International Journal of ChemTech Research
CODEN(USA): IJCRGG ISSN : 0974-4290
Vol.2, No.1, pp 606-614, Jan-Mar 2010

3D- Quantitative Structure-Activity Relationship Analysis of Some 2-Substituted Halogenbenzimidazoles Analogues with Antimycobacterial activity

M.C.SHARMA*[1], SMITA SHARMA[2]

[1]School of Pharmacy, Devi Ahilya Vishwavidyalaya, Takshshila Campus, Khandwa Road, Indore- 452 001, Madhya Pradesh, India

[2]Department of Chemistry Yadhunath Mahavidyalya Bhind- 477001 Madhya Pradesh, India

Corres.author: mukeshcsharma@yahoo.com
Tel. : +91 9826372944

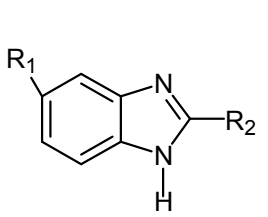
ABSTRACT: A three dimensional quantitative structure-activity relationship study using the k-nearest neighbor analysis method was performed on a series of -polyfluoroalkyl and 2-Nitrobenzylsulphanyl benzimidazoles for their Antimycobacterial, activity. This study was performed using 28 compounds, for which k-nearest neighbor analysis models were developed using a training set of 22 compounds. Database alignment of all 28 compounds was carried out by root-mean-square fit of atoms and field fit of the steric and electrostatic molecular fields. The resulting database was analyzed by partial least squares analysis with cross-validation; leave one out and no validation to extract optimum number of components. Further comparison of the coefficient contour maps with the steric and electrostatic properties of the receptor has shown a high level of compatibility and good predictive capability.
KEY WORDS: *3DQSAR, kNN, method Antimycobacterial, Halogenbenzimidazoles*

INTRODUCTION

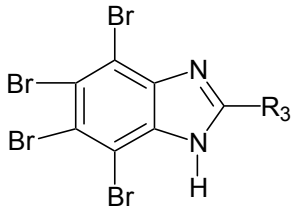
Despite the availability of highly potential antitubercular agents, tuberculosis remains primary cause of comparatively high mortality worldwide. The statistics shows that around three million people throughout the world die annually from tuberculosis [1, 2] and today more people die from tuberculosis than ever before [3]. Therefore, the development of new drugs with activity against multi drug-resistant (MDR) TB, extensively drug-resistant (XDR) TB, and latent TB is a priority task. Although new agents that will shorten the duration of current chemotherapy are also needed. A special interest has been focused on five membered heterocyclic compounds like imidazole, pyrrole, oxadiazole, specially triazoles and other heterocyclic system derivatives have been reported [4]. The azole antitubercular may be regarded as a new class providing truly effective drugs, which is reported to inhibit bacteria by blocking the biosynthesis of certain bacterial lipids and/or by additional mechanisms [5, 6]. Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, remains the leading cause of mortality due to a bacterial pathogen. WHO estimated that there were 8.8 million new cases of tuberculosis in 2020. No new drug against tuberculosis has been developed in the last 30 years. The main objective of the present study was the search for novel benzimidazole compounds that would show a promise to become useful antimycobacterial agent[7]. A series of compounds of 2-polyfluoroalkyl and 2-Nitro benzyl sulphanyl benzimidazoles was selected as novel antimycobacterial agents for 3D-QSAR studies. Many

different approaches to QSAR have been developed over the years. The rapid increase in three-dimensional structural information (3D) of bioorganic molecules, coupled with the development of fast methods for 3D structure alignment (e.g. active analogue approach), has led to the development of 3D structural descriptors and associated 3D QSAR methods. The most popular 3D QSAR methods are comparative molecular field analysis (CoMFA) [8] and comparative molecular similarity analysis (CoMSIA).[9] The CoMFA methods involves generation of a common three dimensional lattice around a set of molecules and calculation of the steric and electrostatic interaction energies at the lattice points. The interaction energies are numerically very high when a lattice point is very close to an atom and special care needs to be taken in order to avoid problems arising because of this. The CoMSIA method avoids these problems by using similarity function represented as Gaussian. This information around the molecule is converted into numerical data using the partial least squares (PLS) method that reduces the dimensionality of data by generating components. However, a major disadvantage is that PLS attempts to fit a linear curve among all the points in the data set. Further, the PLS method does not offer scope for improvement in results. It has been observed from several reports that the predictive ability of PLS method is rather poor due to fitting of a linear curve between the available points. In the case of the CoMSIA method, molecular similarity is evaluated and used instead of molecular field, followed by PLS analysis. Variable selection methods have also been adopted for optimal region selection in 3D QSAR methods and shown to provide improved QSAR models as compared to the original CoMFA technique. For example, GOLPE[10] was developed using chemometric principles, and q2- GRS was developed on the basis of independent analyses of small areas (or regions) of near molecular space to address the issue of optimal region selection in CoMFA.[11] These considerations provide an impetus for the development of fast, generally nonlinear, variable selection methods for performing molecular field analysis. We report here the development of a new method (kNN MFA) that adopts a k-nearest neighbor principle for generating relationships of molecular fields with the experimentally reported activity. This method utilizes the active analogue principle that lies at the foundation of medicinal chemistry.

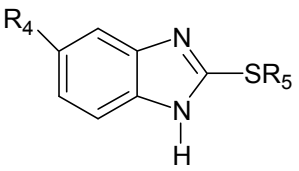
Table I- series of compounds of 2- Substituted Halogenobenzimidazoles with I_{C50} and P_{IC50} values



2a-2l



3a-3d



5a-5l

S.No	Compound	R^1	R^2	IC_{50}	Log IC_{50}
1.	2a	$4,6Cl_2$	Cf_3	32	1.505150
2.	2b	4,6 Cl_2	C_2f_5	16	1.204120
3.	2c	4,6 Cl_2	C_3f_7	8	0.903090
4.	2d	4,6 Cl_2	C_4f_9	8	0.903090
5.	2e	5,6 Cl_2	Cf_3	8	0.903090
6.	2f	5,6 Cl_2	C_2f_5	4	0.602060
7.	2g	5,6 Cl_2	C_3f	8	0.903090
8.	2h	5,6 Cl_2	C_4f_9	4	0.602060
9.	2i	$4,6Br_2$	Cf_3	32	1.505150
10.	2j	$4,6Br_2$	C_2f_5	16	1.204120
11.	2k	$4,6Br_2$	C_3f_7	16	1.204120
12.	2l	$4,6Br_2$	C_4f_9	16	1.204120

		R^3	IC_{50}	Log IC_{50}
13	3a	Cf_3	32	1.505150
14	3b	C_2f_5	8	0.903090
15	3c	C_3f_7	4	0.602060
16	3d	C_4f_9	16	1.204120

		R^4	R^5	IC_{50}	Log IC_{50}
1	5a	5-cl	3,5-dinitrobenzyl	2	0.301030
2	5b	5-Br	3,5-dinitrobenzyl	4	0.301030
3	5c	5-I	3,5-dinitrobenzyl	2	0.602060
4	5d	4,6- Cl_2	4-nitrobenzyl	2	0.301030
5	5e	4,6- Cl_2	2,4-dinitrobenzyl	32	1.505150
6	5f	4,6- Cl_2	3,5-dinitrobenzyl	4	0.602060
7	5g	4,6-Br_2	4-nitrobenzyl	2	0.301030
8	5h	4,6-Br_2	2,4-dinitrobenzyl	16	0.301030
9	5i	4,6-Br_2	3,5-dinitrobenzyl	4	0.602060
10	5j	4,5,6,7-Br_4	4-nitrobenzyl	16	0.602060
11	5k	4,5,6,7-Br_4	2,4-dinitrobenzyl	16	1.204120
12	5l	4,5,6,7-Br_4	3,5-dinitrobenzyl	8	0.903090

MATERIAL AND METHODS

3D QSAR Analysis

The data set used for the QSAR analyses contains 28, 2-polyfluoroalkyl and 2-Nitrobenzylsulphanyl benzimidazoles was selected as novel antimycobacterial agents for QSAR studies. All the structures of the compounds were drawn in 2D-APPL mode of software and exported to 3D model. The chemical structure and their corresponding IC_{50} values were mentioned in Table I. The modeling analyses, calculations, and visualizations for 2D QSAR were performed using the V-Life Molecular Design Suite 3.0 (Vlife MDS) [12]. A set of 28 molecules was selected and divided in training (19) and test set (9).The negative logarithm of IC_{50} values (PIC_{50}) calculated using the IC_{50} values of reported compounds. The biological activity data (IC_{50} in Molar) were converted in to pIC_{50} according to the formula pIC_{50} = (-log (IC_{50}). Thus such studies may help for the design and synthesis of better 2-polyfluoroalkyl and 2-Nitrobenzylsulphanyl benzimidazoles. All the twenty eight compounds were built on workspace of molecular modeling software V-Life MDS 3.5, which is a product VLife Sciences Pvt Ltd., India[12] .The compounds were then subjected to conformational analysis and energy minimization using montocarlo conformational search with RMS gradient of 0.001 kcal/mol and iteration limit of 10000 using a MMFF94 force field. V-Life Molecular Design Suite (VLife-MDS), allows user to choose probe, grid size, and grid interval for the generation of descriptors. The variable selection methods along with the corresponding parameters are allowed to be chosen, and optimum models are generated by maximizing q2. K-nearest neighbor molecular field analysis (kNN-MFA) requires suitable alignment of given set of molecules. This is followed by generation of a common rectangular grid around the molecules. The steric and electrostatic interaction energies are computed at the lattice points of the grid using a methyl probe of charge +1. These interaction energy values are considered for relationship generation and utilized as descriptors to decide nearness between molecules. The term descriptor is utilized in the following discussion to indicate field values at the lattice points. The optimal training and test sets were generated using the sphere exclusion algorithm.[13]This algorithm allows the construction of training sets covering descriptor space occupied by representative points. Once the training and test sets were generated, kNN methodology was applied to the descriptors generated over the grid.

Calculation of field descriptor values: The aligned molecules were transferred to the 3D QSAR worksheet (MS excel type) as one of the modules in MDS. The Gi_50 values were converted to pGi_50 (log10 Gi_50) values and pasted in the next column.

For calculation of field descriptor values, both electrostatic and steric field type with cut offs 10.0 and 30.0 Kcal/mol respectively were selected and charge type was selected as Gasteiger – Marsili. Dielectric constant was set to 1.0 considering the distance dependent dielectric function. Probe setting was carbon atom with charge 1.0 and grid setting as follows:

	From	To	Interval
X:	- 4.130200	10.603500	2.0000
Y:	- 6.636500	12.074900	2.0000
Z:	- 14.37640	6.074400	2.0000

This resulted in calculation of 2080 field descriptors (1040 for each electrostatic and steric) for all the compounds in separate columns. For performing QSAR analysis, all the invariable columns were removed from the work sheet.

Building k-NN MFA model using step-wise variable selection method : The k-NN MFA models for all the antimycobacterial activities were developed using step-wise forward-backward method with cross correlation limit set to 0.5 and term selection criteria as q^2. F-test 'in' was set to 4.0 and F-test 'out' to 3.99. As some additional parameters, variance cut-off was set as 2 Kcal/mol A^0 and scaling and auto scaling, additionally the K-Nearest Neighbor parameter setting was done within the range of 2-5 and prediction method was selected as distance based weighted average.

Nearest Neighbor (kNN) Method

The kNN methodology relies on a simple distance learning approach whereby an unknown member is classified according to the majority of its k-nearest neighbours in the training set. The nearness is measured by an appropriate distance metric (e.g., a molecular similarity measure calculated using field interactions of molecular structures). The standard kNN method is implemented simply as follows:[14]Calculate distances between an unknown object (u) and all the objects in the training set; select k objects from the training set most similar to object u, according to the calculated distances; and classify object u with the group to which the majority of the k objects belongs. An optimal k value is selected by optimization through the classification of a test set of samples or by leave-one out cross-validation.

kNN-MFA with Simulated Annealing

Simulated annealing (SA) is the simulation of a physical process, 'annealing', which involves heating the system to a high temperature and then gradually cooling it down to a preset temperature (e.g., room temperature). During this process, the system samples possible configurations distributed according to the Batsman distribution so that at equilibrium, low energy states are the most populated.

kNN-MFA with Stepwise (SW) Variable Selection

This method employs a stepwise variable selection procedure combined with kNN to
optimize the number of nearest neighbours (k) and the selection of variables from the original pool as described in simulated annealing.

kNN-MFA with Genetic Algorithm

Genetic algorithms (GA) first described by Holland[15] mimic natural evolution and selection. In biological systems, genetic information that determines the individuality of an organism is stored in chromosomes. Chromosomes are replicated and passed onto the next generation with selection criteria depending on fitness.The 3D QSAR for molecular field analysis was performed using the k Nearest Neighbour method using software V-LIFE MDS 3.0. A set of 28 molecules was selected and divided in training (19) and test (9) sets for 3D QSAR studies. All the molecules were aligned based on templet. The model was generated by kNN method which showed the cross validated squared correlation coefficient value Q2 = 0.6765 and Pred r2 = -0.0419 and number of k Nearest Neighbour = 2. the model showed the better correlation with biological activity. The steric and electrostatic contribution showed the effect of substitution on biological activity. Subtituted polyfluoroalkyl substituents at R-2 are essential for biological activity. The dichloro substituent at R-1 gives compounds with better biological activity than dibromo substituent. The study on substituted benzylsulphanylbenzimidazole indicated that of various modified benzyl substituent tested, the respective nitro compounds were the most potent antimycobacterials.The 3, 5-dinitro compounds were several times more effective against M.tuberculosis than the respective isomeric 2,4-dinitro derivatives.

Statistical analysis

Models were generated by using three significant statistical methods, namely, partial least square analysis, multiple regressions, and principle component analysis. The cross-validation analysis was performed using the leave-one-out method. In the selected equations, the cross-correlation limit was set at 0.5, the

number of variables at 10, and the term selection criteria at r^2. An F value was specified to evaluate the significance of a variable. The higher the F value, the more stringent was the significance level: F test ''in'' as 4 and F test ''out'' as 3.99. The variance cutoff was set at 0, and scaling was auto scaling in which the number of random iterations was set at 100.The following statistical parameters were considered for comparison of the generated QSAR models: correlation coefficient (r), squared correlation coefficient (r2), predictive r2 for external test set (pred r2) for external validation, and Fischer's (F).The predicted r2 (pred_r2) value was calculated using Eq. 1, where yi and y^i are the actual and predicted activities of the i^{th} molecule in the test set, respectively, and ymean is the average activity of all molecules in the training set. Both summations are over all molecules in the test set. The pred_r^2 value indicates the predictive power of the current model for the external test set as follows

$$\text{pred\_}r^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{mean})^2} \quad (1)$$

Internal validation was carried out using leave-one-out (q2, LOO) method. For calculating q2, each molecule in the training set was eliminated once and the activity of the eliminated molecule was predicted by using the model developed by the remaining molecules. The q2 was calculated using the equation which describes the internal stability of a model:

$$q^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{mean})^2} \quad (2)$$

Where y_i, and $\hat{y}_i$ are the actual and predicted activity of the $_i$th molecule in the training set, respectively, and y_{mean} is the average activity of all molecules in the training set.

Randomization Test.

To evaluate the statistical significance of the QSAR model for an actual data set, we have
employed a one-tail hypothesis testing. The robustness of the QSAR models for experimental training sets was examined by comparing these models to those derived for random data sets. Random sets were generated by rearranging biological activities of the training set molecules. The significance of the models hence obtained was derived based on calculated Zscore.[17-18]

Evaluation of the QSAR Models.

The QSAR models were evaluated using following statistical measures: *n*, number of
observations (molecules); Vn, number of descriptors; *k*, number of nearest neighbours; *q*2, cross validated *r*2 (by the leave-one-out method); pred_r2, predicted *r*2 for the external test set; Zscore, the Z score calculated by *q*2 in the randomization test; best_ran_q2, the highest *q*2 value in the randomization test; and R, the statistical significance parameter obtained by the randomization test.

RESULT AND DISCUSSION

It is known that the CoMFA method provides significant value in terms of a new molecule design, when contours of the PLS coefficients are visualized for the set of molecules. Similarly, the kNN-MFA models provide direction for the design of new molecules in a rather convenient way. 3D QSAR studies of 2-substituted halogenobenzimidazoles as a novel anti mycobacterial activity were performed, using the software V- LIFE MDS 3.0.

Model 1

k Nearest Neighbor=3;n = 28 ;Degree of freedom = 24 ;q2 = 0.724; q2_se = 0.2164;Predr^2 = 0.6491 ;pred_r2se = 0.362

Another statistically significant model model 2A was obtained for antimycobacterial activity through SW k-NN MFA justified by internal and external predictivity of the model as 64 % (q2=0.6491) and 43% (pred_r2=0.362) respectively.

Model 2

k Nearest Neighbor= 3;n = 28;Degree of freedom = 23 ;q2 = 0.769 ;q2_se = 0.3261 Predr2 = 0.5357; pred_r2se = 0.7130

For tuberculosis activity, out of three developed models, two were found to be statistically significant justified by the values of q2 that explained 98% internal predictivity for model 3

Model 3

k Nearest Neighbor= 3; n = 28 ;Degree of freedom = 23 ;q2 = 0.8215 ;q2_se = 0.2791 Predr2 = 0.6375; pred_r2se = 0.8833

Model 3C, the best model developed through k-NN MFA had a value of q2=0.8215 and that of pred_r2=0.6375 that explained 82% of total variance (internal predictivity) and 63% predictive power for the external test set.

Model 4

k Nearest Neighbor= 4;n = 28 ;Degree of freedom = 23 ;q2 = 0.8968 ;q2_se = 0.5441 Predr2 = 0.7738; pred_r2se = 0.4852

Another statistically significant model, model 4 was generated for activity against antimyobacterial cells having a value of q2=0.8968 and that of pred_r2=0. 0.7738 ,Plot of the k-NN MFA which shows the relative position and ranges of the corresponding important electrostatic/steric fields in the model provides guidelines for new molecule design as follows:

1) Electrostatic field

(a) Negative range indicates that negative electrostatic potential is favorable for increase in the activity and hence a more electro negative substituent group is preferred in that region.

(b) Positive range indicates that positive electrostatic potential is favorable for increase in the activity and hence a less electronegative substituent group is preferred in this region.

2) Steric field

(a) Negative range indicates that negative steric potential is favorable for increase in the activity and hence less bulky substituent group is preferred in that region.

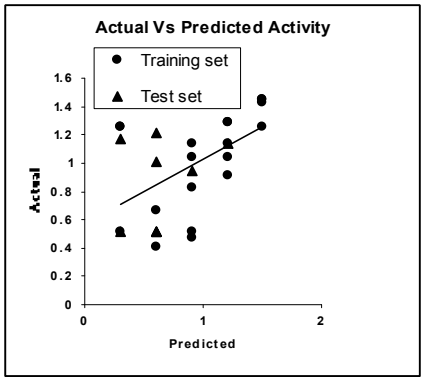
(b) Positive range indicates that positive steric potential is favorable for increase in the activity and hence more bulky substituent group is preferred in that region.

Model-A, the atom based alignment shows a *q2* (cross validated *r2*) of 0.92 with four descriptors namely S 231, S482 , E1853 and S651 . A non-cross-validated *r2* of 0.96, *F* value of 97.71and number nearest neighbors *k* of 2 were observed with this model. i.e all the values are proved statistically significant. The steric and electrostatic contributions were 25and 75 %, respectively and exhibited good external prediction with *r2 pred* of 0.6215. Statistical significance of the model indicated by *Z score* value of 4.432 and α of >0.0001.

Model-B, the kNN-MFA model generated from template based alignment showed *q2* (cross validated *r2*) of 0.7371 with four descriptors namely S762, E970, S643 and S1531. A non-crossvalidated *r2* of 0.73, *F* value of 34.00 and number nearest neighbors *k* of 2 were observed with this model. The steric and electrostatic contributions were 62 and 38%, respectively and exhibited good external prediction with *r2 pred* of 0.63. Statistical significance of the model indicated by *Z score* value of 3.66 and α of >0.001.

kNN-MFA electrostatic points (blue) indicate that areas in which electronegative subsistent might have a favorable (negative range) or disfavorable (positive range) effect on the activity of an inhibitor. results in increase of inhibitory activity of compounds. An important feature of kNN-MFA model is that the electrostatic points are dominated by the region disfavorable to positive charges.

Actual and Predicted values for model-1 and model-2 for 2D QSAR analysis



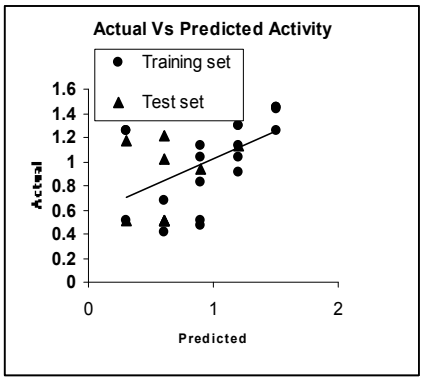
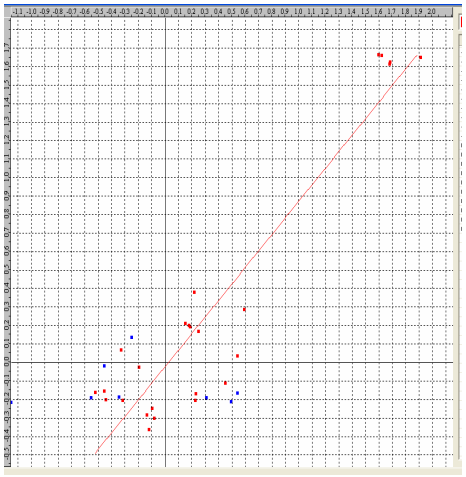


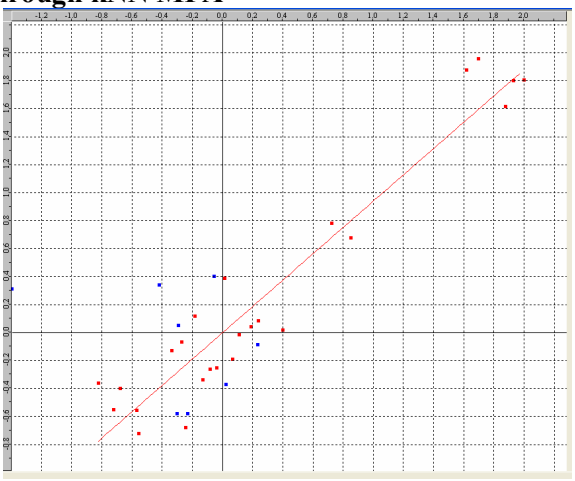
Table- Actual and predicted activities of training and test set compounds in statistically significant models

S.No.	Molecule	Actual Activity	Predicted Activity-1	Predicted Activity-2
1.	2a	1.505150	1.450495	1.450495
2.	2b	1.204120	1.292837	1.292837
3.	2c	0.903090	1.135178	1.135178
4.	2d	0.903090	1.036642	1.036642
5.	2e	0.903090	0.826724	0.826724
6.	2f	0.602060	0.669066	0.669066
7.	2g	0.903090	0.511407	0.511407
8.	2h	0.602060	0.412871	0.412871
9.	2i	1.505150	1.450495	1.450495
10.	2j	1.204120	1.292837	1.292837
11.	2k	1.204120	1.135178	1.135178
12.	2l	1.204120	1.036642	1.036642
13.	3a	1.505150	1.430788	1.430788
14.	3b	0.903090	0.943552	0.943552
15.	3c	0.602060	1.013997	1.013997
16.	3d	1.204120	0.91546	0.91546
17.	5a	0.301030	0.510417	0.510417
18.	5b	0.301030	0.510417	0.510417
19.	5c	0.602060	0.510417	0.510417
20.	5d	0.301030	1.253422	1.253422
21.	5e	1.505150	1.253422	1.253422
22.	5f	0.602060	0.510417	0.510417
23.	5g	0.301030	1.253422	1.253422
24.	5h	0.301030	1.174593	1.174593
25.	5i	0.602060	0.510417	0.510417
26.	5j	0.602060	1.214008	1.214008
27.	5k	1.204120	1.135178	1.135178
28.	5l	0.903090	0.471002	0.471002

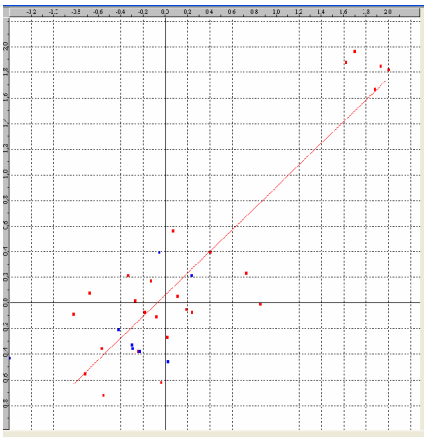
Fig: Graph of observed v/s predicted activities of statistically significant models Obtained through kNN MFA



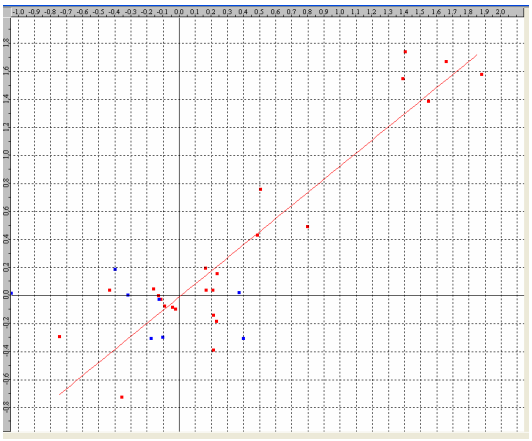
Model 1



Model 2

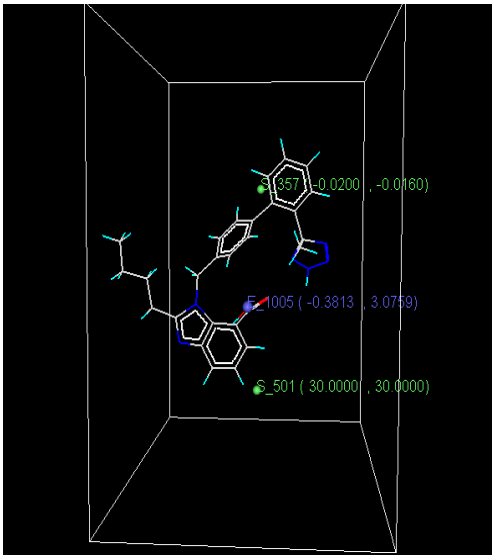


Model 3

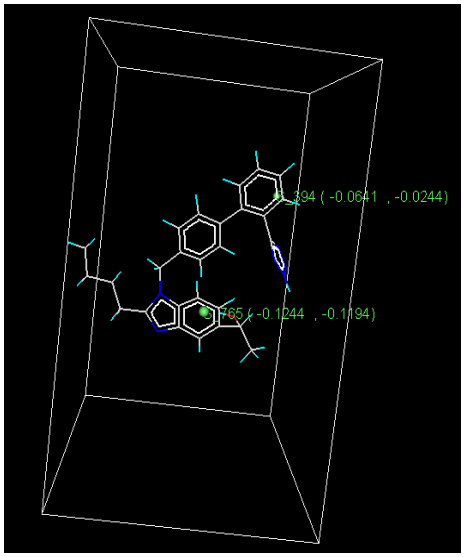


Model 4

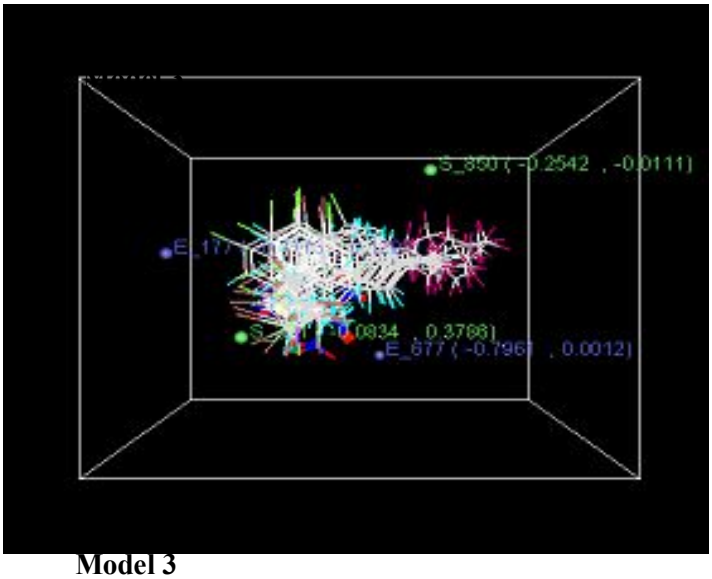
Alignment of 3D structures of the series of 2- Substituted halobenzimidazoles with Steric and electrostatic involvement



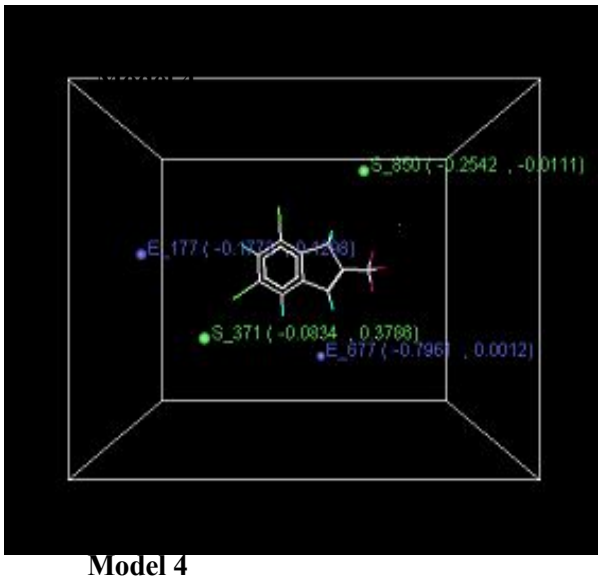
Model 1



Model 2



Model 3



Model 4

Table- Actual Activity, Predicted Activity and Residual values of test set Compounds

S.No.	Molecule	Actual Activity	Predicted Activity-1	Predicted Activity-2
1	5d	0.301030	1.253422	1.253422
2.	5e	1.505150	1.253422	1.253422
3.	5f	0.602060	0.510417	0.510417
4.	5g	0.301030	1.253422	1.253422
5.	5h	0.301030	1.174593	1.174593
6	5i	0.602060	0.510417	0.510417
7.	5j	0.602060	1.214008	1.214008
8.	5k	1.204120	1.135178	1.135178
9	5l	0.903090	0.471002	0.471002

REFERENCES

[1] Snider D.E.,Raviglione M., Kochi A., in B. Bloom (Ed.), Tuberculosis Pathogenesis, Protection and Control: Global Burden of Tuberculosis. ASM Press, Washington, DC, 1994, 3.

[2] Ballell L., Field R.A., Duncan.K., Young R.J., Antimicrob. Agents Chemother. 2005, (6) 49, 2153–2163.

[3] (a). Neu H.C., Science 1992,257, 1064–1072 (b) Spigelman M.K., New tuberculosis therapeutics a growing pipeline, J. Infect. Dis. 2007, 196,828–834. (c) Cegielski J.P., Chin D.P., Espinal M.A., Frieden T.R., Rodriquez C. R., E.A. Talbot D.E. , Weil R,. Zaleskis M.C., The global tuberculosis situation. Progress and problems in the 20th century, prospects for the 21st century, Infect. Dis. Clin. North Am. 2002,161, 1–58.

[4] Walczak K., Gondela A., Suwinski J., Eur. J. Med. Chem. 2004,39,849–853

[5] (a) Babaoglu K., Page M.A., V.C. Jones, McNeil M.R., Dong C.,. Naismith J.H.,. Lee R.E., Bioorg. Med. Chem. Lett.2003,13 (19) 3227–3230. (b) Shiradkar M.R., Murahari K.K., Gangadasu H.R., Suresh T., Kalyan C.A., Panchal D., Kaur R., Burange P., Ghogare J., Mokale V., Raut M., Bioorg. Med. Chem. Lett. 2007 ,15,3997–4008.

[6] (a) Shiradkar M.R., Bhandari S.V., Kale R.P., Laghate A., Rathi A.Asian J. Chem.
2006,18 (4) 2700–2704. (b) Shiradkar M.R., Gorentia Venkata S., Dasari V., Tatikonda S., Akula K.C, Shah R., Eur. J. Med. Chem. 2007, 42 (6) ,807–816;

[7]. Kazimierczuk., Andrzejewska M., Kaustova J.,Klimesova V., Eur. J. Med. Chem. 2005, 40,203-208.

[8] .Cramer R. D.,. Patterson D. E, Bunce J. D.J. Am. Chem. Soc., 1988,110, 5959.

[9] Klebe G., Abraham U., Mietzner T., J. Med. Chem., 1994,37,24.

[10]. Baroni M., Costantino G., Cruciani G., Riganelli D., Valigi R., Clementi S., An Advanced Chemometric Tool for Handling 3D-QSAR Problems, Quant. Struct.-Act. Relat., 1993,12, 9.

[11] Cho S. J., and Tropsha A., J. Med. Chem., 1995,**38**, 1060.

[12]. V-life MDS 3.0 vlife science technologies pvt.Ltd,1 akshay residency, plot 50, Anand park, Aundh, Pune- 411007

[10]. Chakraborti A.K., Gopalakrishnan B.,. Sobhia M.E, Malde A., Eur. J. Med. Chem. 2003,38, 975-982.

[11].W.H.O Report Global Tuberculosis Control, Surveillance, planning, Financing, Geneva, 2004.

[12] .Coi A.,. Fiamingo F.L, Livi O., Bioorg. Med. Chem. 2009,17, 319–325.

[13] .Golbraikh A., Tropsha A., J. Comp. Aid. Molec. Des. 200216, 357-369.

[14] Golbraikh A., J. Chem. Inf. Comput.Sci. 2000,40, 414-425.

[15]. Guha R., Serra J.R., Jurs P.C., J. Mol. Graph. Model 2004, 23, 1-14.

[16] Golbraikh A. and Tropsha A., J. Chem. Inf. Comput. Sci., 2003**43**, 144

[17] Sharaf M. A., Illman D. L., Kowalski B. R., Chemometrics, Wiley, New York, 1986.

[18]. J. Holland, Adaptation in Natural and Artificial Systems, University of Michigan
Press, 1975.

*****