Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 7 July 2021: 12767 – 12778

Research Article

Computational Development of Pyrazole Derivatives by Docking, Virtual Screening, and Admet Predictions

Swapna. B[1], M. Vijey Aanandhi[*2]

Abstract

Heterocyclic chemistry is an excellent origin of medicinal chemistry, where a wide range of molecules with a potent biological activity was identified. Most of these molecules have complex structures with heteroatoms that pose a high affinity or inhibition towards the human protein, most likely for the biological activity towards many diseases. Pyrazole is an outstanding pharmacophore with a wide range of pharmacological representations. In the current study, a library of pyrazole molecules were developed and structurally confirmed by various spectral techniques, including chromatographic analysis. These molecules are also digitally screened for estimation of the biological potency by virtual screening. Molecular docking is a helpful tool to predict probable biological activity towards human biological systems. The ADMET prediction tool is used to estimate protein adsorption/activity. Distribution, metabolism, excretion, and toxicity were predicted between ligand and human cell lines called cyclin-dependent kinase (CDK). A human variant 1G1I is obtained from protein bank **ref** sources is developed here. These molecules are also estimated for their anticancer and Antioxidant activity.

Keywords: *Pyrazole, Cyclin-dependent kinase (CDK), Virtual screening, and Molecular docking.*

[1]Research Scholar, Department of Pharmaceutical Chemistry and Analysis,
School of Pharmaceutical Sciences, Vels Institute of Science, Technology & Advanced Studies (VISTAS) Chennai - 600 117.Tamil Nadu, India
Email: swapna.klsss@gmail.com

[2]Department of Pharmaceutical Chemistry and Analysis,
School of Pharmaceutical Sciences, Vels Institute of Science, Technology & Advanced Studies (VISTAS) Chennai - 600 117.Tamil Nadu, India
Email: hodpchemistry@velsuniv.ac.in

Introduction

Intermolecular interactions play a significant role in studying the protein-ligand interactions in virtual databases (Cadavid et al., 2000). The interactions are interrelated to the binding site and un-identified forces beyond the binding region (Ganatra and Gurjar, 2012). Such interactions between bimolecular collisions and adhesive property with cell membranes control the molecule's collisions and the protein (Maria and Leslie, 2005; Areias et al., 2008). The protein and flexible targeted molecules interact with proteins, small molecules, and nucleic acids with the binding site (Carro et al., 2009). Pyrazole is the smallest and most delicate molecule that interacts with the body proteins in significant areas. Thus the pyrazoles inhibit protein control and their metabolic functions (Kouraklis and Theocharis, 2006).

Experimental Procedure

Protein preparation:

Cyclin-dependent kinase (CDK) is the most important protein that inhibits several cell division functions, and CDK inhibitors control the negative progression of cell division. CDK controls the cell cycle progression in cell division from one stage to another stage. Due to mutational changes in the cell cycle leads to perturbed cell division (Reyes A, 2004), which led to the proliferation and cellular death. The CDK-II mutations in humans lead responsible for 90% of cancers.

Identification of protein for the research work

1. **Selection of species:** The present is designed for the biological activity over humans, so protein with species of Homo sapiens is preferred.
2. **Strength of protein:** The protein strength can be determined by the method, which the structure of the protein is confirmed. X-ray diffraction method has opted for the protein structure confirmation.
3. **Protein resolution:** Resolution of a protein determines the ability of the protein as low as the resolution of the protein more the strength of the protein.
4. **Domain strength:** The completeness of the protein shall determine the strength of the protein.

Table 1

The results of protein sequences

S.No	Sequence ID	Species	Resolution factor in A^o	Experiment	Amino acid length
1	1BCK	HUMAN (H.Sapiens)	2.80	X ray	288
2	1GCL	HUMAN (H.Sapiens)	2.70	X ray	288

3	1JKT	HUMAN (H.Sapiens)	2.50	X ray	1012
4	1FHN	HUMAN (H.Sapiens)	2.20	X ray	1016
5	1JQU	HUMAN (H.Sapiens)	2.60	X ray	742
6	1AI1	HUMAN (H.Sapiens)	2.40	X ray	392
7	1GII	HUMAN (H.Sapiens)	2.10	X ray	298

Preparation of ligands

Ligand selection: The pharmacophore pyrazole is selected to prepare a library of molecules with primary active binding sites (R^1, R^2, R^3) as designed below in Fig. 1.

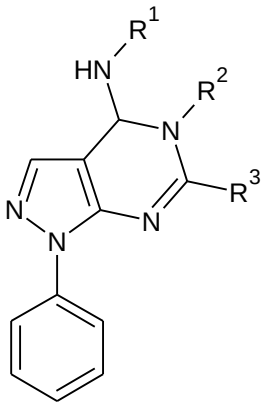

Fig. 1. Pharmacophore for docking

Active ligand pharmacophore: The functional groups at the sites (R^1, R^2, R^3) are substituted for functional groups are listed below in table 2.

Table 2

Combinations of the ligand library

R^1	CH_3, C_2H_5, $_nC_3H_7$, $_{iso}C_3H_7$, $_nC_4H_9$ $_{iso}C_4H_9$	6 COMBINATIONS
R^2	CH_3	1 COMBINATIONS
R^3	CH_3, C_2H_5, $_nC_3H_7$, $_{iso}C_3H_7$, $_nC_4H_9$ $_{iso}C_4H_9$	6 COMBINATIONS

Ligands stabilization: The molecules designed are structurally modified for further computational studies such as ADMET predictions, docking, and virtual screening studies. Chemsketch is utilized for the preparation of ligand structures. Ligand designed was drawn in the 2D format in chemsketch, and output files were stored (.mol).

Energy minimization of the Ligands: Drug discovery studio visualizer 4.0 is utilized for the 3D formation of ligands. Drug discovery studio visualizer 4.0 was developed by Accelrys. The powerful software was utilized to convert the 2D structure into a 3D structure format and energy was optimized, and the force field was applied to the protein data bank format (.pdb).

Protein stabilization: The protein sequences were obtained from the open-source databank, and protein 1GII sequences are selected for this work.10 The sequence was obtained from the same source and extracted in the form of (.pdb). Then 1GII.pdb is explored in the Auto dock 4.0 software, as shown in Fig. 2.

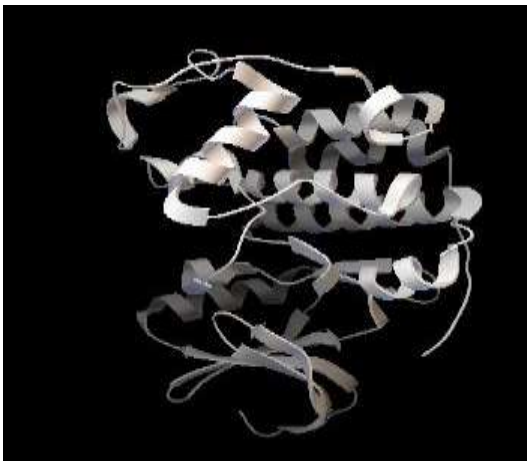

Fig. 2. 1GII Protein structure and the selected protein 1GII are explored in AutoDock 4.0 to perform the Insilico studies.

Bonds optimization: Protein bands were optimized and built by the bond distance.

Atoms optimization: Atom parameters were optimized by AutoDock 4 to enable the protein (AD4 parameter) suitability.

Hydrogen optimization: Missing hydrogen in the protein is added and adjusted. The non-polar center located in between the hydrogen atoms was selected and merged. +1 charge was attributed to all the histidine hydrogen and protonated. Finally, the protein is fixed with .pdb parameter, and the protein is adjusted accordingly.

Charges optimization: Protein is processed further so that all the required charges for protein were attributed. A total of -191.711 Kollman charges is attributed to the protein. Protein has not attributed any gasteiger charges. Finally, Kollman charges are fixed to the protein. Thus, the

protein sequence is optimized, and all parameters required for the study are attributed and recorded in .pdbqt format.

Virtual screening: A computational tool is used in synthetic chemistry, and drug discovery, where the library of ligand binds with a protein and the ligand-protein binding energies were calculated. The method is majorly classified into two techniques.

(1) Ligand-based: A group of ligands designed structurally shall be bonded to a protein receptor and explore the consolidated results. These ligands are developed based on the pharmacophore model opted for the research work. The pharmacophore bindings and compatability to the candidate ligands were estimated for their binding and inhibition capabilities (Scott A, 2011).

(2) Structure-based: Ligand in this method is docked a protein sequence target. The scoring function of the protein-ligand-based is estimated—protein likelihood towards the ligand bounds to protein at high affinity. "PyRx python prescription 0.8 version" is the software used to analyze virtual screening and its results. Virtual screening was performed in three major stages.

a. Protein optimization
b. Ligand library preparation and optimization.
c. Virtual Screening.

Protein optimization: An optimized protein with specified parameters is already prepared at 3.3.4, and the same had been utilized for the virtual screening studies.

Ligand library preparation and optimization: Chemsketch is utilized for the preparation of ligand structures. Ligand designed was drawn in the 2D format in chemsketch, and output files were stored as (.mol). The designed molecules are structurally modified and prepared for further computational studies such as ADMET predictions, docking, and virtual screening studies.

Energy minimization of the ligands: Drug discovery studio visualizer 4.0 is utilized for the 3D formation of ligands. The optimized 2D structures converted to the 3D structure. The energy was optimized, and a force field was applied to develop a protein data bank (.pdb).

Virtual screening: The macromolecule protein is selected and explored in the virtual screening software. Then, the ligand data bank was added to the database, and ligands stabilized accordingly for the virtual screening. Then prepared ligands were quantitatively optimized, and the prepared file was stored as protein data bank quantitative format (.pdbqt). Autodock, Autogrid, Autovina data were added to the database with suitable parameters (mention here data deposition number). The protein and ligand complex were aligned in the grid box graphically so that the total protein and ligand were fixed for the screening. Thus, the infinite research area was fixed to a specific region of ligand and protein. The virtual screening was performed after fixing

the macromolecule, ligand library in a grid box and stored, and virtual screening results are listed in table 3.

Table 3

Results of virtual screening

S.NO	Ligand	Target	Binding Energy	Info
1	10_uff_E=457.23	1GII	-8.8	Vina
2	13_uff_E=384.06	1GII	-8.7	Vina
3	16_uff_E=448.64	1GII	-8.7	Vina
4	19_uff_E=392.45	1GII	-8.7	Vina
5	22_uff_E=310.95	1GII	-8.7	Vina
6	25_uff_E=348.63	1GII	-8.7	Vina
7	28_uff_E=429.07	1GII	-8.6	Vina
8	6_uff_E=444.80	1GII	-8.6	Vina
9	3_uff_E=375.92	1GII	-8.6	Vina
10	31_uff_E=309.53	1GII	-8.6	Vina
11	9_uff_E=444.46	1GII	-8.6	Vina
12	33_uff_E=419.69	1GII	-8.5	Vina
13	1_uff_E=366.55	1GII	-8.5	Vina
14	2_uff_E=366.07	1GII	-8.5	Vina
15	14_uff_E=307.64	1GII	-8.5	Vina
16	17_uff_E=400.40	1GII	-8.5	Vina
17	20_uff_E=435.66	1GII	-8.5	Vina
18	23_uff_E=392.53	1GII	-8.5	Vina
19	26_uff_E=433.51	1GII	-8.5	Vina
20	7_uff_E=346.60	1GII	-8.5	Vina
21	4_uff_E=325.89	1GII	-8.5	Vina
22	32_uff_E=355.13	1GII	-8.5	Vina
23	10_uff_E=375.89	1GII	-8.5	Vina
24	34_uff_E=354.28	1GII	-8.5	Vina
25	15_uff_E=470.95	1GII	-8.5	Vina
26	18_uff_E=338.53	1GII	-8.4	Vina
27	21_uff_E=344.48	1GII	-8.4	Vina
28	24_uff_E=414.51	1GII	-8.4	Vina
29	27_uff_E=429.81	1GII	-8.4	Vina
30	8_uff_E=423.65	1GII	-8.4	Vina
31	29_uff_E=395.51	1GII	-8.4	Vina
32	5_uff_E=388.81	1GII	-8.4	Vina

33	35_uff_E=409.05	1GII	-8.4	Vina
34	36_uff_E=436.69	1GII	-8.4	Vina
35	11_uff_E=441.25	1GII	-8.4	Vina
36	12_uff_E=360.94	1GII	-8.4	Vina

Admet predictions: These are the predictions developed to virtually analyze the drug absorption, distribution, metabolism, excretion, and toxicity studies in the human body.

Lipinski's rule: The rule, developed by Pfizer`s, is an optimized calculation of a group of parameters that exhibit suitable ADMET parameters. The rule is also known as the Pfizer rule and five (RO5), a thumb rule to estimate the drug-likeness of a drug molecule. Rule of five is a basic parameter for the calculation proposed by Lipinski, also known as Lipinski`s rule.

Rule of five

1. Log P of the molecule should be lesser than or equal to 5.
2. The molecular weight of the molecule should be less than or equal to 500 Daltons.
3. Hydrogen acceptors in the molecule should be less than or equal to 10.
4. Hydrogen donor in the molecule should be less than or equal to 5.

Drug molecules that violate the above parameters are found to be of poor bioavailability. It is called as "Rule of 5", as the border values are 5, 500, 2x5, and 5. The software shall be suitable to calculate the optimized parameters and estimate the bioavailability for the research are listed. Data Warrior, Moleinspitation, etc. Data warrior is the suitable database opted for the research work, which reveals all parameters mentioned in the Lipinski rule.

Admet predictions

Ligands optimization for admet: All the molecules (protein data bank .pdf format) have opted for a single stretch of screening to predict ADMET properties. Then the molecules are optimized and prepared for the structure data file (.sdf). These (.sdf) molecules are merged into a single database file by using a database of Discovery studio 4.0.

Osiris data explorer: Osiris data explorer is being opted for the preparing structure data file and explored all the molecules. All the molecules are screened for suitable parameters for the calculation of ADMET properties. Total mol weight in g/mol, cLogP, H-Acceptors, H-Donors, TSA Total surface area, PSA Polar surface area, Molecular weight of largest fragment in g/mol, and Druglikeness. Molecule (Ligand) Toxicity & Efficacy Ligand efficacy (LE), Ligand efficacy Lipophilic Price (LELP), Lipophilic Ligand efficacy (LLE), Mutagenicity, Tumorigenicity, and Irritant. The ligand with optimal Lipinski rule parameters are established by examining the molecular weight above 500 daltons, log p greater than 5, hydrogen donors greater than 5, and hydrogen acceptors greater than 10. Finally, molecules that are highlighted and found to exhibit violation of ADMET properties are virtually removed from the database. Further Insilco studies

ligand with excellent ADMET properties are listed in table 4 and opted for further studies like docking, synthesis, and biological screening.

Table 4

Results of ADMET predictions.

S.NO	Total Molweight	Absolute Weight	cLogP	H-Acceptors	H-Donors	Total Surface Area	Drugliken ess	LE from Molecule Name	LLE from Molecule Name	LELP from Molecule Name
1	355.355	355.0735	3.4351	4	1	334	0.3857	0.54984	5.5549	3.7473
3	370.383	370.0893	3.7108	4	0	339	0.35055	0.5957	5.9883	4.543
3	354.384	354.0943	3.1347	3	0	335	0.31835	0.51539	5.3983	5.0775
4	354.384	354.0943	3.1347	3	0	330	0.31835	0.50537	5.3733	5.1533
5	358.311	358.1099	3.5403	3	0	339	0.034994	0.5594	4.7507	5.3175
6	355.373	355.0895	3.1035	4	1	333	0.43547	0.59355	5.1183	3.5433
7	385.354	385.0537	1.8593	5	0	344	-4.8385	0.53374	5.3957	3.4899
8	319.153	317.9893	3.505	3	0	334	-1.5395	0.58453	4.5909	5.9959
9	374.703	374.0397	3.3858	3	0	331	0.34337	0.58094	4.559	5.8399
10	370.383	370.0893	3.7108	4	0	334	0.35055	0.54875	5.3893	4.9399
11	385.383	385.0841	3.3551	5	1	341	0.3857	0.51993	5.5935	4.549
13	300.309	300.0998	3.5408	5	0	375	0.35055	0.49393	5.38	5.3455
13	384.31	384.1049	3.0547	4	0	355	0.31835	0.51518	4.8314	5.9394
17	349.179	347.9997	3.435	4	0	353	-1.5395	0.50757	4.3335	5.7595
18	304.738	304.0503	3.3158	4	0	350	0.34337	0.50594	4.4379	5.5557
19	355.355	355.0735	3.4351	4	1	331	0.3857	0.55751	5.3851	4.3578
31	385.383	385.0841	3.3551	5	1	344	0.3857	0.50157	5.3137	4.7154
33	370.383	370.0893	3.779	4	1	338	0.359	0.53535	4.8785	5.3907
38	354.384	354.0943	3.1347	3	0	335	0.31835	0.54535	4.4381	5.7397
30	384.31	384.1049	3.0547	4	0	345	0.31835	0.49145	4.4583	5.3157

Docking: The ligands with good ADMET predictions are selected for the docking to identify prominent and standard sources for ligand selection. Selection of suitable software for the docking of ligands.

AutoDock 4.0: Auto dock is an excellent suite for performing computational docking studies for identifying the efficacy of the molecules in a 3D structural format. The present version of the autodock comprises two updated software sources, such as Auto dock Vina and Auto dock 4.0, with excellent result attributes. The Autodock 4.0 performs in two major stages, such as auto grid where infinite space of the structural database was aligned for specific region, the autodock bind ligand protein in the aligned grid binding energies were calculated.

Docking procedure: Protein-ligand stabilization, grid parameter alignment, docking parameter file alignment, and docking results are analyzed.

Grid parameter alignment: The protein macromolecule 1GII is explored in the 3D space as a rigid file, ligand in the form of .pdbqt is exposed in the 3D space as nonrigid alignment. The grid box was aligned from the center of the macromolecule and adjusted so that the whole molecule is within the limits of the grid box, as shown below.

Docking parameter file alignment: The protein macromolecule 1GII is exposed in 3D space as a rigid file, a ligand in the form of .pdbqt is exposed in the 3D space as nonrigid alignment. Genetic algorithm, default docking parameters such as autodock4.2 parameter for RM calculation have also opted, and output was extracted as a Lamirikan Genetic algorithm and saved as docking parameter file (.dpf).

Docking: Docking of ligand to the macromolecule was performed in two stages.

1. Running of the Auto Grid
2. Running of the Auto Dock.

Running of Auto Grid: The standard grid path database was obtained online from the official website of autodock and linked grid path file. Thus, the obtained grid parameter file was connected to the output file, and the autogrid was performed; obtained autogrid was a database as (.glg) grid log file.

Running of Auto Dock: Standard dock path database was obtained online from the autodock and linked dock path file. Thus obtained docking parameter files were connected to output file, and the autodock was performed, obtained autodock result was databased as (.dlg) dock log file.

Analyzing results:

Table 5
Results of Docking studies.

S.NO	MOLECULE NO	MOLECULAR STRUCTURE	BINDING ENERGY	IC 50	IC 50 UNITS	NO.OF CONFIRMATIONS
1	**1**	4-[1-amino-2-(4-aminophenyl)ethyl]-1-phenyl-1*H*-	**-9.53**	**103.7**	**NANO**	**7**

		pyrazol-5-amine			**MOLAR**	
2	**2**	4-{1-amino-2-[4-(methylamino)phenyl]ethyl}-1-phenyl-1*H*-pyrazol-5-amine	**-9.34**	**141.87**	**NANO MOLAR**	**6**
3	**3**	*N*-[amino(5-amino-1-phenyl-1*H*-pyrazol-4-yl)methyl]-*N*'-ethylbenzene-1,4-diamine	**-9.21**	**176.58**	**NANO MOLAR**	**10**
4	**4**	*N*-[amino(5-amino-1-phenyl-1*H*-pyrazol-4-yl)methyl]-*N*'-propylbenzene-1,4-diamine	**-9.05**	**232.38**	**NANO MOLAR**	**7**
5	**5**	*N*-[amino(5-amino-1-phenyl-1*H*-pyrazol-4-yl)methyl]-*N*'-(propan-2-yl)benzene-1,4-diamine	**-8.29**	**840.64**	**NANO MOLAR**	**10**
6	**6**	*N*-[amino(5-amino-1-phenyl-1*H*-pyrazol-4-yl)methyl]benzene-1-amine-4-aminophenyl	**-8.26**	**888.6**	**NANO MOLAR**	**8**
7	**7**	1-phenyl-6-[(phenylamino)methyl]-1*H*-pyrazolo[3,4-*d*]pyrimidin-4-(phenylamino)pyrimidine	**-8.22**	**936.1**	**NANO MOLAR**	**7**
8	**8**	*N*-methyl-*N*'-[(1-phenyl-1*H*-pyrazolo[3,4-*d*]pyrimidin-6-yl)methyl]benzene-1,4-diamine	**-8.22**	**946.25**	**NANO MOLAR**	**7**
9	**9**	*N*-[amino(5-amino-1-phenyl-1*H*-pyrazol-4-yl)methyl]-*N*'-ethylbenzene-1,4-diamine	**-8.19**	**996.33**	**NANO MOLAR**	**8**
10	**10**	*N*-[amino(5-amino-1-phenyl-1*H*-pyrazol-4-yl)methyl]-*N*'-propylbenzene-1,4-diamine	**-8.13**	**1.09**	**MICRO MOLAR**	**9**
11	**11**	*N*-[amino(5-amino-1-phenyl-1*H*-pyrazol-4-yl)methyl]-*N*'-(propan-2-yl)benzene-1,4-diamine	**-8.12**	**1.11**	**MICRO MOLAR**	**9**

Conclusion

Selection of protein is performed, and 1GII is confirmed as the best fit protein for CDK inhibitor. Ligands library is prepared based on the active pharmacophore. Virtual screening is performed, and all the ligands are found to be with sufficient binding energy. ADMET predictions are performed, and ligands with toxicity, mutagenicity, tumorigenicity, irritant nature, and ligands beyond Lipinski ”Rule of five” are removed from the ligand library. Docking studies further screen the remaining ligands, and binding energy and inhibition energy are calculated. The best fit ligands are tabulated for the synthesis of the most active ligands.

Acknowledgement

The authors are thankful to Vels Institute of Science Technology and Advanced Studies (VISTAS), Chennai Tamil Nadu for providing all the facilities to make this work complete. The authors are pleased to express sincere thanks to K.P Labs, Hyderabad for valuable support.

Conflict of Interest

The authors declare that they have no conflict of interest for this study.

Funding Support

The authors declare that they have no funding support for this study.

References

1. Areias, F., Aranda, R., et al. 2008. Synthesis, Binding Affinity, And Molecular Docking Analysis Of New Benzofuranone Derivatives As Potential Antipsychotics. J.Med.Chem., 51(19):6085–6094.

2. Cadavid, M., Ravina, E., et al. 2000. Conformationally Constrained Butyrophenones with Affinity for Dopamine (D 1 , D 2 , D 4) and Serotonin (5-HT 2A , 5-HT 2B , 5-HT 2C) Receptors: Synthesis of Aminomethylbenzo[b]furanones and Their Evaluation as Antipsychotics. J.Med.Chem., 43(24):4678–4693.

3. Carro, L., Ravina, E., et al. 2009. The Synthesis And Binding Affinity Of Potential Atypical Antipsychotics With The Tetrahydroquinazolinone Motif. J Bio Org Med Lett, 19(21):6059–6062.

4. Ganatra, S., Gurjar, S. 2012. Structure-Based Studies of 2-Bezylidine Benzofuran-3-One Class of Compounds as the Cyclin-Dependent Kinases (CDKs) Inhibitor. 5(6):712–717.

5. Kouraklis, G., Theocharis, S. 2006. Histone deacetylase inhibitors: a novel target of anticancer therapy. 15(2):489–494.

6. Maria, I. Z., Leslie, A. K. 2005. Side-chain flexibility in protein ligand binding: The minimal rotation hypothesis. Protein science: a publication of the protein society, 14(4):1104–1114.

7. Scott, A., Wildmanl, X. Z., et al. 2011. Drug-like Leads for Steric Discrimination between Substrate andInhibitors of Human Acetylcholinesterase. Chemi cal Biology and Drug Design, 78(4):495–504.