

## DISCOVERY OF NOVEL AND SELECTIVE C-JUN NH<sub>2</sub>-TERMINAL KINASES 2 INHIBITORS BY TWO-DIMENSIONAL QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP MODEL DEVELOPMENT, MOLECULAR DOCKING AND ABSORPTION, DISTRIBUTION, METABOLISM, ELIMINATION PREDICTION STUDIES: AN *IN SILICO* APPROACH

ASHIMA NAGPAL\*, SARVESH PALIWAL

Department of Pharmacy, Banasthali Vidyapith, Banasthali – 304 022, Rajasthan, India. Email: ashima\_nagpal@yahoo.com

Received: 06 December 2017, Revised and Accepted: 19 January 2017

### ABSTRACT

**Objective:** Incited by the dearth of selective c-Jun NH<sub>2</sub>-terminal kinases (JNK) inhibitors, efforts have been made to design novel JNK2 inhibitors with good selectivity profile by utilizing a set of *in silico* tools.

**Methods:** The present study involved 2D QSAR model development through multiple linear regression and partial least square methods. Further, the information unveiled through the above model was meticulously utilized to design novel JNK2 inhibitors (compound a and compound b). The selectivity of the novel molecules was ascertained through the molecular docking experiments. Determination of Tanimoto similarity index and absorption, distribution, metabolism, and elimination (ADME) properties was performed to ascertain the novelty and drug-like properties of the designed molecules, respectively.

**Results:** Four explanatory variables or descriptors, moment of inertia 2 length (subst.1), moment of inertia 3 length (subst.3), Kier Chi<sub>4</sub> (path/cluster) index (whole molecule), and vamp highest occupied molecular orbital (whole molecule), were found to possess profound influence on the biological activity with the values of standard parameters, s: 0.28, r: 0.96, r<sup>2</sup>: 0.92, r<sup>2</sup>CV: 0.82, and f=59.69. Tanimoto index values were found to be 0.095 and 0.075 for compound a and compound b, respectively.

**Conclusion:** The values of the statistical parameters proved the developed model to be of excellent quality. The results obtained for ADME prediction studies proved both the compounds to be better than the most potent compound (compound 45) of the set of JNK2 inhibitors selected for model development. In addition, extremely low Tanimoto similarity index values for both the compounds provided sufficient evidence for the novelty of the designed molecules.

**Keywords:** QSAR, Multiple linear regression, Partial least square, Tanimoto index, Absorption; distribution; metabolism; and elimination.

© 2018 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>) DOI: <http://dx.doi.org/10.22159/ajpcr.2018.v11i5.24157>

### INTRODUCTION

c-Jun NH<sub>2</sub>-terminal kinases (JNKs), one of the stress-activated kinases, are the member of mitogen-activated protein kinase family. These are known to be evolutionarily conserved. Insights into the physiological roles, through studies conducted so far, of the JNKs have capacitated us to understand its way and extent of participation in diverse events taking place in the human body. Plethora of studies on JNKs has unveiled its crucial involvement in many diseases including neurodegenerative diseases [1-8], diabetes [9,10], and cell apoptosis [11], making it a crucial target for the discovery of novel scaffolds for the treatment of an array of diseases caused through the activation of this pathway. JNKs are known to be encoded by a set of three genes (JNK1, JNK2, and JNK3). Among them, JNK1 and JNK2 manifest wide expression throughout the body [12]. On the other hand, JNK3 shows its presence specifically in the brain and to a somewhat lower extent in the heart and the testis. A different splicing mechanism for the three JNK genes results in 10 different isoforms. When the COOH terminus of the JNK genes is spliced alternatively, it gives rise to the proteins with the size of either 46 or 54 kDa [12-14]. Most *in vitro* study models utilize c-Jun as a substrate to evaluate the nature and mode of JNK-substrate binding affinity as well as specificity. All the three isoforms of JNKs causes activation of c-Jun by targeting its delta domain for binding and carry out phosphorylation of its serine 63 and 73 residues [14]. Binding and affinity studies of JNK-substrate complex have shown JNK2 to possess 25 times greater affinity toward c-Jun substrate than JNK1 [14]. Even though, different isoforms of JNK1 and JNK2 bind to and phosphorylate the substrate to a

different extent, the efficiency of JNK1 isoforms to phosphorylate c-Jun is shown to be better [15]. Well established evidences are available, through various studies, in the area of autoimmune diseases, depicting the crucial involvement of JNK2 in chronic inflammatory diseases through the expression of cytokines and certain proteases [16,17]. For an instance, metalloproteases, when expressed in rheumatoid arthritis are known to contribute toward the destruction of the joints, and meticulous analysis suggested the role of JNK signaling in this pathway [18]. The second most common chronic inflammatory disease, atherosclerosis, is also known to be regulated by JNK pathway [19,20]. This fact was proved by the studies carried out on the animals deficient in JNK2 genes, and it has been observed that those particular animals were less prone to develop atherosclerosis than the rest of the animals used in the study. Furthermore, JNK1 deficient animals did not show any reduction in atherosclerosis developing tendency and the extent of the level of atherosclerosis in the animals remained the same, as in non-JNK1 deficient animals. The possible mechanism of the disease, most probably, is through the JNK2-facilitated phosphorylation of the modified lipid-binding and internalizing scavenger receptor-A that, after getting phosphorylated, promotes the uptake of lipids in macrophages resulting in the formation of the foam cells. The above findings indicate that the complications, which occur during atherosclerosis, can be mitigated using specific JNK2 inhibitors. The therapeutic efficiency of an inhibitor is directly related to how effectively, as well as how selectively, it binds to the binding domain of the receptor. The drug molecule, if not selective, may lead to undesirable side effects by targeting a

group of enzymes and not just one. However, the profound extent of homology in the amino acid sequencing of the binding domain of the kinases makes the designing of selective inhibitors very much difficult. However, the involvement of JNKs in controlling multifarious functions makes it an interesting enzyme to work on. To address this issue of non-availability of specific JNK2 inhibitors as therapeutic agents, a set of *in silico* tools, including multiple linear regression (MLR) and partial least square (PLS), have been employed to develop a QSAR model. The information extracted through the obtained model helped us to design novel molecules as JNK2 inhibitors. The prediction of their selective inhibitory activity has been made on the basis of their binding affinity on JNK2 receptor, evaluated through molecular docking experiments. Furthermore, the designed molecules were declared to be novel only after calculation of their Tanimoto similarity index.

## METHODS

### Sketching of preliminary structures

The selection of the data set of tri-substituted thiophenes [21] was made on the basis of diversity in the structural architecture of compounds [22], exhibiting profound variation in JNK2 inhibitory profile reported as  $IC_{50}$  values. The reported  $IC_{50}$  values of individual compounds were first converted into their respective  $\log IC_{50}$  to overcome the usual issues related to the possibility of data to be skewed. ChemDraw ultra 8.0 software was employed to sketch the two-dimensional structures of the compounds of the selected series followed by cleaning of each structure to transcend the ambiguity barrier usually encountered during the substituent defining step.

### Defining of substituents

To alleviate the ambiguity pertaining to the effects that various groups tend to have on the biological activity of the compound, it becomes necessary to define substitution positions around a moiety that is common among all the molecules. The present study employed TSAR software version 3.3 ([www.accelrys.com](http://www.accelrys.com)) to perform all the computational work required to build the presented QSAR model. First, the two-dimensional structures, along with their  $\log IC_{50}$  values, were imported in the TSAR worksheet, after their conversion into molfile format. After which, three substituents (R1, R2, and R3) in each molecule, using "define substituent" option in TSAR, were defined around N-(3,4-dimethylthiophen-2-yl)acetamide nucleus, found to be common in every molecule of the series (Fig. 1). After defining of the substituents, the partial charges of each molecule were calculated through an option called Charge 2, primarily known for performing the action of deriving charges. As we know, a careful analysis of the molecular structure is the key to deduce the importance of various physicochemical properties in a molecule accountable for certain activity profile. However, certain aspects of the relation between the structure of a molecule and its properties can be best understood through 3D conformation. Therefore, Corina make 3D program, which generates structure in the three-dimensional form, was employed for the conversion of chemical structures of the molecules and their corresponding substituents into their respective 3D form. Further, the

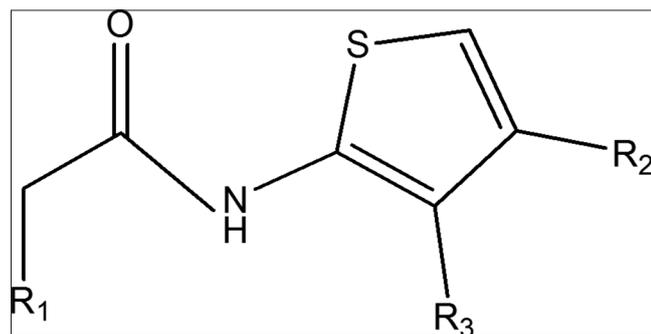


Fig. 1: Substitution pattern around N-(3,4-dimethylthiophen-2-yl)acetamide nucleus

generated 3D structures of the molecules were subjected to energy optimization process by employing cosmic module. Cosmic during flexible optimization, through the supply of force field for calculations of energy component of all molecular structures, ensure the selection of only those conformations which are found to be more energetically realistic.

### Data set preparation and data reduction

The onset of 2D QSAR model development was marked by the calculation of more than 200 descriptors or explanatory variables by TSAR software version 3.3. TSAR, an integrated analysis package, makes the investigation of interactions of quantitative structure-activity relationships, possible. It performs the calculations of numerical value of descriptors for all molecular structures and uses a statistical approach to define a correlation. A number of descriptors from atom counts, molecular indices, molecular attributes, and VAMP parameters were calculated [23,24]. The TSAR method is basically based on the assumption that an adequate sampling of these structure-based descriptors reveal all the evidence required to understand their biological properties [25]. However, redundancy in data can pose a serious problem in developing a reliable and robust model. Therefore, to alleviate the chances development of a low-quality QSAR model, data reduction was performed using correlation matrix technique, in which the intercorrelation among descriptors was evaluated. Among the descriptors showing correlation with each other, the one with lesser correlation with the biological activity was removed from the model. Intercorrelation among the descriptors is highly undesirable in a model because if the descriptors share any correlation with each other, then they are not considered to be independent in true manner and chances that they might give a false prediction about the quality of the model, increases.

### Model development

After curtailing the data to the required level, arbitrary division of compounds into training set and test set was done where 23 molecules of the data set were included in the training set, and the rest were placed in the test set. Training set compounds were then used to perform MLR and PLS for model development, and the test set was employed to determine the predictability of the generated model through the application of same methods. The MLR technique is based on the generation of the equation to evaluate the fitting of the model followed by a selection of the best model that fits well in it. Several MLR equations, based on a single dependent variable, which is usually biological activity, and variable descriptors, which constitutes the independent variable, are used to produce the equation when combined in different patterns. The quality of QSAR model is measured by the values obtained for certain statistical parameters such as standard error (SE) of estimate (s), Fischer's ratio (F), and conventional regression coefficient ( $r^2$ ). The model generated through MLR method was further investigated for its robustness as well as predictability through PLS approach that utilized the very training set compounds that were employed to perform MLR. PLS analysis is considered as an alternative approach to MLR that is applied to enhance the information content in the generated model and to alleviate the chances of overfitting [26].

### Validation

It is likely for a model, in certain cases, to be generated by chance. Therefore, to mitigate these chances, it is highly recommended to validate the developed model through diverse methods. The proposed study has incorporated two validation techniques through which the reliability and the predictability of the QSAR model can be ensured. First method was the application of cross-validation analytical technique that uses "leave-one-out" approach [27] that involves the removal of single compound at a time from the training set and, thereby, determining the predictive power of the model by employing the remaining compounds left in the training set, defining the model through the previously mentioned statistical parameters. Second approach was the internal validation in which the predictability of the generated model was assured through employing the test set compounds that were not made

part of the model building and were used as an external means of QSAR model validation.

### Designing of the novel compounds

Through the careful analysis of the structures of the selected series collectively with the insights bestowed by the descriptors, that built the QSAR model, and their correlation with the JNK2 inhibitory activity, a set of novel compounds (not reported here), in unison with the Lipinski's rule of five [28], was successfully designed and their screening was done by molecular docking experiments. Libdock score was used as a parameter to define the affinity of molecule toward JNK2 binding domain.

### Molecular docking experiments

Furthermore, to get insights into the potential interactions between the designed lead compounds and the binding domain of the receptor; molecular docking studies were performed. In the present study, high-resolution JNK2 receptor protein data bank (PDB) entry "37e0" was employed to study the binding interactions of the designed molecules. Molecular docking experiments, being easily reproducible, are among the most reliable techniques to study the putative binding of the ligand with the enzymatic receptor. This method is a very economical way of drug discovery as the irrelevant expenditure of money to carry out synthesis, and intrinsic enzymatic assays can be prevented and only those compounds exhibiting strong binding with the receptor are chosen for synthesis.

### Tanimoto similarity index

Tanimoto similarity index, being a reliable measure of the novelty of a molecule, was calculated for the designed molecules. It is an inbuilt tool in discovery studio that, through fingerprinting approach, compares the reported ligands obtained from PDB and calculates the extent of similarity between these ligands and the molecule under probe.

### Absorption, distribution, metabolism, and elimination (ADME) prediction studies

Since ADME properties are known to play a crucial role in bioavailability and, hence, the effectiveness of the drug molecule, their prediction is of paramount importance to develop a drug molecule with adequate pharmacokinetic

and safety profile. The designed molecules were, therefore, analyzed for their ADME properties using a publically available software program "swissADME." It is a reliable ADME properties predictor that evaluates the molecular structure and estimates values of different parameters utilizing which we can clearly predict the drug-likeness of an individual molecule [29].

## RESULTS AND DISCUSSIONS

Eventually, after completion of data reduction step, QSAR model was built by a set of four explanatory variables or descriptors including moment of inertia 2 length (Subst. 1), moment of inertia 3 length (Subst. 3), Kier Chi4 (path/cluster) index (whole molecule), and vamp highest occupied molecular orbital (HOMO) (whole molecule) (Fig. 2). All the four descriptors, included in the final model, manifested excellent correlation with the biological activity. Out of these, two descriptors: Inertia moment 2 length and inertia moment 3 length exhibited high correlation values of 0.7 and 0.8, respectively, with the inhibitory activity. The other two descriptors Kier Chi4 (path/cluster) index (whole molecule) and vamp HOMO (whole molecule) also showed significantly, but somewhat lower, values of correlation than the inertia moment descriptors and are shown in the correlation matrix in Table 1. Moreover, no correlation was found among the obtained descriptors, further, providing the evidence that if these are used in unification, model with a good fit can be obtained. The importance of the aforementioned descriptors was also estimated by the calculation of their t-test values, jackknife SE, coefficient values, and covariance SE values (Table 2). In addition, the predictability of the developed model was confirmed through iterative shuffling of the training set and the test set. Furthermore, the reliability of the generated model was assessed statistically by its  $r^2$  value. Among all the generated model, the one exhibiting the best values (Table 3) for the all the crucial statistical parameters was finally selected for the application of MLR and PLS method. The best model was composed of 23 compounds in the training set, and 8 compounds in the test set Tables 4 and 5.

A set of 23 training set compounds when subjected to MLR analysis (Graph 1) utilizing TSAR 3.3, generated the following regression

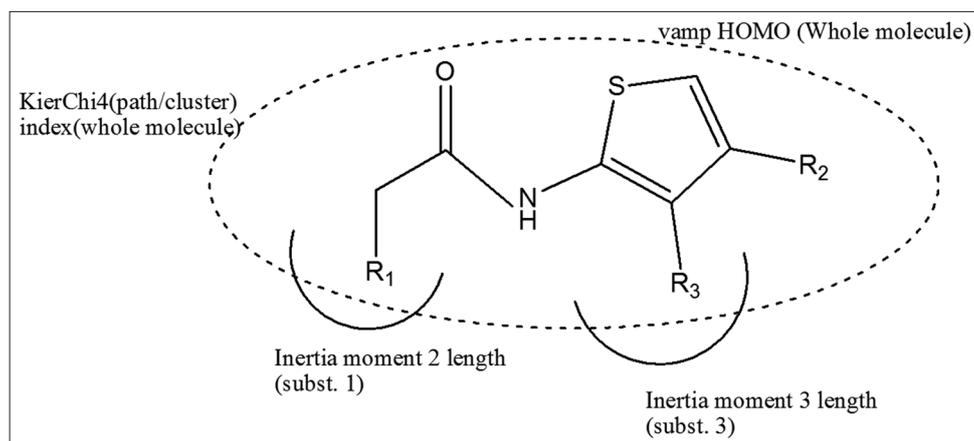


Fig. 2: Correlation of obtained descriptors with different substitution positions

Table 1: Representing the correlation matrix describing the relationship of the obtained descriptors with JNK2 inhibitory activity

Variables	$-\log IC_{50}$	Inertia moment 2 length ( $R_1$ )	Inertia moment 3 length ( $R_3$ )	Kier Chi4 (path/cluster) index (Whole molecule)	vamp HOMO (whole molecule)
$-\log IC_{50}$	1	0.7448	-0.85683	0.57121	-0.52284
Inertia moment 2 length	0.74448	1	-0.65012	0.47593	-0.17457
Inertia moment 3 length	-0.85683	-0.65012	1	-0.33181	0.50122
Kier Chi4 (path/cluster) index	0.57121	0.47593	-0.33181	1	0.10703
vamp HOMO (whole molecule)	-0.52284	-0.17457	0.50122	0.10703	1

JNK2: c-Jun NH2-terminal kinases, HOMO: Highest occupied molecular orbital

equation, equation 1, through relating the data set obtained by the descriptors left in the model with the biological activity:

$$Y = 0.4219918 * X_1 - 4.5166173 * X_2 + 1.4543194 * X_3 - 1.6898502 * X_4 - 18.533213$$

Where, X<sub>1</sub> = moment of inertia 2 length (substituent 4), X<sub>2</sub> = moment of inertia 3 length (whole molecule), X<sub>3</sub> = Kier Chi4 (path/cluster) index (whole molecule), and X<sub>4</sub> = vamp HOMO (whole molecule) and Y represents the biological activity.

In addition, PLS method was employed to derive the regression equation. It was performed to ensure that there is the least deviation between the results obtained from MLR and PLS. The equation obtained through PLS was also subjected to evaluation by considering the values of statistical parameters like r<sup>2</sup> to assure the quality of the generated model (Graph 2).

Regression equation obtained by PLS method generated Equation 2:

$$Y = 0.613X_1 - 4.036X_2 + 0.984X_3 - 1.922X_4 - 20.086$$

Experimentally obtained log IC<sub>50</sub> values for the compounds, the training set compounds and the test set compounds, along with their respective predicted values from both MLR and PLS methods are shown in Tables 4 and 5, respectively.

After successfully building, and performing validation, of the QSAR model, the results were analyzed and the information deciphered by the descriptors, included in the final model, was utilized to understand the influence of substitution variations on the biological activity (Fig. 3). This was done as it was extremely important to know which groups are contributing toward increment and which are detrimental to the inhibitory activity so that, they can be replaced to get optimized JNK2 inhibitors.

Moment of inertia 2 length (Subst. 1), which defines the optimal bulk and shape required to have a molecule aligned with the active site of the receptor and show binding affinity toward it is positively contributing toward the model suggesting the increase in its value will lead to augmentation in the biological activity of the compounds. Based on such information we can infer that those groups, contributing toward the substitution pattern of the molecule, which can increase the shape as well as mass distribution, at the R<sub>1</sub> position, are expected to make a significant contribution toward enhancing the activity profile of the molecule.

On the other hand, moment of inertia 3 length was found to be negatively correlating to the 3<sup>rd</sup> substitution position, R<sub>3</sub>, which implies that minimizing the shape or bulk at this position will have a positive impact on the activity of the molecules under study.

The shape of a molecule plays a crucial role in its orientation and, thereby, its interaction with the binding site of the receptor. On the other

hand, the mass of the molecule is accountable for ADME properties as well as its toxicity. Therefore, a drug to possess both safety and efficacy must have an optimal shape and mass.

Kier Chi4 (path/cluster) index (whole molecule) belongs to the molecular connectivity indices class of descriptors. This descriptor was developed and refined by the scientists, Kier and Hall. It is presented by a series of numbers known as "orders" and by the type of subgraph (path, cluster, path/cluster, and chain). Normally, steric hindrance due to the presence of bulky substituents in a molecule poses a barrier in establishing bonds with a receptor binding site, but at the same time an optimal bulk and branching aids in the desirable orientation of the compound, thereby, increasing its chances of effective bonding interactions with the receptor.

As it is clear from the obtained t-values of the descriptors, included in the final model, that Kier Chi4 (path/cluster) index is having the highest t-values among other descriptors entered in the model. This predicts its importance in the designed model, and as it is positively correlated to the whole molecular structure, addition of the bulky groups or enhancing the branching can be expected to augment the biological activity of the compounds.

Vamp HOMO descriptor was included in the final model and describes the highest level of energy in molecules with electrons. It finds its significance in the reactivity and other properties of a molecule. The values of this descriptor describe a molecule either to be a strong or a weak nucleophile. As it can be understood from the statistical data, obtained through the physicochemical parameters that were used to build our model, that the vamp HOMO is correlating negatively to the biological activity for the whole molecule, as inferred from the t-value of this descriptor, therefore, an overall decrease in the nucleophilic character is desirable to improve the biological activity of the selected compounds.

After in-depth analysis of the results and structures of the selected series, intriguing facts of the structures, showing a remarkable correlation with the descriptors, came into our notice. As we observed that moment of inertia 2 length (Subst. 1), through its positive correlation with the activity of the molecules, indicated that if the bulk and the shape of the substituents, at the R<sub>1</sub> position, is increased, then an increase in activity can be achieved. This fact was found to be true when meticulous study of the different substitution pattern around the molecules was done. For an instance, when an aryl group, from analog 4, was replaced by a fused heterocyclic substituent and when bulk, as well as branching, was increased at this position, as in analog 9, a dramatic increase in potency was observed. This fact became quite clear from the structures of compounds 4, 5, 6, 8, 16, 18, 27, 28, and 29 of the series, which exhibited a lower value of this descriptor (Table 6). Furthermore, unlike rest of the molecules of the series, these are not substituted with fused cyclic structures such as quinolones, isoquinolines, and quinoxaline

**Table 2: Depicting the significance of descriptors used for building the QSAR model**

Descriptors	Coefficient <sup>a</sup>	Jackknife <sup>b</sup>	Covariance SE <sup>c</sup>	t-value <sup>d</sup>	t-probability <sup>e</sup>
Inertia moment 2 length	0.42199	0.1802	0.14549	2.9004	0.00953
Inertia moment 3 length	-4.5166	0.68783	1.01	-4.4719	0.00029
Kier Chi4 (path/cluster) index	1.4543	0.24587	0.24117	6.0302	1.059e-005
vamp HOMO (whole molecule)	-1.6899	0.53453	0.58935	-2.8673	0.10243

<sup>a</sup>Represents the regression coefficient for individual variables in the QSAR equations. <sup>b</sup>Represents an estimate of the SE on each regression coefficient derived from a jackknife method on the final regression model. <sup>c</sup>Represents an estimate of the SE on each regression coefficient derived from covariance matrix. <sup>d</sup>Is the measure of the significance of each variable included in the final model. <sup>e</sup>Represents statistical significance for t-values. SE: Standard error. HOMO: Highest occupied molecular orbital

**Table 3: Obtained values of significant statistical parameters for the developed model**

Model type	Test set compounds	r	r <sup>2</sup>	r <sup>2</sup> CV	s value	f value
Final model	4,6,,23,29,43,44,50	0.96	0.92	0.82	0.28	59.69

Table 4: Training set compounds included in the final model

Name of the compound	IC <sub>50</sub> (nm)	Predicted values	
		MLR	PLS
5	22500	-4.0704	-3.9940
9	28400	-4.3828	-4.2531
15	1838	-3.3731	-3.5167
16	4524	-3.3671	-3.4136
17	613	-3.43	-3.4658
18	2571	-3.4323	-3.4635
19	4150	-3.3478	-3.3175
20	765	-3.0392	-2.9726
21	409	-2.7746	-2.6623
22	863	-2.7933	-2.6800
27	824	-3.3891	-3.4146
28	1300	-3.43	-2.7766
30	230	-2.1641	-2.1702
31	81	-2.0460	-2.0487
32	125	-2.0344	-1.9499
33	42	-1.7387	-2.0887
34	31	-1.4689	-1.6786
45	8	-0.9754	-1.1355
46	50	-1.6935	-1.5671
47	43	-1.8163	-1.7088
48	26	-1.5397	-1.4035
49	90	-1.7169	-1.5848
51	41	-1.5456	-1.4344

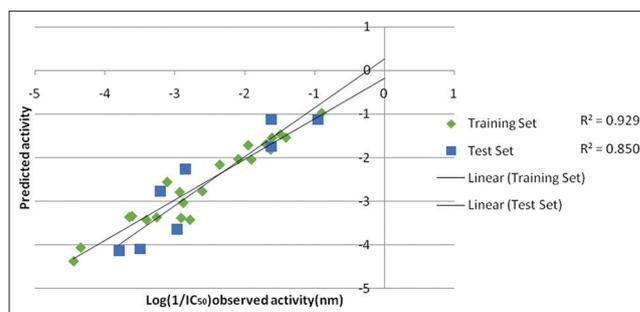
MLR: Multiple linear regression, PLS: Partial least square

Table 5: Test set compounds excluded from the model and used for validation

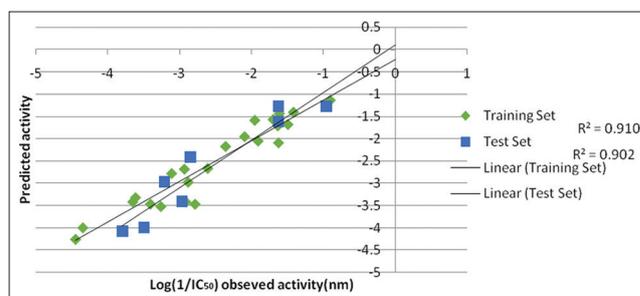
Name of the compound	IC <sub>50</sub> (nm)	Predicted values	
		MLR	PLS
4	3140	-4.0933	-3.9932
6	6290	4.1371	-4.0736
8	930	-3.6397	-3.4066
23	703	-2.2614	-2.4157
29	1620	-2.7728	-2.9688
43	42	-1.1138	-1.2771
44	9	-1.1266	-1.2841
50	42	-1.7436	-1.6162

MLR: Multiple linear regression, PLS: Partial least square

rings at the R<sub>1</sub> position and, probably, due to this reason manifested a lower activity profile. Likewise, those compounds, such as 15 and 17, which are not substituted with the fused cyclic substituents, but with somewhat bulkier substituents consisting of multiple numbers of rings and chains, also exhibited activity profile comparable to the compounds with fused rings substituted at the R<sub>1</sub> position. Whereas, moment of inertia 3 length, showing negative correlation at the R<sub>3</sub> position, implied that an increase in bulkiness as well as the shape of a molecule at this position will eventually lead to the decrease in the activity and hence, potency of the compounds. This was also found to be realistic because when ester group was replaced by a triazole ring at this specific position, it gave a remarkable hike in the activity of the compounds. Interestingly, both the ester and the triazole ring have comparable molecular mass and, thus, no additional bulk was introduced to the molecule. Therefore, just lowering the bulk at this substitution position cannot be accounted for an augmented activity profile. The probable cause for improved activity, perhaps, can be the flat structure of the triazole ring, which led to the smaller shape of the compound that helped it to easily enter into the binding domain and fit closely with the walls of the active site. The importance of Kier Chi (path/cluster) descriptor of the order four can be understood if we again notice an increase in activity profile, facilitated by an increase in bulk at the R<sub>1</sub> position. This substitution pattern, where fused ring replaced the single aryl ring, eventually led to



Graph 1: Multiple linear regression plots between the observed and the predicted activity of the test set and the training set compounds



Graph 2: Partial least square plot between the observed and the predicted activity of the test set and the training set compounds

an increase in overall bulk of the molecule that, in turn, resulted in the enhanced potency of the molecules. Furthermore, for an instance, when methyl group was added to the 5<sup>th</sup> position on the triazole ring, as in compound 31 and 33, an increased potency was seen when compared to an analog bearing hydrogen atom at this position, as in compound 32 and 34. This augmentation in the activity profile of the molecules, through increasing the mass distribution on the entire molecule, made us to infer that the presence of this descriptor in the final model is of immense relevance. Finally, the significance of the vamp HOMO descriptor came into notice when a nitrile group introduced on the R<sub>2</sub> position, in analog 9 with methoxyphenyl group substituted at R<sub>1</sub> position, brought about a remarkable increase in the potency whereas, other methoxyphenyl substituted analogs did not show an increment in the activity, to such an extent, when methyl or ethyl groups, as in analogs 4,5,6,8, were introduced at the same position (R<sub>2</sub>). In addition, this methoxyphenyl analog with cyano group substitution at R<sub>2</sub> was found to have greater value for vamp HOMO descriptor, as can be seen from Table 6, than other analogs with methyl or ethyl group at R<sub>2</sub> position and similar substitutions, as that in analog 9, at the other two positions (R<sub>1</sub> and R<sub>3</sub>). A similar hike in activity profile was witnessed when halogen groups or a nitrile group replaced methyl group at the same position, R<sub>2</sub>. These analogs, with cyano or halogen groups placed at R<sub>2</sub> position, constitute the group of compounds exhibiting the best JNK2 inhibitory activity in this series of compounds (Fig. 3), further enhancing the reliability of the developed model, which indicated an increase in the activity profile of the molecules with decrease in their nucleophilic character or increase in the electrophilic behavior. The compounds, with electronegative groups such as chloro and cyano, were found to exhibit excellent JNK2 inhibitory activity but, the compounds with bromo group were found to be comparatively less potent. The remarkable hike in the inhibitory activity must be attributed to the overall decrease in the nucleophilicity, due to the presence of highly electronegative groups, on the R<sub>2</sub> position of the molecules. It has been assumed that this decrease in nucleophilicity capacitated the molecule to accept electron pairs and to form covalent bonds with the receptor binding domain. It is believed that a molecule, that is capable of making hydrogen bonds as well as covalent bonds, exhibits better bioactivity

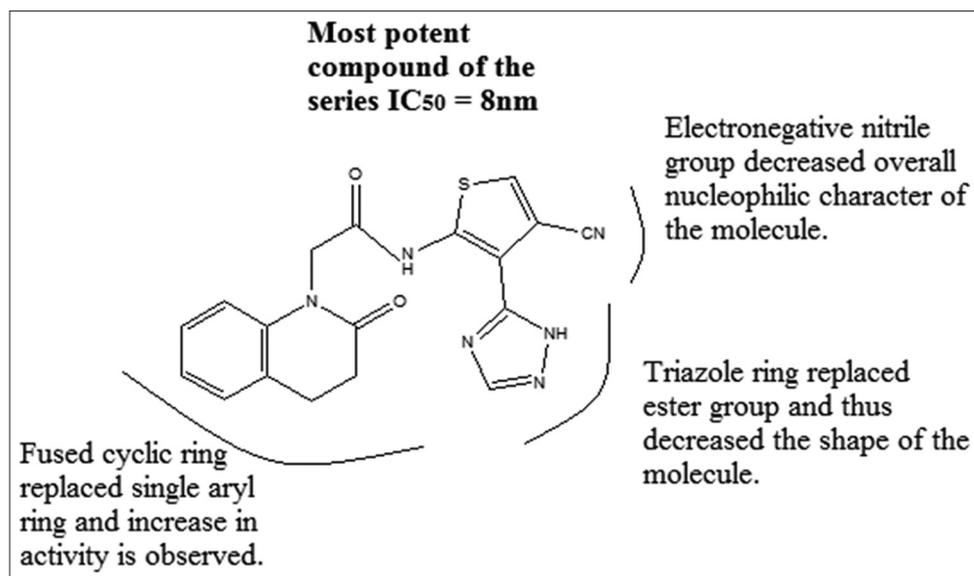


Fig. 3: Effects of changes in substitution on molecular properties of compound 45 and, hence, inhibitory activity

Table 6: Value of descriptors for individual compounds

Name of the compound	Inertia moment 2 length (R <sub>1</sub> )	Inertia moment 3 length (R <sub>2</sub> )	Kier Chi4 (path/cluster) index (whole molecule)	vamp HOMO (whole molecule)
4	2.47324	1.29969	3.12238	-8.71408
5	2.37532	1.29058	3.12797	-8.72288
6	2.52301	1.32989	3.22743	-8.66602
8	2.39292	1.33508	3.05893	-9.1517
9	2.36135	1.33771	3.05893	-8.72688
15	2.22349	1.29253	3.57239	-8.79622
16	2.57845	1.29035	3.4956	-8.77136
17	2.71522	1.2913	3.4956	-8.70254
18	2.74655	1.29203	3.4956	-8.6953
19	3.55085	1.29265	3.58211	-8.47165
20	3.51664	1.29121	3.58211	-8.65896
21	3.59092	1.29071	3.58211	-8.79565
22	3.59643	1.29327	3.58211	-8.79007
23	3.5534	1.29895	4.09138	-8.69248
27	2.54694	1.23303	3.35928	-8.73031
28	2.54225	1.23175	3.80929	-8.83065
29	2.53047	1.23033	3.73251	-8.77073
30	3.71936	1.23557	3.81901	-8.77356
31	3.64884	1.2324	3.81901	-8.85264
32	3.7584	1.08427	3.46985	-8.73672
33	3.35507	1.22423	4.32829	-8.64769
34	3.5395	1.08306	3.97912	-8.68449
43	3.52006	1.08319	3.97912	-8.89982
44	3.5729	1.0836	3.97912	-8.88018
45	3.43332	1.08271	3.98471	-8.99727
46	3.71775	1.08474	3.46985	-8.94985
47	3.7028	1.08477	3.46985	-8.88099
48	3.67914	1.08254	3.47544	-9.0398
49	3.78601	1.08459	3.46985	-8.91857
50	3.77593	1.08483	3.46985	-8.90593
51	3.50014	1.08226	3.47544	-9.08026

HOMO: Highest occupied molecular orbital

due to the fact that the bonding exhibited by it will be stronger but at the same time expected to be reversible. Both these features make a molecule selective, and at the same time less toxic.

Through unification of the information, derived from in-depth study of the descriptors, we deduced that an optimal increase in the bulk, which facilitates the proper alignment and binding of the molecule to the receptor, at certain positions of the molecule and by introducing electrophilic groups and, thereby, decreasing its overall nucleophilic nature, resulted in

remarkable increase in the JNK2 inhibitory activity. An in-depth study of the relationship between these descriptors and the substitution pattern of the selected set of molecules, exhibiting potent inhibitory activity against JNK2, provided us with the insights, which were minute but extremely significant if the structural dependency of biological activity had to be deciphered. The present work decoded the fact that not only the substitution but, its position also plays an important role in the activity, exhibited by a molecule. This point was inferred when we observed a mismatch between the functions of the two descriptors, Kier Chi4 (whole molecule) and the moment of inertia

3 length (subs  $R_3$ ). Where, on the one hand, positive correlation of Kier Chi4 predicted an increase in bioactivity of the compounds, through enhancing the mass distribution and, hence, bulk of the entire molecule, on the other hand, negatively correlating moment of inertia 3 length pointed toward reducing the bulk and shape of the molecule, at position  $R_3$ , to improve the biological activity of the selected compounds. This clearly states that even if Kier Chi4 descriptor predicted an increase in activity with an overall increase in bulk of the molecule but, if bulk is placed at  $R_3$  position, instead of  $R_1$ , then, detrimental effects are expected to be observed on the JNK2 inhibitory activity of the molecules, as predicted by inertia moment 3 length descriptor. Therefore, an optimal increase in the bulk or branching, at only certain positions, will bring about an increase in bioactivity of the molecules. The above results explicitly indicates that all the descriptors, that entered the final 2D QSAR model, were crucial and their correlation with the biological activity concurred well to the substitutions executed in the structures of the selected series of the JNK2 inhibitors. An in-depth study of these physicochemical parameters has provided substantial insights to design better chemical scaffolds in terms of selectivity and efficacy.

By utilizing the information, deciphered by the derived descriptors, two molecules, compound a and compound b, were designed (Fig. 4). Interestingly, both the molecules exhibited better ADME profile, than the most active compound (compound 45) in the selected series of JNK2 inhibitors.

A crystal structure of JNK2 enzyme, PDB entry name "3e70," with the excellent resolution was selected to perform the molecular docking experiments. All the designed compounds were docked on to the binding domain of the enzymatic receptor. The docking experiments were proved to be fruitful and the two compounds, compound a and compound b, exhibited high Libdock score of 129.74 and 138.16, respectively. Deeper insights, into the binding modes of the molecules,

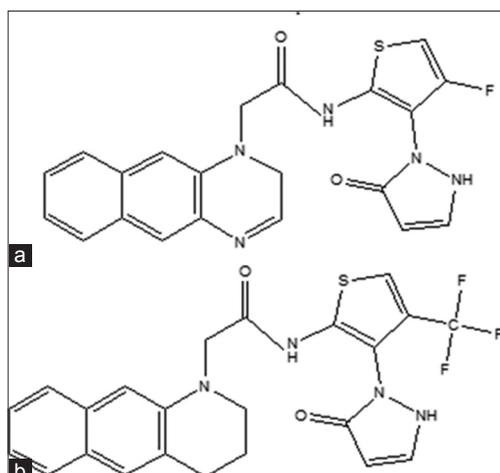


Fig. 4: Discovered optimized compounds, compound a and compound b.

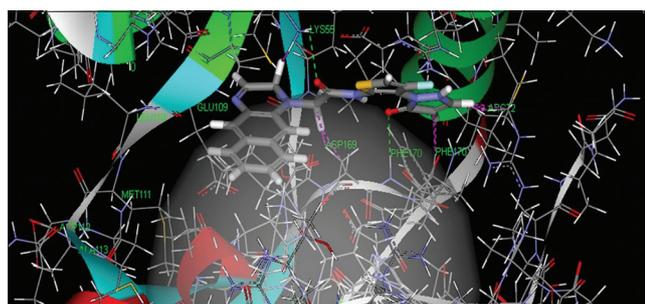


Fig. 5: Docking studies of compound a on c-Jun NH2-terminal kinases receptor PDB entry 3e70

revealed that compound a fits itself in adenine binding region, near the gatekeeper residue MET111 and its tail facing toward phosphate binding region of the receptor (Fig. 5). It was observed to be held firmly, through strong hydrogen bonds, with Phe170 and Lys55 amino acids. Additional Van der Waals bonds were formed with ARG72 and ASP69 amino acid residues. Binding mode evaluation of compound b revealed its binding position to be the same as that observed for compound a and was found to be forming Van der Waals bonds with exactly the same amino acids (Fig. 6). In conclusion, both the compounds get themselves seated deep in the receptor binding site, through the strong hydrogen bonds as well as Van der Waals forces bonding with amino acids in both the planes, further enhancing the stable fitting of the molecules to the receptor and, therefore, effective enzyme inhibition. Docking of compound 45 has also been shown in Fig. 7.

The novelty of the aforementioned compounds, compound a and compound b, was ascertained through evaluating the Tanimoto index for both the compounds. Moreover, the values of Tanimoto index for compound a and compound b were found to be 0.095 and 0.075, respectively, clearly suggesting that the designed compounds are highly novel.

In addition to this, an online available ADME prediction software "swissADME" was employed to determine the pharmacokinetic properties of the designed molecule. According to the results obtained (Table 7), the drug-likeness of both the molecules was ascertained. Interestingly, both the molecules were predicted as non-Pgp substrate whereas compound 45 was found to be a Pgp substrate. Furthermore, the designed compounds were not found to violate any of the laws proposed to decide the extent of drug-likeness of a molecule whereas, compound 45 manifested violations to Egan *et al.*[30] as well as Veber *et al.* rules[31]. In addition, the absorption of the designed molecules, from gastrointestinal tract, is predicted to be high whereas, for the compound 45, it is low. This further strengthens the probability of designed molecules to show a better pharmacokinetic profile as compared to the reported inhibitors.

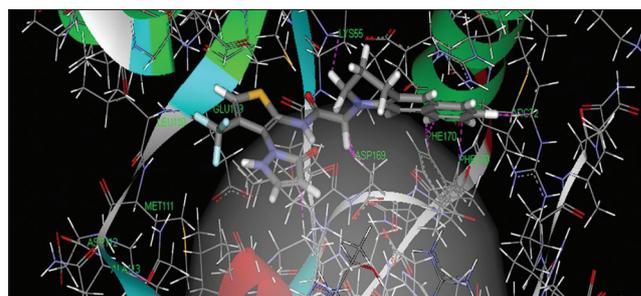


Fig. 6: Docking studies of compound b on c-Jun NH2-terminal kinases receptor PDB entry 3e70

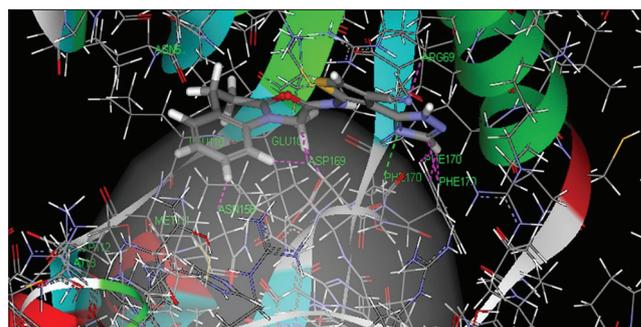


Fig. 7: Docking studies of compound 45 on c-Jun NH2-terminal kinases receptor PDB entry 3e70

**Table 7: The values of diverse parameters predicting ADME properties of designed compounds (compound a and compound b)**

Parameters	Compound a	Compound b	Compound 45
Molecular weight	474.5	423.46	378.41
Hydrogen bond donors	2	2	2
Hydrogen bond acceptors	6	5	5
iLOGP	3.13	2.92	1.79
XLOGP	3.8	1.69	1.75
GI absorption	High	High	Low
Pgp substrate	NO	NO	YES
Lipinski's violation	0	0	0
Ghose violations	0	0	0
Weber violations	0	0	1
Egan violations	0	0	1

ADME: Absorption, distribution, metabolism, and elimination

## CONCLUSION

The QSAR model, developed through the application of MLR and PLS methods, exhibited excellent values for standard statistical parameters and, thereby, proved to be highly reliable and predictable. The descriptors that were included in the final model, and their correlation with the JNK2 inhibitory activity was exploited to carry out replacements of the original groups, present around the N-(3,4-dimethylthiophen-2-yl)acetamide moiety, with different groups that were predicted to bestow optimized activity profile to the existing JNK2 inhibitors. The designed molecules, when imposed on the JNK2 enzymatic receptor, unraveled the potential binding interactions which were found to be better than the most active compound of the selected series of compounds. Eventually, two molecules, exhibiting excellent Libdock score, novelty and predicted ADME properties, have been designed for synthesis and further biological evaluation studies.

## ACKNOWLEDGMENT

The authors are extremely grateful to the vice-chancellor, Banasthali Vidyapith, for being so cooperative and for providing the best possible resources to carry out the presented work. Authors also thankful to our colleagues for extending a hand of support whenever they required.

## AUTHORS CONTRIBUTIONS

Both the authors have contributed equally to the manuscript. Research, as well as writing of this manuscript, is the result of sheer dedication of the contributing authors.

## CONFLICTS OF INTEREST

Authors have no conflicts of interest.

## REFERENCES

1. Yang DD, Kuan CY, Whitmarsh AJ, Rincón M, Zheng TS, Davis RJ, *et al.* Absence of excitotoxicity-induced apoptosis in the hippocampus of mice lacking the jnk3 gene. *Nature* 1997;389:865-70.
2. Hunot S, Vila M, Teismann P, Davis RJ, Hirsch EC, Przedborski S, *et al.* JNK-mediated induction of cyclooxygenase 2 is required for neurodegeneration in a mouse model of Parkinson's disease. *Proc Natl Acad Sci U S A* 2004;101:665-70.
3. Chang L, Jones Y, Ellisman MH, Goldstein LS, Karin M. JNK1 is required for maintenance of neuronal microtubules and controls phosphorylation of microtubule-associated proteins. *Dev Cell* 2003;4:521-33.
4. Bjorkblom B, Ostman N, Hongisto V, Komarovski V, Filen JJ,

- Nyman TA, *et al.* Constitutively active cytoplasmic c-Jun N-terminal kinase 1 is a dominant regulator of dendritic architecture: Role of microtubule-associated protein 2 as an effector. *J Neurosci* 2005;25:6350-61.
5. Tararuk T, Ostman N, Li W, Björkblom B, Padzik A, Zdrojewska J, *et al.* JNK1 phosphorylation of SCG10 determines microtubule dynamics and axodendritic length. *J Cell Biol* 2006;173:265-77.
6. Morishima Y, Gotoh Y, Zieg J, Barrett T, Takano H, Flavell R, *et al.* Beta-amyloid induces neuronal apoptosis via a mechanism that involves the c-Jun N-terminal kinase pathway and the induction of Fas ligand. *J Neurosci* 2001;21:7551-60.
7. Fogarty MP, Downer EJ, Campbell V. A role for c-Jun N-terminal kinase 1 (JNK1), but not JNK2, in the beta-amyloid-mediated stabilization of protein p53 and induction of the apoptotic cascade in cultured cortical neurons. *Biochem J* 2003;371:789-98.
8. Kimberly WT, Zheng JB, Town T, Flavell RA, Selkoe DJ. Physiological regulation of the beta-amyloid precursor protein signalling domain by c-Jun N-terminal kinase JNK3 during neuronal differentiation. *J Neurosci* 2005;25:5533-43.
9. Ozcan U, Cao Q, Yilmaz E, Lee AH, Iwakoshi NN, Ozdelen E, *et al.* Endoplasmic reticulum stress links obesity, insulin action and Type 2 diabetes. *Science* 2004;306:457-61.
10. Jaeschke A, Rincón M, Doran B, Reilly J, Neuberger D, Greiner DL, *et al.* Disruption of the jnk2 (Mapk9) gene reduces destructive insulinitis and diabetes in a mouse model of type I diabetes. *Proc Natl Acad Sci U S A* 2005;102:6931-5.
11. Li Y, Schwabe RF, DeVries-Seimon T, Yao PM, Gerbod-Giannone MC, Tall AR, *et al.* Free cholesterol-loaded macrophages are an abundant source of tumor necrosis factor-alpha and interleukin-6: Model of NF-kappaB- and map kinase-dependent inflammation in advanced atherosclerosis. *J Biol Chem* 2005;280:21763-72.
12. Mohit AA, Martin JH, Miller CA. P493F12 kinase: A novel MAP kinase expressed in a subset of neurons in the human nervous system. *Neuron* 1995;14:67-78.
13. Dérjard B, Hibi M, Wu IH, Barrett T, Su B, Deng T, *et al.* JNK1: A protein kinase stimulated by UV light and ha-ras that binds and phosphorylates the c-jun activation domain. *Cell* 1994;76:1025-37.
14. Kallunki T, Su B, Tsigelny I, Sluss HK, Dérjard B, Moore G, *et al.* JNK2 contains a specificity-determining region responsible for efficient c-Jun binding and phosphorylation. *Genes Dev* 1994;8:2996-3007.
15. Gupta S, Barrett T, Whitmarsh AJ, Cavanagh J, Sluss HK, Dérjard B, *et al.* Selective interaction of JNK protein kinase isoforms with transcription factors. *EMBO J* 1996;15:2760-70.
16. Karin M, Gallagher E. From JNK to pay dirt: Jun kinases, their biochemistry, physiology and clinical importance. *IUBMB Life* 2005;57:283-95.
17. Karin M, Lawrence T, Nizet V. Innate immunity gone awry: Linking microbial infections to chronic inflammation and cancer. *Cell* 2006;124:823-35.
18. Han Z, Boyle DL, Chang L, Bennett B, Karin M, Yang L, *et al.* C-jun N-terminal kinase is required for metalloproteinase expression and joint destruction in inflammatory arthritis. *J Clin Invest* 2001;108:73-81.
19. Sumara G, Belwal M, Ricci R. "Jnking" atherosclerosis. *Cell Mol Life Sci* 2005;62:2487-94.
20. Ricci R, Sumara G, Sumara I, Rozenberg I, Kurrer M, Akhmedov A, *et al.* Requirement of JNK2 for scavenger receptor A-mediated foam cell formation in atherosclerosis. *Science* 2004;306:1558-61.
21. Bowers S, Truong AP, Neitz RJ, Neitzel M, Probst GD, Hom RK, *et al.* Design and synthesis of a novel. Orally active, brain penetrant, tri-substituted thiophene based JNK inhibitor. *Bioorg Med Chem Lett* 2011;21:1838-43.
22. Zhivkova Z, Doytchinova I. *In silico* quantitative structure-pharmacokinetic relationship modeling on acidic drugs: Half Life. *Int J Pharm Pharm Sci* 2014;6:283-9.
23. Cronin MT, Schultz TW. Development of quantitative structure-activity relationships for the toxicity of aromatic compounds of *Tetrahyena pyriformis*: Comparative assessment of the methodologies. *Chem Res Toxicol* 2001;14:1284-95.
24. Dessalew N. Investigation of the structural requirement for inhibiting HIV integrase: QSAR study. *Acta Pharm* 2009;59:31-43.
25. Kovatcheva A, Buchbauer G, Golbraikh A, Wolschann P. QSAR modelling of r-campholenic derivatives with sandalwood odor. *J Chem Inf Comput Sci* 2003;43:259-66.
26. Kubinyi H. QSAR and 3D-QSAR in Drug Design Part 1: Methodology. *Drug Discov Today* 1997;2:457-67.
27. Jagdale D, Ramaa CS. Quantitative structure-activity relationship

- analysis of novel pyrazoline derivatives using k nearest neighbour molecular field analysis method. *Int J Pharm Pharm Sci* 2017;9:87-91.
28. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 2001;46:3-26.
  29. Daina A, Michielin O, Zoete V. Swiss ADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep* 2017;7:42717.
  30. Egan WJ, Kenneth MM, Baldwin JJ. Prediction of drug absorption using multivariate statistics. *J Med Chem* 2000;43:3867-77.
  31. Veber DF, Stephen R, Cheng JH, Smith BR, Ward KW, Kopple KW. Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* 2002;45:2615-623.