ISSN NO. 2320-5407

Journal homepage: http://www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH

RESEARCH ARTICLE

Molecular modeling in drug design: A case study on discovery of new PDK1 inhibitors as anticarcinogenic agents

Sayalee R. Chavan[1*], Radha Charan Dash[1], M. Sarwar Alam[2], Raj R Hirwani[1]

1. CSIR Unit for Research and Development of Information Products, "Tapovan", NCL Campus, S.No.113, 114, Pashan, Pune 411 008

2. Department of Chemistry, Faculty of Science, Jamia Hamdard, Hamdard Nagar, New Delhi 110062, India.

Manuscript Info

Manuscript History:

Received: 15 December 2014
Final Accepted: 26 January 2015
Published Online: February 2015

Key words:

3-Phosphoinositide-dependent protein kinase-1 (PDK1), Pharmacophore, 3D-QSAR, Virtual Screening, Anticancer.

Corresponding Author

Sayalee R. Chavan

Abstract

3-Phosphoinositide-dependent protein kinase-1 (PDK1) is a member of serine/threonine kinase family which plays an important role in signaling pathways activated by several growth factors and hormones. Recent data have revealed that the alteration of PDK1 is a critical component of oncogenic phosphoinositide 3-kinase signaling in breast cancer, suggesting that inhibition of PDK1 can inhibit breast cancer progression. Thus, targeting PDK1 may be a valuable anticancer strategy that may improve the efficacy of chemotherapeutic strategies in breast cancer patients. In this study, pharmacophore model was generated by using previously reported 2-aminopyrimidine based PDK1 inhibitors. The best pharmacophore model generated consisted of five features ADDRR: one hydrogen bond acceptor (A), two hydrogen bond donors (D) and two aromatic rings (R). Based on derived pharmacophore, an atom based 3D-QSAR model having r^2 =0.944 and q^2 = 0.718 was developed to evaluate the structure activity relationships. Docking study was performed which validated pharmacophore and 3D-QSAR models. Two potent hits were identified by screening in-house natural product database by structure based virtual screening approach. In silico toxicity prediction study was performed to evaluate safety of the retrieved hits.

INTRODUCTION

3-Phosphoinositide-dependent protein kinase-1 (PDK1) is a potential target for developing novel anticancer drugs. PDK1 is a protein kinase belonging to the cAMP-dependent, cGMP-dependent protein kinase C (AGC) kinase family (Toker et.al, 2000). PDK1 has an N-terminal kinase domain and a C-terminal pleckstrin homology (PH) domain. The PH domain is essential for interaction with the cell membrane, and the kinase domain is involved in the phosphorylation and activation of downstream kinases. PDK1 phosphorylates the 3-position of phosphatidylinositol (PI), PI(4)P, and PI(4,5)P2 to give rise to three signalling phospholipids: PI(3)P, PI(3,4)P2, and PI(3,4,5)P3, p70 ribosomal S6 kinase, serum- and glucocorticoid induced protein kinase, and protein kinase C isoforms. Hence, PDK1 is known as the master regulator of AGC kinases (Mora et.al, 2004). Accumulating pharmacologic and genetic evidence supports the potential role of PDK1 as a promising anticancer target (Liang et.al, 2006, Zeng et.al, 2002, Flynn et.al 2000, Bayascas et.al, 2005). It has been reported that overexpression of PDK1 in mammary cells resulted in their transformation in vitro and tumour formation in vivo. Elevated levels of PDK1 phosphorylation

were also reported in metastasized breast tumors (Lin et.al, 2005). Another study in which PDK1 is targeted with antisense oligonucleotides has reported a marked reduction of cell proliferation and survival and also an increased rate of apoptosis than that observed in PDK1 or PKB inhibition. PDK1-expressing cells were shown to have enhanced invasion and cell growth (Xie et.al, 2006). PDK1 was also reported as a potential target for sensitizing breast cancer cells to chemotherapeutic agents. Knock down of PDK1 was recently reported to enhance the antitumor effect of an epidermal growth factor receptor inhibitor (Zhang et.al, 2006). This also underlines the role of PDK1 in the development of resistance to current cancer therapies. PDK1 hypomorphic mice expressing only 10% of the normal level of PDK1 were reported to be viable and fertile (Lawlor et.al, 2002). This finding has shown that inhibition of PDK1 can be achieved without severe toxicity. A more advanced study using PDK1 hypomorphic mice has recently revealed that the reduced PDK1 expression in PTEN mice markedly protected the animal from a wide range of tumours. Thus, PDK1 has become a well-validated anticancer target. Kinase inhibitors, such as Imatinib (Gleevec) and Erlotinib (Tarceva), have recently been approved for clinical use in the market as anticancer agents (Garber, 2006). The success of these new drugs has given a new impetus for developing better anticancer agents that target particular pathways with a specific kinase overexpressed and critical to tumor progression (Kim, 2003). Because PDK1 was found to be overexpressed in cancers and has a crucial role in the survival pathway, it is an ideal target for drug development. The development of PDK1 inhibitors could pave the way for the development of better anticancer therapeutics. Compared to other kinases, only a few inhibitors of PDK1 have been reported in the literature as potent nanomolar inhibitors, (Komander et.al, 2003) including staurosporine, UCN-01 and amino pyrimidine derivatives (Feldman et.al, 2003). More recently, a new series of indolinone derivatives were reported as PDK1 inhibitors. This new series of compounds were developed using an initial hit identified from a high-throughput-screening study (Islam et.al, 2007 a, b). Structure-based as well as ligand-based computational approaches have been found to be valuable in further optimization and the development of novel inhibitors. The 3D-QSAR modelling is useful to predict the activity of new molecules to be synthesized (Martin 1998). Molecular docking is used to study how a ligand is interacting with its biological target and to support the conclusions of QSAR studies. Virtual screening is a productive and cost-effective technology in drug design for the identification of novel lead compounds that allow chemists to reduce a huge virtual library to a more manageable size. In our investigation, computer-aided drug design techniques such as pharmacophore modelling, 3D-QSAR study, molecular docking, and virtual screening and toxicity prediction were used to investigate potent inhibitors of PDK1.

Material and Methods

Dataset Collection

The literature was extensively surveyed to collect diverse PDK1 inhibitors. A dataset of 46 6-(2-aminopyrimidine-4-yl)-1H-indazol-3-amine derivatives, (Table 1) was used for pharmacophore modelling and subsequent QSAR analysis. The negative logarithm of the measured IC50 value (pIC50) against PDK1 in an isolated enzyme assay data was used in this study. (Medina et.al, 2011)

Pharmacophore generation and 3D-QSAR model

PHASE (Dixon et.al, 2006) implemented in the Maestro software package of Schrodinger was used to generate pharmacophore and 3D-QSAR models for 2-aminopyrimidine derivatives as PDK1 inhibitors. PHASE is a highly comprehensive, self-contained system for pharmacophore perception, 3D-QSAR model development and 3D database screening**.** By employing a novel, tree-based partitioning algorithm, PHASE exhaustively identifies spatial arrangements of functional groups that are common and essential to the biological activity for a set of high affinity ligands. Combination of activity data with given hypothesis creates a 3D- QSAR model capable of identifying important aspects of molecular structure that govern activity. This model along with hypothesis can mine a chemical database to identify hits with potential activity towards the target.

The 3D structure of each compound was built using Build module with the default Maestro settings. The 3D structures were minimized by default Universal force field within Maestro. A maximum of 2,500 conformers were generated per structure using minimization of 100 steps. Combinations of Monte-Carlo Multiple Minimum (MCMM) and Low Mode (LMOD) were used to study the conformational space of all the molecules. In order to eliminate high-energy or redundant conformers, the maximum relative energy difference was set to 10.0 kcal mol−1, and the cut off root mean square deviation (RMSD) was 2Å. As active compounds are normally considered when developing common pharmacophore hypothesis, ―pharmaset was defined by setting threshold of pIC50 $\geq$ -1.77 for active molecules and pIC50 $\leq$ -1.92 for inactive molecules. PHASE provides a standard set of six pharmacophore features, hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively ionizable (N), positively ionizable (P), and aromatic ring (R). Common pharmacophore hypothesis was identified from molecules of active pharmaset having identical set of features, same spatial arrangement and with a minimum inter site

distance 2.0 Å in a final box size of 1.0 Å. A five point common pharmacophore hypothesis was identified from all conformation of the active ligands having identical set of features. Generated common pharmacophore hypotheses were examined by scoring alignment of actives against a reference ligand by using default settings for Score Actives to identify the pharmacophore from each box that yielded the best alignment of the active ligands. The scoring procedure provided a ranking of different hypotheses from which further investigation was carried out for appropriate hypothesis with rational choice. The hypotheses were ranked according to survival values for actives and inactive compounds.

An atom-based 3D-QSAR model is more useful in explaining the structure activity relationship than pharmacophore-based 3D-QSAR as latter does not consider ligand features beyond the pharmacophore model. In atom-based 3D-QSAR, a molecule is treated as a set of overlapping van der Waals spheres. Each feature is categorized according to a simple set of rules: hydrogens attached to polar atoms are classified as hydrogen bond donors (D); carbons, halogens, and C–H hydrogens are classified as hydrophobic/non-polar (H); atoms with an explicit negative ionic charge are classified as negative ionic (N); atoms with an explicit positive ionic charge are classified as positive ionic (P); N, O and hydrogen-bond acceptors are classified as electron withdrawing (W) and all other types of atoms are classified as miscellaneous (X). For purposes of 3D-QSAR development, Van der Waals models of the aligned training set molecules were placed in a regular grid of cubes, with each cube allotted zero or more bits to account for the different types of atoms in the training set that occupy the cube. This representation gives rise to binary-valued occupation patterns that can be used as independent variables to create partial least-squares (PLS) 3D-QSAR models. Atom-based 3D-QSAR models were generated for the hypothesis using the 33-member training set using a grid spacing of 1.0 Å. The best 3D-QSAR model was validated by predicting activities of the 13 test set compounds. 3D-QSAR models containing one to six PLS factors were generated. The 3D-QSAR was evaluated by cross validated r^2, q^2, SD, and RMSE Person-R. The predicted 1/log IC_{50} at 4th PLS factor are tabulated in (Table 2).

Molecular Docking

Protein Databank (PDB) entry 3NUN was used as the source of PDK1 structure for docking studies. The hydrogen atoms were added to the proteins and further minimization was performed using protein preparation wizard. During the docking process, Glide initially performs a complete systematic search of the conformation, orientation, and position of a compound in the defined binding site and eliminates unwanted poses using scoring and energy optimization. Finally the conformations were further refined via a Monte Carlo sampling. The grid boxes were defined by the centre of the bound inhibitors of the PDK1 proteins. The enclosing box and binding box dimensions were fixed to 14 and 10 Å, respectively. The top 10 poses were collected for each compound. Docking poses were energy minimized using the OPLS-2005 force field.

Docking studies were performed using Glide (Friesner et.al, 2004, Halgren et.al, 2004, Friesner et.al, 2006) (Schrödinger, New York, USA). It performs grid-based ligand docking with energetic and searches for favourable interaction between one or more typically small ligand molecules and a typically larger receptor molecule, usually a protein. Glide provides three different levels of docking precision, HTVS high-throughput virtual screening; SP, standard precision, XP, extra precision. We carried out our calculations in SP first, and then XP mode, a refinement tool designed only on good ligand poses. All molecules were built within maestro by using build, exhaustive conformational search carried out for all molecules using OPLS_2005 force field, and imposing a cut-off of allowed value of the total conformational energy compared to the lowest-energy state. Minimization cycle for conjugate gradient and steepest descent minimizations was used with default value 0.05Å for initial step size and 1.00Å for maximum step size. In convergence criteria for the minimization, both the energy change criteria and gradient criteria, was used with default values 10−7 and 0.001 kcal/mol, respectively.

Virtual screening

Virtual screening refers to a range of in silico techniques used to search large compound databases to select a smaller number for biological testing. Virtual screening workflow included in Glide module of Schrodinger was used for screening. Glide includes series of hierarchical filters which are used to rapidly eliminate pose of ligand which can not correspond to a well-docked solution. There are 3 modes of docking HTVS, SP and XP. For the exploration of novel hits with potential anti-cancer activity, an in silico screening of in-house natural product database which consists of 10,000 natural compounds was performed using the structure based virtual screening approach.

Toxicity prediction of Hits

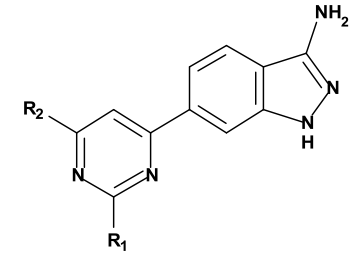
To eliminate or optimize the hits having toxicity in early drug discovery stage, toxicity prediction of retrieved hits were carried out by using OSIRIS property explorer (Thomas Sander, 2001). The OSIRIS property explorer includes 3,300 traded drugs as well as 15,000 commercially available chemicals with associated drug likeliness. Based on a

list of all available fragments of these drugs, this program calculates the drug likeliness of candidate molecule. Overall, the drug score was calculated by multiplying contributions of the individual properties viz drug likeliness, cLogP, logS, molecular weight and toxicity risk providing a confidence range (0–1) for a compound to qualify as a drug. Results are coded with different coloration. The green coloration represents that the compound is non-toxic, yellow and red coloration represents moderate and severe toxicity of the chemicals, respectively.

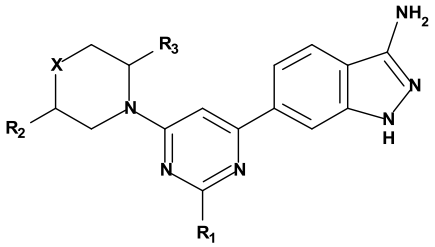
Result and Discussion

In an effort to discover new inhibitors of PDK1, ligand based pharmacophore hypothesis was generated which was validated by 3D-QSAR analyses. Docking study was performed to study the binding interactions between each active molecule and PDK1 binding site. In silico screening of in-house database of natural products was carried out to identify potent inhibitors of PDK1.

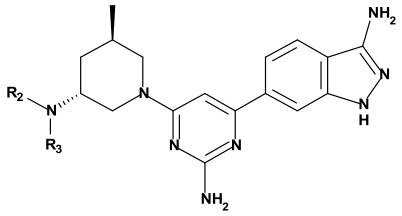
Table 1: Dataset analyzed with experimental and predicted activities



A



B



C

#	Core	R1	R2	R3	X	Expt pIC_{50} (nM)	Pred pIC_{50} (nM)	Fitness
1	A	NH_2	H	---	---	-1.806	-1.78	2.83
2[a]	A	-$NHCH_3$	H	---	---	-1.778	-1.79	2.49
3[t, a]	A	$N(CH3)_2$	H	---	---	-1.672	-1.77	2.76
4[t, a]	A	$NHCH_2CH_3$	H	---	---	-1.708	-1.75	2.76
5[a]	A	OCH3	H	---	---	-1.69	-1.71	2.76
6	A	NH_2	CH3	---	---	-1.839	-1.83	2.8
7	A	NH_2	i-Pr	---	---	-1.82	-1.8	2.89
8	A	NH_2	Ph	---	---	-1.826	-1.83	2.83
9	A	NH_2	OCH3	---	---	-1.82	-1.81	2.48
10	A	NH_2	NHCH3	---	---	-1.851	-1.81	2.92
11[a]	A	NH_2	NHCH2CH3	---	---	-1.748	-1.8	3
12[t]	A	NH_2	CH2NHCH(CH3)2	---	---	-1.839	-1.83	2.95
13[t]	A	NH_2	Pending	---	---	-1.863	-1.85	2.89
14	A	NH_2	NHPh	---	---	-1.886	-1.88	2.81
15[t]	A	NH_2	N(CH3)2	---	---	-1.799	-1.82	2.88
16[t]	A	NH_2	N-morpholine	---	---	-1.813	-1.82	2.83
17	A	NH_2	Pyrrolidine	---	---	-1.82	-1.84	2.85
18	A	NH_2	Piperidine	---	---	-1.82	-1.82	2.83
19	A	NH_2	(R)-3-methylmorpholine-4-yl	---	---	-1.857	-1.86	2.8
20[t, a]	A	NH_2	(S)-3-methylmorpholine-4-yl	---	---	-1.771	-1.8	2.84

21	A	NH_2	(R)-3-ethylmorpholine-4-yl	---	---	-1.875	-1.87	2.75
22[t]	A	NH_2	(R)-3-isopropylmorpholine-4-yl	---	---	-1.826	-1.86	2.72
23	A	NH_2	(R)-2-methylpiperidine	---	---	-1.839	-1.85	2.81
24[t]	B	NH_2	(S)- N-phenylamide	---	O	-1.886	-1.88	2.62
25[a]	B	NH_2	(R)- N-phenylamide	---	O	-1.778	-1.76	2.62
26	B	NH_2	(S)- N-phenylamide	---	CH_2	-1.875	-1.87	2.62
27	B	NH_2	(R)- N-phenylamide	---	CH_2	-1.839	-1.83	2.62
28	B	NH_2	(S)- N-cyclohexylamide	---	CH_2	-1.857	-1.86	2.61
29[t]	B	NH_2	(S)- N-cyclopentylamide	---	CH_2	-1.839	-1.86	2.63
30[t]	B	NH_2	(S)- N-phenylamide	(R)-CH_3	CH_2	-1.929	-1.92	2.6
31[a]	B	NH_2	(R)- N-phenylamide	(S)-CH_3	CH_2	-1.771	-1.78	2.63
32	B	NH_2	(R)- N-phenylamide	(R)-CH_3	CH_2	-1.785	-1.79	2.63
33	B	$NHCH_3$	(S)- N-phenylamide	(R)-CH_3	CH_2	-1.944	-1.94	2.6
34	B	NH_2	(S)- N-cyclohexylamide	(R)-CH_3	CH_2	-1.908	-1.91	2.61
35	B	$NHCH_3$	(S)- N-cyclohexylamide	(R)-CH_3	CH_2	-1.934	-1.93	2.59
36	B	NH_2	(S)- N-phenylamide	(R)-CH_3	O	-1.919	-1.93	2.61
37	B	$NHCH_3$	(S)- N-phenylamide	(R)-CH_2CH_3	O	-1.929	-1.92	2.58
38[t]	B	$NHCH_3$	(S)- N-cyclohexylamide	(R)-CH_3	O	-1.929	-1.92	2.59
39	B	$NHCH_3$	(S)- N-cyclohexylamide	(R)-CH_2CH_3	O	-1.914	-1.91	2.57
40[a]	B	NH_2	(S)-tert-butyl carbamate	H	CH_2	-1.778	-1.79	2.62
41	B	NH_2	(R)-tert-butyl carbamate	H	CH_2	-1.881	-1.87	2.64
42	B	NH_2	(S)-tert-butyl carbamate	(R)-CH_3	CH_2	-1.792	-1.8	2.63
43[t]	B	NH_2	(R)-tert-butyl carbamate	(S)-CH_3	CH_2	-1.903	-1.83	2.65
44	C	---	tert-butyl carboxylate	H	---	-1.914	-1.92	2.63
45	C	---	tert-butyl carboxylate	CH_3	---	-1.959	-1.96	2.61
46	C	---	3,3-dimethylketone	CH_3	---	-1.964	-1.98	2.6

a = active pharma set, i= inactive pharma set, t = test set.

Table 2: Statistical values for common pharmacophore feature with 3D-QSAR model generated by PLS

Training set	**Test set**
m = 6	
n = 34	n_T = 12
r^2 = 0.944	q^2 = 0.718
SD = 0.0166	RMSE = 0.04
F = 119 , P = 3.67e-017	Pearson-R = 0.8914

m = number of PLS factors in the model; n = number of molecules in the training set;

nT = number of molecules in test set; r2 = coefficient of determination;
q2 = r2 for test set; SD = standard deviation of regression; RMSE = root-mean squared error;
F = variance ratio; P = statistical significance; Pearson-R = Pearson correlation coefficient

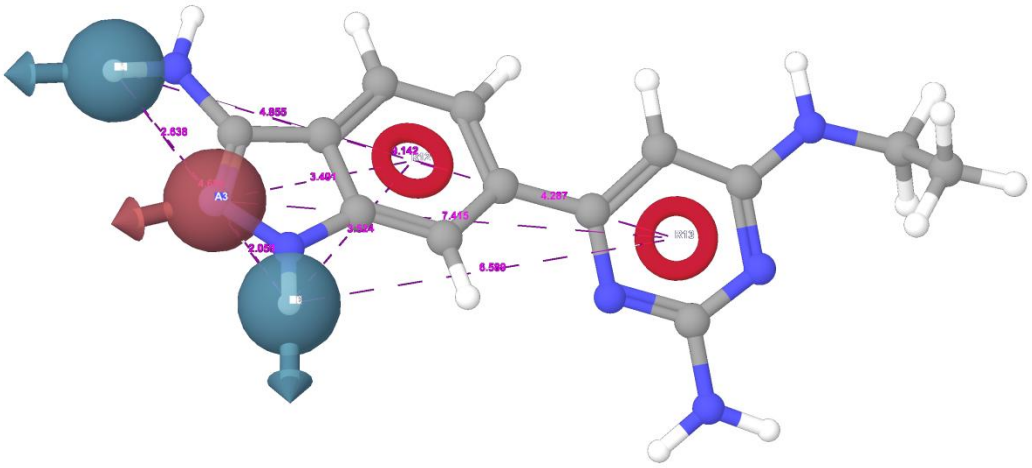


Fig 1: Pharmacophore hypothesis and distance between pharmacophoric sites

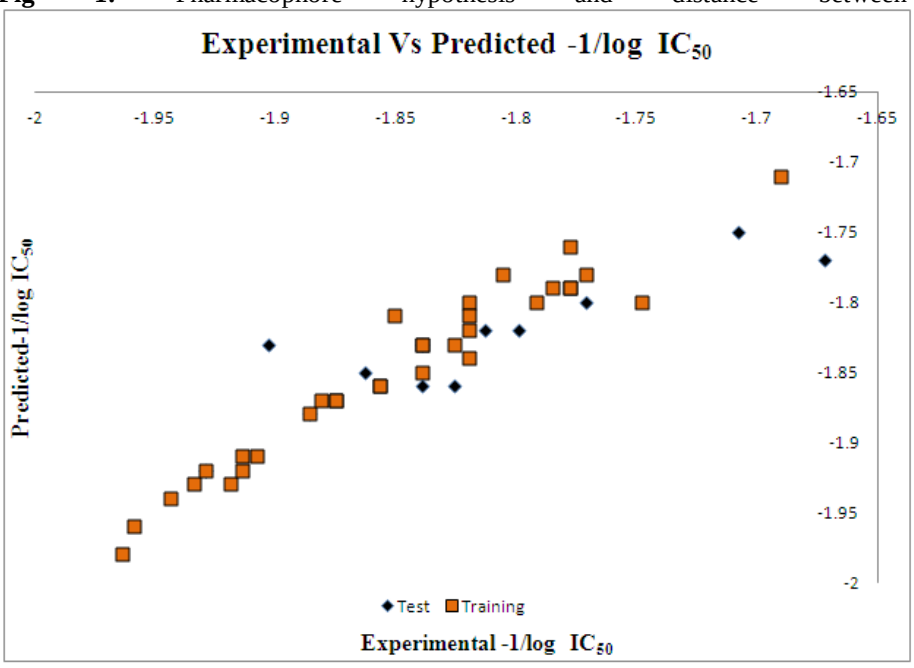


Fig 2: The correlation graph between predicted and actual pIC50 for training and test set

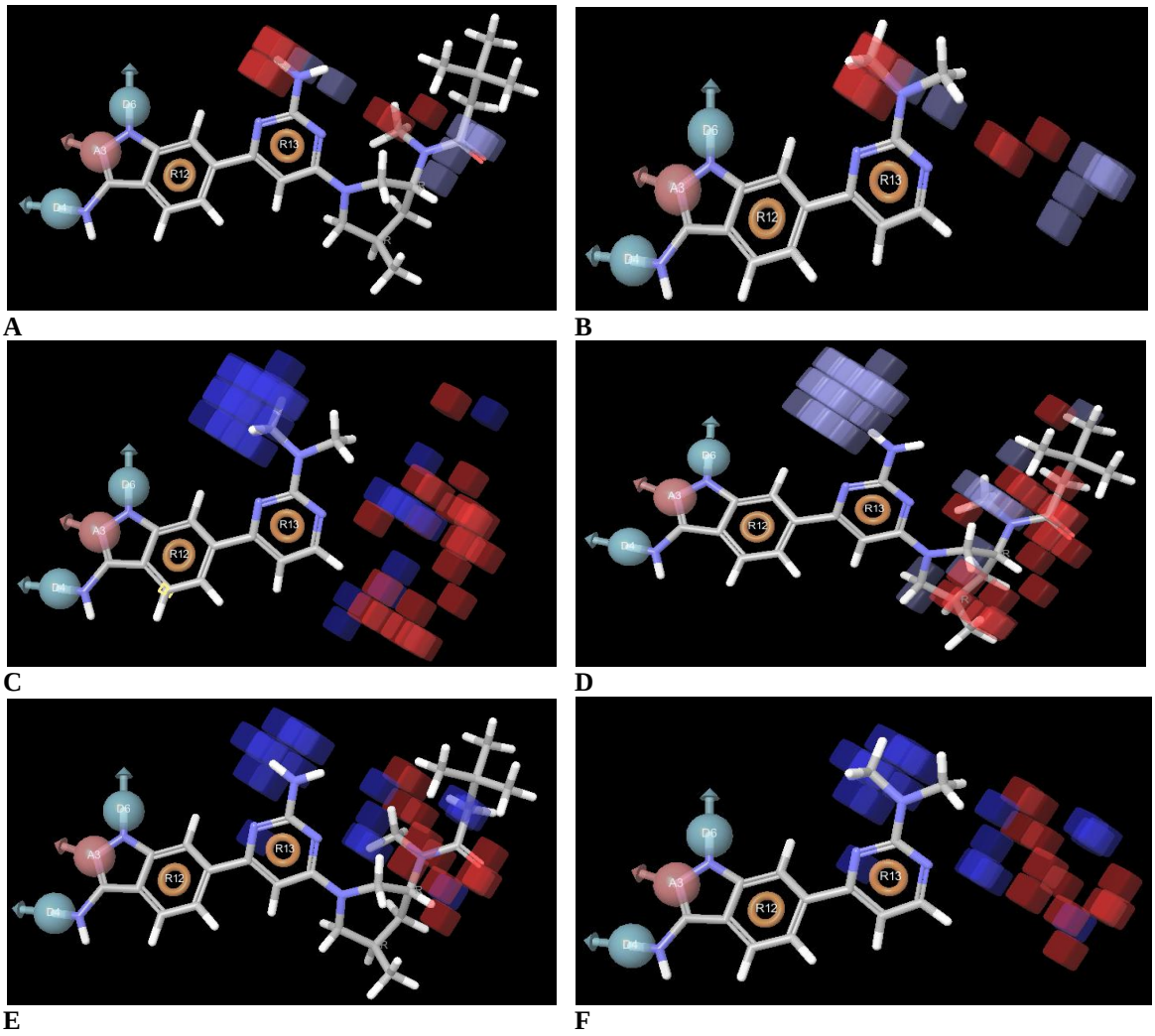
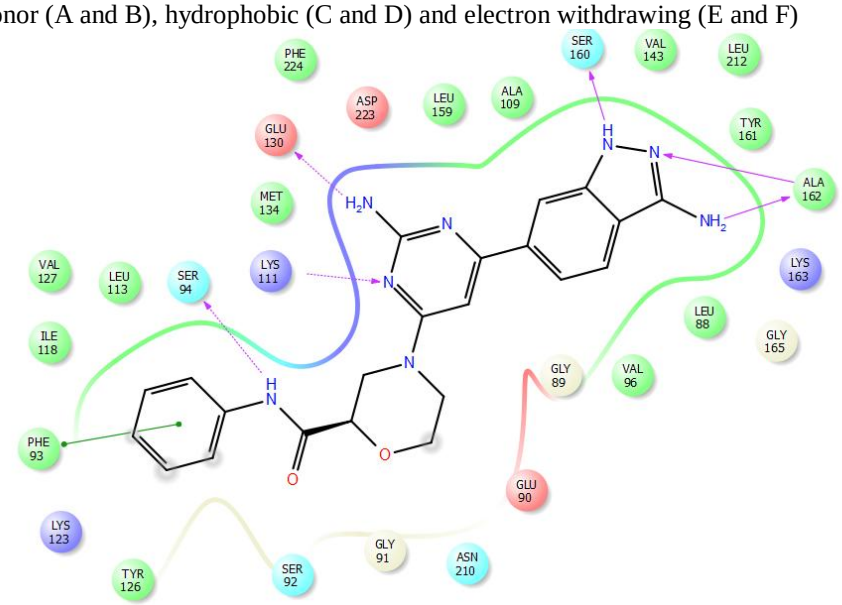


Fig 3: Comparison of most active (3) and least active (46) molecules by 3D-QSAR visualization; Effects for H-bond donor (A and B), hydrophobic (C and D) and electron withdrawing (E and F)



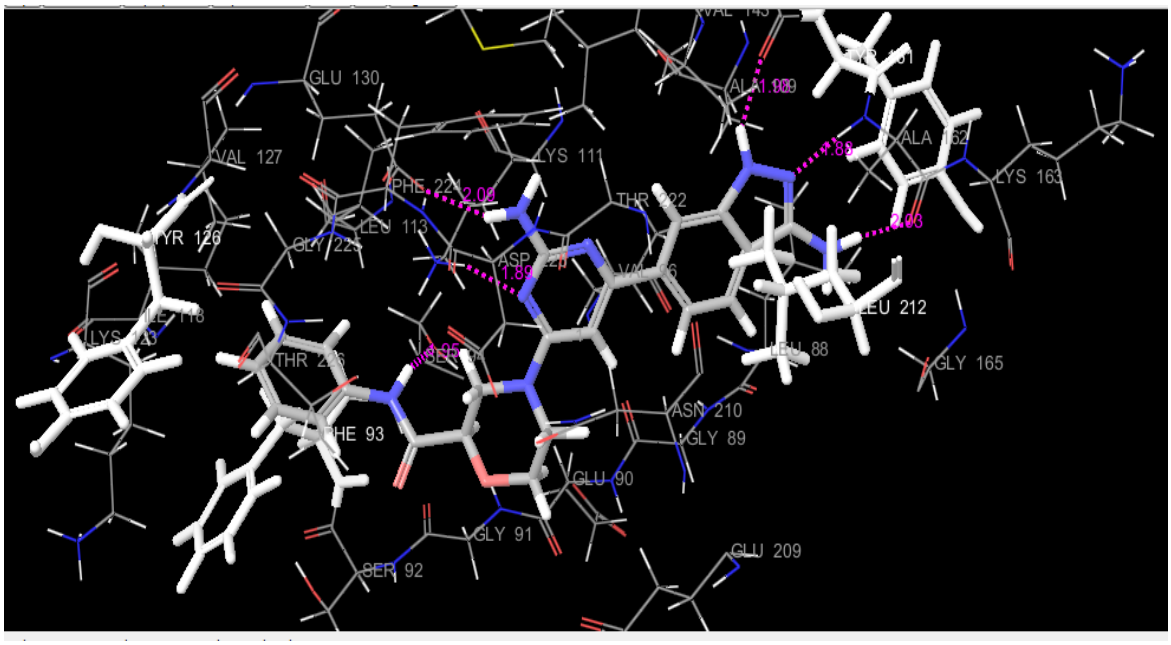
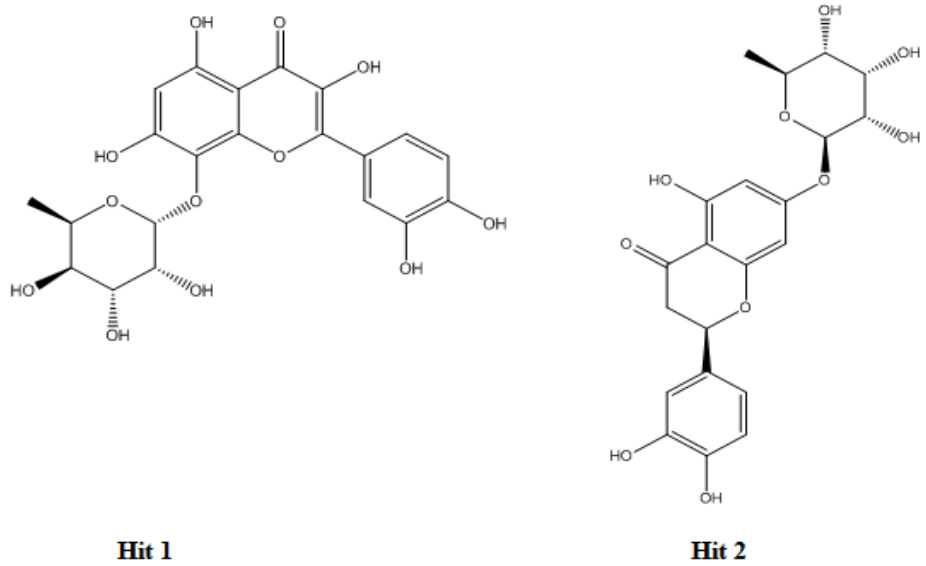
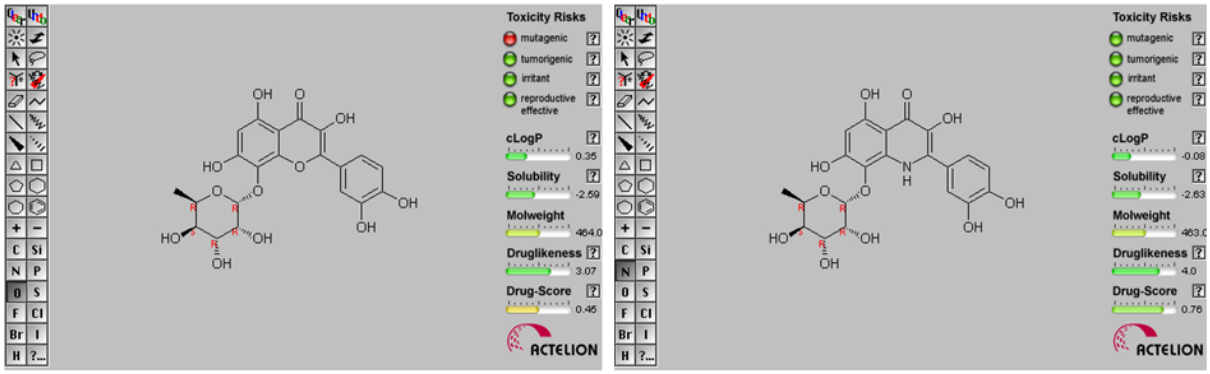


Fig 4: Interaction of compound 25 with PDK 1 active site



Hit 1 **Hit 2**

Fig 5: Hits identified from the virtual screening

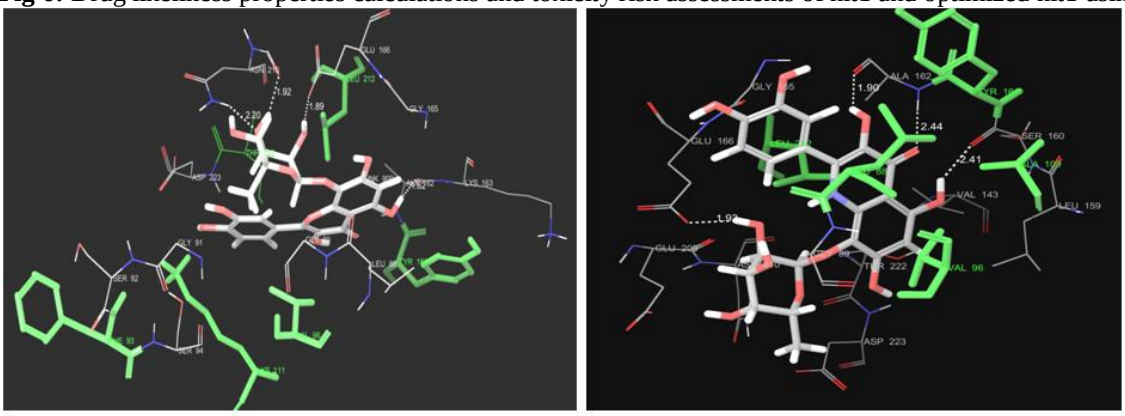


Properties with high risks of undesired effects are shown in red. Whereas a green color indicates drug-conform behaviour.

cLogP=0.347
molweight:464.0)
LogS=-2.588)
drug likeness=3.067)
High Risk of Mutagenicity, Score: 0.6
No Risk of Tumorigenicity, Score: 1.0
No Risk of Irritating Effects, Score: 1.0
No Risk of Reproductive Effects, Score
The drug Score is 0.449

cLogP=-0.082
molwt: 463.0
LogS=-2.63
drug likeness=4.002
No Risk of Mutagenicity, Score: 1.0
No Risk of Tumorigenicity, Score: 1.0
No Risk of Irritating Effects, Score: 1.0
No Risk of Reproductive Effects, Score
The drug Score is 0.76

Fig 6: Drug likeliness properties calculations and toxicity risk assessments of hit1 and optimized hit1 using OSIRIS



Hit 1 **Optimized Hit 1**

Fig 7: Comparison of docking interactions of hit1 and optimized hit

Pharmacophore generation and 3D-QSAR model building

To find common pharmacophore hypothesis, the data sets were divided into active and inactive sets (Golbraikh et al., 2003) depending upon the observed activity. Molecules having pIC50 value higher than -1.78 were considered to be active, and those with pIC50 value below -1.92 were considered to be inactive.

Active molecules were selected for the generation of common pharmacophore hypotheses. Using a tree-based partition algorithm requiring that all active compounds must match, five-featured probable common pharmacophore hypotheses were generated from the list of variants. Scoring function such as score active, score inactive and rescore were applied for all five-featured common pharmacophore hypotheses. The dataset was randomly divided into 72% of total population containing 33 compounds in the training set and 28% of total population containing 13 compounds in the test set. Training and test set compounds were aligned on the common pharmacophore hypotheses and analyzed through PLS values. The predictivity of each hypothesis was analyzed by the test set compounds. Alignment of active compounds shows that all active compounds match with the hypothesis.

The most predictive QSAR model, based on r^2 and q^2 values, was found to be associated with the five-featured hypothesis, which contains one hydrogen bond acceptor, two hydrogen bond donors, and two aromatic rings (ADDRR). The pharmacophore contains one aromatic ring feature mapping on the phenyl part of the indazole ring, and another aromatic ring feature is of pyrimidine ring, one hydrogen bond donor feature mapping on the secondary amine groups of the indazole ring and another hydrogen bond donor feature mapping on primary amine group attached to the indazole ring. One hydrogen bond acceptor mapping on the tertiary amine group of the indazole ring. The pharmacophore hypothesis showing distance between pharmacophoric sites is depicted in (Fig. 1).

The PHASE descriptors serve as independent variables and activity values as dependent variables in deducing 3D-QSAR models by PLS regression analysis method (Jain et al., 2010). The PHASE predicted activity of the ligand molecules were given in (Table 1). Good predictability as well as statistical significance (low—SD, P, and RMSE values; high—r^2, Pearson-R, and F values) turned out as typical features of the developed models satisfying primary requirements for a QSAR study. The statistical results of atom-based 3D-QSAR study are summarized in (Table 2). High r^2 is necessary but not sufficient condition for a predictive QSAR model (Tropsha, 2010). Besides the consideration of high r^2, the best QSAR model should be chosen based on its predictive ability, so the best model should have high q^2 also. The hypothesis ADDRR yielded a statistically significant 3D-QSAR model with PLS factors (r^2 = 0.944, SD = 0.0166, F = 119, P = 3.67e-017) for training set of 33 compounds and PLS factors (q^2 = 0.718, RMSE = 0.04, Pearson-R = 0.8914) for test set of 13 compounds. Plots of predicted vs. actual pIC50 for training and test set were reported in (Fig. 2). It shows that all values were plotted around the best fit line indicating the significance of predicted model.

Analysis of atom-based 3D-QSAR model

By visualizing the hypothesis and various ligands in the context of the QSAR model, additional insight may be gained about the relevance of each feature in explaining activity.

We visualized and analyzed the 3D-QSAR models based on active reference molecule 11 using hydrogen-bond donor, hydrophobicity, and electron withdrawing ionic features. The volume occlusion map obtained from 3D-QSAR study is useful to identify the importance of the above mention properties with respect to the biological activity of the ligand. The volume occlusion maps generated from 3D-QSAR study is shown in (Fig. 3). In these maps, blue cubes area indicate favourable regions while red cubes area indicate unfavourable regions for inhibitory activity.

The volume occlusion map of H-bond donor feature of least active ligand 46 (Fig. 3A) showed red unfavourable area near the N atom attached to pyrimidine ring (R3) which indicates that presence of H-bond donor group resulted in decrease in biological activity. In contrast, absence of H-bond donor group at that position increases the biological activity as in most active compound 3 (Fig. 3B). In the hydrophobic contribution map (Fig. 3C); blue region near N atom attached to pyrimidine ring indicates that presences of bulky group have favourable effect on biological activity. This explains the activity of compounds 3, 33, 35, 37, 38 and 39. Absence of such a bulky group decreases the biological activity as in compound 46 (Fig. 3D). Similarly bulky group near piperidine ring have negative effect on biological activity. It explains the activity of compounds 45-48. In electron withdrawing volume occlusion map red region near piperidine ring suggest that presence of electron withdrawing group at that position may lead to decrease in the biological activity. It explains the activity of compound 43-46 (Fig. 3E). While, absence of electron withdrawing group has favorable effect on biological activity as in compound 3 (Fig. 3F).

Docking analysis

To identify the bonding interactions of protein with inhibitors docking study was performed. The crystal structure of the target protein was obtained from PDB bank (PDB code 3NUN). Protein was prepared for docking by using

protein preparation wizard. Active pharmaset were docked with the crystal structure of PDK1 (PDB code 3NUN). Fig. 4 depicted the complex structure of active compound 25 with PDK1 receptor. As per pharmacophore secondary amino group of indazole ring acts as H-bond donor and forms H-bonds with SER160 with bond distance 1.94 Å and primary amine group attached to indazole ring forms H-bond with ALA 162 with bond distance 1.98 Å. While, tertiary nitrogen of indazole ring acts as H-bond acceptor and forms H bond with ALA 162 with bond angle 2.1 Å. one of the nitrogen atom of pyrimidine ring acts as H-bond acceptor and forms H-bonds with LYS111 with bond distance 1.8 Å. Primary amine attached to pyrimidine ring acts as H- bond donor and forms H bonds with GLU 130 with bond distance 2.0 Å. Amide nitrogen atom acts as a H-bond donor and forms hydrogen bond with SER 94 with bond distance 1.95 Å. Aniline ring shows hydrophobic interaction with PHE93, VAL127, LEU113, ILE118., TYR126. Indazole ring shows hydrophobic interactions with VAL143, LEU212, TYR161, and LEU88.

Virtual screening

To identify the potential inhibitors of PDK1, structure based virtual screening was performed. In-house natural product database containing 10,000 compounds were screened by using various filters. Primary screening was done by using Lipinski rule. It retrieved 650 hits. These hits were further screened based on HTVS followed by SP docking score which retrieved 200 hits. XP docking were carried out for the top 10 retrieved hits. Finally two hits were retrieved having XP docking score ≤ -10.0 and are shown in Fig. 5.

Toxicity prediction of Hits

To eliminate or optimize the hits having toxicity in early drug discovery stage, toxicity prediction of both the hits were carried out by using OSIRIS property explorer. The OSIRIS property explorer includes 3,300 traded drugs as well as 15,000 commercially available chemicals with associated drug likeliness. Based on a list of all available fragments of these drugs, this program calculates the drug likeliness of candidate molecule. Out of the two hits, Hit 1 shows mutagenic property, therefore that hit was optimized by structural modification in order to remove its toxicity. The optimized hit shows no toxicity. Hit 2 fulfills drug likeness properties and do not have toxicity risk. Fig. 6 demonstrates the predicted drug score along with the toxicity risk parameters for one of the retrieved hit (hit1).

Docking of Hits

To validate the structural modifications of hit 1, XP docking was performed and compared with the original Hit. PDK1 crystal structure (PDB code: 3NUN) was used for docking. The structurally modified hit shows similar docking score as original as shown in Fig 7.

Conclusions

In conclusion, pharmacophore and 3D-QSAR models were generated using 2-aminopyrimidine analogues as inhibitors of PDK1. A five-point pharmacophore model having one hydrogen bond acceptor (A), two donors and two aromatic rings (R) were generated. Docking was performed to validate the pharmacophore model and 3D-QSAR predictions which show good complementarity between the pharmacophoric features and the proposed 3D-QSAR model. Further, to evaluate potential and safe inhibitors of PDK1, structure based virtual screening of natural product database was performed. 2 potent hits having xp docking score ≤ -10.0 were retrieved from virtual screening which further subjected to the toxicity prediction and optimization to evaluate the safety. The pharmacophore model, 3D-QSAR study and discovery of potent inhibitor of PDK1 by virtual screening process will be useful in the drug discovery process which saves considerable amount of labor and time and also cost-effective.

Acknowledgments This work was supported by the Council of Scientific and Industrial Research-Unit for Research and Development of Information Products.

References

Bayascas, J. R., Leslie, N. R., Parsons, R., Fleming, S., Alessi, D. R. (**2005):** Hypomorphic mutation of PDK1 suppresses tumorigenesis in PTEN (+/-) mice. Curr. Biol., 15, 1839–1846.

Bohm, M., Sturzebecher, J., Klebe, G. **(1999):** Three dimensional quantitative structure activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. J. Med. Chem., 42: 458–477.

Dixon, S.L.; Smondyrev, A.M.; Knoll, E.H.; Rao, S.N.; Shaw, D.E.; Friesner, R.A., (2006), "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., 20: 647-671

Dixon, S.L.; Smondyrev, A.M.; Rao, S.N.,**(2006)** "PHASE: A Novel Approach to Pharmacophore Modeling and 3D Database Searching," Chem. Biol. Drug Des., 67: 370-372

Feldman, R. I., Wu, J. M., Polokoff, M. A., Kochanny, M. J., Dinter, H., Zhu, D., Biroc, S. L., Alicke, B., Bryant, J., Yuan, S., Buckman, B. O., Lentz, D., Ferrer, M., Whitlow, M., Alder, M., Finster, S., Chang, Z., Arnaiz, D. O. (**2005):** Novel small molecule inhibitors of 3-phosphoinositide dependent kinase-1. J. Biol. Chem., 280: 19867–19874.

Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shaw, D. E.; Shelley, M.; Perry, J. K.; Francis, P.; Shenkin, P. S., **(2004)** "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy," J. Med. Chem., 47:1739–1749

Friesner, R.A.; Murphy, R.B.; Repasky, M.P.; Frye, L.L.; Greenwood, J.R.; Halgren, T.A.; Sanschagrin, P.C.; Mainz, D.T., **(2006)** "Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes," J. Med. Chem., 49, 6177–6196

Flynn, P., Wong, M., Zavar, M., Dean, N. M., Stokoe, D. (**2000):** Inhibition of PDK-1 activity causes a reduction in cell proliferation and survival. Curr. Biol., 10: 1439–1442.

Garber, K. (**2006):** The second wave in kinase cancer drugs. Nat. Biotechnol., 24: 127–130.

Golbraikh, A.; Tropsha, A. **(2003)** QSAR Modelling Using Chirality Descriptors Derived From Molecular Topology. J. Chem. Inf. Comput. Sci. 43:144-154

Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., **(2004)** "Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening," J. Med. Chem., 47, 1750–1759

Islam, I., Bryant, J., Chou, Y. L., Kochanny, M. J., Lee, W.;Phillips, G. B., Yu, H. Y., Adler, M., Whitlow, M., Ho, E.; Lentz, D., Polokoff, M. A., Subramanyam, B., Wu, J. M., Zhu, D. G., Feldman, R. I., Arnaiz, D. O. (**2007)** Indolinone based phosphoinositidedependent kinase-1 (PDK1) inhibitors. Part 1: Design, synthesis and biological activity. Bioorg. Med. Chem. Lett., 17: 3814–3818.

Islam, I., Brown, G., Bryant, J., Hrvatin, P., Kochanny, M. J., Phillips, G. B., Yuan, S. D., Adler, M., Whitlow, M., Lentz, D., Polokoff, M. A., Wu, J., Shen, J., Walters, J., Ho, E., Subramanyam, B., Zhu, D. G., Feldman, R. I., Arnaiz, D. O. (**2007):** Indolinone based phosphoinositide-dependent kinase-1 (PDK1) inhibitors. Part 2: Optimization of BX-517. Bioorg. Med. Chem. Lett., 17:3819–3825.

Jain, A.N. **(2010)**. QMOD: Physically Meaningful QSAR. J Computer-Aided Molecular Design, 24:865-878

Kim, J. A. (**2003):** Targeted therapies for the treatment of cancer. Am. J. Surg., 186: 264–268.

Komander, D., Kular, G. S., Bain, J., Elliott, M., Alessi, D. R., Van Aalten, D. M. F. (**2003):** Structural basis for UCN-01 (7-hydroxy staurosporine) specifically and PDK1 (3-phosphoinositide dependent proteinkinase-1) inhibition. Biochem. J., 375: 255–262.

Lawlor, M. A., Mora, A., Ashby, P. R., Williams, M. R., Murraytrait, V., Malone, L., Prescott, A. R., Lucocq, J. M., Alessi, D. R. (**2002):** Essential role of PDK1 in regulating cell size and development in mice. EMBO J., 21: 3728–3738.

Liang, K., Lu, Y., Li, X., Zeng, X., Glazer, R. I., Mills, G. B., Fan, Z. (**2006):** Differential roles of phosphoinositide-dependent protein kinase-1 and Akt1 expression and phosphorylation in breast cancer cell resistance to plaxitaxel, doxorubicin and gemcitabine. Mol. Pharmacol., 70: 1045–1052.

Lin, H.-J.; Hsieh, F.-C.; Song, H.; Lin, J. (**2005):** Elevated phosphorylation and activation of PDK-1/AKT pathway in human breast cancer. Br. J. Cancer., 93: 1372–1381.

Medina, J. R., Becker, C. J., Blackledge, C.W., Duquenne, C, Feng, Y, Grant, S. W., Heerding, D, Li, W.H., Miller, W.H., Romeril, S.P., Scherzer, D, Shu, A, Bobko, M. A., Chadderton, A.R, Dumble, M, Gardiner, C. M., Gilbert, S, Liu, Q, Rabindran, S, K., Sudakin, V, Xiang, H, Brady, P.G., Campobasso, N, Ward, P, Axten, J. M. **(2011)** Structure-Based Design of Potent and Selective 3-Phosphoinositide-Dependent Kinase-1 (PDK1) Inhibitors. J. Med. Chem., 54:1871–1895.

Mora, A., Komander, D., van Aalten, D. M. F., Alessi, D. R. (**2004):** PDK1, the master regulator of AGC kinase signal transduction. Semin. Cell DeV. Biol., 15: 161–170.

Toker, A., Newton, A. C. **(2000):** Cellular signaling: Pivoting around PDK-1. Cell., 103: 185–188.

Tropsha, A, **(2010)** Best Practices for QSAR Model Development, Validation, and Exploitation, Mol. Inf. 29:476 – 488

Watts, K.S.; Dalal, P.; Murphy, R.B.; Sherman, W.; Friesner, R.A.; Shelley, J.C., (**2010)**"ConfGen: A Conformational Search Method for Efficient Generation of Bioactive Conformers," J.Chem. Inf. Model., 50, 534-546

Xie, Z., Yuan, H., Yin, Y., Zeng, X., Bai, R., Glazer, R. I. (**2006):** 3-Phosphoinositide-dependent protein kinase-1 (PDK1) promotes invasion and activation of matrix metalloproteinases. BMC Cancer. 6: 77.

Zeng, X., Xu, H., Glazer, R. I. (**2002):** Transformation of mammary epithelial cells by 3-phosphoinositide dependent protein kinase-1 (PDK-1) is associated with the induction of protein kinase C. Cancer Res.,62: 3538–3543.

Zhang, Q., Thomas, S. M., Lui, V. W. Y., Xi, S., Siegfried, J. M., Fan, H., Smithgall, T. E., Mills, G. B., Grandis, J. R. (**2006):** Phosphorylation of TNF-R converting enzyme by gastrin-releasing peptide induces amphiregulin release and EGF receptor activation. Proc. Natl. Acad.Sci. U.S.A., 103: 6901–6906.