ScienSage

Journal of Advanced Scientific Research

Available online through http://www.sciensage.info

ISSN
0976-9595
Research Article

COMPUTATIONAL APPROACH IN THE DISCOVERY OF POTENTIAL THIAZOLIDINE-2, 4-DIONE DERIVATIVES AS ANTICANCER AGENTS

Pawan Kumar, Ankur Choubey*

Madyanchal professional University, Bhopal, Madhya Pradesh, India

**Corresponding author: pawanbajpai01@gmail.com*

ABSTRACT

In the present report, we have utilized atom-based 3D-QSAR method to analyze the structural aspects of a series of thiazolidine-2,4-dione derivatives. In this approach the experimental dataset was divided into training (75%) and test (25%) sets and the best model was chosen for the calculation of statistical parameters such as Q^2 and R^2 values. This approach led us to short-list most active derivatives such as compounds 4a, 5c, 5g, 17g, 17n, 17p and 17q with the incorporation of more than one structural feature in a single molecule. Furthermore, 3DQSAR study showed that the thiazole ring substituted with hydrogen bond acceptor is important for activity. Phenyl ring substituted with other heterocyclic rings with electron withdrawing groups may increase the anticancer activity. The validation of experimental results given the idea about further development of pharmacophore based potent anticancer compounds.

Keywords: Anticancer, Pharmacophore modeling, Thiazolidine-2, 4-dione, 3D QSAR, Docking

1. INTRODUCTION

In the drug discovery of potential antitumor molecules, significant efforts have been made for the synthesis of new heterocyclic motifs as the main structural design. Such a heterocyclic scaffold

'thiazolidine-2,4-dione' has been recognized as a 'Master Scaffold' considering for their broad spectrum of biological profiles and affinities towards different targets [1,2]. The molecules based on thiazolidine-2,4-dione derivatives been reported to have potential anticancer activities against different types of cancers [3-5]. Bendamustine [6] and Veliparib [7] are thiazole based drugs approved for the cancers treatment. The different substitutions on this ring and its derivatives have been luring researchers throughout the world to investigate their therapeutic potential [8-11]. On the other hands, thiazolidinedione templates are privileged structural fragments in modern medicinal chemistry and reported to have a diverse range of biological activities [12,13]. The molecules based on this skeleton act as antidiabetic agents via peroxisome proliferator-activated receptor-g (PPAR-g, pioglitazone and its analogues) [14]. Apart from antidiabetic activity, these scaffolds have shown a broad spectrum of activities such as anti-inflammatory [15], antimicrobial [16], wound-healing [17] etc. GSK1059615 is a thiazolidinedione derivative known to exert its anticancer effect through the inhibition of PI3K-a [18]. Liu et al. [19] and Jung et al. [20] reported thiazolidinediones as potential anticancer agents through inhibition of Raf/MEK/ERK and PI3K/Akt signaling pathways. The research on thiazolidinedione for their anticancer activity has recently gained interest and proven promising for cancer treatment. Furthermore, the different heterocyclic ring systems like pyrrolidine, morpholine, piperidine etc have been recognized as the fundamental building blocks of the most drugs in today's market [21]. These nitrogen-containing heterocycles have the capability to form hydrogen-bonding for selective protein binding and enhanced drug-likeness properties [22].

The significance of these rings are well understood in modern medicinal chemistry and drug discovery, since they are known to play a noteworthy role in molecular properties like threedimensionality, lipophilicity or polarity, metabolic stability, and toxicity. From the past few decades, the amalgamation of two dissimilar chemical entities/pharmacophores with two (or more than two) tructural domains having same/different biological functions is being under constant escalation for the exploration of novel and highly active compounds [23-26]. The advantage of using hybridization approach is to turn on different targets by a single molecular entity, thereby increasing therapeutic efficacy and exerting synergistic action with reduced side effects. Hence, in

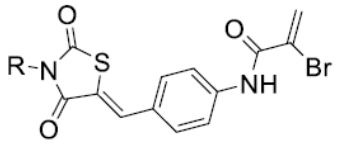
continuation of our research interest [27-29] in discovering and developing new potential anticancer agents, we herein designed new thiazolidinedione derivatives as potential cytotoxic agents. In this view, we performed pharmacophore development and 3DSAR study for the generation of potential thiazolidinedione derivatives.

2. MATERIAL AND METHODS

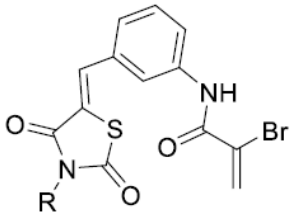
2.1. Collection and preparation of data set

In the present studies, a reported series of 38 thiazolidine-2, 4-dione derivatives [30, 31] has been taken for anticancer activity on selected human cancer cell lines, Cervical cancer cells (HeLa). The biological activity values [IC_{50} (μM)] given in literature were converted to their molar units (M) and then further to negative logarithmic scale (pIC_{50}) and subsequently used as the dependent variable for the QSAR analysis (Table 1). 3D-QSAR method utilizes division of entire dataset into the training and test sets to calculate the predicted pIC_{50} values. The entire dataset was shuffled randomly in such a manner that training and test sets have reasonably uniform distribution of structurally divergent compounds from each cluster with a wide range of pIC_{50} values. Furthermore, for selecting the final splitting of compounds in training and test sets, stepwise random distributions were carried out in the range of 60–90% compounds in a training set with an increment of 5%. Final clustering was performed at this distribution with a PLS factor of 5 and a large number of (100 numbers) random trials were carried out for an unbiased hypothesis of our study. Among these, two trials (trial 45 and 87) were short-listed that showed high Q^2 and R^2 values. From these data, a final model (trial 87) was chosen with seven representative compounds in test set having wide distribution of their experimental activity profiles. Interestingly, the activity profile of compounds in test set was distributed in such a manner that there was at least one data point present in the training set against each data point in test set, which was found to be an important aspect for the QSAR model validation [32,33].

Table 1: Compounds taken for SAR study with their IC_{50} and PIC_{50} values

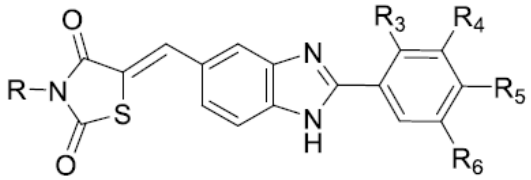


4a to 4k

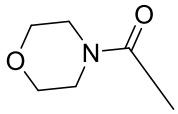
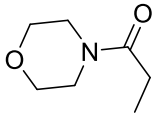
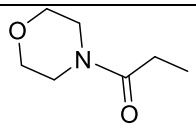
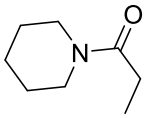
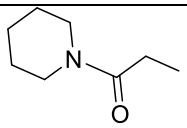
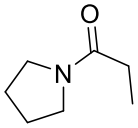
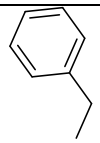
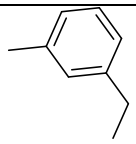
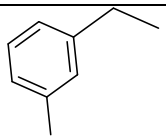
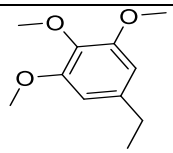


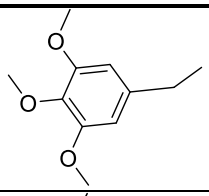
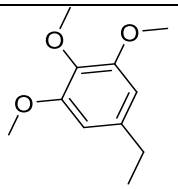
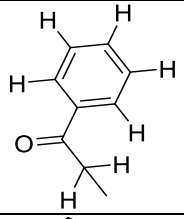
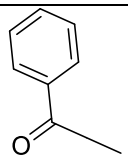
5a to 5i

Compound No.	S No.	Compounds	IC_{50}	p IC_{50}
1	**4a**	CH_3	0.87	6.060481
2	**4b**	C_2H_5	1.4	5.853872
3	**4c**	$C_6H_5CH_2$	1	6
4	**4d**	$C_6H_5(CH_2)_2$	4.5	5.346787
5	**4e**	4-F-$C_6H_4CH_2$	2	5.69897
6	**4f**	3-F-$C_6H_4CH_2$	5.9	5.229148
7	**4g**	4-Cl-$C_6H_4CH_2$	2.3	5.638272
8	**4h**	3-Cl-$C_6H_4CH_2$	8.9	5.05061
9	**4i**	4-CF_3-$C_6H_4CH_2$	13	4.886057
10	**4j**	4-CH_3-$C_6H_4CH_2$	8.9	5.05061
11	**4k**	4-$C(CH_3)_3$-$C_6H_4CH_2$	2.7	5.568636
12	**5a**	$C_6H_5CH_2$	1.4	5.853872
13	**5b**	$C_6H_5(CH_2)_2$	3.5	5.455932
14	**5c**	4-F-$C_6H_4CH_2$	0.8	6.09691
15	**5d**	3-F-$C_6H_4CH_2$	2.7	5.568636
16	**5e**	4-Cl-$C_6H_4CH_2$	2.7	5.568636
17	**5f**	3-Cl-$C_6H_4CH_2$	2.5	5.60206
18	**5g**	4-CF_3-$C_6H_4CH_2$	0.82	6.086186
19	**5h**	4-CH_3-$C_6H_4CH_2$	1.2	5.920819
20	**5i**	4-$C(CH_3)_3$-$C_6H_4CH_2$	2.8	5.552842



17a – 17t

		R3	R4	R5	R6	R	IC_{50}	PIC_{50}
21	**17a**	H	OCH_3	OCH_3	OCH_3	H	1.67	5.777284
22	**17b**	H	OCH_3	OCH_3	H	H	7.83	5.106238
23	**17c**	OCH_3	H	OCH_3	OCH_3	H	15.4	4.812479
24	**17d**	H	OCH_3		H	H	5.22	5.282329
25	**17e**	H	OCH_3	OCH_3	OCH_3		16.3	4.787812
26	**17f**	H	OCH_3	OCH_3	H		7.51	5.12436
27	**17g**	OCH_3	H	OCH_3	OCH_3		0.51	6.29243
28	**17h**	H	OCH_3	OCH_3	OCH_3		23.27	4.633204
29	**17i**	OCH_3	H	OCH_3	OCH_3		6.89	5.161781
30	**17j**	H	OCH_3	OCH_3	OCH_3		4.78	5.320572
31	**17l**	H	OCH_3	OCH_3	OCH_3		0.95	6.022276
32	**17m**	H	OCH_3	OCH_3	OCH_3		1.44	5.841638
33	**17n**	OCH_3	H	OCH_3	OCH_3		0.18	6.744727
34	**17o**	H	OCH_3	OCH_3	OCH_3		2.24	5.649752

35	**17p**	H	OCH_3	OCH_3	H		0.32	6.49485
36	**17q**	H	OCH_3		H		0.25	6.60206
37	**17r**	H	OCH_3	OCH_3	OCH_3		26.26	4.580705
38	**17t**	H	OCH_3	OCH_3	H		7.48	5.126098

2.2. Ligand preparation for QSAR study

The three-dimensional (3D) structures were collected from database and they were cleaned to retain their own conformations. The energy minimization was performed using LigPrep (OPLS 2005 force field), incorporated in PHASE [34, 35]. Conformers were generated using a rapid torsion angle search approach followed by minimization of each generated structure using OPLUS 2005 force field with an implicit GB/SA solvent model [36]. A maximum of 1000 conformers were generated per structure keeping their stereo-chemistry intact and the higher energetic conformers than the threshold value were discarded for the minimization of error level. The superimposition of all molecules based on minimizing root mean square deviation (RMSD), shown in Fig. 1.

2.3. Creating pharmacophore sites

Each compound, selected in the present study, has certain chemical features that may facilitate their non-covalent binding with target. PHASE provides built-in six pharmacophore features such as hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), ring aromaticity (R), positively ionizable (P) and negatively ionizable (N) groups. The rules that are applied to map the positions of pharmacophore sites are defined as 'feature definition'. This feature definition was applied to sets of compounds to determine the common pharmacophore. In this connection, PHASE has created one hydrogen bond donor site (D4) and one hydrogen bond acceptor regions (A2) at the thiazolidine-2,4-dione ring, respectively. Additionally, the benzyl ring present in these compounds was detected as a ring aromaticity (R12) (Fig.1) [37].

2.4. Finding common pharmacophores and scoring hypothesis

The maximum and minimum number of sites was set up to 5 which generate 46 variants for the development of common pharmacophore hypothesis. Common pharmacophores were searched from 12 active compounds with a final box size of 1 Å and minimum inter site distance of 2 Å. The aim to find common pharmacophore is directly related to the identification of sets of features which are very similar in spatial arrangements amongst all ligands. These ligands should contain at least one common pharmacophore on workspace [38].

The pharmacophore hypotheses were scored by using the score hypothesis to identify the pharmacophore from each surviving n-dimensional box that yields the best alignment of the chosen actives. This pharmacophore provides a hypothesis to explain how the active molecules bind to the receptor. Due to the presence of many boxes, different hypotheses were generated. This scoring procedure provides a ranking of the different hypotheses which allow to make rational choices about

which hypotheses are the most appropriate for further investigation [39].

The scoring procedure was started with score actives with the inclusion of weighting factors which defines the survival score of the hypotheses. Once the hypotheses have been scored on the basis of the alignment of the chosen actives, an adjusted score calculated based on the alignment of the chosen inactive (specify a weight for the inactive score). Some different scoring function for survival hypothesis is rescored which generate post-hoc score. In post-hoc score the most active compound used as the reference and a penalty assigned for hypotheses in which the reference ligand shows a high relative conformational energy. Finally hypotheses were clustered on the basis of similarity in variant and their identical scores. Here the actives which showed good fitness scores were selected and used for the generation of 3D-QSAR models. The quality of each pharmacophore is measured in three ways based on the alignments to the input structures: (1) the alignment score, which is the root-mean squared deviation (RMSD) in the site-point positions; (2) the vector score, which is the average cosine of the angles formed by corresponding pairs of vector features (acceptors, donors, and aromatic rings) in the aligned structures; (3) a volume score which measures how well each ligand overlays with the reference ligand and is based on the overlap of van der Waals models of the non-hydrogen atoms in each pair of structures [40-42].

2.5. Building of 3D-QSAR model

The active and inactive scores have been successfully calculated and atom-based 3D-QSAR model was selected to complete the calculation. PLS regression was carried out using PHASE domain with a maximum number of N/5 PLS factors (N = number of ligands in training set, and a grid spacing of 1.0 Å) [43,44].

2.6. ADME property predictions

QikProp from Schrödinger was used for ADME properties predictions allowing us to select ligands with drug-like properties as extra criteria. Specifically, Lipinski's rules of five [MW<500, log P <5, HBA <10, and HBD <5] and Jorgensen's rule of three [log S >−5.7, Caco-2 permeability >22 nm/s, number of primary metabolites <7] were applied to predict oral bioavailability. Default settings were employed for these calculations.

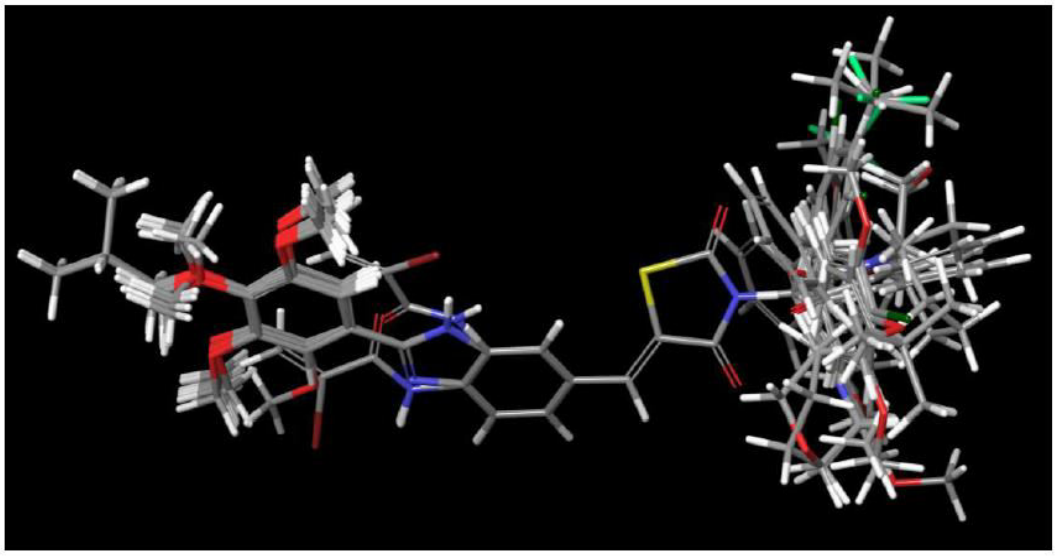


Fig. 1: Pharmacophore based alignment geometry with common pharmacophore features.

3. RESULTS AND DISCUSSION

3.1. Determination of pharmacophore and 3D-QSAR models

In the 3D-QSAR method the entire experimental dataset was shuffled randomly in such a manner that reasonably uniform distribution of structurally divergent compounds with a wide range of pIC_{50} values was present in training and test sets. The final splitting of compounds in training and test sets was decided based on the highest Q^2 and R^2 values with 75% compounds (28 compounds) in training set and the remaining compounds (10 compounds) in test set (Table 1). The

clustering of compounds was performed at this distribution with a PLS factor of 5 and after a large number of (90 numbers) random trials. Among these, two models were short-listed that showed high Q^2 and R^2 values. From these data, a final model was chosen with ten representative compounds in test set having wide distribution of their experimental activity profiles. All the selected compounds from the database were screened by tree-based partition algorithm to obtain five probable common pharmacophore features from the list of variants. No common pharmacophore hypothesis was obtained with five or more common features. Therefore, on applying the scoring function for four-featured common pharmacophore hypothesis using default parameter values, AADHR hypothesis survived. The training set compounds were aligned on this common pharmacophore hypothesis and analyzed by five PLS factors as described in PHASE. PLS factors were selected by trial and error method to get the best statistical model with maximum training set.

In our present study we have calculated the statistical parameters of two common pharmacophore models including their survival score listed in Table 1. As the reliable predictions can preferably come from statistically valid QSAR model, we have selected several statistical parameters such as SD, R2, F, P, RMSE, Q^2 and Pearson-R to evaluate the robustness of QSAR model. As shown in Table 1, all the models have significantly high R^2 values (0.92 and 0.86). These data indicate that all the models can satisfactorily interpret the SAR of the series of training set molecules.

It has been reported by Torpsha that a model with high R^2 is necessary but not a sufficient condition to judge an ideal QSAR model and therefore the best QSAR model should be chosen based on the predictive ability. Based on that hypothesis, we have chosen the best model considering Q^2 values. As shown in Table 1, the Q^2 value of model 1 (0.82) was found to be higher than that of models 2 (0.71), respectively. Therefore, model 1 was chosen as the superior model for carrying out QSAR analysis. Additionally, model 1 has the lowest RMSE value (0.12) and high Pearson-R value (0.91) among all the models. Overall, considering the parameters such as SD, R^2, F, P, RMSE, Q^2 and Pearson-R, the best model was found to be model 1. Therefore, compounds 1-38 used in the present study are screened using model 1 and the relevant parameter such as predicted pIC_{50} values were correlated with the experimentally obtained pIC_{50} values showing residual values and degrees of fitness.

Table 1: Summary of partial least-square (PLS) analysis results for models 1-2 with survival scores and stability.

Statistical parameter	Model 1 (AADHR.33) (run81)	Model 2 (AADHR.49) (run81)
Test	10	28
Training	10	28
Q^2	0.82	0.71
R^2	0.92	0.86
P	5.48e-010	4.58e-010
Pearson-R	0.91	0.89
Stability	0.71	0.82
RMSE	0.12	0.14
Survival	3.62	3.76
SD	0.26	0.27
F	111.8	121.7

SD: standard deviations of the regression; R^2: value of R^2 regression; F: variance ratio; P: significance level of variance ratio; RMSE: root-mean square error; Q^2: value of Q^2 for the predictive activity; Pearson-R: correlation between the predicted and observed activity for the test set.

3.2. Analysis of atom-based PHASE 3D-QSAR model

The pharmacophore alignment geometry obtained from 3D-QSAR model 1 leads to the identification of suitable molecular features, which are important for an effective protein–ligand interaction (Fig. 2). As shown below, four types of common pharmacophore features such as hydrogen bond acceptors (A7, A8), hydrogen bond donors (D10), ring aromaticity (R18) and electron withdrawing group (EWG) (H16) were identified from the entire dataset used here. A7, A8, D10, and H16 were aligned in trapezium geometry and the ring aromaticity R18 is located almost in the center of the formed trapezium. The field distances between A7-A8, A7-D10, A7-H16, A7-R18, A8-D10, A8-H16, A8-R18 and D10-H16 were found to be 4.50 Å, 2.54 Å, 2.45, 5.82, 2.55, 3.06, 6.92 and 2.79, respectively (Fig. 2). Some of the representative field distances of different pharmacophore features are shown in Fig. 2. These features are assumed to play crucial roles for the inhibitory potencies of compounds towards the target enzyme. In addition to these, the effect of EWG and hydrophobic interactions arising out of the flexible hydrocarbon chains present in compounds seem to be significantly important for their relative potencies. The

structures of some of the potent compounds are superimposed to visualize the common pharmacophore screened by our QSAR study. The graph between experimental and predicted activities is given in Fig. 3.

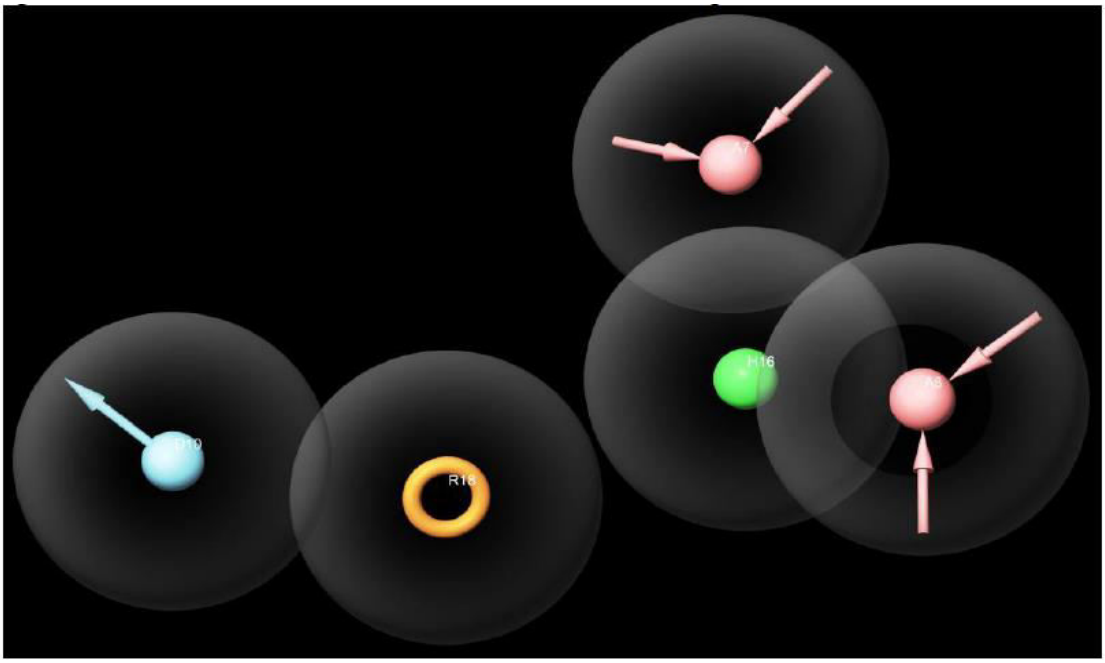


Fig. 2: Pharmacophore alignment for field atomic distances

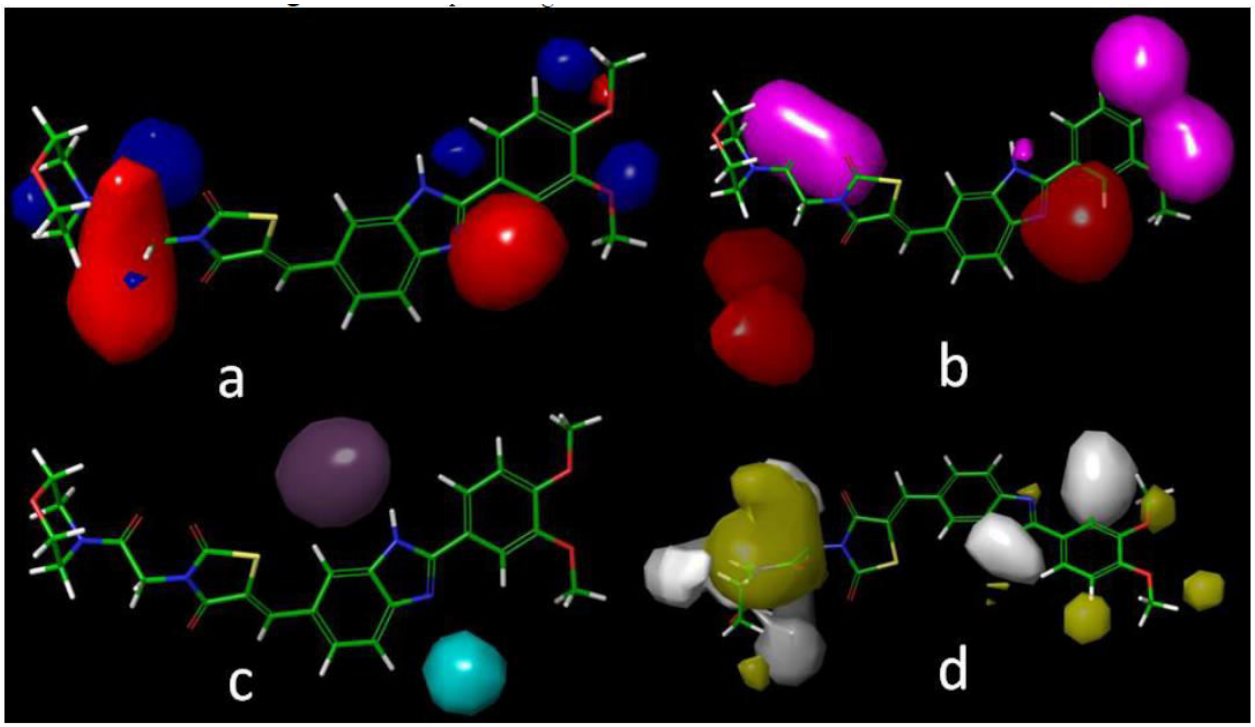


Fig. 3: Field contour maps based on test set compounds; (a) Gaussian electrostatic fields: favoured electropositive (blue) and disfavoured electronegative (red). (b) Gaussian hydrogen bond acceptor field: favoured (red) and disfavoured (magenta). (c) Gaussian hydrogen bond donor field: favoured (purple) and disfavoured (cyan). (d) Gaussian hydrophobic field: favoured (yellow) disfavoured (white)

3.3. Effects of substituent's

As shown below, all compounds were designed using various functional groups/moieties of varying electronic effects at various positions of thiazolidine-2,4-dione. Based on the experimental results, the presence of electron withdrawing groups such as fluorine and CF_3 at N on thiazolidine-2,4-dione found to be very crucial for activity 5c and 5g. Further in the second series of compounds 17g, 17n, 17p and 17q showed potential activity in which morphilino group attached to nitrogen and methoxy group (at para position) which attached to phenyl group connected with benzimidazole moiety. This substitution produced most active compound of the series. In the present report, we have carried out detailed 3D-QSAR study (Fig.4) to predict their activities and to understand the importance of different functionalities/moieties present in compounds to correlate with their experimentally observed potencies. Compound 5g with trifluoro substitution at 4-position of benzene ring (4-CF_3 group) is considered as the standard compound with moderately good inhibition potency (Exp-pIC_{50}: 6.08; Pred-pIC_{50}: 6.03) as shown in Table 2. Replacement of the 4-CF_3 group in compound 5g with other electro donating group such as methyl responsible for decrease in activity. Fig .4 showed the correlation graph between experimental and predicted inhibitory activity of thiazolidine-2,4-dione.

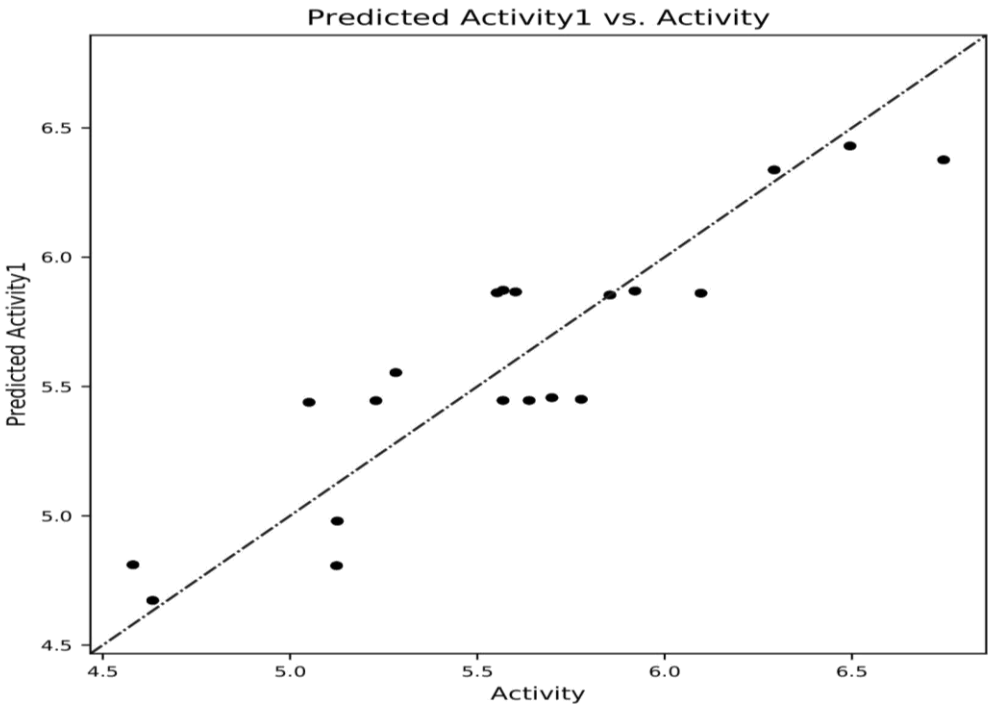


Fig. 4: Correlation graph between experimental and predicted inhibitory activity of thiazolidine-2,4-dione using pharmacophore-based QSAR model for training and test set

Table 2: Comparative inhibitory data (experimental and predicted) of the compounds employed in the present QSAR study along with their residual values and degrees of fitness.

Compound No	Compound name	IC50	Experimental pIC50	Predicted pIC50	Residual value	Fitness
1	4a	0.87	6.06	6.06	0.00	2.69
2	4b	1.4	5.85	5.58	0.27	2.65
3	4c	1	6.00	5.64	0.36	2.68
4	4d	4.5	5.35	5.44	-0.09	2.64
5	4e	2	5.70	5.67	0.03	2.57
6	4f	5.9	5.23	5.20	0.03	2.66
7	4g	2.3	5.64	5.67	-0.03	2.69
8	4h	8.9	5.05	5.17	-0.11	2.69

9	4i	13	4.89	5.13	-0.24	2.65
10	4j	8.9	5.05	5.58	-0.53	2.68
11	4k	2.7	5.57	5.74	-0.17	2.64
12	5a	1.4	5.85	5.76	0.10	2.57
13	5b	3.5	5.46	5.33	0.13	2.66
14	5c	0.8	6.10	5.76	0.34	2.69
15	5d	2.7	5.57	5.76	-0.19	2.69
16	5e	2.7	5.57	5.76	-0.19	2.65
17	5f	2.5	5.60	5.77	-0.17	2.68
18	5g	0.82	6.09	6.03	0.03	2.64
19	5h	1.2	5.92	5.75	0.17	2.57
20	5i	2.8	5.55	5.68	-0.13	2.66
21	17a	1.67	5.78	5.46	0.32	2.69
22	17b	7.83	5.11	5.03	0.07	2.69
23	17c	15.4	4.81	4.65	0.16	2.65
24	17d	5.22	5.28	5.41	-0.13	2.68
25	17e	16.3	4.79	4.89	-0.10	2.64
26	17f	7.51	5.12	4.98	0.15	2.57
27	17g	0.51	6.29	6.38	-0.09	2.66
28	17h	23.27	4.63	4.75	-0.12	2.69
29	17i	6.89	5.16	5.23	-0.07	2.69
30	17j	4.78	5.32	5.32	0.00	2.65
31	17l	0.95	6.02	6.22	-0.19	2.68
32	17m	1.44	5.84	5.38	0.46	2.64
33	17n	0.18	6.74	6.57	0.18	2.57
34	17o	2.24	5.65	5.71	-0.06	2.66
35	17p	0.32	6.49	6.65	-0.16	2.69
36	17q	0.25	6.60	6.51	0.09	2.57
37	17r	26.26	4.58	4.78	-0.20	2.66
38	17t	7.48	5.13	4.94	0.19	2.69

3.4. External validation

External validation in QSAR study is essential step for determining the robustness of the developed model given in Table 3. In 3D QSAR model development squared correlation coefficient (r^2) was calculated by the using of Eq. (1). Internal validation was carried out using 'leave-one-out' (LOO-q^2) method [45]. The cross-validated coefficient, (r^2cv) (also called the LOO-q^2) was calculated by the using of Eq. (2).

$$R^2 \text{ or } r^2 = 1 - \frac{\sum(\tilde{Y}_{i\,(training)} - Y_{i(training)})^2}{\sum(Y_{i\,(training)} - \bar{Y}_{(training)})^2} \quad (1)$$

$$r^2_{cv} \text{ or } q^2 = 1 - \frac{\sum(Y_{i(training)} - \tilde{Y}_{i(training)(LOO)})^2}{\sum(Y_{i(training)} - \bar{Y}_{o(training)})^2} \quad (2)$$

Where, Y_i and $\tilde{Y}_i$ are the observed and predicted activities of the training set molecules, respectively, and $\bar{Y_o}$ and $\bar{Y_p}$ are the average values of the observed and predicted pIC_{50} values of the training set molecules.

However, a high q^2 value was not necessarily a suitable representation of the real predictive power of the model evaluating PIM-1inhibitory activity. Hence, an external validation was carried out to determine the external predictive power of the model using the test set molecules, which were not included in the QSAR model development, denoted by r^2pre represented in Eq. (3). External validation was necessary to evaluate

the true predictive abilities of the established models. The predictive ability of the selected model was also confirmed by $(r^2- R_0^2)/r^2$, $(r^2- R'^2_0)/r^2$, k and k′, r_m^2 and R_p^2. Based on previous findings and developments, the properties of external validation of a 3D-QSAR model could be acceptable when the 3D-QSAR model satisfies the following statistical parameters: $r^2cv > 0.5$, $r^2 > 0.6$, the value of R_0^2 or R'^2_0 should be similar to that of r^2, i.e., $[(r^2 - R_0^2)/r2] < 0.1$ or $[(r^2 - R'^2_0)/r^2] < 0.1$, where, here, the corresponding values of k and k′ should be $0.85 \leq k \leq 1.15$ or $0.85 \leq k' \leq 1.15$, $r_m^2 > 0.5$, and $R_p^2 > 0.5$ [46-49].

By taking regressions in the scatter plots of Y_i vs $\tilde{Y}_i$ or $\tilde{Y}_i$ vs Y_i through the origin, where $Y^{ro} = k\tilde{Y}$ and $\tilde{Y}^{ro} = kY$ (with the intercept set to 0), respectively, [46] the slopes k and k' were calculated by Eq. (5) and (6). Both correlation coefficients for the regression lines through the origin (Y^{ro}) can be defined by R_0^2 and R'^2_0, which were calculated by Eq. (7) and (8).

$$r^2_{pred} = 1 - \frac{\sum(\tilde{Y}_{i(test)} - Y_{i(test)})^2}{\sum Y_{i(test)} - \bar{Y}_{o(training)})^2} \quad (3)$$

$$R = \frac{\sum(Y_{i(test)} - \bar{Y}_{o(test)})(\tilde{Y}_{i(test)} - \bar{Y}_{p(test)})}{\sqrt{\sum(Y_{i(test)} - \bar{Y}_{o(test)})^2 \sum(\tilde{Y}_{i(test)} - \bar{Y}_{p(test)})^2}} \quad (4)$$

$$K = \frac{\sum Y_{i(test)} \tilde{Y}_{i(test)}}{\sum \tilde{Y}_{i(test)}^2} \quad (5)$$

$$K' = \frac{\sum Y_{i(test)} \tilde{Y}_{i(test)}}{\sum Y_{i(test)}^2} \quad (6)$$

$$R_o^2 = 1 - \frac{\sum(\tilde{Y}_{i(test)} - Y_i^{ro})^2}{\sum(\tilde{Y}i_{(test)} - \bar{Y}_{P(test)})^2} \quad (7)$$

$$R'^2_o = 1 - \frac{\sum(Y_{i(test)} - \tilde{Y}_i^{ro})^2}{\sum(Y_{i(test)} - \bar{Y}_o)^2} \quad (8)$$

$$r_m^2 = r^2(1- \sqrt{|r^2 - R_o^2|}) \quad (9)$$

$$r'^2_m = r^2(1- \sqrt{|r^2 - R'^2_o|}) \quad (10)$$

Table 3: External Statistical Validation of Quantitative Structure−Activity Relationship (QSAR) for Hypothesis AADHR_1

External validation	Results obtained	Limitations
Cross-validated coefficient, r^2_{cv}	0.93	$r_{cv}^2 > 0.5$
Correlation coefficient between the actual and predicted activities R and r^2_{pre}	0.85 and 0.70	R~1 and r2 > 0.5
Slope values of regression lines K and K'	0.99 and 1.01	$0.85 \leq k \leq 1.15$
Correlation coefficients for the regression lines through the origin R_0^2 and R'^2_0 $(r^2 - R_0^2)/r^2$ and $(r^2 - R'^2_0)/r^2$	0.97 and 0.97 -0.066 and -0.069	Close to r^2 < 0.1
Modified squared correlation coefficient for the "leave-one-out" method, r_m^2(LOO) r_m^2 and r'^2_m	0.686 and 0.682	> 0.5

The external predictability of the selected model was also checked by modified r^2 (r_m^2, r'^2_m), as proposed by Roy and Roy (2007) [50], by the using Eqs. (9), (10). The entire set considering LOO-predicted values for the training set and predicted values of the test set compounds. The r_m2 statistic equation for overall test and training set values were generally used for selection of the best predictive models. For the calculation of r_m^2, r'^2_m values r^2 taken as the non-cross-validated correlation coefficient obtained from the PLS process whereas R_0^2 and R'^2_0 is calculated from Eq. (7) and (8).

3.5. ADME Prediction

The ADME properties of all previously synthesized compounds have been calculated in comparison with active molecules (Table 4). It has been observed that though all the ADME properties of compounds are well within the acceptable range. Here compound 29 showed high percentages of human oral absorption and partition coefficient. Furthermore compound 29 showed good cell permeability compared to other compounds. From all these results it has been observed that these analogs by making no changes in the core fused scaffold and isosteric/ bioisosteric and knowledge based side chains and fragment attachment came up with very potential lead molecules with favorable drug-like profile which can emerge as a potential drug molecule in further development.

[a]Total solvent accessible surface area (SASA) in square angstromsusing a probe with a 1.4 Å radius (range=300-1000);

[b]Predicted octanol/water partition coefficient (Range= -2.0-6.5);

[c]Predicted apparent Caco-2 cell permeability in nm/sec. Caco-2 cells are a model for the gut blood barrier (,<25% is poor, >500 great.

[d]Prediction of binding to human serum albumin (Range –1.5 – 1.5);

[e]Predicted human oral absorption on 0 to 100% scale. >80% is high,<25% is poor.

3.6. Optimization of novel ligands

The 3D-QSAR studies may be used for the development of novel compounds against as anticancer agents. Here we have pictorially represented the structure activity relationships (SARs) of thiazolidine-2,4-diones scaffold with different possible substituent's on the basis of results obtained by 3D-QSAR study (Fig. 5).

Table 4: ADME and drug-likeness property of most active five compounds

Componds	Mol_MW	SASA[a]	HB donor	HB accpt	QPlogPo/w[b]	QPPCaco[c]	QPlogKhsa[d]	Percent Human Oral Absorption[e]
33	323.80	521.16	3	5	1.35	46.04	-0.05	64.60
35	303.38	517.07	3	5	1.18	49.42	-0.03	64.20
36	289.35	497.11	3	5	0.88	46.03	-0.14	61.85
1	289.35	498.07	3	5	0.87	44.33	-0.14	61.48
14	323.80	508.25	2	5.5	1.44	87.06	-0.03	70.08
18	389.47	674.06	2	6.5	3.10	259.93	0.41	88.32

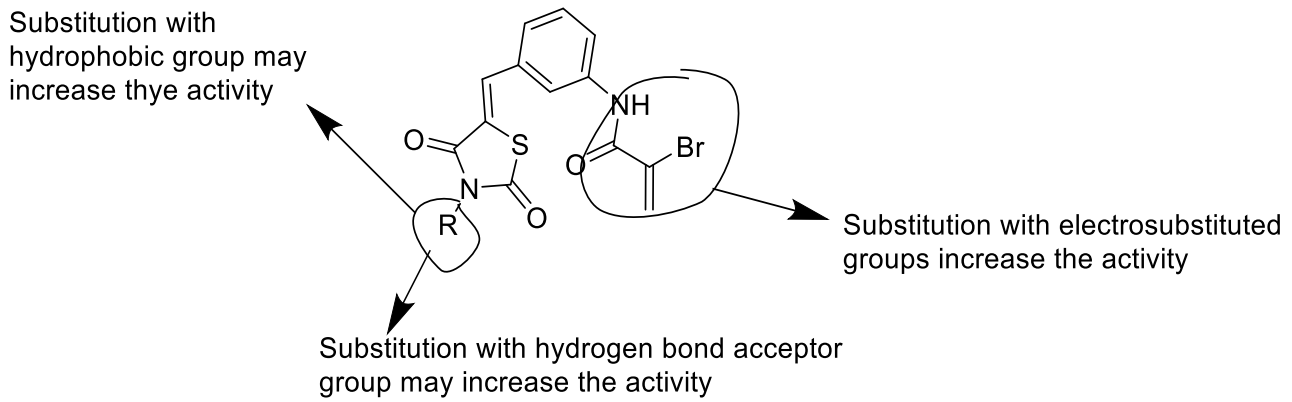


Fig. 5: Ligand scaffold with different features optimized by 3D-QSAR study for the development of novel compounds.

Here we can see the thiazole ring substituted with hydrogen bond acceptor is important for activity. Phenyl ring substituted with other hetercyclic rings with electron withdrawing groups may increase the activity. Furthermore, hydrophobic group substitution on thiazole ring may also increase the anticancer activity.

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. In summary, the atom-based 3D-QSAR studies has been performed on thiazolidine-2,4-dione derivatives. The common pharmacophore model predicted by 3D-QSAR method implies the presence of several important pharmacophore features such as the presence of electron withdrawing group, aromaticity, EDG and hydrophobic long chain along with the presence of H-bond donors and acceptors for their inhibitory potencies towards cancer. Resemblance of these pharmacophore features as predicted by 3DQSAR study with the experimentally outlined structural features in the thiazolidine-2,4-dione analogs indicates the suitability of 3D-QSAR approach to validate the experimental results. This approach led us to short-list most active derivatives such as compounds 4a, 5c, 5g, 17g, 17n, 17p and 17q with the incorporation of more than one structural feature in a single molecule. The description of 3D QSAR study showed the thiazole ring substituted with hydrogen bond acceptor is important for activity. Phenyl ring substituted with other hetercyclic rings with electron withdrawing groups may increase the anticancer activity. Furthermore, hydrophobic group substitution on thiazole ring may also increase the anticancer activity. From the overall analyses, we conclude that the Model 1 (AADHR_1) pharmacophore truly reflects the features of potent inhibitors and this pharmacophore could be used as fast and accurate tool to assist discovery of novel PIM-1 inhibitors

Declaration of interest

The authors report no declarations of interest.

5. REFERENCES

1. Yadav G, Ganguly S. Eur. *J. Med. Chem.,* 2015; **97**:419-443.
2. Akhtar W, KhanMF, Verma G, Shaquiquzzaman M, Rizvi MA, Mehdi SH, et al. *Eur. J. Med. Chem.,* 2017; **126**:705-753.
3. Gu W, Miao TT, Hua DW, Jin XY, Tao XB, Huang CB, Wang SF. *Bioorg. Med. Chem. Lett,* 2017; **27**:1296-1300.
4. Kim MK, Shin H, Park KS, Kim H, Park J, Kim K, et al. *J. Med. Chem.* 2015; **58**:7596-7602.
5. Sharma P, Reddy TS, Thummuri D, Senwar KR, Kumar NP, Naidu VGM, et al. *Eur. J. Med. Chem.* 2016; **124**:608-621.
6. Kath R, Blumenstengel K, Fricke H J, H€offkenK. *J. Cancer Res. Clin. Oncol.,* 2001; **127**:48-54.
7. Wagner L M.*Onco. Targets Ther.,* 2015; **8**:1931-1939.
8. (a) Chen, Liu GZ, Zhang Y, Shan X, Jiang L, Zhao Y et al. *ACS Med. Chem. Lett.* 2013; **4**:69-74;

(b) Kamal A, Suresh P, Ramaiah MJP, Reddy TS, Kapavarapu RK, Imthiajali S, et al. *Bioorg. Med. Chem.* 2013; **21**:5198-5208.

(c) Kamal A, Reddy TS, Vishnuvardhan MVPS, Nimbarte VK, Rao AVS, SrinivasuluV et al. *Bioorg. Med. Chem.*, 2015; **23**:4608-4623.

9. Khan I, Tantray MA, Hamid H, AlamMS, Kalam A, Dhulap A. *Bioorg. Med. Chem. Lett.,* 2016; **26**:4020-4024.
10. Reddy TS, Kulhari H, Reddy VG, Bansal V, Kamal A, Shukla R. *Eur. J. Med.Chem.,* 2015; **101**:790-805.
11. Sharma P, Thummuri D, Reddy TS, Senwar KR, Naidu VGM, Srinivasulu G, et al. *Eur. J. Med. Chem.,* 2016; **122**:584-600.
12. Panigrahy D, Shen LQ, Kieran MW, Kaipainen A. *Expert Opin. Investig. Drugs,* 2003; **12**:1925-1937.
13. Jain VS, Vora DK, Ramaa CS. *Bioorg. Med. Chem.,* 2013; **21**:1599-1620.
14. Pfützner A, Weber MM, ForstT. *Expert Opin. Pharmacother,* 2007; **8**:1985-1998.
15. Faine LA, Rudnicki M, esar FAC, Heras BL, Bosc L, Souza ES, et al. *Curr. Med. Chem.*, 2011; **18**:3351-3360.
16. Guo M, Zheng C-J,Song M-X, Wu Y, Sun L-P, Li Y-J, Liu Y, Piao H-R. *Bioorg. Med. Chem. Lett.*, 2013; **23**:4358-4361.
17. Wu Y, Karna S, Choi CH, Tong M, Tai H-H, Na DH, et al. *J. Med. Chem.,* 2011; **54:**5260-5264.
18. Knight SD, Adams ND, Burgess JL, Chaudhari AM, Darcy MG, Donatelli CA, et al. *ACS Med. Chem. Lett.*, 2010; **1**:39-43.
19. Liu K, Rao W, Parikh H, Li Q, Guo TL,Grant S, et al. *Eur. J. Med. Chem.,* 2012; **47**:125-136.
20. Jung K-Y, Samadani R, Chauhan J, Nevels K, Yap JL, Zhang J, et al. *Org. Biomol. Chem.* 2013; **11:**3706-3732.

21. Taylor RD, Maccoss M, Lawson ADG. *J. Med. Chem* 2014; **57**:5845-5859.
22. Vitaku E, Smith DT,Njardarson JT. *J. Med. Chem.*, 2014; **57**:10257-10274.
23. Bolognese A, Correale G, Manfra M, Lavecchia A, Novellino E, Barone V. *Org. Biomol. Chem.,* 2004; **2**:2809-2813.
24. (a) Nekkanti S, VeeramaniK, Kumar NP, Shankaraiah N. *Green Chem.*, 2016; **18**:3439-3447.
(b) Shankaraiah N, MarkandeyaN, Moraga M-E, Kamal A, Santos LS. *Synthesis* 2009; 2163-2170.
(c) Nekkanti S, Veeramani K, Kumari SS, Tokala R, Shankaraiah N. *RSC Adv.,* 2016; **6**:103556-103566.
25. Meunier B. *Acc. Chem. Res.*, 2008; **41**:69-77.
26. Shaveta S, Mishra P, Singh. *Eur. J. Med. Chem.,* 2016; **124**:500-536.
27. Sharma P, Senwar KR, Jeengar MK, Reddy TS, Naidu VGM, Kamal A, et al. *Eur. J. Med. Chem.,* 2015; **104**:11-24.
28. Senwar KR, Reddy TS, Thummuri D, Sharma P, Bharghava SK, Naidu VGM, et al. *Bioorg. Med. Chem. Lett.,* 2016; **26**:4061-4069.
29. Senwar KR, Sharma P, Reddy TS, Jeengar MK, Nayak VL, Naidu VGM, et al. *Eur. J. Med. Chem.,* 2015; **102**:413-424.
30. Romagnoli R, Baraldi PG, Salvador MK, Camacho ME, Balzarini J, Bermejo J, et al. *Eur J Med Chem.*, 2013; **63**:544-57.
31. Sharma P, Reddy TS, Kumar NP, Senwar KR, Bhargava SK, Shankaraiah N. *Eur J Med Chem.*, 2017; **138**:234-245.
32. Leonard JT, Roy K. *QSAR Comb. Sci.*, 2006; **25**:235.
33. Roy K, Paul S. *QSAR Comb. Sci* 2009; **28**:406.
34. Ligprep. version 2.5, Schrödinger, LLC, New York, NY., 2012.
35. Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. *J. Comput. Aided Mol. Des.*, 2006; **20**:647.
36. Kleinjung J, Fraternali F. *Curr. Opin. Struct. Biol.,* 2014; **25 :**126.
37. Phase, version 4.4, Schrödinger, LLC, New York, NY, 2012.
38. Phase 4.4 Quick Start Guide, Schrödinger, LLC, May 2013.
39. Sallam AA, Houssen WE, Gissendanner CR, Orabi KY, Foudah AI, El Sayed KA. *Med. Chem. Comm.,* 2013; **4**:1360-1369.
40. Hall MD, Salam NK, Hellawell JL, Fales HM, Kensler CB, Ludwig JA, et al. *J. Med. Chem.,* 2009; **52**:3191-3204.
41. Teli MK, Rajanikant GK. *Org. Med. Chem. Lett.,* 2012; **2**:1-10.
42. Kamaria P, Kawathekar N. *Med. Chem. Res.,* 2013; **23**:25-23.
43. Golbraikh A, Tropsha A. *J. Comput. Aid. Mol. Des.,* 2002; **16:**357-369.
44. Tanwar OP, Saha R, Alam MM, Akhtar M. *Med Chem Res.,* doi:10.1007/s00044-010-9523-y.
45. Cramer RD, Patterson DE, Bunce JD. *J. Am. Chem. Soc.,* 1988; **110**:5959-5967.
46. Golbraikh A, Tropsha A. *J. Mol. Graphics Modell.,* 2002; **20**:269-276.
47. Roy PP, Roy K. *QSAR Comb. Sci.,* 2008; **27**:302-313.
48. Roy K, Roy PP. *Eur. J. Med. Chem.*, 2009; **44**:2913-2922.
49. Nayagam VM, Wang X, Tan YC, Poulsen A, Goh KC, Ng T, et al. *J. Biomol. Screening,* 2006; **11**:959-967.
50. Roy PP, Roy K. *Expert Opin. Drug Discovery***,** 2007; **2**:1567-1577.