due to the presence of a bulky chloro group.

The electrostatic contour plots are shown in Figure 7. In these figures, the red contours represent regions of decreased tolerance for positive charge (20% contribution), while the blue contours represent regions of increased tolerance for positive charge (80% contribution). A predominant feature of the electrostatic plot is the presence of a blue contour surrounding R', i.e., 3- and 4-positions of the 3,5-disubstited carbamoylphenyl moiety. Therefore, a low electron density in this area will have a positive effect on the biological activity. Compounds containing electron-withdrawing substituents at these positions show more activity (compound **21**, Figure 7). The red contour present near the 4-position of the 4-phenyl ring indicates the requirement of increased electron density in this area. The blue contours present near the 3,5-dicarbamoylphenyl rings coincide with the sterically unfavorable yellow contours. Therefore, the substituents present at these positions should be electron-withdrawing but with limited steric bulk. The steric and electrostatic contour plots of compound **21** and compound **11** are shown for reference.

Table 9. QSAR Equations Generated Using GFA for the Training Set of Molecules

no.	eq	LOF	r^2	F value	r^2_{cv}	r^2_{pred}
1	activity = 6.2764 − 10.5279 (Density) + 0.4485 (HbondAcc) + 0.004036 (PMI_X)	0.310	0.777	20.924	0.657	0.424
2	activity = −5.3487 + 0.5432 (RotlBonds) + 0.02970 (Foct) + 0.003497 (PMI_X)	0.321	0.769	19.986	0.657	0.551
3	activity = −5.6129 + 0.03111 (Fh2o) + 0.5679 (RotlBonds) + 0.003343 (PMI_X)	0.323	0.768	19.824	0.654	0.549
4	activity = −4.2602 + 0.2817 (HbondAcc) + 0.3468 (RotlBonds) − 0.1202 (Dipole_mag)	0.333	0.761	19.095	0.677	0.544
5	activity = −5.05404 − 0.1103 (Dipole_Z) + 9.6037 (FPSA_P) + 0.008149 (PMI_X)	0.223	0.840	31.456	0.781	0.061
6	activity = −5.7508 + 5.2509 (FPS_PA) + 0.003547 (PMI_X) + 0.4268 (RotlBonds)	0.284	0.796	23.444	0.691	0.560
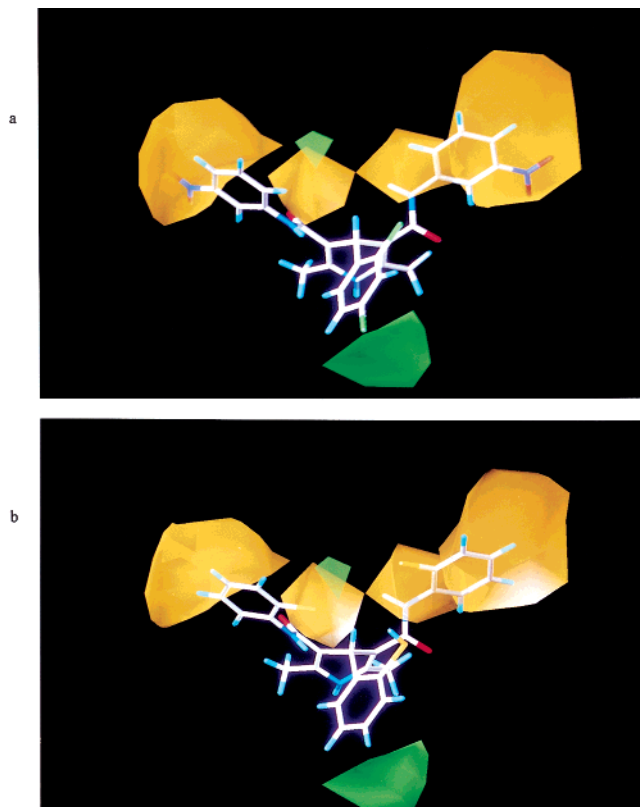
Figure 6. Steric maps from the CoMSIA model using alignment III. (a) Compound **21** shown inside the field. (b) Compound **11** shown inside the field. Green contours (contribution level of 80%) represent areas where steric bulk will enhance activity, and yellow contours (contribution level of 20%) highlight areas, which should be kept unoccupied for increased activity.

Figure 7. Electrostatic maps from the CoMSIA model using alignment III. (a) Compound **21** shown inside the field. (b) Compound **11** shown inside the field. Blue contours (contribution level of 80%) represent regions where an increase of positive charge will enhance activity, and red contours (contribution level of 20%) highlight areas where more negative charge is favored.

GFA. The QSAR models generated using GFA technique are shown in Table 9. Initially, different sets of equations were generated by altering the chain length of the equations. The generated equations were evolved by repeating the GFA runs to check the stability of the GFA models. The final sets of GFA models were analyzed statistically to select the best model. Observation of the variable usage graph indicated that the terms RotlBonds, PMI_X, HbondAcc, and Density contribute more significantly than all other descriptors. The best equation from the set of equations was selected on the basis of predictivity, variables, and LOF value. Equations 1−4 (Table 9) show more or less similar internal predictivity, but eq 2 shows better external predictivity for test set molecules. RotlBonds, PMI_X, and Foct contribute to this equation and explain about 75% variance in the activity. These variables show low correlation among themselves indicating low probability of chance correlation. Equation 2 with a better-predicted r^2 value of 0.551 describes the QSAR model for antitubercular activity of these compounds. The equation has a spatial parameter, PMI_X, structural parameter, RotlBonds, and thermodynamic parameter, Foct, which contribute significantly to the activity. PMI_X, principal moment of inertia X component, represents the orientation or distribution of the total molecular mass in the 3D space. It reflects that the conformation of molecules is important for activity. Because PMI_X shows positive correlation with activity, it indicates that the proper spatial orientation of the carbamoylphenyl groups is important for interaction with some receptor site as these groups lie along the *X*-axis. Foct is 1-octanol desolvation free energy, and positive correlation of this term with activity indicates that increase in the Foct value would result in increase in activity. RotlBonds is the number of rotatable bonds present in the molecule. It is positively correlated with the activity; the higher the number of rotatable bonds, the higher is the activity. In compounds with higher activity groups such as $N(CH_3)_2$, OCH_3, SCH_3, etc., which increase the number of rotatable bonds in the molecule, are present at
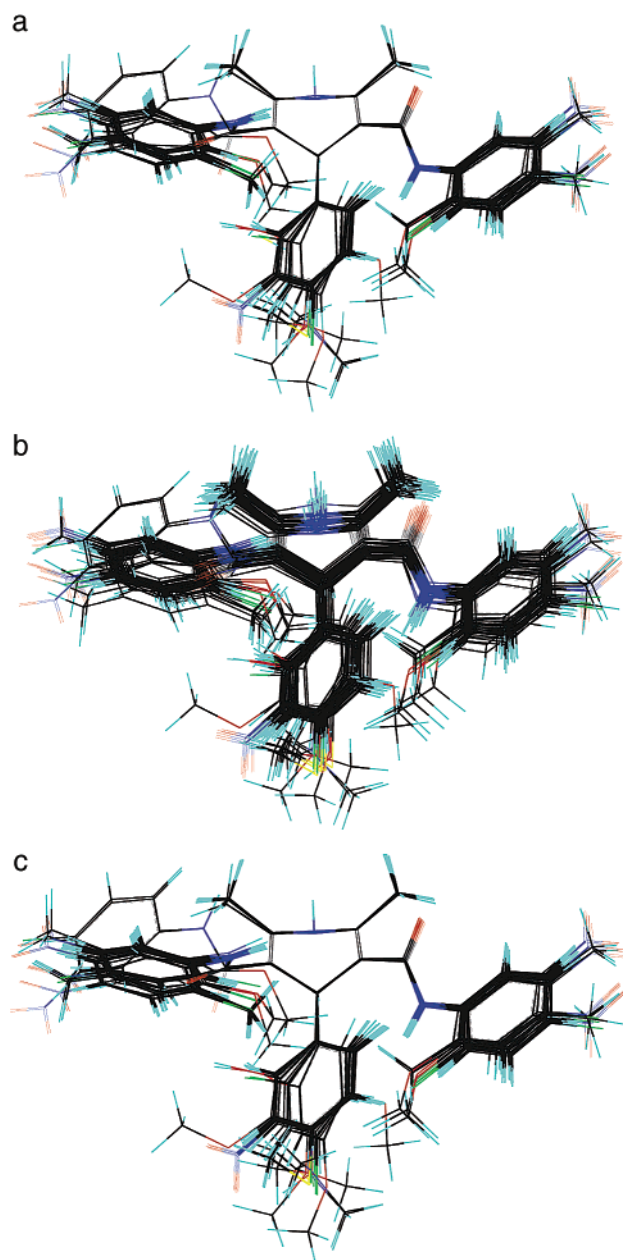
Figure 8. Superposition of molecules (training and test sets) using (a) RMS fitting (alignment I), (b) field fit (alignment II), and (c) flexible fitting (multifit) (alignment III).

position 4 of the 4-substituted phenyl ring, which may act as H-bond acceptor groups.

On the putative mode of action of these agents, we reasoned that the rate of amide bond hydrolysis might be the major factor affecting the antitubercular activity. The CoMFA and CoMSIA studies indicated that the presence of a substituent at position 2 of the carbamoyl phenyl ring is sterically unfavorable for activity. This might be related to the reduction in the polar solvent accessible surface area around the amide bond. To confirm this hypothesis, we have used the fraction of polar solvent accessible surface area to the total solvent accessible surface area (FASA_P) as one of the descriptors.[38] When this descriptor was used along with others in the GFA run, the resulting variable usage graph indicated the dominant contribution of FASA_P. The eqs 5 and 6 (Table 9) indicate the two best models. Among

these, eq 5, although it exhibited high internal predictivity, lacked external predictivity. Equation 6 with high r^2_{cv} and r^2_{pred} values coupled with low LOF is proposed as a model that best describes the antitubercular activity of this series of compounds, and the terms RotlBonds, PMI_X, and FASA_P explain about 80% variation in biological activity. This equation contains a parameter FASA_P, which is a fraction of polar solvent accessible surface area to the total solvent accessible surface area. It shows positive correlation with activity indicating that the higher the value of this parameter, the higher is the activity. The presence of groups such as Cl and CH_3 at position 2 of the substituted phenylcarbamoyl groups decreases the solvent accessible surface area around the amide bond, which may reduce the rate of hydrolysis of the amide bond—probably a crucial step for the antitubercular activity of these compounds. This is supported by the observation that the presence of bulky groups at these positions is detrimental for activity. The observed and calculated biological activities of the training set and test set molecules are given in Tables 2 and 4, respectively. The plots of (i) calculated and observed and (ii) predicted and observed biological activities for the training and test sets, respectively, are shown in Figure 4.

Conclusions

Two 3D QSAR methods, i.e., CoMFA and CoMSIA, were applied to rationalize the antitubercular activity of a set of 33 molecules belonging to the 1,4-dihydropyridine class. The 3D QSAR models obtained using three alignment rules showed a high correlative and predictive ability. A high bootstrapped r^2 value and small standard deviations indicate that a similar relationship exists in all compounds. Inclusion of LUMO energies that reflect the donor–acceptor interactions or ClogP, a lipophilicity parameter, did not improve the significance of both the CoMFA and the CoMSIA models. The results show good correlations between steric and electrostatic fields and antitubercular activity. The contour diagrams obtained for the CoMSIA field contributions can be mapped back onto the structural features accounting for the activity trends in the series.

A QSAR study was performed for the same set of molecules using different types of physicochemical descriptors. The QSAR models generated using GFA technique revealed that spatial properties and conformational flexibility of side chains play important roles for antitubercular activity. Inclusion of fractional polar solvent accessible surface area as a descriptor in the model generation resulted in models with significant statistics, which may support the possible mode of action of these compounds. On the basis of the results obtained from these studies, novel molecules can be designed to possess better biological activity.

References

(1) Nakajima, H. Tuberculosis: A Global Emergency. *World Health* **1993**, *46*, 3.

(2) Dolin, P. J.; Raviglione, M. D.; Kochi, A. Global Tuberculosis Incidence and Mortality during 1990–2000. *Bull. WHO* **1994**, *72*, 213–220.

(3) Rouchi, A. M. Tuberculosis: A Tough Adversary. *Chem. Eng. News* **1999**, *77*, 52–69.

(4) Nunn, P.; Kochi, A. A Deadly Duo—TB and AIDS. *World Health* **1999**, *46*, 7–8.

(5) Raviglione, M. D.; Snider, D. E., Jr.; Kochi, A. Global Epidemiology of Tuberculosis: Morbidity and Mortality of a Worldwide Epidemic. *JAMA* **1995**, *273*, 220–226.

(6) Johnsson, K.; Schultz, P. G. Mechanistic Studies of the Oxidation of Isoniazid by Catalase Peroxidase from *Mycobacterium tuberculosis*. *J. Am. Chem. Soc.* **1994**, *116*, 7425–7426.

(7) Wächter, G. A.; Davis, M. C.; Martin, A. R.; Franzblau, S. G. Antimycobacterial Activity of Substituted Isosteres of Pyridine- and Pyrazinecarboxylic Acids. *J. Med. Chem.* **1998**, *41*, 2436–2438.

(8) Desai, B.; Sureja, D.; Naliapara, Y.; Shah, A.; Saxena, A. Synthesis and QSAR Studies of 4-Substituted phenyl-2,6-dimethyl-3,5-bis-N-(substituted phenyl)-carbamoyl-1,4-dihydropyridines as Potential Antitubercular Agents. *Bioorg. Med. Chem.* **2001**, *4*, 1993–1997.

(9) Cramer, R. D., III.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.

(10) Kulkarni, S. S.; Kulkarni, V. M. Three-Dimensional Quantitative Structure–Activity Relationship of Inteleukin 1-β Converting Enzyme Inhibitors: A Comparative Molecular Field Analysis Study. *J. Med. Chem.* **1999**, *42*, 373–380.

(11) Gokhale, V. M.; Kulkarni, V. M. Comparative Molecular Field Analysis of Fungal Squalene Epoxidase Inhibitors. *J. Med. Chem.* **1999**, *42*, 5348–5358.

(12) 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*, 961–978.

(13) Talele, T. T.; Kulkarni, S. S.; Kulkarni, V. M. Development of Pharmacophore Alignment Models as Input for Comparative Molecular Field Analysis of a Diverse Set of Azole Antifungal Agents. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 958–966.

(14) Klebe, G.; Abraham, U.; Mietzner, T. Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules to Correlate and Predict Their Biological Activity. *J. Med. Chem.* **1994**, *37*, 4130–4146.

(15) Bohm, M.; Sturzebecher, J.; Klebe, G. Three-Dimensional Quantitative Stucture–Activity Relationship Analyses using Comparative Molecular Field Analysis and Comparative Molecular Similarity Indices Analysis to Elucidate Selectivity Differences of Inhibitors Binding to Trypsin, Thrombin, and Factor Xa. *J. Med. Chem.* **1999**, *42*, 458–477.

(16) Klebe, G. Comparative Molecular Similarity Indices Analysis: CoMSIA. *Perspect. Drug Discovery Des.* **1998**, *12*, 87–104.

(17) Rogers, D.; Hopfinger, A. J. Application of Genetic Function Approximation to Quantitative Structure–Activity Relationships and Quantitative Structure–Property Relationships. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 854–866.

(18) Shi, L. M.; Fan, Y.; Myers, T. G.; O'Connor, P. M.; Paull, K. D.; Friend, S. H.; Weinstein, J. N. Mining the NCI Anticancer Drug Discovery Databases: Genetic Function Approximation for the QSAR Study of Anticancer Ellipticine Analogues. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 189–199.

(19) Venkatarangan, P.; Hopfinger, A. J. Prediction of Ligand–Receptor Binding Thermodynamics by Free Energy Force Field Three-Dimensional Quantitative Structure–Activity Relationship Analysis: Applications to a Set of Glucose Analogue Inhibitors of Glycogen Phosphorylase. *J. Med. Chem.* **1999**, *42*, 2169–2179.

(20) Tokarski, J. S.; Hopfinger, A. J. Constructing Protein Models for Ligand-Receptor Binding Thermodynamic Simulation: An Application to a Set of Peptidometic Renin Inhibitors. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 779–791.

(21) Tokarski, J. S.; Hopfinger, A. J. Prediction of Ligand–Receptor Binding Thermodynamics by Free Energy Force Field (FEFF) 3D-QSAR Analysis: Application to a Set of Peptidometic Renin Inhibitors. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 792–811.

(22) Hahn, M.; Rogers, D. Receptor Surface Models. 2. Application to Quantitative Structure–Activity Relationship Studies. *J. Med. Chem.* **1995**, *38*, 2091–2102.

(23) Gokhale, V. M.; Kulkarni, V. M. Understanding the Antifungal Activity of Terbinafine Analogues using Quantitative Structure–Activity Relationship Models. *Bioorg. Med. Chem.* **2000**, *8*, 2487–2499.

(24) Karki, R. G.; Kulkarni, V. M. Three-Dimensional Quantitative Relationship (3D-QSAR) Study of 3-Aryloxazolidin-2-one Antibacterials. *Bioorg. Med. Chem.* **2001**, *9*, 3153–3160.

(25) Collins, L. A.; Franzblau, S. G. Microplate Alamar Blue Assay versus BACTEC 460 System for High Throughput Screening of Compounds Against *Mycobacterium tuberculosis* and *Mycobacterium avium*. *Antimicrob. Agents Chemother.* **1997**, *41*, 1004–1009.

(26) Björkroth, J.-P.; Pakkanen, T. A.; Lindroos, J. Comparative Molecular Field Analysis of Some Clodronic Acid Esters. *J. Med. Chem.* **1991**, *34*, 2338–2343.

(27) *SYBYL Molecular Modeling System*, Version 6.6; Tripos, Inc.: St. Louis, MO, 63144-2913.

(28) Clark, M.; Cramer, R. D., III.; Van Opdenbosh, N. Validation of the General-Purpose Tripos 5.2 Force Field. *J. Comput. Chem.* **1989**, *10*, 982–1012.

(29) MOPAC 6.0 is available from Quantum Chemistry Program Exchange, Indiana University.

(30) Stewart, J. J. P. MOPAC: A Semiempirical Molecular Orbital Program. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1–103.

(31) Kearsley, S. K.; Smith, G. M. An Alternative Method for the Alignment of Molecular Structures: Maximizing Electrostatic and Steric Overlap. *Tetrahedron Comput. Methodol.* **1990**, *3*, 615–633.

(32) Dunn, W. J.; Wold, S.; Edelund, U.; Helberg, S. Multivariate Structure–Activity Relationships between Data from a Battery of Biological Tests and an Ensemble of Structure Descriptors: The PLS Methodol. *Quant. Struct.-Act. Relat.* **1984**, *3*, 131–137.

(33) Cerius2 version 3.5 is available from Molecular Simulations Inc., 9685 Scranton Road, San Diego, CA, 92121.

(34) Rappe, A. K.; Goddard, W. A., III. Charge Equilibration for Molecular Dynamics Simulations. *J. Phys. Chem.* **1991**, *95*, 3358–3363.

(35) Hopfinger, A. J. A Quantitative Structure–Activity Relationship Investigation of Dihydrofolate Reductase Inhibition by Baker Triazines Based upon Molecular Shape Analysis. *J. Am. Chem. Soc.* **1980**, *102*, 7196–7206.

(36) Cramer, R. D., III; 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.

(37) *ClogP*, version 1.0.0; Biobyte Corp.: 201 West, 4th St., Suite 204, Claremont, CA 91711.

(38) FASA_P was calculated using QuaSAR_Descriptors Functionality in Molecular Operating Environment (MOE) 2001.01 available from Chemical Computing Group Inc., 1010 Sherbrooke Street West, Suite 910, Montreal, Canada H3A 2R7.