



# Discovery and design of new PI3K inhibitors through pharmacophore-based virtual screening, molecular docking, and binding free energy analysis

Saikiran Reddy Peddi<sup>1</sup> · Sree Kanth Sivan<sup>2</sup> · Vijjulatha Manga<sup>1</sup>

Received: 24 April 2018 / Accepted: 2 July 2018 / Published online: 18 July 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

Phosphoinositide 3-kinase (PI3K)/protein kinase B (AKT)/the mammalian target of rapamycin (mTOR) signaling pathway is one of the sought after therapeutic target for treating human cancers. This pathway is often hyper activated in cancers. In the present study, pharmacophore-based virtual screening, molecular docking, and binding free energy calculations were performed on a series of quinoline derivatives which were reported to be effective against PI3K $\alpha$ . A five-point pharmacophore hypothesis with one hydrogen bond acceptor (A), one hydrogen bond donor (D), one hydrophobic group (H), and two aromatic rings (R) was developed with acceptable  $R^2$  and  $Q^2$  values of 0.93 and 0.60 respectively. Eventually, common pharmacophore hypothesis-based screening was conducted against TOSLab, CPP, and ASINEX macrocyclic databases, and potential hits were identified which were further subjected to rigorous docking process in order to screen out drug like molecules having crucial interactions with the target PI3K $\alpha$ . Finally, binding free energy analysis was carried out for the top hits obtained from docking process. We also designed new 1, 3, 4-oxadiazole-based cyclic peptides by incorporating the structural features of the hits obtained from the above databases. Among the designed cyclic peptides, the cyclic peptide with tryptophan moiety showed good interactions and free binding energy values. On the whole, this study helped us in identifying new promising molecules as PI3K $\alpha$  inhibitors which can be explored further to generate greater number of compounds with better pharmacokinetic properties.

**Keywords** Phosphoinositide 3-kinase · Pharmacophore-based virtual screening · Molecular docking · Binding free energy · Quinoline derivatives · Cyclic peptides

## Introduction

Phosphoinositide 3-kinase (PI3K) signaling pathway is commonly deregulated in human cancers. PI3Ks are a family of signal transducer enzymes which are involved in various cellular processes such as cell quiescence, proliferation, differentiation, motility, and intracellular trafficking [1–3].

Basically PI3K family is divided into three classes (classes I, II, and III) based on structural characteristics and substrate specificity [4, 5]. The class I PI3K family consists of  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$  isoforms that generate phosphatidylinositol 3, 4, 5-triphosphate, a potent secondary messenger that triggers the activation of several downstream effectors including Akt [6]. Previous studies have shown that many tumors fostering somatic genetic aberrations can result in the immanent activation of the PI3K/Akt/mTOR signaling network in human cancers [6]. Therefore, targeting PI3K pathway is an attractive strategy for treating cancers. Many researchers have reported various quinoline derivatives as PI3K inhibitors. These derivatives have attracted researchers as many compounds which are under clinical trials at present possessed quinoline core (GSK2126458, BEZ235, and BGT226) and displayed exceptional potency in vitro and in vivo against PI3K/mTOR [7, 8]. Hence, we focused our attention on quinoline derivatives and carried out molecular modeling studies on them.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11224-018-1154-9>) contains supplementary material, which is available to authorized users.

✉ Vijjulatha Manga  
vijjulathamanga@gmail.com

<sup>1</sup> Molecular Modeling and Medicinal Chemistry Group, Department of Chemistry, University College of Science, Osmania University, Hyderabad, Telangana 500007, India

<sup>2</sup> Department of Chemistry, Nizam College, Osmania University, Hyderabad 500001, India

Molecular modeling techniques which are developed in recent times have paved the way for quick and cost effective designing and discovery of novel drugs especially in terms of pharmaceutical research. Virtual screening in particular has been employed as a tool to ingress new drug-like candidates from large data base [9, 10]. Over a period of time, numerous *in silico* techniques such as pharmacophore modeling, molecular dynamics have been developed and utilized in the pharmacology. Interestingly, the *in silico* approaches for lead identification and drug discovery have gained popularity in recent times [11–15]. Finding a common pharmacophore hypothesis in conjunction with atom-based 3D QSAR (three dimensional quantitative structural activity relationship) model with reliable statistical values are very useful tools for finding and developing small molecules which are capable of becoming drugs [16–20]. Therefore, in the present study, pharmacophore-based virtual screening and docking studies were performed on a series of quinoline derivatives in order to discover and design new leads with improved selectivity and potency towards PI3K $\alpha$ . The pharmacophore model generated in this study was used as a query and screened against TOSLab, CPP (Cell Penetrating Peptides with peptides length less than or equal to 30 amino acids) and ASINEX macrocycles databases in order to identify potential hits. The obtained hits were subjected to rigorous docking process, and drug-like candidates having key interactions with target PI3K $\alpha$  were screened out. Further binding energy calculations were also carried out using MM/GBSA. Furthermore, based on the outcome of results, new oxadiazole-based cyclic peptides were designed. All in all, in the present study, an attempt was made to identify new PI3K $\alpha$  inhibitors using integrated pharmacophore-based screening, molecular docking, and MM/GBSA approaches.

## Materials and methods

The common pharmacophore and 3D QSAR models was generated using PHASE (Pharmacophore Alignment and Scoring Engine) module of Schrodinger Suite [21]. Basically, PHASE identifies the steric and structural features of the compounds that are common and promising for the target inhibition. It also predicts the inhibitory activity values of the compounds [19]. In PHASE, there are five sequences of steps to build pharmacophores and 3D QSAR models. They are preparation of ligands, pharmacophore site creation, common pharmacophore identification, scoring hypotheses, and building 3D QSAR model.

### Dataset for analysis

A series of quinoline derivatives were taken from literature [6, 22–24]. The IC<sub>50</sub> values of these derivatives were converted

to corresponding pIC<sub>50</sub>. A total of 44 compounds were used as a data set of which 30 compounds were randomly chosen as training set, and 14 compounds were selected as test set in order to generate structural diversity in model generation.

### Ligand preparation

The structures of all the derivatives were drawn in maestro build panel. The sketched structures were converted into 3D and prepared using LigPrep module. Finally, the generated low energy conformers along with activity values were imported to the Phase project workflow.

### Generation of pharmacophore hypotheses

A set of points in 3D space defines the ligand structure which actually facilitates the non-covalent binding between the ligand and its target receptor. For the creation of pharmacophores, generally PHASE supplies a built-in set of six pharmacophore features, H-bond acceptor (A), H-bond donor (D), hydrophobic group (H), negative/anionic (N), positive/cationic (P), and aromatic ring (R). PHASE generates the plausible pharmacophore sites for a given data set based upon the structural similarity and the common pharmacophoric features. The generated features in the PHASE will be further assigned for geometrical entities of the data set in order to define physical characteristics.

### Finding a common pharmacophore

PHASE examines the pharmacophore from all conformations of the ligand and groups together the pharmacophore which contains identical sets of features with related spatial arrangements and finally gives a common pharmacophore. A tree-based partitioning technique is used to identify the common pharmacophores which actually groups together similar pharmacophores according to their inter site distances [19]. The intersite distance was set to 2 Å with maximum tree depth of five; initial and final box sizes were set to 32 and 1 Å respectively. Finally, common pharmacophore containing sites were generated by applying active and inactive thresholds which was further taken for scoring the hypothesis.

### Scoring hypotheses

In this step, common pharmacophores are examined, and a pharmacophore which yields best alignment of active set of ligands is identified by applying a scoring procedure. This particular pharmacophore furnishes a hypothesis to demonstrate how the active compounds bind to the receptor. The scoring procedure ranks different hypotheses and allows making realistic choices about the hypotheses which is most relevant for further investigation. Ligand activity was expressed

as  $-\log(IC_{50})$  and incorporated into the score. The vector, site, and volume scores were set to 1.0 while scoring actives. The hypotheses which were obtained from this process were scored with respect to inactive, applying a weight of 1.0. The inactives were scored in order to check the alignment of these compounds with respect to the pharmacophore hypothesis so as to take a decision on the selection of the hypothesis. For a good hypothesis, the difference between the scores of actives and inactives should be high.

### Building QSAR model

QSAR models can be built in PHASE using the activities of the ligands which match a given hypothesis. These QSAR models are based on PLS regression that are applied to a huge set of binary valued variables. The QSAR models were derived based on PLS regression and enforced on a set of binary variables. The grid space which is occupied by training set of compounds provides independent variables for QSAR model. Depending upon the type of grid space occupied, atom types were assigned into one of the six categories, i.e., D (H-bond donor), H (hydrophobic or non polar), P (positive), N (negative), A (H-bond acceptor/electron-withdrawing groups), M (miscellaneous). Atom-based QSAR models were developed for ADHRR hypothesis using 30 training set of ligands with grid spacing of 1.0 Å. QSAR models were generated consisting of one to three PLS factors. The generated QSAR model is validated both internally and externally using training and test set ligands respectively.

### Screening of 3D databases using pharmacophore model

The generated PHASE model was used as a query and TOSLab, CPP, and ASINEX macrocycles databases were screened in order to identify the new leads that match with the best pharmacophore hypothesis and estimate the predicted activities using 3D QSAR model.

### Molecular docking

The X-ray crystal of PI3K $\alpha$  (PDB ID: 4KYN) was downloaded from protein data bank [25]. It was prepared, refined, and minimized using protein preparation wizard by applying OPLS 2005 force field. Later, receptor grid was generated around the active site of the protein using GLIDE 5.6. During grid generation, the van der Waals scaling was set to 0.9 [26], and ligand molecule was picked so that it could be excluded from grid generation. Subsequently, the ligands were docked in a step-wise manner using extra precision (XP) docking mode. The dimensions of the grid box were 10 Å × 10 Å × 10 Å. In initial phase of docking, 5000 poses per ligand were taken out of which top 800 poses per ligand were passed

on for energy minimization. During energy minimization, the maximum number of minimization steps were set to 100, and distance-dependent dielectric constant was set to 2.0.

### Dock-based virtual screening

The ligands obtained from TOSLab screening were initially docked into the active site using standard precision protocol. The 50% of the top scoring ligands were then taken and subjected to XP docking. Finally, binding free energies were calculated for 10% of the top scoring ligands obtained from XP docking. On the other hand, the ligands retrieved from screening of CPP data base could not dock into the active site as the binding pocket in the protein active site was not very large enough to accommodate these peptides. Interestingly, the ligands obtained from ASINEX macrocycles database screening fitted comfortably in the active site of PI3K $\alpha$ . Initially, these ligands were also docked using standard precision protocol. Later, 50% of top scoring ligands were taken and subjected to XP docking. Finally, binding free energies were calculated for 10% of the top scoring ligands.

### Binding free energy calculations

Molecular dynamics simulations are computationally very expensive especially for larger protein systems like PI3K $\alpha$ . MM/GBSA (Molecular Mechanics with Generalized Born Surface Area) on the other hand is an alternate approach to refinement that utilizes a continuum representation of solvent. One of the major advantages of continuum approach is that it uses conformational search method as opposed to MD simulations, that is conveniently employed in conjunction with a continuum solvation model. The method is efficient and relatively inexpensive for calculating the global free energy [27]. Therefore, MM/GBSA was used for calculating binding free energies. The best ranking molecules obtained from extra precision docking were subjected to binding free energy calculations using Prime MM/GBSA available in the Schrodinger Suite. The equations for estimating binding free energy are as follows:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - [G_{\text{ligand(unbound)}} + G_{\text{receptor(unbound)}}] \\ = \Delta E_{\text{MM}} + \Delta G_{\text{solv}} + \Delta G_{\text{SA}}$$

where  $\Delta E_{\text{MM}}$  is the difference in energy between the protein ligand complex and the sum of the energies of the protein with and without ligand,  $\Delta G_{\text{solv}}$  is the difference in GBSA solvation energy of the protein ligand complex and the sum of the solvation energies for the ligand and unliganded protein,  $\Delta G_{\text{SA}}$  is the difference in the surface energy for the protein ligand complex and the sum of the surface area energies of the ligand and uncomplexed protein [28]. VSGB 2.0 model was used in MM/GBSA calculations. It actually approximates the

solvation free energy with an optimized implicit model [27]. The exterior and interior dielectric constants were set to 80 and 1 respectively, during MM/GBSA calculations. In VSGB model, polar and non-polar solvent solute contributions can be described by a polarization term and a hydrophobic term, respectively. Basically, this polarization term depends on partial charges, solvent, and internal dielectric constants and  $f_{GB}$  which can be given by the following equation:

$$G_{\text{pol}} = \frac{1}{2} \left( \frac{1}{\epsilon_{\text{in}}(ij)} - \frac{1}{\epsilon_{\text{sol}}} \right) \sum_{i < j} q_i q_j f_{GB}$$

Where  $f_{GB}$  is a function of generalized Born radii ( $\alpha_i$  and Equation  $j$ ) and distance between two atoms ( $r_{ij}$ ) as described by Still et al. [29].

## Results and discussion

The typical chemical structures of all PI3K $\alpha$  inhibitors are shown in Table 1. Among these 44 inhibitors, the inhibitors with  $\text{pIC}_{50}$  values greater than 8.5 were assigned as active while inhibitors with  $\text{pIC}_{50}$  values less than 7.5 were assigned as inactive pharm sets to create pharmacophore site points. The  $\text{pIC}_{50}$  values and Fitness scores of all the data set compounds are shown in Table 1.

### Determination of pharmacophore model

A total of four variant combinations, viz. AADRR, AADHR, AAHHR, and ADHRR consisting of hydrogen bond acceptor (A), hydrophobic group (H), H-bond donor (D), and aromatic ring (R) features were derived in the common pharmacophore identification step. Twenty-two five-point pharmacophore hypotheses were generated after scoring these four common pharmacophores. The generated hypotheses were ranked according to the survival active, inactive, post-hoc alignment, vector, and volume scores. The hypothesis ADHRR associated with five pharmacophore site points was taken for further analysis. The geometry of ADHRR is depicted in Fig. 1, where the pink sphere with vectors represents H-bond acceptor feature (A4), blue sphere with a vector represents H-bond donor feature (D10), green sphere represents hydrophobic feature (H11), and two orange tori (ring-shaped surfaces) represents aromatic ring features (R15 and R18).

### Building 3D QSAR model and validation

Atom-based 3D QSAR models were developed for top four pharmacophore hypotheses. PLS regression was performed with three maximum PLS factors where PHASE descriptions

acted as independent variables while  $\text{pIC}_{50}$  values served as dependent variables. The results are tabulated in Table 2. Among the four hypotheses, ADHRR gave good statistical model with high values of correlation coefficient;  $R^2 = 0.93$ , low standard deviation;  $SD = 0.20$ , variance ratio;  $F = 119.1$ , high predictive coefficient;  $Q^2 = 0.60$ , low RMSE = 0.47 and Pearson's  $R$  value = 0.81. Therefore, it is evident that the developed 3D QSAR model has sterling statistical criteria [30, 31] and can be used for further optimization and exploration. The scatter plot of experimental versus predicted  $\text{pIC}_{50}$  values of both training and test set ligands are shown in Fig. 2. The graph showed a positive correlation between predicted and experimental values. Hence, it can be confirmed that the generated 3D QSAR model is robust and significant.

### 3D QSAR visualization of most active compound

The 3D visualization of the best hypothesis ADHRR and the selected quinoline ligands in the aspect of developed QSAR helps in better understanding of structure and activity relationship (SAR) in construing the activity [20]. The 3D QSAR model visualization of the best active compound ( $\text{pIC}_{50} = 9.00$ ) is depicted in Fig. 3. In Fig. 3a, b, and c, green, blue, and red cube areas represent favorable regions where as violet, orange, and green cubes areas represent unfavorable regions for hydrophobic, H-bond donor, and H-bond acceptors respectively.

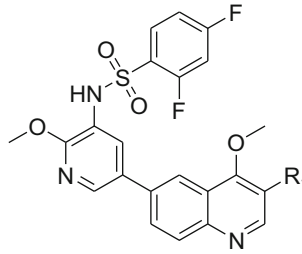
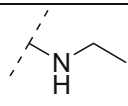
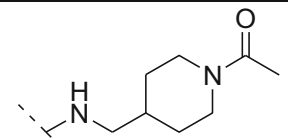
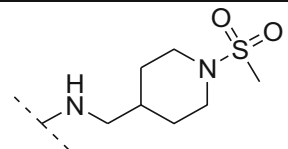
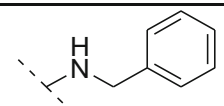
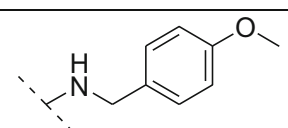
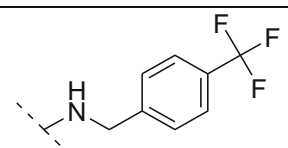
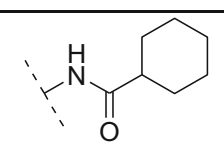
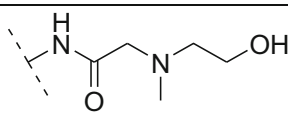
In Fig. 3a, it can be observed that there are three hydrophobic features superposed on fluorine-attached benzene, and benzene of quinoline moiety though the hypothesis has only one of them, i.e., H11. The atom-based 3D QSAR model takes into account the steric clashes with the receptor while predicting the activity apart from pharmacophore features, whereas the pharmacophore-based 3D QSAR model predicts the activity directly by considering pharmacophore sites and their locations [32]. The H-bond donor feature superposed on “NH” attached to sulfonyl group is depicted in Fig. 3b. It exactly matched with the pharmacophoric feature, i.e., D10. On the other hand, H-bond acceptor features superposed on oxygen atom of sulfonyl group, fluoro benzene, and near pyridine of quinoline moiety are shown in Fig. 3c. Interestingly, the hypothesis showed only one H-bond acceptor feature (A4). The presence of additional hydrophobic and H-bond acceptor features is beneficial for the activity.

### Pharmacophore-based virtual screening

The three databases namely TOSLab, CPP, and ASINEX macrocycles databases were screened using the generated pharmacophore model (ADHRR) in order to search for potent PI3K $\alpha$  inhibitors. CPP and ASINEX macrocycles databases were also taken so as to screen out cyclic compounds in addition to linear molecules. During this process, the sites of the

**Table 1** Structures of PI3K $\alpha$  inhibitors along with experimental and predicted pIC<sub>50</sub> values of both training and test set compounds based on pharmacophore hypothesis ADHRR. The inhibitory activity and fitness

score values of both best active (compound 12) and least active (compound 41) compounds are shown in bold.

					
Compound	R <sub>1</sub>	Expt. pIC <sub>50</sub>	Pred. pIC <sub>50</sub>	Fitness Score	Binding Energy (Kcal/mol)
1		8.03	8.26	2.88	-63.35
2		8.43	8.37	2.75	-80.43
*3		8.33	8.35	2.71	-80.56
*4		7.44	7.96	1.51	-71.36
*5		7.57	7.91	1.49	-69.63
*6		7.52	7.91	1.48	-67.33
7		7.82	7.80	2.46	-83.03
8		8.89	8.74	1.81	-78.09

\*Test set compounds

Table 1 (continued)

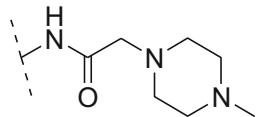
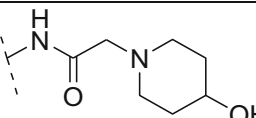
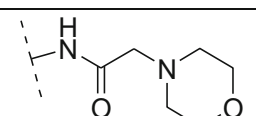
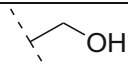
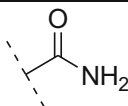
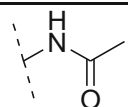
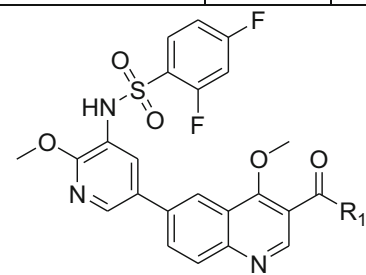
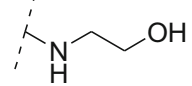
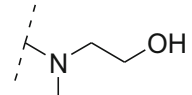
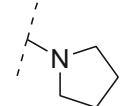
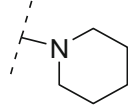
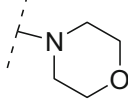
*9		8.50	7.89	2.43	-77.01
*10		8.70	8.29	2.76	-59.76
*11		8.77	8.58	1.00	-83.22
<b>12</b>		<b>9.00</b>	<b>8.52</b>	<b>2.92</b>	<b>-104.17</b>
13		8.59	8.46	3.00	-81.74
14		8.08	7.94	2.54	-67.99
					
15		8.80	8.70	1.96	-89.19
16		8.10	8.15	2.58	-75.59
17		7.59	8.06	2.87	-66.82
*18		7.57	8.42	2.80	-41.59
19		7.48	7.45	1.45	-73.96

Table 1 (continued)

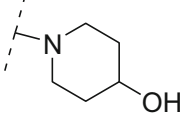
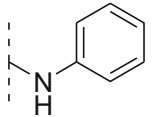
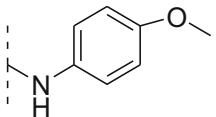
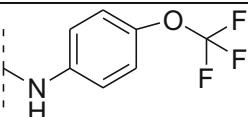
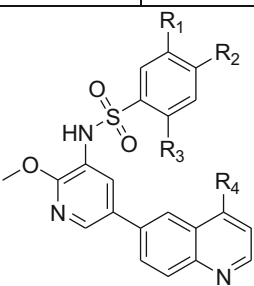
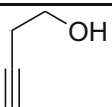
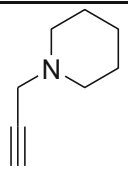
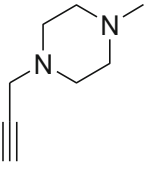
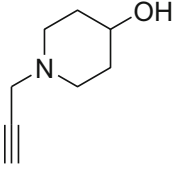
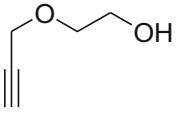
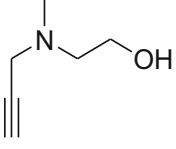
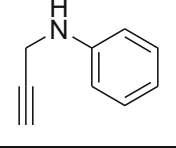
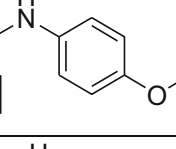
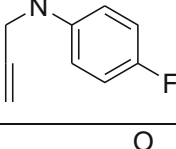
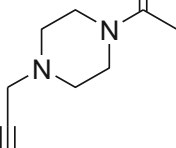
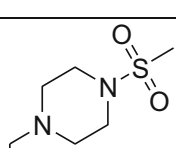
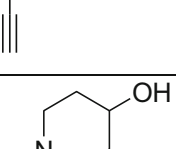
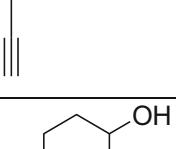
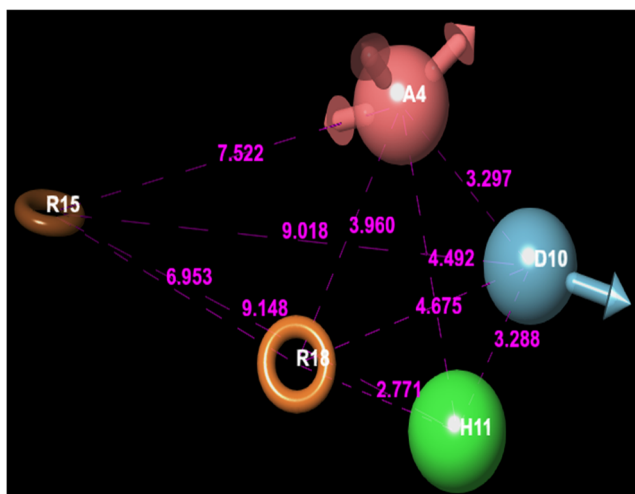
20		7.92	8.03	2.54	-68.93			
21		8.27	8.00	2.48	-70.93			
*22		7.96	8.39	2.76	-63.52			
23		7.14	7.37	2.41	-65.56			
								
Compound	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	Expt. pIC <sub>50</sub>	Pred. pIC <sub>50</sub>	Fitness Score	Binding Energy (Kcal/mol)
24	H	F	F	H	7.85	8.19	2.71	-65.13
25	H	F	F		8.33	8.16	1.36	-66.69
26	H	F	F		7.56	7.57	1.53	-74.95
*27	H	F	F		7.62	7.59	1.17	-71.72
28	H	F	F		8.79	8.83	1.60	-75.88

Table 1 (continued)

29	H	F	F		8.63	8.50	1.41	-79.67
*30	H	F	F		8.57	8.16	1.74	-87.24
31	H	F	F		7.56	7.53	1.26	-71.89
32	H	F	F		7.49	7.45	1.24	-79.68
33	H	F	F		7.66	7.51	1.54	-69.51
34	H	F	F		8.70	8.77	1.57	-57.89
35	H	F	F		8.46	8.50	1.05	-61.37
36	H	F	H		7.40	7.43	1.29	-87.54
37	F	H	H		7.37	7.60	1.27	-86.67





**Fig. 1** The illustration of pharmacophore model ADHRR where the pink sphere with vectors represent H-bond acceptor feature, blue sphere with vector represent H-bond donor feature, green sphere represents hydrophobic feature and two orange rings represent aromatic ring features

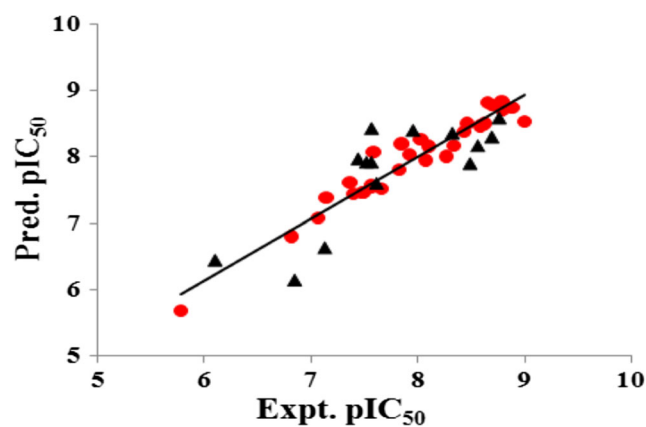
hypothesis will be agnated with the conformers of the ligands which are pre-computed [19]. In the current study, the ligands were read to match maximum five sites of the hypothesis. The search process restituted 273 hits out of 39,988 molecules from TOSLab, 16 hits out of 1411 molecules from CPP database, and 204 hits out of 31,500 molecules from ASINEX macrocycles database. Since the pharmacophore model was derived from known PI3K $\alpha$  inhibitors, it typifies the traits that are required for binding to PI3K $\alpha$ . Therefore, it is clear that the hits obtained from the search process will definitely have greater probability of becoming PI3K $\alpha$  inhibitors.

### Molecular docking-based virtual screening using Glide

The hit molecules obtained from the three databases (TOSLab, CPP, and ASINEX macrocycles) were docked into the active site of PI3K $\alpha$  using the receptor grid which was generated during the docking process. The docking process was carried out in two stages in our analysis. Initially, about 273 molecules obtained from pharmacophore-based screening of TOSLab database were taken for SP docking. Later, the top

**Table 2** Statistical results of top four pharmacophore hypotheses generated through PHASE

ID	SD	$R^2$	$F$	RMSE	$Q^2$	Pearson- $R$
ADHRR	0.20	0.93	119.1	0.47	0.60	0.81
AAHHR	0.19	0.94	127.8	0.59	0.46	0.73
AADHR	0.16	0.96	192.0	0.60	0.36	0.60
AADRR	0.20	0.93	120.8	0.86	0.33	0.22



**Fig. 2** Scatter plot of experimental versus predicted pIC<sub>50</sub> values of training (red dots) and test (black triangles) set compounds

137 molecules, i.e., 50% of the molecules which showed high dock scores were subjected to XP docking. Similarly, the hit molecules obtained from CPP database were subjected to SP docking. However, these molecules did not dock into the protein due to their large size. Finally, 204 hits obtained from ASINEX macrocycles database were docked using SP mode. About 102 molecules (i.e., 50%) which showed high dock scores were passed on to the next stage of XP docking. The top 10%, i.e., 14 molecules of TOSLab and 10 molecules of ASINEX macrocycles database which exhibited good dock scores were taken further for binding free energy analysis. The schematic representation of entire virtual screening process is shown in Fig. 4.

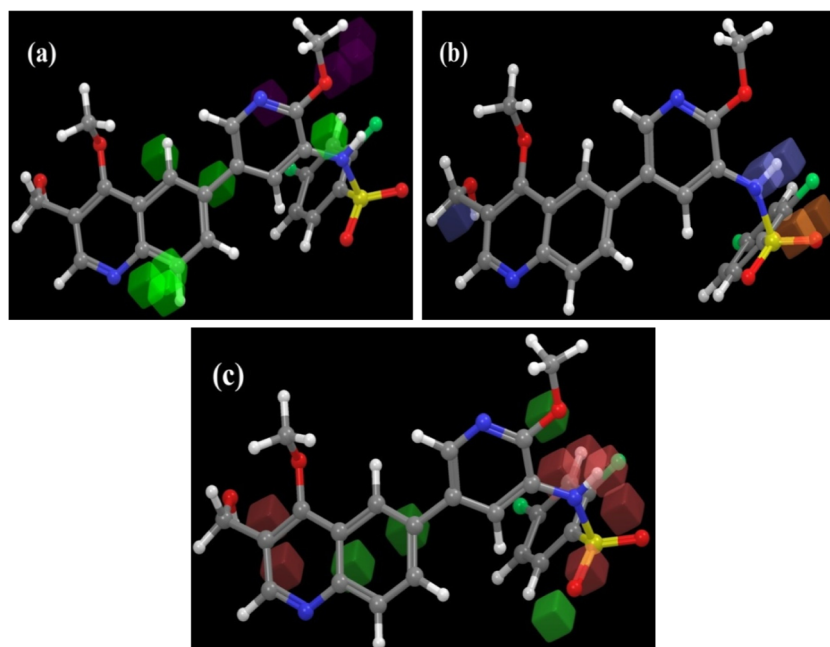
### Calculation of binding free energies

The binding free energy values of the top 10% of the hit molecules obtained from both the databases are shown in Table 3. These values were compared with the binding free energy value of the best active compound in the data set (compound 12, Table 1). From Table 3, it can be observed that some of the hits showed good binding free energy values. Molecules T2, T4, T6, and T10 from TOSLab database and molecules A3, A5, A6, and A10 from ASINEX macrocycles database in particular showed good binding energy values of  $-94.83$ ,  $-101.90$ ,  $-123.77$ ,  $-105.32$ ,  $-95.12$ ,  $-109.99$ ,  $-97.77$  and  $-97.01$  Kcal/mol respectively which are in the range of best active compound 12 ( $-104.17$  Kcal/mol). Molecules T6, T10, and A5 possessed high binding energy values which are even greater than the best active compound. Therefore, it is evident that the molecules T6, T10, and A5 can definitely act as potent inhibitors against PI3K $\alpha$ .

### Interaction studies of screened hits

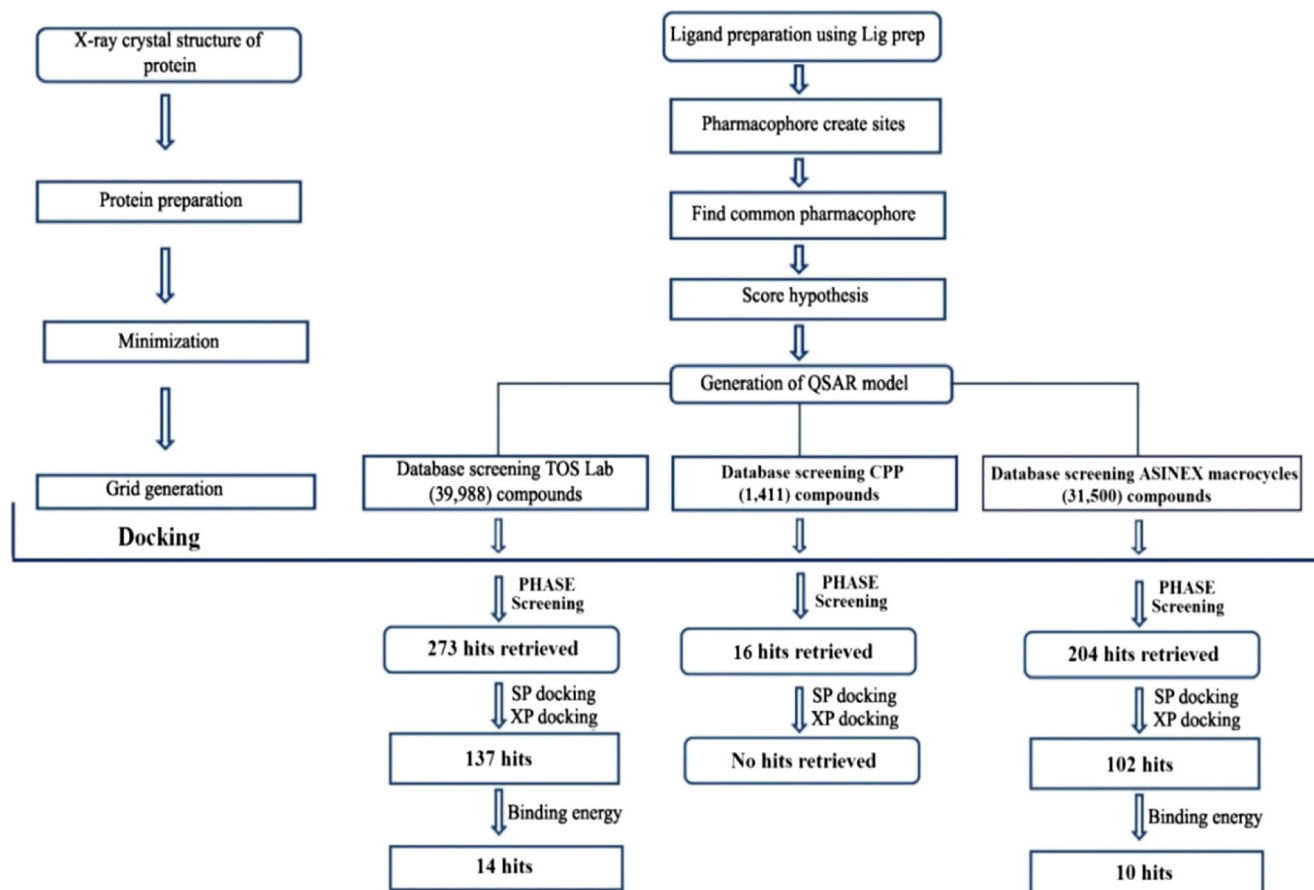
Ligand interaction diagram (LID) available in the Schrodinger Suite was employed to explore the interaction pattern of the screened hits. The results are depicted in Fig. 5. The pink

**Fig. 3** 3D QSAR model visualization in connection with the best active compound (compound 12), illustrating the effect of **a** hydrophobic, **b** H-bond donor, and **c** H-bond acceptor features



colored lines in Fig. 5 represent hydrogen bonds, green colored lines represent  $\pi$ - $\pi$  stacking interactions, and red colored lines represent  $\pi$ -cationic interactions. Initially, the best active

compound in complex with PI3K $\alpha$  was analyzed with the help of LID. The result is depicted in Fig. 5. It showed five hydrogen bond interactions with amino acid residues Ser



**Fig. 4** Flowchart showing the virtual screening workflow for identification of hit molecules

**Table 3** Binding free energies and predicted pIC<sub>50</sub> values of screened hits obtained from TOSLab (T1-T14) and ASINEX macrocycles database (A1-A10)

SNo	Compound	Binding energy (Kcal/mol)	Predicted pIC <sub>50</sub>
1	T1	-77.40	7.52
2	T2	-94.83	7.78
3	T3	-90.25	7.65
4	T4	-101.90	7.89
5	T5	-90.29	7.63
6	T6	-123.77	7.64
7	T7	-79.45	7.91
8	T8	-82.41	7.83
9	T9	-69.03	7.62
10	T10	-105.32	7.66
11	T11	-89.83	7.60
12	T12	-89.35	7.60
13	T13	-79.76	7.81
14	T14	-79.60	7.82
15	A1	-78.57	7.27
16	A2	-82.38	7.69
17	A3	-95.12	7.66
18	A4	-87.95	7.49
19	A5	-109.99	7.75
20	A6	-97.77	7.21
21	A7	-84.39	7.43
22	A8	-81.55	7.55
23	A9	-84.45	7.68
24	A10	-97.01	7.65

1774, Lys 1802, Val 1851, Ser 1854, and Asp 1933. The hits obtained from TOSLab database showed  $\pi$ - $\pi$  stacking,  $\pi$ -cationic interactions and formed salt bridges apart from hydrogen bonds. They showed hydrogen bond interactions with Arg 1770, Ser1774, Lys 1776, Lys 1802, Glu 1849, Val 1851, Asn 1853, Ser 1854, Gln 1859, Ser 1919 and Asp 1933;  $\pi$ - $\pi$  stacking interactions with Trp 1780, Tyr 1836, HIS 1855, and HIS 1917;  $\pi$ -cationic interactions with Lys 1802 and formed salt bridges with Glu 1798 and Lys 1802. Among the screened hits of TOSLab, molecules T6, T7, T10, and T11 showed five interactions (Fig. 5) and the remaining hits exhibited two to three interactions (Supplementary information, Fig. S1). On the other hand, the hits obtained from ASINEX macrocycles database showed hydrogen bond interactions with Ser 1773, Ser 1774, Lys 1776, Lys 1802, Val 1851, Gln 1859, Asn 1920, and Asp 1933 and  $\pi$ - $\pi$  stacking interactions with Trp 1780, Tyr 1836, and HIS 1917. The  $\pi$ -cationic interactions and salt bridges were absent in these molecules. Among the screened hits of ASINEX macrocycles, molecule A5 showed four interactions (Fig. 5); molecules A2, A3, and A4 showed three interactions, and remaining ones showed only one to two

interactions (Supplementary information, Fig. S1). It is therefore clear that both linear and cyclic molecules can form interactions with PI3K $\alpha$  and may act as good inhibitors.

### Design of 1, 3, 4-oxadiazole-based cyclic peptides

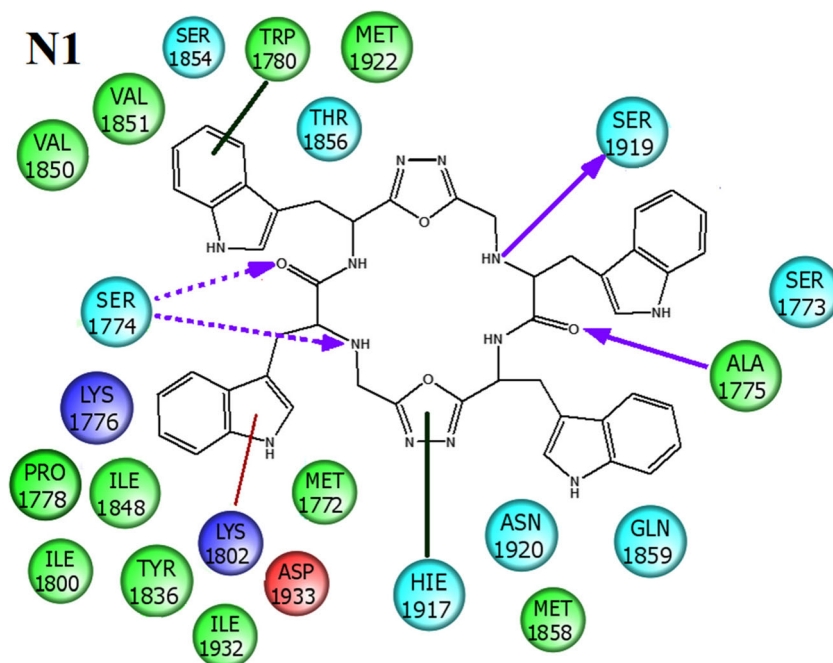
The hits obtained from both TOSLab and ASINEX macrocycles databases showed good interactions and binding free energy values. The hits T6, T10, and A5 could even compete with the best active compound in the dataset (compound 12) in terms of interactions and binding free energy values. They exhibited binding free energy values greater than the best active compound (-123.77, -105.32 and -109.99 respectively) and formed four to five interactions (T6 and T10 exhibited five and A5 exhibited four interactions). However, they showed only two to three hydrogen bond interactions which are considered to be strong interactions. Therefore, we have designed certain 1,3,4-oxadiazole-based cyclic peptides by imbibing the structural features of the hits retrieved from both the databases which could possess good interactions apart from binding free energy values. Among the designed 1, 3, 4-oxadiazole-based cyclic peptides, the cyclic peptide with tryptophan moiety exhibited seven interactions (four hydrogen bond, two  $\pi$ - $\pi$  stacking, and one  $\pi$ -cationic interaction Fig. 6) and showed binding free energy value of -132.24Kcal/mol. It satisfied both the conditions of interactions and binding free energy values. Actually, we have designed and reported certain 1, 3, 4-oxadiazole-based cyclic peptides in our earlier studies against HIV-1 TAR RNA [28]. Most of the TOSLab hits possessed indole moiety (T1, T6, T9, T11, and T12) in their core structure, and few compounds possessed benzimidazole and benzothiazole moieties in their core structure (Supplementary information, Fig. S1). The aromatic system in these compounds particularly formed  $\pi$ - $\pi$  stacking interactions. On the other hand, the cyclic peptides which were docked into the active site of PI3K $\alpha$  (ASINEX macrocycles database hits) showed that the cavity was large enough to accommodate bulky molecules. Further, the pharmacophore model generated in the present study had five features, namely, one acceptor group (A), one donor group (D), one hydrophobic group (H), and two aromatic rings (ADHRR). Furthermore, cyclic peptides are resistant to proteases, and by reducing entropic penalties, they bind to specific receptors with high affinity. Hence, all the above features were taken into consideration and integrated together to design the oxadiazole-based cyclic peptides.

### Molecular docking and binding free energy analysis of 1, 3, 4-oxadiazole-based cyclic peptides

The designed cyclic peptides were docked into the active site of PI3K $\alpha$  using the same grid which was used for docking the dataset and database compounds. The extra precision docking



**Fig. 6** Ligand interaction diagram of newly designed 1, 3, 4-oxadiazole-based cyclic peptide with tryptophan moiety



(XP) protocol was used for docking these molecules. The designed cyclic peptides showed hydrogen bond interactions with Ile 1771, Ser1774, Ala 1775, Glu 1768, Arg 1770, Glu 1798, Gln 1859, Ser 1919 and Asp 1933;  $\pi$ - $\pi$  stacking interactions with Trp 1780, HIS 1917, and HIS 1936; and  $\pi$ -cationic interactions with Arg 1770 and Lys 1802. Among the designed molecules, except molecule N4, all molecules showed five interactions (Supplementary information, Fig. S2). The only molecule which showed more than five interactions was molecule N1. It showed four hydrogen bond interactions with Ser 1774, Ala 1775, and Ser 1919, two  $\pi$ - $\pi$  stacking interactions with Trp 1780 and HIS 1917, and one  $\pi$ -cationic interaction with Lys 1802 (Fig. 6).

Binding free energy values were also calculated for the designed cyclic peptides. These values varied from  $-63.67$  to  $-132.24$  Kcal/mol. The binding free energy values of all 1, 3, 4-oxadiazole-based cyclic peptides are shown in Table 4. Among the designed cyclic peptides, tryptophan- and phenyl alanine-based 1, 3, 4 -oxadiazole cyclic peptides showed good

binding energy values which are greater than  $-100$  Kcal/mol (Table 4). Therefore, these two molecules may act as good inhibitors against PI3K $\alpha$ .

## Conclusions

The objective of the present study was to discover and design new potent inhibitors against PI3K $\alpha$ . Pharmacophore-based virtual screening, molecular docking, and binding free analysis were employed to achieve this purpose. A five-point common pharmacophore hypothesis (ADHRR) was developed using 44 PI3K $\alpha$  inhibitors and was applied to screen three different databases, i.e., TOSLab, CPP, and ASINEX macrocycles. The identified hits were docked into the active site of PI3K $\alpha$  and further subjected to binding free energy analysis. The results revealed that the compounds possessing indole and benzothiazole moieties with prescribed pharmacophoric features can act as potent inhibitors against PI3K $\alpha$ . It was also evident that even cyclic compounds can act as good inhibitors against PI3K $\alpha$  on par with linear compounds. Further, the results obtained from docking and binding free energy analysis of 1, 3, 4-oxadiazole-based cyclic peptides disclosed that the cyclic peptides with tryptophan moiety can definitely act as good inhibitors against PI3K $\alpha$ . On the whole, the results obtained in this study suggest that the combined 3D QSAR, molecular docking, and binding free energy protocols can be helpful in identifying and designing new PI3K $\alpha$  inhibitors. We hope that the inferences drawn in this work can provide some insights to researchers to discuss and design new PI3K $\alpha$  inhibitors with greater activity.

**Table 4** Binding free energies of newly designed 1, 3, 4-oxadiazole-based cyclic peptides

S No	Compound	Binding energy (Kcal/mol)	Predicted pIC <sub>50</sub>
1	N1	$-132.24$	7.82
2	N2	$-69.72$	7.11
3	N3	$-63.67$	7.31
4	N4	$-100.91$	7.69
5	N5	$-88.39$	7.55
6	N6	$-69.48$	7.01

**Acknowledgements** We greatly acknowledge Schrödinger LLC, New York for providing the software and author Saikiran Reddy Peddi (IF150172) would like to thank DST for a research fellowship.

**Funding information** The study is financially supported by the DST PURSE-II/2017.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

1. Vanhaesebroeck B, Stephens L, Hawkins P (2012) PI3K signalling: the path to discovery and understanding. *Nat Rev Mol Cell Biol* 13: 195–203
2. Vanhaesebroeck B, Guillermet-Guibert J, Graupera M, Bilanges B (2010) The emerging mechanisms of isoform-specific PI3K signalling. *Nat Rev Mol Cell Biol* 11:329–341
3. Engelman JA (2009) Targeting PI3K signalling in cancer: opportunities, challenges and limitations. *Nat Rev Cancer* 9:550–562
4. Fruman DA, Meyers RE, Cantley LC (1998) Phosphoinositide kinases. *Annu Rev Biochem* 67:481–507
5. Katso R, Okkenhaug K, Ahmadi K, White S, Timms J, Waterfield MD (2001) Cellular function of phosphoinositide 3-kinases: implications for development, homeostasis, and cancer. *Annu Rev Cell Dev Biol* 17:615–675
6. Peng W, Zheng-Chao T, Long Z-J, Liu Q, Gui L (2016) Discovery of 2-(2-aminopyrimidin-5-yl)-4-morpholino-N-(pyridin-3-yl)quinazolin-7-amines as novel PI3K/mTOR inhibitors and anticancer agents. *Eur J Med Chem* 108:644–654
7. Lv X, Ma X, Hu Y (2013) Furthering the design and the discovery of small molecule ATP-competitive mTOR inhibitors as an effective cancer treatment. *Expert Opin Drug Discovery* 8:991–1012
8. Patil SA, Patil SA, Patil R, Hashizume R (2016) Imidazoquinolines: recent developments in anticancer activity. *Mini Rev Med Chem* 16:309–322
9. Patel H, Kukol A (2016) Evaluation of a novel virtual screening strategy using receptor decoy binding sites. *J Negat Results Biomed* 15:15
10. Lionta E, Spyrou G, Vassilatis DK, Cournia Z (2014) Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem* 14:1923–1938
11. Kubinyi H (1997) QSAR and 3D QSAR in drug design part 1: methodology. *Drug Discov Today* 2:457–467
12. Perkins R, Fang H, Tong W, Welsh WJ (2003) Quantitative structure-activity relationship methods: perspectives on drug discovery and toxicology. *Environ Toxicol Chem* 22:1666–1679
13. Puzyn T, Leszczynski J, Cronin MTD (2010) In: Cronin MTD (ed) recent advances in QSAR studies: methods and applications. Springer, Netherlands
14. Winkler DA (2002) The role of quantitative structure–activity relationships (QSAR) in biomolecular discovery. *Brief Bioinform* 3: 73–86
15. Mustyala KK, Chitturi AR, Naikal James PS, Vuruputuri U (2012) Pharmacophore mapping and in silico screening to identify new potent leads for A(2A) adenosine receptor as antagonists. *J Recept Signal Transduct* 32:102–113
16. Acharya C, Coop A, Polli JE, Mackerell Jr AD (2011) Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr Comput Aided Drug Des* 7:10–22
17. Güner OF (2002) History and evolution of the pharmacophore concept in computer-aided drug design. *Curr Top Med Chem* 2:1321–1332
18. Lee CH, Huang HC, Juan HF (2011) Reviewing ligand-based rational drug design: the search for an ATP synthase inhibitor. *Int J Mol Sci* 12:5304–5318
19. Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA (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
20. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, Dearden J, Gramatica P, Martin YC, Todeschini R, Consonni V, Kuz'min VE, Cramer R, Benigni R, Yang C, Rathman J, Terfloth L, Gasteiger J, Richard A, Tropsha A (2014) QSAR modeling: where have you been? Where are you going to? *J Med Chem* 57:4977–5010
21. Schrödinger LLC (2010) Glide, Version 5.6. New York, NY
22. Zhang J, Lv X, Ma X, Hu Y (2017) Discovery of a series of N-(5-(quinolin-6-yl)pyridin-3-yl)benzenesulfonamides as PI3K/mTOR dual inhibitors. *Eur J Med Chem* 127:509–520
23. Lv X, Ying H, Ma X, Qiu N, Wu P, Yang B, Hu Y (2015) Design, synthesis and biological evaluation of novel 4-alkynyl-quinoline derivatives as PI3K/mTOR dual inhibitors. *Eur J Med Chem* 99: 36–50
24. Zhang J, Ma X, Lv X, Li M, Zhao Y, Liu G, Zhan S (2017) Identification of 3-amidoquinoline derivatives as PI3K/mTOR dual inhibitors with potential for cancer therapy. *RSC Adv* 7: 2342–2350
25. Tsitsanou KE, Drakou CE, Thireou T, Vitlin Gruber A, Kythreoti G, Azem A, Fessas D, Eliopoulos E, Iatrou K, Zographos SE (2013) Crystal and solution studies of the “Plus-C” odorant-binding protein 48 from *Anopheles gambiae*: control of binding specificity through three-dimensional domain swapping. *J Biol Chem* 288: 33427–33438
26. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 47:1739–1749
27. Li J, Abel R, Zhu K, Cao Y, Zhao S, Friesner RA (2011) The VSGB 2.0 model: a next generation energy model for high resolution protein structure modelling. *Proteins* 79:2794–2812
28. Peddi SR, Sivan SK, Manga V (2018) Molecular dynamics and MM/GBSA-integrated protocol probing the correlation between biological activities and binding free energies of HIV-1 TAR RNA inhibitors. *J Biomol Struct Dyn* 36:486–503
29. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) Semi analytical treatment of solvation for molecular mechanics and dynamics. *J Am Chem Soc* 112:6127–6129
30. Debnath T, Majumdar S, Kalle A, Aparna V, Debnath S (2015) Identification of potent histone deacetylase 8 inhibitors using pharmacophore-based virtual screening, three-dimensional quantitative structure–activity relationship, and docking study. *Res Rep Med Chem* 5:21–39
31. Veerasamy R, Rajak H, Jain A, Sivadasan S, Varghese CP, Agrawal RK (2011) Validation of QSAR models—strategies and importance. *Int J Drug Des Discov* 2:511–519
32. Ihle NT, Paine-Murrieta G, Berggren MI, Baker A, Tate WR, Wipf P, Abraham RT, Kirkpatrick DL, Powis G (2005) The phosphatidylinositol-3-kinase inhibitor PX-866 overcomes resistance to the epidermal growth factor receptor inhibitor gefitinib in A-549 human non-small cell lung cancer xenografts. *Mol Cancer Ther* 4:1349–1357