WORLD JOURNAL OF PHARMACY AND PHARMACEUTICAL SCIENCES
SJIF Impact Factor 7.421
Volume 8, Issue 2, 662-678 Research Article ISSN 2278 – 4357

2D- AND 3D-QSAR STUDIES ON THIAZOLIDINE-2,4-DIONE DERIVATIVES

Manoj Upadhyay, Md. Asif, Kohit Sahu, Vivek Asati, Jagdish Chandra Rathi*

NRI Institute of Pharmaceutical Sciences, Bhopal.

Article Received on
09 Dec. 2018,

Revised on 29 Dec. 2018,
Accepted on 19 Jan. 2019

DOI: 10.20959/wjpps20192-13043

***Corresponding Author**
Jagdish Chandra Rathi
NRI Institute of Pharmaceutical Sciences, Bhopal.

ABSTRACT

A series of 21 thiazolidine-2,4-dione derivatives were used for quantitative structure–activity relationship (QSAR) studies. These compounds were introduced into two-dimensional (2D-QSAR), and three-dimensional (3D-QSAR) studies to find the structural requirements for PIM-2 kinase inhibitory activity. The stepwise forward-backward (SW-FB) variable selection method has been applied for the generation of 2D-QSAR. The best 2D-QSAR model generated by partial least squares regression method having r^2 = 0.78; q^2 = 0.63 with pred_r^2 = 0.78. The k-nearest neighbor (kNN) with genetic algorithm(GA) method has been applied for the generation of The 3D-QSAR models having q^2 = 0.64 and pred_ r^2 = 0.94 values. The docking study showed the binding orientations of these inhibitors at active site amino acid residues (ASP 124, ASP 182, PHE 43 and GLU 83) of PIM-2 enzyme (PDB ID: 3IWI). The results showed the important sites of thiazolidine-2,4-dione analoges for the generation of potent compounds against PIM-2 kinase.

KEYWORDS: G-QSAR, kNN-MFA, Thiazolidine-2,4-dione, 2D/3D QSAR.

1. INTRODUCTION

PIM kinases regulate the cell signaling pathways via Janus kinase (JAK)/signal transducer and activator of transcription (STAT) pathway including proliferation, migration and metabolism. The PIM kinases generally work as weak oncogenes when expressed as transgenes. The oncogenic potential enhanced on co-expression with c-Myc, a transcription factor plays an important role in cell growth and differentiation [Forshell et al., 2011; Möröy et al., 1991; Van Lohuizen et al., 1989; Zhang et al., 2008].

PIM-1 and PIM-2 kinases are over--expressed in varieties of cancer such as multiple myeloma, lymphomas, leukemia and prostate cancer. PIM-3 over-expression is linked with solid tumors in pancreatic, prostate, colon and other organales [Brault et al., 2010; Nawijn et al., 2011]. Consequently, the PIM kinases may be considered as a potential target for cancer therapy. Furthermore, no severe side effects have been observed after inhibiting all these kinases in an experiment on mice [Mikkers *et al.*, 2004]. Novel substituted benzylidene-1,3-thiazolidine-2,4-diones (TZDs) have been identified as potent and highly selective inhibitors of the PIM kinases [Dakin *et al.*, 2012].

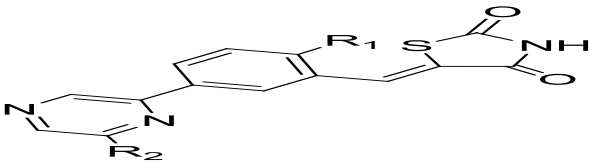
Heterocycles play an important role in cancer therapy paticularly five-membered ring heterocycles which contain three carbon atoms, one nitrogen atom, and one sulfur atom, known as thiazoles are of considerable interest in different areas of medicinal chemistry [Asati *et al.*, 2014]. Thiazolidine-2,4-dione (TZD), one of the most important heterocyclic systems has therapeutic importance and when combined with other heterocyclic rings it may produce better anticancer activity.

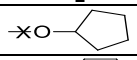
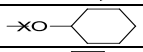
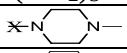
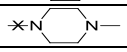
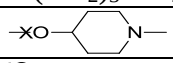
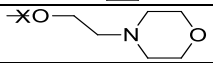
In literature several thiazolidine-2,4-diones have been synthesized and evaluated for their anti-cancer activity. In the present study, QSAR analysis was performed for 21 previously synthesized 5-benzylidenethiazolidine-2,4-dione analogues [Lee *et al.*, 2014] to establishing quantitative relationship between biological activity of derivatives and their physicochemical/structural properties. The aim of the present work is to generate best predictive and validated 2D-, 3D- and G-QSAR models which may help to medicinal chemist for designing and development of novel thiazolidine-2,4-dione derivatives. In this work widely used technique viz. stepwise forward-backward (SW-FB) with partial least square (PLS) analysis has been applied for the development of 2D- and G-QSAR models as variable selection method. k-nearest neighbor (kNN) analysis with Genetic algorithm (GA) has been applied for the development of 3D-QSAR model. The generated models may provide insights into the influence of various interactive fields on the activity and thus, can help in designing and forecasting the inhibitory activity of novel anticancer agents.

2. MATERIALS AND METHODS

2.1. Dataset: The QSAR studies (2D- and 3D-QSAR) were performed by using the VLife Molecular Design Suite (MDS 4.4). Data set and their biological activity are discribed in Table **1** where R_1 and R_2 are various substituents. Biological data presented as IC_{50} (μM) were converted into log ($1/IC_{50}$) for computational work.

Table. 1. Structure and anticancer activity (IC_{50}) of thiazolidine-2,4-dione derivatives.



S No.	R_1	R_2	pIC_{50} (µM)
1	OH		8.03
2	OH		7.84
3	H		7.01
4	OH	–O$(CH_2)_2NMe_2$	8.66
5	H	–O$(CH_2)_2NMe_2$	7.49
6	OH	–O$(CH_2)_3$NMe2	8.77
7	H	–O(CH2)$_3NMe_2$	7.77
8	OH	–NH$(CH_2)_2NMe_2$	8.74
9	H	–NH$(CH_2)_2NMe_2$	8.24
10	OH	–NH(CH2)$_3NMe_2$	8.68
11	H	–NH$(CH_2)_3NMe_2$	7.98
12	OH		7.80
13	H		7.88
14	OH	–O$(CH_2)_2NEt_2$	8.66
15	OH	–O$(CH_2)_3NEt_2$	8.68
16	OH		8.55
17	OH		8.05
18	OH	–NH$(CH_2)_2NEt_2$	8.55
19	OH	–NH$(CH_2)_3NEt_2$	8.79
20	OH	–N(Me)$(CH_2)_2NMe_2$	8.05
21	OH	–N(Me)$(CH_2)_2NEt_2$	7.49

X = Position of attachment

2.2. 2D-QSAR: VLife MDS 4.4 software was used for the building of molecular structures of all the 21 compounds. All 2D structures were converted into 3D by using Chem Bio Draw and saved in .mol2 format. All the compounds were batch optimized using standard Merck Molecular Force Field (MMFF) and Gasteiger-marsili charge followed by taking into account distance-dependent dielectric constant at 1.0, convergence criterion or root-mean-square (RMS) gradient at 0.01 kcal/mol Å and the iteration limit to 10,000 [Halgren, 1996; Ghosh and Bagchi, 2009; Sahu *et al.*, 2011]. The most stable structure for each compound was generated and used for the calculation of various 2D descriptors like physicochemical and Baumann alignment-independent topological descriptors [Baumann, 2002]. The energy-minimized geometry was used for the calculation of the various 2D descriptors including Individual, Chi, ChiV, Path count, ChiChain, ChiVChain, Chain path count, Cluster, Path

cluster, Kappa, Element Count, Estate number, Estate contribution, Semi-empirical and Polar surface area etc. Alignment-independent descriptors used ‘‘T-attribute’’ which characterize the topology of the molecule. In this study to calculate alignment independent descriptors, we have used following attributes, 2 (double bonded atom), 3 (triple bonded atom), T(any), C, N, O, S, H, F, Cl, Br and I, the distance range of 0–7 was considered as independent variables. Preprocessing of the independent variables (i.e. descriptors) was done by removing the invariables (descriptors that are constant for all the molecules), which resulted in 212 descriptors in the descriptor pool.

In order to evaluate the QSAR model externally, data sets were divided into training and test set using Random selection method, Manual data selection method and Sphere Exclusion methods. Training sets are used as known biological activity data in QSAR model development. Test set is used to evaluate the QSAR model which has been developed by training set. Data set selected randomly given best 2D-QSAR results included five compounds, namely, 5, 7, 9, 14 and 16 were used as test set for model 1 while remaining 16 molecules were used as the training set (Table 1). In order to assess the similarity of the distribution pattern of the molecules in the generated sets, statistical parameters (with respect to the biological activity), i.e. mean, maximum, minimum and standard deviation were calculated for the training and test sets (Table **2**). For the prediction statistics to be reliable, the test set must include at least four compounds [Golbraikh and Tropsha, 2002].

Table. 2. Unicolumn statistics of the training and test sets for thiazolidine-2,4-dione derivatives in different models.

Data set	Average	Max.	Min.	SD	Sum
2D-QSAR MODE- 1					
Training	8.14	8.90	7.06	0.54	120.32
Test	8.57	8.81	8.47	0.08	40.26
3D-QSAR MODE- 3					
Training	8.15	8.78	7.02	0.66	128.21
Test	8.37	8.71	7.78	0.43	40.59

2.3. 3D-QSAR: For 3D-QSAR molecular modeling has been done by a process including allignment of molecules, descriptor genearation with the selection of training and test set and development of model.

2.3.1. Allignment of molecules: The 3D-QSAR studies were performed using V Life Molecular Design Suite software. In this study molecules of the dataset are aligned by

template-based method [Ajmani *et al.*, 2006]. In this method a template structure is defined and used as a basis for alignment of a set of molecules. The template structure, i.e. 5-(3-pyrazin-2-ylbenzylidene)-1,3-thiazolidine-2,4-dione is used for alignment by considering common elements of the series and is shown in Fig. **1**. The reference molecule, compound 19 has been chosen in such a way that it is the most active among the series of molecules. Reference molecule is chosen on which the other molecules of the dataset get aligned considering the chosen template. After optimizing, the template structure and reference molecule were used to superimpose all molecules. The superimposition of all molecules based on minimizing root mean square deviation (RMSD), shown in Fig. **2**.

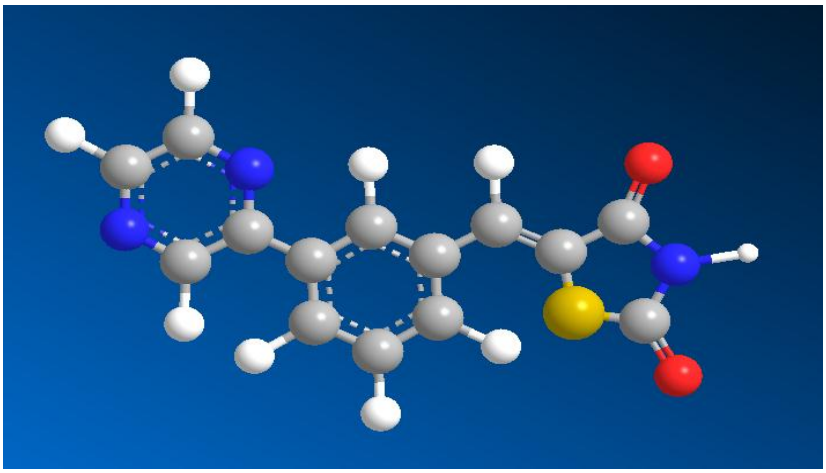


Fig. 1: 5-(3-pyrazin-2-ylbenzylidene)-1,3-thiazolidine-2,4-dione ring as a template.

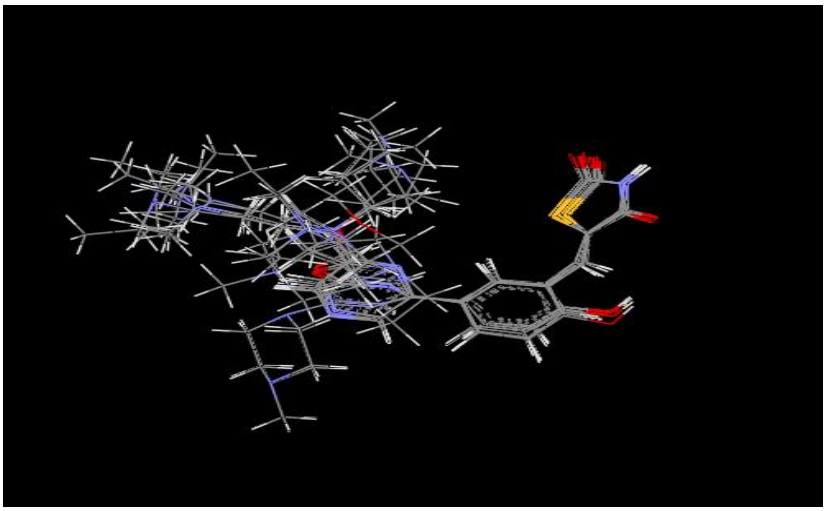


Fig. 2: Template based alignment of all the 21 thiazolidine-2,4-dione derivatives.

2.3.2. Generation of field descriptors: The results of molecular field analysis (MFA) may provide predictive and sufficiently reliable information to medicinal chemist for design and

development of novel anticancer agents. This approach is effective for the analysis of data sets, where activity information is available but the structure of the receptor site is unknown. It attempts to postulate and represent the essential features of a receptor site from the aligned common features of the molecules that bind to it. The MFA calculates probe interaction energies on a rectangular grid around a bundle of active molecules. The atomic coordinates of the contributing models were used to compute field values on each point of a 3D grid. Fields of molecules were represented using grids and energy associated with each grid point which can serve as input for the calculation of 3D-QSAR. These energies are added to the study table to form new columns headed according to the probe type. The molecular field is created using methyl group as probe, which represent steric, electrostatic and hydrophobic fields, respectively. For calculation of field descriptor values electrostatic, steric and hydrophobic fields with cutoff values 10.0 and 30.0 kcal/mol, respectively, were selected and charge was selected as Gasteiger and Marsili [Shen *et al.*, 2003]. The dielectric constant was set to 1.0 considering the distance dependent dielectric function. Probe setting was carbon atom with charge 1.0. This resulted in calculation of 4,711 field descriptors (electrostatic, steric and hydrophobic) for all the compounds in separate columns after removing descriptors having zero values or same values. Data set selected manually given best result in which five compounds, namely, 2, 4, 8, 12 and 18 were used as test set while the remaining 16 molecules as the training set. The unicolumn statistics has been given in Table **2**.

2.3.3. Model development: 3D-QSAR model development was performed using k-Nearest Neighbour Molecular Field Analysis (kNN-MFA) methodology which relies on a simple distance learning approach whereby an unknown member is classified according to the majority of its k-nearest neighbours in the training set [Ajmani *et al.*, 2006; Sharaf *et al.*, 1986]. The nearness is measured by an appropriate distance metrics (e.g. a molecular similarity measure calculated using field interactions of molecular structures). Given a suitable distance metric a k-nearest neighbor (kNN) algorithm only requires that a suitable value of k be chosen. In many cases setting k to 1 provides reasonably good predictive performance for classification purposes. In general, optimal values of k are obtained via trial and error. A more systematic approach is to use a cross-validation scheme to obtain the best value of k for a given dataset. The kNN-MFA models were developed by the using of genetic algorithm approach where cross-correlation limit set to 1.0 and the term selection criterion set as r^2. Some parameters like Population and number of generations were set up to 100 and 500 correspondingly. As some additional parameters, variance cutoff was set at 0.0, and scaling to

autoscaling; additionally, kNN parameter setting was done within the range of 2–5 and the prediction method was selected as the distance-based weighted average.

2.4. Model evaluation and validation

This is done to test the internal stability and predictive ability of the QSAR models. Internal validation was carried out using 'leave-one-out' (q^2, LOO) method. The cross validated coefficient, q^2, was calculated using the following Eq. 1:

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

where y_i and $\hat{y}_i$ are the actual and predicted activity of the i^{th} molecule in the training set, respectively, and y_{mean} is the average activity of all molecules in the training set. For external validation, activity of each molecule in the test set was predicted using the model generated from the training set. The pred_r^2 value is calculated by using Eq. 2.

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

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

The robustness of the QSAR models for experimental training sets was examined by comparing these models to those derived for random data sets. Random sets were generated by rearranging biological activities of the training set molecules. The significance of the models hence obtained was derived based on calculated Z-score. Z-score can be defined as the absolute difference between the value of the model and the activity field, divided by the square root of the mean square error of the dataset Eq. 3.

$$Z_{score} = \frac{q^2_{org} - q^2_a}{q^2_{std}} \quad (3)$$

The developed QSAR models were evaluated using the following statistical measures: r^2 (coefficient of determination), q^2 (cross-validated r^2 by LOO), pred_r^2 (r^2 for external test set), F test (Fischer's value for statistical significance), Z_{score} (Z_{score} calculated by q^2 in the randomization test), best_ran_q^2 (highest q^2 value in the randomization test), a (statistical significance parameter obtained by the randomization test), r^2_se (SEE, standard error of

estimate of the model) q^2_se (CV_SE, standard error of cross-validation) and pred_r^2se (pred_SE, standard error of external test set prediction). However, a QSAR model is considered to be predictive, if the following conditions are satisfied: r^2>0.6, q^2>0.6 and pred_r^2> 0.5.

2.5. Molecular docking: Molecular docking study was performed by using AutoDock Vina. AutoDock Vina provided improves average accuracy of the binding mode predictions. AutoDock Vina has been tested against a virtual screening benchmark called the Directory of Useful Decoys by the Watowich group, and was found to be better than other programs. Some important attributes for ligand and macromolecular receptor are essential before doing the docking process through autodock vina as follows [http://vina.scripps.edu, 2014; Trott and Olson, 2010]. For ligand (i) Add all hydrogens, compute Gasteiger charges, and merge non-polar H (ii) Ensure total charge corresponds to tautomeric state (iii) Choose torsion tree root & rotatable bonds. For macromolecule (i) Add all hydrogens, compute Gasteiger charges, and merge non-polar H (ii) Assign Stouten atomic solvation parameters (iii) Optionally, create flexible residues PDBQT in addition to the rigid PDBQT file.

The final evaluation was done with AutoDock Score (docking score). Each docking produced multiple docked conformations of the ligand as well as corresponding binding energy scores which were computed using AutoDock scoring function [Seeliger and Groot, 2010]. The conformations were ranked based on the scores; a lower scoring conformation was ranked higher. Since an experimentally derived conformation of the bound ligand (true conformation) is available, for each docked conformation of the ligand, a RMSD value was also computed. The RMSD value measures the distance between the docked conformation and the true conformation [Chang *et al.*, 2010; Osterberg *et al.*, 2002]. The conformations were also ranked based on the RMSD values, a conformation with lower RMSD value was ranked higher. Autodock generated different file format that loaded in Python Molecular Viewer (PyMOL) and visualized different binding site properties which may provide valuable insights for structure-based drug design. A RMSD value inferior or close to 2 Å was considered as a successful docking [Stigliani *et al.*, 2012].

3. RESULTS AND DISCUSSION

QSAR study of a series of thiazolidine-2,4-dione derivatives was performed by using VLife MDS 4.4 software. In this study biological activity (pIC_{50}) was used as dependent variable and various physiochemical, topological and 3D descriptors have been taken as independent

variable. A data set of 21 compounds were divided into training (16 molecules) and test sets (5 molecules) using random selection method. Selection of molecules 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 test set lie within the maximum and minimum value range of biological activities of training set of compounds. A Uni-Column statistics for training set and test set were generated to check correctness of selection criteria for trainings and test set molecules (Table **2**). After regression analysis the best models were selected, on the basis of r^2, q^2 and pred_r^2 values (Table **3**).

Table. 3. Statistical results of best QSAR models of thiazolidine-2,4-dione derivatives.

S No.	Statistical parameter	2D-QSAR (Model 1)	3D-QSAR (Model 3)
1	r2	0.92	-
2	q2	0.71	0.74
3	pred_r2	0.83	0.91
4	r2_se	0.13	-
5	q2_se	0.18	0.32
6	pred_r2se	0.36	0.13
7	F_test	68.35	-
8	Z_{score}	3.70	-
9	Best-rand_q2	0.50	-
10	α_rand_ q2	0.001	-
12	n_{traing}	16	16

3.1. Interpretation of 2D-QSAR

The 2D-QSAR study of 21 compounds (divided into 5 test and 16 training) for anticancer activity through PLSR analysis coupled with SW-FB variable selection method resulted model summarized in Table **4**. Comparative observed and predicted activities of thiazolidine-2,4-dione derivatives by the best 2D-QSAR model is presented in Table **6**. The inter-correlation matrix between four descriptors with the biological activity for the models 1 is presented in Tables **5**. Contribution plot of descriptor for models 1 is depicted in Fig. **3A**. The graph of actual versus predicted activity for the PLSR analysis is given in Fig. **3B**. The present QSAR model reveals that Baumann's alignment-independent topological descriptors have a major contribution in explaining variation in activity. In general, a descriptor T_X_Y_Z can be defined as a count of fragments formed with atom types X and Y separated by topological distance of Z bonds. The definitions for the descriptors that were found in the developed 2D-QSAR models are given below.

H-DonorCount: Number of hydrogen bond donor atoms.

T_2_C_6: This is the count of number of double bounded atoms (i.e. any double bonded atom, T_2) separated from carbon atom by 6 bonds.

The descriptors obtained in the best 2D model suggested that number of hydrogen bond donor groups (~40%) such as –OH contributed most influential part for the anticancer activity. The positive contribution of this showed that increase in the values would be beneficial for the anticancer activity of thiazolidine-2,4-dione derivatives. Another alignment-independent descriptor T_2_C_6 (negative ~35%) showed that mimimum number of double bounded atoms separated from carbon beneficial for activity. DistTopo signifies distance based topological index. XKMost Hydrophobic Hydrophilic Distance descriptor signifies distance between most hydrophobic and hydrophilic point on the vdW surface.

Table. 4. List of predictive 2D model generated from various regression methods

Model No.	Method	Equation
01	2D/Random/ PLSR/SWFB	IC 50 = + 0.511 H-DonorCount -0.135 T_2_C_6 -0.385 DistTopo -0.084 XKMostHydrophobicHydrophilicDistance + 13.4038

Table. 5. Correlation matrix for descriptors used in different models.

Model 1				
	H-DonorCount	T_2_C_6	Dist topo	XKMostHydrophobicHydrophilicDistance
H-DonorCount	1	-0.04	0.004	-0.199
T_2_C_6	-0.04	1	-0.105	0.274
Dist topo	0.004	-0.105	1	0.001
XKMostHydrophobic HydrophilicDistance	-0.199	0.274	0.001	1
Model 2				
	S_556	E_849	S_1381	
S_556	1.000	0.344	-0.054	
E_849	0.344	1.000	-0.655	
S_1381	-0.054	-0.655	1.000	

Table. 6. Comparative observed and predicted activities of thiazolidine-2,4-dione derivatives by best QSAR models.

Compounds	Exp. pIC_{50} (M)	2D-QSAR	3D-QSAR
		PLSR	kNN
1_opt.mol2	8.03	8.13	8.11
2_opt.mol2	7.84	7.82	7.77
3_opt.mol2	7.01	7.15	7.66
4_opt.mol2	8.66	8.43	8.72
5_opt.mol2	7.49	7.76	7.23
6_opt.mol2	8.77	8.58	8.67
7_opt.mol2	7.77	7.91	8.11
8_opt.mol2	8.74	8.79	8.67
9_opt.mol2	8.24	8.12	7.9
10_opt.mol2	8.68	8.79	8.43
11_opt.mol2	7.98	8.12	8.07
12_opt.mol2	7.80	7.97	7.92
13_opt.mol2	7.88	7.30	8
14_opt.mol2	8.66	8.28	8.72
15_opt.mol2	8.68	8.58	8.72
16_opt.mol2	8.55	8.05	8.74
17_opt.mol2	8.05	8.35	7.77
18_opt.mol2	8.55	8.49	8.74
19_opt.mol2	8.79	8.79	8.37
20_opt.mol2	8.05	7.97	7.76
21_opt.mol2	7.49	7.67	8.05

M = Molar concentration

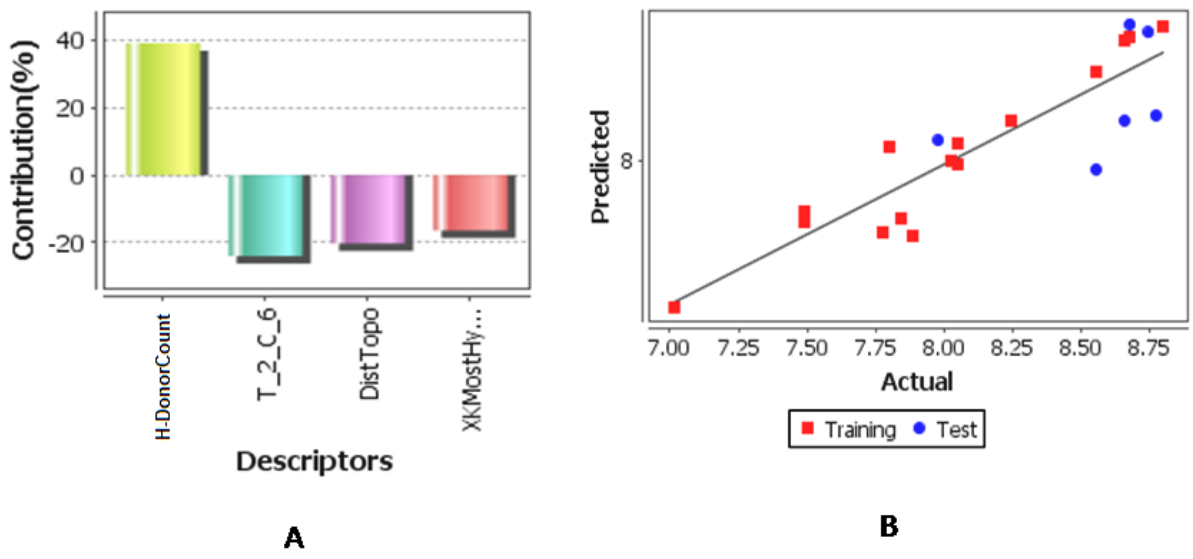


Fig. 3 A Contribution plot between selected descriptors for model 1 B. Fitness plot between actual and predicted activities for model 1.

3.2. Interpretation of 3D-QSAR

The 3D-QSAR studies were performed by using of k-nearest neighbor (kNN) molecular field analysis approach with GA variable selection method. The models were selected on the basis of statistical parameters and the values of the best model 3 were 0.64 for q^2 and 0.93 for pred_r^2 (Table **3**). The careful analysis of comparison of biological activities (pIC_{50}) verses predicted activities for training and test set molecules presented in Table **6**. Therefore it may be said that the predictive abilities of the model 2 is good. The results are supported by the statistical values presented in Table **3**. The graphs of actual verses predicted activity for the series are plotted in Fig. **4** for model 2 which showed good correlation coefficient. Inter-correlation matrix between descriptors with the biological activity for the models 2 is presented in Table **5**.

In 3D-QSAR studies, 3D data points generated around thiazolidine-2,4-dione pharmacophore in model 2 are S_556 (-0.0929 -0.0812), S_1381 (-0.3207 -0.3034) and E_849 (0.2828 0.3184) that is, steric and electrostatic interaction field at lattice points 556, 1381 and 849 respectively (Figs. **5**, **6**). This helps in identification of various molecular features responsible for activity variation and hence aid in design of novel anticancer agents. The steric descriptor like S_556 and S_1381 with negative coefficients showed a region where bulky substituents are disfavored for activity. Electrostatic field descriptor like E_849 with positive coefficient represents regions where electropositive (electron-withdrawing) groups are favorable for activity. The above model is validated by predicting the biological activities of the test molecules, as indicated in Table **6**.

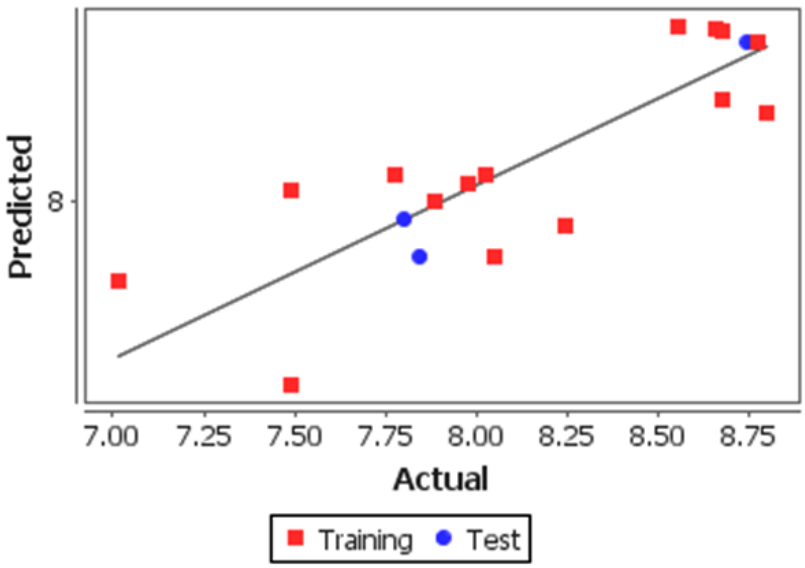


Fig. 4 Fitness plot between actual and predicted activities for model 3

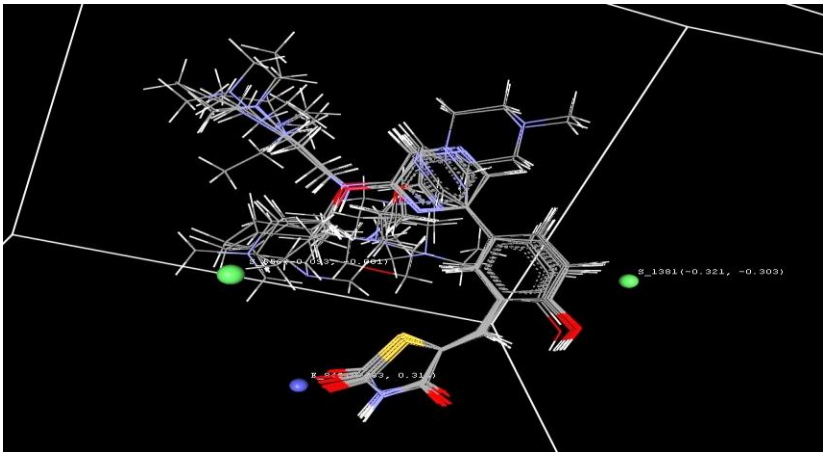


Fig. 5 Contribution plot for steric and electrostatic interactions of all compounds (model 3)

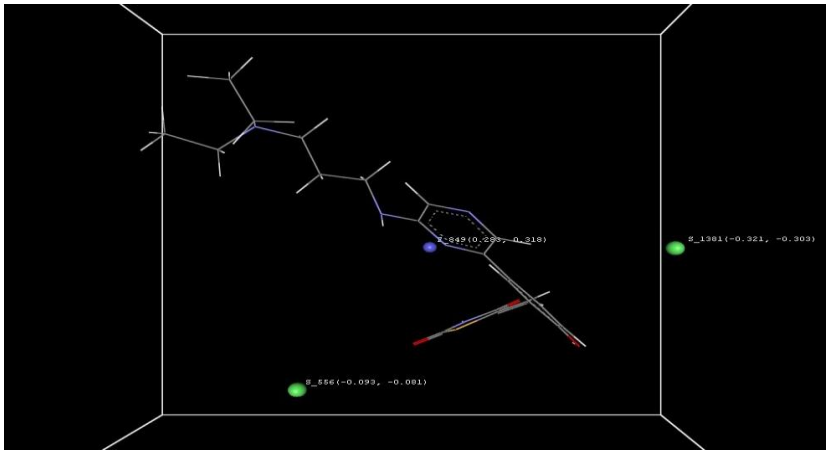


Fig. 6: Contribution plot for steric and electrostatic interactions of most active compound 19 in series (model 3).

3.3. Binding mode analysis by molecular docking: The intermolecular interaction between ligand and targeted kinase PIM-2 was confirmed by docking study through Autodock Vina software. Here active compound 1 binds with different sites on PIM-2 given information for further structural optimization. In this binding mode it was found that most active compound 19 interacts closely with the receptor site and bind with different amino acid residues like PHE-43, GLU-83, ASP-182 and ASP-124 of PIM-2 enzyme through hydrogen bonds (Fig. **7**). Lowest RMSD value was 4.169 Å with -7.8 kcal/mol binding affinity proved good inhibitory activity against PIM-2 kinase. The hydrogen bond distance between PHE-43 to carbonyl, GLU-83 to carbonyl, GLU-83 to hydrogen, ASP-182 to carbonyl and ASP-124 to -NH groups were consecutively 6, 4.5, 6.4, 4.8 and 5.5 Å.

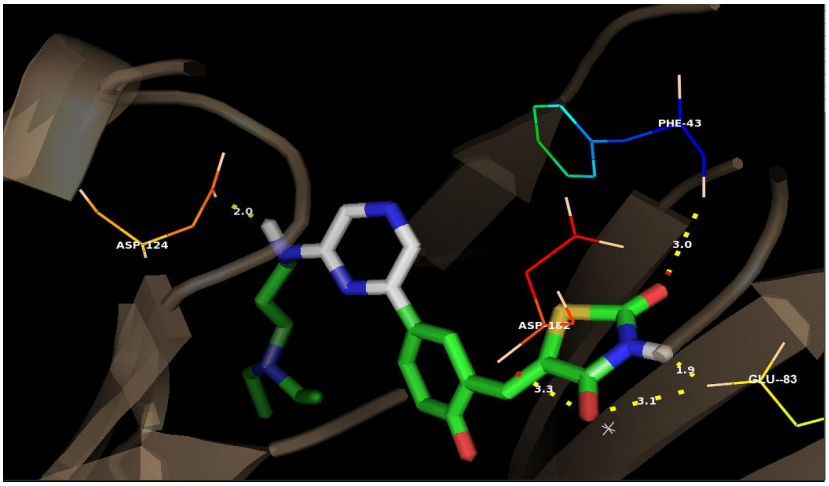


Fig. 7: The docking study of most active compound 19 with protein binding site.

4. CONCLUSION

In the present study an attempt has been made to identify the necessary structural requirements of thiazolidine-2,4-dione derivatives for potential anticancer activity. From QSAR analysis, different models of 2D- and 3D-QSAR have been developed for comparative study between them. Here three best models were generated among which any one can be used for the development of novel compounds as anticancer agents.

Descriptors generated in 2D-QSAR equation proved/highlighted the importance of thiazolidine-2,4-dione scaffold for anticancer activity of compounds. The 2D-QSAR model indicates that the descriptors are statistically significant and agreeable with a high correlation coefficient and reliable predictability. These models generated various descriptors like H-DonorCount, T_2_C_6 Dist Topo and XK Most Hydrophobic Hydrophilic Distance contribute to biological activity. The negative coefficient value of descriptors count on the biological activity indicated that a lower value leads to better anticancer activity. Positive coefficient value of descriptors indicates that a higher value leads to better anticancer activity. The results obtained from 3D-QSAR studies were used to optimize the electrostatic, steric and hydrophobic requirements around the thiazolidine-2,4-dione scaffold for enhancing the anticancer activity. 3D-QSAR studies suggested that substitution with less bulk and more electropositive groups around thiazolidine-2,4-dione core increase anticancer activity. The results of present work may be useful for (medicinal) chemists in understanding the relationship of physiochemical parameters with structure and biological activity and will be helpful to select the suitable substituent for the development of more potent, effective and selective anticancer agents.

ACKNOWLEDGMENT

The authors gratefully acknowledge NRI college of Pharmacy for providing the facilities for conducting research work.

DECLARATION OF INTEREST

The authors report no declarations of interest.

REFERENCES

1. Ajmani S, Jadhav K, Kulkarni SA Three-dimensional QSAR using the k-nearest neighbor method and its interpretation. J Chem Inf Model., 2006; 46: 24-31.
2. Ajmani S, Jadhav K, Kulkarni SA Group-based QSAR (GQSAR): mitigating interpretation challenges in QSAR. QSAR Comb Sci., 2009; 28: 36-41.
3. Ajmani S, Jadhav K, Kulkarni SA. A comprehensive structure–activity analysis of protein kinase B-alpha (Akt1) inhibitors. J Mol Graph Model, 2010; 28: 683-694.
4. Asati V, Mahapatra DK, Bharti SK Thiazolidine-2,4-diones as multi-targeted scaffold in medicinal chemistry: Potential anticancer agents. Eur. J. Med. Chem, 2014; 87: 814-833.
5. Azizmohammadi M, Khoobi M, Ramazani A, Emami S, Zarrin A, Firuzi O, Miri R, Shafiee A, 2H-Chromene derivatives bearing thiazolidine-2,4-dione, rhodanine or hydantoin moieties as potential anticancer agents, Eur. J. Med. Chem, 2013; 59: 15-22.
6. Bataille CJ, Brennan MB, Byrne S, Davies SG, Durbin M, Fedorov O, Huber KV, Jones AM Knapp S, Liu G, Nadali A, Quevedo CE, Russell AJ, Walker RG, Westwood R, Wynne GM, Thiazolidine derivatives as potent and selective inhibitors of the PIM kinase family, Bioorganic & Medicinal Chemistry (2017), doi: http://dx.doi.org/10.1016/j.bmc, 2017.02.056.
7. Baumann K An alignment-independent versatile structure descriptor for QSAR and QSPR based on the distribution of molecular features. J Chem Inf Comput Sci., 2002; 42: 26-35.
8. Brault L, Gasser C, Bracher F, Huber K, Knapp S, Schwaller J (2010) PIM serine/threonine kinases in the pathogenesis and therapy of hematologic malignancies and solid cancers. Haematologica, 2010; 95: 1004–1015.
9. Chang MW, Ayeni C, Breuer S, Torbett BE Virtual Screening for HIV Protease Inhibitors: A Comparison of AutoDock 4 and Vina. PLoS ONE, 2010; 5(8): e11955.
10. Dakin LA, Block MH, Chen H, Code E, Dowling JE, Feng X, Ferguson AD, Green I, Hird AW, Howard T, Keeton EK, Lamb ML, Lyne PD, Pollard H, Read J, Wu AJ, Zhang

T, Zheng X Discovery of novel benzylidene-1,3-thiazolidine-2,4-diones as potent and selective inhibitors of the PIM-1, PIM-2, and PIM-3 protein kinases. Bioorg Med Chem Lett, 2012; 22: 4599–4604.

11. Flanders Y, Dumas S, Caserta J, Nicewonger R, Baldino M, Lee CS, Baldino CM, A versatile synthesis of novel pan-PIM kinase inhibitors with initial SAR study, Tetrahedron Letters, 2015; 56: 3186–3190.
12. Forshell LP, Li Y, Forshell TZ, Rudelius M, Nilsson L, Keller U, Nilsson J The direct Myc target PIM-3 cooperates with other PIM kinases in supporting viability of Myc-induced B-cell lymphomas. Oncotarget, 2011; 2: 448–460.
13. Ghosh P, Bagchi MC QSAR modeling for quinoxaline derivatives using genetic algorithm and simulated annealing based feature selection. Curr Med Chem, 2009; 16: 4032-4048.
14. Golbraikh A, Tropsha A Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. J Comput Aided Mol Des., 2002; 16: 357–369.
15. Halgren TA Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. Journal of Computational Chemistry, 1996; 17: 553–586.
16. Hasegawa K, Kimura T, Funatsu K GA strategy for variable selection in QSAR studies: enhancement of comparative molecular binding energy analysis by GA-based PLS method. Quantitative Structure-Activity Relationships, 1999; 18: 262–272.
17. http://vina.scripps.edu/ (accessed on Sept. 2, 2014).
18. Lee J, Park J, Hong VS (2014) Synthesis and Evaluation of 5-(3-(Pyrazin-2-yl)benzylidene)thiazolidine- 2,4-dione Derivatives as Pan–Pim Kinases Inhibitors. Chem Pharm Bull, 2014; 62(9): 906–914.
19. Mikkers H, Nawijn M, Allen J, Brouwers C, Verhoeven E, Jonkers J, Berns A Mice deficient for all PIM kinases display reduced body size and impaired responses to hematopoietic growth factors. Mol Cell Biol., 2004; 24: 6104–6115.
20. Möröy T, Verbeek S, Ma A, Achacoso P, Berns A, Alt F E mu N and E mu L-myc cooperate with E mu pim-1 to generate lymphoid tumors at high frequency in double-transgenic mice. Oncogene, 1991; 6: 1941–1948.
21. Nagarapu L, Yadagiri B, Bantu R, Kumar CG, Pombala S, Nanubolu J, Studies on the synthetic and structural aspects of benzosuberones bearing 2, 4- thiazolidenone moiety as potential anti-cancer agents, Eur. J. Med. Chem, 2014; 71: 91-97.

22. Nawijn MC, Alendar A, Berns A for better or for worse: the role of Pim oncogenes in tumorigenesis. Nat Rev Cancer, 2011; 11: 23–34.
23. Osterberg F, Morris GM, Sanner MF, Olson AJ, Goodsell DS Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in autodock. Proteins: Struct., Funct., Genet, 2002; 46: 34–40.
24. Rego M.J.B., Galdino-Pitta M.R., Pereira D.T.M., Silva J.C., Rabello M.M., Lima M.C.A., Hernandes M.Z., Pitta I.R., Galdino S.L., Pitta M.G.R., Synthesis, in vitro anticancer activity and in silico study of new disubstituted thiazolidinedione derivatives. Medicinal Chemistry Research, 2014; 23(6): 3220–3226.
25. Sahu NK, Shahi S, Sharma MC, Kohli DV QSAR studies on imidazopyridazine derivatives as PfPK7 inhibitors. Mol Simulat, 2011; 37: 752–765.
26. Seeliger D, Groot BL Ligand docking and binding site analysis with PyMOL and Autodock/Vina. J Comput Aided Mol Des., 2010; 24: 417–422.
27. Sharaf MA, Illman DL, Kowalski BR Chemometrics. New York: Wiley, 1986; 332p.
28. Shen M, Xiao Y, Golbraikh A, Gombar VK, Tropsha A Development and validation of k-nearest-neighbor QSPR models of metabolic stability of drug candidates. J Med Chem, 2003; 46: 3013–3020.
29. Stigliani JL, Bernardes-Génisson V, Bernadoua J, Pratviela G Cross-docking study on InhA inhibitors: a combination of Autodock Vina and PM6-DH2 simulations to retrieve bio-active conformations. Org Biomol Chem, 2012; 10: 6341-6345.
30. Trott O, Olson AJ AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comp Chem, 2010; 31: 455-461.
31. Van Lohuizen M, Verbeek S, Krimpenfort P, Domen J, Saris C, Radaszkiewicz T, Berns A Predisposition to lymphomagenesis in pim-1 transgenic mice: cooperation with c-myc and N-myc in murine leukemia virus-induced tumors. Cell, 1989; 56: 673–682.
32. VLife Sciences Technology Pvt. Ltd. Pune-411045, Web: www.vlifesciences.com
33. Wold S (1995) PLS for multivariate linear modelling. In: van de Waterbeemd H, ed. QSAR: chemometric methods in molecular design. Weinheim: Wiley-VCH 195–218.
34. Zhang Y, Wang Z, Li X, Magnuson NS Pim kinase-dependent inhibition of c-Myc degradation. Oncogene, 2008; 27: 4809-4819.