Chapter - 3

3D QSAR BASED DESIGNING OF PHARMACOPHORE AND PHARMACOPHORE BASED VIRTUAL SCREENING

3.1 Introduction

Pharmacophore based approaches have become one of the major tools in drug discovery after the past century's development (Yang., 2010). Pharmacophore based drug designing approach is an important concept for the rational drug designing bases on the ideas that any drug or small molecules are active against particular receptors only in the presence of certain key features (Functional groups) which favourably binds to the receptor binding sites of the target protein

In order to apply this approach, one should have wide activity range of training set ligands basically grouped in three main category, Highly active compounds, moderate active compound and least active compounds. The main objective of the this method is to identify the common features or essential 3 dimensional geometrical arrangements of atoms or functional groups to exhibit biological response (Faulon et al., 2008; Guner, 2002; Balatsos et al., 2012; Dalkas et al., 2013).Since the biological active conformer of particular active compound is not known, training set data should have a set of low energy conformation of each compounds (Li et al., 2009).This is also advantageous as the lowest active

conformer does not necessarily need to be active conformer but the energy of the active conformer is not much greater than the energy of global minima.

Recognition process between ligand and model is based on spatial distribution of certain structural features of active site being complimentary to those of the interacting ligands; and the features common to the ligands would provide the information about the active site. A pharmacophore mapping is the essential step towards understanding of receptor-ligand recognition process and is established as one of the successful computational tools in rational drug design (Kurogi et.al., 2001; Guner et al., 2004). This involves the identification of a three dimensional arrangement of functional groups which a molecule must possess to be recognized by the receptor. Further, a model is generated by finding chemically important functional groups that are common to the molecules that bind. Pharmacophore can be derived by direct analysis of the structure of known ligand either in the most stable conformer or in the form observed for complexing with the target protein.

In the present study, a three-dimensional pharmacophore model for PDK-1 kinase inhibitor has been developed. The generated model is further utilized for screening of potentially active candidates from NCI (http://dtp.nci.nih.gov/screening.html) and Maybridge (www.maybridge.com) databases. The efficacy of these compounds is further validated by molecular docking method.

3.2 Material and methods

3.2.1 General methodology

All pharmacophore models generation and Hypo1-based virtual screening were performed using the:-

3.2.1.1 Hypogen:- Implemented in Catalyst **(Catalyst 4.1, Molecular Simulations Inc., San Diego, CA).** Hypogen attempts to derive the SAR models for set of molecules for which activity value (K_i)on a given biological target are available. Hypogen optimizes the hypothesis that are present in the highly active compounds in the training set, missing among the least active (or inactive) ones. It constructs the simplest hypothesis that best correlates the test set data.

3.2.1.2 Fisher Randomization Test **:-** Fisher randomization was done by **CatScamble** program implemented in Catalyst. It was used to evaluate the statistical relevance of Hypo1 (Best Hypothesis)

3.2.1.3 Lipinski Filtration :- Was performed by using Pipeline Pilot Studio (**SciTegic, Inc., San Diego, CA)**. Lipinski rules are set of guidelines based on structural properties of compounds, used for fast calculation of drug like properties of a molecule.

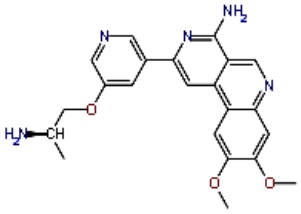
3.2.1.4 LigandFit **:-** Docking studies were achieved using Discovery Studio 2.5 **(Accelrys Inc., San Diego, CA)**. LigandFit is shape based docking method used for more accurately docking of ligands in to the active sites of protein.

3.2.2 Data set for pharmacophore analysis

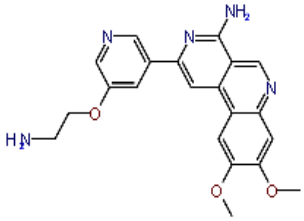
A set of 83 different compounds has been collected from different references (Kyung et al., 2009; Thomas et al., 2010; Zhu et al., 2004; Gopalsamy et al., 2007), which are reported to be inhibitors of PDK-1 kinase. The inhibitory activity of these compounds, expressed as IC_{50} (i.e., concentration of compound required to inhibit 50% of PDK-1 kinase activity) was taken for the whole process. The IC_{50} values spanned across a wide range from 3.0 to 65, 000 nM. Amongst 83 compounds, 21 compounds were selected as training set compounds and the rest compounds were taken as a test set compounds. Chemical structures of all training set compounds are shown in Figure 3.1.The selection of the training set and test set were according to the following rules: a) structural diversity among molecules. b) Both training set and set cover a wide range of activity. c) Highest active compounds should include in the training set because they provide critical information for pharmacophore generation. The geometry of all compounds was built by using AccelrysDiscoveryStudio2.5, (Accelrys, San Diego, Calif, USA, 2009). All the compounds were minimized using the steepest descent algorithm (Fetcher et al., with a convergence gradient value of 0.001 kcal/ mol and a family of representative conformations was generated by fast conformational analysis methods using Poling minimize algorithm (Smellie et al.,1995) and CHARMM force field parameters (Brooks et al., 1983). A large number of conformations of each compound were generated within an energy threshold of 20.0 kcal/ mol above the global energy minimum.

3.2.3 Pharmacophore modelling

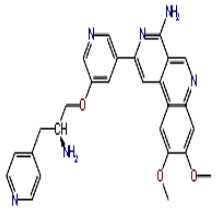
Based on the conformations for each compound, HypoGen module of DiscoveryStudio2.5 was used to construct the possible pharmacophore models (Li et al., 2000). Instead of using the lowest energy conformation of each compound, all the conformational models for each compound in training set were used in DiscoveryStudio2.5 for pharmacophore hypothesis generation. The training set compounds (21 in number) associated with their conformations were submitted to Discovery Studio 2.5. 3D QSAR pharmacophore Generation (HypoGen).The HypoGen module generates hypothesis with features common in active molecule and missing from inactive molecule.



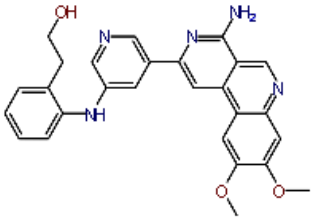
Compound _1 (IC_{50}= 3)



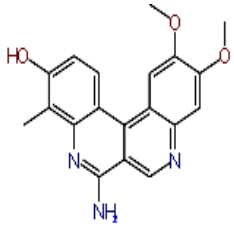
Compound_2 (IC_{50}=5)



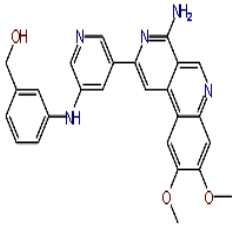
Compound_3 (IC_{50}=9)



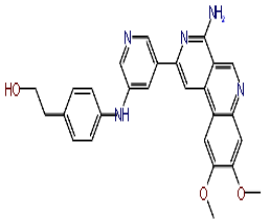
Compound _4 (IC_{50}=24)



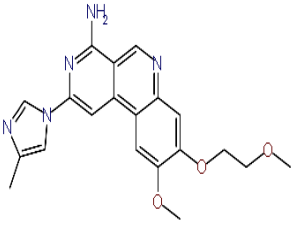
Compound _5 (IC_{50}=40)



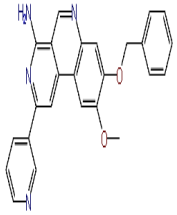
Compound_6 (IC_{50}=67)



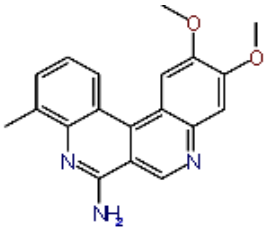
Compound_7 (IC_{50}= 91)



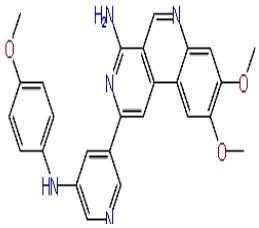
Compound_8 (IC_{50}=110)



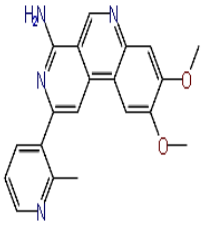
Compound_9 (IC_{50}=150)



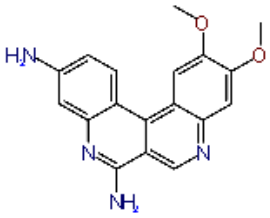
Compound_10 (IC_{50}=260)



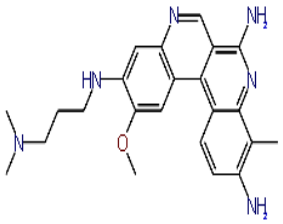
Compound_11 (IC_{50}=380)



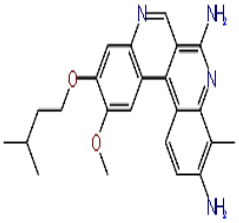
Compound_12 (IC_{50}= 530)



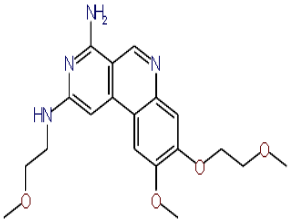
Compound _13(IC_{50}=670)



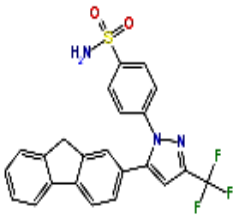
Compound_14(IC_{50}= 970)



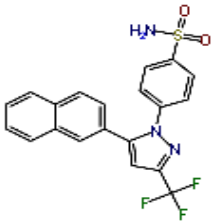
Compound_15(IC_{50}=1180)



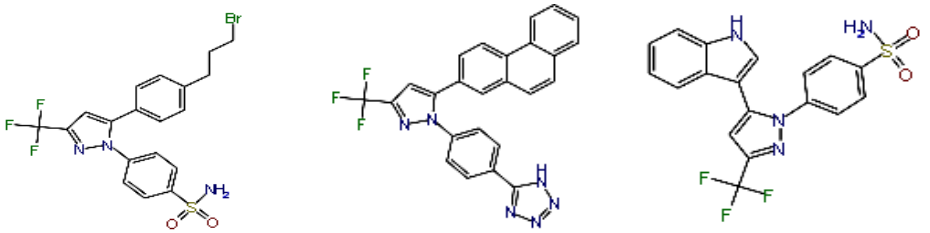
Compound_16 (IC_{50}=7500)



Compound_17 (IC_{50}=16,000)



Compound _18 (IC_{50}=24,000)



Compound _19(IC_{50}=38,000) Compound _20(IC_{50} =52,000) Compound_21(IC_{50}=65,000)

Figure. 3.1: Chemical structures and activity data (IC_{50} values, nM) of 21 training set molecules Applied for HypoGen pharmacophore generation

3.2.4 Model validation

The statistical parameters, such as the cost value determine the significance of the model. . The best model was selected on the basis of significant statistical parameters, like high correlation (r), lowest total cost, and lower value of RMSD, and the value of the total cost should be closer to the fixed cost and much away from null cost. Another parameter, configuration cost is also important for determination of significance of model. It should be <17.

Best hypothesis Hypo1 was also validated by test set validation method, Fischer's randomization validation and decoy set method. Ligand pharmacophore mapping protocol is used for estimating the activity of the entire 63 test set compounds.

3.2.5 Decoy Set validation

Results of test set validation method could only indicate that the generated pharmacophore model (Hypo 1) has high efficiency in picking the active molecules but it may not be confirmed, whether this also shows efficiency in picking the inactive molecules. For outcome this, decoy sets validation method was used to evaluate the efficiency of Hypo1 by calculating the GH (goodness of hit list) and EF (enrichment factor). A data set of small molecule has been generated by decoy set finder 1.1 which including 1980 molecules with unknown activity and 20 active molecules were taken to prepared a decoy set of 2000 molecules.

GH (goodness of hit list) and EF (enrichment factor) were calculated by the equations given below:

$$EF = (Ha/Ht)/(A/D)$$

$$GH = \{[Ha * (3A + Ht)]/4HtA\} * [1 - (Ht - Ha)/(D - A$$.....Eq.1

Where Ht=total no of molecules in hit list, Ha=total active molecules present in the hit list, A= total active molecules present in database, D= total molecules present in decoy set. The range of GH score varies from 0 to 1. GH score 0 means a null model was generated while the GH score 1 means generation of an ideal model. Although when the GH score is higher than 0.7 reflect the generation of a very good model. The EF and GH are found to be 69.23 and 0.73, (shown in Table 3.1) indicates that the generated pharmacophore model is rational for virtual screening.

Table 3.1: Statistical parameters from the validation of the pharmacophore model by mean of decoy set

No	Parameters	Values
1	Total number of molecule in database (D)	2,000
2	Total number of active molecules in database (A)	20
3	Total no of hit molecules from database (Ht)	26
4	Total no of active molecules in hit list (Ha)	18
5	% Yield of actives [(Ha/A) 100]	69.23
6	% Ratio of actives in the hit list [(Ha/A) x 100]	90
7	Enrichment factor (EF) [Ha x D)/(Ht x A)]	69.23
8	False negatives [A- Ha]	2
9	False positives [Ht-Ha]	8
10	GH score[a] (goodness of hit list)	0.73

[a] [(Ha/4HtA) (3A+Ht) x [1-(Ht- Ha)/(D-A)]]

3.2.6 Virtual screening and ADMET analysis

The validated 3D QSAR pharmacophore model Hypo1 was used as a 3D structural query for retrieving potent compounds from NCI database (Milne et al., 1994) and Maybridge database having 23,8819 molecules and 2,000 molecules respectively. A systematic diagram of virtual screening was shown in Figure. 3.2

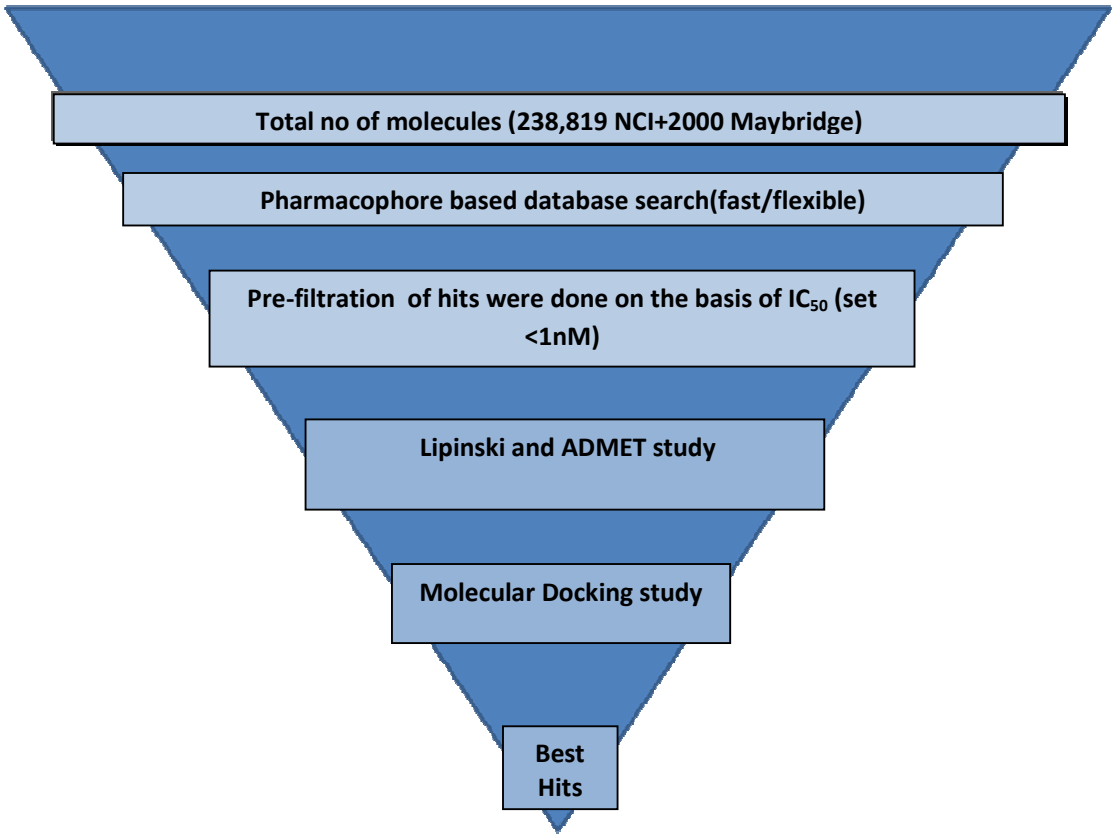


Figure 3.2: Diagrammatic representation of Virtual screening protocol.

3.3. Results and Discussion

3.3.1 Pharmacophore Modelling

Ten hypotheses were produced by 3D QSAR Pharmacophore generation module of AccerlysDiscoveryStudio2.5 through 21 training sets compounds (Table 3.2).Hypo1 was the most significant hypothesis characterized by high cost difference (168.48433), lowest root mean square deviation (RMSD=1.0719), and best correlation coefficient (r=0.96906). The fixed cost and the null cost values were 77.5618 and

258.686, respectively, with total cost value 90.2017 for Hypo1.This observation was much lower than null cost and closer to the fixed cost.

Table 3.2: Information of statistical significance and predictive power presented in cost values measured in bits for top 10 hypotheses as a result of automated HypoGen pharmacophore generation process.

Hypothesis no.	Total cost	Cost difference[a] (Total cost-null cost)	Error	RMS	Correlation (r)	Features[b]
1	90.2017	168.4843	73.0296	1.0719	0.96906	HBA, HBA, HBD, HyA
2	91.6361	167.0499	74.8623	1.1505	0.96423	HBA, HBA , HyA, HyA
3	92.4311	166.2549	74.9287	1.1532	0.96414	HBA, HBA , HBD, HyA
4	93.3537	165.3323	76.4684	1.2151	0.96002	HBA, HBA , HBD, HyA
5	93.5794	165.1066	76.5881	1.2198	0.95971	HBA, HBA , HyA, HyA
6	95.5213	163.1647	78.8279	1.3043	0.95376	HBA, HBA , HyA,HyA
7	96.8603	161.8257	79.8688	1.34181	0.951035	HBA, HBD, HyA, HyA
8	99.4115	159.4565	79.8688	1.34181	0.951035	HBA, HBD , HyA, HyA
9	99.7217	158.9643	83.011	1.44906	0.942609	HBA, HBD, HyA, HyA
10	99.7794	158.9066	83.1717	1.45431	0.942169	HBA, HBA , HBD, HyA

[a] The cost difference between null cost and total cost ;null cost is 258.686 bits ;fixed cost is 77.5618 bits; configuration cost is 15.4729 bits.

[b] Abbreviation used for features: HBA; H-bond acceptor, HBD; H-bond donor; HyA, hydrophobic aliphatic.

The best hypothesis (Hypo1) consist of four features, i.e. two hydrogen bond acceptor (HBA), one hydrogen bond donor (HBD) and one hydrophobic aliphatic feature (HyA). Figure. 3.3 3a & 3b represent features of best pharmacophore Hypo1 and distance and the angular constraints between the features in the best pharmacophore (Hypo1). The experimental and estimated activities by the best pharmacophore hypothesis (Hypo1) for the 21 training set compounds are shown in (Table 3.3). Figure 3.4a represents the top scoring hypothesis Hypo1, mapped to the most active compound 1 (IC_{50} = 3 nM) and Figure. 3.4b represents the least active compound 21(IC_{50} = 65,000 nM) of the training set.

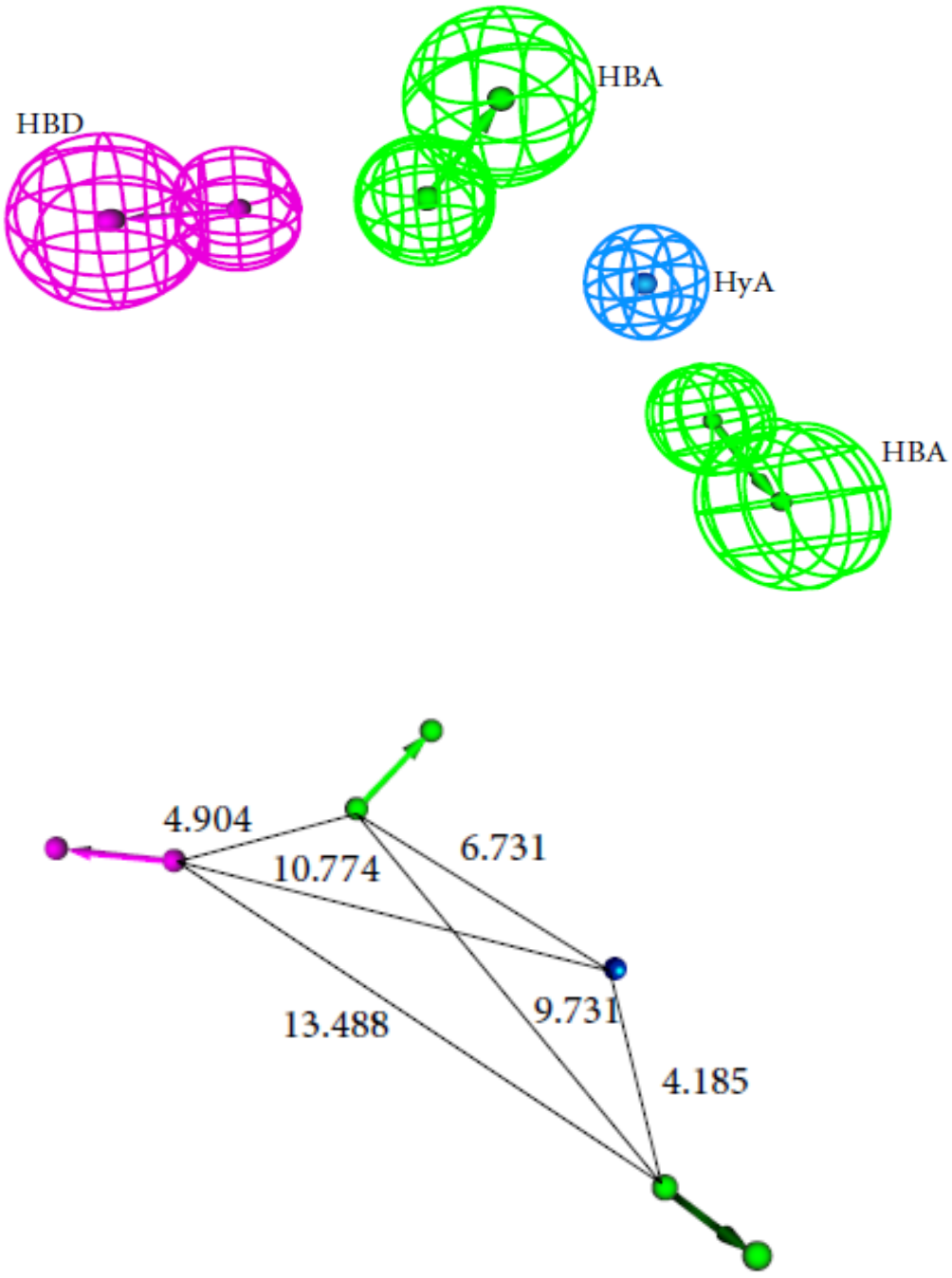
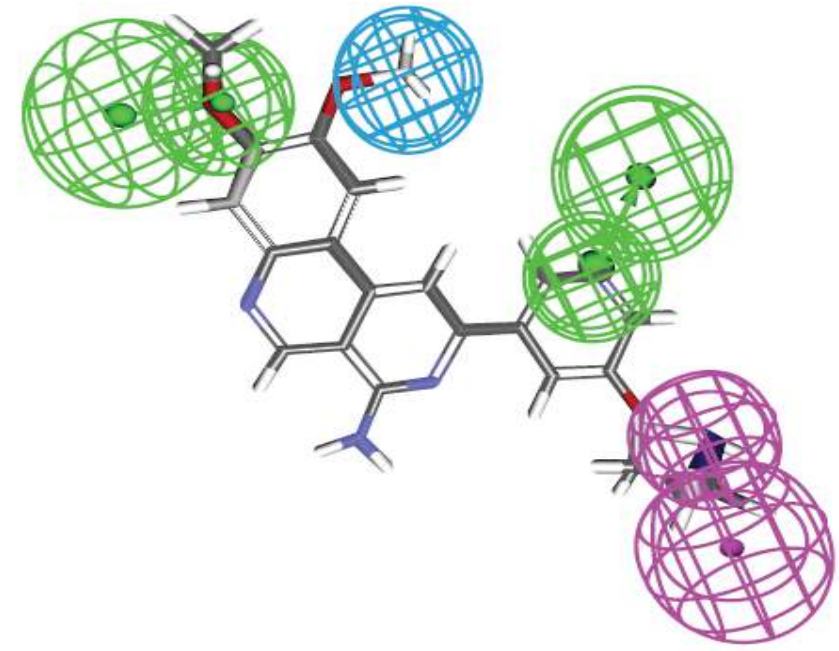


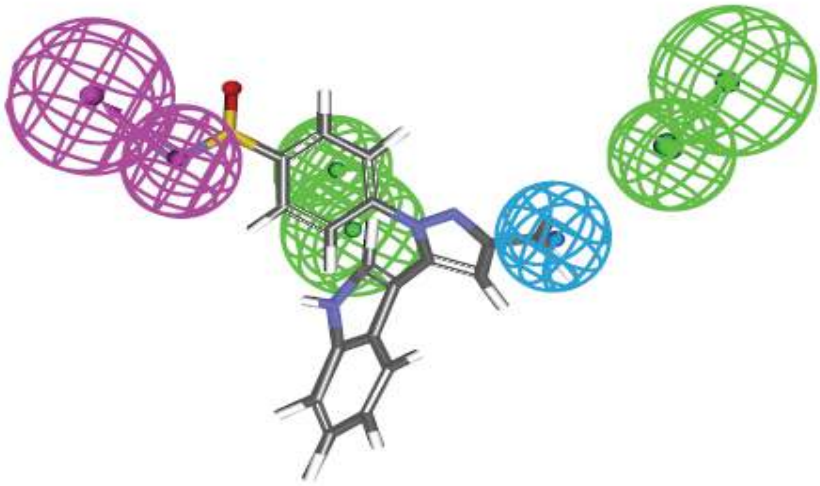
Figure 3.3: (a) Hypo1(best pharmacophore) generated by Hypogen (3D QSAR pharmacophore protocol);(b) pharmacophore model with distance between chemical features.

Table 3.3: Experimental biological data and estimated IC_{50} of training set molecules based on pharmacophore model Hypo1

Compound no	IC50 value(nM)		Error	Uncertainty	Fit Value
	Experimental	Expected			
1	3	2.4	-1.3	2	8.60
2	5	4.9	-1.0	2	8.29
3	9	14	+1.6	2	7.83
4	24	20	-1.2	2	7.67
5	40	130	+3.3	2	6.86
6	67	39	-1.8	2	7.40
7	91	110	+1.2	2	6.94
8	110	290	+2.6	2	6.52
9	150	230	+1.6	2	6.61
10	260	1000	+4.0	2	5.97
11	380	160	-2.4	2	6.71
12	530	1100	+2.0	2	5.94
13	670	1000	+1.5	2	5.97
14	690	420	-2.3	2	6.36
15	1200	390	-3.1	2	6.39
16	7500	27000	+2.7	2	5.55
17	16000	21000	+1.6	2	4.65
18	24000	21000	-1.1	2	4.65
19	38000	21000	-1.8	2	4.65
20	52000	93000	+1.8	2	4.01
21	65000	21000	-3.0	2	4.05



(a)



(b)

Figure 3.4 (a) Highest active compound (compound_1, IC_{50}= 3.0 nM) mapped in best pharmacophore model; (b) Least active compound (Compound_21, IC_{50} = 65,000) mapped on best pharmacophore model (Hypo1)

The most active compound exhibited a good fit with all features of the pharmacophore hypothesis, Hypo1, where in the least active compound had hydrogen bond acceptor feature missing. Based on this, it may be concluded that two HBA features are important for PDK-1 kinase inhibitory activity.

3.3.2 Cost analysis

In addition to generating a hypothesis, Hypogen also provides two theoretical costs (represented in bit units) to help assess the validity of the hypothesis. The first is fixed cost (cost of an ideal hypothesis), which represents the simplest model that fits all data perfectly. The second one is the null cost (cost of null hypothesis), which represents the highest cost of a pharmacophore with no features and which estimates activity to be the average of the activity data of the training set molecules. They represent the upper and lower limits for the hypothesis that are generated. A meaningful pharmacophore hypothesis may be generated when the difference between null hypothesis and the fixed hypothesis is large; a value of 40-60 bits may indicate that it has 75-90% probability of correlating the data. Other two parameters that also determine the quality of any pharmacophore are configuration cost or entropy cost and error cost. The configuration cost depends on the complexity of the pharmacophore and should have value <17 whereas the error cost dependent on the root mean square difference between the estimated and the actual activity of the training set. The difference between total fixed cost and the null cost of the Hypo1 is 168.4843, which is more than 40-60, and defines that there is more than 90% probability of data correlation. Noticeably, the total cost of Hypo1 was much closer to the

fixed cost than to the null cost. Furthermore, a high correlation coefficient of 0.96906 was observed with RMS value of 1.0719 and the configuration cost of 15.4729, demonstrating the development of a reliable pharmacophore model with high predictivity.

3.3.3 Validation of Pharmacophore Model

3.3.3.1 Test Set Validation

The test set method is for examining whether the pharmacophore model is capable of predicting the activities of external compounds of the test set series. The test set contains 62 compounds structurally different from the training set molecules. All the test set molecules were prepared in the same way as that for the training set molecules. Test set validation was done using Ligand Pharmacophore mapping protocol. The test set of 62 compounds were mapped on the Hypo1.It was observed that pharmacophore model performed well in estimation of activity of test set compounds, with a significant predictive correlation value (r= 0.87) for 62 test set compounds shown in Figure 3.5. The experimental and estimated activities of test set compounds mapped on the best hypothesis (Hypo1) are shown in (Table 3.4).

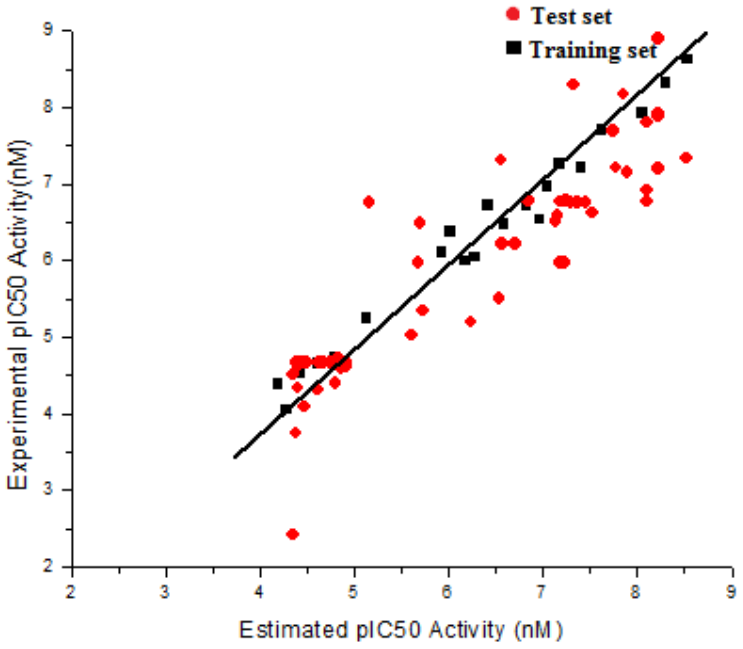


Figure. 3.5: Regression plot of 62 test set molecule against Hypo 1

Table 3.4: Experimental biological data and estimated IC_{50} of test set molecules based on pharmacophore model Hypo 1.

Compound no	IC50(nM)		Error	Unc	Fit Value
	Experimental	Expected			
Compound_22	6	1.3	-4.6	2	8.87
Compound_23	47	5.2	-9.0	2	8.26
Compound_24	14	6.8	-2.0	2	8.14
Compound_25	6	11.9	+1.8	2	7.89
Compound_26	6	13.3	+2.2	2	7.85
Compound_27	8	15.9	+1.9	2	7.78
Compound_28	18	20.6	+1.4	2	7.66
Compound_29	3	46.8	+15.6	2	7.30
Compound_30	283	49.0	- 5.8	2	7.29
Compound_31	17	62.2	+3.6	2	7.18
Compound_32	6	63.9	+10.5	2	7.17
Compound_33	13	71.0	+5.5	2	7.13
Compound_34	8	121.3	+15.2	2	6.89
Compound_35	144	159.2	+1.1	2	6.79
Compound_36	142	163.7	+1.1	2	6.76
Compound_37	57	166.9	+2.9	2	6.75
Compound_38	8	169.0	+21.1	2	6.75
Compound_39	65	173.1	+2.7	2	6.74
Compound_40	6,923	173.4	-39.9	2	6.74
Compound_41	35	173.7	+4.9	2	6.74
Compound_42	51	174.9	+3.4	2	6.73
Compound_43	43	176.8	+4.1	2	6.73

Compound_44	30	238.0	+7.9	2	6.60
Compound_45	70	264.1	+3.8	2	6.56
Compound_46	74	308.3	+4.2	2	6.49
Compound_47	2,033	332.1	+6.1	2	6.46
Compound_48	270	604.6	+2.2	2	6.20
Compound_49	200	604.6	+3.0	2	6.19
Compound_50	2,120	1,063.9	-1.9	2	5.95
Compound_51	60	1,070.4	+17.8	2	5.95
Compound_52	66	1,084.5	+16.4	2	5.94
Compound_53	290	3,107.8	+10.7	2	5.48
Compound_54	1,864	4,488.0	+2.4	2	5.33
Compound_55	580	6,336.0	+10.9	2	5.18
Compound_56	2,490	9,421.4	+3.8	2	5.00
Compound_57	15,000	19,013.5	+1.3	2	4.70
Compound_58	32,000	21,429.9	-1.5	2	4.65
Compound_59	38,000	21,429.8	-1.8	2	4.65
Compound_60	18,000	21,429.7	-1.2	2	4.65
Compound_61	32,000	21,430.2	-1.5	2	4.65
Compound_62	42,000	21,430.7	-1.9	2	4.65
Compound_63	32,000	21,430.9	-1.5	2	4.65
Compound_64	23,000	21,433.3	-1.1	2	4.65
Compound_65	15,000	21,434.2	-1.4	2	4.65
Compound_66	18,000	21,435.1	+1.2	2	4.65
Compound_67	22,000	21,436.7	-1.0	2	4.65
Compound_68	24,000	21,444.1	-1.1	2	4.65
Compound_69	21,000	21,446.6	+1.0	2	4.65

Compound_70	12,000	21,449.9	+1.8	2	4.65
Compound_71	42,000	21,453.4	+1.3	2	4.65
Compound_72	17,000	21,454.8	+1.3	2	4.65
Compound_73	16,000	21,493.5	+1.3	2	4.64
Compound_74	34,000	21,538.8	+1.6	2	4.64
Compound_75	18,000	21,619.6	+1.2	2	4.64
Compound_76	12,000	24,152.4	+2.0	2	4.59
Compound_77	40,000	24,273.3	- 1.6	2	4.59
Compound_78	14,000	25,463.7	+1.8	2	4.57
Compound_79	45,000	30,669.3	-1.5	2	4.49
Compound_80	16,000	39,286.2	+2.4	2	4.38
Compound_81	40,000	46,181.3	+1.1	2	4.31
Compound_82	25,000	48,798.6	+1.9	2	4.29
Compound_83	34,000	80,163.1	+2.3	2	4.07

Further, another validation method was used to characterize the quality of the hypothesis using error ratio, which is the difference between estimated activity and experimental activity. Also an error ratio ≤ 10 depicts that there is no more than one order difference between estimated and experimental activity values not more than one order. The best hypothesis (Hypo1) exhibited an error value ≤ 10 for 53 compounds out of 62 compounds. Only 9 compounds (compound_29, compound_32, compound_34, compound_38, compound_40, compound_51, compound_52, compound_53, compound_55) with values >10 were considered as outliers and rejected. The most potent compound_22 of test set (IC_{50}=6 nM) was mapped with Hypo1 (Figure 3.6). It was observed

that best hypothesis (Hypo1) mapped very well, also all the chemical features of this compound matched and the estimated activity of this compound had an IC_{50} value of 1.3 nM. Based on these results, it was confirmed that one HBD, two HBA and one HyA (hydrophobic aliphatic) features are essential for PDK-1 inhibitory activity.

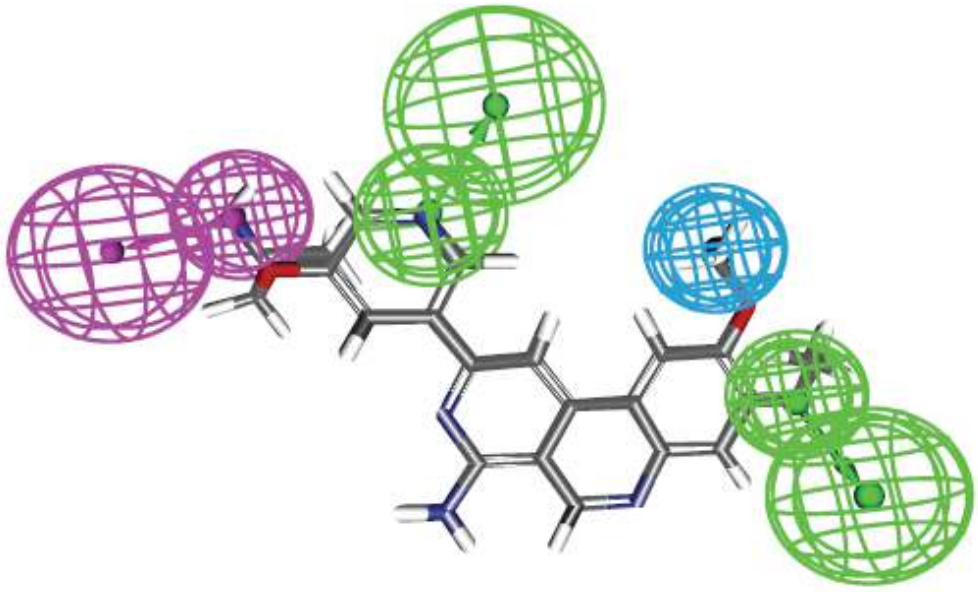


Figure. 3.6: Most active test set compound (compound_22) mapped in best pharmacophore model (Hypo1)

3.3.3.2 Fisher Validation

Fischer randomization test method was used to evaluate the statistical relevance of Hypo1 by using the CatScramble program. The confidence level was set to 95%. The CatScramble program generated 19 random spreadsheets to construct hypothesis using exactly the same conditions as used in generating the original pharmacophore hypothesis. Total cost of

19 pharmacophore hypothesis generated randomly and the original pharmacophore hypothesis are also presented in (Figure. 3.7). In Figure 3.7 it is observed that an original hypothesis is far more superior to the 19 random hypotheses. This result provides the 95% confidence of the proposed hypothesis.

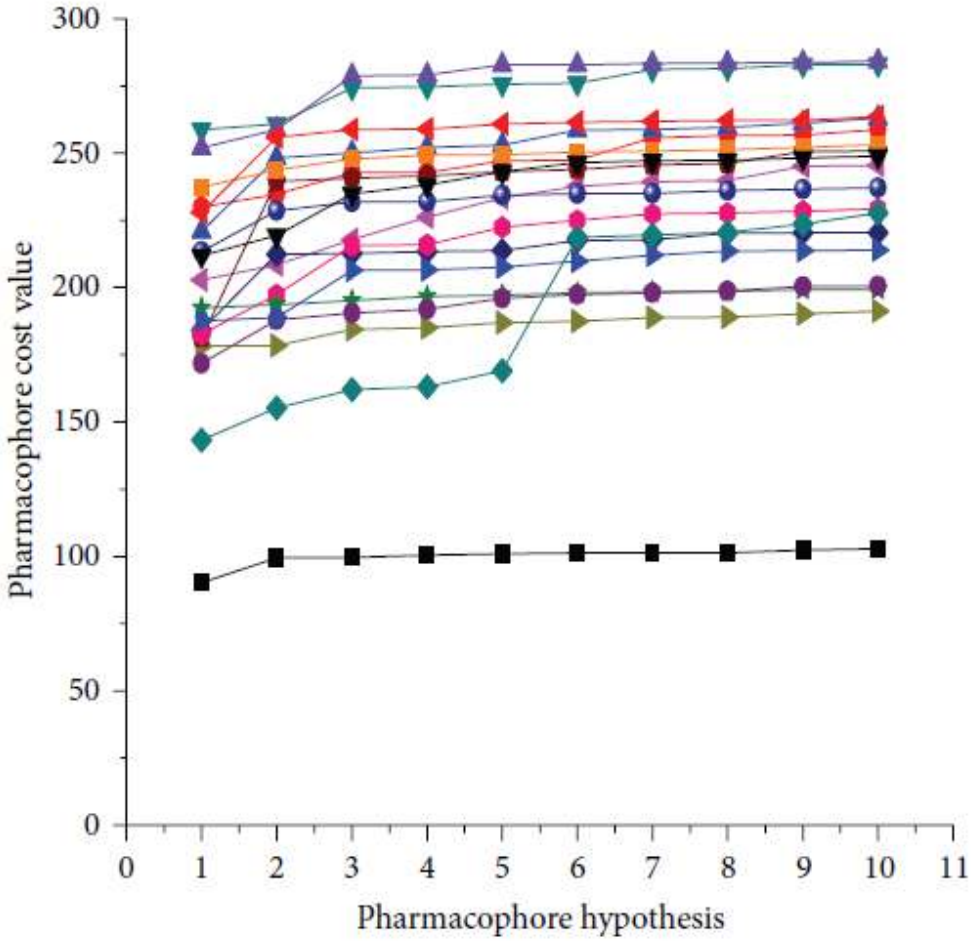


Figure. 3.7: The difference in costs between the Hypogen runs and scrambled runs. The 95% confidence level was selected.

3.4 Pharmacophore Based Virtual Screening

The validated 3D QSAR pharmacophore model Hypo1 was used as a 3D structural query for retrieving potent compounds from NCI database and Maybridge database having 23,8819 molecules and 2,000 molecules respectively. A total of 8,833 compounds exhibited good mapping with Hypo1 using fast and flexible search method. Out of 8,833 compounds 8,530 compounds were from NCI and 333 compounds were from Maybridge database. Out of these 8,833 molecules, 2033 molecules having their IC_{50}<1 μM has been selected for study. These hit compounds were further screened by using Lipinski's rule of five, to evaluate them drug similarity, and a total of 1,613 molecules passed this evaluative process. These 1,613 molecules were further used for the ADME studies. Only 842 molecules were passed from the ADMET filtration. We selected only those molecules for further molecular docking study those having estimated activity <=0.5 μM. Only 43 molecules were satisfied this conditions, hence further molecular docking study has been done for these successful molecules.

3.5 Molecular Docking studies

For further refining the retrieved hits and evaluating the binding mode between compounds and proteins, all compounds and compound_1 were docked into the binding site of PDK-1(Medina et al., 2010) (PDB entry:1UU7) (Komander et al., 2004) by using LigandFit (Venkatachalam et al., 2003) docking method implemented in Discovery Studio 2.5 program package. Before docking of all the molecules, compound_1

(most active compound of the training set) was docked into the active site of PDK-1.

i) Compound_1 docking Details

Compound_1 has shown the docking energy of - 64.5 kcal/mol and RMSD value of 0.841. This depicts that LigandFit docking method reproduced the original binding mode, so for the further docking study we used the LigandFit docking method. It also showed the hydrogen bond interactions with important residues of like Lys111, Asp 230, Ala 162 & Tyr 161 as shown in Figure 3.8 (a).

ii) Other Compounds docking Details

All 43 molecules were docked in to the active sites PDK-1 kinase, Only top 7 molecules those having high docking energy, different scaffolds, better hydrogen bond interactions with active site residues as well as lower estimated activity (<=0.19μM) were selected. The estimated activity, interaction energy as well as LignadFit score of all seven compounds along with compounds_1 are listed in Table 3.5. Finally, the three compounds (NSC_218341, NSC_24871 and NSC_211930) were selected for further analysis. Further compound NSC_211930, NSC_218341 and NSC_24871 were mapped all the features of the Hypo1. Amino group of Compound NSC_211930 formed the hydrogen bonding with Ala162 a hinge region amino acid. While the amide group formed the hydrogen bond with Asp223. Lys111 involved in cation-pi interaction. Compound NSC_24871 formed the hydrogen bond interaction with Lys111, Ser160 and Ala 162. The phenyl ring of compound is sandwiched in between the phenyl rings of Tyr161 and

Phe93 and they formed the pi-pi interaction. Tyr161 formed pi-pi interactions with phenyl ring of Compound_218342. While the carboxyl group involved in formation of two hydrogen bond with Lys111and Phe94.Phenolic oxygen was involved in formation of hydrogen bond with Ser162 & Ala162 amino acids. In all the cases Try 161 involved in forming pi-pi interaction with the phenyl ring of the compounds. 2D representation of molecular docking results of all three compounds is shown in the Figure 3.8 (b), (c) & (d). Lys111 formed two hydrogen bonds with the two different oxygen atom of phenyl groups of the Compound NSC_24871.Apart from this one phenolic oxygen formed the two hydrogen bonds with the two hinge regions in amino acids i.e. Ser160 & Ala162. These three compounds were retrieved from two databases (NCI& Maybridge) and exhibited good interactions with important amino acids in the active sites. Among all the three compounds, Compound NSC_218342 retrieved from the NCI database, showed good estimated activity, fit values and docking score as well as hydrogen bond interactions. Molecular docking results support that these molecules can be further taken as the potential leads for designing of novel PDK-1 inhibitors in the future.

Table 3.5: The estimated activity, interaction energy and LigandFit scoring results of top ranked four compounds obtained from the combination of Hypo1 based virtual screening and molecular docking studies.

Name	Est.(nM)	Interaction energy (kcal/mol)	Lig-1	Lig-2	-PLP1	-PLP	-PMF
Compound_1	3.00	64.40	4.23	4.98	73.60	68.90	103.10
NSC_218342	31.40	65.39	4.93	4.25	61.42	59. 79	83.91
NSC_24871	74.10	51.18	5.03	5.77	65.14	67.09	95.66
NSC_211930	82.00	57.30	6.17	6.23	79.00	71.43	75.76
NSC_84044	87.70	51.11	4.23	5.31	65.40	55.20	92.87
NSC_325657	93.80	55.47	5.84	6.16	75.72	73.50	106.54
NSC_343659	94.50	9.16	4.41	4.47	51.82	53.26	82.51
SB_01794	160.00	46.36	3.67	4.04	53.22	56.12	107.10

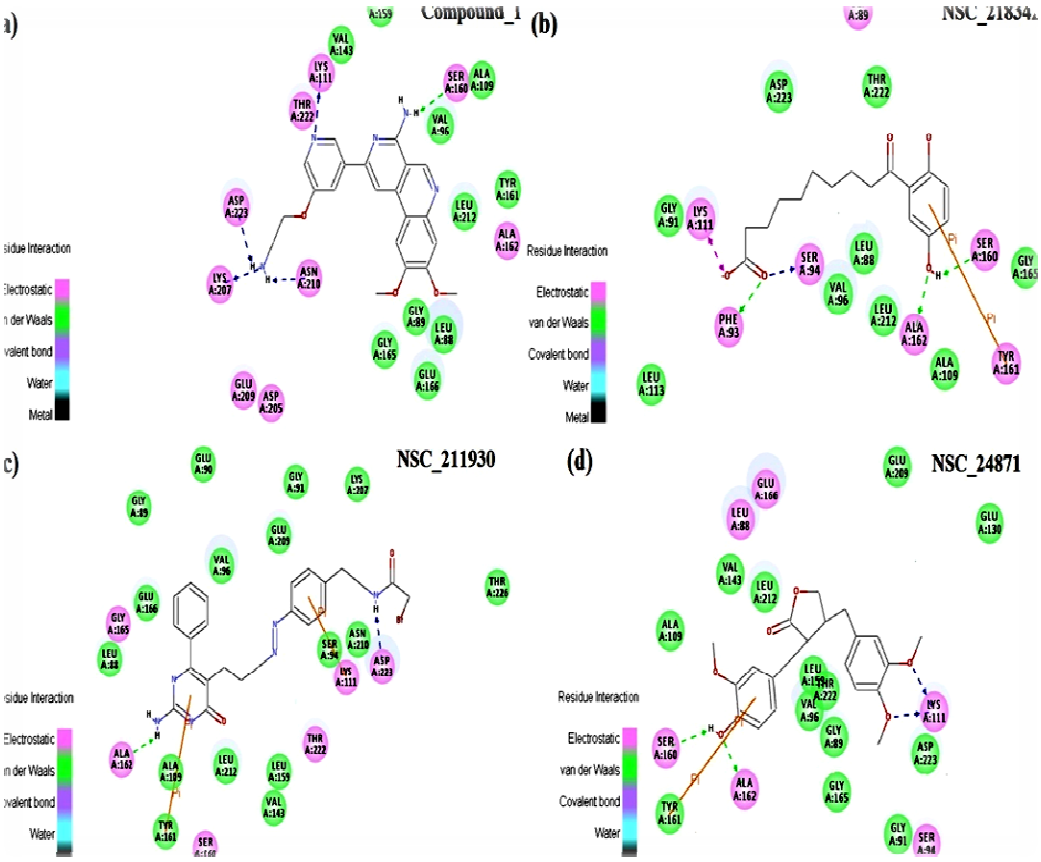


Figure 3.8: 2 D representation of top docking hits retrieved from database and most active compound (Compound_1)