

3D-QSAR STUDY OF BENZYLIDENE DERIVATIVES AS SELECTIVE CYCLOOXYGENASE-2-INHIBITORS.

J. D. FEGADE*, S. S. RANE, R. Y. CHAUDHARI, V. R. PATIL

*Department of Pharmaceutical Chemistry, TVES's Hon. Madhukarrao Chaudhari
College of Pharmacy, Nehru Vidyanagar, Savda Road, Faizpur, (MS), India-425
503.*

The urgent need for novel cyclooxygenase-2-inhibitors has provided an impetus for understanding the structural requisites at the molecular level. Towards this objective, k-nearest neighbor molecular field analyses of 19 COX-2 inhibitors were performed. The 3D-QSAR studies were based on lowest energy conformer of most active compound (**TR-11**), employing atom and template based alignment methods. Results generated from the atom-based model was found superior ($r^2=0.96$, $q^2=0.92$), to those obtained by the template-based model ($r^2=0.90$, $q^2=0.71$), with four components. The predictive ability of the models was validated using structurally diversified test set of four compounds that had not been included in a preliminary training set of 15 compounds. The predictive r^2 value for atom-based kNN-MFA model was 0.75, while the corresponding predictive r^2 value for template-based kNN-MFA model was 0.68. The potency of the benzylidene derivatives was interpreted based on kNN-MFA steric and electrostatic point distribution maps. The 3D-QSAR model (**Model-A**) was found to accurately predict the Cyclooxygenase-2 inhibitory activity of structurally diverse test set compounds and to yield reliable clues for further optimization of the benzylidene derivatives in the data set.

(Received January 27, 2009; accepted February 13, 2009)

Keywords: 3D-QSAR, kNN-MFA, COX-2, Benzylidene derivatives, PLS

1. Introduction

Nonsteroidal anti-inflammatory drugs (NSAID) are among the most frequently prescribed medications being the drugs of the first choice for treatment of the inflammatory and rheumatic diseases. The common mechanism of NSAIDs involves the nonselective inhibition of cyclooxygenases (COXs) thereby preventing the biosynthesis of prostaglandins (PG) which are the important lipid mediators of inflammation as well as numerous homeostatic physiological functions.¹ As it is now well appreciated, COXs exist in two isoforms, namely COX-1 and COX-2,² while the existence of a third isoform (COX-3) is still into debate. In general terms, COX-1 is the constitutive isoform providing normal production of PGs having roles in homeostasis and gastroprotection, whereas the COX-2 is induced by proinflammatory stimuli at inflammatory sites.³ The discovery of inducible COX-2 at sites of inflammation led to the development of selective COX-2 inhibitors with the hope of diminished gastrointestinal side effects associated with traditional NSAIDs.^{4,5} However, recent studies have shown that COX-2 inhibitors are associated with increased thromboembolic phenomena in specific patient populations such as cardiovascular disease patients challenging the benefits of selective COX-2 inhibition.⁶⁻⁸ Moreover, there is currently no clear evidence that COX-2 inhibitors represent an independent risk factor in patients at low demographic risk of cardiovascular diseases and therefore, clinical rationale for developing compounds with selective COX-2 inhibition still remains to be established.^{8,9} Meantime,

*Corresponding author: jitufegade@gmail.com

considerable interest in the further potential clinical utilities of COX-2 inhibitors has emerged.¹⁰⁻¹² Recent studies indicating the place of COX-2 inhibitors in cancer chemotherapy and neurological diseases such as Alzheimer's and Parkinson diseases still continues to attract investigations on development of COX-2 inhibitors.^{13,14}

There are at least four mechanisms of COX inhibition competitive, tight binding/time dependent, weak binding/ mixed and covalent binding.¹⁵ Some NSAIDs inhibits COX- 1 and COX-2 by similar mechanisms, whereas other NSAIDs have distinct inhibition mechanisms for each isoform. For example, celecoxib has been reported as a reversible competitive inhibitor of COX-1, while demonstrating time-dependent irreversible inhibition of COX-2.¹⁶

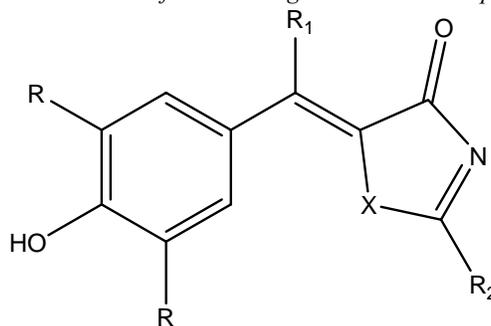
In the present study, we have applied k-nearest neighbor molecular field analysis (kNN-MFA),¹⁷⁻¹⁹ 3D-QSAR methodologies to the Benzylidene,²⁰ scaffold of cyclooxygenase-2 inhibitors. The predictiveness of each of our optimized model was evaluated using test set of four compounds that were not included in the model. The distribution point maps derived from kNN-MFA 3D-QSAR models permitted an understanding of the steric and electrostatic requirements for ligand binding.

2. Results and discussion

Results

The kNN-MFA technique was used to derive 3D-QSAR model for Benzylidene derivatives which inhibits cyclooxygenase-2. The in vitro inhibitory activity (IC_{50} values) in μM , were converted to pIC_{50} , was used as dependant variable. Relative alignment of all the energy minimized molecules was then carried out by using two techniques namely atom and template based for better results and better assessment between both. The kNN-MFA models were generated by using training set of 15 compounds (Table 1). The 3D-QSAR models were validated using a test set of 04 compounds (Table 1).

Table 1. Structures of the Training and Test Set Compounds



Training Set Compounds					
Sr.No.	R	R ₁	R ₂	X	IC ₅₀ (μM) [#]
TR 1	<i>t</i> -Bu	H	NHOH	S	1.5
TR 2	<i>t</i> -Bu	H	NHOET	S	3.2
TR 3	<i>t</i> -Bu	H	NHO-allyl	S	2.7
TR 4	<i>t</i> -Bu	H	NHC(=NH)NH ₂	S	21
TR 5	<i>t</i> -Bu	H	NCH ₃ OCH ₃	S	18
TR 6	<i>t</i> -Bu	H	NHCN	S	55
TR 7	<i>i</i> -Pr	H	NHC(=NH)NH ₂	S	1.8
TR 8	<i>t</i> -Bu	H	SCH ₃	S	1.8
TR 9	<i>t</i> -Bu	H	NHOET	O	79
TR 10	<i>t</i> -Bu	H	NHO-allyl	O	26

TR 11	<i>t</i> -Bu	H	SH	NH	0.26
TR 12	<i>t</i> -Bu	H	OH	S	4.7
TR 13	<i>t</i> -Bu	H	OH	O	34
TR 14	<i>t</i> -Bu	CH ₃	NHC(=NH)NH ₂	S	12
TR 15	<i>t</i> -Bu	H	SH	S	1.5
Test Set Compounds					
T 1	<i>t</i> -Bu	H	NHOCH ₃	S	1.5
T 2	<i>i</i> -Pr	H	NHOCH ₃	S	3.2
T 3	<i>t</i> -Bu	H	NHC(=NH)NH ₂	O	2.7
T 4	<i>t</i> -Bu	H	NHC(=NH)NH ₂	NCH ₃	21

#Observed activity pIC50= (-logIC50 (μM))

External predictions were used to select best model. Results from PLS analysis are reported in Table 2.

Table 2. Results of kNN-MFA Models by PLS analysis

Parameter	Model A	Model-B
n	15	15
r ²	0.9638	0.9000
q ²	0.9186	0.7129
F test	97.7116	54.0052
r ² _{pred}	0.75	0.68
ZScore	2.84706	2.66063
α	>0.0001	>0.001
k/vn	2/4	2/4
Components (Contribution)	E_709 (23), E_852(31), E_1006(29), S_701(17)	E_641(-35), E_836 (23), S_513 (-26), S_801(16)

To ascertain the true predictivity of the model, applying leave-one-out method of cross validation using weighted k-nearest neighbor was performed for all the analysis.

Model-A, the atom based alignment shows a q^2 (cross validated r^2) of 0.92 with four descriptors namely E709, E852, E1006 and S701. A non-cross-validated r^2 of 0.96, F value of 97.71 and 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 17 and 83 %, respectively and exhibited good external prediction with r^2_{pred} of 0.75. Statistical significance of the model indicated by Z score value of 2.847 and α of >0.0001.

Model-B, the kNN-MFA model generated from template based alignment showed q^2 (cross validated r^2) of 0.71 with four descriptors namely E641, E836, S513 and S801. A non-cross-validated r^2 of 0.90, F value of 54.00 and number nearest neighbors k of 2 were observed with this model. The steric and electrostatic contributions were 58 and 42 %, respectively and exhibited good external prediction with r^2_{pred} of 0.68. Statistical significance of the model indicated by Z score value of 2.66 and α of >0.001.

With the view of all above also based on the predictive (Table 4 and 5) ability of two kNN-MFA models, analysis A, the model generated with atom based alignment and four components exhibits good predictive ability r^2_{pred} of 0.75.

Table 4. Actual Activity, Predicted Activity and Residual values of Training set Compounds.

Sr. No	pIC ₅₀	pIC ₅₀ ^a	pIC ₅₀ ^b	Residuals ^c
TR 1	-0.176	-0.116	-0.876	-0.7
TR 2	-0.505	-0.554	-0.823	-0.318
TR 3	-0.431	-0.387	-0.925	-0.494
TR 4	-1.322	-1.383	-1.272	0.05
TR 5	-1.255	-1.558	-1.161	0.094
TR 6	-1.740	-1.717	-1.061	0.679
TR 7	-0.255	-0.173	-0.308	-0.053
TR 8	-0.255	-0.278	-0.308	-0.053
TR 9	-1.898	-1.548	-1.407	0.491
TR 10	-1.494	-1.465	-1.297	0.197
TR 11	0.585	0.374	0.065	-0.52
TR 12	-0.672	-1.070	-0.308	0.364
TR 13	-1.531	-1.059	-1.478	0.053
TR 14	-1.079	-1.256	-1.419	-0.34
TR 15	-0.176	-0.010	0.379	0.555

Observed activity pIC₅₀= (-logIC₅₀ (μM)), ^apredicted activity for atom based alignment (Model - A), ^bpredicted activity for template based alignment (Model -B), ^cresiduals considering best model-A

Table 5. Actual Activity, Predicted Activity and Residual values of test set Compounds.

Sr No	pIC ₅₀	pIC ₅₀ ^a	pIC ₅₀ ^b	Residuals ^c
T 1	-0.230	-1.388	-0.133	-0.097
T 2	0.244	-1.047	-1.034	1.278
T 3	-0.707	-1.478	-3.265	2.558
T 4	-1.908	-0.308	-1.095	-0.813

Observed activity pIC₅₀= (-logIC₅₀ (μM)), ^apredicted activity for atom based alignment (Model - A), ^bpredicted activity for template based alignment (Model -B), ^cresiduals considering best model-A

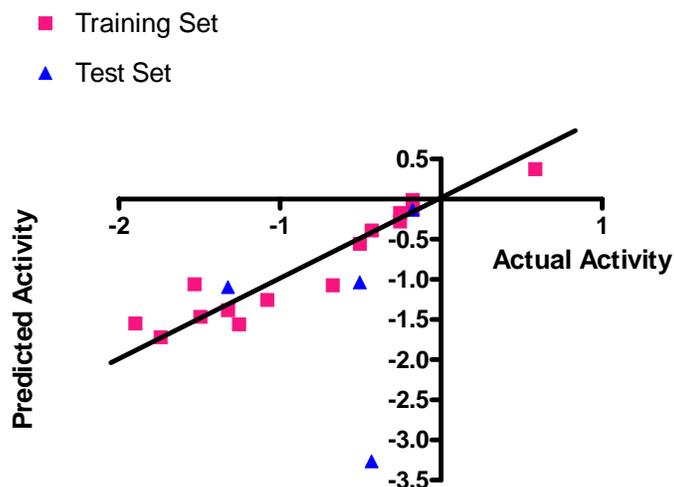


Fig. 1. A graph of actual verses predicted activities of the training and test set molecules from atom-based alignment (Model-A).

Negative values of electrostatic field descriptors (blue) indicates that negative electronic potential is required to increase activity and more electronegative substituents group is preferred in that position, positive range indicates that group that imparting positive electrostatic potential is favorable for activity so less electronegative group is preferred in that region.

Steric descriptors (green), negative range indicates that negative steric potential is favorable for activity and less bulky substituents group is preferred in that region, Positive value of steric descriptors reveals that positive steric potential is favorable for increase in activity and more bulky group is preferred in that region.

Discussion

Three-dimensional quantitative structure activity relationship by *k*-nearest neighbor molecular field analysis were performed on series of 19 molecules belonging to benzylidene derivatives which inhibits cyclooxygenase-2. The lowest energy conformation of most active compound **TR-11** was used as template for 3D-QSAR studies.

The alignment of the compounds is one of the critical inputs for kNN-MFA studies. The alignments define the putative pharmacophore for the series of ligands. In the present study we have aligned the ligands onto a template structure compound **TR-11** using two alignment rules. These alignments were validated using kNN-MFA studies. The kNN-MFA models were validated by predicting the activity of external test set.

The PLS analyses on two alignments (atom and template based) are reported in Table 2. The atom based alignment shows better r^2 values than template based alignment. This indicates that all ligands have to be superimposed by the template structure used for alignment. This superimposition produced a good external prediction. The template based fitting of ligands did not improve the predictiveness of the model. This shows that exact superimposition of ligands is essential for good predictions.

A Cross-validation analysis was also performed by applying leave-one-out technique using weighted *k*-nearest neighbor method.

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. The points which contribute to the kNN-MFA model in data set are termed as the distribution point map. The range of property values for the chosen points may aid in the design of new potent molecules. The range is based on the variation of the field values at the chosen points using the most active molecule and its nearest neighbor set. These points show regions that are important for variation in activity of these data sets.

The 3D-kNN-MFA distribution point maps are shown in Figure 2. Align molecules are displayed in map to assist in illustration. kNN-MFA steric points (green) indicate that areas in which steric bulk subsistent might have a favorable (positive range) or unfavorable (negative range) effect on the activity of an inhibitor. A positive contribution of steric descriptor near the substituent at 2 position of thiazolone/ oxazolone ring indicates that sterically bulky group is required in this reign. This can be also seen with compound **T-4** and **TR-3**, in which replacing sterically less bulky hydrozonoformamide group with sterically bulky allyloxyamino group, results in increase of inhibitory activity of compound. In fact most of favorable regions are found surrounding the thiazolone/ oxazolone ring and R₄ substituents, which define open pocket of the active site of cyclooxygenase-2, indicating a possibility of possible structural modifications of the functional COX-2 inhibitors may prove for improve inhibitory activity.

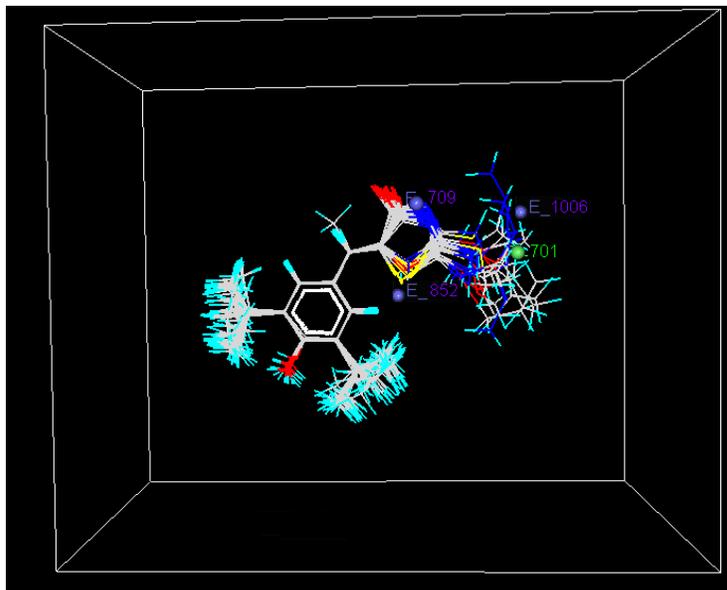


Fig. 2. Distribution of chosen points in kNN-MFA method for atom-based alignment (Model-A).

kNN-MFA electrostatic points (blue) indicate that areas in which electronegative substituent might have a favorable (negative range) or unfavorable (positive range) effect on the activity of an inhibitor. A positive contribution of electrostatic descriptors near the thiazolone/oxazolone and R₄ substituents indicates less electronegative substituent are favorable in that region, that effect can be seen with compounds **TR-9** and **TR-11**, in which replacing more electronegative oxygen with nitrogen and ethoxyamino group with thiol group, respectively, 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. Such a region is mostly observed surrounding the thiazolone/oxazolone rings, functional mimetic in the catalytic binding cavity, indicating that compounds with less electron density are preferred in the active binding pocket.

Conclusions

The 3D-QSAR study of 19 benzylidene derivatives which inhibits cyclooxygenase-2 was carried out using kNN-MFA method with the abet of atom based and template base methods. We find that the atom based alignment shows this superimposition produced a good external prediction. The kNN-MFA model obtained from atom based alignment having better r^2 values than template based alignment. This indicates that all ligands have to be superimposed by the template structure used for alignment. Also it showed good correlation with biological and predictive ability. Steric and electrostatic fields were found important for cyclooxygenase-2 inhibitory activity as exemplified by the higher predictive power of the kNN-MFA model. The results obtained from the 3D-QSAR models were found to accurately predict the cyclooxygenase-2 inhibitory activity of structurally diverse test set of compounds and to yield reliable clues for further optimization of benzylidene derivatives in the data set.

Experimental Section

Data Set and Biological Activity:

The training and test sets used to comprise a series of benzylidene derivatives, which inhibits cyclooxygenase-2 (COX-2) enzyme. The IC₅₀ values, in μM , were converted to pIC₅₀ (-log IC₅₀) values, which were used as dependant variables in the 3D-QSAR study. Training set (15

compounds) and the test set (04 compounds) were selected by considering the fact that the test set compounds represents structural diversity and a range of biological activities similar to that of training set. Compounds in test set allowed us to use one test compounds per three training compounds thus resulting in more rigorous validation of the training model. In addition, a wide range of structural diversity of compounds in the test set permit us to evaluate the extrapolative accuracy of the QSAR models. The structures of the compounds in the training and test sets are shown in Table 1.

Molecular Modeling:

The 3D-QSAR computations were carried out using VLife Molecular Design Suite.²¹ All the molecules were drawn and converted to 3D structures in draw module of VLife MDS. Energy minimization were performed using the MMFF94,²² force field and Gasteiger-marsili,²³ charges followed by AM-1 (Austin Model-1) Hamiltonian method available in MOPAC module with the convergence criterion 0.001 kcal/mol Å.

Alignment Rules:

The position of each atom is important for kNN-MFA because the descriptors were calculated based on the 3D space grid. Thus, the method to determine the conformation of each molecule and the way to align molecules together are two sensitive input parameters to build reasonable model.²⁴ In the present study two deferent alignment rules were adopted.

Alignment 1:

This alignment involves RMS fitting (atom-based) of heavy atoms of ligands. The compounds were fitted to the template molecule, compound **TR-11**. Figure 1, The alignment maximizes the overlap of the heavy atoms.

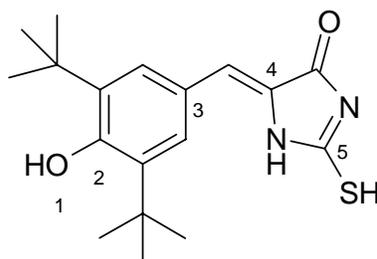


Fig. 1. compound TR-11 used as a template for atom based alignment. The atoms for alignment are numbered 1-5. (Note- the atom numbering does not follow IUPAC rules). a) Oxygen atom of hydroxyl group (1). b) C₁ atom of aromatic ring (2). c) C₄ atom of aromatic ring (3). d) C₃ atom of heterocyclic ring (4) e) C₂ atom of heterocyclic ring (5).

Alignment 2:

Alignment of the molecules was carried out by flexible fitting of the atoms of the ligands to the template molecule, compound **TR-11**.

k-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 neighbors 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 follows²⁵: (1) calculate distances between an unknown object (u) and all the objects in the training set; (2) select k objects from the training set most similar to object u , according to the calculated distances and (3) 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. The variables and optimal k values were chosen using different variable selection methods as described below.

kNN-MFA 3D-QSAR Models:

To derive the kNN-MFA descriptor fields, a 3D cubic lattice grid in x , y and z directions, was created to encompass the aligned molecules. kNN-MFA descriptors were calculated using an

sp³ carbon probe atom with a van der Waals radius of 1.52 Å and a charge of +1.0 to generate steric field energies and electrostatic fields with the distance dependant dielectric at each lattice point. The steric and electrostatic energy values were truncated at a default value of 30 kcal/mol.

PLS analysis:

The partial least squares method (PLS),²⁶⁻²⁸ was used to derive a linear relationship and cross-validation was performed using leave-one out method,^{29,30} to check consistency and predictiveness.

Cross-Validation Using Weighted k-Nearest Neighbor:

The standard leave-one-out procedure was implemented as summarized as follows.

(1) A molecule in the training set was eliminated, and its biological activity was predicted as the weighted average activity of the *k* most similar molecules (eq 1). The similarities were evaluated as the inverse of Euclidean distances between molecules (eq 2) using only the subset of descriptors corresponding to the current trial solution.

$$w_i = \frac{\exp(-d_j)}{\sum \exp(-d_j)}$$

k-nearest neighbors

$$\hat{y}_i = \sum w_i y_i \quad (1)$$

$$d_{ij} = \left[\sum_{k=1}^N (X_{i,k} - X_{j,k})^2 \right]^{1/2} \quad (2)$$

(2) Step 1 was repeated until every molecule in the training set has been eliminated and its activity predicted once. (3) The cross-validated *r*²(*q*₂) value was calculated using eq 3, where *y_i* and *y[^]_i* are the actual and predicted activities of the *i*th molecule, respectively, and *y*_{mean} is the average activity of all molecules in the training set. Both summations are over all molecules in the training set. Since the calculation of the pairwise molecular similarities, and hence the predictions, were based upon the current trial solution, the *q*₂ obtained is indicative of the predictive power of the current kNN-MFA model.

$$q^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{\text{mean}})^2} \quad (3)$$

(4) Steps 1-3 were repeated for *k* 2, 3, 4, etc. Formally, the upper limit of *k* is the total number of molecules in the data set. However, the best value has been empirically found to lie between 1 and 5. The *k* value that led to the highest *q*₂ value was chosen for the current kNN-MFA model.³¹

External Validation:

The following procedure was applied for external validation.

(1) Predict the biological activity of a molecule in the test set as the weighted average activity of the *k* most similar molecules in the training set (eq 1). The similarities were evaluated as the inverse of Euclidean distances between molecules (eq 2) as calculated using the descriptors determined by the current model.

(2) Step 1 was repeated for every molecule in the test set.

(3) The predicted *r*² (pred•*r*²) value was calculated using eq 4, where *y_i* and *y[^]_i* are the actual and predicted activities of the *i*th molecule in test set, respectively, and *y*_{mean} 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 is indicative of the predictive power of the current kNN-MFA model for external test set.

$$\sum (y_i - \hat{y}_i)^2$$

$$\text{pred}_r^2 = 1 - \frac{\sum (y_i - y_{\text{mean}})^2}{\dots} \quad (4)$$

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 Z score.^{31,32}

Evaluation of the QSAR Models:

The QSAR models were evaluated using following statistical measures: n , number of observations (molecules); V_n , number of descriptors; k , number of nearest neighbors; q^2 , cross-validated r^2 (by the leave-one-out method); pred_r^2 , predicted r^2 for the external test set; Z score, the Z score calculated by q^2 in the randomization test; $\text{best}_{\text{ran}} q^2$, the highest q^2 value in the randomization test; and R, the statistical significance parameter obtained by the randomization test.

For predicting the activity of a molecule, regression methods use the following equation

$$\text{Activity} = C_0 + C_1 D_1 + C_2 D_2 + \dots + C_N D_N$$

where C_i 's are coefficients and D_i 's are descriptors.

In the case of the kNN-MFA method, the activity of a molecule is predicted using

$$\text{Activity} = C_1 A_1 + C_2 A_2 + \dots + C_k A_k$$

where C_i 's are weights and A_i 's are activities of the k-nearest neighbors in the training set. The nearest neighbors of any molecule are obtained from calculating the distance between the descriptors selected from various variable selection methods, described above. Thus, kNN-MFA prediction uses an interpolative method, and hence predicted activities of new designed molecules will be within the range of activities of molecules in training set. Since the kNN method is based on distances of descriptors, their interpretation is quite difficult compared to the regression models. Although several models are generated by the kNN-MFA method, the time required for obtaining results is significantly more than for the CoMFA method.

References

- [1] J. R. Vane, R. M. Botting, *Am J Med.* **104** (3A), 2S-8S (1998); discussion 21S-22S.
- [2] J.R. Vane, Y.S. Bakhle, R.M. Botting, *Annu Rev Pharmacol Toxicol.* **38**, 97-120 (1998).
- [3] D.L. Simmons, R.M. Botting, T. Hla, *Pharmacol Rev.* **56** (3), 387-437 (2004).
- [4] J. van Ryn, G. Trummlitz, M. Pairet, *Curr Med Chem.* **7** (11), 1145-1161, (2000).
- [5] G. Coruzzi, N. Venturi, S. Spaggiari, *Acta Biomed.* **78** (2), 96-110, (2007).
- [6] J.A. Mitchell, T.D. Warner, *Nat Rev Drug Discov.* **5** (1), 75-86, (2006).
- [7] G.A. Fitzgerald, *N Engl J Med.* **351** (17), 1709-1711, (2004).
- [8] C.D. Funk, G.A. Fitzgerald, *J Cardiovasc Pharmacol.* **50** (5), 470-479, (2007).
- [9] G.A. FitzGerald, *Nat Rev Drug Discov.* **2** (11), 879-890, (2003).
- [10] P. Patrignani, S. Tacconelli, M.G. Sciuilli, M.L. Capone, *Brain Res Brain Res Rev.* **48**(2), 352-359, (2005).
- [11] M. Breinig, P. Schirmacher, M.A. Kern, *Curr Pharm Des.* **13**(32), 3305 (2007).
- [12] G. Aparicio Gallego, S. Diaz Prado, P. Jimenez Fonseca, R. Garcia Campelo, J. Cassinello Espinosa, L.M. Anton Aparicio, *Clin Transl Oncol.* **9**(11), 694 (2007).
- [13] W.A. van Gool, P.S. Aisen, P. Eikelenboom, *J Neurol.* **250** (7), 788 (2003).
- [14] P.S. Aisen, K.A. Schafer, M. Grundman, E. Pfeiffer, M. Sano, K.L. Davis, M.R. Farlow, S. Jin, R.G. Thomas, L.J. Thal, *Jama.* **289**(21), 2819, (2003).
- [15] J. K. Gierse, C. M. Koboldt, M. C. Walker, K. Seibert, P.C. Isakson, *Biochem.*

- J., **339**, 607 (1999).
- [16] M.C. Walker, R. G. Kurumbail, J. R. Kiefer, K. T. Moreland, C. M. Koboldt, P. C. Isakson, K. Seibert, J. K. Gierse, *Biochem. J.*, **357**, 709, (2001).
- [17] W. Zheng, A. Tropsha, *J. Chem. Inf. Comput. Sci.* **40**, 185-194, (2000).
- [18] M. Shen, A. Le Tiran, Y. Xiao, A. Golbraikh, H. Kohn, A. Tropsha *J. Med. Chem.* **45**, 2811-2823, (2002).
- [19] S. Ajmani, K. Jadav, S. A. Kulkarni, *J. Chem. Inf. Model.* (2006).
- [20] Y. Song, D. T. Connor, R. Doubleday, R. J. Sorenson, A. D. Sercel, P. C. Unangst, B. D. Roth, R.B. Gilbertsen, K. Chan, D. J. Schrier, A. Guglietta, D.A. Bornemeier, R. D. Dyer. *J. Med. Chem.* **42**, 1151-1160, (1999).
- [21] VLifeMDS2.0; Molecular Design Suite (Evaluation Version), Vlife Sciences Technologies Pvt. Ltd., Pune, India, (2004) www.vlifesciences.com.
- [22] T. A. Halgren, R. B. Nachbar, *J. Comput. Chem.* **17**, 587 (1996).
- [23] J. Gasteiger, M. Marsili, *Tetrahedron.* **36**, 3219-3228, (1980).
- [24] R. D. Cramer, D.E. Patterson, J. D. Bunce, *J. Am. Chem. Soc.* **110**, 5959-5967, (1988).
- [25] M. A. Sharaf, D. L. Illman, B. R. Kowalski, *Chemometrics*. Wiley, New York, (1986).
- [26] S. Wold, C. Albano, W. J. Dunn, U. Edlund, K. Esbenson, P. Gelad, S. Hellberg, W. Lindberg, M. Sjostrom, In; *Chemometrics*, Int. Ed. Kowalski, B., Reidel: Dordrecht, The Netherlands, (1984), 17.
- [27] W. J. Dunn, S. Wold, U. Edlund, S. Hellberg, J. Gasteiger. *Quant. Struct. Act. Relat. Chem. Bio.* **3**, 31, (1984).
- [28] P. Geladi, *J. Chemom.* **2**, 231, (1998).
- [29] S. Wold, *Technometrics.* **4**, 397, (1978).
- [30] R. D. Cramer III, J. D. Bunce, D.E. Patterson, *Quant. Struct. Act. Relat.* **7**, 18 (1998).
- [31] W. Zheng, A. Tropsha, *J. Chem. Inf. Comput. Sci.* **40**, 185-194, (2000).
- [32] M. Shen, A. LeTiran, Y. Xiao, A. Golbraikh, H. Kohn, A. Tropsha; *J. Med. Chem.* **45**, 2811-2823, (2002).