Research Paper

Ligand-based Pharmacophore Modeling; Atom-based 3D-QSAR Analysis and Molecular Docking Studies of Phosphoinositide-Dependent Kinase-1 Inhibitors

P. KIRUBAKARAN, K. MUTHUSAMY*, K. H. D. SINGH AND S. NAGAMANI
Department of Bioinformatics, Science Block, Alagappa University, Karaikudi–630 004, India

Phosphoinositide-dependent kinase-1 plays a vital role in the PI3-kinase signaling pathway that regulates gene expression, cell cycle growth and proliferation. The common human cancers include lung, breast, blood and prostate possess over stimulation of the phosphoinositide-dependent kinase-1 signaling and making phosphoinositide-dependent kinase-1 an interesting therapeutic target in oncology. A ligand-based pharmacophore and atom-based 3D-QSAR studies were carried out on a set of 82 inhibitors of PDK1. A six point pharmacophore with two hydrogen bond acceptors (A), three hydrogen bond donors (D) and one hydrophobic group (H) was obtained. The pharmacophore hypothesis yielded a 3D-QSAR model with good partial least square statistics results. The training set correlation is characterized by partial least square factors (R^2 = 0.9557, SD = 0.2334, F = 215.5, P = 1.407e-32). The test set correlation is characterized by partial least square factors (Q^2 ext = 0.7510, RMSE = 0.5225, Pearson-R =0.8676). The external validation indicated that our QSAR model possess high predictive power with good value of 0.99 and value of 0.88. The docking results show the binding orientations of these inhibitors at active site amino acid residues (Ala162, Thr222, Glu209 and Glu166) of phosphoinositide-dependent kinase-1 protein. The binding free energy interactions of protein-ligand complex have been calculated, which plays an important role in molecular recognition and drug design approach.

Key words: External validation, molecular docking and prime MM-GBSA, PDK1, PHASE, 3D-QSAR

Phosphoinositide-dependent kinase-1 (PDK1) is a serine/threonine kinase protein, plays an important role in various signaling cascades[1,2]. Biochemical and genetic studies have shown that PDK1 is a major player in the phosphatidylinositol 3-kinases (PI 3-kinases or PI3Ks) signaling pathway that regulates cell cycle, gene expression, growth and proliferation by phosphorylation of the numerous AGC kinase families (PKB/AKT, PKC, SGK, p70S6K and PDK1 itself)[3]. Approximately 50% of common tumor (breast, lung, ovarian, gastric and prostate) types possess mutations in genes that regulate PIP3 (phosphatidylinositol trisphosphate) production and this imparts these cancer cells with abnormally high levels of this lipid second messenger[4,5]. As a result, PIP3 causes over stimulation of PDK1 and constitutively activates the AGC kinases. Consequence of this regulation by PDK1 results in various effects such as reduced apoptosis, enhanced tumor cell proliferation and angiogenesis[6,7]. The inhibitors which were identified for PDK1 showed better activity against Wnt/β-catenin signaling in medulloblastoma tumorigenesis[8].

PDK1 has an N-terminal kinase and C-terminal pleckstrin homology (PH) domains. The interaction of PDK1 with cell membrane was carried out by PH domain and the phosphorylation and activation of downstream kinases were carried out by kinase domain[3]. Accumulating pharmacologic and genetic evidence supports the vital role of PDK1 as a promising anticancer target[9,10]. Till date 14 small-molecule inhibitor-PDK1 complexes are available in the protein data bank (PDB)[11], providing extensive structural information in terms of ligand-protein interactions[12].

Some of the inhibitors of PDK1, which have been reported in the literature as potent nanomolar inhibitors which are belong to the various groups of chemicals sach as bisindolylmaleimides, thieno [3,2-c] pyridine-7-carboxamides, N-phenylpyrimidin-2-amines, diazepinone derivatives, celecoxib derivatives and

***Address for correspondence**
E-mail: mkbioinformatics@gmail.com

7-aminopyrazolo[1,5-a]pyrimidines (figs. 1-3) were selected for QSAR studies[13]. Till now there is no evidence for QSAR studies using these inhibitors against PDK1 protein and QSAR studies have been reported only on indolinone derivatives.

The current study deals with pharmacophore hypothesis generation, correlation analysis by 3D-QSAR and receptor-guided molecular docking techniques to gain further insight into their structure activity relationship.

The pharmacophore modeling was carried out using pharmacophore alignment and scoring engine (PHASE). PHASE performs systematic explorations of rotatable bonds and calculates the associated conformational energies, retaining only the most reasonable conformations. High-dimensional, tree-based partitioning algorithm was used for plausible mapping which performs different conformations are placed in multi-dimensional boxes. The common pharmacophore was represented by each boxes and if it contains a

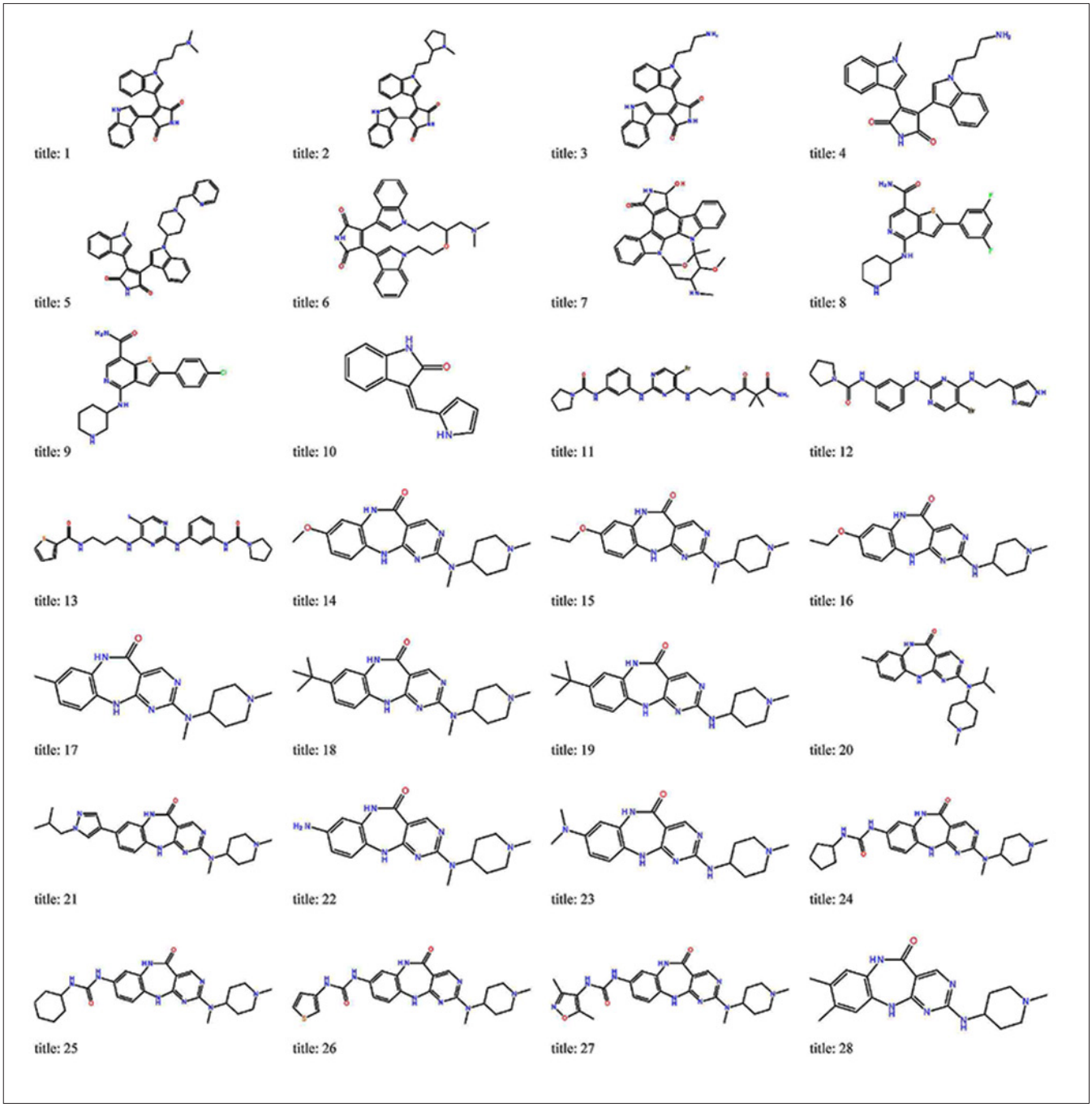

Fig. 1: 2D repesentation of the compounds used for QSAR studies.

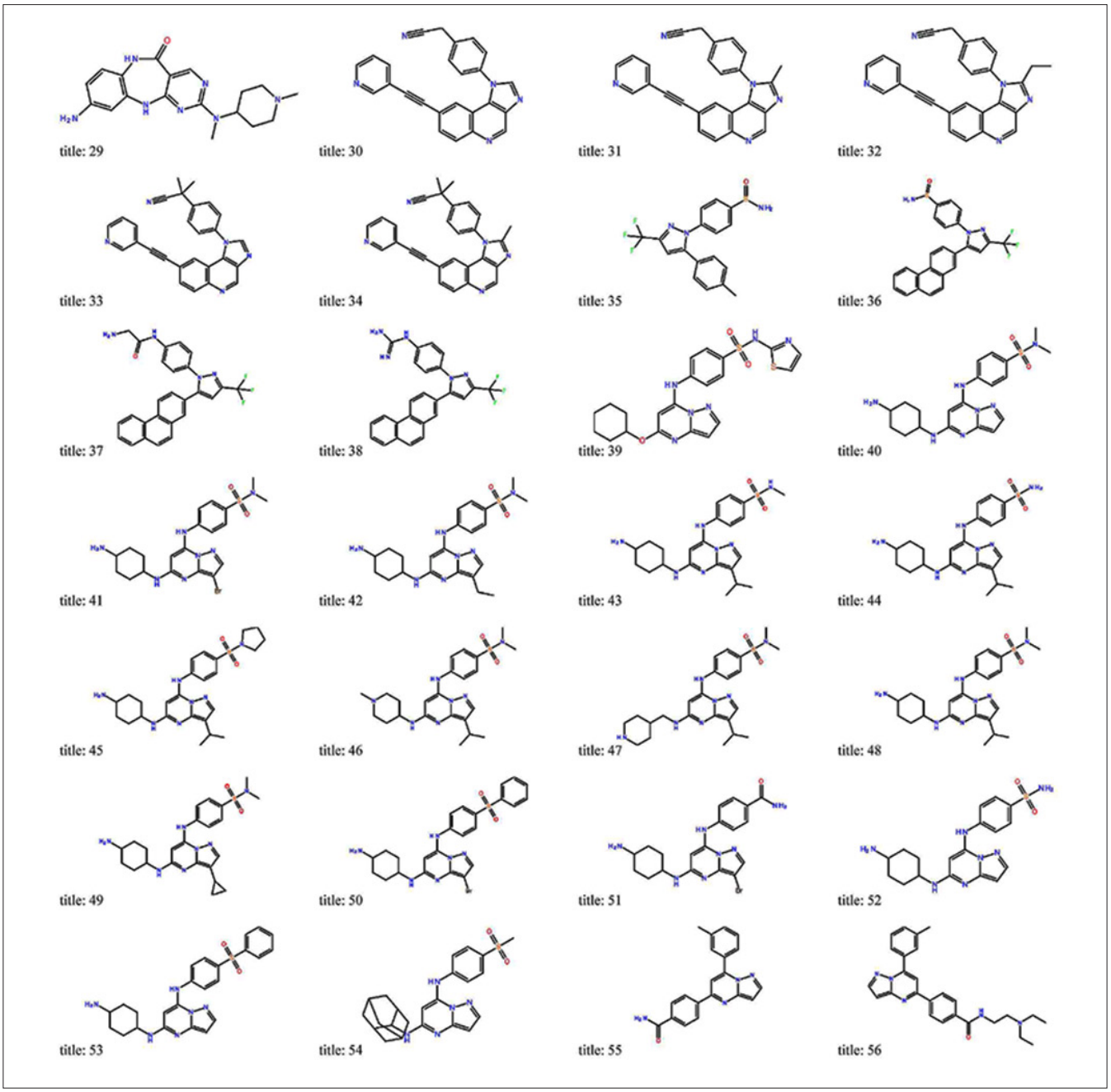

Fig. 2: 2D repesentation of the compounds used for QSAR studies.

sufficient number of active ligands. Finally PHASE performing a partial least-square (PLS) regression, resulting in prediction a significant model[14,15]. The binding mode of the active molecule with the active site amino acid residues of PDK1 was analyzed by docking using Glide XP. The constructed hypothesis and 3D-QSAR models not only can be used in rapid and accurate prediction of the activities of newly designed inhibitors, it also provide the better tools for rational design of promising PDK1 inhibitors having greater therapeutic safety and efficacy[16,17].

MATERIALS AND METHODS

All computational and molecular modeling studies were carried out on a Red Hat 5.1 Linux platform in Lenovo Intel core 2 duo processor RAM 2 GB using the molecular modeling software package Schrodinger, LLC, New York, 2010.

Dataset analysis:

The experimental information of 82 molecules were collected from the literature[13] and pIC_{50} values were

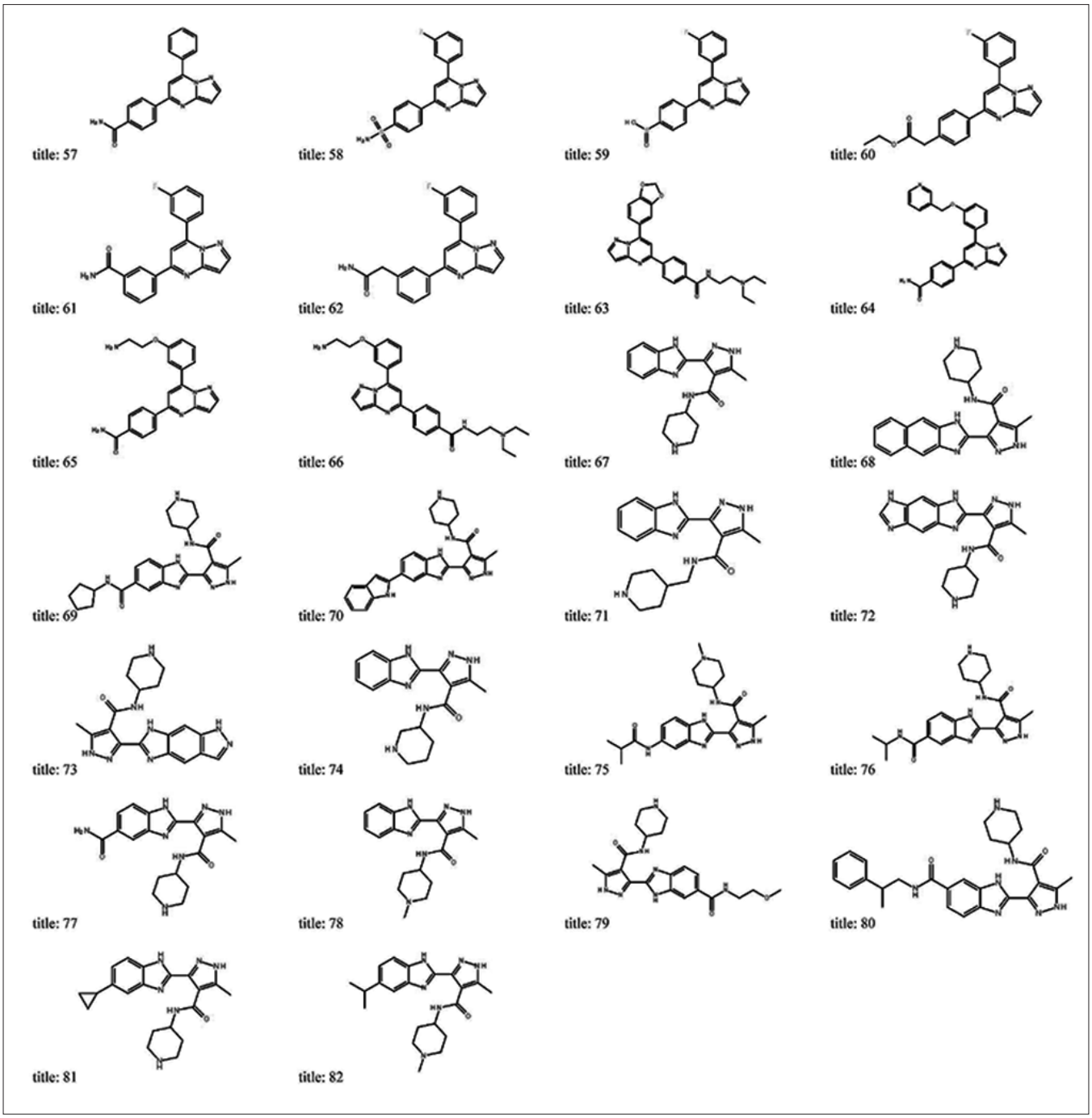

Fig. 3: 2D repesentation of the compounds used for QSAR studies.

calculated (-log IC_{50}, where IC_{50} is the concentration of compound required for 50% inhibition of PDK1 activity) for the QSAR analysis. The dataset has been classified into various different chemical classes i.e. bisindolylmaleimides, thieno[3,2-c] pyridine-7-carboxamides, N-phenylpyrimidin-2-amines, diazepinone derivatives, celecoxib derivatives and 7-aminopyrazolo[1,5-a]pyrimidines. The two dimensional structures of the 82 molecules were drawn using MarvinSketch version 5.2, 2009. The three dimensional conversion and minimization of ligands were performed using 'LigPrep' incorporated in PHASE. Conformers were generated using a rapid torsion angle search approach followed by minimization of each generated structure using the MMFF (Merck Molecular Force Field**)**, with an implicit GB/SA (Generalized-Born/Surface Area) solvent model. Ligand-based and receptor-guided 3D-QSAR models were developed using pharmacophore identification and molecular docking method, respectively.

Generation of common pharmacophore hypothesis: The common pharmacophore hypotheses were generated by PHASE, version 3.2, Schrodinger, LLC, New York, NY, 2010. The pharmacophoric features were identified from a set of variants and six pharmacophoric features were built: i.e. hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively charged group (N), positively charged group (P), aromatic ring (R). The PHASE was used to define the chemical features of ligand that facilitates the non covalent bonding between the ligand and its target receptor. Common pharmacophore hypotheses (CPH) were identified based on active analog approach, in which common pharmacophores were culled from the conformations of the set of highly active molecules using a tree based partitioning technique according to their inter site distances. These CPH's were examined by scoring function to get the best alignment of the active ligands using an overall maximum root mean square deviation (RMSD) value of 1.2Å with default options for distance tolerance.

The alignment was measured using a survival score, defined as: $S = WsiteSsite + WvecSvec + WvolSvol + WselSsel + W^{m}rew$, Where W, represents the weights and S as scores; Ssite represents alignment score, Svec represents vector score, Svol represents volume score and Ssel represents selectivity score and accounts for what fraction of molecules are likely to match the hypothesis regardless of their activity towards the receptor. Wsel has default value of 0.0 and Wsite, Wvec, Wvol and Wrew have default values of 1.0. The default values have been used for hypothesis generation. W^{m} rew represents reward weights defined by m–1, where m is the number of actives that match the hypothesis.

QSAR model building:
The data set were randomly divided into a training set of 65 molecules and a test set of 17 molecules by incorporating biological and chemical diversity to generate atom-based QSAR models. A rectangular grid was generated with the spacing of 1.0Å to encompass the space occupied by the aligned training set molecules. Each model contains five or more partial least squares (PLS) factors tend to fit the pIC_{50} values beyond their experimental uncertainty. Each of these models was validated using an external test set of 17 molecules that were not considered during model generation. The statistical parameters (R^2 = coefficient of determination, SD = standard deviation of regression) were calculated the overall significance of model and statistical significance. The parameters were used to evaluate the test set predictions were described as Q^2, root mean squared error and Pearson correlation coefficient (r) value. The 3D-QSAR models with the best predictive power were considered to be those that simultaneously met all these criteria.

External validation:
According to literature[18], a high R^2 value may be the necessary but not the sufficient condition. Even though a model may exhibit a good predictive ability based on the statistics for the test set, it is not always sure that the model will perform well on a new set of data. So, we performed external validation to test the predictive power of the QSAR model. Based on this theory, 3D-QSAR models are accepted if they satisfy all of the following conditions: $r^2>0.5$, $q^2>0.6$, $[(r^2-r_0^2)/r^2]<0.1$, $0.85\leq k\leq 1.15$ and $r_m^2>0.5$

The regression value (R) calculated by the following formula, $R^2 = [\sum(y_i-\bar{y}_o)(\tilde{y}_i-\bar{y}_p)]/\sqrt{\sum(y_i-\bar{y}_o)\sum(\tilde{y}_i-\bar{y}_p)^2}$, In these equations, $\bar{y}_o$ and $\bar{y}_p$ are the average values of the observed and predicted pIC_{50} values of the test set molecules. For the ideal QSAR model, the R^2 value should be close to one. Meanwhile the regression of y against $\tilde{y}$ through origin: $y_i^{ro} = k\tilde{y}_i$ should be characterized by k close to 1. Slope k is calculated as: $k = \frac{\sum y_i\tilde{y}_i}{\sum \tilde{y}_i^2}$

Another essential parameter r_m^2 was defined as follow: $r_m^2 = r^2(1-\sqrt{|r^2-r_o^2|})$. Where, the r^2 was the non-cross-validated correlation coefficient obtained from the PLS process, and the r_o^2 was calculated as follows: $r_o^2 = 1-\frac{\sum(\tilde{y}_i-y_i^{ro})^2}{\sum(\tilde{y}_i-\tilde{y}_p)^2}$. The y_i^{ro} was obtained by this formula[19], $y_i^{ro} = k\tilde{y}_i$

Applicability domain:
The test for applicability domain was performed using SIMCA-P 12.0 demo version[20]. The QSAR applicability domain (AD) is the structural or biological space, physic-chemical, knowledge or

information based on the training set of the model, and for which it is applicable to make predictions for new compounds. One can directly analyze properties of the multivariate descriptor, analyze distance (or similarity) metrics to investigate the AD of a training set. This may be attained by different means of feature selection and successive principal component analysis. The reliable prediction of a compound is unlikely when it is highly dissimilar to all compounds of modeling set. To avoid such an unjustified exploration of activity predictions we used the concept of AD[19,21].

Molecular docking:

Docking studies were carried out using Glide version 5.6[22]. The crystal structure of the human PDK1 complex was obtained from the RCSB PDB (PDB: 3IOP)[11]. The hydrogen atoms were added and unwanted water molecules were removed from the protein structure. The receptor binding site was defined as a Glide enclosing box in the centroid of the co-crystallized ligand molecule and the size was set to a default value of 26Å without any Hydrogen bonding constraint. The compounds selected for QSAR analysis have been evaluated using Glide extra precision mode (Glide XP).

Binding free energy calculation using Prime/MM-GBSA approach:

Prime/MM-GBSA was used to predict the free energy of binding for ligand molecules to receptor. The Glide docked poses were minimized by local optimization features from prime and the energies of complex were calculated using OPLS_2005 force field and GB/SA continuum solvent model. The binding free energy (ΔG_{bind}) was estimated using following equation:

$$\Delta G_{bind} = E_{R:L} - (E_R + E_L) = \Delta G_{solv} + \Delta G_{SA}$$

Where $E_{R:L}$ is energy of the protein ligand complex, $E_R + E_L$ is sum of the energies of the receptor and ligand separately, using the OPLS-2005 force field, ΔG_{solv} and ΔG_{SA} is the difference between GBSA solvation energy or surface area energy of complex and sum of the corresponding energies for the ligand and unliganded protein.

RESULT AND DISCUSSION

Pharmacophore models were generated six different hypotheses using a terminal box size of 1Å with different active molecules, belongs to bisindolylmaleimides, thieno[3,2-c]pyridine-7-carboxamides, N-phenylpyrimidin-2-amines, diazepinone derivatives, celecoxib derivatives and 7-aminopyrazolo[1,5-a]pyrimidines selected (figs. 1-3) using a tree based partition algorithm[23,24]. To find the common pharmacophore hypothesis, the dataset was divided in to active and inactive set. Molecules with pIC_{50} values higher than 7.00 were considered to be active, and those with pIC_{50} values less than 5.20 were considered to be inactive (Table 1), whereas those in-between were considered to be moderately active. Six hypotheses survived to the three different phases of PHASE scoring procedure (survival, survival inactive

TABLE 1: DATA'S USED FOR THE 3D-QSAR STUDIES

Compound no	Actual pIC_{50}	Predicted pIC_{50}	Compound No	Actual pIC_{50}	Predicted pIC_{50}
1[a]	5.04	5.05	42[a]	5.18	5.65
2[a]	4.85	4.83	43[b]	5.35	5.51
3[a]	5.39	5.15	44[a]	5.49	5.5
4[a]	6.00	5.41	45[a]	4.72	5.64
5[a]	6.43	6.5	46[a]	5.58	5.65
6[a]	6.12	6.11	47[a]	5.65	5.51
7[a]	8.22	8.36	48[a]	6.74	5.32
8[a]	6.45	6.68	49[a]	5.97	5.72
9[a]	6.79	6.71	50[a]	5.92	4.99
10[a]	6.28	6.3	51[b]	4.36	5.59
11[a]	7.52	7.66	52[a]	4.85	5.83
12[a]	7.58	7.67	53[a]	5.44	6.13
13[b]	7.95	6.75	54[a]	5.38	6.13
14[a]	6.09	6.4	55[b]	5.21	5.41
15[a]	6.36	6.42	56[a]	5.29	5.51
16[a]	6.28	6.34	57[a]	5.20	4.86
17[b]	6.39	6.35	58[a]	4.44	5.32
18[a]	6.15	6.32	59[a]	5.63	5.32
19[a]	5.20	6.29	60[a]	5.09	4.86
20[a]	6.95	5.52	61[a]	5.95	5.25
21[a]	6.72	6.74	62[a]	5.74	5.34
22[a]	6.72	6.47	63[b]	5.76	5.16
23[a]	6.82	6.39	64[a]	6.60	4.39
24[a]	6.82	6.76	65[a]	6.09	5.51
25[a]	6.79	6.77	66[a]	7.22	5.18
26[b]	6.65	6.43	67[b]	7.15	5.88
27[b]	5.60	6.32	68[a]	7.69	5.74
28[b]	6.29	6.21	69[a]	6.88	5.62
29[a]	4.46	6.45	70[a]	6.92	5.93
30[a]	3.95	4.25	71[a]	8.00	6.14
31[a]	3.48	3.86	72[a]	7.30	7.19
32[a]	3.82	3.44	73[a]	7.15	7.41
33[a]	3.61	4.01	74[a]	7.30	7.31
34[b]	4.31	4.37	75[b]	7.09	7.57
35[b]	5.04	4.53	76[a]	6.85	6.9
36[a]	5.30	4.91	77[a]	7.15	7.31
37[b]	5.69	5.1	78[b]	6.52	7.44
38[a]	5.60	5.7	79[a]	5.54	7.12
39[b]	5.46	5.51	80[a]	7.39	7.37
40[a]	5.89	5.50	81[a]	7.69	7.35
41[a]	5.58	5.64	82[b]	7.39	7.35

[a] - Training set compounds, [b] - Test set compounds

and post-hoc), and therefore these were used for the generation of QSAR models (Table 2).

The top model was found to be associated with the six-point hypotheses, which consists of two hydrogen bond acceptor (A), three hydrogen bond donors (D) and one hydrophobic group (H). This is denoted as A_1, A_2, D_3, D_4, D_5, H_6. The pharmacophore hypothesis shows distance and angles between pharmacophoric sites are depicted in fig. 4.

For the QSAR models generation, non-modeled (inactive or moderately active) molecules in the dataset were then aligned, based on matching with at least three pharmacophore features. The dataset were randomly divided into training set of 65 compounds and 17 compounds in the test set, in order to create the standard of 4:1 training set to test set ratio needed for QSAR study[25]. The best pharmacophore model resulted AADDDH.1670 (R^2 = 0.9557). Plots of predicted *vs.* actual pIC_{50} for training and test set are reported in fig. 5. Histogram of residuals for test set molecules have shown in fig. 6.

The pharmacophore map and QSAR contour maps can be used to design new and more active analogues. A descriptive representation of the contours generated in the QSAR is shown in fig. 7. The major advantages of 3D-QSAR techniques are the cubes generated using PLS regression which could be visualized in 3D space. The activity cubes can be generated for different properties such as hydrogen bond acceptor, hydrogen bond donor, hydrophobic, positive and negative ionic features, which define the noncovalent interactions with receptor. In these generated cubes, blue cubes indicate the favourable features and red cubes indicate the unfavorable features for the biological activity spectrum. The comparison of the most significant favourable and unfavorable interactions, which arise when the 3D-QSAR model was applied to the most active compound 7 and the least active compound 32, which is shown in figs. 7a-f.

Additional insights into the inhibitory activity can be gained by visualizing the 3D-QSAR model in the context of one or more ligands in the series

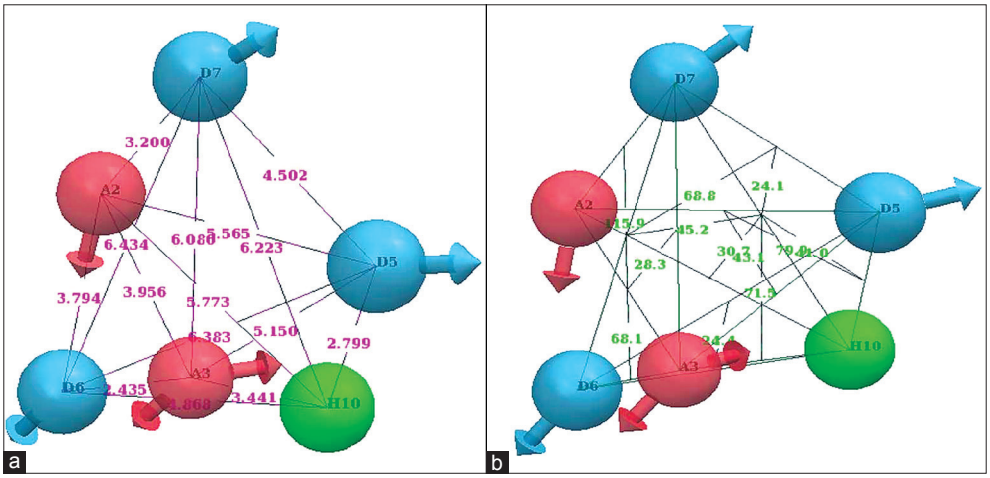

Fig. 4: Geometry of the pharmacophore.
Red spheres with vectors Acceptor feature, blue spheres positive ionic feature, green sphere hydrophobic feature. (a) CPH 1 features and distances. (b) CPH 1 features with angles

TABLE 2: SUMMARY OF PARTIAL LEAST-SQUARES ANALYSIS RESULTS FOR THE SIX BEST COMMON PHARMACOPHORE HYPOTHESES (CPHS)

Hypotheses	Factor	SD	R^2	F	*P*	RMSE	Q^2	Pearson-R
AADDDH	5	0.2334	0.9557	215.5	1.407e-32	0.5225	0.7510	0.8676
ADDHPR	5	0.2057	0.9644	314.2	1.175e-40	0.7081	0.6061	0.8527
AADDHR	5	0.2530	0.9472	211.9	2.397e-36	0.6843	0.6321	0.8383
ADDPRR	5	0.2005	0.9666	329.7	9.966e-41	0.7489	0.5594	0.7813
AAAHPR	5	0.1909	0.9705	381.0	5.298e-43	0.6484	0.6697	0.8425
ADHPRR	5	0.1641	0.9778	519.8	2.043e-47	0.6932	0.6225	0.8079

SD standard deviation of the regression, R^2 value of R^2 for the regression, F variance ratio, *P* significance level of variance ratio, RMSE root mean square error, Q^2 value of Q^2 for the predicted activities, Pearson R correlation between predicted and observed activity for the test set.

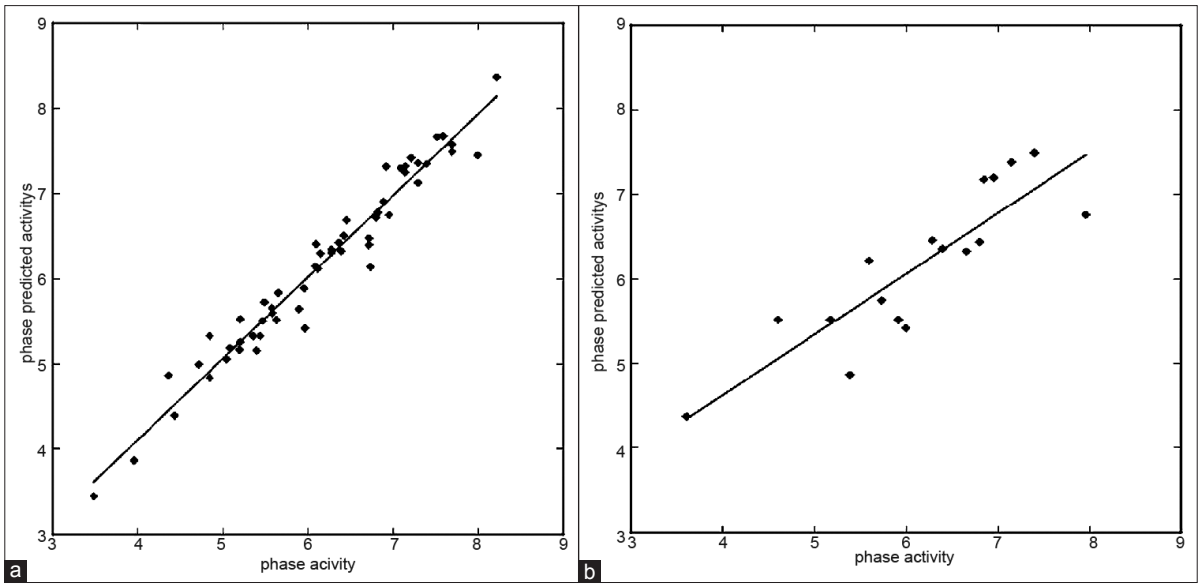

Fig. 5: Scatter plots for the predicted and experimental pIC_{50} values for the PDK1 QSAR model. (a) Training set (b) Test set

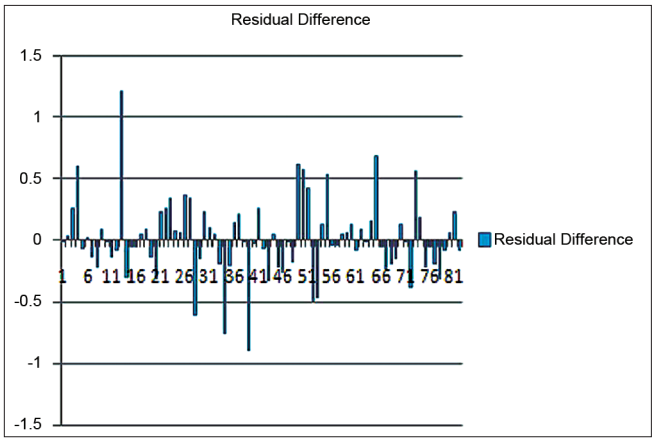

Fig. 6: Histogram of residuals for test set molecules

with different activity. In the generated cubes, blue cubes indicate the favourable features, whereas red cubes indicate the unfavorable features for biological activity. Among the all compounds, compound 7 has a planar 5H-indolo [2,3-a]pyrrolo[3,4-c]carbazol-5-one scaffold which is identical to the core of staurosprine and it showed much potency against PDK1. The pictorial representation of the most active compound and the least active compound are shown in fig. 7 (a and b, respectively). The electron-withdrawing is the most important criteria for the potency of the molecule. The electron-withdrawing group near hydrogen bond donor and hydrogen bond acceptor will improve the activity of the compound 7 (fig. 7c) whereas in the most inactive ligand there is no electron-withdrawing group (fig. 7d). The hydrophobic group on the most active compound

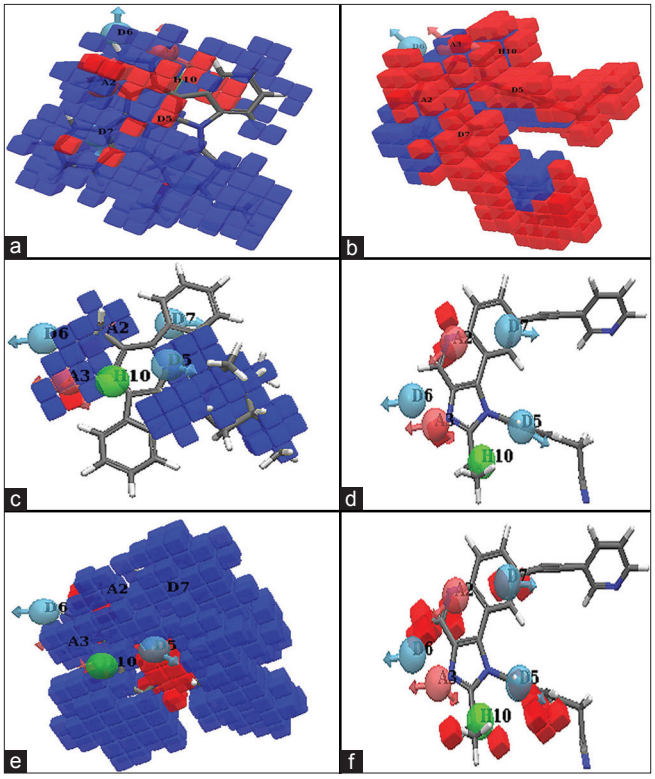

Fig. 7: Pictorial representation of the contours generated using the QSAR model. Blue cubes indicate favorable regions while red cubes indicate unfavorable region for the activity. (a) QSAR model visualized in the context of most active molecule 7. (b) QSAR model visualized in the context of least active molecule 32. (c) QSAR model visualized in the context of hydrogen bond acceptor property and electron withdrawing features with reference ligand 7. (d) QSAR model visualized in the context of hydrogen bond acceptor property and electron withdrawing features with ligand 32. (e) The significant favorable and unfavorable hydrophobic interactions that arise when the QSAR model is applied to reference ligand 7. (f) The significant favorable and unfavorable hydrophobic interactions that arise when the QSAR model is applied to the least active ligand 32

is shown in fig. 7e. The blue color cubes around the pharmacophore positions suggested that these substitutions are essential for the activity of the molecule. The red cubes show the absence of hydrophobic group in most inactive ligand (fig. 7f). This indicated that hydrophobic groups are important for activity of the molecule.

The established QSAR model was externally validated using 17 test set molecules, which gave an excellent value of 0.879 (>0.5) as well as high slope of regression lines through the origin (k) value of 1.00 ($0.85 \leq k \leq 1.15$), the correlation coefficient (R) values of 0.99 (close to 1) and the calculated $r^2(1-\sqrt{|r^2-r_o^2|})$ values of 0.0003 (<0.1) were obtained. The results of the external validation indicated that the QSAR models possessed a high accommodating capacity, they may be reliable for being used to predict the activities of new derivatives.

The validation of AD results indicates that the predictions of test set compounds are quite reliable.

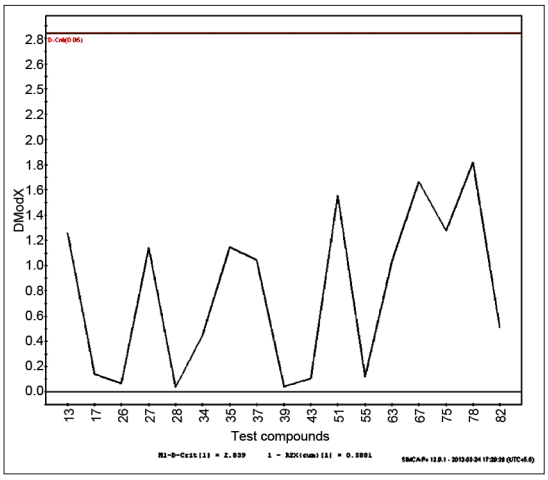

Fig. 8: The DModX values of the 17 test set compounds for prediction of AD

The DModX values of all 17 test set compounds are below the critical value of 2.8. Fig. 8 represents the residual SD of X-residuals (DModX) of the test set compounds for the best model.

We carried out docking studies for all molecules and we analyzed the most active molecule and least active molecule in the receptor ligand binding region. Fig. 9a shows docked model of best active compound (compound 7) within the active site of PDK1. As we explained earlier, the donors (amine group) presented in the northern region were well interacted with back bone of Ala162 (carboxy group) (NH…OC, 2.028 Å) and another two interactions by electron-withdrawing group in ligand molecule with Ala162 (2.028 Å) and Thr222 (2.101 Å). In the southern region the amino group made hydrophobic contact with two different amino acids Glu209 (NH…OC, 2.159 Å) and Glu166 (NH…OC, 1.673 Å). The best active ligand had the docking score of -13.94.

Fig. 9b displayed the docking mode of least active compound in the receptor site of protein molecule. The hydrogen molecule presented in the amino moiety well interacted with the carboxyl group of Glu209 (NH.OC) 1.518 Å. In the other side the cyano group (CN) has interacted with Lys111 (1.855 Å) (CN… HN). The other cyano group of ligand molecule interacted with oxygen atom in amine group in Ala162 (2.138 Å) (CN…HN).

The performance of the Glide dock program was also evaluated by comparing the docked pose of active and inactive ligand to co-crystallize ligand pose in the 3IOP crystal structure. The co-crystallized ligand was removed from the active site and again re-docked along with compound 7 and compound 32. The rmsd values for heavy atoms between the docked poses of compound 7, compound 32 and the native ligand

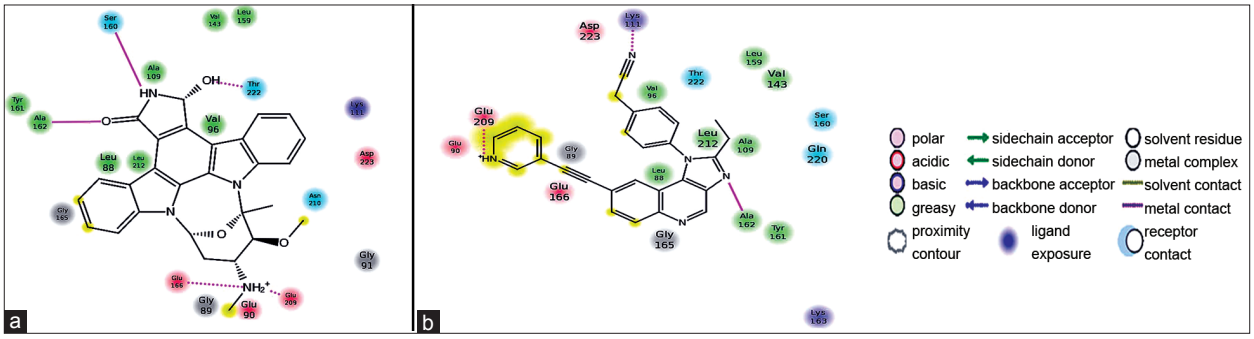

Fig. 9: 2D representation of protein ligand interactions.
(a) Most active compound with PDK1 (b) Least active compound with PDK1

show that all docked poses have rmsd values lower than 2.0 Å. These low rmsd values indicate that Glide dock is able to identifying the native poses in 3IOP crystal structure and can be reliably used to predict the binding free energy calculation for all ligands.

The calculation of ligand binding energies and ligand strain energies for a set of ligands and a single receptor, we used the MM-GBSA method. Here, the inhibitors taken for QSAR studies were docked into the active site of 3IOP protein. It shows all the molecules were binding in the catalytic amino acid residues and these docked files were further subjected for binding energy analysis. The binding free energy calculations were carried out with the protein-ligand complex, protein and ligand energies. The ΔG_{bind} values are ranges from ~ -10 to ~ -154 kcal/mol for all the 82 compounds. The most active compound having more energy (-89.58 kcal/mol) when compare with the least active compound (-30.56 kcal/mol). For all the compounds used for docking and binding free energy calculations are shown in Table 3.

TABLE 3: THE DOCKING RESULTS FOR THE 82 MOLECULES

Ligand	Glide score	ΔG_{bind}(kcal/mol)	HB*	Ligand	Glide score	ΔG_{bind} (kcal/mol)	HB*
1	-12.71	-58.45	4	42	-12.45	-71.88	5
2	-13.12	-60.57	5	43	-12.20	-58.99	5
3	-14.38	-66.28	6	44	-12.51	-87.64	5
4	-14.31	-61.78	5	45	-12.33	-72.83	4
5	-11.93	-104.86	3	46	-10.26	-97.43	3
6	-4.90	-42.96	2	47	-12.46	-73.38	5
7	**-13.93**	**-89.58**	**5**	48	-12.47	-73.30	4
8	-8.42	-83.70	4	49	-12.38	-71.25	5
9	-9.21	-73.00	4	50	-11.52	-62.37	4
10	-9.99	-18.24	2	51	-10.62	-75.08	5
11	-10.12	-85.97	4	52	-11.86	-71.49	5
12	-10.94	-76.84	5	53	-10.55	-63.46	5
13	-6.86	-19.06	1	54	-8.97	-22.23	3
14	-6.89	-89.75	1	55	-7.21	-14.63	2
15	-8.14	-80.66	2	56	-7.86	-82.04	3
16	-7.20	-71.29	3	57	-7.05	-12.49	2
17	-8.43	-86.40	2	58	-8.03	-15.20	1
18	-4.92	-75.17	2	59	-7.61	-45.47	4
19	-6.80	-60.20	3	60	-7.89	-12.66	2
20	-7.99	-72.15	2	61	-7.54	-13.59	2
21	-7.28	-73.21	1	62	-7.47	-11.97	2
22	-9.34	-87.78	3	63	-9.01	-72.98	2
23	-9.31	-87.54	1	64	-7.95	-54.41	2
24	-7.33	-76.20	2	65	-9.77	-72.96	3
25	-6.60	-69.81	2	66	-9.25	-154.19	4
26	-7.45	-88.29	2	67	-10.91	-105.60	4
27	-8.52	-33.10	3	68	-10.74	-92.11	4
28	-7.05	-81.34	1	69	-11.55	-89.03	6
29	-8.58	-85.82	3	70	-11.41	-98.18	5
30	-9.93	-51.97	2	71	-10.80	-101.44	4
31	-8.29	-40.38	3	72	-10.88	-193.36	4
32	-9.48	-30.56	3	73	-10.95	-79.78	4
33	-9.09	-35.80	3	74	-11.54	-104.23	4
34	-8.30	-37.22	3	75	-11.06	-94.92	4
35	-7.94	-10.90	2	76	-11.83	-97.77	5
36	-7.47	-21.00	3	77	-11.86	-93.87	5
37	-9.15	-49.05	4	78	-10.07	-90.15	3
38	-8.24	-69.96	3	79	-11.65	-90.09	6
39	-8.12	-14.48	2	80	-11.14	-90.08	6
40	-11.40	-73.58	5	81	-10.87	-92.46	4
41	-12.14	-70.64	6	82	-10.42	-102.03	3

*Number of hydrogen bonds formed. The best and least active compounds are shown in boldface

The combined ligand-based pharmacophore and atom-based 3D-QSAR models were generated using a training set of 65 molecules which consists of six-point pharmacophore hypothesis has led to the development of satisfactory models for predicting the biological activity of new compounds. Understanding the intermolecular interactions of various derivatives with PDK1 was achieved by integrating statistically significant and predictable models. The atom-based 3D-QSAR visualization model in the context of the structure of molecules understudy provides details of the relationship between structure and activity. This will enhance in-depth information related with modifications of some structural side chains which will help in the designing of new analogs with better activity before synthesis. Molecular docking studies also support by validating the Glide program which shows that binding orientation of these 82 inhibitors are correlating with the native ligand present in the PDK1 protein. These reveal that newly synthesizing inhibitory molecules not only possess higher activity (pIC_{50}) and also fit the binding pocket of PDK1 protein. And last, the binding free energy was calculated from the Prime MM-GBSA method to revel the potency of the electrostatic interactions between protein ligand complexes.

ACKNOWLEDGMENTS

The authors like to thank the Alagappa University, Karaikudi, Tamil nadu, India for its support and providing the facilities for this work. P. Kirubakaran is grateful to Council of Scientific and Industrial Research (CSIR), New Delhi, India for supporting as a Senior Research Fellowships (SRF).

REFERENCES

1. Lawlor MA, Mora A, Ashby PR, Williams MR, Murray-Tait V, Malone L, *et al.* Essential role of PDK1 in regulating cell size and development in mice. EMBO J 2002;14:3728-38.
2. Pullen N, Dennis PB, Andjelkovic M, Dufner A, Kozma SC, Hemmings BA, *et al.* Phosphorylation and Activation of p70s6k by PDK1. Science 1998;5351:707-10.
3. Mora A, Komander D, van Aalten D, Alessi D. PDK1, the master regulator of AGC kinase signal transduction. Sem Cell Dev Biol 2004;2:161-70.
4. Vanhaesebroeck B, Leevers SJ, Ahmadi K, Timms J, Katso R, Driscoll PC, *et al.* Synthesis and function of 3-Phosphorylated Inositol lipids. Annu Rev Biochem 2001;1:535-602.
5. Osaki M, Oshimura M, Ito H. PI3K-Akt pathway: Its functions and alterations in human cancer. Apoptosis 2004;6:667-76.
6. Hanada M, Feng J, Hemmings BA. Structure, regulation and function of PKB/AKT-a major therapeutic target. Biochim Biophys Acta 2004;1-2:3-16.
7. Zhang Q, Thomas SM, Lui VW, Xi S, Siegfried JM, Fan H, *et al.* Phosphorylation of TNF- α converting enzyme by gastrin-releasing peptide induces amphiregulin release and EGF receptor activation. PNAS 2006;18:6901-6.
8. Baryawno N, Sveinbjornsson B, Eksborg S, Chen CS, Kogner P, Johnsen JI. Small-Molecule Inhibitors of Phosphatidylinositol 3-Kinase/Akt Signaling Inhibit Wnt/beta-Catenin Pathway Cross-Talk and Suppress Medulloblastoma Growth. Cancer Res 2010;1:266-76.
9. Zeng X, Xu H, Glazer RI. Transformation of Mammary Epithelial Cells by 3-Phosphoinositide-dependent Protein Kinase-1 (PDK1) Is Associated with the Induction of Protein Kinase C Alpha. Cancer Res 2002;12:3538-43.
10. Flynn P, Wong M, Zavar M, Dean NM, Stokoe D. Inhibition of PDK-1 activity causes a reduction in cell proliferation and survival. Curr Biol 2000;22:1439-42.
11. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, *et al*. The Protein Data Bank. Nucleic Acids Res 2000;1:235-42.
12. PDB codes: 2PE1, 1H1W, 1UVR, 1UU9, 1UU8, 1UU7, 1UU3, 1OKZ, 1OKY, 1Z5M, 3H9O, 3ION, 3IOP.
13. Peifer C, Alessi DR. Small-Molecule Inhibitors of PDK1. Chem Med Chem 2008;12:1810-38.
14. Dixon S, Smondyrev A, Knoll E, Rao S, Shaw D, Friesner R. PHASE: A new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. J Comput Aided Mol Des 2006;10:647-71.
15. Evans DA, Doman TN, Thorner DA, Bodkin MJ. 3D QSAR Methods: Phase and Catalyst Compared. J Chem Inf Model 2007;3:1248-57.
16. Bagchi MC, Maiti BC. On application of atom pairs in drug design. J Mol Struct 2003;1-3:31-7.
17. Nandi S, Vracko M, Bagchi MC. Anticancer Activity of Selected Phenolic Compounds: QSAR Studies Using Ridge Regression and Neural Networks. Chem Biol Drug Des 2007;5:424-36.
18. Golbraikh A, Tropsha A. Beware of q2! J Mol Graph Model 2002;4:269-76.
19. Kar S, Roy K. Development and validation of a robust QSAR model for prediction of carcinogenicity of drugs. Indian J Biochem Biophys 2011;48:111-22.
20. UMETRICS SIMCA-P 12.0, info@umetrics.com: www.umetrics.com, Umea, Sweden, 2002.
21. Mitra I, Saha A, Roy K. Development of multiple QSAR models for consensus predictions and unified mechanistic interpretations of the free-radical scavenging activities of chromone derivatives. J Mol Model 2012;18:1819-40.
22. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, *et al*. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. J Med Chem 2004;7:1739-49.
23. Wang S, Folkes A, Chuckowree I, Cockcroft X, Sohal S, Miller W, *et al*. Studies on Pyrrolopyrimidines as Selective Inhibitors of Multidrug-Resistance-Associated Protein in Multidrug Resistance. J Med Chem 2004;6:1329-38.
24. Wang S, Wan NC, Harrison J, Miller W, Chuckowree I, Sohal S, *et al*. Design and Synthesis of New Templates Derived from Pyrrolopyrimidine as Selective Multidrug-Resistance-Associated Protein Inhibitors in Multidrug Resistance. J Med Chem 2004;6:1339-50.
25. Nagamani S, Karthikeyan M, Kirubakaran P, Singh KD, Gopinath K. Theoretical studies on benzimidazole derivatives as *E. coli* biotin carboxylase inhibitors. Med Chem Res 2012;21:2169-80.

Accepted April 04, 2012
Revised March 26, 2012
Received April 28, 2011
Indian J. Pharm. Sci., 2012, 74 (2): 141-151