Research Article

Combined 3D QSAR Based Virtual Screening and Molecular Docking Study of Some Selected PDK-1 Kinase Inhibitors

Shalini Singh and Pradeep Srivastava

School of Biochemical Engineering, Indian Institute of Technology (Banaras Hindu University), Varanasi 221005, India

Correspondence should be addressed to Pradeep Srivastava; drpradeep19@gmail.com

Received 30 September 2013; Revised 21 December 2013; Accepted 6 January 2014; Published 1 June 2014

Academic Editor: Jeon-Hor Chen

Copyright © 2014 S. Singh and P. Srivastava. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Phosphoinositide-dependent kinase-1 (PDK-1) is an important therapeutic target for the treatment of cancer. In order to identify the important chemical features of PDK-1 inhibitors, a 3D QSAR pharmacophore model was developed based on 21 available PDK-1 inhibitors. The best pharmacophore model (Hypo1) exhibits all the important chemical features required for PDK-1 inhibitors. The correlation coefficient, root mean square deviation (RMSD), and cost difference were 0.96906, 1.0719, and 168.13, respectively, suggesting a good predictive ability of the model (Hypo1) among all the ten pharmacophore models that were analyzed. The best pharmacophore model (Hypo1) was further validated by Fisher's randomization method (95%), test set method ($r = 0.87$), and the decoy set with the goodness of fit (0.73). Further, this validated pharmacophore model Hypo1 was used as a 3D query to screen the molecules from databases like NCI database and Maybridge. The resultant hit compounds were subsequently subjected to filtration by Lipinski's rule of five as well as the ADMET study. Docking study was done to refine the retrieved hits and as a result to reduce the rate of false positive. Best hits will further be subjected to in vitro study in future.

1. Introduction

Protein kinases are critical components of cellular signal transduction cascades [1]. Over 500 protein kinases in the human genome have been reported till date and they are considered as the second largest group of drug targets [2, 3]. Phosphoinositide-dependent kinase-1 (PDK-1), a 63 kDa serine/threonine kinase, is a major player in the PI3-kinase signaling pathway that regulates gene expression, cell cycle, growth, and proliferation [4–11]. PDK-1 is also termed as the “master kinase” because it phosphorylates highly conserved serine or threonine residues in the T-loop (or activation loop) of numerous AGC kinases, including PKB/AKT, PKC, p70S6K, SGK, and PDK-1 itself [12]. Although precise regulatory mechanisms vary, in the case of PKB/AKT, activation by PDK-1 is critically dependent on prior PI3 kinase activation and the presence of phosphatidylinositol-(3,4,5)-triphosphate (PIP3). A significant proportion (40–50%) of all tumors involve mutations in PIP3-3-phosphatase (PTEN) [13–15], which result in elevated levels of PIP3 and enhanced activation of PKB/AKT, p70S6K, and SGK. The inhibitors of

PDK-1 could potentially provide valuable therapeutic agents for the treatment of cancer.

Recognition process between ligand and model is based on spatial distribution of certain structural features of active site being complimentary to those of the interacting ligands, and the features common to the ligands would provide the information about the active site. A pharmacophore mapping is the essential step towards understanding of receptor-ligand recognition process and is established as one of the successful computational tools in rational drug design [16, 17]. This involves the identification of a three-dimensional arrangement of functional groups which a molecule must possess to be recognized by the receptor. Further, a model is generated by finding chemically important functional groups that are common to the molecules that bind. Pharmacophore can be derived by direct analysis of the structure of known ligand either in the most stable conformer or in the form observed when complexed with the target protein.

In the present study, a three-dimensional pharmacophore model for PDK-1 kinase inhibitors has been developed. The generated model is further utilized for screening of

potentially active candidates from NCI and Maybridge database. The efficacy of these compounds is further validated by molecular docking.

2. Material and Methods

2.1. General Methodology. All pharmacophore models generation and Hypo1 based virtual screening were performed using the following tools.

HypoGen. It was implemented in Catalyst (Catalyst 4.1, Molecular Simulations Inc., San Diego, CA).

Fisher Randomization Test. It was done by CatScamble program implemented in Catalyst.

Lipinski Filtration. It was performed using Pipeline Pilot Studio (SciTegic Inc., San Diego, CA).

Ligandfit. Docking studies were achieved using Discovery Studio 2.5 (Accelrys Inc., San Diego, CA).

2.2. Data Set for Pharmacophore Analysis. A set of 83 different compounds were collected from different references [18–21], which have been identified and reported to be inhibitors of PDK-1 kinase. The inhibitory activity of these compounds, expressed as IC_{50} (i.e., concentration of compound required to inhibit 50% of PDK-1 kinase activity), was studied for all compounds. The IC_{50} values spanned across a wide range from 3.0 to 65,000 nM. Amongst 83 compounds, 21 compounds were selected as training set compounds and the rest of compounds were taken as a test set compounds. The chemical structures of all training set compounds are shown in Figure 1. The selection of the training set and test set were according to the following rules: (i) there is structural diversity among molecules, (ii) both training set and test set cover a wide range of activity, and (iii) the highest active compounds were included in the training set because they provide critical information for pharmacophore generation. The geometry of all compounds was built by using Accelrys Discovery Studio 2.5 [22]. All the compounds were minimized using the steepest descent algorithm with a convergence gradient value of 0.001 kcal/mol and a family of representative conformations was generated by fast conformational analysis methods using poling minimize algorithm [23] and CHARMM force field parameters [24]. A large number of confirmations of each compound were generated within an energy threshold of 20.0 kcal/mol above the global energy minimum.

2.3. Pharmacophore Modeling. Based on the conformations for each compound, HypoGen module of Discovery Studio 2.5 was used to construct the possible pharmacophore models [25]. Instead of using the lowest energy conformation of each compound, all the conformational models of each compound in the training set were used in Discovery Studio 2.5 for pharmacophore hypothesis generations. The training set compounds (21 in number) associated with their conformations were submitted to Discovery Studio 2.5

for 3D QSAR pharmacophore Generation (HypoGen). The HypoGen module generated hypothesis with features common in active molecule and missing from inactive molecule.

2.4. Model Validation. The statistical parameters, such as the cost value, determine the significance of the model. The best model was selected on the basis of significant statistical parameters, like high correlation (r), predicted the lowest total cost, and lower value of RMSD, and the value of the total cost should be closer to the fixed cost and far away from null cost. Another parameter, configuration cost, is also important for the determination of significance of the model, and it should be <17 . The best hypothesis Hypo1 was also validated by test set validation method, Fischer's randomization validation, and decoy set method. Ligand pharmacophore mapping protocol was used for estimating the activity of all 62 test set compounds.

2.5. Decoy Set Validation. Results of test set validation method could only indicate that the generated pharmacophore model (Hypo1) has high efficiency in picking the active molecules but is not conformity as it also picked the inactive molecules. To further evaluate this, decoy set validation method was used to evaluate the efficiency of Hypo1 by calculating the GH (goodness of hit list) and EF (enrichment factor). A data set of small molecules was generated by decoy set finder 1.1 which included 1980 molecules with unknown activity and 20 active molecules were making a decoy set of 2,000 molecules. GH (goodness of hit list) and EF (enrichment factor) were calculated by the following equations:

$$EF = \frac{(Ha/Ht)}{(A/D)}, \quad (1)$$

$$GH = \left\{ \frac{[Ha * (3A + Ht)]}{4HtA} \right\} * \left[1 - \frac{(Ht - Ha)}{(D - A)} \right],$$

where Ht = total number of molecules in hit list, Ha = total active molecules present in the hit list, A = total active molecules present in database, and D = total molecules present in decoy set.

The range of GH score varies from 0 to 1. GH score 0 means a null model, while the GH score 1 means generation of an ideal model.

Although when the GH score is higher than 0.7, it reflect the generation of a very good model. The EF and GH were found to be 69.23 and 0.73 (shown in Table 1) indicating that the generated pharmacophore model had a rationale for virtual screening.

2.6. Virtual Screening and ADMET Analysis. The final validated hypothesis (Hypo1) was used as a 3D structural query for retrieving potent compounds from NCI database and Maybridge database having 23,8819 molecules and 2,000 molecules, respectively. A systematic diagram of virtual screening protocol is shown in Figure 2.

(a)

FIGURE 1: Continued.

FIGURE 1: Chemical structures of and activity data (IC_{50} values, nM) of 21 training set molecules applied for HypoGen pharmacophore generation.

FIGURE 2: Diagrammatic representation of virtual screening protocol.

TABLE 1: Statistical parameters from the validation of the pharmacophore model by mean of decoy set.

Number	Parameters	Values
1	Total number of molecules in database (D)	2,000
2	Total number of active molecules in database (A)	20
3	Total number of hit molecules from database (Ht)	26
4	Total number of active molecules in hit list (Ha)	18
5	% yield of actives $[(Ha/A) \times 100]$	69.23
6	% ratio of actives in the hit list $[(Ha/A) \times 100]$	90
7	Enrichment factor (EF) $[(Ha \times D)/(Ht \times A)]$	69.23
8	False negatives [$A - Ha$]	2
9	False positives [$Ht - Ha$]	8
10	GH score ^a (goodness of hit list)	0.73

^a $[(Ha/4HtA)(3A + Ht) \times [1 - (Ht - Ha)/(D - A)]]$.

3. Results and Discussion

3.1. Pharmacophore Modeling. Ten hypotheses were produced by 3D QSAR pharmacophore generation module of Accelrys Discovery Studio 2.5 through 21 training sets compounds (Table 2). Hypo1 was the most significant hypothesis characterized by high cost difference (168.48433), the lowest root mean square deviation ($RMSD = 1.0719$), and the best correlation coefficient ($r = 0.96906$). The fixed cost and the null cost values were 77.5618 and 258.686, respectively, with total cost value 90.2017 for Hypo1. This observation was much lower than null cost and closer to the fixed cost.

The best hypothesis (Hypo1) consists of four features, that is, two hydrogen bonds acceptor (HBA), one hydrogen bond donor (HBD), and one hydrophobic aliphatic feature (HyA). Figures 3(a) and 3(b) represent features of the best pharmacophore (Hypo1) and the distance and angular

TABLE 2: Information of statistical significance and predictive power presented in cost values measured in bits for top 10 hypotheses as a result of automated HypoGen pharmacophore generation process.

Hypothesis number	Total cost	Cost difference ^a (total cost-null cost)	Error	RMS	Correlation (<i>r</i>)	Features ^b
1	90.2017	168.4843	73.0296	1.0719	0.96906	HBA, HBA, HBD, and HyA
2	91.6361	167.0499	74.8623	1.1505	0.96423	HBA, HBA, HyA, and HyA
3	92.4311	166.2549	74.9287	1.1532	0.96414	HBA, HBA, HBD, and HyA
4	93.3537	165.3323	76.4684	1.2151	0.96002	HBA, HBA, HBD, and HyA
5	93.5794	165.1066	76.5881	1.2198	0.95971	HBA, HBA, HyA, and HyA
6	95.5213	163.1647	78.8279	1.3043	0.95376	HBA, HBA, HyA, and HyA
7	96.8603	161.8257	79.8688	1.34181	0.951035	HBA, HBD, HyA, and HyA
8	99.4115	159.4565	79.8688	1.34181	0.951035	HBA, HBD, HyA, and HyA
9	99.7217	158.9643	83.011	1.44906	0.942609	HBA, HBD, HyA, and HyA
10	99.7794	158.9066	83.1717	1.45431	0.942169	HBA, HBA, HBD, and HyA

^aThe cost difference between null cost and total cost; null cost is 258.686 bits; fixed cost is 77.5618 bits; configuration cost is 15.4729 bits.

^bAbbreviation used for features: HBA: H-bond acceptor; HBD: H-bond donor; HyA: hydrophobic aliphatic.

FIGURE 3: (a) Hypo1 (best pharmacophore) generated by HypoGen (3D QSAR pharmacophore protocol); (b) pharmacophore model with distance between chemical features.

constraints in the best pharmacophore (Hypo1). The experimental and estimated activities of the best pharmacophore hypothesis (Hypo1) for 21 training set compounds are shown in (Table 3). Figure 4(a) represents the top scoring hypothesis Hypo1, mapped on the most active compound 1 ($IC_{50} = 3$ nM) and Figure 4(b) represents the mapping of least active compound_21 ($IC_{50} = 65,000$ nM) of the training set.

3.2. Cost Analysis. In addition to generating a hypothesis, HypoGen also provides two theoretical costs (represented in bit units) to help assess the validity of the hypothesis. The first is fixed cost (cost of an ideal hypothesis), which represents the simplest model that fits all data perfectly, and the second is the null cost (cost of null hypothesis), which represents the highest cost of a pharmacophore with no features and which estimates activity to be the average of the activity data of the training set molecules. They represent the upper and lower limits for the hypothesis that

are generated. A meaningful pharmacophore hypothesis may be generated when the difference between null hypothesis and the fixed hypothesis is large; a value of 40–60 bits may indicate that it has 75–90% probability of correlating the data. Other two parameters that also determine the quality of any pharmacophore are configuration cost or entropy cost and error cost. The configuration cost depends on the complexity of the pharmacophore and should have value <17 whereas the error cost is dependent on the root mean square difference between the estimated and the actual activity of the training set. The difference between total fixed cost and the null cost of the Hypo1 was observed to be 168.4843, which is more than 40–60, which depicts more than 90% probability of data correlation. Noticeably, the total cost of Hypo1 was much closer to the fixed cost than to the null cost. Furthermore, a high correlation coefficient of 0.96906 was observed with RMS value of 1.0719 and the configuration cost of 15.4729, demonstrating the development of a reliable pharmacophore model with high predictivity.

TABLE 3: Experimental biological data and estimated IC_{50} of training set molecules based on pharmacophore model Hypo1.

Compound no	IC_{50} value (nM)		Error	Uncertainty	Fit value
	Experimental	Expected			
1	3	2.4	-1.3	2	8.60
2	5	4.9	-1.0	2	8.29
3	9	14	+1.6	2	7.83
4	24	20	-1.2	2	7.67
5	40	130	+3.3	2	6.86
6	67	39	-1.8	2	7.40
7	91	110	+1.2	2	6.94
8	110	290	+2.6	2	6.52
9	150	230	+1.6	2	6.61
10	260	1000	+4.0	2	5.97
11	380	160	-2.4	2	6.71
12	530	1100	+2.0	2	5.94
13	670	1000	+1.5	2	5.97
14	690	420	-2.3	2	6.36
15	1200	390	-3.1	2	6.39
16	7500	27000	+2.7	2	5.55
17	16000	21000	+1.6	2	4.65
18	24000	21000	-1.1	2	4.65
19	38000	21000	-1.8	2	4.65
20	52000	93000	+1.8	2	4.01
21	65000	21000	-3.0	2	4.05

FIGURE 4: (a) The highest active compound (compound_1, $IC_{50} = 3.0$ nM) mapped on the best pharmacophore model; (b) the least active compound (Compound_21, $IC_{50} = 65,000$ nM) mapped on the best pharmacophore model (Hypo1). The most active compound exhibits a good fit with all features of the pharmacophore hypothesis, Hypo1, whereas in the least active compound it had hydrogen bond acceptor feature missing. Based on this, it may be concluded that two HBA features are important for PDK-1 kinase inhibitory activity.

3.3. Validation of Pharmacophore Model

3.3.1. Test Set Validation. The test set method is for examining whether the pharmacophore model is capable of predicting the activities of external compounds of the test set series. The test set contains 62 compounds structurally different

from the training set molecules. All the test set molecules were prepared in the same way as that for the training set molecules. Test set validation was done using ligand pharmacophore mapping protocol. The test set of 62 compounds were mapped on the Hypo1. It was observed that pharmacophore model performed well in estimation of activity of test set

FIGURE 5: Regression plot of 62 test set molecules against Hypo1.

compounds, with a significant predictive correlation value ($r = 0.87$) between experimental and estimated activities (shown in Figure 5). The experimental and estimated activities of test set compounds mapped on the best hypothesis (Hypo1) were shown in (Table 4).

Further, another validation method was used to characterize the quality of the hypothesis using error ratio, which is the difference between estimated activity and experimental activity. Also an error ratio ≤ 10 depicts that there is no more than one order difference between estimated and experimental activity values, not more than one order. The best hypothesis (Hypo1) exhibited an error value ≤ 10 for 53 compounds out of 62 compounds. Only 9 compounds (compound_29, compound_32, compound_34, compound_38, compound_40, compound_51, compound_52, compound_53, and compound_55) with values > 10 were considered as outliers and rejected. The most potent compound_22 of the test set ($IC_{50} = 6$ nM) was mapped with Hypo1 (Figure 6). The best hypothesis (Hypo1) mapped very well, also all the chemical features of this compound matched and the estimated activity of this compound had an IC_{50} value of 1.3 nM. Based on these results, it was confirmed that one HBD, two HBA, and one HyA (hydrophobic aliphatic) features are essential for PDK-1 inhibitory activity.

3.3.2. Fisher's Validation. Fischer's randomization test method was used to evaluate the statistical relevance of Hypo1 by using the CatScramble program. The confidence level was fixed at 95%. The CatScramble program generated 19 random spreadsheets to construct hypothesis using exactly the same conditions as used in generating the original pharmacophore hypothesis. Total cost of 19 pharmacophore hypothesis generated randomly and the original pharmacophore hypothesis are also presented in Figure 7. It is observed that an original hypothesis (Hypo1) was far more superior to the 19 random hypotheses,

FIGURE 6: Most active test set compound (compound_22) mapped on the best pharmacophore model (Hypo1).

FIGURE 7: The difference in costs between the HypoGen runs and scrambled runs. The 95% confidence level was selected.

suggesting that Hypo1 is not generated by any chance event. These results have provided 95% confidence of the proposed hypothesis.

3.4. Pharmacophore Based Virtual Screening. The validated 3D QSAR pharmacophore model Hypo1 was used as a 3D structural query for retrieving potent compounds from NCI database and Maybridge database having 23,8819 molecules and 2,000 molecules, respectively. A total of 8,833 compounds

TABLE 4: Experimental biological data and estimated IC_{50} of test set molecules based on pharmacophore model Hypo1.

Compound no	IC_{50} (nM)		Error	Uncertainty	Fit value
	Experimental	Expected			
Compound_22	6	1.3	-4.6	2	8.87
Compound_23	47	5.2	-9.0	2	8.26
Compound_24	14	6.8	-2.0	2	8.14
Compound_25	6	11.9	+1.8	2	7.89
Compound_26	6	13.3	+2.2	2	7.85
Compound_27	8	15.9	+1.9	2	7.78
Compound_28	18	20.6	+1.4	2	7.66
Compound_29	3	46.8	+15.6	2	7.30
Compound_30	283	49.0	-5.8	2	7.29
Compound_31	17	62.2	+3.6	2	7.18
Compound_32	6	63.9	+10.5	2	7.17
Compound_33	13	71.0	+5.5	2	7.13
Compound_34	8	121.3	+15.2	2	6.89
Compound_35	144	159.2	+1.1	2	6.79
Compound_36	142	163.7	+1.1	2	6.76
Compound_37	57	166.9	+2.9	2	6.75
Compound_38	8	169.0	+21.1	2	6.75
Compound_39	65	173.1	+2.7	2	6.74
Compound_40	6,923	173.4	-39.9	2	6.74
Compound_41	35	173.7	+4.9	2	6.74
Compound_42	51	174.9	+3.4	2	6.73
Compound_43	43	176.8	+4.1	2	6.73
Compound_44	30	238.0	+7.9	2	6.60
Compound_45	70	264.1	+3.8	2	6.56
Compound_46	74	308.3	+4.2	2	6.49
Compound_47	2,033	332.1	+6.1	2	6.46
Compound_48	270	604.6	+2.2	2	6.20
Compound_49	200	604.6	+3.0	2	6.19
Compound_50	2,120	1,063.9	-1.9	2	5.95
Compound_51	60	1,070.4	+17.8	2	5.95
Compound_52	66	1,084.5	+16.4	2	5.94
Compound_53	290	3,107.8	+10.7	2	5.48
Compound_54	1,864	4,488.0	+2.4	2	5.33
Compound_55	580	6,336.0	+10.9	2	5.18
Compound_56	2,490	9,421.4	+3.8	2	5.00
Compound_57	15,000	19,013.5	+1.3	2	4.70
Compound_58	32,000	21,429.9	-1.5	2	4.65
Compound_59	38,000	21,429.8	-1.8	2	4.65
Compound_60	18,000	21,429.7	-1.2	2	4.65
Compound_61	32,000	21,430.2	-1.5	2	4.65
Compound_62	42,000	21,430.7	-1.9	2	4.65
Compound_63	32,000	21,430.9	-1.5	2	4.65
Compound_64	23,000	21,433.3	-1.1	2	4.65
Compound_65	15,000	21,434.2	-1.4	2	4.65
Compound_66	18,000	21,435.1	+1.2	2	4.65
Compound_67	22,000	21,436.7	-1.0	2	4.65
Compound_68	24,000	21,444.1	-1.1	2	4.65
Compound_69	21,000	21,446.6	+1.0	2	4.65

TABLE 4: Continued.

Compound no	IC ₅₀ (nM)		Error	Uncertainty	Fit value
	Experimental	Expected			
Compound_70	12,000	21,449.9	+1.8	2	4.65
Compound_71	42,000	21,453.4	+1.3	2	4.65
Compound_72	17,000	21,454.8	+1.3	2	4.65
Compound_73	16,000	21,493.5	+1.3	2	4.64
Compound_74	34,000	21,538.8	+1.6	2	4.64
Compound_75	18,000	21,619.6	+1.2	2	4.64
Compound_76	12,000	24,152.4	+2.0	2	4.59
Compound_77	40,000	24,273.3	-1.6	2	4.59
Compound_78	14,000	25,463.7	+1.8	2	4.57
Compound_79	45,000	30,669.3	-1.5	2	4.49
Compound_80	16,000	39,286.2	+2.4	2	4.38
Compound_81	40,000	46,181.3	+1.1	2	4.31
Compound_82	25,000	48,798.6	+1.9	2	4.29
Compound_83	34,000	80,163.1	+2.3	2	4.07

exhibited good mapping with Hypo1 using fast and flexible search method. Out of total 8,833 compounds, 8,530 compounds were from NCI and 333 compounds were from Maybridge database. Out of these 8,833 molecules, 2033 molecules having their IC₅₀ < 1 μ M were selected for further studies. Further sorting of these hits has been done by Lipinski's rule of five, to evaluate their drug similarity. Total 1,613 molecules passed this evaluative process. These 1,613 molecules were further evaluated for the ADMET studies. Only 842 molecules passed the ADMET filtration process. Those molecules were selected for further molecular docking studies, which exhibited estimated activity $\leq 0.5 \mu$ M. It was observed that only 43 molecules satisfied these conditions and further molecular docking study was prepared for these molecules.

3.5. Molecular Docking Studies. Further studies were conducted for selected compounds (retrieved hits) and evaluated the binding mode between compounds and protein. All the compounds and compound_1 were docked into the binding site of PDK-1 [26] (PDB entry: 1UU7) [27] by using LigandFit [28] docking method implemented in Discovery Studio 2.5 program package. Before docking all molecules, compound_1 (most active compound of the training set) was docked into the active site of PDK-1.

(a) Compound_1 Docking Description. Compound_1 has shown the docking energy of 64.5 kcal/mol and RMSD value of 0.841. This depicts that LigandFit docking method reproduced the original binding mode. Hence, for further docking of the LigandFit, docking method was used. It showed hydrogen bond interactions with important residues of amino acids, Lys111, Asp 230, Ala162, and Tyr 161 as shown in Figure 8(a).

(b) Other Compounds Docking Description. All 43 molecules (selected) were docked to the active sites of PDK-1 kinase, and, furthermore, only top 7 molecules, having high

docking energy, better hydrogen bond interactions with active site residues, and lower estimated activity ($\leq 0.19 \mu$ M) were selected. The estimated activity, interaction energy, and LignadFit scores of all seven compounds along with compounds_1 are listed in Table 5. Finally, the three compounds which were selected for further analysis are NSC.218341, NSC.24871, and NSC.21193. Compound NSC.211930, NSC.218341, and NSC.24871 mapped to all features of the Hypo1. Compound NSC.211930 formed the hydrogen bonding with Ala162 a hinge region amino acid. While the amide group formed the hydrogen bond with Asp223, Lys111 was involved in the cation-pi interaction. Compound NSC.24871 formed the hydrogen bond interaction with Lys111, Ser160, and Ala162. It was observed that the phenyl ring of compound was sandwiched between the phenyl rings of Tyr161 and Phe93 and they formed the pi-pi interaction. Tyr161 formed pi-pi interactions with phenyl ring of Compound.218342, while the carboxyl groups were involved in the formation of two hydrogen bonds with Lys111 and Phe94. Phenolic oxygen was involved in formation of hydrogen bond with Ser162 and Ala162 amino acids. In all the cases, Try 161 was involved in forming pi-pi interaction with the phenyl ring of the compounds. 2D representation of molecular docking results of all three compounds was shown in the Figures 8(b), 8(c), and 8(d). Lys111 formed two hydrogen bonds with the two different oxygen atoms of phenyl groups of the Compound NSC.24871. Also one phenolic oxygen atom formed the two hydrogen bonds with the two hinge regions in amino acids, that is, Ser160 and Ala162. These three compounds retrieved from two databases (NCI & Maybridge), exhibited good interactions with important amino acids in the active sites. Among all three compounds, Compound NSC.218342 retrieved from the NCI database was observed to exhibit good estimated activity, fit values and docking score, and hydrogen bond interactions. Molecular docking results support that these molecules can be further taken as the potential leads for designing novel PDK-1 inhibitors in the future.

Name	Est. (nM)	Interaction energy (kcal/mol)	LigScore1	LigScore2	-PLP1	-PLP2	-PMF
Compound_1	3.00	64.40	4.23	4.98	73.60	68.90	103.10
NSC_218342	31.40	65.39	4.93	4.25	61.42	59.79	83.91
NSC_24871	74.10	51.18	5.03	5.77	65.14	67.09	95.66
NSC_211930	82.00	57.30	6.17	6.23	79.00	71.43	75.76
NSC_84044	87.70	51.11	4.23	5.31	65.40	55.20	92.87
NSC_325657	93.80	55.47	5.84	6.16	75.72	73.50	106.54
NSC_343659	94.50	59.16	4.41	4.47	51.82	53.26	82.51
SB_01794	160.00	46.36	3.67	4.04	53.22	56.12	107.10

A ligand based computational method was used to identify molecular structural features required for effective PDK-1 inhibitors for discovery of drugs to prevent and cure wide variety of cancers. A data set of 83 compounds of selective PDK-1 inhibitors with their respective activities ranging over a wide range of magnitude has been used to generate pharmacophore hypothesis and to predict the activity successfully and accurately. A highly predictive pharmacophore model was generated based on 21 training set molecules, which had hydrogen bond acceptor, hydrogen bond donor, and

hydrophobic aliphatic groups as chemical features which described their activities towards PDK-1 kinase. The validation of the model was based on 62 test set molecules, which finally showed that the model was able to differentiate various classes of PDK-1 inhibitors with a high correlation coefficient of 0.87 between experimental and predicted activities accurately. Further validation of Hypo1 was done by a decoy set method. The decoy set method exhibited GH score of 0.73 which depicts that designed model has very high efficiency in screening the molecules from database. Hypo1 was used as a 3D query to screen the potential molecules from the NCI database as well as Maybridge database. The

hit compounds were filtered subsequently by Lipinski's rule of five and ADMET filtration. Further, molecule selection was refined by docking study. After the docking studies, it was observed that the 3 molecules (NSC.218342, NSC.24871, and NSC.211930) with different scaffolds exhibited better docking energy as well as better interaction. To conclude, the defined drug candidate is further evolved using in vitro and in vivo studies as anticancer molecule.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank Dr. J. A. R. P. Sarma, Senior Vice President, and Dr. K. V. Radhakishan, Ex-Director of GVK Biosciences Pvt., Ltd., for their cooperation and providing software facilities.

References

- [1] M. A. Fabian, W. H. Biggs III, D. K. Treiber et al., "A small molecule-kinase interaction map for clinical kinase inhibitors," *Nature Biotechnology*, vol. 23, no. 3, pp. 329–336, 2005.
- [2] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam, "The protein kinase complement of the human genome," *Science*, vol. 298, no. 5600, pp. 1912–1934, 2002.
- [3] I. Akritopoulou-Zanze and P. J. Hajduk, "Kinase-targeted libraries: the design and synthesis of novel, potent, and selective kinase inhibitors," *Drug Discovery Today*, vol. 14, no. 5-6, pp. 291–297, 2009.
- [4] S. L. Anderson, D. Stokoe, H. Erdjument-Bromage et al., "Protein kinase B kinases that mediate phosphatidylinositol 3,4,5-trisphosphate-dependent activation of protein kinase B," *Science*, vol. 279, no. 5351, pp. 710–714, 1998.
- [5] D. R. Alessi, M. Deak, A. Casamayor et al., "3-Phosphoinositide-dependent protein kinase-1 (PDK-1): structural and functional homology with the *Drosophila* DSTPK61 kinase," *Current Biology*, vol. 7, pp. 776–789, 1997.
- [6] R. A. Currie, K. S. Walker, A. Gray et al., "Role of phosphatidylinositol 3,4,5-trisphosphate in regulating the activity and localization of 3-phosphoinositide-dependent protein kinase-1," *Biochemical Journal*, vol. 337, no. 2-3, pp. 575–583, 1999.
- [7] T. Kobayashi and P. Cohen, "Activation of serum- and glucocorticoid-regulated protein kinase by agonists that activate phosphatidylinositol 3-kinase is mediated by 3-phosphoinositide-dependent protein kinase-1 (PDK-1) and PDK2," *Biochemical Journal*, vol. 339, no. 2, pp. 319–328, 1999.
- [8] T. Kobayashi, M. Deak, N. Morrice, and P. Cohen, "Characterization of the structure and regulation of two novel isoforms of serum- and glucocorticoid-induced protein kinase," *Biochemical Journal*, vol. 344, no. 1, pp. 189–197, 1999.
- [9] J. Park, M. L. Leong, P. Buse, A. C. Maiyar, G. L. Firestone, and B. A. Hemmings, "Serum and glucocorticoid-inducible kinase (SGK) is a target of the PI 3-kinase-stimulated signaling pathway," *The EMBO Journal*, vol. 18, no. 11, pp. 3024–3033, 1999.
- [10] N. Pullen, P. B. Dennis, M. Andjelkovic et al., "Phosphorylation and activation of p70^{s6k} by PDK-1," *Science*, vol. 279, no. 5351, pp. 707–710, 1998.
- [11] D. R. Alessi, M. T. Kozlowski, Q. P. Weng, N. Morrice, and J. Avruch, "3-phosphoinositide-dependent protein kinase 1 (PDK-1) phosphorylates and activates the p70 S6 kinase *in vivo* and *in vitro*," *Current Biology*, vol. 8, no. 2, pp. 69–81, 1998.
- [12] A. Mora, D. Komander, D. M. F. van Aalten, and D. R. Alessi, "PDK-1, the master regulator of AGC kinase signal transduction," *Seminars in Cell and Developmental Biology*, vol. 15, no. 2, pp. 161–170, 2004.
- [13] P. L. Dahia, "PTEN, a unique tumor suppressor gene," *Endocrine-Related Cancer*, vol. 7, no. 2, pp. 115–129, 2000.
- [14] I. Sansal and W. R. Sellers, "The biology and clinical relevance of the *PTEN* tumor suppressor pathway," *Journal of Clinical Oncology*, vol. 22, no. 14, pp. 2954–2963, 2004.
- [15] M. Cully, H. You, A. J. Levine, and T. W. Mak, "Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis," *Nature Reviews Cancer*, vol. 6, no. 3, pp. 184–192, 2006.
- [16] Y. Kurogi and O. F. Guner, "Pharmacophore modeling and three-dimensional database searching for drug design using catalyst," *Current Medicinal Chemistry*, vol. 8, no. 9, pp. 1035–1055, 2001.
- [17] O. F. Guner, O. Clement, and Y. Kurogi, "Pharmacophore modeling and three dimensional database searching for drug design using catalyst: recent advances," *Current Medicinal Chemistry*, vol. 11, no. 22, pp. 2991–3005, 2004.
- [18] H. K. Kyung, W. Allan, B. F. J. Middleton et al., "Benzo[c][2,7]naphthyridines as inhibitors of PDK-1," *Bioorganic and Medicinal Chemistry Letters*, vol. 19, no. 17, pp. 5225–5228, 2009.
- [19] N. Thomas, G. D. Russell, I. Charles et al., "The identification of 8,9-dimethoxy-5-(2-aminoalkoxy-pyridin-3-yl)-benzo[c][2,7]naphthyridin-4-ylamines as potent inhibitors of 3-phosphoinositide-dependent kinase-1 (PDK-1)," *European Journal of Medicinal Chemistry*, vol. 45, no. 4, pp. 1379–1386, 2010.
- [20] J. Zhu, J.-W. Huang, P.-H. Tseng et al., "From the cyclooxygenase-2 inhibitor celecoxib to a novel class of 3-phosphoinositide-dependent protein kinase-1 inhibitors," *Cancer Research*, vol. 64, no. 12, pp. 4309–4318, 2004.
- [21] A. Gopalsamy, M. Shi, D. H. Boschelli et al., "Discovery of dibenzo[c,f][2,7]naphthyridines as potent and selective 3-phosphoinositide-dependent kinase-1 inhibitors," *Journal of Medicinal Chemistry*, vol. 50, no. 23, pp. 5547–5549, 2007.
- [22] *Accelrys Discovery Studio 2.5*, Accelrys, San Diego, Calif, USA, 2009, <http://www.accelrys.com/>.
- [23] A. Smellie, S. Teig, and P. Towbin, "Poling: promoting conformational variation," *Journal of Computational Chemistry*, vol. 16, no. 2, pp. 171–187, 1995.
- [24] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, "CHARMM: a program for macromolecular energy, minimization, and dynamics calculations," *Journal of Computational Chemistry*, vol. 4, no. 2, pp. 187–217, 1983.
- [25] H. Li, J. Sutter, and R. Hoffmann, *Pharmacophore Perception and Development and Use in Drug Design*, International University Line, La Jolla, Calif, USA, edited by O. F. Guner, 2000.
- [26] J. R. Medina, C. W. Blackledge, D. A. Heerding et al., "Aminoin-dazole PDK-1 inhibitors: a case study in fragment-based drug

- discovery,” *ACS Medicinal Chemistry Letters*, vol. 1, no. 8, pp. 439–442, 2010.
- [27] D. Komander, G. S. Kular, A. W. Schüttelkopf et al., “Interactions of LY333531 and other bisindolyl maleimide inhibitors with PDK-1,” *Structure*, vol. 12, no. 2, pp. 215–226, 2004.
- [28] C. M. Venkatachalam, X. Jiang, T. Oldfield, and M. Waldman, “LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites,” *Journal of Molecular Graphics and Modelling*, vol. 21, no. 4, pp. 289–307, 2003.