

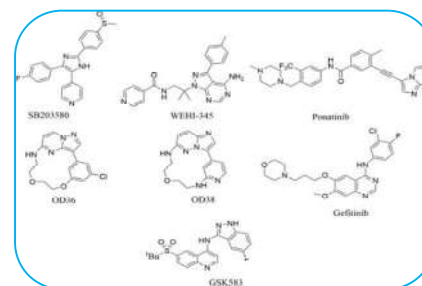


QSAR ANALYSIS OF PYRIDINES AND PYRIMIDINES DERIVATIVES OF LRRK2 INHIBITORS

Dr. Garima

ABSTRACT

Parkinson's disease is the most common neurodegenerative disorder. These include dyspepsia, hyposmia, dysphagia, mood swings, constipation, cognitive decline, orthostatic hypertension, and sleep disturbances. Gly2019Ser, often known as G2019S, is a common mutation of LRRK2 that is observed in people with autosomal leading Parkinson's disease (PD) and outward irregular PD, which is clinically similar to idiopathic PD. The majority of the features of the very large protein LRRK2, which belongs to the ROCO superfamily, include the GTPase domain, the RAS domain, the WD40 domain, the leucine-rich repeat (LRR) domain, the DFG-like motif, the kinase domain, the mixed lineage kinase (MLK) like domain, and so on. A QSAR model provides information that is highly helpful for drug design and medicinal chemistry. QSAR plays a significant role in the discovery of new drugs and finds many applications in predicting the activity of novel compounds by mathematical expression that determines the relationship between their chemical structure to their biological activity. According to current research, CADD based on QSAR has proved crucial in the development of novel medications for treating a variety of illnesses. The statistical parameter values for R², R²_{adj}, Q²_{loo}, R²_{ext}, and CCC_{ext} were, respectively, 0.8261, 0.8120, 0.7769, 0.7998, and 0.8941 in this model. To determine the ideal protein-ligand interaction, a molecular docking analysis of the most active molecule (D21026) and created compounds (H1-H5) was conducted. In short, this study helps in development of various drugs for treatment of LRRK2 inhibitors related diseases.



KEYWORDS: LRRK2, PD, G2019S, QSAR, descriptors, Molecular docking.

INTRODUCTION

In the year 1817, a scientist named James Parkinson described a disease known as Parkinson's disease (PD). It is the most common neurodegenerative disorder. It is a neurodegenerative brain disorder characterized by four fundamental motor indications: postural instability, bradykinesia, rigidity and resting tremor. A number of nonmotor indications are, however, progressively recognized as being portion of the illness manifestation. These include constipation, hyposmia, problems with speech, mood disorders and swallowing, cognitive impairment, orthostatic hypertension, and sleep disorders.

In the year 2004, it is found that alterations in LRRK2 can be the key cause of Parkinson's disease. Even not a single alteration of LRRK2, in fact diversity of inherent ties to this illness¹. The popular mutation of LRRK2 is Gly2019Ser or G2019S, are noticed in patients with autosomal leading PD and in those with outward irregular PD, who are clinically indistinguishable from those with idiopathic PD². Typical of multifactorial illnesses, the existence of PD rises with age, with a predictable 0.3% distressed at age 50 increasing to 4.3% by age of 85³. This mutation lies in a preserved part of the protein kinase domain that starts the activation loop, which, as the name suggests, normalizes catalytic

enzyme activity. PD results in neuronal disfunction and advanced loss of dopamine-producing neurons situated in the substantia nigra pars compacta area of the midbrain. Foremost among these are mitochondrial dysfunction, autophagy/lysosomal dysfunction, and inflammation, all of which are highly combined and complexly controlled. The detection of pathogenic LRRK2 alterations in 2004 opened a new opening for PD therapy, with a focus on emerging LRRK2 inhibitors. However, drug expansion has established to be highly challenging with absence of knowledge concerning LRRK2 biology, blood-brain barrier (BBB) permeability restriction, and a lack of preclinical models that authentically summarize PD phenotypes, among many stimulating factors that need to be overcome to advance drugs for potential first-in-human hearings⁴. Numerous LRRK2 inhibitors have been described, but many of them deficiency of selectivity or the capability to pass the blood-brain barrier⁵. One of the most promising and actively pursued targets for the forthcoming pharmaceutical action of PD is LRRK2. Huge efforts are being made in this area from the academic community and the pharmaceutical sector with the goal of developing selective and brain-permeable LRRK2 inhibitors as a treatment for Parkinson's disease.

Most of the features of uncommonly big protein, LRRK2, which is classified as a member of the ROCO superfamily are leucine-rich repeat (LRR) domain, a DFG-like motif, a kinase domain, a mixed lineage kinase (MLK) like domain, a GTPase domain, a RAS domain, and a WD40 domain etc. Though, it is related to the outer membrane of mitochondrial, the protein is primarily found in the cytoplasm. LRRK2's physiological function is unclear, and several of its substrates. Nevertheless, numerous LRRK2 inhibitors are utilised as neuroprotective medicines for Parkinson's disease (PD), and it has been hypothesised that they may be useful for avoiding neurodegeneration. Furthermore, multiple investigations revealed that LRRK2 the amount of -synuclein that aggregates in dopaminergic neurons exposed to -synuclein fibrils have been found increased by mutations. Drug virtual screening or optimisation heavily rely on QSAR/QSPR modelling, which has emerged as one of the core computational molecular modelling technique. QSAR models make it possible to pinpoint connections between a physicochemical or biological feature under investigation and the structural details of chemical substances (molecular descriptors). These techniques are frequently employed in place of experimental research nowadays to envisage the activity of molecules grounded on their structure. Particularly, over the past few decades, machine learning techniques have seen widespread use in this discipline. A minimal amount of QSAR investigations in LRRK2 have been available in the scholarly literature. In addition to this, the majority of research papers revealed a little predictive activity for the datasets used for external validation. The information and findings reported in a conference article presented at the 12th International Conference on Practical Applications of Computational Biology & Bioinformatics are extended in that paper with new QSAR models for envisaging potential inhibitors of the LRRK2 protein. Specifically, a number of regression and classification QSAR models are compared for accuracy and model complexity, along with their accuracy performances⁶. LRRK2 also exhibits GTPase activities, playing a crucial role in the regulation of intracellular processes. Leucine-rich repeat kinase 2 (*LRRK2*) is located on human chromosome 12 with 51 exons and encodes a protein of 286 kDa. Over-expression of mutant LRRK2 has been found to be toxic in neurons. Many studies have been undertaken to screen the mutation of this gene in several specific regions, in order to design inhibitors for suppressing PD and LRRK2 toxicity in neurons. One of the most repetitive mutations in *LRRK2* occurs in the G2019S substitution, and is attracting growing attention to target LRRK2 toxicity by creating specific inhibitors targeting this site. Several approaches have been developed to design various LRRK2 inhibitors to treat PD⁷.

In QSAR modeling activity or property are correlated with their chemical parameters. Primarily QSAR procedures depend on experimental data quality, accurate chemical structures, distribution of response/activity and variability of the structures between compounds dataset. A perfect QSAR model is categorized by predefined endpoint, proper applicability domain, correct indicator of goodness of fit and robustness. QSAR models have been extensively used to develop and design antioxidant, anticancer, anti-inflammatory, antitubercular and antiviral reagents.⁸ QSAR study is a beneficial tool to find relationship between molecular descriptors and biological activity of diverse classes of compound.

QSAR play an important role in novel drug discovery and it finds various applications in predicting the activity of novel compounds by mathematical expression which figure out the connection between chemical structure to their biological activity and a QSAR model give information that is very useful for drug design and medicinal chemistry. In recent studies, CADD base on QSAR has been very important to develop original medicines for the handling of diverse ailments.⁹

MATERIALS AND METHODS

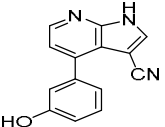
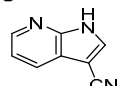
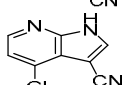
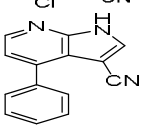
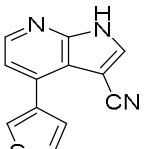
Dataset Preparation

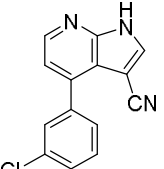
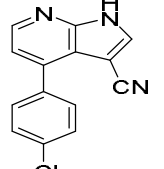
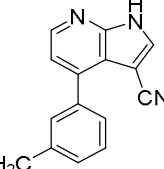
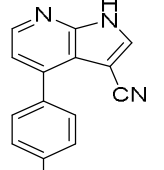
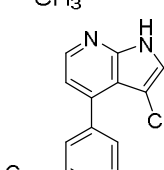
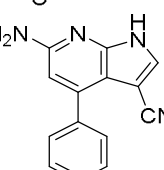
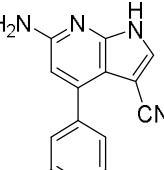
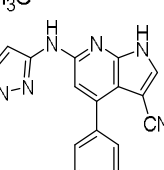
A dataset of 58 compounds of pyrrolo[2,3-b]pyridines, Pyrrolo[2,3-d]pyrimidines, and Pyrrolo[3,2-c]pyridine were taken into consideration for the present study. These LRRK2 inhibitors were taken because their inhibitor activity was noted from the same lab using enzymatic protocol.¹⁰⁻¹¹ The reported IC_{50} values of the compounds were converted into pIC_{50} by taking negative logarithms of IC_{50} values i.e., $pIC_{50} = -\log_{10}(IC_{50})$. The structures were drawn by using Marwin Sketch software.¹² The inhibition potential (pIC_{50}) of 58 pyridines and pyrimidines derivatives with their structure and descriptors values were shown in Table 1.

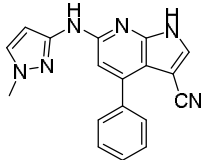
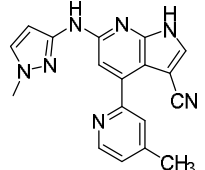
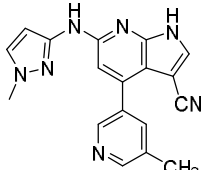
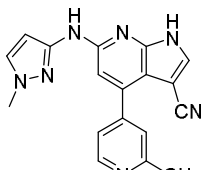
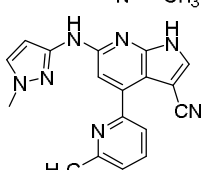
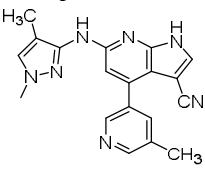
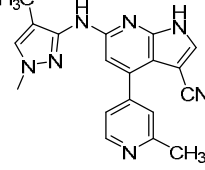
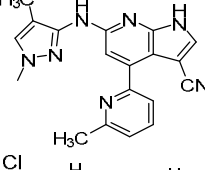
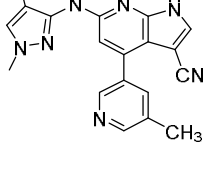
Calculation of Descriptors

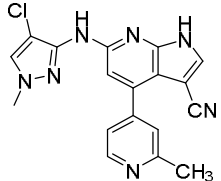
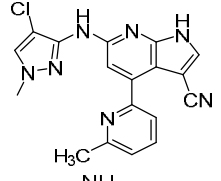
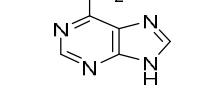
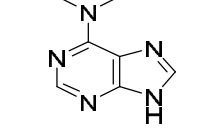
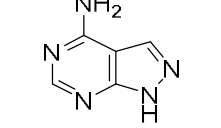
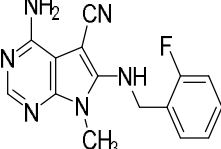
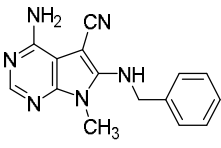
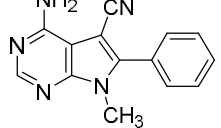
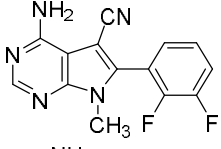
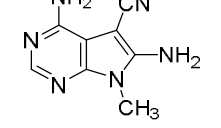
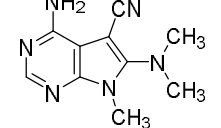
All the OECD guidelines were followed in this research. Marwin sketch software was used to drawing the structures and saved them in form of MDL mol format. Before developing a QSAR model all these structures details were converted into numerical values i.e., in the form of descriptors. PaDEL descriptor software was used for calculation of descriptors.¹³⁻¹⁵ All these specific classes of descriptors were used for development of model which includes connectivity, topological, functional, E-state indices, constitutional, 2D autocorrelation, 2D atom pairs, atom centered fragments, ring and molecular property descriptors. The calculated descriptors were pretreated to decrease noisy and redundant data; intercorrelated ($|r| > 0.95$) variables and constant (variance < 0.0001) were removed with the help of a software available at <http://dtclab.webs.com/software-tools> before developing model^{13-14, 16}.

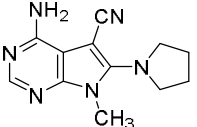
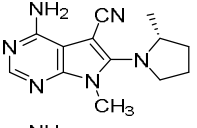
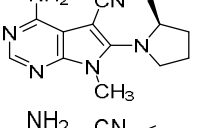
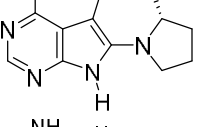
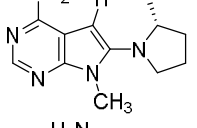
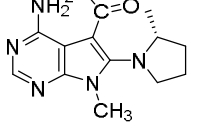
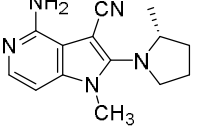
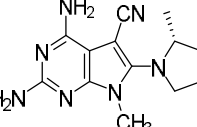
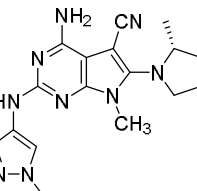
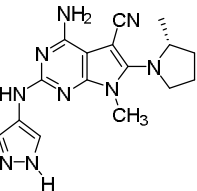
Table:1 Structures of the compounds along with the pIC_{50} value and details of the descriptors

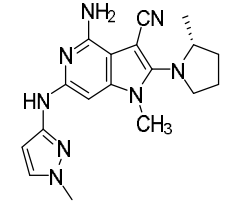
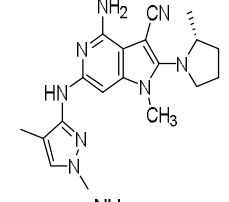
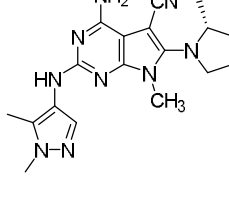
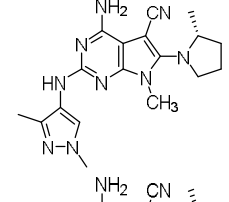
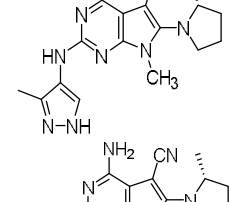
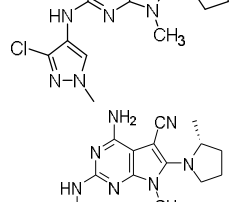
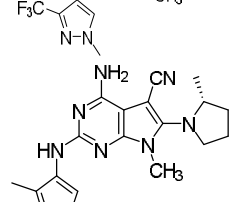
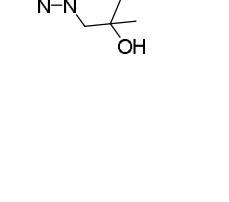
Sr.No	Paper Sr.No.	Structure of the Compound	G2019S IC_{50} (nM)	G2019S pIC_{50}	ATSC7i	ATSC8v	piPC7	Ref
01	D17009		415	6.381952	-1.760448	-9.057821	7.074818	10
02	D17010		46522	4.332342	1.917014	0.000000	5.885561	10
03	D17011		37765	4.422910	2.033808	0.000000	5.995383	10
04	D17012		1178	5.928855	2.600434	188.838347	7.008406	10
05	D17013		870	6.060481	3.997203	197.545046	6.944614	10

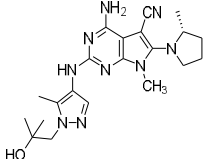
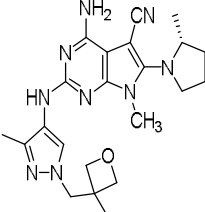
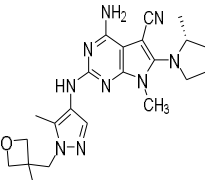
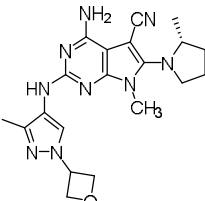
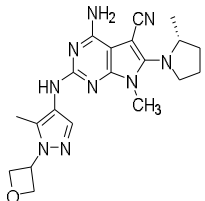
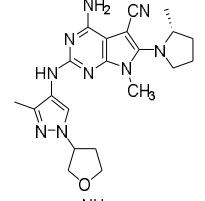
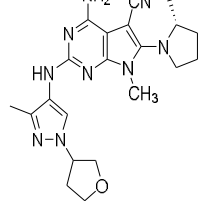
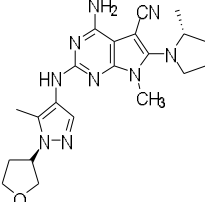
07	D17015		1275	5.894490	1.584154	-49.900403	7.074818	10
08	D17016		1983	5.702677	4.980930	292.005427	7.068372	10
10	D17018		195	6.709965	-	-272.667281	7.074818	10
					13.789183			
11	D17019		1639	5.785421	-0.968441	-35.144997	7.068372	10
12	D17020		354	6.450997	-1.248629	-236.557068	7.112646	10
13	D17021		176	6.754487	-0.370661	251.637167	7.118814	10
14	D17022		33	7.481486	-	-759.795860	7.184262	10
					17.897473			
15	D17023		8	8.096910	-	-885.839505	7.405524	10
					18.055600			

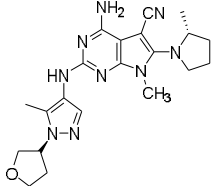
16	D17024		112	6.950782	-1.922605	-151.442454	7.353412	10
17	D17025		324	6.489455	-	-808.666719	7.405524	10
18	D17026		37	7.431798	-	-	7.405524	10
19	D17027		55	7.259637	-6.814510	-642.844344	7.405524	10
20	D17028		41	7.387216	-	-808.666719	7.405524	10
21	D17029		48	7.318759	-	-	7.420794	10
22	D17030		108	6.966576	-9.735154	-755.862052	7.420794	10
23	D17031		50	7.301030	-	-997.713844	7.420794	10
24	D17032		68	7.167491	-9.840467	-781.554250	7.420794	10

25	D17033		158	6.801343	-5.899018	-335.218667	7.420794	10
26	D17034		57	7.244125	-	-582.015706	7.420794	10
27	D21007		15809	4.801096	0.000000	0.000000	5.550200	11
28	D21008		3355	5.474307	3.717716	0.000000	5.808823	11
29	D21009		26630	4.574629	0.000000	0.000000	5.550200	11
30	D21010		1739	5.759700	-	-16.517645	6.615992	11
31	D21011		3532	5.451979	-1.899041	81.958147	6.590440	11
32	D21012		483	6.316053	-5.906541	169.616115	7.202807	11
33	D21013		159	6.798603	8.287690	121.089670	7.327549	11
34	D21014		8564	5.067323	5.159937	0.000000	6.245092	11
35	D21015		3764	5.424350	2.396690	1014.255503	6.407872	11

36	D21017		1909	5.719194	6.941728	-37.904411	6.586312	11
37	D21018		11	7.958607	11.404943	-700.207922	6.635752	11
38	D21019		3083	5.511026	11.404943	-700.207922	6.635752	11
39	D21020		201	6.696804	-2.295964	-657.981760	6.544732	11
40	D21021		6387	5.194703	1.215395	-787.836168	6.363472	11
41	D21022		2240	5.649752	9.574995	15.226358	6.635752	11
42	D21024		28	7.552842	-5.156219	-420.175727	6.328160	11
43	D21025		217	6.663540	8.339716	-968.344733	6.777976	11
44	D21026		5	8.301030	36.599306	-	7.062064	11
45	D21027		6	8.221849	21.977961	-	7.044558	11

46	D21029		206	6.686133	-8.272744	-451.888614	6.778643	11		
47	D21030		1287	5.890421	-	15.889061	-620.365859	6.801213	11	
48	D21031		43	7.366531	26.662685	-	1325.723554	7.082109	11	
49	D21032		30	7.522879	26.662685	-	1325.723554	7.082109	11	
50	D21033		27	7.568636	12.105751	-	1189.442458	7.064953	11	
51	D21034		21	7.677781	36.450662	-	1042.471990	7.082109	11	
52	D21035		358	6.446117	-	21.990524	-	1052.410379	7.111442	11
53	D21036		42	7.376751	16.359180	1.436714	-	7.137260	11	

54	D21037		35	7.455932	42.428740	- 1000.467637	7.133227	11
55	D21038		51	7.292430	14.914531	20.589104	7.161992	11
56	D21039		10	8.000000	50.473493	-975.947161	7.156889	11
57	D21040		25	7.602060	16.645397	- 1014.015772	7.147934	11
58	D21041		37	7.431798	9.322089	- 1014.015772	7.147049	11
59	D21042		12	7.920819	15.033333	-662.022070	7.154988	11
60	D21043		28	7.552842	15.033333	-662.022070	7.154988	11
61	D21044		12	7.920819	17.290257	-968.362792	7.156451	11

62	D21045		14	7.853872	17.290257	-968.362792	7.156451	11
----	--------	---	----	----------	-----------	-------------	----------	----

Dataset Division

In the present work, the recommended standard protocol and all the OECD guidelines as followed by different researchers have been followed. The dataset used for QSAR model development was divided into training and test set. Data set splitting is a significant step in QSAR modeling. It leads to the formation of predictive powers of the model. Due to non-availability of required datasets researchers faces problems in QSAR modeling.¹⁶

Both these training and test sets in QSAR study were used for different purposes. Training set was utilized to build a model on the other hand, prediction (test) set was used to validate the model during QSAR study.¹⁷

Model Development of QSAR

The main aim of analysis by QSAR method is to recognize structural features that effects the activity of series of molecules and to identify activity in prior to actual synthesis of a compound. QSAR model can be quantitative or qualitative both.¹⁸ The primary principle and application of QSAR model development was to obtain maximum information of the activity which was related to structural features and before real synthesis and bio-screening molecule's desired activity was to be predicted. Hence to attain these goals multiple QSAR models will be developed by using alienated dataset and easily understandable descriptors were selected during model generation. The dataset was converted into training (70%) and prediction (30%) set by means of Random faster method of division. The different QSAR modelling approach confirmed that for molecular descriptors maximum information can be gained that direct the biological profile of the molecules. QSARIN Chem 2.2.1 software was used to develop diverse QSAR models.¹⁹⁻²¹

Validation of the Model

Model validation was an important feature in QSAR model building. The GA-MLR equations statistical validity were recognized by means of internal or cross validation by LMO and LOO procedure; with the help of test set; Y-scrambling or data randomization and by checking that the following conditions of statistical parameters were satisfied or not:

$$R_{tr}^2 \geq 0.6, Q_{loo}^2 \geq 0.5, Q_{LMO}^2 \geq 0.6, R^2 > Q^2, R_{ext}^2 \geq 0.6, RMSE_{tr} < RMSE_{cv}, \Delta K \geq 0.05, Q^2 - F^n \geq 0.60, CCC \geq 0.80, r_m^2 \geq 0.6, \frac{(1-r^2)}{r_0^2} < 0.1, 0.9 \leq k \leq 1.1 \text{ or } \frac{(1-r^2)}{r_0'^2} < 0.1, 0.9 \leq k' \leq 1.1, |r_0^2 - r_0'^2| < 0.3 \text{ with RMSE and MAE close to zero.}$$

The values of statistical parameters within this range ensures the external predictive validity and robustness of the developed model.²²

Tropsha and Golbraikh parameters

Predictivity assessment criteria as given by Tropsha and Golbraikh was also checked for the developed PLS models. According to these criteria the condition for acceptance of QSAR model is:

$$\frac{r^2 - r_0^2}{r^2} < 0.1 \text{ or } \frac{r^2 - r_0'^2}{r^2} < 0.1 \quad (1)$$

$$0.85 < k < 1.15 \text{ or } 0.85 < k' < 1.15 \quad (2)$$

Y-Randomization

Y-randomization was performed from the test set. It was an external validation technique used to find out that a developed QSAR model was reliable, strong and not formed by luck. If the values of Q^2 and R^2 were low then it indicates that model was very reliable and robust.⁹ Y-randomization was performed from the test set. It was an external validation technique used to find out that a developed QSAR model was reliable, strong and not formed by luck. If the values of Q^2 and R^2 were low then it indicates that model was very reliable and robust. Here, to perform the Y-randomization of the developed QSAR model, DTC software is used. This is done in order to check whether the model is established by accidentally or not. The number of arrangements may differ. According to this method, Y and X variables could be transposed as per their fit into recorded models and therefore, various models could be formed by taking dissimilar groupings. These number of combinations might differ according to different study. But here in the present study, only Y-variables were permuted upon 50 combinations. If the value of statistical metrics of the randomized model were lesser than the real one, then it can be supposed that the model was not established coincidentally.²³

Molecular Docking

The various interactions like H-bonding, hydrophobic and electrostatic interactions play very important roles in the model. Molecular docking expects the interactions and binding mode between the receptor protein and ligand.²⁴ The initial information about interaction between ligand and protein is provided by Molecular docking. Before performing docking study the complex is prepared using pymol.²⁵ Molecular docking is one of the well-recognized and extensively used in-silico structure-based detection methods. Docking predicts or describes protein-ligand interactions at the molecular level. Docking studies are used to discover configurations using ligands inside the binding pocket of the protein (macromolecule).²⁶ It is generally preferred when there is the availability of sufficient evidence and knowledge about the receptor or target protein with which the drug interacts. A molecular docking study is performed by the AutoDock software tool. The crystal structure of the receptor having **PDB ID:7BK2** is retrieved from the RCSB site. First of all, the ligands and water molecules were deleted from the PDB protein file. Both the protein and ligand files were converted into pdbqt format by using OpenBabel software. This is a very important step before docking.²⁷ The side and terminal chains of the receptor (protein) were repaired. Polar hydrogen atoms and Kollman charges were added. Non-polar hydrogen atoms and Gasteiger partial charges were assigned for the inhibitors and all the single bonds were set to flexible. This software had various graphics tools to display different docking poses of the molecule.²⁸⁻²⁹ Discovery Studio Visualizer was used to visualize the docking results. A grid box was prepared by taking grid parameters X=70, Y=60, and Z=60 having a grid spacing of 0.375 was generated.³⁰⁻³¹ The population size of 150 was set having a mutation rate of 0.02 developed for 10 existences. Lamarckian Genetic Algorithm was considered.²⁰ The validity of the docking protocol is checked by calculating the RMSD value. If this value is less than 2Å then the docking protocol is good. This result validates the docking.³²⁻³³ Different visualization software like Pymol³⁴, UCSF ChimeraX³⁵, and Discovery Studio Files was used for visualization of docking results.³⁶⁻³⁷

RESULTS AND DISCUSSION

QSAR Models

The compounds of dataset used in the present study are either functional or positional isomers. The best QSAR model is selected from all the models on the basis of statistical performance and applying first among equal approach.³⁸ To find out the connection between chemical activities and structures of the compound QSAR computational modeling method is used. As a result, every compound can be expressed in form of numerical values called descriptors. These descriptors can be attained by structural calculation and can be utilized in form of independent variables (predictors or X variable) to guess genomic movement (Y variable) of the compounds.³⁹ SFS (Subjective Feature Selection) method was used for QSAR modeling. For applying this SFS method various methods like GA (Genetic Algorithm), Stepwise regression etc. methods are used. All these methods result in a good QSAR model

with similar statistical parameters and different descriptors. ‘First Among Equal Approach’ was used for selection of best QSAR model. In the current study dataset was small but the compounds were either structural or positional isomers. It is very difficult for a QSAR modeler to develop model if the dataset was small as some compounds were holding out for validation purposes. The strategy adopted in this study for the QSAR model development was shown in Figure.2.

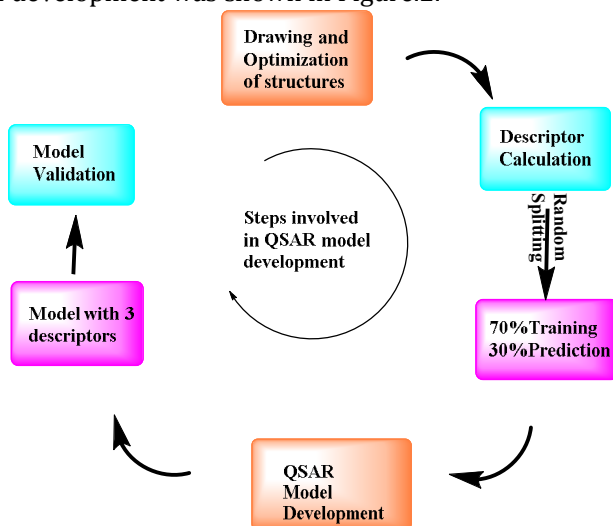


Figure.1 Strategy adopted for development of QSAR model

This figure clearly explained that one QSAR model was developed in the present study. A dataset of 58 compounds of pyrrolo[2,3-b]pyridines, Pyrrolo[2,3-d]pyrimidines, and Pyrrolo[3,2-c]pyridine were taken into consideration for the present study.¹⁰⁻¹¹ From this dataset 41 compounds were used as training set and 17 compounds were used as test set. The MLR model development for internal validation showed a good result for the coefficient of determination (R²) and Q² LOO i.e. leave-one-out squared correlation coefficient. In this model, three descriptors (ATSC7i, ATSC8v, piPC7) were applied.

The detailed value of descriptors along with their description is shown in Table 2. The model equation for the obtained descriptor is shown below:

Model: (Divided dataset)

$$pIC_{50} = -4.0857(\pm 2.4307) + 0.0145(\pm 0.0092) \times ATSC7i - 0.0007(\pm 0.0003) \times ATSC8v + 1.4898(\pm 0.357) \times piPC7$$

Table 2: Descriptor along with their type and correlation with the model

Descriptor	Description	Type	Correlation with Model
ATSC7i	Centered Broto-Moreau autocorrelation - lag 7 / weighted by first ionization potential	2D autocorrelation descriptor	Positive
ATSC8v	Average centred Broto-Moreau autocorrelation - lag 8 / weighted by van der Waals volumes	2D autocorrelation descriptor	Negative
piPC7	Conventional bond order ID number of order 7 (ln(1+x))	2D path counts descriptor	Positive

ATS is defined as autocorrelation of a topological structure. These descriptors don't give 3D information.⁴⁰ A topological descriptor proposed by Moreau and Broto also encircled numerical

properties in addition to structure of the molecules.⁴¹ **ATSC7i** was positively correlated to this model. Autocorrelation descriptors are used in QSAR studies because they are specific for a particular geometry. These descriptors are very sensitive if there is slight change in conformation. One of the drawbacks of these descriptors might be that about the molecular structure we can't recreate original data. The researchers named Moreau and Broto were the first who applied an autocorrelation function on the molecular graph for measuring the atomic properties like electronegativities, charges etc. These descriptors have some positive benefits in QSAR/QSPR studies like fragment independent and also invariant to roto-translation. With the help of these descriptors, we can convert uniqueness, atom types, electronegativities etc.⁴²

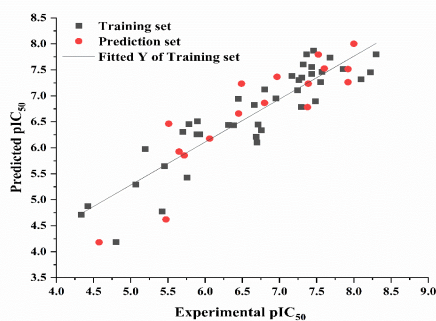
ATSC8v is negatively correlated in the model. It is the descriptor which is weighted by vander waals volumes. **piPC7** is a path counts descriptor positively correlated in the model. The numerical values of descriptor obtained from this model for the LRRK2 inhibitors along with their experimental *pIC₅₀* is shown in Table 3.

Table 3: Detailed value of descriptor obtained for LRRK2 inhibitors

Sr. No.	Name as per lit.	G2019S <i>pIC₅₀</i>	ATSC7i	ATSC8v	piPC7
V.1	D17009	6.3819519	-1.7604477	-9.0578209	7.0748178
V.2	D17010	4.3323416	1.9170144	0.0000000	5.8855613
V.3	D17011	4.4229105	2.0338077	0.0000000	5.9953826
V.4	D17012	5.9288547	2.6004342	188.83835	7.0084063
V.5	D17013	6.0604807	3.9972028	197.54505	6.9446144
V.6	D17015	5.8944898	1.5841543	-49.900403	7.0748178
V.7	D17016	5.7026773	4.9809301	292.00543	7.0683716
V.8	D17018	6.7099654	-13.7891836	-272.66728	7.0748178
V.9	D17019	5.7854210	-0.9684410	-35.144997	7.0683716
V.10	D17020	6.4509967	-1.2486287	-236.55707	7.1126457
V.11	D17021	6.7544873	-0.3706612	251.637167	7.1188136
V.12	D17022	7.4814861	-17.8974733	-759.79586	7.1842619
V.13	D17023	8.0969100	-18.0555997	-885.83950	7.4055242
V.14	D17024	6.9507820	-1.9226050	-151.44245	7.3534122
V.15	D17025	6.4894550	-19.7868609	-808.66672	7.4055242
V.16	D17026	7.4317983	-10.6465474	-1068.0901	7.4055242
V.17	D17027	7.2596373	-6.8145097	-642.84434	7.4055242
V.18	D17028	7.3872161	-19.7868609	-808.66672	7.4055242
V.19	D17029	7.3187588	-13.5145897	-1168.8841	7.4207941
V.20	D17030	6.9665762	-9.7351543	-755.86205	7.4207941
V.21	D17031	7.3010300	-22.576634	-997.71384	7.4207941
V.22	D17032	7.1674911	-9.8404666	-781.55425	7.4207941
V.23	D17033	6.8013429	-5.8990185	-335.21867	7.4207941
V.24	D17034	7.2441251	-19.1430287	-582.01571	7.4207941
V.25	D21007	4.8010956	0.0000000	0.0000000	5.5502002
V.26	D21008	5.4743075	3.7177165	0.0000000	5.8088226
V.27	D21009	4.5746288	0.0000000	0.0000000	5.5502002
V.28	D21010	5.7597004	-24.3291840	-16.517645	6.6159920
V.29	D21011	5.4519793	-1.8990406	81.958147	6.5904405
V.30	D21012	6.3160529	-5.9065405	169.61612	7.2028066
V.31	D21013	6.7986029	8.2876898	121.08967	7.3275495
V.32	D21014	5.0673233	5.1599373	0.0000000	6.2450918
V.33	D21015	5.4243504	2.3966897	1014.2560	6.4078722

V.34	D21017	5.7191941	6.9417285	-37.904400	6.5863117
V.35	D21019	5.5110265	11.4049429	-700.20792	6.6357517
V.36	D21020	6.6968039	-2.2959637	-657.98176	6.5447321
V.37	D21021	5.1947031	1.2153955	-787.83617	6.3634725
V.38	D21022	5.6497520	9.5749955	15.226358	6.6357517
V.39	D21025	6.6635403	8.3397158	-968.34473	6.7779758
V.40	D21026	8.3010300	36.5993057	-1169.2417	7.0620644
V.41	D21027	8.2218487	21.9779610	-1014.5374	7.0445577
V.42	D21029	6.6861328	-8.2727438	-451.88861	6.7786427
V.43	D21030	5.8904214	-15.8890610	-620.36586	6.8012135
V.44	D21031	7.3665315	26.6626847	-1325.7235	7.0821090
V.45	D21032	7.5228787	26.6626847	-1325.7235	7.0821090
V.46	D21033	7.5686362	12.1057510	-1189.4425	7.0649526
V.47	D21034	7.6777807	36.4506625	-1042.4720	7.0821090
V.48	D21035	6.4461170	-21.9905243	-1052.4104	7.1114420
V.49	D21036	7.3767507	16.3591804	1.436713872	7.1372598
V.50	D21037	7.4559320	42.4287403	-1000.467637	7.1332274
V.51	D21038	7.2924298	14.9145311	20.58910396	7.1619916
V.52	D21039	8.0000000	50.4734934	-975.9471611	7.1568894
V.53	D21040	7.6020600	16.6453966	-1014.015772	7.1479341
V.54	D21041	7.4317983	9.3220893	-1014.015772	7.1470490
V.55	D21042	7.9208187	15.0333334	-662.0220696	7.1549876
V.56	D21043	7.5528420	15.0333334	-662.0220696	7.1549876
V.57	D21044	7.9208187	17.2902568	-968.3627918	7.1564508
V.58	D21045	7.8538720	17.2902568	-968.3627918	7.1564508

The results obtained from the QSAR model are presented in Tables 4. The values of all statistical parameters were obtained within the required limits, representing a successful and consistent model. The values of experimental and predicted endpoints were nearly equal which proves the result. The numerical values of R^2 for the training and test set were 0.8261 and 0.7998. The values of all the required statistical parameters like R^2 , Q^2 , and MAE, including fitting criteria, internal and external validation criteria, and predictions by LOO and model equation are shown in Tables 4. The developed QSAR model was found to have reliable and satisfactory numerical values for various validation matrices which include R^2 , Q^2 , Q^2_{LMO} , $R^2_m(LOO)$, $\Delta R^2_m(LOO)$, MAE and RMSE. Figure 2 displays the graph between the predicted versus experimental pIC_{50} and residual versus experimental pIC_{50} values. The applicability domain was studied using William's plot ($h^* = 0.2927$). All compounds fall inside the domain of applicability (Figure 3).



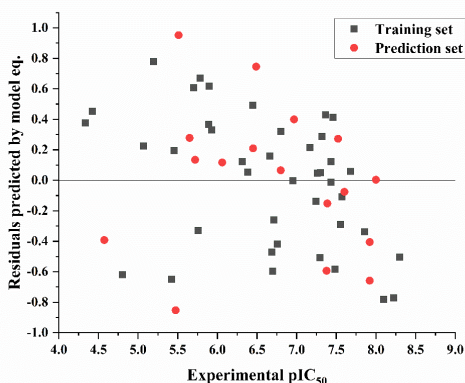


Figure 2: (i) Description of plot of Predicted versus experimental end point (ii) Description of predicted model res. equation versus experimental end point

Table 4: Values of statistical parameters obtained from QSAR model

Statistical parameters	Result	Statistical parameters	Result
N_{tr}	41	No. of descriptors	3
N_{ex}	17	θ^*	-2.6018°
Fitting Criteria			
R^2_{tr}	0.8261	$RMSE_{tr}$	0.4249
R^2_{adj}	0.812	MAE_{tr}	0.3603
$R^2_{tr} - R^2_{adj}$	0.0141	RSS_{tr}	7.4023
LOF	0.2478	CCC_{tr}	0.9048
Kxx	0.2472	s	0.4473
ΔK	0.2102	F	58.5977
Internal Validation Criteria		External Validation Criteria	
$R^2_{cv}(Q^2_{loo})$	0.7769	$RMSE_{ext}$	0.4674
$R^2 - Q^2_{loo}$	0.0492	MAE_{ext}	0.3711
$RMSE_{cv}$	0.4813	$PRESS_{ext}$	3.7139
MAE_{cv}	0.4052	R^2_{ext}	0.7998
$PRESS_{cv}$	9.4984	$Q^2 - F^1$	0.784
CCC_{cv}	0.88	$Q^2 - F^2$	0.7837
Q^2_{LMO}	0.7701	$Q^2 - F^3$	0.7896
R^2_{Yscr}	0.0776	CCC_{ext}	0.8941
Q^2_{Yscr}	-0.1368	$R^2_{m aver.}$	0.7161
$RMSE^{AV}_{Yscr}$	0.9782	$R^2_m \text{ delta}$	0.0321
Predictions by LOO			
Exp(x) vs. Pred(y) R^2	0.7789	Pred(x) vs. Exp(y) R^2	0.7789
R^2_o	0.7414	R^2_o	0.7769
k'	0.9952	k	0.9996
Clos'	0.0481	Clos	0.0025
R^2_m	0.6281	R^2_m	0.7443
External predictions by model equation			
Exp(x) vs. Pred(y) R^2	0.7998	Pred(x) vs. Exp(y) R^2	0.7998
R^2_o	0.7926	R^2_o	0.7842
k'	0.9985	k	0.9967
Clos'	0.009	Clos	0.0195
R^2_m	0.7321	R^2_m	0.7

R^2 : coefficient of determination; R^2_{adj} : adjusted R^2 , LOF: lack of fit; CCC_{tr} : concordance correlation coefficient for training set; CCC_{cv} : concordance correlation coefficient of cross-validation; F:

Fischer's statistics; R^2_{Yscr} : response scrambling coefficient; Q^2_{Yscr} : cross-validation response scrambling coefficient; Q^2_{LOO} : leave one out cross validation coefficient; Q^2_{LMO} : leave many out cross validation coefficient; Q^2F_1 , Q^2F_2 , Q^2F_3 : External validation criteria; Delta K: difference in the correlation; RMSE: root mean square error; MAE: mean absolute error; S: standard error

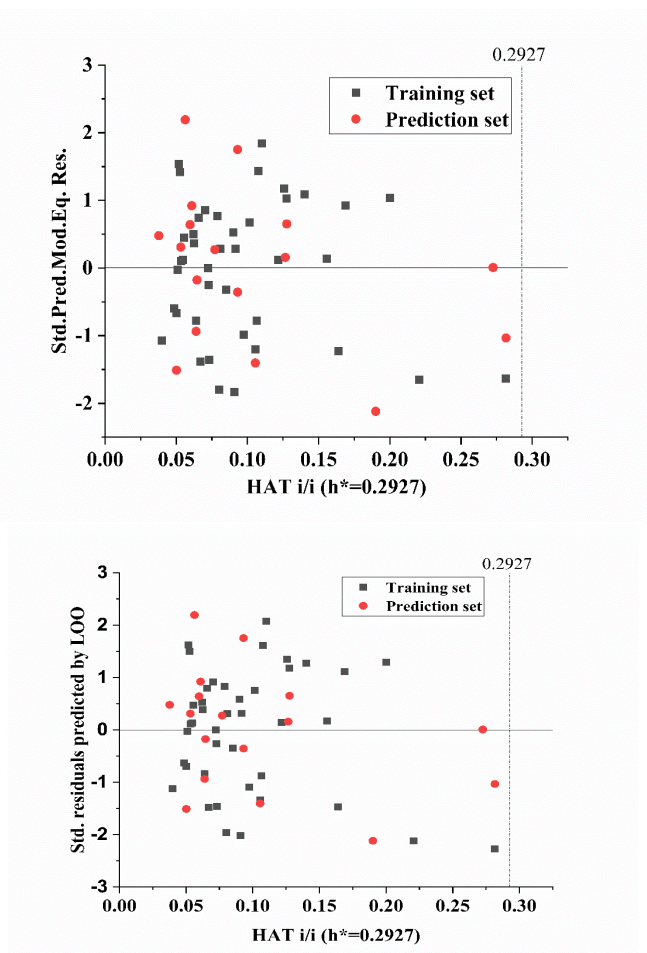


Figure 3: Williams plots for LRRK2 inhibitor

Molecular Docking

In order to fully understand the binding interactions between the receptor protein and ligand as they are simulated by 3D-QSAR models, a complete technique for this research is called molecular docking.^{4,5} Using the Autodock tool, the potential binding configuration was investigated. The RMSD ($<2 \text{ \AA}$) value of the redocked conformation was frequently used to assess the reliability of the docking process. The internal ligand's RMSD value was determined to be less than 2 \AA , indicating the docking protocol's dependability and reproducibility.⁶ Figure 4 and Figure 5 also depicts a potential interaction between the internal ligand and the receptor's active site. The hydrophobic contact between different amino acid residues, including LYS 38, MET 84, ALA 147, LEU 59, and the most effective molecule D21026, was deduced from the fact that this interaction occurs. Molecular docking was used to clarify the binding mechanism of the most effective drug, D21026, within the active site of 7BK2 following the successful validation of the docking procedure. Through the hydrogen of the OH group, the most powerful molecule, D21026, forms a hydrogen bond with the important residue GLU 17. Overall,

D21026 interactions mirrored those observed between co-crystallized ligand and 7BK2. The robust binding conformation of the powerful combination was therefore confirmed by the docked position.

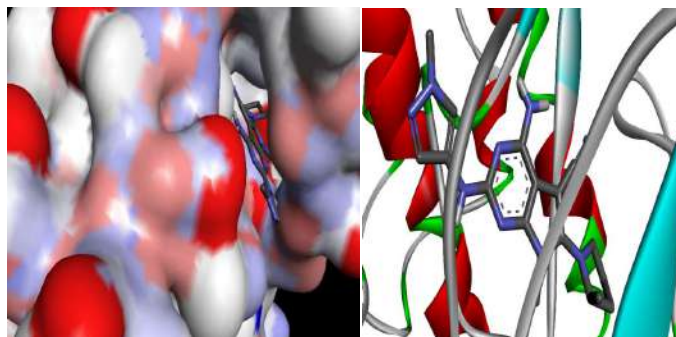


Figure 4 : 3D sturcture and cartoon structure of ligand D21026 and receptor 7BK2

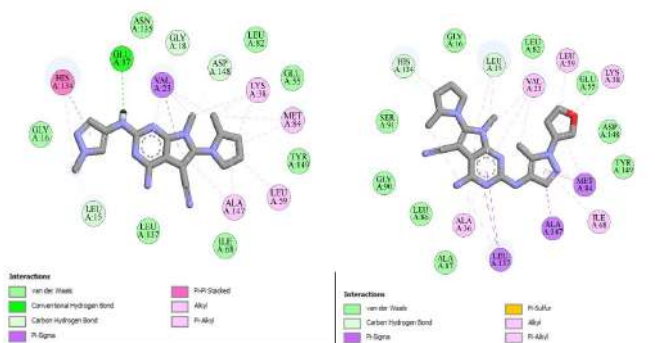
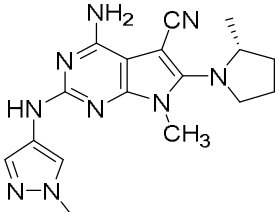
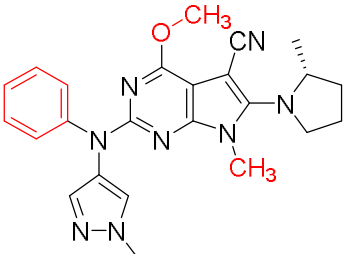
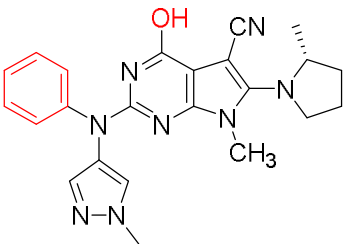
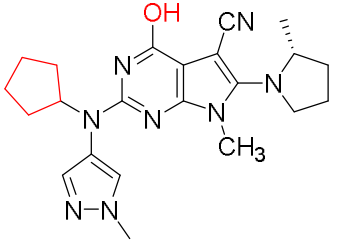
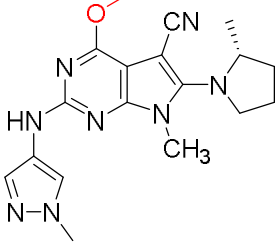
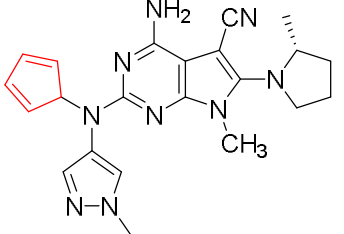


Figure 5: 2D structure of 34S and UOK with receptor 7BK2

Details of Designed Compounds

Some drugs (H1-H5) were created (Table 5) using the molecular modelling methods used in this investigation. The machine learning techniques utilised in this study were then able to forecast these compounds' inhibitory potential (pIC_{50}) (Table 5). In order to investigate the manner of interactions in the active site of LRRK2 (PDB ID: 7BK2), the developed compounds (H-1 to H-5) were further examined through molecular docking research. All of these substances fit perfectly into the active site of LRRK2 (PDB ID: 7BK2), according to a molecular docking analysis. Similarly to their projected pIC_{50} , the docking score or binding affinity was also discovered. In Figure 6, the developed compounds' 2D structures are displayed.

Table 5: Predicted pIC_{50} of the designed compounds with their structure and binding affinity

S.No.	Structure	Predicted pIC_{50}	Binding energy(kcal/mol)
D21026		8.301	-6.97
V.H1		9.424953	-8.91
V.H2		8.531437	-8.77
V.H3		8.439507	-8.40
V.H4		8.332397	-7.83
V.H5		8.322227	-7.52

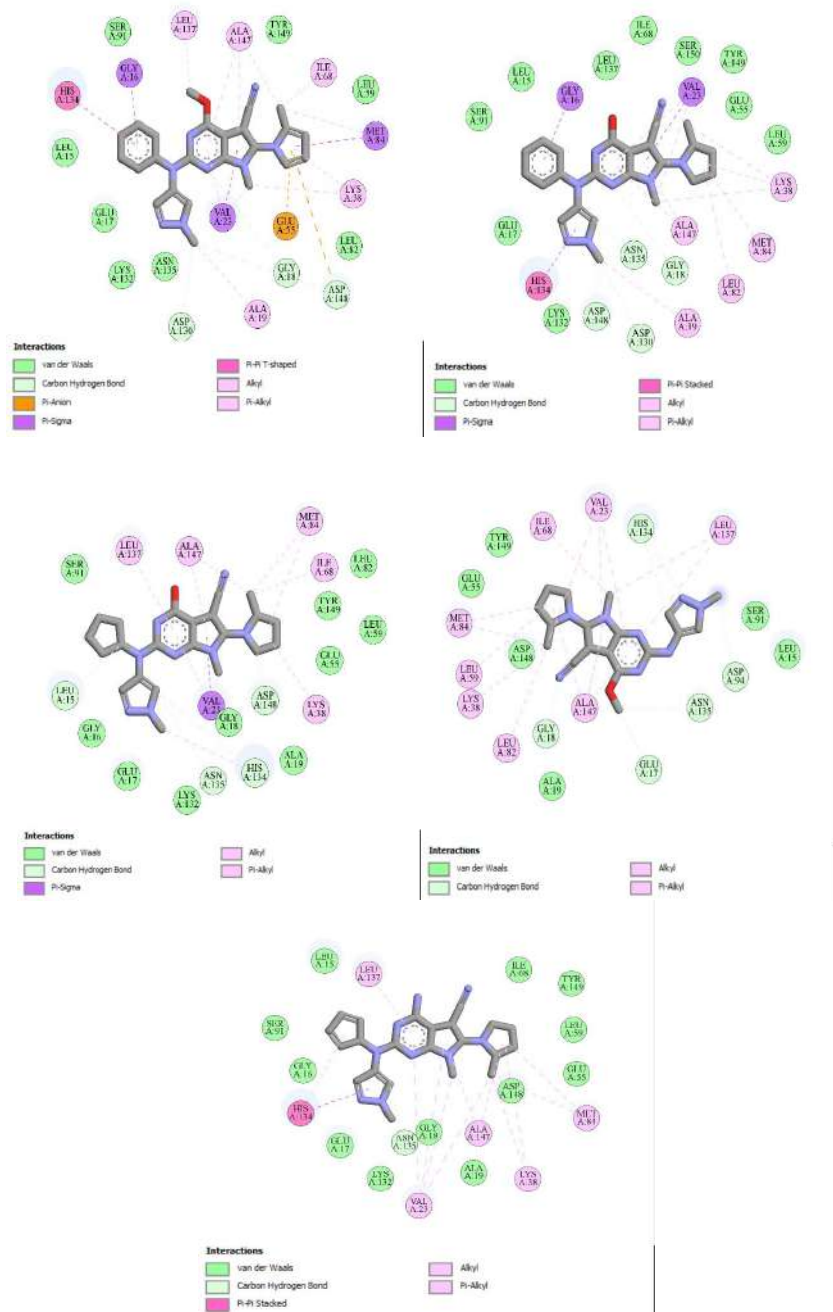


Figure 6: 2D structure of designed compounds from H1-H5 respectively

CONCLUSION

The values of statistical parameter for this model were 0.8261, 0.8120, 0.7769, 0.7998 and 0.8941 for R^2 , R^2_{adj} , Q^2_{loo} , R^2_{ext} , and CCC_{ext} respectively. A molecular docking study of the most active compound (D21026) and designed compounds (H1-H5) was done to find the best interaction between protein and ligand. The docking results found that the designed compounds showed similar pattern of interaction as that of the most active compound(D21026). The hydrophobic contact between different amino acid residues, including LYS 38, MET 84, ALA 147, LEU 59, and the most effective molecule D21026, was deduced from the fact that this interaction occurs. Molecular docking was used to clarify

the binding mechanism of the most effective drug, D21026, within the active site of 7BK2 following the successful validation of the docking procedure. Through the hydrogen of the OH group, the most powerful molecule, D21026, forms a hydrogen bond with the important residue GLU 17. Some compounds (H1-H5) were created based on the molecular modelling methods used in this investigation. The machine learning techniques employed in this work subsequently predicted the inhibitory potential (pIC_{50}) of these compounds. Hence, the results of the present investigation may be employed to identify and develop effective inhibitors for the treatment of LRRK2-related pathophysiological disorders.

Abbreviations used

PD- Parkinson disease; **QSAR**-Quantitative structure activity relationship; **MD**-Molecular Dynamics; **QSARINS**- QSAR Insubria; **MLR**-Multiple Linear Regression; **GA**-Genetic Algorithm; **CCC**- Concordance Correlation Coefficient; **OECD**-Organization for Economic Co-operation and development; **AD**- Applicability Domain; **LOO**-Leave-one-out; **LOO-CV**-Leave-one-out-cross-validation; **MAE**-Mean Absolute Error; **PLS**- Partial Least Squares; **RMSE_{ext}** -Root Mean Square Error -external dataset; **RMSE_{cv}**-Root Mean Square Error- cross- validation; **RMSE**-Root Mean Square Error; **R²_{cv}**-Coefficient of determination-cross-validation; **R²_{ext}**-Coefficient of determination-external dataset; **R²**-Coefficient of determination

Conflicts of Interest

The author(s) pronounces that there is no known source of competing financial interests or personal relationship that could have seemed to impact the work stated in this paper.

Acknowledgements

The author is thankful to Dr Parvin Kumar for providing softwares for the calculation of QSAR model development. Dr Parvin Kumar is grateful to Prof. Paola for providing a license for QSARINS. The author is obliged to the establishments of the respective university for providing the needed facilities.

Data availability statement

The data will be made available on request.

REFERENCES:

1. Paisan-Ruiz, C.; Lewis, P. A.; Singleton, A. B., LRRK2: cause, risk, and mechanism. *J Parkinsons Dis* **2013**, *3* (2), 85-103.
2. Tolosa, E.; Vila, M.; Klein, C.; Rascol, O., LRRK2 in Parkinson disease: challenges of clinical trials. *Nat Rev Neurol* **2020**, *16* (2), 97-107.
3. Mata, I. F.; Wedemeyer, W. J.; Farrer, M. J.; Taylor, J. P.; Gallo, K. A., LRRK2 in Parkinson's disease: protein domains and functional insights. *Trends Neurosci* **2006**, *29* (5), 286-93.
4. Atashrazm, F.; Dzamko, N., LRRK2 inhibitors and their potential in the treatment of Parkinson's disease: current perspectives. *Clin Pharmacol* **2016**, *8*, 177-189.
5. Gilsbach, B. K.; Messias, A. C.; Ito, G.; Sattler, M.; Alessi, D. R.; Wittinghofer, A.; Kortholt, A., Structural Characterization of LRRK2 Inhibitors. *J Med Chem* **2015**, *58* (9), 3751-6.
6. Sebastian-Perez, V.; Martinez, M. J.; Gil, C.; Campillo, N. E.; Martinez, A.; Ponzoni, I., QSAR Modelling to Identify LRRK2 Inhibitors for Parkinson's Disease. *J Integr Bioinform* **2019**, *16* (1).
7. Pourbasheer, E.; Aalizadeh, R., 3D-QSAR and molecular docking study of LRRK2 kinase inhibitors by CoMFA and CoMSIA methods. *SAR QSAR Environ Res* **2016**, *27* (5), 385-407.
8. Begum, S.; Jaswanthi, P.; Venkata Lakshmi, B.; Bharathi, K., QSAR studies on indole-azole Analogues using DTC tools; imidazole ring is more favorable for aromatase inhibition. *Journal of the Indian Chemical Society* **2021**, *98* (1), 100016.

9. Olasupo, S. B.; Uzairu, A.; Shallangwa, G.; Uba, S., QSAR analysis and molecular docking simulation of norepinephrine transporter (NET) inhibitors as anti-psychotic therapeutic agents. *Heliyon* **2019**, *5* (10), e02640.
10. Williamson, D. S.; Smith, G. P.; Acheson-Dossang, P.; Bedford, S. T.; Chell, V.; Chen, I. J.; Daechsel, J. C. A.; Daniels, Z.; David, L.; Dokurno, P.; Hentzer, M.; Herzig, M. C.; Hubbard, R. E.; Moore, J. D.; Murray, J. B.; Newland, S.; Ray, S. C.; Shaw, T.; Surgenor, A. E.; Terry, L.; Thirstrup, K.; Wang, Y.; Christensen, K. V., Design of Leucine-Rich Repeat Kinase 2 (LRRK2) Inhibitors Using a Crystallographic Surrogate Derived from Checkpoint Kinase 1 (CHK1). *Journal of Medicinal Chemistry* **2017**, *60* (21), 8945-8962.
11. Williamson, D. S.; Smith, G. P.; Mikkelsen, G. K.; Jensen, T.; Acheson-Dossang, P.; Badolo, L.; Bedford, S. T.; Chell, V.; Chen, I. J.; Dokurno, P.; Hentzer, M.; Newland, S.; Ray, S. C.; Shaw, T.; Surgenor, A. E.; Terry, L.; Wang, Y.; Christensen, K. V., Design and Synthesis of Pyrrolo[2,3-d]pyrimidine-Derived Leucine-Rich Repeat Kinase 2 (LRRK2) Inhibitors Using a Checkpoint Kinase 1 (CHK1)-Derived Crystallographic Surrogate. *Journal of Medicinal Chemistry* **2021**, *64* (14), 10312-10332.
12. Masand, V. H.; Patil, M. K.; El-Sayed, N. N. E.; Zaki, M. E. A.; Almarhoon, Z.; Al-Hussain, S. A., Balanced QSAR analysis to identify the structural requirements of ABBV-075 (Mivebresib) analogues as bromodomain and extraterminal domain (BET) family bromodomain inhibitor. *Journal of Molecular Structure* **2021**, *1229*, 129597.
13. Yap, C. W., PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J Comput Chem* **2011**, *32* (7), 1466-74.
14. Danishuddin; Khan, A. U., Descriptors and their selection methods in QSAR analysis: paradigm for drug design. *Drug Discov Today* **2016**, *21* (8), 1291-302.
15. Pramanik, S.; Roy, K., Modeling bioconcentration factor (BCF) using mechanistically interpretable descriptors computed from open source tool "PaDEL-Descriptor". *Environmental Science and Pollution Research* **2014**, *21*, 2955-2965.
16. De, P.; Roy, K., Nitroaromatics as hypoxic cell radiosensitizers: A 2D-QSAR approach to explore structural features contributing to radiosensitization effectiveness. *European Journal of Medicinal Chemistry Reports* **2022**, *4*, 100035.
17. Nath, A.; De, P.; Roy, K., In silico modelling of acute toxicity of 1, 2, 4-triazole antifungal agents towards zebrafish (*Danio rerio*) embryos: Application of the Small Dataset Modeller tool. *Toxicol In Vitro* **2021**, *75*, 105205.
18. Masand, V. H.; El-Sayed, N. N. E.; Bambole, M. U.; Patil, V. R.; Thakur, S. D., Multiple quantitative structure-activity relationships (QSARs) analysis for orally active trypanocidal N-myristoyltransferase inhibitors. *Journal of Molecular Structure* **2019**, *1175*, 481-487.
19. Masand, V. H.; El-Sayed, N. N. E.; Mahajan, D. T.; Mercader, A. G.; Alafeefy, A. M.; Shibi, I. G., QSAR modeling for anti-human African trypanosomiasis activity of substituted 2-Phenylimidazopyridines. *Journal of Molecular Structure* **2017**, *1130*, 711-718.
20. Edache, E. I.; Uzairu, A.; Mamza, P. A.; Shallangwa, G. A., Structure-based simulated scanning of rheumatoid arthritis inhibitors: 2D-QSAR, 3D-QSAR, docking, molecular dynamics simulation, and lipophilicity indices calculation. *Scientific African* **2022**, *15*, e01088.
21. Manisha; Chauhan, S.; Kumar, P.; Kumar, A., Development of prediction model for fructose- 1,6-bisphosphatase inhibitors using the Monte Carlo method. *SAR QSAR Environ Res* **2019**, *30* (3), 145-159.
22. Masand, V. H.; Mahajan, D. T.; Maldhure, A. K.; Rastija, V., Quantitative structure-activity relationships (QSARs) and pharmacophore modeling for human African trypanosomiasis (HAT) activity of pyridyl benzamides and 3-(oxazolo[4,5-b]pyridin-2-yl)anilides. *Medicinal Chemistry Research* **2016**, *25* (10), 2324-2334.
23. Seth, A.; Roy, K., QSAR modeling of algal low level toxicity values of different phenol and aniline derivatives using 2D descriptors. *Aquat Toxicol* **2020**, *228*, 105627.
24. Ding, L.; Wang, Z. Z.; Sun, X. D.; Yang, J.; Ma, C. Y.; Li, W.; Liu, H. M., 3D-QSAR (CoMFA, CoMSIA), molecular docking and molecular dynamics simulations study of 6-aryl-5-cyano-pyrimidine

- derivatives to explore the structure requirements of LSD1 inhibitors. *Bioorg Med Chem Lett* **2017**, 27 (15), 3521-3528.
25. Babu Singh, M.; Jain, P.; Tomar, J.; Kumar, V.; Bahadur, I.; Arya, D. K.; Singh, P., An In Silico investigation for acyclovir and its derivatives to fight the COVID-19: Molecular docking, DFT calculations, ADME and td-Molecular dynamics simulations. *Journal of the Indian Chemical Society* **2022**, 99 (5), 100433.
 26. Jawarkar, R. D.; Bakal, R. L.; Zaki, M. E. A.; Al-Hussain, S.; Ghosh, A.; Gandhi, A.; Mukerjee, N.; Samad, A.; Masand, V. H.; Lewaa, I., QSAR based virtual screening derived identification of a novel hit as a SARS CoV-229E 3CL(pro) Inhibitor: GA-MLR QSAR modeling supported by molecular Docking, molecular dynamics simulation and MMGBSA calculation approaches. *Arab J Chem* **2022**, 15 (1), 103499.
 27. Yang, R.; Zha, X.; Gao, X.; Wang, K.; Cheng, B.; Yan, B., Multi-stage virtual screening of natural products against p38alpha mitogen-activated protein kinase: predictive modeling by machine learning, docking study and molecular dynamics simulation. *Heliyon* **2022**, 8 (9), e10495.
 28. T. N. , M. M.; K. , S.; Asiri, A. M.; Sobahi, T. R.; Asad, M., Green synthesis of chromonyl chalcone and pyrazoline as potential antimicrobial agents – DFT, molecular docking and antimicrobial studies. *Journal of Molecular Structure* **2023**, 1271, 133993.
 29. Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **2010**, 31 (2), 455-61.
 30. Wang, F.; Yang, W.; Zhou, B., Studies on the antibacterial activities and molecular mechanism of GyrB inhibitors by 3D-QSAR, molecular docking and molecular dynamics simulation. *Arabian Journal of Chemistry* **2022**, 15 (6), 103872.
 31. Mishra, D.; Maurya, R. R.; Kumar, K.; Munjal, N. S.; Bahadur, V.; Sharma, S.; Singh, P.; Bahadur, I., Structurally modified compounds of hydroxychloroquine, remdesivir and tetrahydrocannabinol against main protease of SARS-CoV-2, a possible hope for COVID-19: Docking and molecular dynamics simulation studies. *J Mol Liq* **2021**, 335, 116185.
 32. Li, L.; Peng, C. e.; Wang, Y.; Xiong, C.; Liu, Y.; Wu, C.; Wang, J., Identify promising IKK- β inhibitors: A docking-based 3D-QSAR study combining molecular design and molecular dynamics simulation. *Arabian Journal of Chemistry* **2022**, 15 (5), 103786.
 33. Jana, S.; Dalapati, S.; Ghosh, S.; Guchhait, N., Binding interaction between plasma protein bovine serum albumin and flexible charge transfer fluorophore: A spectroscopic study in combination with molecular docking and molecular dynamics simulation. *Journal of Photochemistry and Photobiology A: Chemistry* **2012**, 231 (1), 19-27.
 34. Yuan, S.; Chan, H. C. S.; Hu, Z., Using PyMOL as a platform for computational drug design. *WIREs Computational Molecular Science* **2017**, 7 (2), e1298.
 35. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **2004**, 25 (13), 1605-12.
 36. Gao, W.; Ma, X.; Yang, H.; Luan, Y.; Ai, H., Molecular engineering and activity improvement of acetylcholinesterase inhibitors: Insights from 3D-QSAR, docking, and molecular dynamics simulation studies. *J Mol Graph Model* **2022**, 116, 108239.
 37. Kasralikar, H. M.; Jadhavar, S. C.; Goswami, S. V.; Kaminwar, N. S.; Bhusare, S. R., Design, synthesis and molecular docking of pyrazolo [3,4d] thiazole hybrids as potential anti-HIV-1 NNRT inhibitors. *Bioorg Chem* **2019**, 86, 437-444.
 38. Masand, V.; Mahajan, D.; Maldhure, A.; Rastija, V., Quantitative structure-activity relationships (QSARs) and pharmacophore modeling for human African trypanosomiasis (HAT) activity of pyridyl benzamides and 3-(oxazolo[4,5-b]pyridin-2-yl)anilides. *Medicinal Chemistry Research* **2016**, 25.
 39. Pingaew, R.; Prachayasittikul, V.; Worachartcheewan, A.; Thongnum, A.; Prachayasittikul, S.; Ruchirawat, S.; Prachayasittikul, V., Anticancer activity and QSAR study of sulfur-containing thiourea and sulfonamide derivatives. *Heliyon* **2022**, 8 (8), e10067.

-
40. Puzyn, T.; Leszczynski, J.; Cronin, M. T., Recent advances in QSAR studies: methods and applications. **2010**.
 41. Hollas, B., An Analysis of the Autocorrelation Descriptor for Molecules. *Journal of Mathematical Chemistry* **2003**, *33*, 91-101.
 42. Danishuddin; Khan, A. U., Descriptors and their selection methods in QSAR analysis: paradigm for drug design. *Drug Discovery Today* **2016**, *21* (8), 1291-1302.