



King Saud University  
Arabian Journal of Chemistry

www.ksu.edu.sa  
www.sciencedirect.com



ORIGINAL ARTICLE

2nd Heterocyclic Update

# An approach to design potent anti-Alzheimer's agents by 3D-QSAR studies on fused 5,6-bicyclic heterocycles as $\gamma$ -secretase modulators using kNN–MFA methodology



Kamlendra Singh Bhadoriya <sup>a,\*</sup>, Mukesh C. Sharma <sup>b</sup>, Smita Sharma <sup>c</sup>,  
Shailesh V. Jain <sup>d</sup>, Mandar H. Avchar <sup>a</sup>

<sup>a</sup> Drug Design and Development Department, R.C. Patel Institute of Pharmaceutical Education and Research, Shirpur, District-Dhule, Maharashtra 425405, India

<sup>b</sup> School of Pharmacy, Devi Ahilya Vishwavidyalaya, Takshashila Campus, Ring Road, Indore, M.P. 452017, India

<sup>c</sup> Department of Chemistry, Choudhary Dilip Singh Kanya Mahavidyalaya, Bhind, M.P. 477001, India

<sup>d</sup> Department of Pharmaceutical Chemistry, Institute of Pharmacy, Nirma University, S.G. Highway, Ahmedabad 382481, Gujarat, India

Received 14 August 2012; accepted 1 February 2013

Available online 16 February 2013

## KEYWORDS

Alzheimer's disease;  
 $\gamma$ -Secretase modulators;  
Fused 5,6-bicyclic heterocycles;  
3D-QSAR;  
kNN–MFA;  
Hydrophobic and steric descriptors

**Abstract** Alzheimer's disease (AD) is a chronic neurodegenerative disease. Current therapies of AD are only symptomatic, therefore the need for the development of new therapies to treat Alzheimer's disease effectively. To achieve this objective quantitative structure–activity relationship (QSAR) studies were carried out as it provides the rationale for the changes in the structure to have more potent  $A\beta_{42}$  inhibitors or anti-Alzheimer's agents. Quantitative structure–activity relationship (QSAR) studies were carried out on a series of 34 fused 5,6-bicyclic heterocycles to investigate the structural requirements of their inhibitory activity against  $A\beta_{42}$ . The statistically significant best 3D-QSAR model having cross-validated squared correlation coefficient  $q^2 = 0.8457$  with external predictive ability of  $\text{pred}_r^2 = 0.7556$  was developed by SW-kNN. Developed kNN–MFA model highlighted the importance of shape of the molecules, i.e., hydrophobic and steric descriptors at the grid points H\_83 and S\_183, S\_227 for  $\gamma$ -secretase binding interaction. This model (3D) was found to yield reliable clues for further optimization of fused 5,6-bicyclic heterocycles in the data set. The information

\* Corresponding author. Tel.: +91 7879605622; fax: +91 2563251808.

E-mail addresses: kamlendra.bhadoriya@gmail.com, kamlendra.bhadoriya@yahoo.com (K.S. Bhadoriya).

Peer review under responsibility of King Saud University.



Production and hosting by Elsevier

rendered by the 3D-QSAR model may lead to a better understanding of the structural requirements of  $\gamma$ -secretase modulators and can also help in the design of novel potent  $\gamma$ -secretase modulators.

© 2013 Production and hosting by Elsevier B.V. on behalf of King Saud University.

## 1. Introduction

AD is a neurodegenerative disorder associated with difficulties in memory, judgment, abstraction, and language (Kim and Kim, 2008). More than 35 million people suffer from AD worldwide, with an estimated annual cost of over \$600 billion, and the AD population may increase to more than 115 million by the year 2050 according to a report from Alzheimers Disease International (Sun et al., 2012). Alzheimer's disease (AD) symptoms are dementia, apraxia, aphasia, depression, short attention span, visuospatial navigation deficits, anxiety and delusions (Harvey et al., 2011). The majority of AD cases are sporadic, with disease onset after 65 years of age (Iijima-Ando and Iijima, 2010). Key molecules involved in AD, include the presenilins, amyloid precursor protein, tau, and  $\beta$ -amyloid (Balaraman et al., 2006).

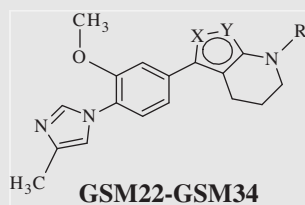
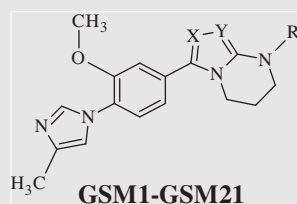
Alzheimer's disease (AD) is the biggest unmet medical need in neurology due to the lack of disease-modifying anti-Alzheimer's drugs (DMAADs) (Kreft et al., 2008). Over the last decade,  $\gamma$ -secretase emerged as a promising target for the treatment of AD (Xin et al., 2011). It has been postulated that modulation of  $\gamma$ -cleavage to favor the production of shorter fragments, while not affecting total A $\beta$  levels, might be a safe approach to a disease-modifying therapy (Fischer et al., 2011).  $\gamma$ -Secretase activity can be controlled by the inhibition of the active site of PS1 or by interference with complex assembly or substrate recognition, the latter resulting in allosteric modulation or inhibition. The allosteric mechanisms are particularly attractive targets for drug development (Narlawar et al., 2007), as they may produce shorter, soluble, and non-toxic peptides (e.g., A $\beta$ 36–A $\beta$ 40) instead of the highly insoluble and neurotoxic A $\beta$ 42, without interfering in the processing of Notch and other substrates (Caldwell et al., 2010).  $\gamma$ -Secretase modulators (GSMs) modulate the cleavage of the APP C-terminal fragment such as C-99 to decrease A $\beta$ 42 and increase the shorter A $\beta$  fragments (e.g., A $\beta$ 37/38) while not affecting the cleavage of other substrates such as Notch (Fischer et al., 2011). The fact that an increase as small as 30% in the levels of A $\beta$ 42 can cause familial AD suggests that lowering A $\beta$ 42 by a similar amount could have a disease-modifying effect. This suggests that this degree of Notch-sparing selectivity may be sufficient to avoid mechanism-based toxicity (Kreft et al., 2008).

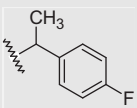
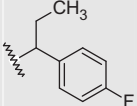
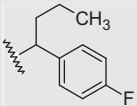
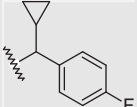
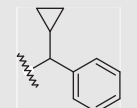
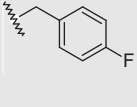
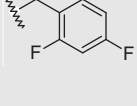
$\gamma$ -Secretase modulation is more desirable than inhibition from a therapeutic perspective and may reduce the risk of mechanism-based toxicities (Xin et al., 2011). Such compounds, called Notch-sparing GSIs (NS-GSIs) or  $\gamma$ -secretase modulators (GSMs), would be good candidates for AD therapeutics (Kurosumi et al., 2010).

Computational chemistry has developed into an important contributor to rational drug design. Molecular modeling study is an approach used to narrow down a library containing an extraordinarily high number of random molecules into a smaller list of potentially effective inhibitors. The techniques of QSAR are valuable molecular modeling tools for drug design. The quantitative structure activity relationship (QSAR) approach became very useful and largely widespread for the

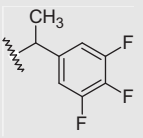
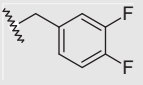
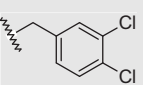
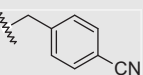
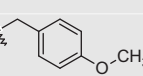
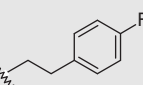
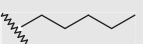
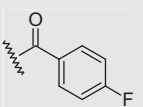
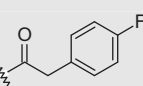
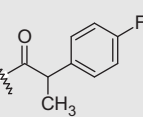
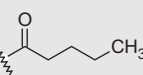
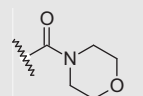
prediction of biological activities, particularly in drug design. The QSAR method provides a rational basis for understanding mechanisms of biological performance and shows how to improve performance by altering chemical structures of ligands, which has been proved to be one of the most embraced computational approaches in modern drug discovery (Bhadoriya et al., 2012b; Chen et al., 2009; Jain et al., 2012a,b,c). Quantitative structure activity relationship (QSAR) is one of the major tools in drug discovery to explore ligand–receptor/enzyme interactions, especially when either the structural details of the target are not known or protein binding data of ligand is unavailable. Quantitative structure–activity relationship (QSAR) studies leading to models in terms of chemical structures and their biological activities produce useful information for drug design and medicinal chemistry. 2D-QSAR does not involve complex alignment or assumptions on conformations (Joshi et al., 2010). 3D-QSAR is a broad term encompassing all those QSAR methods which correlate macroscopic target properties with computed atom-based descriptors derived from the spatial (3D) representation of the molecular structures (Verma et al., 2010). The 3D-QSAR protocols have been selected with a view to understand the ligand–receptor interaction in the light of steric, electrostatic and hydrophobic properties (Ghosh and Bagchi, 2009a). Three dimensional quantitative structure activity relationship (3D-QSAR) is a widely used tool to identify the steric, electrostatic, and hydrophobic structural requirements of various drugs acting via receptor modulation for exerting biological activity. Development of the 3D-QSAR model based on the biological activity of compounds enables ligand-based drug design that guides experimental chemical synthesis of compounds with higher potency even when the 3D structure of the biological target is unknown (Bhadoriya et al., 2012a,b,d; Mathura et al., 2010; Sharma et al., 2012).

Previously, we have reported 2D-QSAR studies on a series of fused 5,6-bicyclic heterocycles such as  $\gamma$ -secretase modulators (GSMs) with classical 2D descriptors. MLR method was used to generate statistically significant 2D-QSAR models (Bhadoriya et al., 2012c). Now, in continuation with our earlier work, we report 3D-QSAR studies on such a series of fused 5,6-bicyclic heterocycles with classical 3D descriptors. The present work is an attempt to generate predictive 3D-QSAR models based on 3D-QSAR methods and to find the structural features of fused 5,6-bicyclic heterocycles as  $\gamma$ -secretase modulators (GSMs) required for A $\beta$ <sub>42</sub> inhibitory activities to guide the rational synthesis of novel  $\gamma$ -secretase modulators (GSMs). 3D-QSAR field descriptors i.e. steric, electrostatic and hydrophobic are useful for the better understanding of molecular modeling studies of this series of compounds in terms of ligand–receptor interactions. In this investigation, a widely used technique, viz. stepwise (SW) has been applied for descriptor optimization, and kNN–MFA analysis has been applied for 3D-QSAR model development. The developed model provides insight into the influence of various interactive fields on the activity and, thus, can help in designing and forecasting the A $\beta$ <sub>42</sub> inhibitory activities of fused 5,6-bicyclic heterocycles.

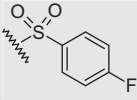
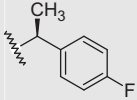
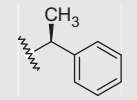
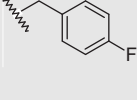
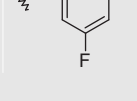
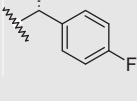
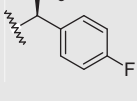
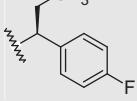
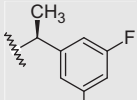
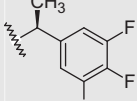
**Table 1** Structures of dataset used for kNN-MFA QSAR analysis with corresponding observed and predicted A $\beta$ <sub>42</sub> inhibitory activities of fused 5,6-bicyclic heterocycles as  $\gamma$ -secretase modulators.

Compound_ID	R	X	Y	Observed activity A $\beta$ <sub>42</sub> IC <sub>50</sub> (nM)	Observed activity (A $\beta$ <sub>42</sub> pIC <sub>50</sub> )	Stepwise forward kNN-MFA	
						Predicted Activity (A $\beta$ <sub>42</sub> pIC <sub>50</sub> )	Residual
GSM1	H	N	N	20,000	4.70	4.7	0
GSM2		N	N	509	6.29	6.31	-0.02
GSM3 <sup>a</sup>		N	N	156	6.81	6.94	-0.13
GSM4		N	N	145	6.84	6.94	-0.1
GSM5		N	N	16,374	4.79	4.7	0.09
GSM6		N	N	20,000	4.70	4.79	-0.09
GSM7 <sup>a</sup>		N	N	649	6.19	6.21	-0.02
GSM8		N	N	342	6.47	6.79	-0.32

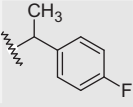
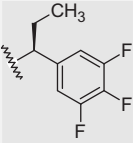
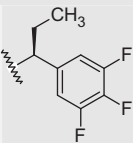
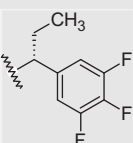
**Table 1** (continued)

Compound_ID	R	X	Y	Observed activity $A\beta_{42}IC_{50}$ (nM)	Observed activity ( $A\beta_{42}pIC_{50}$ )	Stepwise forward kNN-MFA	
						Predicted Activity ( $A\beta_{42}pIC_{50}$ )	Residual
GSM9		N	N	116	6.94	6.84	0.1
GSM10 <sup>a</sup>		N	N	552	6.26	6.31	-0.05
GSM11 <sup>a</sup>		N	N	126	6.90	6.35	0.55
GSM12		N	N	2531	5.60	5.65	-0.05
GSM13		N	N	1409	5.85	6.21	-0.36
GSM14		N	N	619	6.21	5.85	0.36
GSM15		N	N	1406	5.85	6.06	-0.21
GSM16		N	N	19,316	4.71	5.65	-0.94
GSM17 <sup>a</sup>		N	N	1919	5.72	5.6	0.12
GSM18		N	N	485	6.31	6.29	0.02
GSM19		N	N	2069	5.68	5.6	0.08
GSM20		N	N	20,000	4.70	4.7	0

**Table 1** (continued)

Compound_ID	R	X	Y	Observed activity A $\beta$ <sub>42</sub> IC <sub>50</sub> (nM)	Observed activity (A $\beta$ <sub>42</sub> pIC <sub>50</sub> )	Stepwise forward kNN-MFA	
						Predicted Activity (A $\beta$ <sub>42</sub> pIC <sub>50</sub> )	Residual
GSM21		N	N	2215	5.65	5.6	0.05
GSM22		O	N	465	6.33	6.27	0.06
GSM23 <sup>a</sup>		O	N	537	6.27	6.03	0.24
GSM24 <sup>a</sup>		O	N	356	6.45	6.7	-0.25
GSM25		O	N	440	6.36	6.84	-0.48
GSM26 <sup>a</sup>		N	O	1392	5.86	6.03	-0.17
GSM27		N	O	448	6.35	6.03	0.32
GSM28		N	O	875	6.06	5.85	0.21
GSM29		N	O	935	6.03	6.35	-0.32
GSM30		N	O	540	6.27	6.33	-0.06

**Table 1** (continued)

Compound_ID	R	X	Y	Observed activity $A\beta_{42}IC_{50}$ (nM)	Observed activity ( $A\beta_{42}pIC_{50}$ )	Stepwise forward kNN-MFA	
						Predicted Activity ( $A\beta_{42}pIC_{50}$ )	Residual
GSM31		N	NH	164	6.79	6.84	-0.05
GSM32		N	NH	201	6.70	6.91	-0.21
GSM33		N	NH	122	6.91	6.7	0.21
GSM34		N	NH	209	6.68	6.27	0.41

<sup>a</sup> Indicates the compounds considered in the test set, rest of the compounds considered in the training set for the 3D-QSAR study.

## 2. Experimental work

The molecular modeling studies (3D-QSAR) were carried out on a Windows XP workstation using the molecular modeling software package VLife Molecular Design Suite (VLifeMDS) version 3.5 (VLife MDS 3.5, 2008).

### 2.1. Biological activity dataset for analysis

A set of 34 fused 5,6-bicyclic heterocycles as  $\gamma$ -secretase modulators (GSMs) reported by Qin et al. (Qin et al., 2011), were selected for  $A\beta_{42}$  inhibitory activity requirements. The structures, biological activity and predicted activity data are given in Table 1.  $A\beta_{42}$  inhibitory activity was reported as  $IC_{50}$  in nM units. For the 3D-QSAR study the reported  $IC_{50}$  was converted to negative logarithm ( $pIC_{50}$ ) in molar units and subsequently used as the dependent variable for the 3D-QSAR analysis.

### 2.2. Computational details

The structures of the fused 5,6-bicyclic heterocycles considered for 3D-QSAR study were sketched on Chem sketch 12.0 software (ACD/Chemsketch 12.0, 2009). All structures were cleaned and 3D optimized. The energy optimization of the molecules was performed by batch calculation in VLife MDS 3.5 software (VLife MDS 3.5, 2008) using Merck Molecular

Force Field (MMFF) (Halgren, 1996a,b,c,d, 1999a,b) with distance dependent dielectric function and an energy gradient of 0.001 kcal/mol Å.

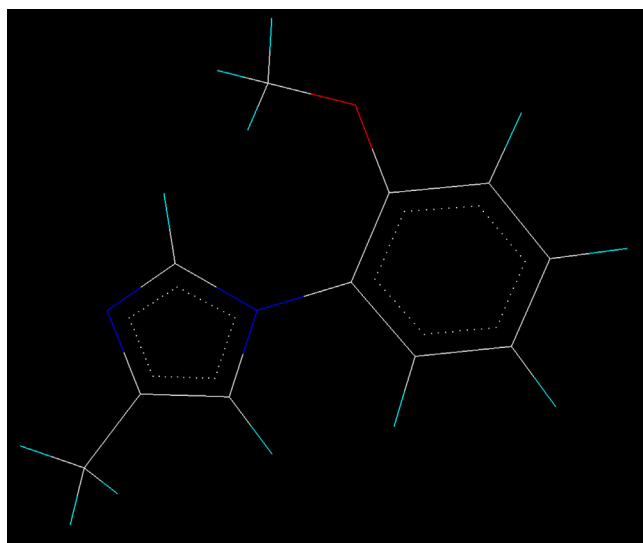
### 2.3. Conformation generation of molecules

It is a well recognized fact that each compound containing one or more single bonds is existing at each moment in many different so-called rotamers or conformers. Although small molecules may have only a single lowest energy conformation large and flexible molecules do exist in multiple conformations at physiological conditions. Therefore, it becomes necessary to include various such conformations of the molecules in a 3D-QSAR study (Verma et al., 2010). Multiple conformations of each molecule were generated using a Monte Carlo conformational search with an RMS gradient of 0.001 kcal/mol using a MMFF. Monte Carlo search method is a random search method for finding conformations of molecules (Metropolis et al., 1953). The low energy conformer for each compound was selected for further 3D-QSAR study.

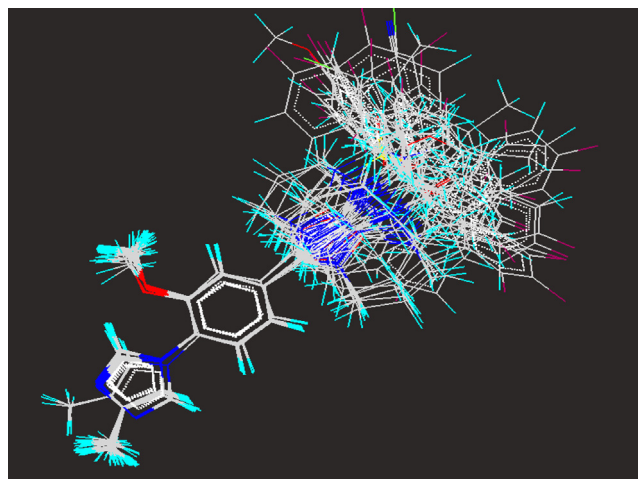
### 2.4. Alignment of molecules

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 (Vyas and Ghate, 2012). One of the most crucial problems in most of the align-

ment-based 3D-QSAR methods is that their results are highly sensitive to the manner in which the bioactive conformations of all the molecules are superimposed over each other. In cases, where all the molecules in a data set have a common rigid core structure, molecules can be aligned easily using the least-square fitting procedure. However in the case of structural heterogeneity in the dataset, alignment of highly flexible molecules becomes quite difficult and time consuming (Verma et al., 2010). The position of each atom is important for the kNN-MFA study because the descriptors calculation is 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 and determining factors to build a reasonable and a reliable model (Bhadoriya et al., 2012d). A proper alignment of the structures is a crucial, critical and decisive step for obtaining effective and valid 3D-QSAR models. Improper or incorrect alignment of molecules can create models providing little information relating to the main orientation of the molecule in the active site (Telvekar et al., 2010). Energy-minimized and geometry-optimized structures of molecules were aligned by the template-based method (Ajmani et al., 2006), where a template structure is defined and used as a basis for the alignment of a set of molecules, and a reference molecule is chosen on which the other molecules of the data set get aligned considering the chosen template. The template structure, i.e. 1-(2-methoxyphenyl)-4-methyl-1*H*-imidazole ring, was used for the alignment by considering the common elements of the series as shown in Fig. 1. The reference molecule is chosen in such a way that it is the most active among the series of molecules considered. The compound GSM9 possessed very high A $\beta$ <sub>42</sub> inhibitory activity which made it a valid lead molecule and, therefore, was chosen as a reference molecule. After optimizing, the template structure and the reference molecule were used to superimpose all molecules from the series using the template alignment method in VLife MDS 3.5 software (VLife MDS 3.5, 2008) to obtain optimal alignment between the molecular structures necessary for ligand–receptor/enzyme interactions. The superimposition of all molecules based on minimizing RMS deviation is shown



**Figure 1** 1-(2-Methoxyphenyl)-4-methyl-1*H*-imidazole ring as a common template used for template based alignment of fused 5,6-bicyclic heterocycles as  $\gamma$ -secretase modulators.



**Figure 2** Stereo view of template based alignment of fused 5,6-bicyclic heterocycles on the base template.

in Fig. 2. The resulting alignments of molecules were used for building kNN-MFA 3D-QSAR models.

## 2.5. Calculation of molecular field descriptors for 3D-QSAR analysis

Molecular field analysis (MFA) was employed to derive the 3D-QSAR model in this study. Molecular field analysis (MFA) model is predictive and sufficiently reliable to guide the chemist in designing novel compounds. This approach is effective for the analysis of data sets, where activity information is available but the structure of the receptor site is unknown (Silakari et al., 2011). After superimposition, the overlaid set of molecules is positioned in the center of a lattice or grid box, to calculate interaction energies between the ligands and different probe atoms placed at each intersection of the lattice (Verma et al., 2010). The aligned biologically active conformations of fused 5,6-bicyclic heterocycles were used for the calculation of molecular fields. Molecular fields are the electrostatic, steric and hydrophobic interaction energies which were computed at the lattice points of the grid using a methyl probe of charge +1 considering Gasteiger–Marsili charges (Gasteiger and Marsili, 1980). 10.0 kcal/mol electrostatic cut-off and 30.0 kcal/mol steric cut-off was employed to perform descriptor calculation. The term descriptor is utilized to indicate field values at the lattice points. These interaction energy values are considered for relationship generation and utilized as descriptors to decide nearness between molecules. A value of 1.0 is assigned to the distance-dependent dielectric constant. A total of 5577 three dimensional descriptors were calculated using VLife MDS software. These included electrostatic, steric and hydrophobic field descriptors (1859 for each electrostatic, steric and hydrophobic descriptor) for all the compounds in separate columns.

## 2.6. Design of the training and test sets using SEM (sphere exclusion method)

In order to obtain a validated QSAR model for the purpose of meaningful prediction, an available dataset should be divided into training and test sets. For the prediction statistics to be reliable, the test set must include at least five compounds



(Golbraikh and Tropsha, 2002b). The segregation or separation of the data set of 34 molecules into training and test sets comprising of 26 and 8 molecules, respectively, was performed by the rational sphere exclusion (SE) method with a dissimilarity value of 2.23. The dissimilarity value gives the sphere exclusion radius. For sphere exclusion algorithm,  $\text{pIC}_{50}$  activity was used as a dependent variable and various calculated electrostatic, steric and hydrophobic field 3D molecular descriptors as independent descriptors. Sphere exclusion algorithm allows constructing training sets covering all descriptor space areas occupied by representation points. The test set compounds should represent structural diversity and a range of biological activities similar to that of the training set. Compounds in the test set allowed us to use one test compound over 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 permitted us to evaluate the extrapolative accuracy of the 3D-QSAR models. The test set molecules captured structural features of the training set molecules, thus their activities could be well predicted. Training set of 26 compounds and a test set of eight compounds were used for generating 3D-QSAR models and for validating the quality of the models, respectively. Sphere exclusion approach resulted, eight compounds, namely GSM3, GSM7, GSM10, GSM11, GSM17, GSM23, GSM24, and GSM26 as the test set and the remaining 26 compounds as the training set (Table 1).

#### 2.7. Stepwise forward as variable (feature) selection method

The problem that is faced frequently by a researcher in developing QSAR models is that of a small number of observations (molecules) compared with a large number of molecular parameters in the descriptor pool. Recent trends in QSAR have focused on the development of procedures that allow the selection of optimal variables with low complexity and good predictive accuracy from the available pool of descriptors of chemical structures, i.e. the ones that are most meaningful and statistically significant in terms of correlation with biological activity. This process forms the basis of a technique known as feature selection or variable selection (Guyon and Elisseeff, 2003; Ghosh and Bagchi, 2009b; Sahu et al., 2011). Our first objective was to determine the best variables which produce the most significant 3D-QSAR models linking the structure of compounds with their binding affinity. Among several search algorithms, SW forward variable selection algorithm coupled with kNN-MFA was used to establish the 3D-QSAR models. In the SW forward variable selection algorithm, the search procedure begins with developing a trial model step by step with a single independent variable and to each step, independent variables are added one at a time, examining the fit of the model by using the kNN-MFA procedure. Thus, the model is repeatedly altered from the previous one by adding or removing a predictor variable in accordance with the 'stepping criteria' (in this case,  $F = 4$  for inclusion for the forward selection method). The method continues until there is no significant variable remaining outside the model.

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 variance cut-off was set at 0.0, and scaling as none. Additionally the kNN parameter setting was done in

which the number of maximum neighbors and the number of minimum neighbors were set at 5 and 2, respectively and the prediction method was selected as the distance-based weighted average.

#### 2.8. *k*-Nearest neighbor molecular field analysis method for generating 3D-QSAR models

In order to predict binding domain of  $\gamma$ -secretase, a *k*-nearest neighbor (kNN) classification model was developed. The *k*-nearest neighbor (kNN) method is one of the simplest machine learning algorithms, most commonly used for classifying a new pattern (e.g. a molecule) (Verma et al., 2010). The kNN technique is a conceptually simple approach to pattern recognition problems (VLife MDS 3.5, 2008). *k*-Nearest neighbor molecular field analysis (kNN-MFA) is the 3D-QSAR method which has been used to produce the 3D models to indicate the regions that affect biological activity with a change in the chemical substitution (Sharma et al., 2012). The kNN methodology relies on a simple distance learning approach whereby an unknown/new 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 follows:

- (1) Calculate the 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 (Sharaf et al., 1986). 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 stepwise variable selection method. This method employs a stepwise variable selection procedure combined with kNN to optimize

- (i) The number of nearest neighbors ( $k$ ) and
- (ii) The selection of variables from the original pool.

#### 2.9. *k*-Nearest neighbor QSAR (*k*NN weighted average method)

The kNN method was also used to develop a QSAR model using continuous variable, i.e. using activity as  $\text{pIC}_{50}$  values. In this case, by using a developed kNN QSAR model the activity of a molecule can be predicted using weighted average activity (Eq. (1)) of  $k$  most similar molecules in the training set.

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

where  $y_i$  and  $\hat{y}_i$  are the actual and predicted activities of the  $i$ th molecule respectively, and  $w_i$  is weight calculated using (Eq. (2)).

$$w_i = \frac{\exp(-dj)}{\sum_{j=1}^k \exp(-dj)} \quad (2)$$



The similarities were evaluated as the inverse of Euclidean distances ( $d_j$ ) between molecules (Eq. (3)) using only the subset of descriptors corresponding to the model, where,  $k$  is the number of nearest neighbors in the model.

$$d_{i,j} = \left[ \sum_{m=1}^{Vn} (X_{i,m} - X_{j,m})^2 \right]^{1/2} \quad (3)$$

where,  $X$  is the matrix of the selected descriptors ( $Vn$ ) for the kNN QSAR model.

### 2.10. Model validation and evaluation

This is done to test the internal stability and predictive ability of the QSAR models.

#### 2.10.1. Internal and external validations

Internal validation was carried out using the leave-one-out ( $q^2$ , LOO) method. For calculating  $q^2$ , each molecule in the training set was sequentially eliminated, the model refit using same descriptors, and the biological activity of the eliminated molecule predicted using the refit model. The cross-validated coefficient,  $q^2$ , was calculated using Eq.(4).

$$q^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{\text{mean}})^2} \quad (4)$$

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

However, a high  $q^2$  value does not necessarily give a suitable representation of the real predictive power of the model for  $\gamma$ -secretase modulators. So, an external validation is also carried out in this study. The external predictive power of the model is assessed by predicting pIC<sub>50</sub> value of eight test set molecules, which are not included in the 3D-QSAR model development. The predictive ability of the selected model is also confirmed by pred\_r<sup>2</sup>.

For external validation, activity of each molecule in the test set was predicted using the model generated from the training set. The pred\_r<sup>2</sup> value is calculated as follows (Eq. (5))

$$\text{Pred}_r^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - y_{\text{mean}})^2} \quad (5)$$

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

Both summations are over all molecules in the test set. Thus the pred\_r<sup>2</sup> value is indicative of the predictive power of the current kNN-MFA model based on the external test set.

#### 2.10.2. Randomization test

Y-randomization (randomization of response) is a widely used approach to establish the model robustness. It consists of repeating the calculation procedure with randomized activities and subsequent probability assessment of the resultant statistics. Frequently, it is used along with cross-validation (Golbraikh and Tropsha, 2002b). The robustness of the models for training sets was examined by comparing these models to those derived for random data sets. Random sets were generated by rearranging the activities of the molecules in the training set. The significance of the models hence obtained was derived based on a calculated  $Z$ score; Eq. (6).

A  $Z$ score value is calculated using the following formula:

$$Z\text{score} = \frac{(h - \mu)}{\sigma} \quad (6)$$

where  $h$  is the  $q^2$  value calculated for the actual data set,  $\mu$  is the average  $q^2$ , and  $\sigma$  is its standard deviation calculated for various iterations using models built by different random data sets.

The probability ( $\alpha$ ) of significance of the randomization test is derived by using calculated  $Z$ score value as given in the literature (Shen et al., 2003).

#### 2.10.3. Evaluation of the quantitative model

The developed 3D-QSAR model was evaluated using the following statistical measures:  $N$ , number of observations (molecules) in the training set; number of nearest neighbors, number of  $k$ -nearest neighbors in the model;  $q^2$ , cross-validated  $r^2$  (by leave one out) which is a relative measure of quality of fit; pred\_r<sup>2</sup>,  $r^2$  for external test set;  $q^2$ se, standard error of cross-validation and pred\_r<sup>2</sup>se, standard error of external test set prediction. However, a QSAR model is considered to be predictive, if the following conditions are satisfied:  $q^2 > 0.6$  and pred\_r<sup>2</sup>  $> 0.5$  (Golbraikh and Tropsha, 2002a). The low standard error of pred\_r<sup>2</sup>se and  $q^2$ se shows absolute quality of fitness of the model. The high pred\_r<sup>2</sup> and low pred\_r<sup>2</sup>se show high predictive ability of the model.

The  $q^2$  and pred\_r<sup>2</sup> values were used as deciding factors in selecting the optimal models.

## 3. Results and discussion

The 3D-QSAR study of 34 fused 5,6-bicyclic heterocycles for A $\beta$ <sub>42</sub> inhibitory activity (Table 1) through kNN methodology, based on SW feature selection method using VLife MDS 3.5 software (VLife MDS 3.5, 2008), resulted in the following statistically significant model, considering the term selection criterion as  $q^2$  and pred\_r<sup>2</sup>. The training and test set compounds (Table 1) for this group of compounds were selected by the sphere exclusion method, and the model was validated by both internal and external validation procedures.

Selection of compounds in the training set and test is a key and important feature of any QSAR model. Therefore care was taken in such a way that biological activities of all compounds in the test lie within the maximum and minimum value range of biological activities of the training set of compounds. The UniColumn Statistics of test and training sets further reflected the correct selection of test and training sets (Table 2).

The maximum and minimum values in the training and test sets were compared in a way that:

1. The maximum value of pIC<sub>50</sub> of the test set should be less than or equal to the maximum value of pIC<sub>50</sub> of the training set.
2. The minimum value of pIC<sub>50</sub> of the test set should be higher than or equal to the minimum value of pIC<sub>50</sub> of the training set.

This observation showed that the test set was interpolative and derived within the minimum–maximum range of the training set. The mean and standard deviation of pIC<sub>50</sub> values of sets of training and test provide insights into the relative differ-

**Table 2** Unicolumn statistics of the training and test sets for A $\beta$ <sub>42</sub> inhibitory activity.

Data set	Column name	Average	Max.	Min.	SD	Sum
Training	A $\beta$ <sub>42</sub> pIC <sub>50</sub>	5.9912	6.9400	4.7000	0.7349	155.7700
Test	A $\beta$ <sub>42</sub> pIC <sub>50</sub>	6.3075	6.9000	5.7200	0.4117	50.4600

Max., maximum; Min., minimum; SD, standard deviation.

ence of mean and point density distribution (along mean) of the two sets. The mean of the test sets was higher than the training sets that indicates the presence of relatively more active molecules as compared to the inactive ones. Also, a relatively higher standard deviation in training sets indicates that training sets had widely distributed activity of the molecules as compared to the test sets (VLife MDS 3.5, 2008).

Statistically significant best 3D-QSAR model was selected, considering the term selection criterion as  $q^2$ , the internal predictive ability of the model and  $\text{pred}_r^2$ , the ability of the model to predict the activity of the external test set. Best 3D-QSAR model is chosen for discussion.

### 3.1. 3D-QSAR modeling and its validation

Various 3D-QSAR models were developed using the kNN method. In the present study, kNN coupled with stepwise variable selection method was used to develop 3D-QSAR models of fused 5,6-bicyclic heterocycles with reported A $\beta$ <sub>42</sub> inhibitory activities based on hydrophobic and steric fields. Several statistically significant 3D-QSAR models were generated, of which the corresponding best model is reported herein. 3D-QSAR model was selected based on the value of statistical parameters and the best kNN-MFA 3D-QSAR model with 26 training set compounds have a  $q^2 = 0.8457$  and  $\text{pred}_r^2 = 0.7556$  (Table 3).

The descriptors selected for 3D-QSAR modeling A $\beta$ <sub>42</sub> inhibitory activity of fused 5,6-bicyclic heterocycles are summarized in Table 4 and the correlation matrix between the hydrophobic and steric descriptors influencing the A $\beta$ <sub>42</sub> inhibitory activity is presented in Table 5.

The actual/observed activities, predicted activities by 3D-QSAR model and residuals of both training and test sets compounds are given in Table 1. The plots of observed versus predicted activity of both training and test sets compounds helped in cross-validation of the kNN-MFA QSAR model and are depicted in Fig. 3. The residuals of the kNN-MFA calculated values of pIC<sub>50</sub> are plotted against the observed activity (A $\beta$ <sub>42</sub>-pIC<sub>50</sub>) values in Fig. 4. The propagation of residuals at both

**Table 3** Statistical results of the 3D-QSAR model generated by stepwise forward variable selection kNN-MFA method for fused 5,6-bicyclic heterocycles as  $\gamma$ -secretase modulators.

S. No.	Statistical parameter	3D-QSAR results
1	$q^2$	0.8457
2	$q^2_{se}$	0.2887
3	$\text{pred}_r^2$	0.7556
4	$\text{pred}_r^2_{se}$	0.2634
5	$N_{\text{training}}$	26
6	Nearest neighbor	2
7	Degree of freedom	22
8	Contributing descriptors	1. H_83 (0.3416, 0.3466) 2. S_183 (−0.3355, −0.3273) 3. S_227 (−0.0931, −0.0911)

**Table 4** Descriptors used in 3D-QSAR model with values.

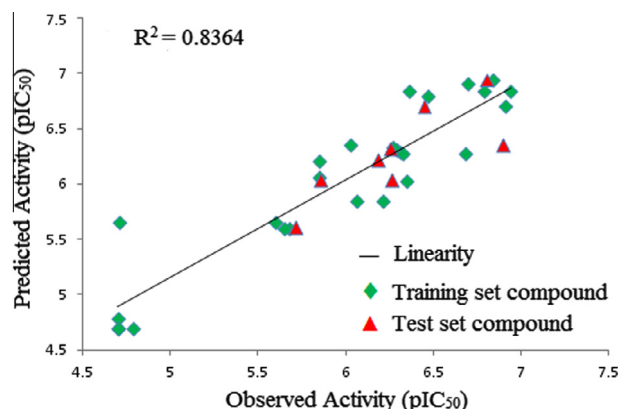
Compound_ID	Observed activity (A $\beta$ <sub>42</sub> pIC <sub>50</sub> )	H_83	S_183	S_227
GSM1	4.70	0.191522	−0.326565	−0.090766
GSM2	6.29	0.305327	−0.331884	−0.089395
GSM3	6.81	0.332995	−0.332539	−0.089892
GSM4	6.84	0.341568	−0.335532	−0.091112
GSM5	4.79	0.31945	−0.338717	−0.080566
GSM6	4.70	0.323797	−0.330989	−0.088111
GSM7	6.19	0.286387	−0.331192	−0.087782
GSM8	6.47	0.346581	−0.32725	−0.093112
GSM9	6.94	0.338179	−0.331278	−0.094365
GSM10	6.26	0.31265	−0.334265	−0.092369
GSM11	6.90	0.365571	−0.331393	−0.092395
GSM12	5.60	0.271679	−0.328204	−0.09097
GSM13	5.85	0.287434	−0.334026	−0.08554
GSM14	6.21	0.285316	−0.332226	−0.08714
GSM15	5.85	0.307373	2.682489	−0.085977
GSM16	4.71	0.279406	−0.328525	−0.08907
GSM17	5.72	0.26086	−0.331234	−0.088706
GSM18	6.31	0.312223	−0.32969	−0.091833
GSM19	5.68	0.276221	0.986976	−0.091342
GSM20	4.70	0.189483	−0.338925	−0.091308
GSM21	5.65	0.274999	−0.332262	−0.09314
GSM22	6.33	0.389735	−0.331284	−0.093743
GSM23	6.27	0.360056	−0.334452	−0.092105
GSM24	6.45	0.373484	−0.334879	−0.09559
GSM25	6.36	0.334969	−0.335777	−0.083753
GSM26	5.86	0.35857	−0.339599	−0.083049
GSM27	6.35	0.358021	−0.327903	−0.089254
GSM28	6.06	0.391041	3.574147	−0.093438
GSM29	6.03	0.355705	−0.332605	−0.087771
GSM30	6.27	0.400326	−0.334155	−0.095092
GSM31	6.79	0.347824	−0.333944	−0.092523
GSM32	6.70	0.380233	−0.333618	−0.083544
GSM33	6.91	0.371048	−0.329226	−0.082761
GSM34	6.68	0.403634	−0.336324	−0.083096

**Table 5** Correlation matrix for steric and hydrophobic descriptors influencing the A $\beta$ <sub>42</sub> inhibitory activity (3D-QSAR model).

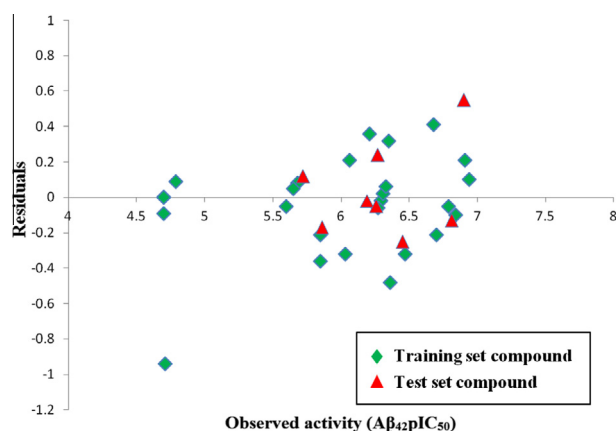
	S_183	S_227	H_83	Score
S_183	1	−0.10075	0.118567	3
S_227	−0.10075	1	0.088562	3
H_83	0.118567	0.088562	1	3

sides of the zero line indicates that no systematic error exists in the development of the kNN-MFA model.

The model selection criterion is the value of  $q^2$ , the internal predictive ability of the model and that of  $\text{pred}_r^2$ , the ability of the model to predict the activity of the external test set. For A $\beta$ <sub>42</sub> inhibitory activity, selected model was found to be statistically most significant, especially with respect to the internal predictive ability ( $q^2 = 0.8457$ ) of the model. As the cross-



**Figure 3** Comparison of observed activity versus predicted activity for the training set & test set compounds according to the 3D-QSAR SW-kNN-MFA model.



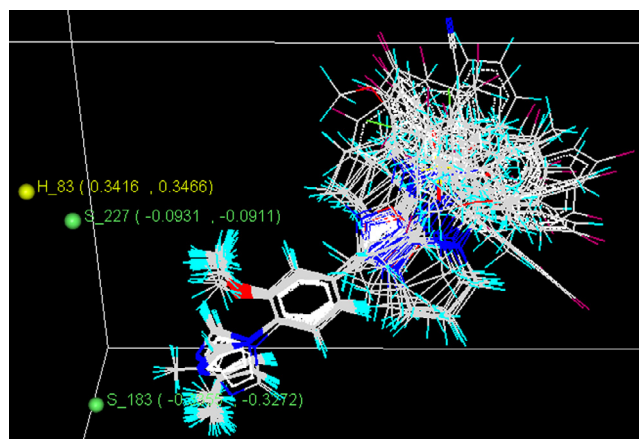
**Figure 4** Scatter plot of the observed activity ( $A\beta_{42}pIC_{50}$ ) versus residuals for training set & test set compounds.

validated correlation coefficient ( $q^2$ ) is used as a measure of the reliability of prediction, the correlation coefficient suggests that our model is reliable and accurate. A data set of eight compounds was selected as the test set from the original data of 34 compounds for the validation experiments. The value of  $pred\_r^2$  was obtained for the test set and gave better results, with a value of 0.7556, which means 76% predictive power for the external test set. Thus, our model displays good predictivity in regular cross-validation (Table 3.).

H\_83, S\_227, and S\_183 are the hydrophobic and steric field energy of interactions. The above model is validated by predicting the biological activities of the test set molecules, as indicated in Table 1.

The plot of observed versus predicted activities for the test compounds is represented in Fig. 3. From Table 1, it is evident that the predicted activities of all the compounds in the test set are in good agreement with their corresponding experimental activities and optimal fit is obtained.

The external predictability of the above 3D-QSAR model using the test set was determined by  $pred\_r^2$ , which is 0.7556. So, the above results indicate that 3D-QSAR model for  $\gamma$ -secretase modulators generate 84.57% and 75.56% internal and external model predictions, respectively.



**Figure 5** SW-kNN-MFA model field plot for hydrophobic and steric field interactions.

### 3.1.1. Hydrophobic and steric field plot

The plot of contributions of hydrophobic and steric field interactions (Fig. 5) indicates relative regions of the local fields (hydrophobic and steric field) around the aligned molecules. Yellow and green balls represent hydrophobic and steric field effects, respectively.

In the QSAR model, hydrophobic descriptor with positive coefficients represents regions of high hydrophobic tolerance. Steric descriptors with negative coefficients indicate regions where the bulky substituent is disfavored.

From the 3D-QSAR model (Fig. 5) it is observed that the steric field with negative coefficient (S\_183) at the methyl position of the imidazole ring indicate that bulky substituents are unfavorable on this site and the presence of bulky substituents decrease the activity of fused 5,6-bicyclic heterocycles. Presence of another steric field with negative coefficient (S\_227) at the methoxy position of the benzene ring suggests that bulky substituents are unfavorable on this site and the presence of bulky substituents decrease the activity of fused 5,6-bicyclic heterocycles. Hydrophobic field with positive coefficient (H\_83) near the methoxy position of the benzene ring indicates that more hydrophobic groups in this position would have a beneficial effect on the anti-Alzheimer's activity of these fused 5,6-bicyclic heterocycles. Thus, the contribution plot arising out of 3D-QSAR studies provide some useful insights for better understanding of the structural features of these compounds responsible for producing significant  $A\beta_{42}$  inhibitory activities.

## 4. Conclusions

In conclusion, the present study highlights the importance of the structural features responsible for the  $A\beta_{42}$  inhibitory activity. A 3D-QSAR study of a set of fused 5,6-bicyclic heterocycles has been performed. Stepwise forward variable selection algorithm has been applied for variable selection and the models were developed by the kNN-MFA method. Reliability of the models was confirmed by several statistical analyses. For the dataset of 34 fused 5,6-bicyclic heterocycles, shape of the substituents i.e., steric and hydrophobic descriptors of the molecules appear to be the governing factor for the  $A\beta_{42}$  inhibitory activity. The present investigation will guide the synthetic medicinal chemist to design and synthesize new novel more potent  $A\beta_{42}$  inhibitors (anti-Alzheimer's agents) with in-

creased biological activity in comparison to the reported compounds. In future, continuation with this work docking studies will be performed on this particular dataset of 34 fused 5,6-bicyclic heterocycles by using receptor–ligand-co-crystal structures available in the protein data bank.

### Acknowledgements

The authors gratefully acknowledge “the Department of Science and Technology, Govt. of India” for providing funding to the institute for the Vlife MDS software. Authors would like to acknowledge the Principal of the institute for providing facilities to carry out the work.

### References

- ACD/Chemsketch 12.0, 2009.
- Ajmani, S., Jadhav, K., Kulkarni, S.A., 2006. *J. Chem. Inf. Model.* 46, 24–31.
- Balaraman, Y., Limaye, A.R., Levey, A.I., Srinivasan, S., 2006. *Cell. Mol. Life Sci.* 63, 1226–1235.
- Bhadoriya, K.S., Jain, S.V., Bari, S.B., Chavhan, M.L., Vispute, K.R., 2012a. *J. Chem.* 9 (4), 1753–1759.
- Bhadoriya, K.S., Kumawat, N.K., Bhavthankar, S.V., Avchar, M.H., Dhumal, D.M., Patil, S.V., Jain, S.V., 2012b. *J. Saudi Chem. Soc.*, <<http://dx.doi.org/10.1016/j.jscs.2012.11.001>>.
- Bhadoriya, K.S., Sharma, M.C., Jain, S.V., Kad, S.A., Raghuvanshi, D., 2012c. *J. Pharm. Res.* 5 (8), 4127–4132.
- Bhadoriya, K.S., Sharma, M.C., Jain, S.V., Raut, G.S., Rananaware, J.R., 2012d. *Med. Chem. Res.* <http://dx.doi.org/10.1007/s00044-012-0226-4>.
- Caldwell, J.P., Bennett, C.E., McCracken, T.M., Mazzola, R.D., Bara, T., Buevich, A., Burnett, D.A., Chu, I., Cohen-Williams, M., Josein, H., Hyde, L., Lee, J., McKittrick, B., Song, L., Terracina, G., Voigt, J., Zhang, L., Zhu, Z., 2010. *Bioorg. Med. Chem. Lett.* 20, 5380–5384.
- Chen, K.-X., Li, Z.-G., Xie, H.-Y., Gao, J.-R., 2009. *Lett. Drug Des. Discovery* 6, 193–200.
- Fischer, C., Zultanski, S.L., Zhou, H., Methot, J.L., Brown, W.C., Mampreian, D.M., Schell, A.J., Shah, S., Nuthall, H., Hughes, B.L., Smotrov, N., Kenific, C.M., Cruz, J.C., Walker, D., Bouthilllette, M., Nikov, G.N., Savage, D.F., Jeliakova-Mecheva, V.V., Diaz, D., Szwczak, A.A., Bays, N., Middleton, R.E., Munoz, B., Shearman, M.S., 2011. *Bioorg. Med. Chem. Lett.* 21, 4083–4087.
- Gasteiger, J., Marsili, M., 1980. *Tetrahedron* 36, 3219–3228.
- Ghosh, P., Bagchi, M.C., 2009a. *Curr. Med. Chem.* 16, 4032–4048.
- Ghosh, P., Bagchi, M.C., 2009b. *Mol. Simul.* 35, 1185–1200.
- Golbraikh, A., Tropsha, A., 2002a. Beware of  $q^2$ ! *J. Mol. Graphics Modell.* 20, 269–276.
- Golbraikh, A., Tropsha, A., 2002b. *J. Comput. Aided Mol. Des.* 16, 357–369.
- Guyon, I., Elisseeff, A., 2003. *J. Mach. Learning Res.* 3, 1157–1182.
- Halgren, T.A., 1996a. *J. Comput. Chem.* 17, 520–552.
- Halgren, T.A., 1996b. *J. Comput. Chem.* 17, 553–586.
- Halgren, T.A., 1996c. *J. Comput. Chem.* 17, 616–641.
- Halgren, T.A., 1996d. *J. Comput. Chem.* 17, 490–519.
- Halgren, T.A., 1999a. *J. Comput. Chem.* 20, 720–729.
- Halgren, T.A., 1999b. *J. Comput. Chem.* 20, 730–748.
- Harvey, B.K., Richie, C.T., Hoffer, B.J., Airavaara, M., 2011. *J. Neural Transm.* 118, 27–45. <http://dx.doi.org/10.1007/s00702-010-0476-6>.
- Iijima-Ando, K., Iijima, K., 2010. *Brain Struct. Funct.* 214, 245–262. <http://dx.doi.org/10.1007/s00429-009-0234-4>.
- Jain, S.V., Bhadoriya, K.S., Bari, S.B., 2012a. *Med. Chem. Res.* 21, 1665–1676. <http://dx.doi.org/10.1007/s00044-011-9681-6>.
- Jain, S.V., Bhadoriya, K.S., Bari, S.B., Sahu, N.K., Ghate, M., 2012b. *Med. Chem. Res.* 21, 3465–3484. <http://dx.doi.org/10.1007/s00044-011-9889-5>.
- Jain, S.V., Ghate, M., Bhadoriya, K.S., Bari, S.B., Chaudhari, A., Borse, J.S., 2012c. *Org. Med. Chem. Lett.* 2, 22. <http://dx.doi.org/10.1186/2191-2858-2-22>.
- Joshi, P., Tanwar, O., Rambhade, S., Bhaisare, M., Jain, D., 2010. *Med. Chem. Res.* <http://dx.doi.org/10.1007/s00044-010-9529-5>.
- Kim, S.D., Kim, J., 2008. *Cell Stress and Chaperones* 13, 401–412. <http://dx.doi.org/10.1007/s12192-008-0046-0>.
- Kreft, A., Harrison, B., Aschmies, S., Atchison, K., Casebier, D., Cole, D.C., Diamantidis, G., Ellingboe, J., Hauze, D., Hu, Y., Huryn, D., Jin, M., Kubrak, D., Lu, P., Lundquist, J., Mann, C., Martone, R., Moore, W., Oganessian, A., Porte, A., Riddell, D.R., Sonnenberg-Reines, J., Stock, J.R., Sun, S.-C., Wagner, E., Woller, K., Xu, Z., Zhou, H., Jacobsen, J.S., 2008. *Bioorg. Med. Chem. Lett.* 18, 4232–4236.
- Kurosumi, M., Nishio, Y., Osawa, S., Kobayashi, H., Iwatsubo, T., Tomita, T., Miyachi, H., 2010. *Bioorg. Med. Chem. Lett.* 20, 5282–5285.
- Mathura, V.S., Patel, N., Bachmeier, C., Mullan, M., Paris, D., 2010. *Bioinformation* 5 (3), 122–127.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. *J. Chem. Phys.* 21, 1087–1092.
- Narlawar, R., Revuelta, B.I.P., Baumann, K., Schubene, R., Haass, C., Steiner, H., Schmidt, B., 2007. *Bioorg. Med. Chem. Lett.* 17, 176–182.
- Qin, J., Dhondi, P., Huang, X., Mandal, M., Zhao, Z., Pissarnitski, D., Zhou, W., Aslanian, R., Zhu, Z., Greenlee, W., Clader, J., Zhang, L., Cohen-Williams, M., Jones, N., Hyde, L., Palani, A., 2011. *Bioorg. Med. Chem. Lett.* 21, 664–669.
- Sahu, N.K., Shahi, S., Sharma, M.C., Kohli, D.V., 2011. *Mol. Simul.* 37, 752–765.
- Sharaf, M.A., Illman, D.L., Kowalski, B.R., 1986. *Chemometrics, Chemical Analysis Series*. Wiley, New York.
- Sharma, M.C., Sharma, S., Bhadoriya, K.S., 2012. *J. Saudi Chem. Soc.* <<http://dx.doi.org/10.1016/j.jscs.2012.10.001>>.
- Shen, M., Xiao, Y., Golbraikh, A., Gombar, V.K., Tropsha, A., 2003. *J. Med. Chem.* 46, 3013–3020.
- Silakari, P., Srivastava, S.D., Kohli, D.V., Srivastava, S.K., Silakari, G., Vyas, B., Silakari, O., 2011. *Med. Chem. Res.* 20, 158–167. <http://dx.doi.org/10.1007/s00044-010-9301-x>.
- Sun, Z.-Y., Asberom, T., Bara, T., Bennett, C., Burnett, D., Chu, I., Clader, J., Cohen-Williams, M., Cole, D., Czarniecki, M., Durkin, J., Gallo, G., Greenlee, W., Josien, H., Huang, X., Hyde, L., Jones, N., Kazakevich, I., Li, H., Liu, X., Lee, J., MacCoss, M., Mandal, M.B., McCracken, T., Nomeir, A., Mazzola, R., Palani, A., Parker, E.M., Pissarnitski, D.A., Qin, J., Song, L., Terracina, G., Vicarel, M., Voigt, J., Xu, R., Zhang, L., Zhang, Q., Zhao, Z., Zhu, X., Zhu, Z., 2012. *J. Med. Chem.* 55, 489–502.
- Telvekar, V.N., Patel, D.J., Jadhav, N.C., Mishra, S.J., 2010. *Med. Chem. Res.* 19, 1174–1190. <http://dx.doi.org/10.1007/s00044-009-9261-1>.
- Verma, J., Khedkar, V.M., Coutinho, E.C., 2010. *Curr. Top. Med. Chem.* 10, 95–115.
- VLife MDS 3.5, 2008. Molecular design suite. VLife Sciences Technologies Pvt. Ltd., Pune, India. <[www.vlifesciences.com](http://www.vlifesciences.com)>.
- Vyas, V.K., Ghate, M., 2012. *Med. Chem. Res.* 21, 3021–3034. <http://dx.doi.org/10.1007/s00044-011-9837-4>.
- Xin, Z., Peng, H., Zhang, A., Talreja, T., Kumaravel, G., Xu, L., Rohde, E., Jung, M.-Y., Shackett, M.N., Kocisko, D., Chollate, S., Dunah, A.W., Snodgrass-Belt, P.A., Arnold, H.M., Taveras, A.G., Rhodes, K.J., Scannevin, R.H., 2011. *Bioorg. Med. Chem. Lett.* 21, 7277–7280.