Med Chem Res (2012) 21:3021–3034
DOI 10.1007/s00044-011-9837-4

MEDICINAL CHEMISTRY RESEARCH

ORIGINAL RESEARCH

2D and 3D QSAR study on amino nicotinic acid and isonicotinic acid derivatives as potential inhibitors of dihydroorotate dehydrogenase (DHODH)

Vivek K. Vyas · Manjunath Ghate

Received: 12 April 2011 / Accepted: 25 October 2011 / Published online: 6 November 2011
© Springer Science+Business Media, LLC 2011

Abstract Dihydroorotate dehydrogenase (DHODH) is a central enzyme of pyrimidine biosynthesis which catalyzes oxidation of dihydroorotate (DHO) to orotate (ORO). DHODH inhibitors are considered as promising targets for the development of antiproliferative, antiparasitic, and immunosuppressive drugs. The best 2D QSAR model for the prediction of human dihydroorotate dehydrogenase (*h*DHODH) inhibitory activity was obtained by applying MLR giving $r^2 = 0.834$ and $q^2 = 0.756$, PCR giving $r^2 = 0.833$ and $q^2 = 0.756$, and PLS giving $r^2 = 0.864$ and $q^2 = 0.786$. 2D QSAR studies reveal the importance of alignment independent descriptors for predicting *h*DHODH inhibitory activity. The 3D QSAR study was performed by comparative molecular field analysis (CoMFA) to rationalize the structural requirements responsible for the inhibitory activity of these compounds. The best CoMFA model obtained for the training set was statistically significant with cross-validated coefficients (q^2) of 0.630 and conventional coefficients (r^2) of 0.949. The CoMFA contour maps suggest some important structural features-like electronegative substituents at biphenyl ring system for strong inhibitory activity. We believe that these results are helpful in design of more potent and selective *h*DHODH inhibitors.

Keywords Dihydroorotate dehydrogenase (DHODH) · 2D QSAR · CoMFA · Amino nicotinic acid and isonicotinic acid derivatives · *h*DHODH inhibitors

V. K. Vyas (✉) · M. Ghate
Department of Pharmaceutical Chemistry, Institute of Pharmacy, Nirma University, Ahmedabad 382 481, Gujarat, India
e-mail: vicky_1744@yahoo.com; vivekvyas@nirmauni.ac.in

Introduction

Flavoenzyme dihydroorotate dehydrogenase (DHODH) [EC 1.3.99.11] (Liu *et al.*, 2004) is a fourth enzyme of pyrimidine de novo synthesis which catalyses oxidation of intermediate dihydroorotate (DHO) to orotate (ORO) (Fig. 1). Pyrimidines are required for the biosynthesis of DNA, RNA, glycoproteins and phospholipids (Jones, 1980). DHODH catalyzes synthesis of pyrimidines which are necessary for cell growth and proliferation of rapidly growing cells. Requirement of pyrimidine nucleotides are depended on cell type and developmental stage, involvement of de novo pathway is small in resting or fully differentiated cells where cells acquire pyrimidine mainly by the salvage pathways (Mascia *et al.,* 2000). DHODH enzymes are divided into two families based upon their localization, amino acid sequence, substrate/cofactor dependence, and cellular localization (Norager *et al.*, 2002; Bjornberg *et al.,* 1999). Enzymes belongs to family-1 located in the cytosol, electron acceptors involved in second half reaction of redox process are either fumarate or NAD^+ whereas family-2 enzymes transfer electrons to ubiquinone (CoQ), to which *h*DHODH belongs (Bjornberg *et al.,* 1997). DHODH inhibitors blocks the growth of fast proliferating cell whereas cells which grow at normal speed can meet the requirement of pyrimidine bases from normal metabolic cycle. Inhibitors of *h*DHODH have proven efficacy for the treatment of cancer (Shawver *et al.,* 1997; Baumann *et al.,* 2009) and immunological disorders, such as rheumatoid arthritis and multiple sclerosis (Chen *et al.,* 1986; Herrmann *et al.,* 2004; Merrill *et al.,* 2009). Brequinar (Shannon *et al.,* 1999) and leflunomide (Fox *et al.,* 1999; Rozman, 1998) (Fig. 2) are two examples of such compounds. Brequinar is an antitumor and immunosuppressive agent, while leflunomide, which is a prodrug of the active metabolite A771 726 (Williamson *et al.,* 1996), shows immunosuppressive activity.

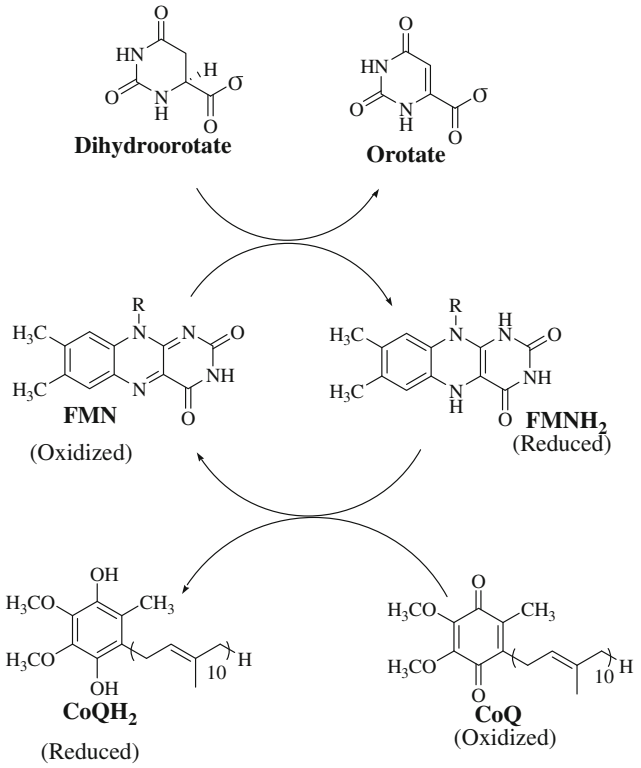


Fig. 1 Reactions catalyzed by DHODH

We have recently compiled the literature pertaining recent advancements in the discovery and development of DHODH inhibitors (Vyas and Ghate, 2011). In search for selective *h*DHODH inhibitors, we have attempted 2D and 3D QSAR (CoMFA) studies on amino nicotinic acid and isonicotinic acid derivatives. QSAR models generated in this work can provide useful information for the design of new compounds with better *h*DHODH inhibitory activity.

Material and method

2D QSAR was performed using Vlife MDS QSAR plus software and 3D QSAR (CoMFA) was performed using the SYBYL × 1.2 software from Tripos Inc., St. Louis, MO, USA on a HP computer with Core2 Duo processor and a window XP operating system.

2D QSAR modeling and data set

The *h*DHODH inhibitory activity data IC_{50} (μM) was taken from the published work of Castro Palomino Laria et al. (Castro *et al.,* 2010). The negative logarithm of the measured IC_{50} (μM) values were converted to pIC_{50} and subsequently used as the dependent variable for QSAR study. Compounds were sketched using the 2D draw application and converted to 3D structures. Energy minimization and geometry optimization were conducted using Merck molecular force field (MMFF) and atomic charges, maximum number of cycles were 1000, convergence criteria (RMS gradient) was 0.01 and medium's dielectric constant of 1 by batch energy minimization method. Conformational search was carried out by a systemic conformational search method. Energy minimized geometry was used for calculation of descriptors, a total of 208, 2D descriptors were calculated which encoded different aspects of molecular structure and consists of electronic, thermodynamic, spatial and structural descriptors, e.g., retention index (chi), atomic valence connectivity index (chiV), path count, chain path count, cluster, path cluster, element count, estate number, semi-empirical, molecular weight, molecular refractivity, logP, and topological index. Various alignment-independent (AI) descriptors were also calculated.

Selection of training and test set

Dataset of 26 molecules (Table 1) was divided into training (22) and test (4) set compounds. Selection of the training set and the test set molecules was done manually 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

Fig. 2 DHODH inhibitors

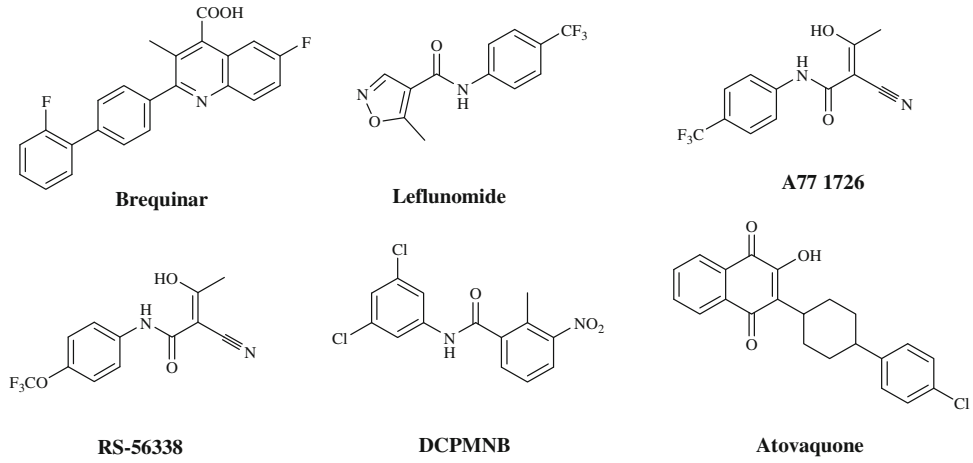


Table 1 Structure, experimental, and predicted activity with residual and 2D descriptors of amino nicotinic acid and isonicotinic acid derivatives

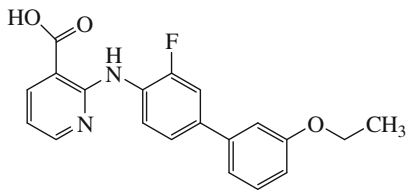
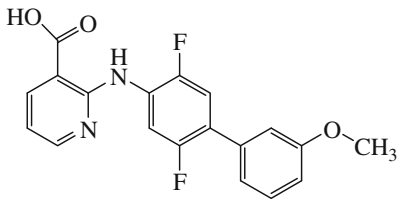
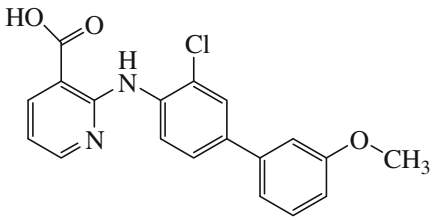
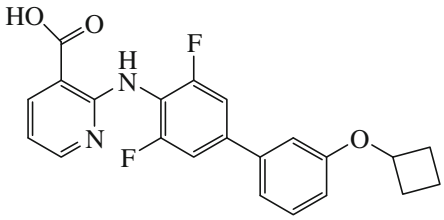
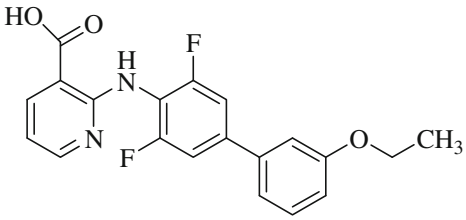
Comp. no.	Structure	IC_{50} (μm)	pIC_{50}	MLR		PCR		PLS		T_N_F_5	4path ClusterCount	T_C_C_6
				[a]Pred	Res	[b]Pred	Res	[c]Pred	Res			
1		200	6.699	6.913	−0.214	6.902	−0.203	6.911	−0.212	0.0	29.0	16.0
2[t]		88	7.056	7.721	−0.665	7.193	−0.137	7.187	−0.131	2.0	36.0	22.0
3		150	6.824	6.689	0.135	6.706	0.118	6.692	0.132	2.0	33.0	18.0
4		90	7.046	7.329	−0.283	7.321	−0.275	7.327	−0.281	1.0	29.0	18.0
5		19	7.721	7.609	0.112	7.589	0.132	7.606	0.115	2.0	33.0	16.0

Table 1 continued

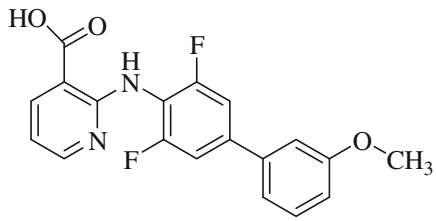
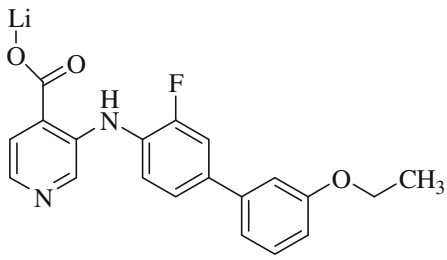
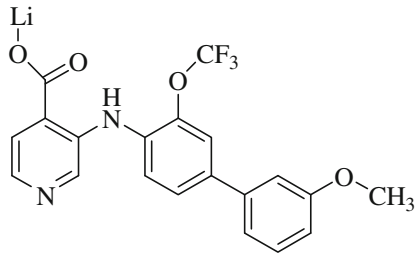
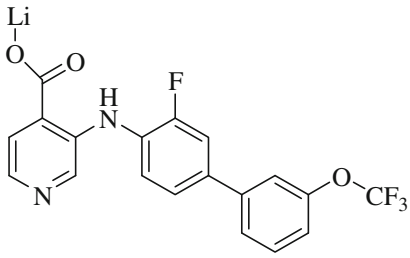
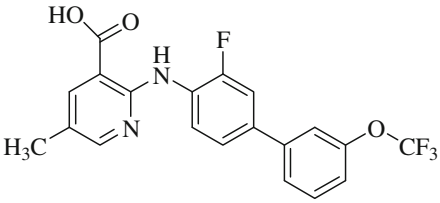
Comp. no.	Structure	IC_{50} (μm)	pIC_{50}	MLR		PCR		PLS		T_N_F_5	4path ClusterCount	T_C_C_6
				[a]Pred	Res	[b]Pred	Res	[c]Pred	Res			
6		15	7.824	7.84	−0.016	7.821	0.003	7.837	−0.013	0.0	30.0	17.0
7[t]		19	7.721	7.634	0.087	7.655	0.066	7.638	0.083	3.0	31.0	19.0
8		14	7.854	7.828	0.026	7.772	0.082	7.817	0.037	0.0	30.0	15.0
9		200	6.699	6.865	−0.166	6.887	−0.188	6.869	−0.17	1.0	31.0	17.0
10		110	6.959	7.149	−0.19	7.148	−0.189	7.149	−0.19	1.0	31.0	17.0

Table 1 continued

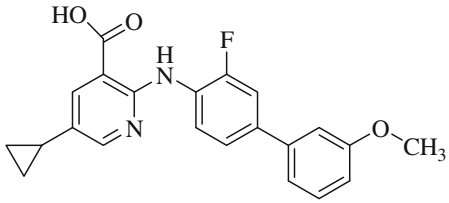
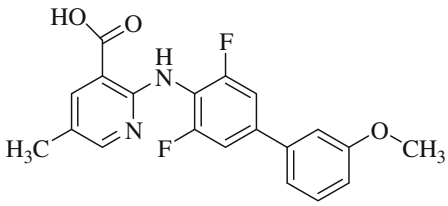
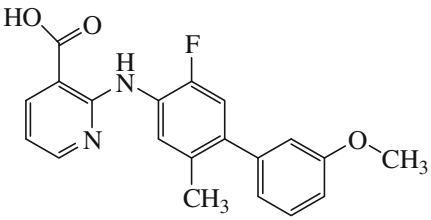
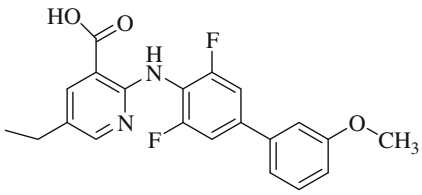
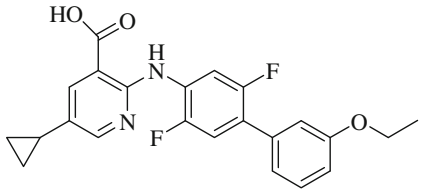
Comp. no.	Structure	IC_{50} (μm)	pIC_{50}	MLR		PCR		PLS		T_N_F_5	4path ClusterCount	T_C_C_6
				[a]Pred	Res	[b]Pred	Res	[c]Pred	Res			
11		33	7.482	7.149	0.333	7.348	0.134	7.349	0.133	1.0	37.0	17.0
12[t]		12	7.921	7.846	0.075	7.736	0.185	7.517	0.404	2.0	35.0	17.0
13		99	7.004	7.386	−0.382	7.834	−0.83	7.844	−0.84	2.0	35.0	17.0
14		12	7.921	8.209	−0.288	7.834	0.087	7.817	0.104	2.0	36.0	17.0
15		23	7.638	7.906	−0.268	7.899	−0.261	7.844	−0.206	1.0	41.0	19.0

Table 1 continued

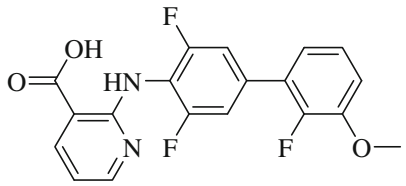
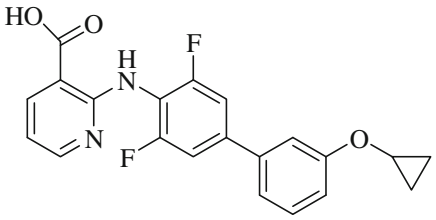
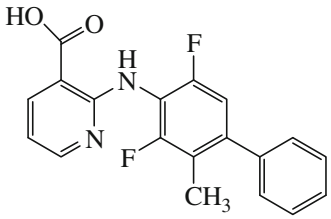
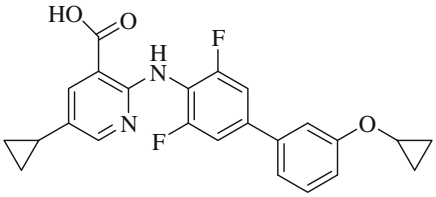
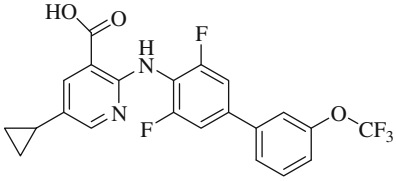
Comp. no.	Structure	IC_{50} (μm)	pIC_{50}	MLR		PCR		PLS		T_N_F_5	4path ClusterCount	T_C_C_6
				[a]Pred	Res	[b]Pred	Res	[c]Pred	Res			
16		53	7.276	7.524	−0.248	7.564	−0.288	7.531	−0.255	2.0	39.0	16.0
17		17	7.769	8.204	−0.435	8.008	−0.239	8.006	−0.237	2.0	36.0	16.0
18		5	8.301	8.023	0.278	8.014	0.287	8.021	0.28	2.0	44.0	21.0
19		6	8.222	7.929	0.293	7.954	0.268	8.148	0.074	2.0	41.0	17.0
20[t]		4	8.398	8.674	−0.276	8.223	0.175	8.212	0.186	2.0	42.0	15.0

Table 1 continued

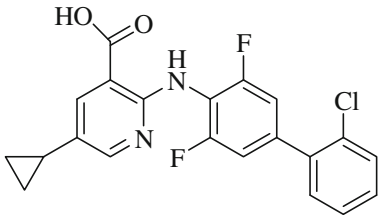
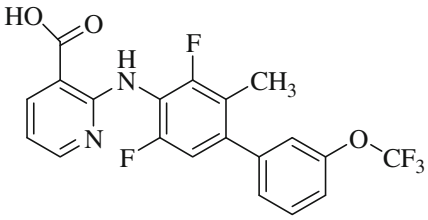
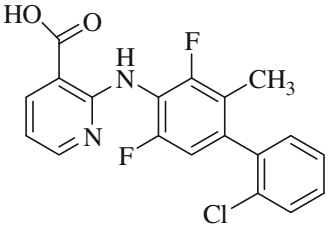
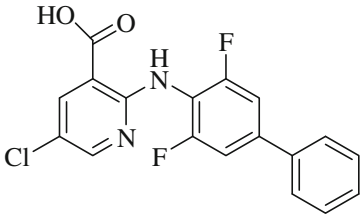
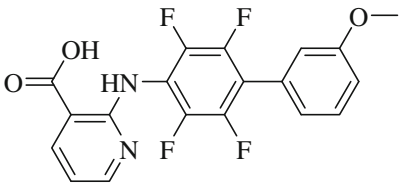
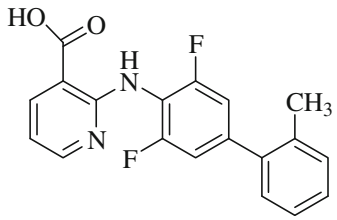
Comp. no.	Structure	IC_{50} (μm)	pIC_{50}	MLR		PCR		PLS		T_N_F_5	4path ClusterCount	T_C_C_6
				[a]Pred	Res	[b]Pred	Res	[c]Pred	Res			
21		5	8.301	8.501	−0.2	8.518	−0.217	8.204	0.097	2.0	39.0	18.0
22		6	8.222	7.973	0.249	7.987	0.235	7.989	0.233	2.0	40.0	16.0
23		4	8.398	8.268	0.13	8.273	0.125	8.267	0.131	2.0	32.0	14.0
24		8	8.097	8.011	0.086	7.987	0.11	8.007	0.09	1.0	33.0	16.0
25		3	8.523	8.568	−0.045	8.597	−0.074	8.573	−0.05	2.0	45.0	16.0

Table 1 continued

Comp. no.	Structure	IC_{50} (μm)	pIC_{50}	MLR		PCR		PLS		T_N_F_5	4path ClusterCount	T_C_C_6
				[a]Pred	Res	[b]Pred	Res	[c]Pred	Res			
26		11	7.959	8.017	−0.058	8.001	−0.042	8.014	−0.055	2.0	34.0	15.0

t test set compounds, *Res* residual values

[a] Pred Predicted pIC_{50} by MLR

[b] Pred Predicted pIC_{50} by PCR

[c] Pred Predicted pIC_{50} by PLS

Table 2 Unicolumn statistics of training and test sets (2D QSAR)

Set	Average	Max	Min	Std Dev.	Sum
Training	7.741	8.523	6.699	0.573	170.31
Test	7.381	7.921	6.824	0.5234	29.52

was a true representative of the training set. This was achieved by arbitrarily setting aside four compounds as a test set with a regularly distributed biological data. Unicolumn statistics of test and training sets (Table 2) showed accurate selection of test and training sets, as maximum of the training set was more than that of test set and the minimum of training set was less than or equal to that of test set.

Statistical computation

Vlife MDS was used to generate 2D QSAR models by multiple linear regression (MLR), principal component regression (PCR) and partial least squares (PLS) regression methods, coupled with forward–backward variable selection method. Statistical measures were used for the evaluation of 2D QSAR models were the number of compounds in regression n, regression coefficient r^2, number of descriptors in a model *k*, *F*-test (Fisher test value) for statistical significance *F*, cross validated correlation coefficient q^2, predictive squared correlation coefficients r^2_{pred}, coefficient of correlation of predicted data set pred_r^2se and standard error of estimation r^2 se and q^2 se.

Multiple linear regression (MLR) analysis

MLR is a regression method used to model linear relationship between a dependent variable Y (*h*DHODH inhibitory activity) and independent variables X (2D descriptors). MLR is based on least squares: the model is fit such that sum-of-squares of differences of observed and a predicted value is minimized. MLR estimates values of regression coefficients (r^2) by applying least squares curve fitting method. The model creates a relationship in the form of a straight line (linear) that best approximates all the individual data points. Regression equation takes the form

$$Y = b1 * x1 + b2 * x2 + b3 * x3 + c$$

where *Y* is dependent variable, '*b*'s are regression coefficients for corresponding '*x*'s (independent variable), '*c*' is a regression constant, or intercept (Kubyani 1994; Croux and Joossens 2005).

Principal component regression (PCR) method

PCR is a data compression method based upon the correlation among dependent and independent variables. PCR provides a method for finding structure in data sets. Its aim is to group correlated variables, replacing the original descriptors by new set called principal components. These principal components uncorrelated and are built as a simple linear combination of original variables. It rotates the data into a new set of axes such that first few axes reflect most of the variations within the data. PCA selects a new set of axes for the data. These are selected in decreasing order of variance within the data. Purpose of PCR is the estimation of values of a dependent variable on the basis of selected principal components (PCs) of independent variables (Huberty, 1984).

Partial least squares (PLS) regression method

PLS analysis is a popular regression technique which can be used to correlate one or more dependent variable (*Y*) to several independent (*X*) variables. PLS relates a matrix *Y* of dependent variables to a matrix *X* of molecular structure descriptors. PLS is useful in situations where the number of independent variables exceeds the number of observation, when *X* data contain colinearties or when *N* is less than

5 M, whares N is number of compound and M is number of dependant variable. Main aim of PLS regression is to predict the activity (Y) from X and to describe their common structure (Wold *et al.,* 2001).

2D QSAR models were generated using pIC_{50} values as dependent variable and various descriptors values as independent variables. The cross-correlation limit was set at 0.5, number of variables in final equation is 4 in MLR, PCR and PLS. Term selection criteria was set as r^2 and F-test, 'in' at 4 and 'out' at 3.99. Variance cutoff was set at 0, scaling to auto scaling and number of random iterations to 10.

Validation of QSAR models

The definitive validity of the model is examined by mean of external validation (q^2), which evaluates how well the equation generalizes. The training set was used to derive an adjustment model that was used to predict the activity of the test set. The predicted power of equations was validated using predictive squared correlation coefficients r^2_{pred}.

3D QSAR modeling and data set

The structures of all the compounds were constructed from the template molecule (compound 25) by using the "SKETCH" option function in SYBYL, and partial atomic charges were calculated by the Gasteiger Huckel method and energy minimizations were performed using the Tripos force field (Gasteiger and Marsili, 1980) with a distance-dependent dielectric and the Powell conjugate gradient algorithm convergence criterion of 0.01 kcal/mol Å (Clark *et al.,* 1989). The total set of inhibitors was divided manually into training set of 22 compounds for generating 3D QSAR model and a test set of 4 compounds for validating the quality of the model.

Molecular modeling and alignment

The alignment of molecules is the process of aligning two or more molecules in 3D space to optimally superimpose specific atoms on each other based on distances. Compound 25 was used as a template because of the highest activity and all other compounds were aligned on the basis of the common structure (Fig. 3). Rigid body alignment of molecules in a Mol2 database was performed using maximum common substructures defined by Distill (without including bond types in rings). Structure of the template compound 25 and common substructure in bold is shown in Fig. 3. Alignment of training and test set compounds is shown in Fig. 4.

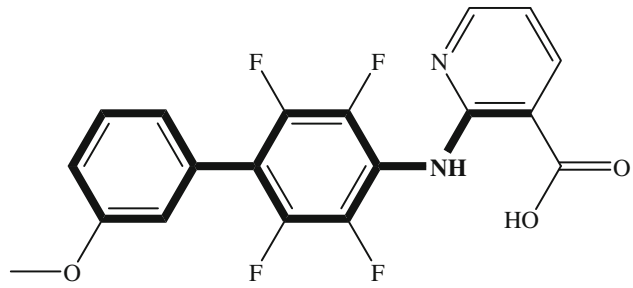


Fig. 3 Structure of the template compound **25**, common substructure is in *bold*

CoMFA model

CoMFA steric and electrostatic interaction fields of each molecule were calculated on a 3D cubic lattice with grid spacing of 2 Å in all the Cartesian directions and CoMFA fields were calculated using the QSAR module of SYBYL. CoMFA descriptors were calculated using sp^3 carbon probe atom with a van der Waals radius of 1.52 Å and a charge of +1.0 to generate steric (Lennarde-Jones 6–12 potential) field energies and electrostatic (Coulombic potential) fields with a distance-dependent dielectric at each lattice point. The SYBYL default energy cutoff of 30 kcal/mol was set for both steric and electrostatic fields. In order to reduce noise and improve efficiency, column filtering (minimum sigma) was set to 2.0 kcal/mol.

Predictive r^2 value

To validate the CoMFA model, predictive abilities for the test set compounds (expressed as r^2_{pred}) was determined using the following equation

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

where, SD is the sum of the squared deviations between the inhibitory activity of molecules in the test set and the mean inhibitory 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.

Analysis of the residuals

The training set was initially checked for outliers for both 2D and 3D QSAR analysis. In general, if the residual of a compound between experimental and predicted pIC_{50} values is greater than 1 logarithm unit, compound is considered as outlier. Examination of the residuals from cross-validated predictions (Tables 1, 3) indicated that there is no outlier in 2D and 3D QSAR models.

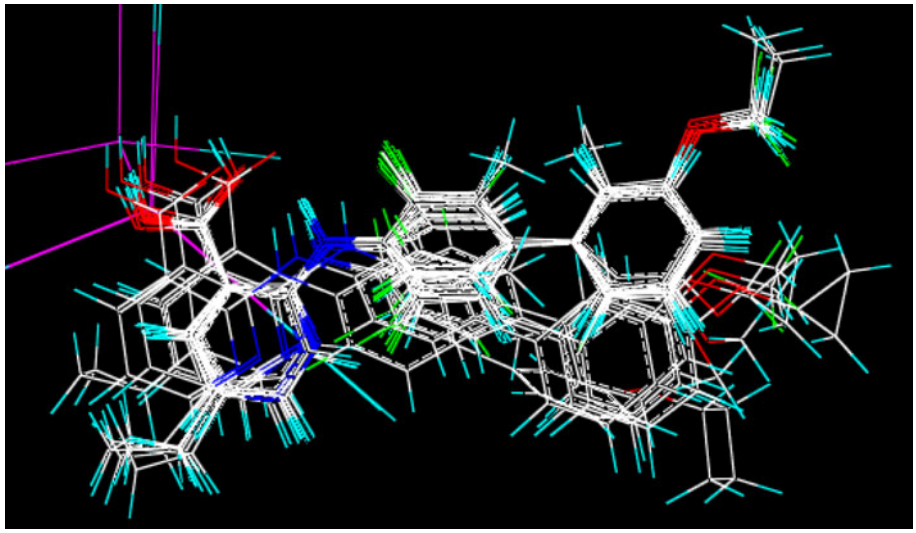


Fig. 4 Alignment of training and test set compounds on compound **25**

Table 3 Experimental and predicted pIC_{50} with residual values using 3D QSAR (CoMFA) model

Compound no.	Experimental pIC_{50}	Predicted pIC_{50}	Residual
1	6.699	6.712	−0.013
2[t]	7.056	7.150	−0.094
3	6.824	6.713	0.110
4	7.046	7.181	−0.136
5	7.721	7.648	0.074
6	7.824	7.981	−0.157
7[t]	7.721	7.786	−0.065
8	7.854	7.530	0.323
9	6.699	6.961	−0.262
10	6.959	6.818	0.140
11	7.482	7.347	0.135
12[t]	7.921	8.046	−0.125
13	7.004	7.012	−0.607
14	7.921	7.612	0.309
15	7.638	7.736	−0.098
16	7.276	7.229	0.047
17	7.769	7.794	−0.025
18	8.301	8.315	−0.014
19	8.222	8.251	−0.029
20[t]	8.398	8.370	0.028
21	8.301	8.173	0.128
22	8.222	8.276	−0.054
23	8.398	8.125	0.273
24	8.097	8.108	−0.011
25	8.523	8.574	−0.051
26	7.959	7.785	0.174

t test set compound

Results and discussion

Results of 2D QSAR study

Generation of 2D QSAR models

2D QSAR study on amino nicotinic acid and isonicotinic acid derivatives resulted in several QSAR models. Statistically significant QSAR models were selected for discussion.

Model-1 (MLR) $pIC_{50} = +0.4500\,(T_N_F_5) + 0.0613\,(4pathClusterCount) - 0.1157\,(T_C_C_6) + 6.7645$ where, $n = 22_{training}$ and 4_{test}, $k = 2$, DF = 19, $r^2 = 0.834$, $q^2 = 0.756$, F test = 47.83, r^2 se = 0.243, q^2 se = 0.295, $r^2_{pred} = 0.793$, pred_r^2se = 0.334.

Model-2 (PCR) $pIC_{50} = +0.4275\,(T_N_F_5) + 0.0647\,(4pathClusterCount) - 0.1156\,(T_C_C_6) + 6.6789$ where, $n = 22_{training}$ and 4_{test}, $k = 2$, DF = 19, $r^2 = 0.833$, $q^2 = 0.774$, F test = 47.34, r^2 se = 0.244, q^2 se = 0.284, $r^2_{pred} = 0.811$, pred_r^2se = 0.304.

Model-3 (PLS) $pIC_{50} = +0.4500\,(T_N_F_5) + 0.0613\,(4path\ ClusterCount) - 0.1157\,(T_C_C_6) + 6.7645$ where, $n = 22_{training}$ and 4_{test}, $k = 2$, DF = 19, $r^2 = 0.864$, $q^2 = 0.786$, F test = 48.83, r^2 se = 0.233, q^2 se = 0.294, $r^2_{pred} = 0.821$, pred_r^2se = 0.304.

In above QSAR models, r^2 is a correlation coefficient that multiply by one hundred gives explained variance in inhibitory activity. Predictive ability of generated QSAR

models was evaluated by q^2 employing leave-one-out method. F value reflects ratio of variance explained by models and variance due to error in regression. High F value indicates that model is statistically significant. Low standard error (SE) of estimation indicted by r^2 se and q^2 se, suggested that models are statistically significant. Predictive ability of QSAR model was also confirmed by external validation of test set compounds denoted by r^2_{pred}. Observed and predicted pIC_{50} is shown in Table 1. Plot of observed versus predicted pIC_{50} is shown in Fig. 5.

Interpretaion of 2D QSAR models

Descriptors used in generation of 2D QSAR models are shown in Fig. 6. 2D QSAR models indicates positive contribution of T_N_F_5 and 4pathClusterCount and negative contribution of T_C_C_6. Alignment independent (AI) topological descriptor (Balaban, 1982) T_N_F_5 contributed positively to QSAR models, where T_C_C_6 contributed negatively. Alignment-independent descriptors can be generated by considering the topology of the molecule, atom type, and bond. For calculation of alignment independent descriptors every atom in the molecule was assigned at least one and at most three attributes. First attribute is 'T-attribute' to thoroughly characterize topology of the molecule. Second attribute is atom type, atom symbol is used here. Third attribute is assigned to atoms taking part in a double or triple bond. After all the atoms have been assigned their respective attributes, selective distance count statistics for all combinations of different attributes are computed. A selective distance count statistic 'XY2' (e.g., 'TOPO2N3') counts all the fragments between start atom with attribute 'X' (e.g., '2' double bonded atom) and end atom with attribute 'Y' (e.g., 'N') separated by graph distance 3. Graph distance can be defined as the smallest number of atoms along the path connecting two atoms in the molecular structure. In this study, to calculate AI descriptors, we used following attributes: 2 (double bonded atom), 3 (triple bonded atom), C, N, O, H, F, and Cl the distance range of 0–7. T_N_F_5 is a count of number of nitrogen atoms separated from any fluorine atom (single or double bonded) by five bond distance, e.g., N_C_N_C_C_C_F. Positive contribution of T_N_F_5 reveals the importance of presence of nitrogen atom in pyridine ring and fluorine atom on first phenyl ring of biphenyl ring template. T_C_C_6 is a count of number of carbon atoms separated from any other carbon atom (single or double bonded) by six distance, e.g., C_C_C_C_C_C_C_C. 4pathClusterCount is a molecular connectivity index which signifies total number of fragments of fourth order path cluster in a molecule. Molecular connectivity index is used to describe electronic environment and bonding configuration of each non-hydrogen atom (heavy atom) in the molecule for example carbon valence connectivity index takes into account only bonds between carbon atoms. 4pathClusterCount reveals the importance of molecular connectivity for heavy atoms and their bonding configuration in the molecules.

Results of 3D QSAR study

The q^2, r^2_{pred}, r^2_{ncv}, F, and SEE values were computed as defined in SYBYL. PLS analysis showed a high q^2 value of

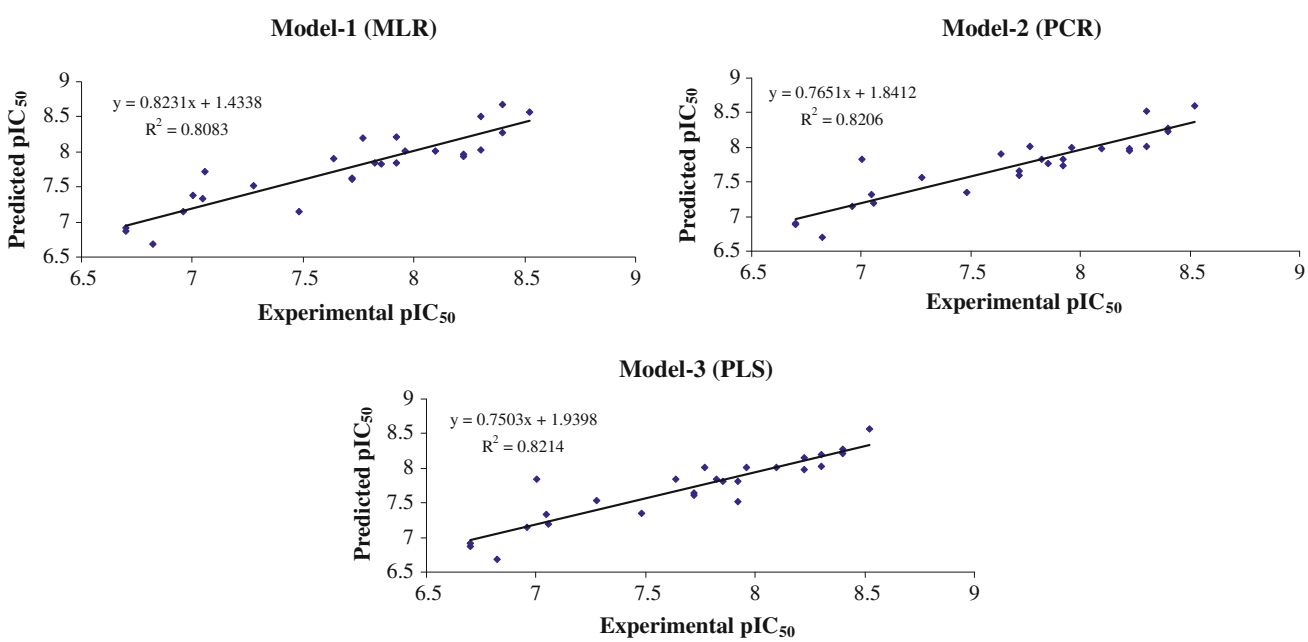


Fig. 5 Graphs of experimental versus predicted pIC_{50} using 2D QSAR models

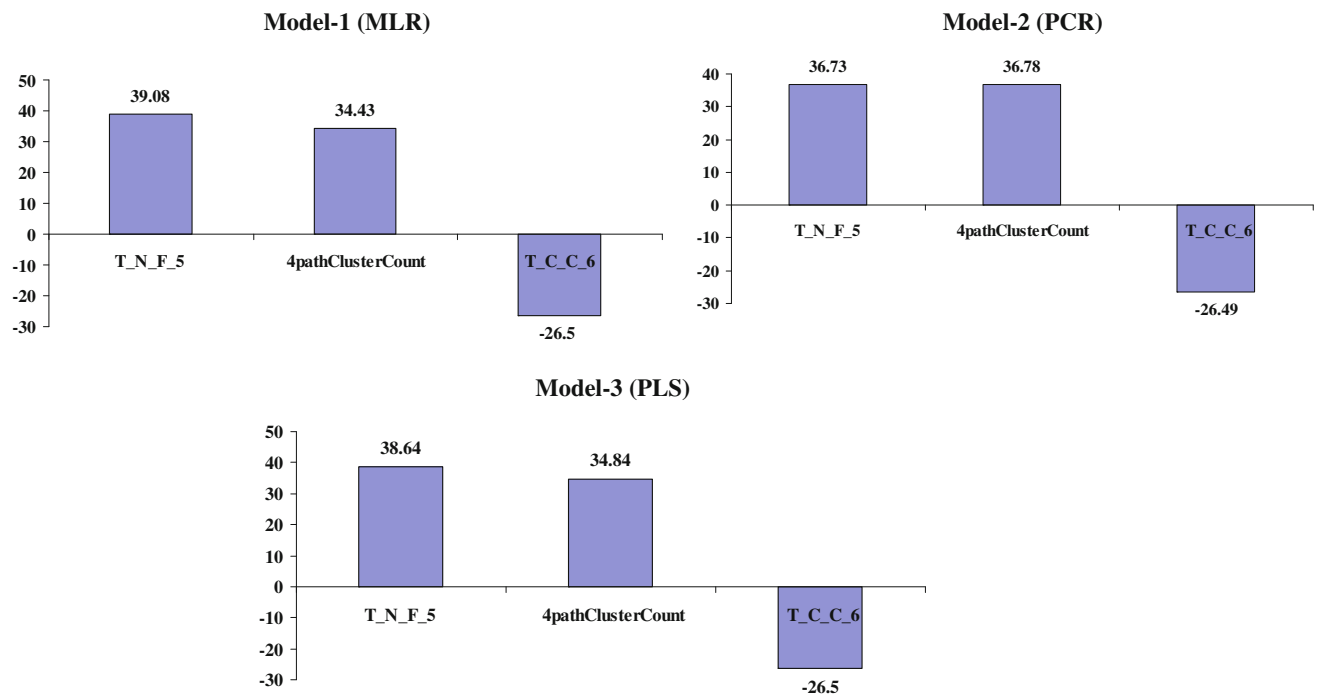


Fig. 6 Contribution charts of 2D QSAR models

0.630 with four components. Cross-validated q^2 (r^2_{cv}) of 0.630 indicated a good predictive ability of the model. The non-cross-validated PLS analysis results in a conventional r^2 of 0.949, $F = 137$ and a standard error of estimation (SEE) of 0.218. In both steric and electrostatic field contributions, the former accounts for 0.554, while the latter contributes 0.446, indicating that these two factors nearly contribute the same to the binding affinities. The high bootstrapped r^2 (0.966) value and low standard deviation (0.032) suggest a high degree of confidence in the analysis. The predicted, experimental activity and the residual value of all the inhibitors are listed in Table 3, and the correlation between predicted and experimental activity is depicted in Fig. 7. The predictive ability of the 3D QSAR model was further validated using an external test set of four compounds not included in the model generation study. The predicted r^2 (r^2_{pred}) values from the CoMFA model was 0.763.

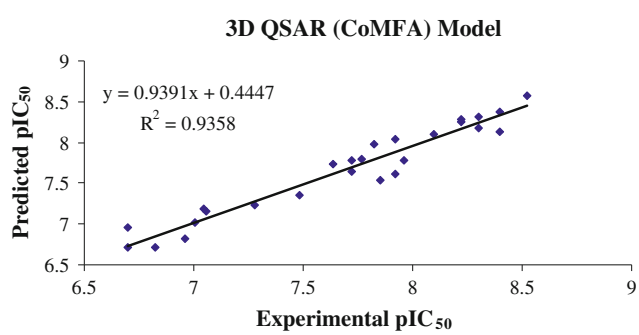


Fig. 7 Plot of experimental versus predicted pIC_{50} using 3D QSAR (CoMFA) model

CoMFA contour maps

Contour maps for the best CoMFA model are shown in Fig. 8. In the contour maps, the steric CoMFA contour plot with the highest active compound 25 is shown in Fig. 8a. The field energies at each lattice point were calculated as the scalar results of the coefficient and the standard deviation associated with a particular column of the data table (std*coeff), as always plotted as the percentages of the contribution of CoMFA equation. In this figure, 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 steric contour of CoMFA showed a large green contour around the first phenyl ring of biphenyl ring template, indicating a favorable effect of steric bulk of fluorine atom for inhibitory activity. This steric favored area is generated by high electron density of fluorine atom. This can be explained by analyzing the structural features and inhibitory activity of 25 (2,3,5,6-tetrafluorophenyl, IC_{50} = 3 μm) and 26 (2,6-difluorophenyl IC_{50} = 11 μm), 13 (2-fluoro, 5-methylphenyl IC_{50} = 99 μm), and 3(2-chlorophenyl IC_{50} = 150 μm). Fluorine atom is larger than hydrogen, thus steric bulk (lipophilicity) in the molecule can be increased by replacing H atom by F atom. A steric unfavorable yellow contour was observed near the C-3′ methoxy at terminal phenyl ring of biphenyl ring template, suggested that bulky

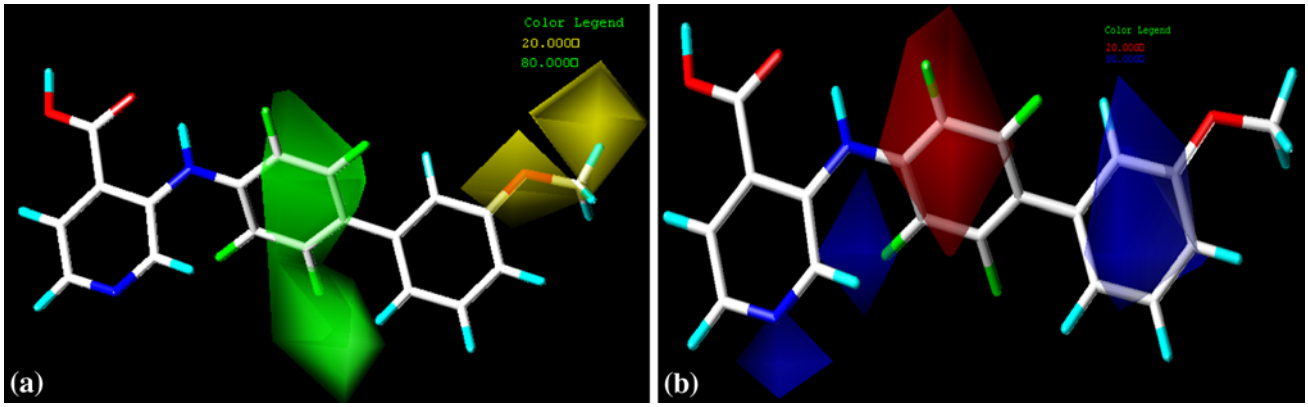


Fig. 8 CoMFA (std*coeff) contour maps. Compound **25** is shown inside the field, **a** CoMFA steric contour map and **b** CoMFA electrostatic contour maps

groups in these region would decrease *h*DHODH inhibitory activity. CoMFA electrostatic contour map is shown in Fig. 8b. Regions where increased positive-charge is favorable for inhibitory activity are indicated in blue (80% contribution), while regions where increased negative-charge is favorable for inhibitory activity are indicated in red (20% contribution). A large region of red contours near the first phenyl ring of biphenyl ring template shows that the presence of electronegative substituent (–F). Fluorine has the highest electron density, thus such an electronegative groups ($–CF_3$, $–OCF_3$) are very important for better *h*DHODH inhibitory activity. It can shows the fact that activity of 13 (2-fluoro, 5-methylphenyl IC_{50} = 99 μm) is less than 2 (2,5-difluorophenyl IC_{50} = 88 μm). A large blue contour is seen in the vicinity of terminal phenyl ring, depicts that positively charged groups, such as hydrogen atoms is beneficial for inhibitory activity. This is indeed the case for 24 (IC_{50} = 8 μm) and 4 (IC_{50} = 90 μm). Second blue polyhedron near the C-2 position of pyridine ring indicate that a low electron density in this area will have a positive effect on the inhibitory activity. Small blue polyhedra located near the nitrogen atom of pyridine ring, indicate that an electropositive group needs to be present in this region.

2D versus 3D QSAR (CoMFA) analysis

The comparison of 2D and 3D QSAR (CoMFA) analysis suggested common structural features responsible for *h*DHODH inhibitory activity. Positive contribution of alignment-independent topological descriptor T_N_F_5 reveals the importance of nitrogen atom in pyridine ring and fluorine atom on first phenyl ring of biphenyl ring template separated by five bond distance. T_N_F_5 is an important descriptor, accounts for highest contribution (Fig. 6) for *h*DHODH inhibitory activity in all the 2D QSAR models. Fluorine is much more lipophilic than hydrogen, so incorporating fluorine atoms in a molecule will make it more lipophilic. Lipophilicity is an important property in describing the affinity of the compounds in terms of their partitioning the biological membranes hence the fluorinated compound has a higher bioavailability. Fluorine is a good leaving group, so it has a potential for covalent bonds to be formed between the molecule and *h*DHODH by loss of fluoride, leading to inhibition of *h*DHODH activity. Lone pair of electron on N-atom in pyridine ring system can form H-bond with CoQ binding site of *h*DHODH. Positive contribution of 4pathCluster-Count reveals the importance of molecular connectivity for heavy atoms and their bonding configuration in the molecules. Analysis of CoMFA steric and electrostatic contour plots offered enough information to understand the binding mode between the inhibitors and binding site (CoQ) of *h*DHODH. The bulky and electronegative group (–F) of compound at the first phenyl ring of biphenyl ring system seems to be penetrating the junction of red (electrostatic) and green (steric) contours indicating the presence of bulkiness as well as electronegativity for the enhancement of *h*DHODH inhibitory activity. 2D and 3D QSAR models suggested that substitution on first phenyl ring especially with –F, $–CF_3$, and $–OCF_3$ and terminal phenyl ring with positively charged groups led toward better inhibitory activity.

Conclusion

2D and 3D QSAR study identifies common features responsible for *h*DHODH inhibitory activity of nicotinic acid and isonicotinic derivatives. 2D QSAR studies revealed that alignment-independent descriptors were major contributing descriptors. CoMFA model is satisfactory according to the statistical results as well as the contour maps

analysis. CoMFA contour plots offered enough information to understand the binding mode between inhibitors and CoQ binding site of *h*DHODH. The most significant feature for better *h*DHODH inhibitory activity is the substitution pattern (–F, $–CF_3$, $–OCF_3$) on biphenyl ring system. QSAR models generated in this study can provide useful information for the design of new compounds and helped in prediction of *h*DHODH inhibitory activity prior to synthesis.

Acknowledgments The authors would like to thank Director General, Nirma University, Ahmedabad, India for funding the project and Vlife Science Technologies Pvt. Ltd Pune.

References

Balaban AT (1982) Highly discriminating distance-based topological index. Chem Phys Lett 89:399–404

Baumann P, Mandl WS, Volkl A, Adam C, Bumeder I, Oduncu F, Schmidmaier RA (2009) Perspective, randomised, controlled, double-blind phase I-II clinical trial on the safety of A-Part(R) gel as adhesion prophylaxis after major abdominal surgery versus non-treated group. Mol Cancer Ther 8:366–375

Bjornberg O, Rowland P, Larsen S, Jensen KF (1997) Active site of dihydroorotate dehydrogenase A from *Lactococcus lactis* investigated by chemical modification and mutagenesis. Biochemistry 36:16197–16205

Bjornberg O, Gruner AC, Roepstorff P, Jensen KF (1999) The activity of *Escherichia coli* dihydroorotate dehydrogenase is dependent on a conserved loop identified by sequence homology, mutagenesis, and limited proteolysis. Biochemistry 38:2899–2908

Castro PLJC, Erra SM, Lozoya TME, Navarro RE (2010) Amino nicotinic acid and isonicotinic acid derivative as DHODH inhibitor. US patent US20100074898

Chen SF, Ruben RL, Dexter DL (1986) Mechanism of action of the novel anticancer agent 6-fluoro-2-(2′-fluoro-1,1′-biphenyl-4-yl)-3-methyl-4-quinolinecarboxylic acid sodium salt (NSC 368390): inhibition of de novo pyrimidine nucleotide biosynthesis. Cancer Res 46:5014–5019

Clark M, Cramer RD, Opdenbosch NV (1989) Validation of the general purpose Tripos 5.2 forcefield. J Comput Chem 10:982–1012

Croux C, Joossens K (2005) Influence of observations on the misclassification probability in quadratic discriminant analysis. J Multivar Anal 96:348–403

Fox RI, Herrmann ML, Frangou CG, Wahl GM, Morris RE, Strand V, Kirschbaum BJ (1999) Mechanism of action for leflunomide in rheumatoid arthritis. Clin Immunol 93:198–208

Gasteiger J, Marsili M (1980) Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. Tetrahedron 36:3219–3228

Herrmann ML, Schleyerbach R, Kirschbaum BJ (2004) Leflunomide: an immunomodulatory drug for the treatment of rheumatoid arthritis and other autoimmune diseases. Immunopharmacology 47:273–289

Huberty CJ (1984) Issues in the use and interpretation of discriminant analysis. Psychol Bull 95:156–171

Jones M (1980) Pyrimidine nucleotide biosynthesis in animals: genes, enzymes, and regulation of UMP biosynthesis. Annu Rev Biochem 49:253–279

Kubyani H (1994) Variable selection in QSAR studies. I. An evolutionary algorithm. Quant Struct Act Relat 13:285–294

Mascia L, Turchi G, Bemi V, Ipata PL (2000) Uracil salvage pathway in PC12 cells. Biochem Biophys Acta 1524:45–50

Merrill JE, Hanak S, Pu SF, Liang J, Dang C, Iglesias BD, Harvey B, Zhu B, McMonagle SK (2009) Teriflunomide reduces behavioral, electrophysiological, and histopathological deficits in the Dark Agouti rat model of experimental autoimmune encephalomyelitis. J Neurol 256:89–103

Norager S, Jensen KF, Bjornberg O, Larsen S (2002) *E. coli* dihydroorotate dehydrogenase reveals structural and functional distinctions between different classes of dihydroorotate dehydrogenase. Structure 10:1211–1223

Rozman B (1998) Clinical experience with leflunomide in rheumatoid arthritis. J Rheumatol Suppl 53:27–31

Shannon PVR, Eichholtz T, Linstead D, Masdin P, Skinner R (1999) Condensed heterocyclic compounds as anti-inflammatory and immunomodulatory agents. International patent WO 99/45926, 1999

Shawver LK, Schwartz DP, Mann E, Chen H, Tsai J, Chu L, Taylorson L, Longhi M, Meredith S, Germain L, Jacobs JS, Tang C, Ullrich A, Berens ME, Hersh E, McMahon G, Hirth KP, Powell TJ (1997) Inhibition of platelet-derived growth factor-mediated signal transduction and tumor growth by N-[4-(trifluoromethyl)-phenyl]-5-methylisoxazole-4-carboxamide. Clin Cancer Res 3:1167–1177

Vyas VK, Ghate M (2011) Recent developments in the medicinal chemistry and therapeutic potential of dihydroorotate dehydrogenase (DHODH) inhibitors. Min Rev Med Chem 11:1039–1055

Williamson RA, Yea CM, Robson PA, Curnock AP, Gadher S, Hambleton AB, Woodward K, Bruneau JM, Hambleton P, Spinella-Jaegle S, Morand P, Courtin O, Sautes C, Westwood R, Hercend T, Kuo EA, Ruuth E (1996) Dihydroorotate dehydrogenase is a target for the biological effects of leflunomide. Transplant Proc 28:3088–3091

Wold S, Sjostrom M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. Chemo Int Lab Sys 58:109–130