

Discovery of Some Potent PDHK Inhibitors Using Pharmacophore Modeling, Virtual Screening and Molecular Docking Studies

Neeta Azad, Mamta Bhandari, Rita Kakkar

Neeta Azad, Mamta Bhandari, Rita Kakkar, Correspondence to: Prof. Rita Kakkar, Computational Chemistry Group, Department of Chemistry, University of Delhi, Delhi-110007, India

Correspondence to: Rita Kakkar, Computational Chemistry Group, Department of Chemistry, University of Delhi, Delhi-110007, India.

Email: rkakkar@chemistry.du.ac.in

Telephone: +911127666313

Received: June 5, 2015

Revised: December 23, 2015

Accepted: December 30, 2015

Published online: March 15, 2016

© 2016 ACT. All rights reserved.

Key words: PDHK; Pharmacophore; Virtual screening; Docking; ADME

Azad N, Bhandari M, Kakkar R. Discovery of some potent PDHK inhibitors using pharmacophore modeling, virtual screening and molecular docking studies. *Journal of Biochemistry and Molecular Biology Research* 2016; 2(1): 139-151 Available from: URL: <http://www.ghrnet.org/index.php/jbmr/article/view/1239>

ABSTRACT

A pharmacophore model has been generated and 3D-QSAR studies performed on a set of fifty fluorinated compounds having known activities for PDHK inhibition (IC_{50} values ranging from $0.1 \mu\text{mol L}^{-1}$ to $100.0 \mu\text{mol L}^{-1}$). Common pharmacophore hypotheses have been developed and a predictive atom based 3D-QSAR model is further generated for all the surviving hypotheses. All the statistical parameters calculated for the obtained model (AHH.2) have been found to be satisfactory. The model obtained suggests that the presence of hydrophobic groups and ability to form hydrogen bonds are the two vital features which play significant roles in the activity of the selected ligands. The binding site of the protein is hydrophobic in nature and hence the presence of -F in the binding compounds enhances their inhibiting property. An excellent correlation between the calculated and experimental IC_{50} values of the compounds has been obtained. Virtual screening of a database comprising 2924 commercially available drug molecules has been performed to ensure the validity of the model developed. A total of 1263 compounds, which show good fit to the model, have been docked to the PDHK protein for further study of their binding affinities and interaction with the protein. A few of these are short-listed for further evaluation.

INTRODUCTION

The pyruvate dehydrogenase complex (PDC) plays an essential role in catalyzing the key step of vital metabolic processes such as oxidation of glucose or decarboxylation of pyruvate through the highly regulated and synchronized action of multiple enzyme subunits^[1]. PDC consists of multiple copies of three enzymes: pyruvate dehydrogenase (E1), dihydrolipoyl acetyltransferase (E2), and dihydrolipoyl dehydrogenase (E3)^[2]. The complex is regulated via reversible phosphorylation, catalyzed by several specific kinases and phosphatase isozymes. PDHK is the isozyme of the pyruvate dehydrogenase kinase family which inactivates the pyruvate dehydrogenase complex (PDC). Four PDHK isozymes (PDHK1, PDHK2, PDHK3 and PDHK4), which are different in respect of their structure, binding to E2 and tissue distribution, are known for regulating the mammalian PDC^[3-5]. The structure of PDHK consists of two domains: the amino-terminal domain (R domain) and the carboxyl-terminal domain (K domain)^[6]. PDHK is bound to the E2 component^[7].

The reduced activity of PDC affects the formation of acetyl-CoA, which in turn raises the blood glucose level, causing type 2 diabetes. Hence, the PDC activators or PDHK inhibitors can increase the glucose utilization in the body and be effective for the treatment of type 2 diabetes^[8]. Therefore, in order to increase

the activity of PDC, PDHKs need to be inhibited. A number of compounds in this category, namely AZ12 (N-{4-[(ethylamino) sulfonyl]-2-methylphenyl}-3,3,3-trifluoro-2-hydroxy-2-methylpropanamide), Nov3r4-((2,5)-dimethyl-4-[3,3,3-trifluoro-2-hydroxy-2-methylpropanoyl]piperazinyl)carbonylbenzotrile, and AZD7545 ((2R)-N-{4-[4-(dimethylcarbamoyl)phenylsulfonyl]-2-chlorophenyl}-3,3,3-trifluoro-2-hydroxy-2-methylpropanamide), have already been evaluated for their PDHK inhibition activity, and their crystal structures in complexation with PDHK isoforms have also been determined^[9-12].

The rapid increase in the study of biologically active fluorinated compounds in the pharmaceutical field has come up because of a greater understanding of the impact of fluorine on the physical and chemical properties of organic molecules. This trend has been aided by the development of new synthetic methodologies and fluorinated reagents for incorporating fluorine or fluorinated substituents into the desired framework. Fluorinated compounds have shown efficacy as antibacterials, antifungicides, antibiotics and anticancer agents^[13].

Pharmacophore modeling and virtual screening have become popular tools to test the commercially available molecules computationally so that they can further be tested for their potency. In the present work, PDHKs have been computationally targeted for developing specific non-toxic inhibitors based on 9-(trifluoromethyl)-9H-fluorin-9-ol and its analogues. The inhibitory concentrations (IC₅₀) against PDHK for the studied compounds are available in the literature^[14], which enabled us to develop quantitative structure activity relationships (QSAR) to predict their biological activities. The pharmacophore model was developed for the active site of PDHK, followed by 3D-QSAR of the inhibitors. The developed pharmacophore model was then used to virtually screen a library of known and commercially available drug molecules to find out more potent and non-toxic PDHK inhibitors. The chosen chemical database has been successfully screened to identify hits through computer-aided screening or virtual screening, which reduces the time and cost of research effectively^[15-19]. The intensity with which a particular molecule can inhibit PDHK2 has been judged on the basis of two important criteria: fitness to the pharmacophore and 3D-QSAR model AHH.2 and good binding affinity within the PDHK active site. Finally, the screened compounds have been docked into the protein and Prime-MM/GBSA calculations performed to evaluate their binding energies.

COMPUTATIONAL DETAILS

A dataset comprising fifty compounds having a wide range of activity (IC₅₀ values ranging from 0.1 $\mu\text{mol L}^{-1}$ to 100.0 $\mu\text{mol L}^{-1}$) was selected from the literature^[14]. In order to build a pharmacophore model, the module PHASE (Pharmacophore Alignment and Scoring Engine) developed by Schrödinger, Inc. was employed. The selected compounds are analogues of 9-(trifluoromethyl)-9H-fluorin-9-ol, and are listed in Table S1 (Supplementary Information, SI). To normalize the inhibitory concentration distribution, all the IC₅₀ values were converted to pIC₅₀ using the relation $\text{pIC}_{50} = -\log_{10}(\text{IC}_{50})$. An activity threshold (in terms of pIC₅₀ values) was set for these molecules so that we could distinguish the active ones from the inactive. This was done using the “Bayes classification” application of the Canvas 1.2 program of Schrödinger. The fifty compounds were divided into three classes (Table S2, SI). Class I, with an activity range of 4 to 4.5, comprises nine compounds, tagged as inactive. Class II, with an activity range of 4.5 to 5.4, contains 21 moderately active compounds and Class III, having an activity range of 5.4 to 6.6, contains 20

active compounds.

The geometries of all compounds were refined using LigPrep. Further details^[20-22] are given in the Supplementary Information. The conformers for each molecule were generated by running a mixed MCMM/LMOD (Mixed Monte Carlo Multiple Minimum/Low Mode) search^[21] with a distance-dependent dielectric solvation treatment and the OPLS-2005 force-field. For each molecule, the conformers with a maximum energy difference of 10 kcal mol⁻¹ relative to the global energy minimum conformer were retained. All these states of the compounds were used for pharmacophore model generation and subsequent scoring^[22].

Common pharmacophoric features (CPH) were then identified from a set of variants. When we tried to find common pharmacophoric hypotheses for 3 - 7 variant lists, the 5- 7 variant lists did not yield any common pharmacophore hypotheses. Only 3 and 4 variant lists yielded common pharmacophore hypotheses. These CPHs were then examined using a scoring function in the “Score Hypotheses” panel to yield the best alignment of the active ligands^[23].

We performed atom-based 3D-QSAR, as it takes into account the entire molecular structure^[23-26]. We evaluated the best scoring hypothesis^[27,28] by generating training and test sets using K-means clustering^[29,30]. The Canvas 1.2 module of Schrödinger was used for this purpose. We divided the set of fifty molecules into three clusters given in Table S3 (SI). The clusters were created from the “atom-pairs” type binary fingerprints of the structures. Only “atom-pairs” type binary fingerprints were able to divide the molecules according to their structural features as well as biological activity. After this, the first molecule of Cluster 1 was selected, followed by every third molecule (30% selection). This selection was continued till the last cluster (Cluster 3). We obtained 18 molecules for our test set and the remaining 32 molecules made up our training set. The PLS regression was carried out with a maximum number of four PLS factors to avoid over-fitting^[22,23].

For virtual screening, the drug database comprising a collection of 2924 drugs was lifted from ZINC (www.zincdocking.org)^[31,32]. Some properties of the 2924 molecules are reported in Table S4 (SI). The energies of these compounds were minimized via the multiple minimization options available in the MacroModel application of Schrödinger Inc. using the OPLS-2005 force-field and force-field defined electrostatic treatment. Using LigPrep, the ionization, tautomeric and stereoisomeric states for each of these molecules were generated before screening them through our pharmacophore model. A total of 4893 states were obtained for the 2924 molecules at physiological pH 7±2. The Lipinski’s filter was applied on these 4893 states, leaving 2852 compounds^[33,34].

As different conformers fit the pharmacophore and 3D-QSAR model differently^[35], we generated conformations for all 2852 molecules using the ConfGen application of Schrödinger using the OPLS-2005 force-field and distance-dependent electrostatic treatment with a dielectric constant of 4.0. The search mode was rapid and an energy window of 100 kJ mol⁻¹ was used for saving conformations. A redundancy check with RMSD of 1.0 Å in the heavy atom positions was applied to remove duplicate conformers. A total of 37,594 conformers were produced for all states of the 2924 compounds. Pharmacophore-based screening was then performed on these compound states using the obtained model. Each representative was matched on all the points, i.e. acceptor and hydrophobic sites. A total of 1263 compounds were obtained as hits along with their fitness score to the model.

For further verification of the inhibiting capabilities of the hits obtained, a ligand library comprising the top 859 compounds was

docked into QM/MM treated PDHK2. First of all, the protein structure was refined using the protein preparation workflow. The bond orders of the residues were adjusted. The missing loops of proteins were filled using the “Fill loops” option of “Protein preparation wizard”. The pre co-crystallized water molecules beyond 5 Å of the active site were removed and hydrogen atoms were added to the structures. The protein structures were first optimized using a molecular mechanics calculation using the OPLS-2005 force-field^[36]. Minimizations were performed until the average root mean square deviation (RMSD) of the non-hydrogen atoms reached 0.3 Å. Additional states for the inhibitors within the proteins were generated (using “Generate Het States”) and the most appropriate states for all the ligands were selected. The refinement of the side chains was done using “Prime side chain refinement” option. We have applied QM on the active site and MM on rest of the prepared protein using QSite. The ligand and the active site residues were treated using density functional theory (DFT), and the rest of the protein was treated by the OPLS-2005 force-field. The B3LYP/LACVP* basis set was used to treat the active site residues. LACVP* employs the 6-31G* basis set for non-transition elements in the active site.

After ensuring that the protein and ligands were in the correct form for docking, the receptor-grid files were generated using a grid-receptor generation program. To soften the potential for the nonpolar parts of the receptor, we scaled the van der Waals radii of receptor atoms by 1.00 with a partial atomic charge of 0.25. A grid box of size 56×56×56 Å³ with coordinates X = 56.09 Å, Y = 62.66 Å and Z = 78.55 Å was generated at the centroid of the ligand. The ligands were docked into the active site first using Glide-SP (Glide “Standard Precision”) and then Glide-XP (Glide “Extra Precision”) for the top 100 Glide-SP scorers. The OPLS-2005 force field was used for this purpose. The choice of the best docked structure for each ligand was made using the model energy score (E_{model}) that combines the Glide score, the nonbonded interaction energy and the excess internal

energy of the generated ligand conformation.

The Gibbs energy of binding between the receptor and ligand was predicted using Prime MM/GBSA. MM/GBSA is a method that combines OPLS molecular mechanics energies (E_{MM}), surface generalized Born solvation model for polar solvation (GSB), and a nonpolar solvation term (G_{NP}). The G_{NP} term comprises the nonpolar solvent accessible surface area and van der Waals interactions. The total Gibbs energy of binding is calculated as:

$$\Delta G_{binding} = G_{protein} - (G_{protein} + G_{ligand})$$

$$G = E_{MM} + G_{SB} + G_{NP}$$

The QikProp program^[37] was used to obtain the absorption, distribution, metabolism and excretion (ADME) properties. A total of forty-four properties were predicted for 102 molecules with the highest Glide-SP scores given in Tables S5-S7 (SI).

RESULTS AND DISCUSSION

PHARMACOPHORE BASED 3D-QSAR MODEL

A tree-based partitioning algorithm was used, which resulted in 93 three-featured and 34 four featured probable common pharmacophore

Table 1 Survival and “Survival-inactive” scores of some selected three-featured hypotheses.

Hypothesis	Survival score	Survival-inactive score
AHH.1	3.695	3.370
AHH.2	3.692	3.373
DHH.2	3.674	3.357
DHH.4	3.644	3.325
HHR.7	3.588	3.262
DHH.5	3.577	3.262
HHR.5	3.571	3.249
DHH.6	3.569	3.248
AHH.3	3.521	3.268
HHR.6	3.497	3.180
HHR.8	3.251	3.000

Table 2 Atom-based QSAR statistics for three-featured hypotheses.

Hypothesis	SD*	R ²	F	p	RMSE [#]	Q ²	Pearson-R
AHH.1	0.2158	0.9410	35.90	1.54×10 ⁻⁵	0.3373	0.6198	0.8468
AHH.2	0.2011	0.9348	50.20	3.77×10 ⁻⁸	0.3004	0.6170	0.9401
DHH.2	0.1600	0.9615	68.60	1.05×10 ⁻⁷	0.3773	0.6130	0.7976
DHH.4	0.2397	0.9060	29.00	4.38×10 ⁻⁶	0.4215	0.5850	0.8182

*Standard Deviation #Root Mean Square Error

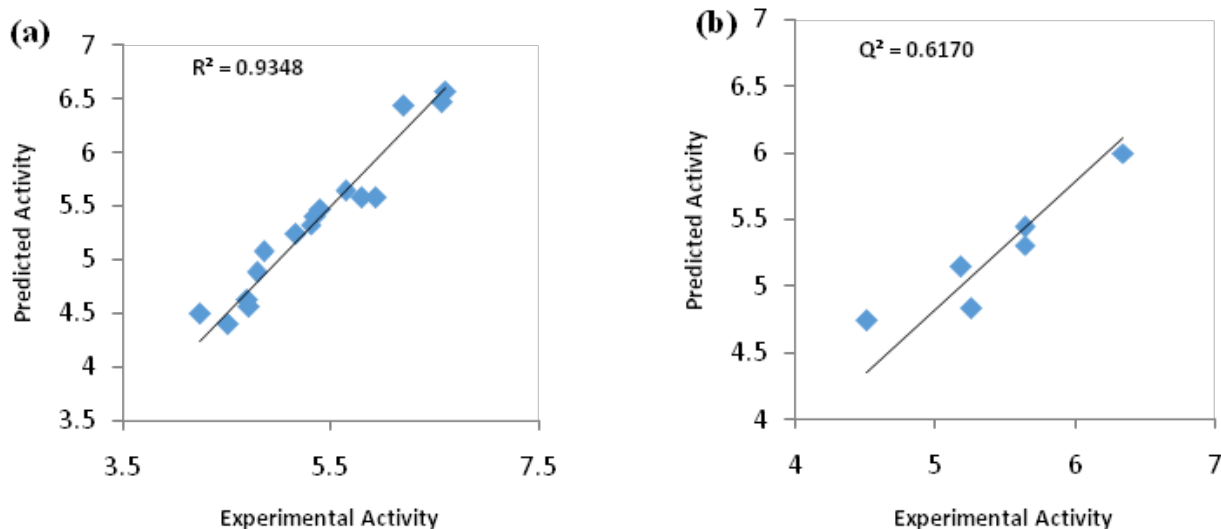


Figure 1 The correlation plots for the experimental activity versus the predicted activity for AHH.2 applied to the (a) training set (b) test set ligands.

hypotheses (CPHs). No CPH was generated for five- and six-featured hypotheses. Amongst these, only 69 three-featured and 19 four-featured CPHs survived the scoring process. We developed the QSAR model for all the three- and four-featured hypotheses which survived the scoring process and analyzed them.

The statistics for the four-featured CPHs was found to be unsatisfactory, as there was found a huge mismatch between the R^2 (regression with the training set molecule) and Q^2 (regression coefficient for the test set molecules) values, which clearly indicates that the developed QSAR model is not acceptable for further study^[22,23]. Also, the Pearson- R values (correlation between the predicted and observed activity for the test set) of three-featured CPHs were found to be higher as compared to those for the four-featured CPHs. Keeping all these observations in mind, we restricted ourselves to the pharmacophore model and 3D-QSAR for three-featured CPHs only. Out of the 93 three-featured variants that survived the scoring process, the CPHs having appreciable survival scores belong to three categories: AHH, DHH and HHR. The eleven CPHs belonging to these three categories, arranged in the order of decreasing values of their survival score, are given in Table 1.

A good hypothesis is one that has high active features and low inactive features. As the survival and survival-inactive scores are all in the range of 3-4, the QSAR model was developed for the top ten hypotheses (Table S5, SI). This selection is based on the criterion of their high survival as well as “survival-inactive” scores.

To select the best hypothesis amongst these, the statistical analysis of all the fifty molecules was carried out for the top four hypotheses. The training and test sets were decided on the basis of K-means clustering method. The training set comprised 32 molecules (64%) and the remaining 18 (36%) molecules formed the test set.

The statistics obtained using cluster analysis (Table 2) point towards AHH.2 as the best of all hypotheses, having appreciably high value of R^2 (>0.6), Q^2 (>0.5), Pearson- R (>0.5) and high F (Table 2). In fact, it has the highest value of Pearson- R . Its p-value is the least. The other hypotheses which performed well are DHH.4, AHH.1 and DHH.2. The plots of the statistical data for the training (Figure 1(a)) and test (Figure 1(b)) sets are linear, with appreciably high regression coefficients, good enough to be accepted for a valid pharmacophore model.

AHH.2-THE PHARMACOPHORE MODEL

The AHH.2 pharmacophore model is shown in Figure 2. It features a hydrogen bond acceptor (A2), having two lone pairs, and two hydrophobic sites (H4 and H5).

On superimposing all the molecules on our proposed model, a good overlap between the molecules and the pharmacophore model was obtained (Figure 3). The basic moiety (9-(trifluoromethyl)-9H-fluorin-9-ol) is the same in all the molecules. The presence of -OH and -CF₃ groups in this moiety indicates the existence of A (acceptor) and H (hydrophobic) features respectively. Another hydrophobic feature (H) comes from the presence of alkyl and other hydrophobic groups in these molecules.

The ligands showing the best and least fitness scores to our model, i.e. ligands C36 and C47, respectively, were selected and then subjected to analysis for finding the favorable and unfavorable sites for substitution (Figure 4). It can be seen that the best fit ligand C36 has its acceptor group (-OH) and hydrophobic group (-CF₃) perfectly aligned with our model, AHH.2. The second hydrophobic group is provided by -Cl. In the worst fit ligand, there is an -OCH₃ group at this hydrophobic site. The middle row of Figure 4 shows the favorable (dark green) and unfavorable (purple) sites for the

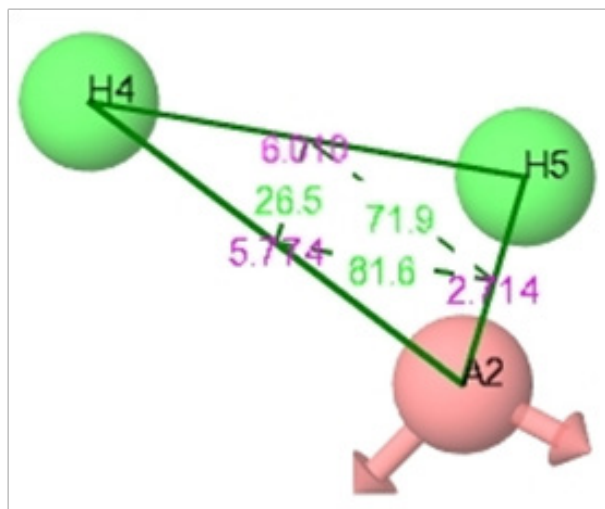


Figure 2 Model AHH.2 representing features, distances (in Å) and angles (in °) between them. Acceptor “A2” is represented as a light red sphere with the lone pair vectors; hydrophobic sites “H4” and “H5” are represented by light green spheres.

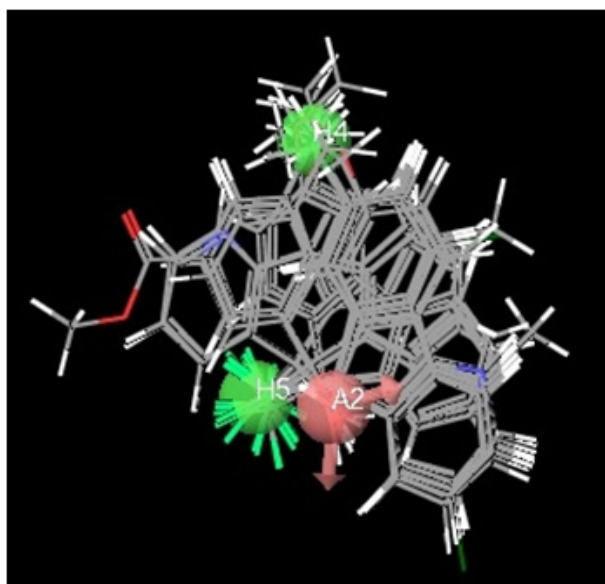


Figure 3 Alignment of all the ligands on AHH.2 showing good overlap of all the ligands on the same features.

substitution of hydrophobic groups. Clearly, there are many more unfavorable sites in C47. In fact, the presence of -OCH₃ makes this site unfavorable, in contradiction to our model, which requires a hydrophobic group at this position. In case of electron withdrawing group substitution, ligand C36 (Figure 4, bottom left) shows the region in the form of orange cubes where substitution of electron withdrawing groups possibly enhances the inhibiting activity.

Using these guidelines, we inserted hydrophobic groups at their preferred site in C36 in order to enhance its activity, and similarly we attached electron withdrawing substituents at the pyridine nitrogen position. The result is gratifying in the sense that there is remarkable improvement in the Glide scores from -5.32 for C36 to a maximum value of -7.29 for two compounds, F32 and S24 (Table S6). However, there is only a marginal improvement in the Glide Score, as C16, the highest Glide Scorer in the original compounds, already had scored -7.23 (Table S1). The top 10 hydrophobic substituents (Table S7) are arylalkyl halides with F or Cl, except one compound, which contains Br. We then attached electron-withdrawing substituents at

the suggested position, but this did not have such a profound effect, only four compounds doing better than the parent compound (Table S8). The highest Glide score is -6.11 for -OH substitution (Tables S8 and S9). The results show that hydrophobic substitution has a more profound effect than electron-withdrawing group substitution.

We then combined the four high scoring electron withdrawing substituents with the 69 high scoring hydrophobic groups which had better scores than C36 to obtain 276 compounds. For these compounds, we again computed the Glide scores and Gibbs binding energies (Table S10). A huge enhancement in the Glide Score to -9.76 is obtained for the best scoring compound. The top twenty molecules have Glide Scores above -8.25. Their Emodel values are also impressive. The structures of the top 15 Glide-XP scorers

are displayed in Table S11. The top Glide-XP scorer L1 contains hydrophobic group F40 (Table S6) and electron withdrawing group E2 (Table S8). F40 is an arylalkyl group ((1,1-difluoro-2-methylpropan-2-yl)benzene) and the E2 group is an aldehyde group (propanal).

To further calculate their fitness to the pharmacophore model, the ligands obtained after hydrophobic and electron withdrawing substitution were screened through the pharmacophore model AHH.2. The fitness scores of hits obtained are reported in Table S11.

These 276 compounds were then subjected to calculation of their ADME properties using QikProp. Out of these, 37 compounds failed. The ADME properties and descriptors of the remaining 239 compounds are given in Tables S12 and S13. One can see that all

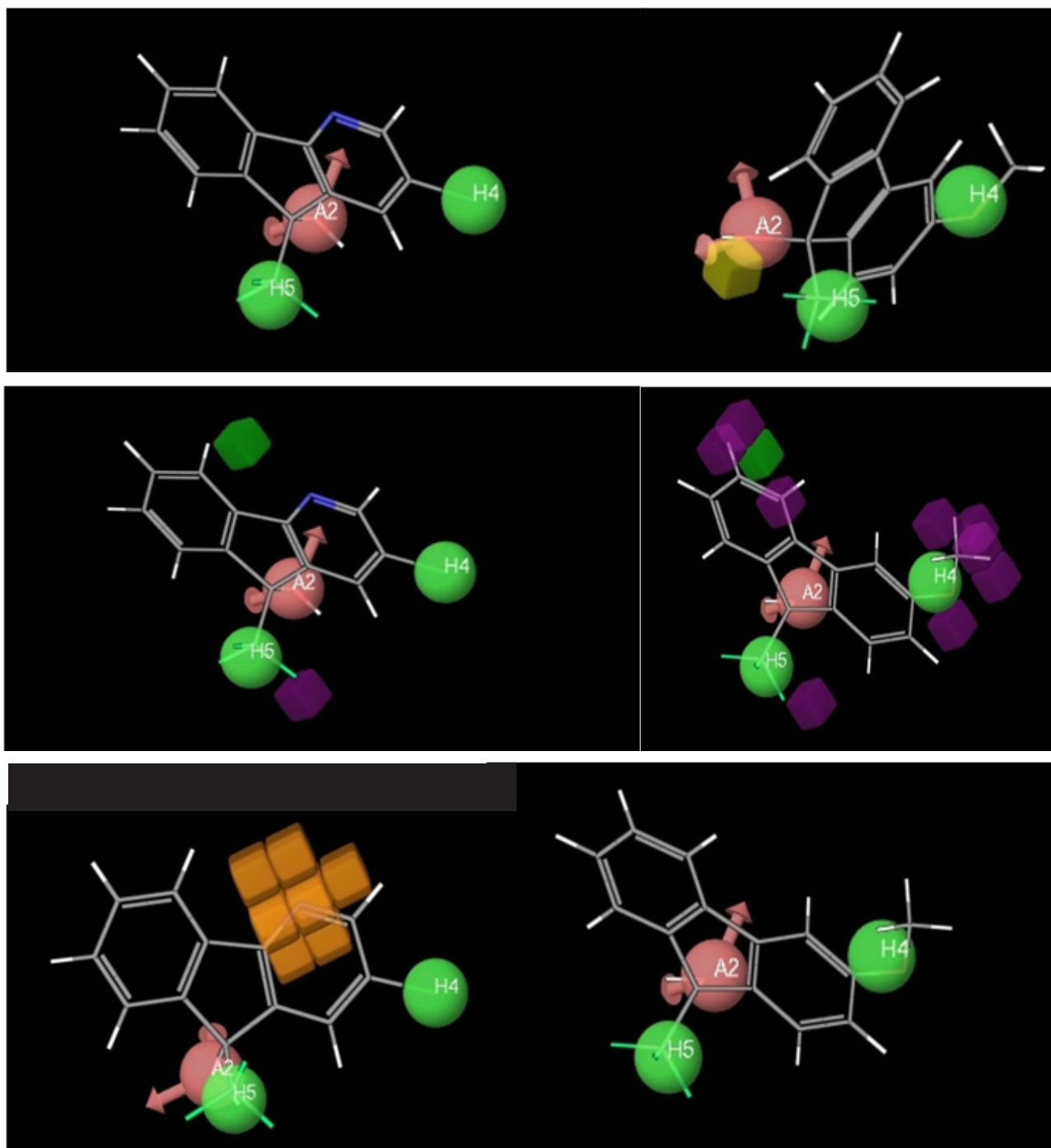


Figure 4 The favorable and unfavorable regions for H-bond donors (top), hydrophobic (middle) and electron withdrawing groups (bottom) for the best fit ligand C36 (left) and the worst fit ligand C47 (right).

the compounds violate at least one of Lipinski's rules which define the drug likelihood of a compound to be orally active. Inclusion of both hydrophobic and electron withdrawing groups increases the molecular weight to above 500, violating one of the conditions for the drug likelihood of a compound. The other rule that is violated by all these compounds is that the logPo/w should be lower than 5. Introducing another hydrophobic group increases the lipophilicity to more than the required value. Ideally a drug should have a balance between its solubility in nonpolar solvents like 1-octanol (so that it can pass through cell membranes, which consist of a phospholipid bilayer, into cells) and in water (so that it can be transported around the body through the bloodstream). Clearly, these molecules do not satisfy the latter criterion.

Though one violation of Lipinski's rules is acceptable, we decided to go for a search for known drugs that fit our pharmacophore model and have suitable ADME properties. Moreover, we would like higher Glide scores and hence we performed virtual screening to explore the possibility.

VIRTUAL SCREENING

The main aim of virtual screening is to propose new and potent inhibitors from a library of already recognized drugs. The two key features for a ligand to be a good inhibitor are its fitness to the pharmacophore model and its binding affinity with the protein binding site. Our model AHH.2 showed satisfactory internal ($R^2 = 0.935$) and external predictivity ($Q^2 = 0.617$), which implies that we can screen a library of known and commercially available drugs to find more potent and novel PDHK inhibitors.

In order to find the hits, a library of 37,594 molecules was first generated and then screened using the pharmacophore model AHH.2. 1263 compounds in all were obtained as hits. The maximum value of their fitness score is 2.31 (77% match with ligand C36) whereas the lowest fitness score value is 0.13 (4% match). Table 3 analyses the distribution of ligands in the various fitness score ranges.

From Table 3, it is clear that out of the 1263 molecules, 991 (79%) show fitness scores above 1.5, i.e. 79% of the total molecules show more than 50% match to our pharmacophore model. This further validates the quality of our model AHH.2.

GLIDE DOCKING

After screening these compounds via model AHH.2, the next step was their docking to the protein binding site. Glide-SP (Glide "Standard Precision") docking of the 1263 molecules was performed on the QM/MM treated PDHK2. The docking scores comprise van der Waals, coulombic and hydrophobic interactions, and so docking is an effective technique to further defend the acceptability of pharmacophore model fitted molecules. Fitness with the model alone is not sufficient to ensure good inhibition of kinase proteins.

From the Glide-SP scores, it was observed that not all the molecules with good fitness scores exhibit good binding with PDHK2. The Glide-SP docked ligands falling in different ranges of their docking scores are displayed in Table 4. The number of ligands having fitness scores ≥ 1.5 falling in a particular range of docking scores is also mentioned there. The maximum value of the Glide-SP docking score is -8.99 and the lowest is -0.89.

The top 102 (with FS > 1.5) ligands were selected for Glide-XP docking. All these ligands have their Glide-SP docking score in the range of -8.99 to -7.00. The complete data for Glide-XP docking of these 102 ligands is given in Table S14 (SI). The structures of the top ten scorers of Glide-XP docking are given in Table 5, along with their docking scores.

Table 3 Analysis of the distribution of ligands according to their fitness scores.

Fitness Score (F.S.) Range	No. of Ligands
$0.0 \leq F.S. \leq 0.50$	14
$0.50 \leq F.S. \leq 1.00$	31
$1.0 \leq F.S. \leq 1.50$	227
$1.5 \leq F.S. \leq 2.00$	858
$2.00 \leq F.S. \leq 2.5$	133
	1263

Table 4 Analysis of the distribution of ligands according to their Glide-SP scores, and the number of molecules with fitness scores ≥ 1.50 for each range of Glide score.

Glide-SP Score (GS)	No. of Ligands	No. of Ligands with fitness score ≥ 1.5
$-9.00 < GS \leq -8.00$	18	17
$-8.00 < GS \leq -7.00$	99	83
$-7.00 < GS \leq -6.00$	454	345
$-6.00 < GS \leq -5.00$	489	376
$-5.00 < GS \leq -4.00$	160	138
$-4.00 < GS \leq -3.00$	30	26
$-3.00 < GS \leq -2.00$	07	06
$-2.00 < GS \leq -1.00$	03	00
$-1.00 < GS \leq 0.00$	01	00
	1261*	991

From the analysis of the top ten Glide-XP scorers, it can be clearly seen that the top five positions in the scoring process are taken up by different conformers of the same compound 'Doxifluridine', which show the best Glide Scores, which improve to more than -10 and fitness score ≥ 1.5 (Table 6). The ligand has fluorine substitution, which enhances its chances of hydrogen bond formation with the residues in the protein binding site, which is very essential for the stability of the protein-ligand complex. From all respects, the various conformers of doxifluridine, with Glide scores way above the rest, appear to be the best candidates. Another suitable candidate is Z9, which has the highest value of E_{model} (Table 6), which is a combination of Glide score, the non-bonded interaction energies and the excess internal energy of the conformer. Z9 also has the best value of Glide Energy.

The ligand Z1 (doxifluridine), as shown in Figure 5, forms three hydrogen bonds with the receptor protein, one between the $-C=O$ group of the ligand and a water molecule trapped inside the protein cavity. The second hydrogen bond is formed between the $-OH$ group present at one end of the ligand and Leu160, and the third between the $-NH$ present in the ring at the other end of the ligand and Gln27. Ligands Z2, Z3, Z4 and Z5 also show good docking results. They all are conformers of doxifluridine; hence, they are also good binders to the protein. The major part of the ligand is embedded in the hydrophobic region of the binding cavity of the protein, which shows that the ligand also has a hydrophobic part, which is helpful in the binding of the ligand to the protein, as predicted by our pharmacophore model, AHH.2.

It has already been reported that the binding site of PDHK2 is hydrophobic in nature^[11] and the same was observed from our site map calculations, performed using Qsite. The hydrophobic part of the PDHK2 binding site has been displayed in orange color in Figure 5 (right). The ligand Z1 is totally surrounded by the hydrophobic surface, and forms three hydrogen bonds with the receptor residues.

Ligand Z9 forms one hydrogen bond with the residue Gln27 and does not completely lie inside the protein (Figure6) like other small ligands, but partially peeps out of the cavity. However, the part which is responsible for binding is completely embedded inside the hydrophobic region. All these observations point towards the fact that hydrophobicity is an important criterion for the ligand to bind firmly to the binding site of PDHK2

Table 5 Top ten Glide-XP scorers, along with their popular names, structures and total charge (Q).

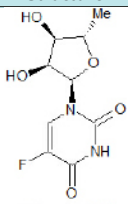
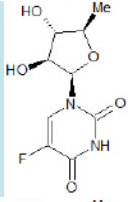
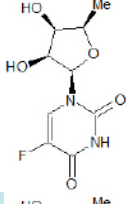
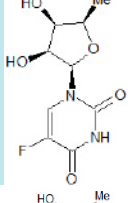
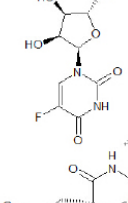
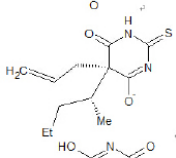
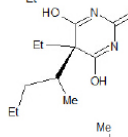
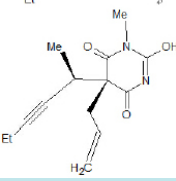
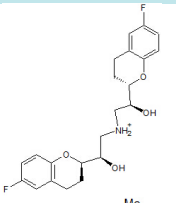
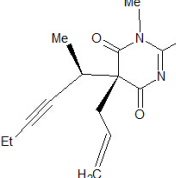
Ligand No.	Title	Common Name	Structure	Glide Score	Q
Z1	ZINC03830726	Doxifluridine		-10.42	0
Z2	ZINC03830727	Doxifluridine		-10.38	0
Z3	ZINC03830725	Doxifluridine		-10.30	0
Z4	ZINC03830725	Doxifluridine		-10.10	0
Z5	ZINC03830726	Doxifluridine		-10.02	0
Z6	ZINC17747847	Thiamylal sodium		-9.77	-1
Z7	ZINC13545636	Pentobarbital		-9.69	0
Z8	ZINC03876069	Methohexital		-9.58	0
Z9	ZINC05844788	Nebivololol		-9.55	1
Z10	ZINC03876069	Methohexital		-9.55	0

Table 6 Glide-XP docking data* for the top 10 virtually screened drug molecules.

L	FS	GS	E_{model}	E_{vdw}	E_{coul}	GE	$E_{internal}$	$hbond$	Lipo
Z1	1.72	-10.42	-58.71	-32.25	-7.56	-39.81	1.70	-0.48	-1.74
Z2	1.65	-10.38	-52.79	-25.60	-9.85	-35.45	7.13	-0.50	-1.49
Z3	2.01	-10.30	-54.49	-31.33	-8.32	-39.65	2.64	-0.45	-1.74
Z4	2.01	-10.10	-51.58	-28.28	-8.38	-36.66	2.49	-0.44	-1.73
Z5	1.72	-10.02	-59.23	-32.13	-7.60	-39.73	1.40	-0.49	-1.74
Z6	2.05	-9.77	-36.74	-22.60	-6.67	-29.28	13.45	-0.61	-2.04
Z7	2.15	-9.69	-44.30	-27.26	-5.78	-33.04	5.86	-0.32	-2.02
Z8	1.72	-9.58	-47.06	-28.54	-6.05	-34.59	5.53	-0.61	-2.27
Z9	1.97	-9.55	-64.87	-33.37	-10.25	-43.62	7.61	0.00	-2.92
Z10	1.73	-9.55	-47.58	-28.34	-6.43	-34.77	4.40	-0.61	-2.25

*Emodeland Einternalare in kcal mol⁻¹; FS = Fitness Score; GE = Glide Energy

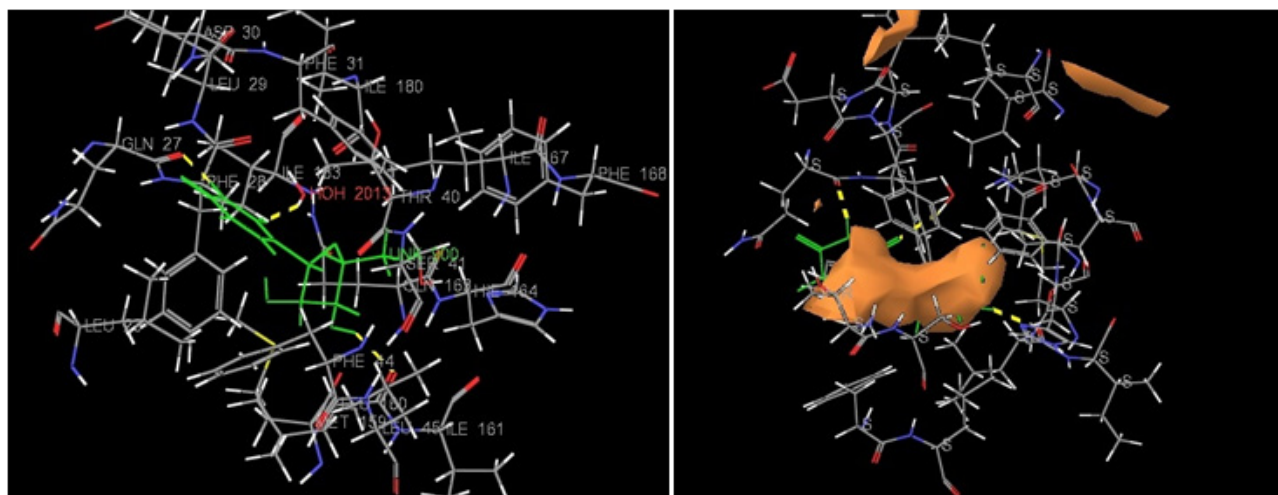


Figure 5 Ligand Z1 shown in green color within the PDHK active site forming three hydrogen bonds with the residues Leu160, Gln27 and H2O (left) and the binding hydrophobic region of the protein containing ligand is shown in orange color surface (right).

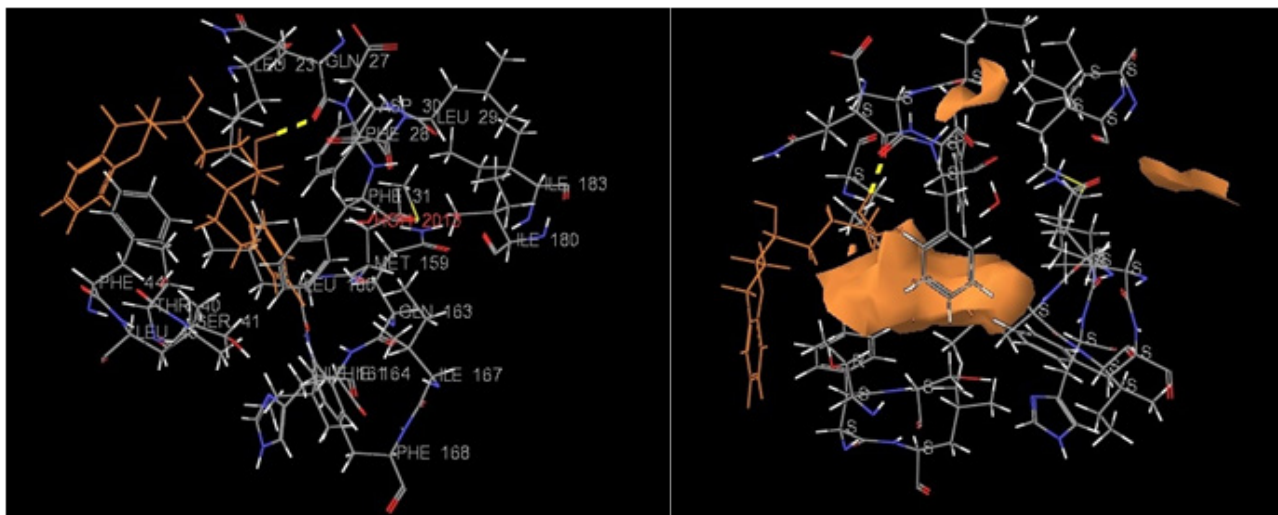


Figure 6 Ligand Z9 (orange color) in the binding site of PDHK2, forming one hydrogen bond with the receptor (left), the hydrophobic region of the protein containing ligand Z9.

ADME SCREENING

The ADME properties were calculated for the Glide-XP docked top 102 ligands. This was done using the QikProp module of Schrodinger Inc. A total of forty-four properties were calculated for all the 102 ligands and analyzed to check whether they fall in the acceptable range (Tables S15 - S17, SI) for drugs.

It is observed that none of these shows violation of Lipinski's rules. The top ten ligands reported here (Table 7) have their molecular weights in the range 200 - 424 g mol⁻¹, which is within the 130-725 range of most drug molecules. Polarizability (Polrz) is also within the range (13 - 70 Å³) of most drugs, but Z9 has a value

close to the middle of the range. The values of logPC16 and logPoct are also in the middle of the range for Z9, but are low for the others. The water/gas partition coefficient (logPw) for all the molecules is in the low side of the acceptable range (4-45). The aqueous solubility (logS) and conformation-independent predicted aqueous solubility (CIlogS) are within the acceptable ranges for most drugs, implying smooth distribution of the drugs in the body. The computed logPo/w values are a measure of the lipophilicity of a molecule, and these are frequently used to estimate the membrane permeability and the bioavailability of compounds. Z1 is slightly hydrophilic, but Z9 is lipophilic.

Table 7 Predicted ADME properties of the compounds.

L	Mol.Wt.	Polrz	logPC16	logPoct	logPw	logPo/w	logs	ClogS
Z1	246.20	21.06	6.91	16.83	14.75	-0.86	-1.94	-1.58
Z2	246.20	21.09	6.98	17.03	14.79	-0.94	-1.95	-1.58
Z3	246.20	20.27	6.82	16.61	14.62	-0.90	-1.77	-1.58
Z4	246.20	20.27	6.82	16.61	14.62	-0.90	-1.77	-1.58
Z5	246.20	21.06	6.91	16.83	14.75	-0.86	-1.94	-1.58
Z6	254.35	25.48	8.28	13.20	7.19	3.27	-3.62	-3.21
Z7	226.28	20.92	6.97	12.23	7.58	2.05	-2.42	-2.52
Z8	262.31	27.00	7.80	12.24	6.22	2.92	-3.47	-3.25
Z9	405.44	41.81	12.52	20.98	12.75	3.79	-4.57	-4.53
Z10	262.31	27.00	7.80	12.24	6.22	2.92	-3.47	-3.25

Table 8 Descriptors calculated for top ten ligand states by QikProp simulation.

HERG	Caco	logBB	MDCK	logKp	logKhsa	QP%	PSA	donorHB	metab	acptHB
-3.24	109.69	-1.14	79.63	-5.05	-0.75	58.42	123.42	3.00	3	8.60
-3.21	84.21	-1.24	59.90	-5.29	-0.75	55.90	123.69	3.00	3	8.60
-2.87	119.22	-1.04	87.05	-5.00	-0.74	58.85	119.30	3.00	3	8.60
-2.87	119.22	-1.04	87.05	-5.00	-0.74	58.85	119.30	3.00	3	8.60
-3.24	109.69	-1.14	79.63	-5.05	-0.75	58.42	123.42	3.00	3	8.60
-1.94	256.90	-0.45	399.66	-2.83	-0.13	89.21	77.95	2.00	1	3.50
-2.38	617.11	-0.62	293.60	-3.48	-0.25	88.88	96.72	2.00	0	4.00
-3.86	1388.36	-0.53	705.31	-2.52	0.09	100.00	85.35	1.00	3	4.00
-6.82	408.49	-0.16	680.49	-3.40	0.36	95.85	67.00	3.00	5	6.40
-3.86	1388.36	-0.53	705.31	-2.52	0.09	100.00	85.35	1.00	3	4.00

Table 9 Pearson correlation matrix for the various scoring functions in Schrödinger.

	ΔG_{1bind}	Glide score	E_{model}	Glide energy
ΔG_{1bind}	1			
Glide score	-0.020	1		
E_{model}	-0.008	0.036	1	
Glide energy	0.026	0.026	0.960	1

Table 10 Descriptors calculated for top ten ligand states by QikProp simulation.

L	CNS	WPSA	glob	FOSA	PISA
Z1	-2	44.58	0.90	151.31	23.56
Z2	-2	44.66	0.90	145.29	18.05
Z3	-2	44.50	0.92	146.70	15.43
Z4	-2	44.50	0.92	146.70	15.43
Z5	-2	44.58	0.90	151.31	23.56
Z6	-1	80.48	0.87	278.79	37.27
Z7	0	0.00	0.95	290.63	0.00
Z8	0	0.00	0.87	376.25	50.05
Z9	0	93.99	0.80	270.14	263.71
Z10	0	0.00	0.87	376.25	50.05

In Table 8, the logHERG values are presented. These represent the predicted IC_{50} values for blockage of HERG K^+ channels. Values below -5 are of concern, and Z9 is automatically rejected. Caco-2 cells are a model for the gut-blood barrier and higher values ($> 500 \text{ nm s}^{-1}$) signify good penetration. The log BB values, which represent the brain/ blood partition coefficients, are also within the acceptable range for 95% of known drugs (-3.0 – 1.2). MDCK cells are good mimics for the blood brain barrier and higher MDCK values ($> 500 \text{ nm s}^{-1}$) are considered good. Here Z9 fits the bill, but values below 25 are considered poor; hence, Z1 is also acceptable. Skin permeability (logKp) is also within the acceptable range (-8 - -1). Similarly, logKhsa, prediction of binding to human serum albumin, donor HB and acptHB are all within the acceptable ranges. The PSA values, which represent the van der Waals surface area of polar nitrogen and oxygen atoms and carbonyl carbon atoms, are also within the acceptable range (7 – 200).

We conclude that Z1 is a suitable candidate for further study. There are three metabolic reactions for Z1 and hence this molecule passes all tests.

REGRESSION ANALYSIS

In virtual screening, the major challenge is to decide the scoring

parameters to rank the inhibiting ability of different ligands towards a protein. Hence, the various scoring functions, viz. Glide score, E_{model} , Glide energy, and Prime MM/GBSA Gibbs energy of binding were analyzed in order to see which of them can best explain the biological activity of the ligands. The biological activity of any molecule can be directly correlated to its structural parameters and the scoring function that shows optimum correlation with the ligand structure parameters can also be the best way to describe its biological activity. The top 102 ligands obtained after Glide-XP docking were subjected to the Prime MM/GBSA minimization procedure, and the Gibbs energy of binding ΔG_{1bind}^1 (kcal mol^{-1}) was generated (Table S18). In Table 9, the Pearson correlation matrix is shown, correlating various scoring parameters for the top 102 Glide-XP docked ligands.

From the data in Table 9, we observe that the Glide energy and E_{model} are highly correlated, and any one of them can be selected. The Pearson's correlation coefficients for others are insignificant, which means Glide score, E_{model} , and Prime MM/GBSA are not interrelated and can be individually explored to find out that which one best describes the biological activity. The three scoring functions were then tested for their dependence on ligand based descriptors.

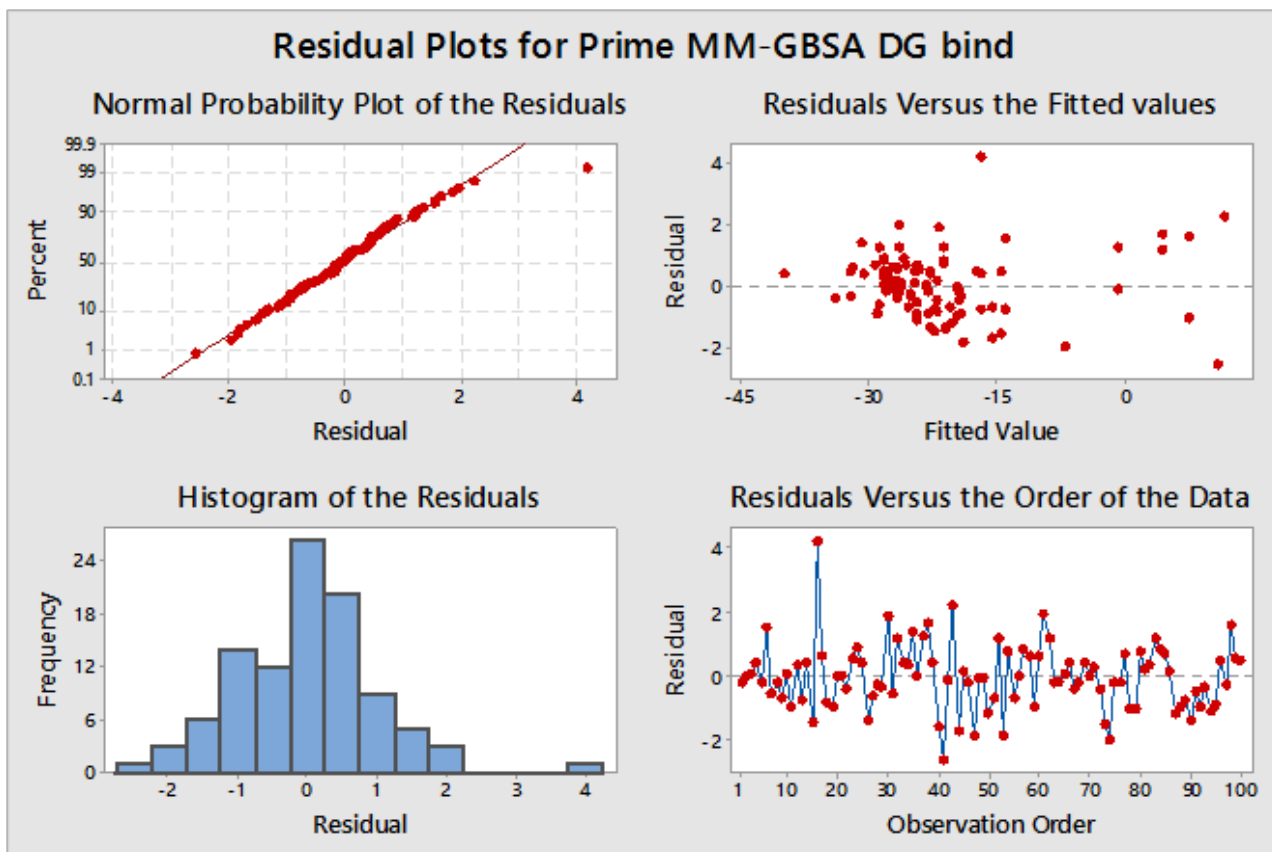


Figure 7 Residual plots for Prime-MM/GBSA Gibbs energy of binding ΔG^{bind} (kcal mol⁻¹).

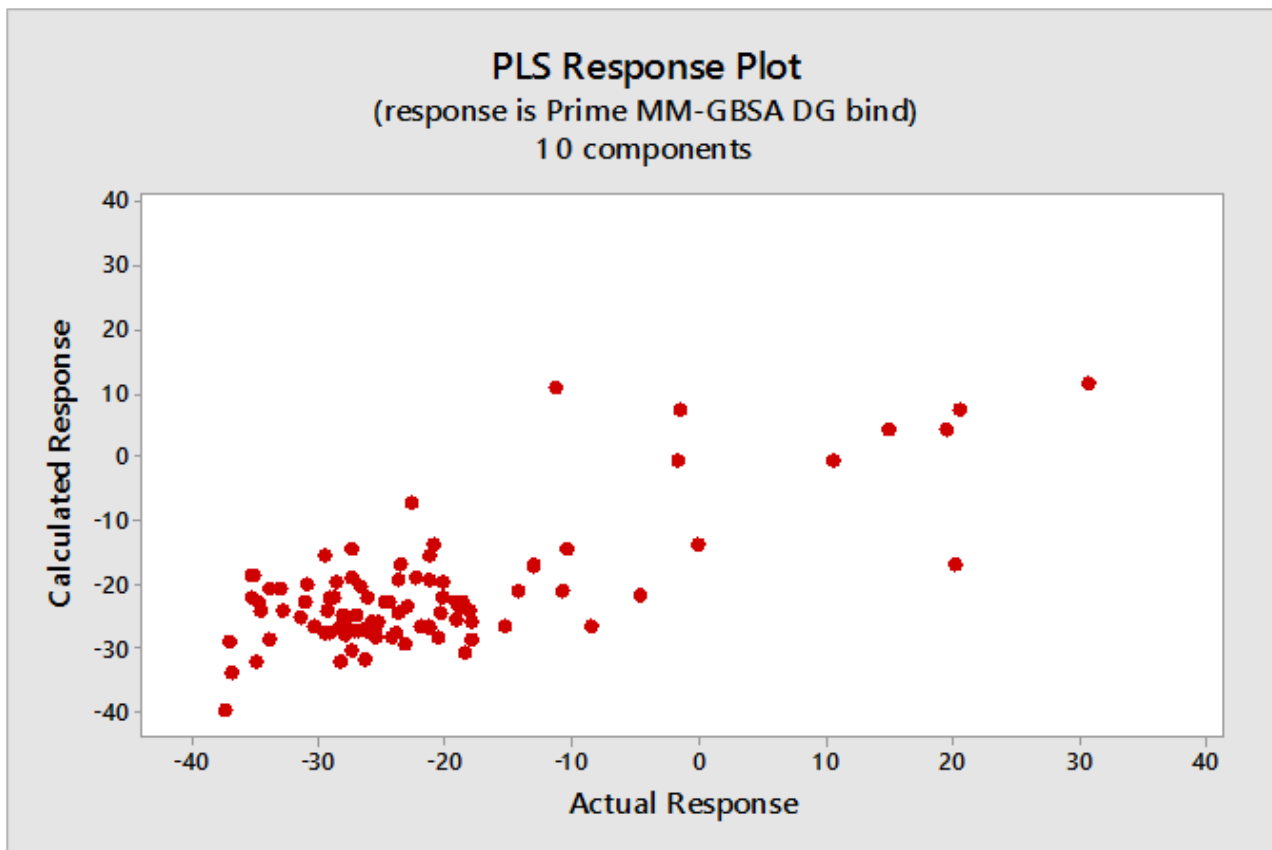


Figure 8 Partial Least Square response plot for Prime-MM/GBSA Gibbs energy of binding ΔG^{bind} (kcal mol⁻¹).

Table 11 Prime-MM/GBSA energy of binding (kcal mol⁻¹) for ligand binding to the receptor.

L	Title	Prime-MM/GBSA ΔG^1_{bind}	Lstrain	Prime-MM/GBSA ΔG^1_{bind}
Z1	ZINC03830726	-32.55	3.11	-29.44
Z2	ZINC03830727	-31.75	5.00	-26.75
Z3	ZINC03830725	-31.08	7.47	-23.61
Z4	ZINC03830725	-33.68	13.39	-20.29
Z5	ZINC03830726	-32.13	3.00	-29.13
Z6	ZINC17747847	-16.14	16.14	0.00
Z7	ZINC13545636	-30.84	1.57	-29.27
Z8	ZINC03876069	-38.73	10.30	-28.44
Z9	ZINC05844788	-39.65	8.32	-31.33
Z10	ZINC03876069	-38.68	12.37	-26.30

All the descriptors calculated via QikProp were considered. The correlation coefficients were calculated for all the descriptors to see their mutual interrelation with each other. If any of the descriptors were found to be strongly inter-correlated (correlation coefficient ≥ 0.9), only one of them was retained. FISA is correlated to PSA and logBB with correlation coefficients 0.905 and -0.923, respectively. LogPC16 is correlated to polrz and logCoct with correlation coefficients 0.963 and 0.903; polrz is also very closely correlated to SASA with correlation coefficient 0.988 and logPw is correlated to acptHB with a correlation coefficient 0.925. Hence, some of these parameters were dropped from the regression analysis.

Stepwise regression analysis for the scoring functions ΔG^1_{bind} , Glide-XP score and E_{model} was performed to see their dependence on the descriptors. The regression coefficients obtained for ΔG^1_{bind} , Glide-XP score and E_{model} were 0.528, 0.328 and 0.732 respectively. Though the regression coefficient is the highest for E_{model} , the *p* value for this correlation is > 0.05 which is considered to be statistically insignificant. Hence we considered ΔG^1_{bind} for further calculations. The Prime-MM/GBSA Gibbs energy of binding is thus the best function that can be related to the structures of ligands and ADME properties. It has been shown in our previous studies too that the biological activity of protein-ligand complexes can be better explained by the Prime-MM/GBSA energy of binding^[38]. The regression equation obtained is:

$$\Delta G^1_{bind} = -225 - 4.56 \text{ donorHB} + 7.60 \text{ logPC16} - 6.88 \text{ CNS} + 0.250 \text{ WPSA} - 6.55 \text{ acptHB} + 208 \text{ glob} - 14.0 \text{ logPo/w} + 0.0964 \text{ FOSA} - 1.75 \text{ logPw} + 0.0777 \text{ PISA} \quad (1)$$

From equation (1), it is very clear that the regression equation obtained for Prime-MM/GBSA contains both structure-based properties and ADME properties. The highest coefficient is for the globularity, which correlates positively with the Gibbs energy of binding. The globularity descriptor is defined as $\text{Glob} = (4\pi r^2)/S_{mol}$, where *r* is the radius of the sphere whose volume is equal to the molecular volume. Smaller values indicate larger deviation from the spherical shape and better binding (more negative Gibbs energy). Hence, size and shape are important factors in determining the binding of the ligand, and the more the ligand differs from a sphere, the better it can bind. The data for the remaining descriptors appearing in equation (1) is presented in Table 10.

Positive values of the log of the octanol-water partition coefficient favor binding, and hence, more lipophilic a ligand, higher is its binding affinity. The presence of hydrogen bond donor and acceptor terms show how significant these factors are for the binding of a drug to the receptor. The coefficients for the hydrogen bond donor and acceptor terms are both negative, which indicates that increase in these terms makes the ΔG^1_{bind} more negative, i.e. the complex becomes more stable. The other terms with reasonably high

Table 12 Predicted pIC₅₀ (μM) of the top ten Glide-scorers for PDHK inhibition.

Ligand	Predicted pIC ₅₀ according to QSAR model AHH.2
Z1	5.31
Z2	4.99
Z3	5.58
Z4	5.58
Z5	5.31
Z6	4.77
Z7	4.81
Z8	4.73
Z9	5.33
Z10	4.68

coefficients are logPC16 (hexadecane/gas partition coefficient) and logPw (water/gas partition coefficient). These are related to the Henry law constants. The positive coefficient of the former and negative of the latter imply that the ligand should have high tendency to partition from the gas phase to hexadecane and small tendency to partition from the gas phase to water, implying that it should be lipophilic. The other three terms are small and positive and imply that WPSA (weakly polar component of the solvent accessible surface area comprising halogens, P and S), FOSA (hydrophobic component of the solvent accessible surface area comprising saturated carbon and hydrogens) and PISA [π (carbon and attached hydrogen) component of the solvent accessible surface area] should be small, but the coefficients of the three terms are relatively small.

The residual plots of the Prime-MM/GBSA Gibbs energy of binding ΔG^1_{bind} are displayed in Figure 7 and the partial least square with 10 components gives a straight line, as shown in Figure 8.

PRIME-MM/GBSA GIBBS ENERGIES OF BINDING

The regression analysis of the various functions in the previous section has shown that the Prime-MM/GBSA binding energy is the best way to explain the biological activity of a ligand towards the protein. The binding energies ΔG^1_{bind} (with ligand strain energy) and ΔG^2_{bind} (without ligand strain energies) are reported in Table 11 for the top 10 Glide-XP scorers. The rest of the values for the top 102 scorers are reported in Table S18 of the Supplementary Information. ΔG^1_{bind} , obtained by adding the strain energy, seems more significant for explaining the ligand binding affinities. In this respect too, Z1 scores the best (Table 11) having competition only from Z9, which we had already rejected because of its lack of metabolic reactions. Moreover, though Z9 shows the strongest binding affinity, there is also a higher strain energy (8.32 kcal mol⁻¹), which implies that this ligand requires more effort to fit into the binding pocket. This reduces the difference in binding energies of Z1 and Z9 to only 1.9 kcal mol⁻¹.

The predicted pIC₅₀ values for the hits, obtained from the AHH.2 model, were calculated using four PLS factors and the pIC₅₀ values for the top 10 Glide-XP scorers are reported in Table 12. The

inhibitory concentration for 50% inhibition is $\sim 5 \times 10^{-6} \mu\text{M}$ for both Z1 and Z9. Though there is no improvement from the original dataset, as the inhibitory concentrations remain in the micromolar range, overall Z1 appears to be a good candidate for further evaluation.

Z1 is doxifluridene (5'-deoxy-5-fluorouridine), a fluoropyrimidine derivative. It is an oral prodrug of the antineoplastic agent 5-fluorouracil (5-FU), which has antitumor activity, but undergoes rapid degradation by dihydropyrimidine dehydrogenase in the gut wall. Doxifluridene is converted into 5-FU in the presence of pyrimidine nucleoside phosphorylase selectively in tumor cells. Doxifluridene has also been shown to be active against exudative age-related macular degeneration (AMD), characterized by choroidal neovascularization (CNV), which is a major cause of visual loss in developed countries^[39].

An interesting observation is that most of the other high Glide scorers (Z6, Thiomyal sodium, Z7, pentobarbital, Z8 and Z10, methohexital) are barbiturates, used for inducing anesthesia. Pentobarbital (Z7) is a barbiturate which slows down the activity of the brain and is used for inducing sleep. Nebivolol (Z9) belongs to a class of medicines known as β -blockers used to treat hypertension. Doxifluridene seems to be the best bet, given that the first five positions in the Glide Score are taken up by different conformations of this drug.

However, a disappointment is that, though we could improve our Glide scores considerably, this did not translate into better pIC_{50} values. None of the new ligands, whether obtained by hydrophobic group and electron-withdrawing substitution or by virtual screening, could better the pIC_{50} values of the parent compounds, though this work could identify the factors responsible for better binding.

CONCLUSION

We selected a series of inhibitors belonging to the 9-(trifluoromethyl)-9H-fluorin-9-ol family to generate a pharmacophore model. The three and four featured hypotheses were generated but only three featured hypotheses were found to be significant. After deciding the training and test sets via Bayes classification, a 3D-QSAR was obtained for each of them and the best model was selected on the basis of the internal and external data set regression analysis. The best regression turned out to be for the AHH.2 model with $R^2 = 0.935$ and $Q^2 = 0.617$. After getting an idea about the favorable sites of ligand C36 for various substitutions, a combinatorial library of 276 compounds was generated by substituting hydrophobic and electron withdrawing groups at suitable positions. This library was then tested via ADME properties calculations but only 239 compounds were found suitable on this scale. Glide-XP docking and Gibbs binding energy calculations were then performed to see their efficacy as inhibitors of PDHK2. Using pharmacophore model AHH.2, their expected activity values were also calculated and analyzed but no appreciable improvement was seen in the activity of C36 after substitution; hence, the alternate was virtual screening of some known drugs in order to get better inhibitors with improved activities. Again, model AHH.2 was used to virtually screen a library comprising of 2924 drug molecules which was already available and tested by ZINC (www.zincdocking.org). Most of the drug molecules in the selected library were non-fluorinated. After generating all possible states for the 2924 molecules at the physiological pH and on applying Lipinski's filter, the final library of 37,594 compounds was obtained. On screening this library via model AHH.2, 1263 hits were obtained. These were then tested for their binding affinities with the protein using Glide docking, Prime-MM/GBSA binding energy calculations and ADME properties calculations. The top ten ligands obtained from Glide-XP

docking were analyzed for their further validation. Out of these ten ligands, doxifluridene, pentobarbital and nebivolol were found to have the appropriate *E_{model}*, Glide Energy and ADME properties. These three ligands proved to be best on the basis of their binding energies ΔG_{bind} as well. The predicted IC_{50} values were also in the acceptable range for the three. However, only doxifluridene passed all tests, including a number of metabolic reactions for its excretion. We conclude that a variety of novel drugs that are non-toxic and easily available can be identified by this method before taking them for wet lab proceedings.

ACKNOWLEDGEMENTS

The authors thank University of Delhi's Scheme to Strengthen Research by Providing Funds to Faculty. NA thanks ARSD College, University of Delhi for granting Study Leave to her.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Hiromasa Y, Hu L, Roche TE. Ligand-Induced Effect on Pyruvate Dehydrogenase Kinase Isoform 2. *J BiolChem* 2006; 281, 12568-12579
2. Roche TE, Hiromasa Y. Pyruvate dehydrogenase kinase regulatory mechanisms and inhibition in treating diabetes, heart ischemia, and cancer. *Cell Mol Life Sci* 2007; 64,830-849
3. Gudi R, Bowker-Kinley MM, Kedishvili NY, Zhao Y, Popov KM. Diversity of the Pyruvate Dehydrogenase Kinase Gene Family in Humans. *J BiolChem* 1995; 270,28989-28994
4. Rowles J, Scherer SW, Xi T, Majer M, Nickle DC, Rommens JM, Popov KM, Harris RA, Riebow NL, Xia J, Tsui LC, Bogardus C, M. Prochazka M. Cloning and characterization of PDK4 on 7q21.3 encoding a fourth pyruvate dehydrogenase kinase isoenzyme in human. *J BiolChem* 1996; 271, 22376-22382
5. Bowker-Kinley MM, Davis WI, Wu P, Harris RA, Popov KM. Evidence for existence of tissue-specific regulation of the mammalian pyruvate dehydrogenase complex. *Biochem J* 1998; 329, 191-196
6. Steussy CN, Popov KM, Bowker-Kinley MM, Sloan Jr RB, Harris RA, Hamilton JA. Structure of pyruvate dehydrogenase kinase. Novel folding pattern for a serine protein kinase. *J BiolChem* 2001; 276, 37443-37450
7. Patel MS, Korotchikina LG. Regulation of mammalian pyruvate dehydrogenase complex by phosphorylation: complexity of multiple phosphorylation sites and kinases. *ExpMol Med* 2001; 33, 191-197
8. Randle PJ. Metabolic fuel selection: general integration at the whole-body level. *ProcNutrSoc* 1995; 54, 317-327
9. Kato M, Chuang JL, Tso SC, Wynn RM, Chuang DT. Crystal structure of pyruvate dehydrogenase kinase 3 bound to lipoyl domain 2 of human pyruvate dehydrogenase complex. *EMBO J* 2005; 24, 1763-1774
10. Knoechel TR, Tucker AD, Robinson CM, Phillips C, Taylor W, Bungay PJ, Kasten SA, Roche TE, Brown DG. Regulatory roles of the N-terminal domain based on crystal structures of human pyruvate dehydrogenase kinase2 containing physiological and synthetic ligands. *Biochemistry* 2006; 45, 402-415
11. Steussy CN, Popov KM, Bowker-Kinley MM, Sloan Jr RB, Harris RA, Hamilton JA. Structure of pyruvate dehydrogenase kinase. Novel folding pattern for a serine protein kinase. *J BiolChem* 2001; 276,37443-37450
12. Mayers RM, Leighton B, Kilgour E. PDH kinase inhibitors: a novel therapy for Type II diabetes? *BiochemSoc Trans* 2005; 33, 367-370
13. Filler R, Saha R. Fluorine in medicinal chemistry: a century of progress and a 60-year retrospective of selected highlights. *Future*

- Med Chem* 2009; 1,777-791
14. Motomura T, Nagamori H, Suzawa K, Ito H, Morita T, Kobayashi S, Shinkai H. Fluorine Compounds and Pharmaceutical Use Thereof, Pub. No. US 2010/0240634 A1, Pub. Date Sep. 23, 2010
 15. Gohlke H, Hendlich M, Klebe G. Predicting binding modes, binding affinities and "hot spots" for protein-ligand complexes using a knowledge-based scoring function. *Perspec Drug Discov Des* 2000; 20,115-144
 16. McInnes C. Virtual screening strategies in drug discovery. *Curr Opin Chem Biol* 2007; 11,494-502
 17. Seidel T, Ibis G, Bendix F, Wolber G. Strategies for 3D pharmacophore-based virtual screening Drug Discovery Today. *Technologies* 2010; 7, e221-e228
 18. Seifert MHJ, Wolf K, Vitt D. Virtual high-throughput in silico screening. *BioSilico* 2003; 1,143-149
 19. Subramaniam S, Mehrotra M, Gupta D. Virtual high throughput screening (vHTS) - A Perspective, *Bioinformation* 2008; 3, 14-17
 20. Nair SB, Teli MK, Pradeep H, Rajanikant GK. Computational identification of novel histone deacetylase inhibitors by docking based QSAR. *Comp Bio Med* 2012; 42,697-705
 21. Watts KS, Dalal P, Murphy RB, Sherman W, Friesner RA, Shelley JC. ConfGen: a conformational search method for efficient generation of bioactive conformers. *J Chem Inf Model* 2010; 50,534-546
 22. Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. 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* 2006B; 20,647-671
 23. Dixon SL, Smondyrev AM, Rao SN. PHASE: a novel approach to pharmacophore modeling and 3D database searching. *Chem Biol Drug Des* 2006A; 67, 370-372
 24. Lather V, Kristam R, Saini JS, Kristam R, Karthikeyan NA, Balaji VN. QSAR models for prediction of glycogen synthase kinase-3 β inhibitory activity of indirubin derivatives. *QSAR Comb Sci* 2008; 27,718-728
 25. Shah UA, Deokar HS, Kadam SS, Kulkarni VM. Pharmacophore generation and atom-based 3D-QSAR of novel 2-(4-methylsulfonylphenyl)pyrimidines as COX-2 inhibitors. *Mol Divers* 2010; 14,559-568
 26. Suganya RP, Kalva S, Saleena LM. International Conference on Bioscience, Biochemistry and Bioinformatics, International Proceedings of Chemical. *Biological & Environmental Engineering, Vol. 5, IACSIT Press, Singapore*, 2011
 27. Drwal MN, Griffith R. Combination of ligand and structure-based methods in virtual screening. *Drug Discov Today: Tech* 2013; 10, 395-401
 28. Sundarapandian T, Shalini J, Sugunadevi S, Woo LK. Docking-enabled pharmacophore model for histone deacetylase 8 inhibitors and its application in anti-cancer drug discovery. *J Mol Graph Model* 2010; 29,382-395
 29. Leonard JT, Roy K. On selection of training and test sets for the development of predictive QSAR models. *QSAR Comb Sci* 2006; 25,235-251
 30. Vijayan RSK, Ghoshal N. Structural basis for ligand recognition at the benzodiazepine binding site of GABAA α 3 receptor, and pharmacophore-based virtual screening approach. *J Mol Graph Model* 2008; 27,286-298
 31. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: A Free Tool to Discover Chemistry for Biology. *J Chem Inf Model* 2005; 45,177-182
 32. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: A Free Tool to Discover Chemistry for Biology. *J Chem Inf Model* 2012; 52, 1757-1768
 33. Lipinski CA, Lombardo F, Dominy BW, Freeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 1997; 23, 3-25
 34. Lipinski CA, Lombardo F, Dominy BW, Freeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 2001; 46, 3-26
 35. Bharatham N, Bharatham K, Lee KW. Pharmacophore identification and virtual screening for methionyl-tRNA synthetase inhibitors. *J Mol Graphics Modell* 2007; 25, 813-823
 36. Jorgensen WL, Duffy EM. Prediction of drug solubility from Monte Carlo simulations. *Bioorg & Med Chem Lett* 2000; 10, 1155-1158
 37. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc* 1996; 118, 11225-11236
 38. Kakkar R. Structure-Based Design of PDHK2 Inhibitors from Docking Studies. *Intl R J of Pharmaceuticals* 2011; 01, 50-58
 39. Yanagi Y, Tamaki Y, Inoue Y, Obata R, Muranaka K, Homma N. Subconjunctival Doxifluridine Administration Suppresses Rat Chorioidal Neovascularization through Activated Thymidine Phosphorylase. *Invest Ophth Vis Sci* 2003; 44, 751-754
 40. Michot JM, Seral C, van Bambeke F, Mingeot-Leclercq MP, Tulkens PM. Influence of efflux transporters on the accumulation and efflux of 4 quinolones (ciprofloxacin, levofloxacin, garenoxacin, and moxifloxacin) in J774 macrophages. *Antimicrob Agents Chemother* 2005; 49,2429-2437
 41. Matijević-Sosa J, Cvetnić Z. Antimicrobial activity of N-phthaloyl-amino acid hydroxamates. *Acta Pharm* 2005; 55, 387-399
- Peer reviewers:** Giuseppe Manco, MD, Professor, Institute of Protein Biochemistry, National Research Council of Italy, Via P. Castellino 111, 80131, Naples, Italy.