Original Article

Quantitative Structure–Activity Relationship Analysis of Thiophene Derivatives to Explore the Structural Requirements for c-Jun NH_2-Terminal Kinase 1 Inhibitory Activity

Ashima Nagpal, Monika Chauhan[1]

Department of Pharmacy, G. D. Goenka University, Sohna, Haryana, [1]Department of Pharmacy, Banasthali Vidyapith, Banasthali, Rajasthan, India

Abstract

Background: With an aim to design a validated two-dimensional quantitative structure–activity relationship (2D QSAR) model, a probe was executed on a series of reported c-Jun NH_2-terminal kinase-1 (JNK1) inhibitors, exhibiting selectivity toward JNKs (and not other members of MAPK family). **Objective:** The present work focused on obtaining valuable insights from the structural architecture of the selected compounds and their effects on JNK1 inhibitory activity. The present work deciphers the importance of descriptive variables, namely Verloop L (Subst. 1), Bond Dipole Moment (Subst. 2), LogP (Subst. 1), Balaban Topological index (Subst. 1), and VAMP Total Dipole (whole molecule), in molecules possessing JNK1 inhibitory profile. **Results:** These explanatory variables, obtained after iteratively reducing the data, did not only provide us with the substantial evidence pertaining to the dependence of bioactivity on the structural features of molecules, but also suggested the measures to optimize the selected compounds so as to obtain potent JNK1 inhibitors with good selectivity profile. Based on these distinct descriptors, exhibiting no apparent intercorrelation and manifesting good correlation with biological activity, a 2D QSAR model was generated. **Conclusion:** Robustness of the developed model was evaluated by performing multiple linear regression, partial least square, and artificial neural network studies. The reliability and predictive ability of the developed model was ascertained through the values of standard statistical parameters, such as s = 0.38, F = 97.22, r = 0.95, r^2 = 0.90, and r^2cv = 0.88, for the training set compounds. The generated model was validated through the test set compounds, as well as by leave one out method.

Keywords: *Artificial neural network, descriptors, insulin receptor substrate, multiple linear regression, partial least square, quantitative structure–activity relationship*

Introduction

An increase in the number of type 2 diabetes cases has been attributed, majorly to a sedentary lifestyle. Changes in insulin sensitivity as well as insulin levels are the two major hallmarks of type 2 diabetes and are responsible for glucose intolerance and increase in blood glucose level. Various studies conducted in this realm have revealed a direct connection of increased pro-inflammatory cytokine concentration in the circulation with an increase in peripheral insulin resistance.[1-3] Additionally, activation of a specific stimulus, such as endoplasmic reticulum stress by adipocyte hypertrophy or rise in plasma levels of free fatty acids (FFAs), also contributes toward peripheral insulin resistance.[4,5] At the molecular level, these stimuli trigger various serine/threonine protein kinases, including nuclear factor-kB kinase inhibitor and the c-Jun NH_2-terminal kinase (JNK). Both of these kinases are known to target the insulin receptor substrate (IRS)-1, for serine phosphorylation. Phosphorylation of IRS-1 prevents it to reach and bind to the insulin receptor that, in turn, prevents the activation of insulin-signaling cascade.[6,7]

JNKs constitute the subfamily of mitogen-activated protein kinases and belong to the class of serine/threonine kinases. These are primarily triggered upon their exposure to environmental stimuli, such as ultraviolet irradiation and certain other stimuli including cytokines and osmotic shock.[8-12] JNKs perform their functions by phosphorylating the N-terminal transactivation domain of c-Jun that results in the activation of c-Jun-dependent transcriptional processes. JNK1/SAPKb, JNK2/SAPKa, and JNK3/SAPKg are the

How to cite this article: Nagpal A, Chauhan M. Quantitative structure–activity relationship analysis of thiophene derivatives to explore the structural requirements for c-Jun NH_2-terminal kinase 1 inhibitory activity. J Rep Pharma Sci 2019;8:115-23.

Address for correspondence:
Dr. Ashima Nagpal, Department of Pharmacy, G. D. Goenka University, Sohna - 122 103, Haryana, India.
E-mail: ashima_nagpal@yahoo.com

Access this article online

Website: www.jrpsjournal.com

DOI: 10.4103/jrptps.jrptps_32_18

Quick Response Code:

three genes that are known to encode the JNK family. Out of these three isoforms, JNK1 and JNK2 are found to be expressed ubiquitously, whereas JNK3 is primarily localized in the brain and at somewhat lower levels in the heart and testes.[8,11,13,14] A number of studies, performed on the JNK1 gene knock-out mouse models, showed a drastic reduction in adiposity as well as a discernible increase in insulin sensitivity.[15] Therefore, the crucial role of JNK in linking inflammatory responses with metabolic disorder, through the modulation of ER stress and IRS, gives the researchers a promising opportunity to design the inhibitors that can help combat type 2 diabetes.[16] According to a recent study, an orally administered small molecule pan-JNK inhibitor had a discernible impact when compared with rosiglitazone and rimonabant on insulin sensitization, glucose levels, and adiposity with negligible effect on liver enzymes.[17]

JNKs represent a paradigm of enzymes that share an exceptionally high degree of homology. They are known to modulate a wide array of cellular functions. Therefore, complete inhibition of the JNK activity by a nonselective inhibitor will have a profound effect on multiple processes including the ones not involved in the pathophysiology of any disease. This can probably lead to undesired side effects that indeed will be more pronounced in case of chronic conditions. The lack of selectivity in JNK inhibitors is an issue of great concern and must be sincerely addressed to avoid undesirable consequences. Therefore, prompted by the dearth of selective JNK1 inhibitors and adverse effects posed by the inhibitors developed so far, multiple linear regression (MLR), partial least square (PLS), and artificial neural network (ANN) methods, which are the *in silico* tools of quantitative structure–activity relationship (QSAR), have been effectively employed to assess the dependence of JNK1 inhibitory activity on the structural design of the selected molecules. This study was conducted to deduce the significant structural modifications that can be worked out to design the optimized JNK1 inhibitors, possessing excellent potency as well as enhanced selectivity profile.

Materials and Methods

Dataset for analysis

The present work employed IC_{50} data from a series of 77 di-substituted thiophene derivatives,[18,19] showing good JNK1 inhibitory activity. Usually, reported biological activities are skewed and therefore, to eliminate this problem, values of the biological activity were converted into their respective pIC_{50} by using the following formula:

$$pIC_{50} = -\log IC_{50}$$

Sketching of the structures of selected di-substituted thiophene derivatives, using Chemdraw Ultra 8.0 software ((www.perkinelmer.com), USA),[20] marked the onset of QSAR model development. The sketched chemical structures were then imported to the new data sheet of TSAR 3.3 software (www.accelrys.com).[21]

Defining substituents

In total, two major substituents (R1 and R2), around an N-methyl acetamide moiety (common to all molecules), were defined through an in-built option called "define substituents," available in the TSAR worksheet's toolbar (version 3.3; Accelrys Inc., Oxford, England) as depicted in Figure 1. These positions were selected on the basis of impact that these substituents manifested through a discernible change in JNK1 inhibitory activity.

Data set preparation

The present study employed TSAR Version 3.3 to develop the model. Structures of the selected set of molecules after their import to the TSAR worksheet were converted into high-quality 3D structures through an in-built option "Corina make 3D."[22] The Cosmic option, present in TSAR, determined the total energy which is the sum total of Van der Waal, coulombic, bond length, bond angle, and torsion angle terms, for an individual set of atoms.[23] Inclusion of valence electrons present in the molecular atoms was also made in these calculations. The calculations were terminated as soon as the energy gradient became smaller than 1×10^{-5} and 1×10^{-10} kcal/mol.

Descriptor calculation

After optimizing the energy of all the imported structures, TSAR sheet containing nearly 200 classical descriptors from structural, electronic, geometrical, and hydrophobic classes was generated through calculating their numerical values. For calculation of the aforementioned descriptors, whole molecules and their substituents were selected and different types of descriptors, such as KierChi, KierChiV, molecular surface area, topological indices, logP, and electronic descriptors, were calculated. Another feature called Vamp, which is a molecular orbital package of semi-empirical type in TSAR 3.3, was employed to estimate the electrostatic properties, such as electronic energy, total energy, nuclear repulsion energy, atomic charge, accessible

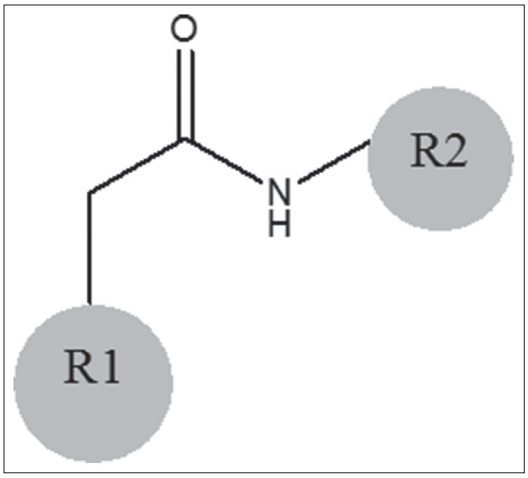


Figure 1: Substituents defined around N-methyl acetamide moiety

surface area, mean polarizability, total dipole, heat of formation, polarizability, and dipole components.

Data reduction

Data redundancy, which is the major cause of deceptive results, is mainly observed when the data are large and lead to ambiguity in choosing the relevant descriptors that actually decipher the enigma, pertaining to the dependence of bioactivity on the structural architecture of the molecules. Therefore, correlation matrix was used to curtail the data to obtain relevant physicochemical parameters exhibiting maximum correlation with the biological activity and no intercorrelation. In this step, a correlation matrix was generated through an in-built option, and the descriptors having intercorrelation were evaluated. Among the intercorrelated descriptors, the one exhibiting higher correlation with the biological activity was retained, whereas the other one was deleted from the sheet. The descriptors that were left after data reduction constituted the final model and were employed to decode the information, encoded by the structures of the molecules under study.

Statistical analysis

The quantification of the relationship between the biological activity and the descriptors that were obtained after data reduction was carried out through the implementation of MLR, PLS, and ANN approaches, available in TSAR 3.3. The criteria "F-to-Enter" and "F-to-leave" particularly explain the significance and insignificance of the role of a variable in the obtained regression equation, respectively, for adding to the equation and removing from the equation. In TSAR Version 3.3, the value for F-to-enter and F-to-leave, by default, is fixed to 4.[24] The evaluation of predictive power of the proposed model was performed through a number of statistical parameters, such as conventional regression coefficient (r), squared regression coefficient (r^2), cross-validation test (r^2cv), standard deviation (s), and Fischer's ratio (F).

Multiple linear regression analysis

MLR involves the calculation of an equation that describes the relation between the biological activity data (dependent Y variable) and the structural descriptors (independent X variable). This method involves fitting of the data, extracted from the dependent as well as independent variables to the derived regression equations.[25]

Partial least square

PLS analysis technique also involves the calculation of the equations explaining the relationship between a dependent variable and a set of descriptors (independent variables). It is considered as a desired tool for surmounting the difficulties common with MLR, owing to redundancy resulted due to a large pool of data or high intercorrelations among descriptors.[26]

Artificial neural network approach

ANN is typically a software-based program. In ANN technique, neurons (processing elements) are connected to each other through links within a structure, which resembles net and form "layers." The features of the ANN are suitable for processing of the data, especially in the cases where the functional relationship between the input and the output is not previously defined or is of nonlinear type.

Model validation

Leave-one-out method was employed for the cross-validation purpose and involved the deletion of one descriptor, at a time, and analyzing the data set values for the obtained model based on the remaining descriptors. The values of r^2 and cross-validated r^2, with least prediction error, were chosen. Additionally, the test set compounds, not included in building of the model, were used to determine the predictability of the developed QSAR model.[27]

Result and Discussion

For the selected set of compounds, approximately 200 descriptors belonging to distinct classes, such as the electronic, shape, lipophilic, and refractivity, were determined numerically. The data set was reduced to eliminate the chances of redundancy. Data reduction provided better understanding of the substitution pattern and how it accounts for the unique behavior of the molecules, against JNK1. Reducing the data, by deleting redundant and intercorrelated descriptive variables, led to the development of QSAR model that consisted of five descriptors, as depicted in Figure 2. After completion of the data reduction step, the selected compounds were partitioned into the training set and the test set Tables 1 and 2. The training set molecules were used to build the model, and their activity values were predicted through MLR, PLS, and ANN analyses. The test set molecules underwent the same fate as that of training set compounds, and their predicted values were employed as a validating tool for the obtained model. Plots between the predicted activity data and the experimental activity data were employed in evaluating the

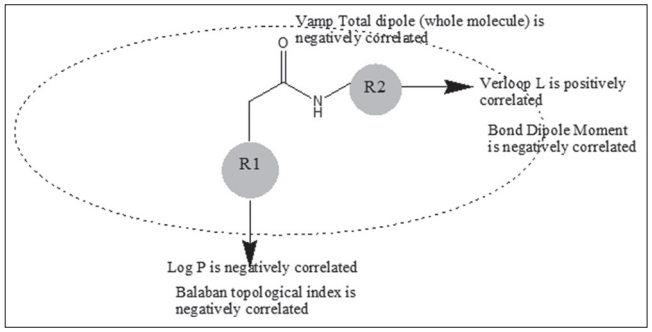


Figure 2: Correlation of the descriptors, used to build the quantitative structure–activity relationship model, with the substituents defined around a common nucleus

Table 1: Representing experimental activity data and the predicted activity values of the training set compounds, obtained from multiple linear regression, partial least square, and artificial neural network methods

Name of the compound	$-\log IC_{50}$ (nm)	Predicted values		
		MLR	PLS	ANN
1	−1.14613	−0.6383	−0.66615	−0.93142
3	−1.23045	−1.09845	−1.09018	−1.34475
5	−1.32222	−1.02597	−1.0994	−1.25406
7	−1.90849	−1.86057	−1.9321	−2.12476
9	−0.95424	−0.58322	−0.46656	−0.59967
10	−0.69897	−1.02571	−0.95994	−0.90798
11	−0.60206	−0.28247	−0.25226	−0.10726
12	−0.60206	−0.73668	−0.75744	−0.72563
13	−0.90309	−0.67813	−0.71735	−0.65627
15	−0.47712	−0.49303	−0.51961	−0.32639
16	−0.47712	−0.48121	−0.52971	−0.29183
18	−0.77815	−0.81184	−0.87788	−0.83374
19	−1.49136	−1.23715	−1.36273	−1.5707
20	−0.30103	−0.67635	−0.77579	−0.55738
21	−1.04139	−0.81524	−0.88425	−0.88541
22	0	−0.5886	−0.54718	−0.55948
23	0	−0.50691	−0.48592	−0.37994
24	−0.8451	−0.89936	−0.8726	−0.76048
26	0	−0.35123	−0.25523	−0.29696
28	−0.60206	−0.80139	−0.88701	−0.43069
29	−0.8451	−0.54057	−0.61941	−0.10918
30	−0.30103	−0.52769	−0.57488	−0.19837
31	−3.25527	−3.04826	−3.10186	−3.08918
32	−2.8451	−2.40113	−2.56012	−2.43824
34	−2.66276	−2.66089	−2.5583	−2.83962
35	−3	−2.81136	−2.67794	−3.10171
36	−3.30103	−4.00967	−4.17952	−3.75619
37	−2.77815	−2.43522	−2.33602	−2.6481
38	−2.60206	−2.79888	−2.8047	−2.78733
40	−2.43136	−2.4773	−2.35853	−2.75302
41	−2.47712	−2.55874	−2.44048	−2.83191
42	−2.8451	−2.92203	−2.93966	−2.92192
43	−3.79934	−3.31448	−3.34463	−3.37084
45	−2.47712	−2.99681	−3.01148	−3.00529
46	−3.41497	−3.48361	−3.51152	−3.58599
48	−2.60206	−2.86457	−2.88905	−2.74429
50	−2.94939	−2.84571	−2.86909	−2.7322
51	−3.32222	−3.37338	−3.35582	−3.25522
53	−3.50515	−3.49821	−3.56918	−3.36268
56	−2.72428	−2.73036	−2.63731	−2.71428
57	−3.25527	−3.62323	−3.68245	−3.52836
59	−2.39794	−2.70808	−2.60183	−2.70823
61	−4.63949	−3.30429	−3.36551	−3.2928
62	−2.91908	−2.7656	−2.77818	−2.7589
65	−2.89763	−3.09412	−3.00602	−3.16571
66	−4.3784	−4.09285	−4.0554	−4.25882
67	−3.04139	−2.95589	−2.96411	−2.94121
69	−2.85126	−2.23194	−2.10801	−2.49101
70	−3.14613	−3.08234	−3.06178	−3.13868
71	−2.61278	−2.53291	−2.44241	−2.65939
72	−1.68124	−2.3376	−2.26953	−2.51079
73	−2.69897	−3.03147	−3.03564	−3.04202
74	−3.53148	−3.09095	−3.03329	−3.00224
75	−2.90309	−2.93155	−2.9423	−3.14016
76	−2.34242	−3.14267	−3.19094	−3.10916

MLR: Multiple linear regression; PLS: Partial least square; ANN: Artificial neural network

predictive power of the developed model. These plots were also utilized to search for the presence of any outliers in the model. However, no such compound, deviating from the idealness and not fitting in the developed model, was found to exist in the selected series of molecules.[28]

The regression equation, obtained through the application of MLR technique on the training set compounds used to build the model, is as follows:

Y = 1.4216795-X1-0.24535894*X2-0.18161716*X3-2.004 7016*X4-0.137721*X5-6.5111089

where X1 = Verloop L (Substituent 1), X2 = Bond Dipole Moment (Subst. 2), X3 = LogP (Subst. 1), X4 = Balaban Topological index (Subst. 1), and X5 = VAMP Total Dipole (whole molecule).

As per the acceptable statistical standard criterion, a minimal value of 0.80 for r^2 is essential for a statistically significant model.[29,30] Moreover, the developed model, with excellent r^2 value of 0.97, yet again proved its statistical significance [Table 3].

In addition to MLR, robustness of the developed QSAR model was evaluated by PLS approach and the equation so obtained is as follows:

Y = 1.276677*X1-0.26492366*X2-0.23346859*X3-2.4868 717*X4-0.11788595*X5-4.7390904

The experimental versus predicted activity plot is shown in Graphs 1 and 2, respectively. Intriguingly, regression values (R^2) of both the plots were close enough that further validated the reliability of the developed model.[31]

Furthermore, not all relationships can be linear. Hence, to deal with nonlinear data, as well as to enrich our findings, the ANN approach was used. The neural trains were used to predict the activity data [Figures 3-7]. The plot between the predicted and the original activity data is represented in Graph 3.

Even for the graph obtained through the neural approach, R^2 values were found to be very close to those obtained from MLR and PLS methods. This further validated the reliability of the developed model.

The descriptors included in the final model were found to be of relevance, as reflected by their coefficient,

Table 2: Representing experimental activity data of the test set compounds and the predicted activity values, obtained through multiple linear regression, partial least square, and artificial neural network methods

Name of the compound	$-\log IC_{50}$ (nm)	Predicted values		
		MLR	PLS	ANN
2	−1.20412	−0.82289	−0.87781	−0.91669
4	−1.14613	−0.79666	−0.82315	−0.87549
6	−2.25042	−1.24876	−1.32691	−1.60956
8	−1.39794	−1.0697	−0.97739	−1.49466
14	−0.90309	−0.35844	−0.28025	−0.44879
17	−0.30103	−0.95633	−0.99739	−1.20789
25	−1.11394	−0.37936	−0.38896	−0.04864
27	−1	−0.01137	0.039817	−0.00149
33	−3.81954	−2.69198	−2.63709	−3.009
39	−3.8451	−2.84804	−2.74428	−3.15235
47	−3.36173	−1.59088	−1.80919	−0.55715
49	−3.32222	−2.42784	−2.52754	−2.37368
52	−3.80618	−2.75367	−2.70432	−2.70847
54	−4.42975	−2.81469	−2.86001	−2.67359
55	−3.50515	−3.02128	−2.96926	−2.92813
58	−4.38202	−3.11803	−3.04371	−3.04314
60	−4.48144	−3.12905	−3.18888	−2.94669
64	−3.53148	−2.52193	−2.47864	−2.61368
68	−3.23045	−2.78865	−2.74748	−2.83919
77	−4.36922	−3.64787	−3.58308	−3.81895

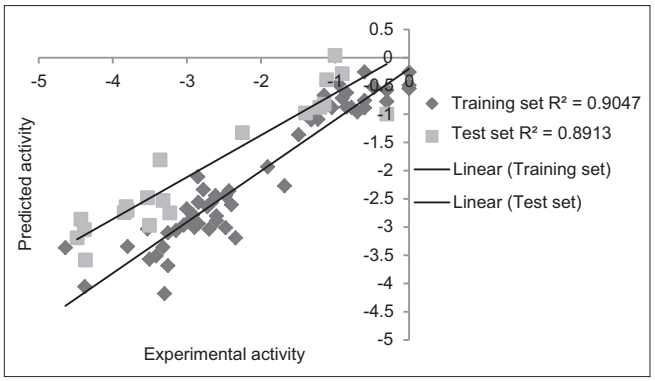
MLR: Multiple linear regression; PLS: Partial least square; ANN: Artificial neural network

Table 3: Representing the values of standard parameters employed to evaluate predictability of the developed quantitative structure-activity relationship model

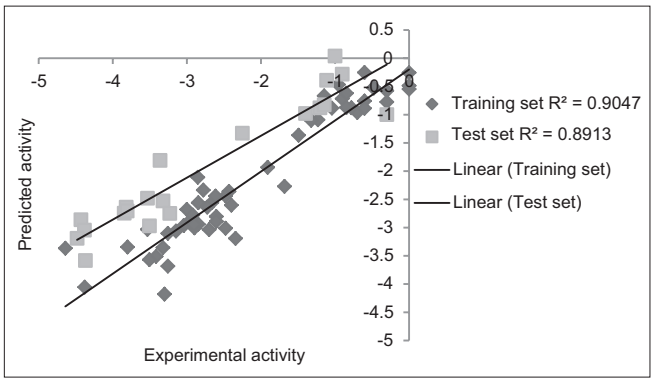
r	r^2	r^2CV	*S*	*F*
0.95	0.90	0.88	0.38	97.22

jackknife, *t*-test, as well as covariance standard error values [Table 4]. Their significance was further estimated through observing the correlation of an individual descriptor with the bioactivity [Table 5]. Among the retrieved descriptors, Verloop L (Subst. 2) and Bond Dipole Moment (Subst. 2) manifested high correlation values of 0.82 and −0.76, respectively. LogP (Subst. 1) displayed moderate correlation of −0.4 with the JNK1 inhibitory activity. However, Balaban Topological index (Subst. 1), as well as VAMP Total Dipole (whole molecule), exhibited low correlation with the biological activity, but their removal from the final model digressed the developed model from idealness (s-value changed from 0.38 to 0.79) which was considered as unacceptable and therefore, these descriptors were retained in the final model.

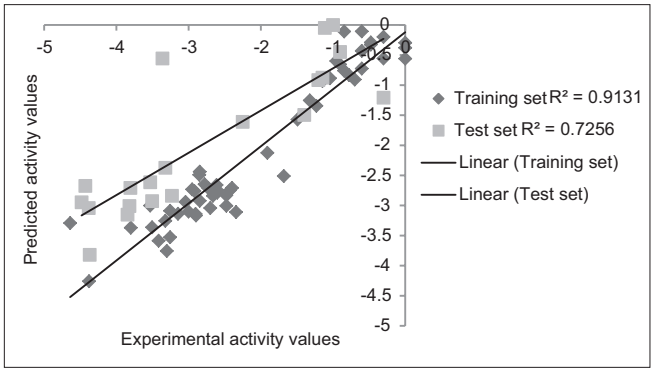
Verloop parameters,[32-34] in general, constitute a group of multidimensional steric descriptors. The VerloopL parameters define the influence of the length of the substituted groups in a molecule under study. As this explanatory variable (VerloopL) is positively correlated



Graph 1: Plot between predicted and experimental biological activity values obtained through multiple linear regression approach



Graph 2: Plot between predicted and experimental biological activity values obtained through partial least square approach



Graph 3: Plot between predicted and experimental biological activity values obtained through artificial neural network

with the JNK1 inhibitory activity, increase in the length of the substituent on R2 position will probably have a positive impact on the inhibitory activity. In general, steric hindrance, owing to the presence of bulky groups, is considered as unfavorable for a molecule to approach and bind with the target receptor. However, at the same time, for a molecule to adequately fit into the binding domain and properly align with the binding site sequence, an optimal amount of bulk, as well as branching, helps it to orient in a better way and thus, augments the likelihood of adequate bonding interactions, between a molecule and a receptor.

Table 4: Representing the values of various parameters employed to determine the relevance of the descriptors in the developed model

Descriptors	Coefficient[a]	Jackknife[b]	Covariance SE[c]	*t*[d]	*t*, *P*[e]
Verloop L (Substituent 2)	1.421	0.17755	0.17681	8.0409	1.6547e, 010
Bond dipole moment (Subst. 2)	−0.24536	0.051168	0.046011	−5.3326	2.4495e, 006
LogP (Subst. 1)	−0.18162	0.063219	0.72628	−2.5007	0.01579
Balaban Topological Index (Subst. 1)	−2.0047	0.663	0.45148	−4.4403	5.1103e, 005
VAMP total dipole (whole molecule)	−0.13772	0.03791	0.03885	−3.545	0.00087483

[a]The regression coefficient for individual variable in the QSAR equations. [b]An estimate of the standard error on individual regression coefficient obtained from a jackknife method on the final regression model, [c]An estimate of the standard error on individual regression coefficient derived from covariance matrix, [d]Measure of the significance of each variable included in the final model, [e]Statistical significance for *t*-test values. SE: Standard error; QSAR: Quantitative structure-activity relationship

Table 5: Correlation matrix describing correlation between the descriptive variables and the experimental biological activity

Variables	$-\log IC_{50}$	Verloop L (Substituent 2)	Bond dipole moment (Subst. 2)	LogP (Subst. 1)	Balaban Topological Index (Subst. 1)	VAMP total dipole (whole molecule)
$-\log IC_{50}$	1	0.82988	−0.76088	−0.40176	−0.14849	0.090375
Verloop L (Subst. 2)	0.82988	1	−0.69038	−0.36035	−0.021717	0.21511
Bond dipole moment (Subst. 2)	−0.76088	−0.69038	1	0.22672	−0.027491	−0.40024
LogP (Subst. 1)	−0.40176	−0.36035	0.22672	1	−0.15066	0.0036989
Balaban Topological Index (Subst. 1)	−0.14849	−0.021717	−0.027491	−0.15066	1	0.12301
VAMP total dipole (whole molecule)	0.090375	0.21511	−0.40024	0.0036989	0.12301	1

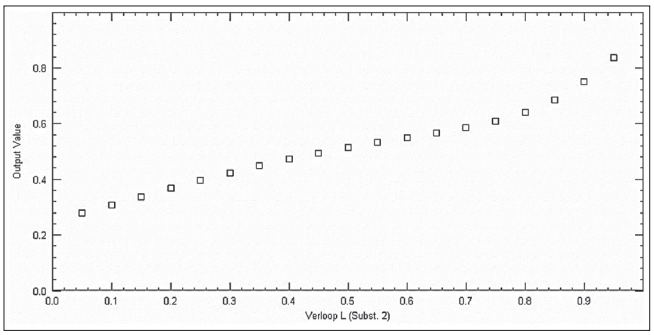


Figure 3: Neural plot of Verloop L (Substituent 2) parameter with the biological activity. As this descriptive variable is positively correlated with the biological activity, the output value is increasing with the increase in the value of this descriptive variable

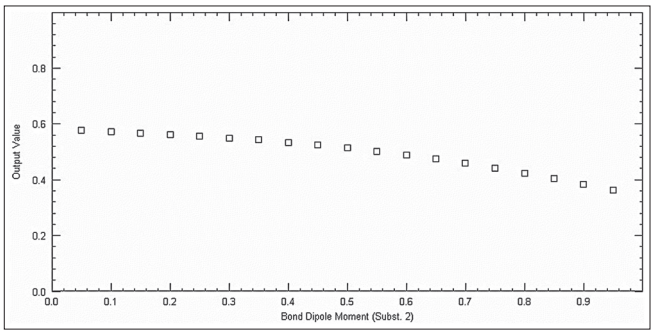


Figure 4: Neural plot of Bond Dipole Moment (Substituent 2) parameter with biological activity. As this descriptive variable is negatively correlated with the biological activity, the output value is decreasing with the increase in the value of this descriptive variable

Parameter logP is known to play a primary role in biochemical cascades and impacts the ADME properties of a drug and thereby, its bioavailability showed negative correlation, at substituent R1, with the biological activity. Thus, a substituent that reduces the hydrophobicity will result in better fit of the molecule.

Bond Dipole Moment (Subst. 2) utilizes the concept of electric dipole moment to compute the polarity of a chemical bond in any molecule.[35] Furthermore, as this parameter is found to be negatively correlated with the inhibitory activity, improvement in JNK1 inhibitory activity is expected by decreasing the polar nature of substituent at this position.

Balaban Topological Index, also known as index "J," typically describes the connectivity among atoms where atoms are considered as vertices and bonds between them as edges.[36] As this variable is negatively correlating with the biological activity at position R1, replacement with less steric substituent may result in augmented inhibitory activity.

Finally, total dipole parameter, calculated through semi-empirical package "VAMP" in TSAR 3.3, was found to negatively correlate with the biological activity. Majorly, this parameter explains the molecular charge distribution, in three dimensions of the molecule. Hence, it can be deduced that, by decreasing the overall electronegativity of the molecule, inhibitory activity can be improved.

Due to minor differences in the structural skeleton of the molecules under study, the significance of these descriptive variables can be easily understood by observing their computed values, for the active and the inactive molecules. For an instance,

the value of Verloop L for active compounds was found to be high and, on the other hand, for inactive compounds, the value of this descriptor was found to be low [Table 6].

However, when the values of the negatively correlating descriptors were observed for the individual molecules of the series, an exactly opposite pattern was noticed. As the rest of the physicochemical parameters were exhibiting a negative correlation with the inhibitory activity, their value for active compounds was low, as compared to the inactive ones [Tables 7-10].

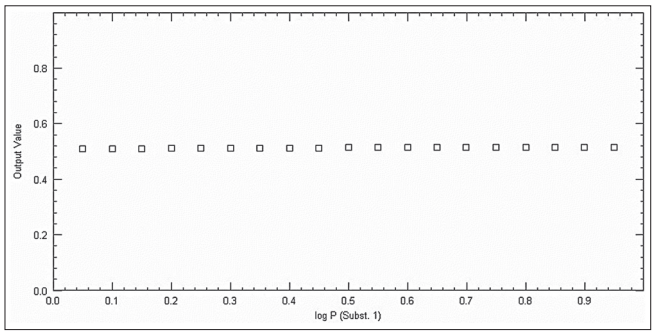


Figure 5: Neural plot of logP (Substituent 1) parameter with biological activity. As this descriptive variable is positively correlated with the biological activity, the output value is somewhat increasing with the increase in the value of this descriptive variable. However, as its correlation with the bioactivity is not that strong, the increment in the output value is not significant

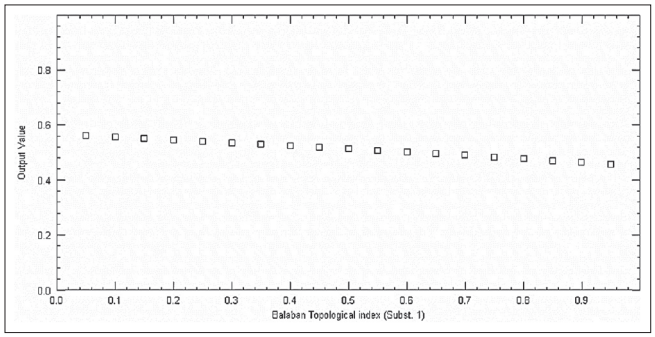


Figure 6: Neural plot of Balaban Topological Index (Substituent 1) parameter with biological activity. As this descriptive variable is negatively correlated with the biological activity, the output value is decreasing with the increase in the value of this descriptive variable

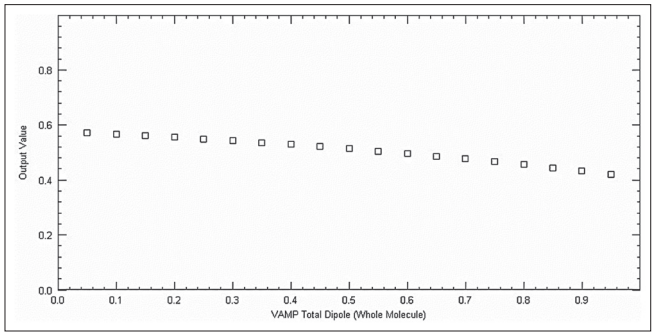


Figure 7: Neural plot of VAMP Total Dipole (whole molecule) with biological activity. As this descriptive variable is negatively correlated with the biological activity, the output value is decreasing with the increase in the value of this descriptive variable

Conclusion

A validated two-dimensional QSAR model was constituted by five major descriptors, namely Verloop L (Subst. 1), Bond Dipole Moment (Subst. 2), LogP (Subst. 1), Balaban Topological index (Subst. 1), and VAMP Total Dipole (whole molecule). An in-depth assessment of the obtained physicochemical descriptors provided explicit knowledge about the dependence of the biological activity on the

Table 6: Correlation of biological activity of active and inactive molecules with Verloop L descriptor

	Name of compound	Biological activity IC_{50} (nM)	Verloop L (R^2)
Active compounds	22	1	6.76
	23	1	6.76
	26	1	6.73
Inactive compounds	58	24,100	5.75
	60	31,500	5.76
	61	43,600	5.74

Table 7: Correlation of biological activity of active and inactive molecules with logP descriptor

	Name of compound	Biological activity IC_{50} (nM)	LogP (R1)
Active compounds	22	1	0.94
	23	1	0.94
	26	1	−0.058
Inactive compounds	58	24,100	2.80
	60	31,500	2.80
	61	43,600	1.54

Table 8: Correlation of biological activity of active and inactive molecules with bond dipole moment descriptor

	Name of compound	Biological activity IC_{50} (nM)	Bond Dipole moment (R2)
Active compounds	22	1	−2.04
	23	1	−1.95
	26	1	−2.30
Inactive compounds	58	24,100	0.039
	60	31,500	−0.96
	61	43,600	2.31

Table 9: Correlation of biological activity of active and inactive molecules with Balaban Topological Index descriptor

	Name of compound	Biological activity IC_{50} (nM)	Balaban Topological Index (R1)
Active compounds	22	1	1.713
	23	1	1.618
	26	1	1.715
Inactive compounds	58	24,100	1.717
	60	31,500	1.717
	61	43,600	1.903

Table 10: Correlation of biological activity of active and inactive molecules with VAMP total dipole descriptors

	Name of compound	Biological activity IC_{50} (nM)	VAMP total dipole (whole molecule)
Active compounds	22	1	4.31
	23	1	3.47
	26	1	4.02
Inactive compounds	58	24,100	2.18
	60	31,500	2.71
	61	43,600	1.51

molecular structure. In addition to this, the values of statistical parameters, such as *r*, r^2, r^2cv, *s*-value, and *F*-value, proved the statistical soundness of the developed model. Therefore, the valuable information retrieved from this model can be applied to alter the substituents in the selected molecules and thereby, optimized JNK1 inhibitors, in terms of potency as well as selectivity, can be successfully designed.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Hotamisligil GS, Shargill NS, Spiegelman BM. Adipose expression of tumor necrosis factor-alpha: Direct role in obesity-linked insulin resistance. Science 1993;259:87-91.
2. Uysal KT, Wiesbrock SM, Marino MW, Hotamisligil GS. Protection from obesity-induced insulin resistance in mice lacking TNF-alpha function. Nature 1997;389:610-4.
3. Ventre J, Doebber T, Wu M, MacNaul K, Stevens K, Pasparakis M, *et al.* Targeted disruption of the tumor necrosis factor-alpha gene: Metabolic consequences in obese and nonobese mice. Diabetes 1997;46:1526-31.
4. 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.
5. Houstis N, Rosen ED, Lander ES. Reactive oxygen species have a causal role in multiple forms of insulin resistance. Nature 2006;440:944-8.
6. Withers DJ, Burks DJ, Towery HH, Altamuro SL, Flint CL, White MF, *et al.* Irs-2 coordinates Igf-1 receptor-mediated beta-cell development and peripheral insulin signalling. Nat Genet 1999;23:32-40.
7. Dresner A, Laurent D, Marcucci M, Griffin ME, Dufour S, Cline GW, *et al.* Effects of free fatty acids on glucose transport and IRS-1-associated phosphatidylinositol 3-kinase activity. J Clin Invest 1999;103:253-9.
8. Dérijard 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.
9. Smeal T, Binetruy B, Mercola D, Grover-Bardwick A, Heidecker G, Rapp UR, *et al.* Oncoprotein-mediated signalling cascade stimulates c-Jun activity by phosphorylation of serines 63 and 73. Mol Cell Biol 1992;12:3507-13.
10. Smeal T, Binetruy B, Mercola DA, Birrer M, Karin M. Oncogenic and transcriptional cooperation with ha-ras requires phosphorylation of c-Jun on serines 63 and 73. Nature 1991;354:494-6.
11. Kyriakis JM, Banerjee P, Nikolakaki E, Dai T, Rubie EA, Ahmad MF, *et al.* The stress-activated protein kinase subfamily of c-Jun kinases. Nature 1994;369:156-60.
12. Behrens A, Sibilia M, Wagner EF. Amino-terminal phosphorylation of c-Jun regulates stress-induced apoptosis and cellular proliferation. Nat Genet 1999;21:326-9.
13. Pulverer BJ, Kyriakis JM, Avruch J, Nikolakaki E, Woodgett JR. Phosphorylation of c-Jun mediated by MAP kinases. Nature 1991;353:670-4.
14. 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.
15. Hirosumi J, Tuncman G, Chang L, Görgün CZ, Uysal KT, Maeda K, *et al.* A central role for JNK in obesity and insulin resistance. Nature 2002;420:333-6.
16. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. Nature 2006;444:840-6.
17. Cho H, Black SC, Looper D, Shi M, Kelly-Sullivan D, Timofeevski S, *et al.* Pharmacological characterization of a small molecule inhibitor of c-Jun kinase. Am J Physiol Endocrinol Metab 2008;295:E1142-51.
18. Hom RK, Bowers S, Sealy JM, Truong AP, Probst GD, Neitzel ML, *et al.* Design and synthesis of disubstituted thiophene and thiazole based inhibitors of JNK. Bioorg Med Chem Lett 2010;20:7303-7.
19. Bowers S, Truong AP, Jeffrey Neitz R, Hom RK, Sealy JM, Probst GD, *et al.* Design and synthesis of brain penetrant selective JNK inhibitors with improved pharmacokinetic properties for the prevention of neurodegeneration. Bioorg Med Chem Lett 2011;21:5521-7.
20. ChemDraw Ultra Version 8.0. Cambridge: Cambridge Scientific Computing; 2004.
21. TSAR Version 3.3. Oxford: Oxford Molecular, Ltd.; 2000.
22. Dalby A, Nourse JG, Hounshell WD, Gushurst AK, Grier DL, Leland BA, *et al*. Description of several chemical structure file formats used by computer programs developed at molecular design limited. J Chem Inf Comput Sci 1992;32:244-55.
23. Wylie WA, Vinter JG. Molecular modelling and drug design. In: Gardner M, editor. Topics in Molecular Biology. London: Macmillan; 1994.
24. Roy K, Mandal AS. Development of linear and nonlinear predictive QSAR models and their external validation using molecular similarity principle for anti-HIV indolyl aryl sulfones. J Enzyme Inhib Med Chem 2008;23:980-95.
25. Saghaie L, Sakhi H, Sabzyan H, Shahlaei M, Shamshirian D. Stepwise MLR and PCR QSAR study of the pharmaceutical activities of antimalarial 3-hydroxypyridinone agents using B3LYP/6-311++G** descriptors. Med Chem Res 2013;22:1679-88.
26. Kubinyi H. QSAR and 3D-QSAR in drug design part 1: Methodology. Drug Discov Today 1997;2:457-67.
27. Nagpal A, Paliwal SK. QSAR model development using MLR, PLS and NN approach to elucidate the physicochemical properties responsible for neurologically important JNK3 inhibitory activity. J Pharm Sci Res 2017;9:1831-43.

28. Kim D, Hong S, Lee D. The quantitative structure-mutagenicity relationship of polycyclic aromatic hydrocarbon metabolites. Int J Mol Sci 2006;7:556-70.
29. Golbraikh A, Tropsha A. Beware of q2! J Mol Graph Model 2002;20:269-76.
30. Tropsha A, Gramatica P, Gombar VK. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. QSAR Comb Sci 2003;23:69-76.
31. Nagpal A, Paliwal S. Discovery of novel and selective c-Jun NH2-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. Asian J Pharm Clin Res 2018;11:100-8.
32. Verloop A, Tipker J. Use of linear free energy related and other parameters in the study of fungicidal selectivity. Pestic Sci 1976;7:379-90.
33. Verloop A, Tipker JA. Comparative study of new parameters in drug design. In: Keverling BJ, editor. Biological Activity and Chemical Structure. Vol. 65. Amsterdam: Elsevier; 1977.
34. Verloop A, Tipker HW. Development and application of new steric substituent parameters in drug design. In: Ariens EJ, editor. Drug Design. Vol. 7. New York: Academic; 1976.
35. Karelson M. Molecular Descriptors in QSAR/QSPR. New York: Wiley Interscience; 2000.
36. Balaban AT. Highly discriminating distance-based topological index. Chem Phys Lett 1982;89:399-4.