Chapter - 5

MOLECULAR DOCKING AND *INSILICO* ADMET STUDY OF MYRICETIN AND THEIR ANALOGUES

5.1 Introduction

The computational based approach is one of the most recent and rapid developing techniques in pharmacokinetics prediction, ADMET (absorption, distribution, metabolism, excretion, toxicity) analysis, drug discovery as well as in toxicity prediction. Insilico quantitative analysis are now possible for several pharmacokinetic (PK) parameters, particularly absorption and distribution. The emerging Insilico prediction approaches are no worse than those made using in vitro tests, with the decisive advantage that much less investment in technology, resources and time is needed. In addition, and of critical importance, it is possible to screen virtual compounds. Some packages are able to handle thousands of molecules in a few hours. However, common experience shows that, in part at least for essentially irrational reasons, there is currently a lack of confidence in these approaches. An effort may be made more transparency, in order to improve the confidence of their consumers. It seems highly probable that in silico approaches may evolve rapidly, as did in vitro methods during the last decade. Past experience with the latter should be helpful in avoiding repetition of similar errors and in taking the necessary steps to ensure effective implementation. A general concern is

the lack of access to the large amounts of data on compounds no longer in development, but still kept secret by the pharmaceutical industry. Controlled access to these data could be particularly helpful in validating new in silico approaches.

The work presented in this chapter, is an attempt to screen 95% Myricetin similar compounds deposited in Pubchem database. Several computational methods like Lipinski filter molecular docking and *Insilico* ADME/Toxicity studies have been incorporated on the screen compounds to predict these molecules behavior as a putative future drug for cancer treatment.

5.2 MATERIAL AND METHODS

5.2.1 Chemical similarity search

The two dimensional chemical structure of natural flavonoid Myricetin (CID5281672) was retrieved from the NCBI PubChem database (http://www.pubchem.ncbi.nlm.nih.gov) and similarity search was performed on the chemical compounds deposited in the Pubchem database to retrieve the related compound and analogues. The search parameters were set at 95% similarity. 2D structures of all screened compounds were downloaded from pubchem database.

5.2.2 Preparation of Ligands

The three dimensional format of all filtered compounds were downloaded from Pubchem database in .sd file formt. Subsequently CharMM (Brooks et al., 1983) based force field was applied and further subjected to single step energy minimization using steepest descent method for 500 steps at RMS gradient of 0.01.

5.2.3 Lipinski Filter

The drug likeliness properties of all the retrieved compounds were evaluated by Lipinski drug filter (Lipinski et al., 2001) implemented in Accelrys Discovery Studio 2.5. This rule basically describes those molecular properties which is essential for a drug's pharmacokinetics in the human body and also provides the information concerning the deployment of the ligands as a drug molecule.

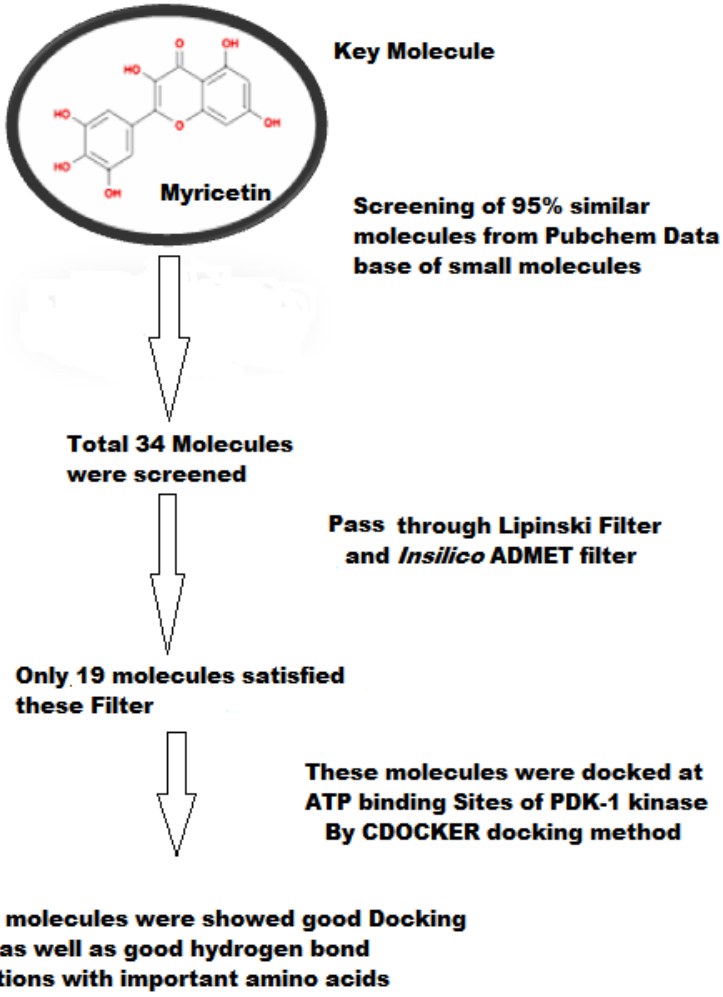


Figure. 5.1 Diagrammatic representation of Working Protocol used for Screening of analogues

5.2.4 Protein Preparation

The X ray crystal structure of PDK-1 (PDB id 1UU7) (Komander et al., 2004) taken in this study was retrieved from protein data bank (http://www.pdb.org), all the HETATMS were removed. Further the protein was subjected to two steps energy minimization to remove the bad steric clashes using steepest descent and conjugate gradient methods for 1000 steps at RMS gradient of 0.1 and 0.05 respectively. During the energy minimization process the backbone and side chain were fixed by applying the fixed atom constraint, and only hydrogen atoms were minimized. The CharmM force field (Brooks et al., 1983) was applied to the receptors. Fixed atom constrained were removed after the minimization. The receptor protein is divided into the protein part and crystal ligand part. The protein part was only selected and selections were made to "define selected molecule as receptor" under define and edit binding site, sub panel of the "Tool panel", where in, the protein is marked as receptor molecule. By selecting only the ligand part and further clicking on "Define sphere from selection" so that the crystal ligand can be used to define the binding site of 15 Angstroms on the receptor molecule. This 'input receptor molecule' is used as input parameter in the CDOCKER (Wu et al., 2003) protocol parameter explorer.

5.2.5 Molecular docking simulation

Molecular docking was performed by the CDOCKER docking method implemented in Discovery Studio 2.5. CDOCKER is a simulated annealing based molecular docking method. In this docking method ligands are treated as fully flexible while protein is kept rigid. The

minimized structure of all compounds were used as input ligand in the protocol explorer of CDOCKER. Each of them is given as input in another parameter meant for 'input ligands' and the protocol was run as many times as the number of inhibitors are selected for the experiment. The various conformations for ligand in this procedure were generated by using molecular dynamics. The generated initial structures for the ligand may be further refined using simulated annealing. The CDOCKER energy (-(protein-ligand interaction energies)) of best configuration docked into the receptor of all the selected natural inhibitors, which were calculated and compared with that of interacting residues at active site region with the crystallized inhibitors, PDK-1 kinase protein. Binding energy of protein and ligands were calculated by following calculation:

$$E_{binding} = E_{complex} - (E_{receptor} + E_{ligand})$$

5.2.6 ADME Study

Insilico ADMET studies has been done by using the ADMET protocol implemented in D.S 2.5 (Accelrys Discovery studio software). Insilico ADME studies solely depends on the chemical structure of molecules. *In silico* ADMET properties such as ADMET BBB level (Egan et al., 2000), absorption, aqueous solubility (Cheng and Merz., 2003) hepatotoxicity (Susnow and Dixon., 2005), CYP2D6 (Dixon and Merz., 2001), AlogP98 (Erlt et al., 2000) and PSA (Waterbeemd et al., 2001) are studied for the standard compounds from standard data set and further evaluation has been done on test set compounds. A standard ADMET model is generated which predict the human intestinal absorption (HIA) after oral administration of the inhibitors tested. The intestinal absorption model

includes 95% and 99% confidence ellipses in the ADMET_PSA_2D and ADMET_AlogP98 plane. There are four prediction levels for the absorption of compounds as good (0), moderate (1), poor (2) and very poor (3). These levels are defined by the 95% (red line) and 99% (green line) confidence ellipsoids .The upper limit of PSA_2D value for the 95% confidence ellipsoid is at 131.62, while the upper limit of PSA_2D value for the 99% confidence ellipsoid is at 148.12.

5.2.7 Toxicity Prediction

Toxicity profiling of all selected ligands were performed by employing Toxicity prediction – extensible protocol implemented in Accelrys discovery studio 2.5.

Toxicity prediction profile includes screening for aerobic biodegradability, developmental toxicity potentials, AMES mutagenicity, carcinogenicity & skin irritancy (Xia et al., 2003).

5.3 RESULTS AND DISCUSSIONS

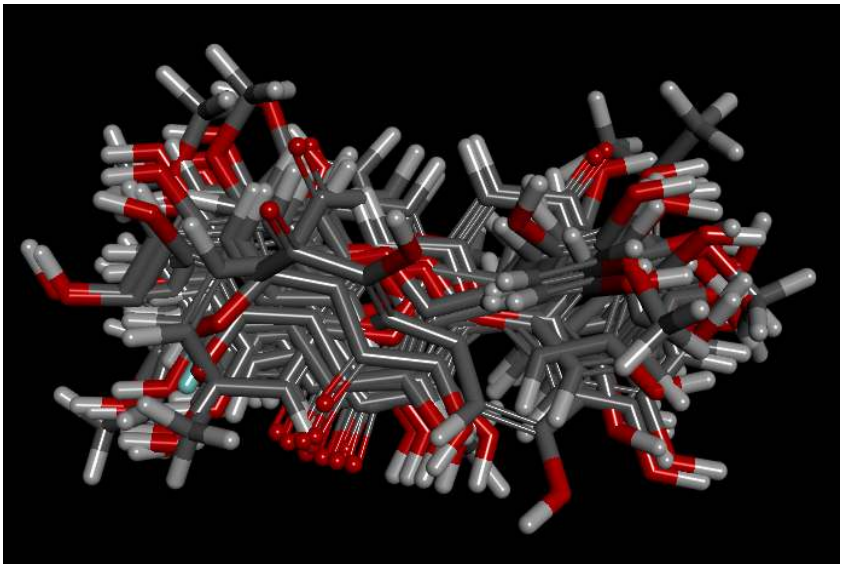


Figure 5.2 Myricetin and screened 95% similar structural analogues

5.3.1 Drug likeness study

Lipinski filter is used to study the drug likeness of all screened molecules. Molecular properties of all the compounds calculated by lipinski were tabulated in Table 5.1. Table 5.1 showed that all the screened analogues were satisfied the lipinski rule of five for being used as a probable drug in future.

Table: 5.1 Molecular properties of Myricetin and analogues molecules

Compounds	ALogP	Molecular Weight	Num_Rings	Num Aromatic Rings	Num_H Acceptor	Num_H Donar	Molecular Fractional Polar Surface area
CID_5281701	1.926	302.236	3	2	7	5	0.47
CID_10517292	1.872	286.236	3	2	6	4	0.419
CID_10636768	2.731	284.263	3	2	5	3	0.335
CID_13964548	2.619	314.289	3	2	6	2	0.289
CID_13964550	2.394	300.263	3	2	6	3	0.349
CID_24721178	1.872	286.236	3	2	6	4	0.419
CID_5281697	2.168	286.236	3	2	6	4	0.419
CID_5281953	1.839	346.288	3	2	8	4	0.391
CID_5315126	3.487	370.353	3	2	7	5	0.361
CID_5318214	2.394	300.263	3	2	6	3	0.349
CID_5320287	2.081	330.289	3	2	7	3	0.342
CID_5322065	2.41	270.237	3	2	5	3	0.358
CID_5393164	2.168	286.236	3	2	6	4	0.419
CID_57402278	2.728	344.315	3	2	7	4	0.364
CID_6477684	2.878	296.274	3	2	5	3	0.321
CID_6477685	2.636	312.274	3	2	6	4	0.377
CID_66574000	2.283	322.217	3	2	6	4	0.393
CID_9839293	2.098	300.263	3	2	6	3	0.349
Myricetin	1.388	318.23	3	2	8	6	0.532

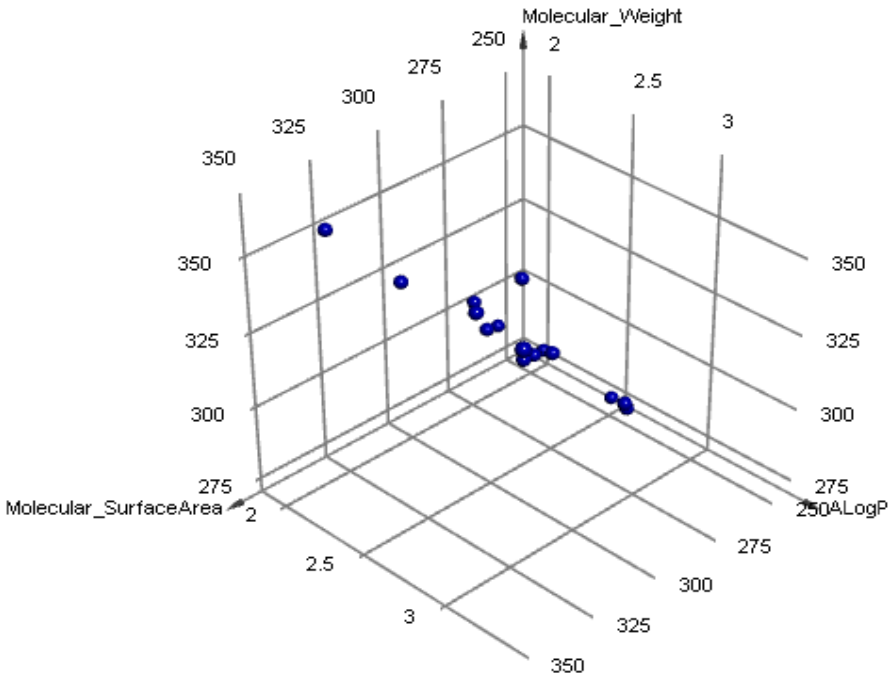


Figure 5.3: Represent the 3D plot of Myricetin and analogues Representing the correlation of Mol. Weight/ Alog P/Molecular_surface Area

5.3.2 ADME Study

Most of drug failures have been reported in early and late pipeline stage due to undesired pharmacokinetics and toxicity problems. If these issues can be addressed early, it would be extremely advantageous for the drug discovery process. The use of *in silico* methods to predict ADMET properties is projected as a first step in this direction to analyze the novel chemical entities to prevent wasting time on lead candidates that would be toxic or metabolized by the body into an inactive form and unable to

cross membranes, and the results of such analysis are herein reported in Table 5.2 together with a biplot (Figure 5.4) and discussed. The pharmacokinetic properties of all the molecules under study was predicted by six predefined ADMET models presented in Discovery Studio 2.5 program. After examine the biplot of the ADMET study, it was observed the biplot represents two analogous at 95% and 99% confidence ellipses corresponding to HIA and BBB models. PSA (polar surface area) is an useful parameter for prediction of drug transportation in different part of the body. The predefined models usually neglect the effect of other descriptors. The drug transportation and permeability has been demonstrated by PSA(plasma surface area).The cell membrane phospholipidbilayer is able to form hydrophobic and hydrophilic interactions as suggested by the fluid mosaic model, so lipophilicity is also play an essential property for drug designing and development. Lipophilicity of any compound could be expressed as the logarithms of the partition coefficient between *n*-octanol and water (log P).Thus the all the information about H-bonding could be govern by both PSA as well as log P calculation .Therefore in all model a plot between descriptors AlogP98 and PSA 2D at 95%and 99% confidence ellipses was considered for the precise prediction for the cell permeability of compounds. The region of 95% confidence ellipse depicts the chemical area well-absorbed compounds (≥90%) 95 out of 100 times. Whereas 99% is a confidence ellipse depicts chemical area of those compounds which having excellent absorption through cell membrane. Compound having an optimum cell permeability should follow the criteria (PSA < 140 ˚A2 and AlogP98 < 5) as describe in the model. The results shows that all the compounds except myricetin (151.23 A2) showed polar

surface area (PSA) < 140 ˚A2.It was shown in Figure. 5.4 that all the compounds have AlogP98 value <5 at 99% and 95% confidence ellipse for both HIA and BBB. Table 5.2 shows that majority of the compounds have low or undefined values for BBB penetration levels (levels 3 and 4) without any violation. The aqueous solubility also plays a critical role in the bioavailability of the candidate drugs, myricetin and compound CID_5315126 having low aqueous solubility level (level 2 and level 3 respectively) as shown in Table 5.2 while others having good aqueous solubility level shows that analogues are more soluble. Solubility play an important role in bioavailability of any drug. For any drug to be absorbed properly it should be more or less completely soluble in water. Further, all compounds have been predicted to have hepatotoxicity level of 1. It means the Myricetin and their screened analogues have some liver toxicity. Further studies are necessary to determine the dose level. Similarly, among all screened compounds only few compounds are showed satisfactory results respect to CYP2D6 liver (with reference to Table 5.2), suggesting that a these compounds are non inhibitors of CYP2D6 . This indicates that these analogues (CID_5281701, CID_13964550, CID_24721178, CID_5315126, CID_6477685 and CID_66574000) are well metabolized in Phase-I metabolism.CYP2D6 is class of Cytochrome 450 class of enzyme, is play an essential role in drug metabolism.

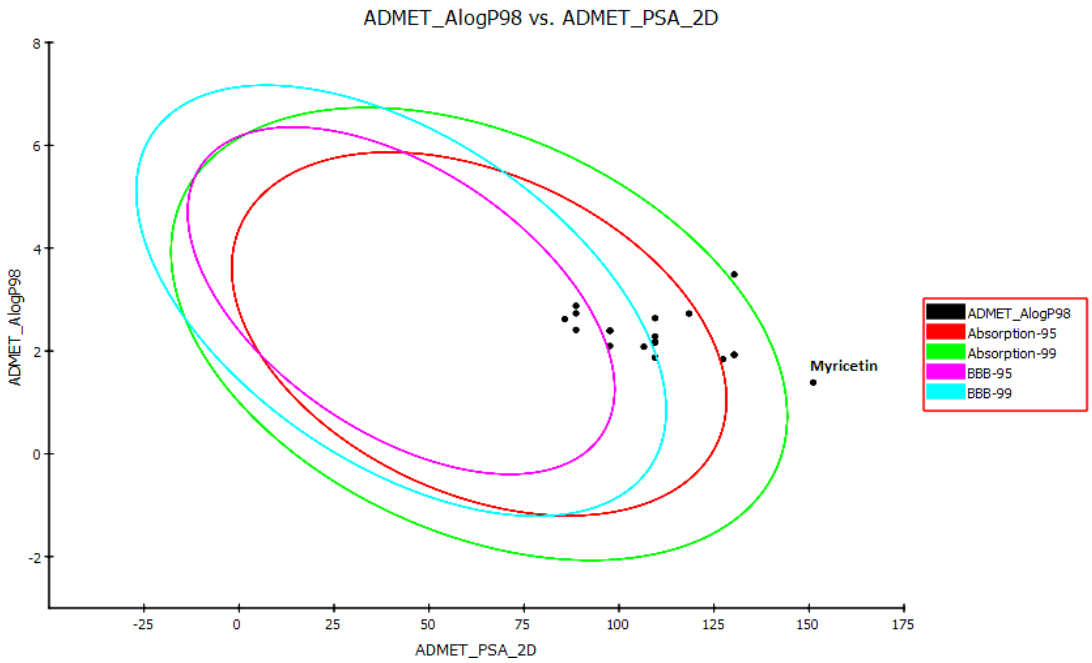


Figure 5.4: ADMET Description plot of Myricetin and their analogues

Table. 5.2 ADMET results of Myricetin and their Analogues

Comp.	ADMETBBB	Absorption Level	ADMET Solubility	ADMET Solubility Level	Hepatotoxicity	Hepatotoxicity_Probability	CYP2D6	CYP2D6_Probability	PPB Level	ADMET_AlogP98	Unknown AlogP98	PSA_2D	ADMET BBB
CID_5281701	4	1	-2.9	3	1	0.96	0	0.207	1	1.93	0	130.31	-1.308
CID_10517292	3	0	-2.6	3	1	0.99	1	0.75	2	1.87	0	109.49	-0.713
CID_10636768	3	0	-3.18	3	1	0.98	1	0.52	2	2.73	0	88.68	-0.701
CID_13964548	3	0	-3.6	3	1	0.91	1	0.74	2	2.62	0	85.72	-0.959
CID_13964550	3	0	-3.22	3	1	0.96	0	0.128	2	2.39	0	97.61	-1.308
CID_24721178	3	0	-2.58	3	1	0.97	0	0.237	2	1.87	0	109.49	
CID_5281697	4	0	-2.86	3	1	0.97	1	0.6	2	2.17	0	109.49	
CID_5281953	4	1	-3.19	3	1	0.96	1	0.52	2	1.84	0	127.35	
CID_5315126	4	2	-4.12	2	1	0.93	0	0.49	0	3.49	0	130.31	
CID_5318214	3	0	-3.21	3	1	0.97	1	0.59	2	2.39	0	97.61	-0.959
CID_5320287	3	0	-3.14	3	1	0.97	1	0.54	1	2.08	0	106.54	-1.196
CID_5322065	3	0	-2.98	3	1	0.94	1	0.72	2	2.41	0	88.68	-0.812
CID_5393164	4	0	-2.86	3	1	0.99	1	0.53	2	2.17	0	109.49	
CID_57402278	4	0	-3.48	3	1	0.94	1	0.56	1	2.73	0	118.42	
CID_6477684	3	0	-3.31	3	1	0.97	0	0.43	2	2.88	0	88.68	-0.668
CID_6477685	4	0	-3.17	3	1	0.99	0	0.35	0	2.64	0	109.49	
CID_66574000	4	0	-3.79	3	1	0.98	0	0.21	1	2.28	0	109.49	
CID_9839293	3	0	-2.99	3	1	0.89	1	0.71	2	2.1	0	97.61	-1.05
Myricetin	4	3	-2.843	3	1	0.966	1	0.504	1	1.388	0	151.123	

5.3.3 Toxicity Prediction

Insilico toxicity profile of all selected ligands were shown in Table 5.3. None of the compounds were show the Ames mutagenecity, skin sensitivity, rodent carcinogenicity. But only few compounds were pass the DTP(developmental toxicity potential parameters). The following results depicts that those compounds which were passed the all parameters of toxicity prediction parameters can be developed as future drug for cancer treat.

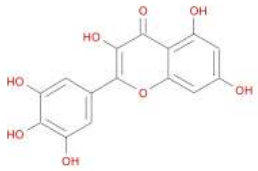
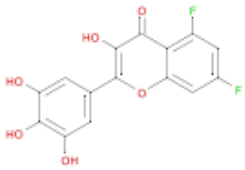
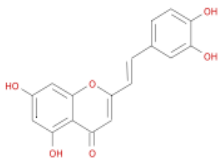
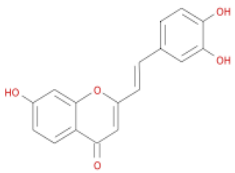
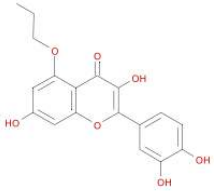
Table 5.3: Toxicity profile of all screened ligands using Toxicity Prediction – Extensible protocol of Accelrys Discovery Studio 2.5

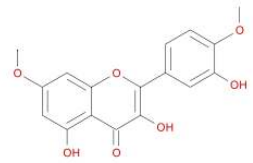
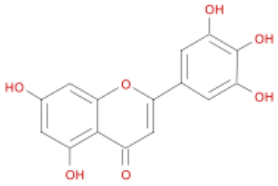
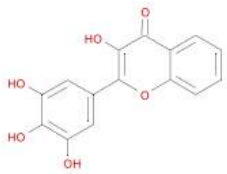
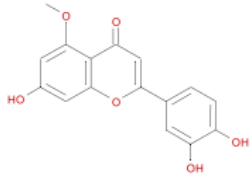
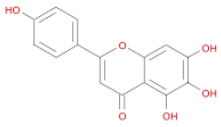
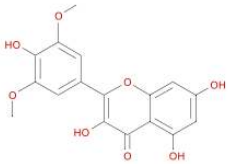
Compound	Rat Oral LD_{50} mg/kg body weight	Develop-mental toxicity potential	Ames mutag-enicity	Daphnia EC50 (mg/L)	Skin sensitization	Rat chronic LOAEL(g/kg body weight)	Fathead minnow LC50 (g/L)	Aerobic biodegradability	Rodent Carcinogenicity			
									Male Mouse	Female Mouse	Male Rat	Female Rat
CID_5281701	982.8	Non Toxic	No	2.6324	None	0.0407373	0.0220374	Non-Degradable	No	No	No	No
CID_10517292	224.6	Toxic	No	2.47664	None	0.13146	0.0566802	Non-Degradable	No	No	No	No
CID_10636768	857	Toxic	No	2.40206	None	0.0621067	0.00373682	Non-Degradable	No	No	No	No
CID_13964548	374.2	Non Toxic	No	1.24791	None	0.0716127	0.00171853	Degradable	No	No	No	No
CID_13964550	632.8	Non Toxic	No	2.8657	None	0.126524	0.00399408	Non-Degradable	No	No	No	No
CID_24721178	164.7	Toxic	No	6.51433	None	0.230328	0.0566802	Non-Degradable	No	No	No	No
CID_5281697	608.8	Toxic	No	2.43336	None	0.0487154	0.0117032	Non-Degradable	No	No	No	No
CID_5281953	982.8	Toxic	No	2.6324	None	0.0407373	0.0220374	Non-Degradable	No	No	No	No
CID_5315126	146.9	Non Toxic	No	1.21026	None	0.0861424	0.0291649	Non-Degradable	No	No	No	No
CID_5318214	866.9	Toxic	No	1.12638	None	0.0665433	0.00220504	Non-Degradable	No	No	No	No
CID_5320287	508.1	Non Toxic	No	2.08257	None	0.102257	0.00399408	Non-Degradable	No	No	No	No
CID_5322065	84	Toxic	No	1.03991	None	0.122494	0.0155985	Non-Degradable	No	No	No	No
CID_5393164	557.1	Toxic	No	2.39837	None	0.058545	0.0061957	Non-Degradable	No	No	No	No
CID_57402278	894.2	Non Toxic	No	2.43336	None	0.0403968	0.0117032	Non-Degradable	No	No	No	No
CID_6477684	430.5	Toxic	No	4.81778	None	0.271069	0.0309312	Non-Degradable	No	No	No	No
CID_6477685	.002	Non Toxic	No	0.791483	None	0.0366196	0.00270037	Non-Degradable	No	No	No	No
CID_66574000	.0016	Non Toxic	No	0.799086	None	0.0303215	0.00507574	Non-Degradable	No	No	No	No
CID_9839293	472.3	Toxic	No	3.08792	None	0.0751443	0.035466	Non-Degradable	No	No	No	No
Myricetin	212.3	Toxic	No	0.859026	None	0.177082	0.0193439	Non-Degradable	No	No	No	No

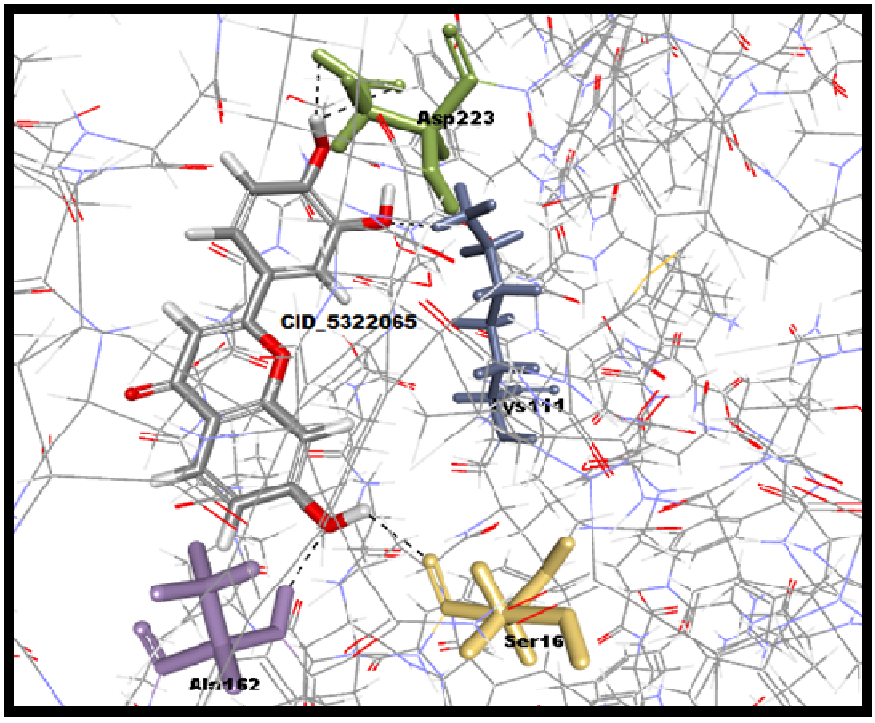
5.3.4 Molecular Docking Results

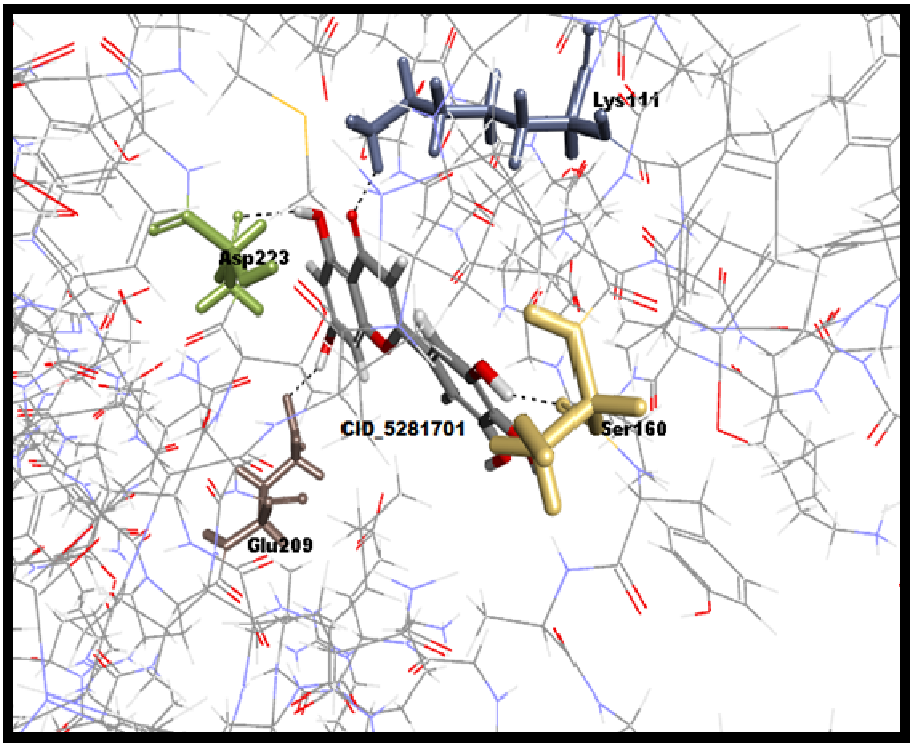
Myricetin (3,5,7-Trihydroxy-2-(3,4,5-trihydroxyphenyl)- 4-chromenone), is natural occurring flavanol (Ong and Khoo.,1997; Ross and Kasum., 2002)present in the plant kingdom as a secondary metabolite. It is the most well defined group of polyphenolic compounds.Myricetin is commonly found as O-glycosides with one of its hydroxyl group is substituted by sugars of various type. Molecular docking study of all compounds within the active sites of the PDK-1 kinase was carried out using CDOCKER docking method implemented in Discovery studio 2.5. Molecular docking results of Myricetin and the their top hits analogues were tabulated in Table 3. Docking energy of top ten docking hits were varies between -42.8 to -35.6 kcal/mol (detailed in Table 3). Negative docking energy indicates a more favorable binding of ligand at the binding pocket of the PDK-1 kinase. The non-covalent interaction of small-molecule to the proteins is governed by a range of inter-atomic contacts. These are mainly electrostatic interaction as well as van der Waals interactions. Hydrogen bonding is one of the most essential type of interaction shown by the protein and ligand molecule. The residues most likely involved in formation of hydrogen bonds were Lys111, Ala162, Ser 160, Glu 130, Thr222 and Asp223 in most of cases. Among these amino acid residues the Ala162 and Ser160 are the Hinge region's amino acid. Figure.5.5 depict the molecular docking interactions of top hits with PDK-1 kinase. Molecular docking results revealed that all the analogues formed a stable complex within the active site of the PDK-1 kinase. It depicts that the screened analogues of Myricetin can also act as a probable drug as PDK-1 kinase inhibitor.

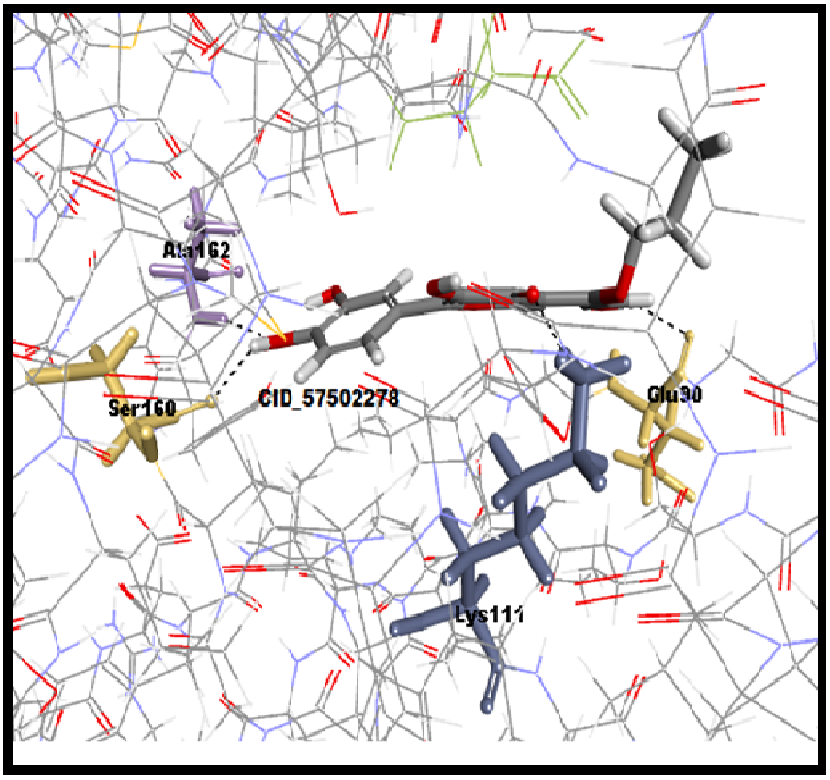
Table:5.4 Top ten docking hits of Myricetin and analogues

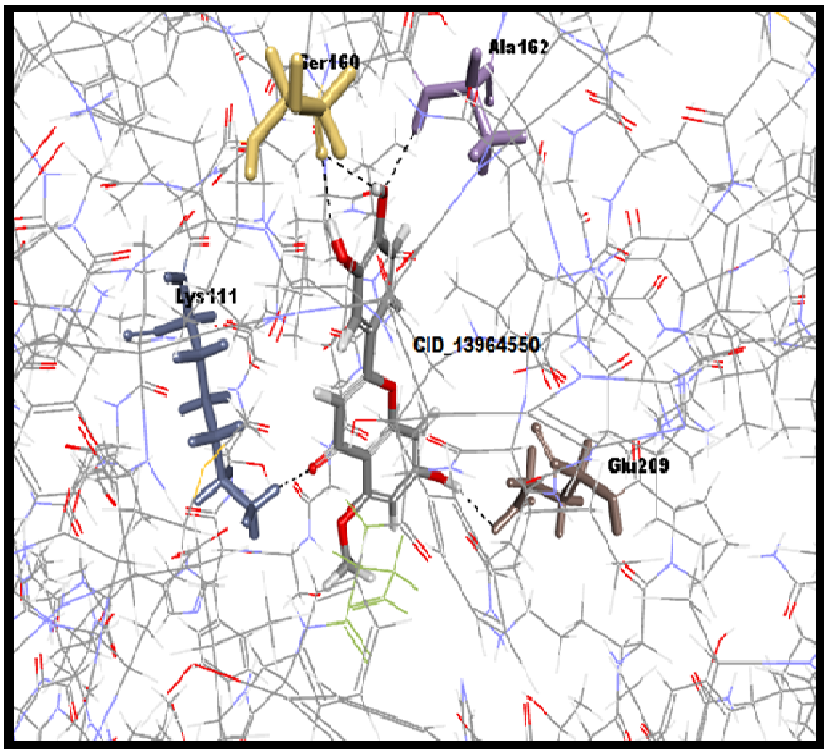
Compounds	Docking Energy Kcal/mole	Hydrogen bonding Residues	Hydrogen Bond distances (Å)
Myricetin 	-41	A:LYS111:HZ1 - 5281672:O5 A:LYS111:HZ3 - 5281672:O5 A:ALA162:HN -5281672:O8 5281672:H28-AGLU:OE2 5281672:H30-AASP223:OD1 5281672:H32 A:SER162:O 5281672:H33 A:SER162:O	2.44 2.19 2.34 2.31 2.09 2.14 1.98
CID_66574000 	-42.8	A:LYS111:HZ3 - 66574000:O8 A:ALA162:HN - 66574000:F1 A:THR222:HG1 - 66574000:O7 66574000:H30 - A:GLU130:OE2 66574000:H31 - A:ASP223:OD1	1.89 2.31 2.24 2.33 2.45
CID_6677685 	-35.6	A:LYS111:HZ2 - 6477685:O6 A:LYS111:HZ3 - 6477685:O6 A:ALA162:HN - 6477685:O4 6477685:H32 - A:ALA162:O 6477685:H34 - A:GLU130:OE2 6477685:H35 - A:ASP223:OD1	2.47 2.10 2.39 1.97 2.22 2.03
CID_6677684 	-37.90	A:LYS111:HZ2 - 6477684:O5 A:LYS111:HZ3 - 6477684:O5 A:ALA162:HN - 6477684:O2 6477684:H33 - A:GLU130:OE2 6477684:H34 - A:ASP223:OD1	2.37 2.12 2.14 2.09 2.03
CID_57402278 	-38.17	A:LYS111:HZ3 - 57402278:O4 A:ALA162:HN - 57402278:O7 57402278:H39 - A:GLU90:O 57402278:H41 - A:SER160:O	1.77 2.33 2.48 1.92

CID_5322065 	-38.98	A:LYS111:HZ3 - 5322065:O3 A:ALA162:HN - 5322065:O5 5322065:H28 - A:ASP223:OD1 5322065:H29 - A:SER160:O 5322065:H30 - A:SER160:O	2.16 1.89 2.04 1.90 2.05
CID_5281701 	-41.87	A:LYS111:HZ3 - 5281701:O3 5281701:H28 - A:ASP223:OD1 5281701:H29 - A:GLU209:OE2 5281701:H31 - A:SER160:O	1.91 2.02 2.09 1.87
CID_10517292 	-39.36	A:LYS111:HZ3 - 10517292:O5 10517292:H28 - A:GLU166:OE2 10517292:H29 - A:ASP223:OD2 10517292:H31 - A:ASP223:OD2 10517292:H29 - A:ASN210:OD1	1.80 2.19 2.33 2.36 1.92
CID_13964550 	-36.82	A:LYS111:HZ3 - 13964550:O3 A:ALA162:HN - 13964550:O6 13964550:H32 - A:GLU209:OE2 13964550:H33 - A:SER160:O 13964550:H34 - A:SER160:O	1.72 2.21 2.14 1.87 2.22
CID_5281697 	-40.31	A:LYS111:HZ3 - 5281697:O3 5281697:H29 - A:ASP223:OD1 5281697:H30 - A:GLU130:OE2 5281697:H31 - A:ALA162:O	2.10 2.35 2.16 1.93
CID_5281953 	-37.8	A:LYS111:HZ3 - 5281953:O40 A:LYS111:HZ3 - 5281953:O7 A:ALA162:HN - 5281953:O8 5281953:H30 - A:GLU166:OE2 5281953:H32 - A:SER94:OG 5281953:H33 - A:SER160:O	1.80 2.02 1.97 2.17 2.03 2.40


Asp223
CID_5322065
Lys111
Ser16


Lys111
Asp223
CID_5281701
Ser160
Glu209


Ala162
Ser160
CID_57502278
Glu90
Lys111


Ala162
Lys111
CID_13964550
Glu209

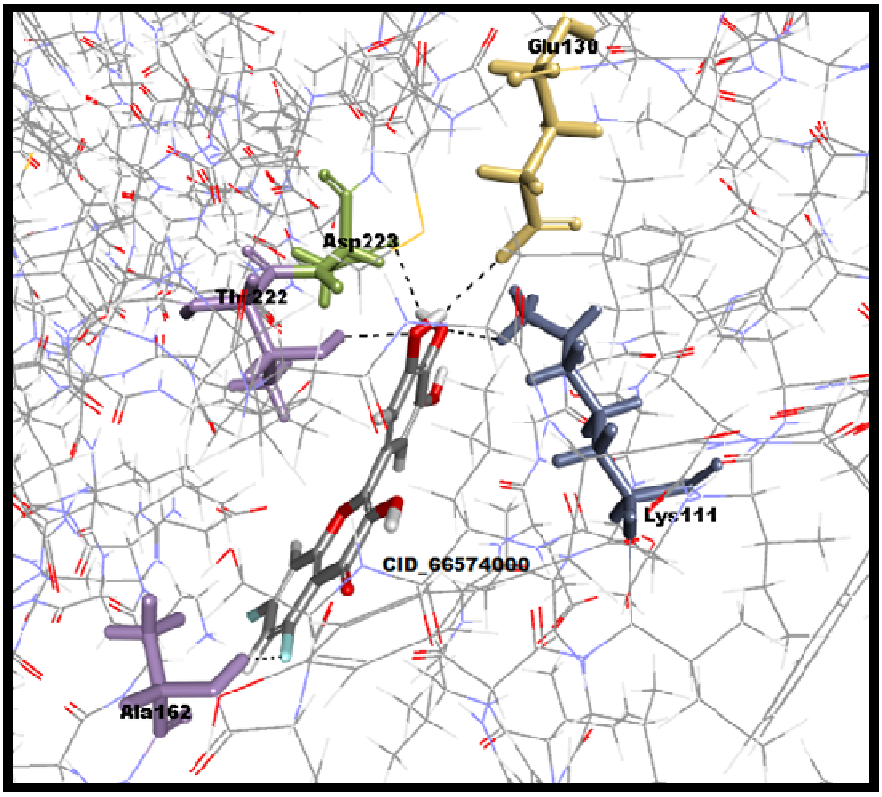


Figure 5.5. Molecular docking results of top hits