Med Chem Res (2014) 23:3705–3713
DOI 10.1007/s00044-014-0955-7

MEDICINAL CHEMISTRY RESEARCH

ORIGINAL RESEARCH

Exploring structural requirement, pharmacophore modeling, and de novo design of LRRK2 inhibitors using homology modeling approach

Sagar S. Bhayye · Kunal Roy · Achintya Saha

Received: 6 September 2013 / Accepted: 11 February 2014 / Published online: 26 February 2014

Abstract A mutation in the gene, encoding leucine rich repeat kinase 2 (LRRK2), is a genetic cause of Parkinson's disease (PD). LRRK2 is a dimeric multidomain protein, largely regulates guanosine triphosphate (GTP). G2019S and I2020T, the mutation encodes in the kinase domain of LRRK2 increase the GTPase activity, are the important regulators in pathogenesis of PD. To design potent LRRK2 inhibitors, pharmacophore modeling approach was employed with a wide chemical diversity of compound's database. The best hypothesis consists of hydrogen-bond acceptor and donor as well as hydrophobic aliphatic and ring aromatic features. The model was validated by the test and decoy sets followed by Fischer's randomization test. The validated model was used to screen the database of compounds, which were designed through de novo approach. Homology model of the kinase domain of LRRK2 was built initially using the crystal structure of Janus kinase 3. The designed molecules were further screened for ADMET properties, and ligand–receptor interaction of top hits was analyzed by molecular docking studies to explore potent LRRK2 inhibitors.

Keywords Pharmacophore · De novo drug design · Homology modeling · Leucine rich repeat kinase 2 · Parkinson's disease

Introduction

Parkinson's disease (PD) is an age-related neurodegenerative disorder (Dexter and Jenner, 2013). Missense mutation in LRRK2 genes is the most common genetic cause of PD. The LRRK2 belongs to the ROCO protein family (Lewis, 2009). It is a large dimeric protein, consisting of 2527 residues and 6 different independent domains including a central catalytic domain that consists of GTPase (ROC) and kinase domain (MAPKKK) (Fig. 1). It is supposed to be surrounded by a series of potential protein–protein interacting domains. GTP binding and hydrolysis occur in ROC domain, and phosphorylation takes place in the kinase domain. Amino acids comprising 1879–2138 residues are functioning as kinase domain, classified under tyrosine kinase subfamily. The mutation occurs in ROC and COR domains decrease the GTPase activity. G2019S and I2020T mutations, which are pathogenic in nature, increase the kinase activity, while mutation in WD40 domain does not affect the kinase activity (Gloeckner *et al.*, 2006). G2019S mutation is not only increasing the kinase activity, but also decreasing the binding capacity of natural substrate, such as ATP (Anand *et al.*, 2009). Additionally, A2016T mutation is active but occurs less frequently in nature. The protein acts as an integrator of multiple signaling pathways that are crucial for proper neuronal functioning. Both protein interaction domains (LRR and WD40) as well as enzymatic domains (ROC and MAPKKK) within LRRK2 serve as a scaffold for assembling of a multiprotein signaling complex.

Electronic supplementary material The online version of this article (doi:10.1007/s00044-014-0955-7) contains supplementary material, which is available to authorized users.

S. S. Bhayye · A. Saha (✉)
Department of Chemical Technology, University of Calcutta, 92 A.P.C Road, Kolkata 700 009, West Bengal, India
e-mail: achintya_saha@yahoo.com

K. Roy
Department of Pharmaceutical Technology, Jadavpur University, Jadavpur, Kolkata 700 032, West Bengal, India

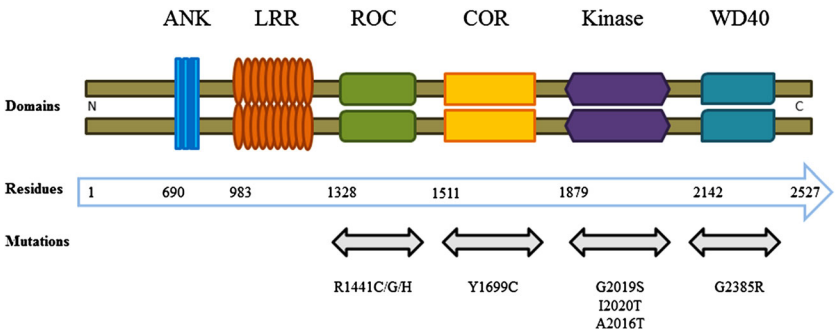

Fig. 1 Domains and mutations of LRRK2. LRRK2 shown as dimer with head to head orientation. Six domains—*ANK* ankyrin-like repeats, *LRR* leucine rich repeat, *ROC* Ras of complex proteins, *COR* carboxy terminal of ROC, kinase and WD40, a β-propeller like-domain—made of WD40 repeats. Starting residue of particular domain shown below the domains. Mutations which occurred in the respective domain shown below residues. Mutation G2019S in the kinase domain increases the kinase activity, which causes pathogenesis

LRRK2 is widely expressed in many body tissues. Under normal body condition, it is predominantly present in neuronal cells. Human genetic, animal model, and biochemical data suggest that there is an association between LRRK2, α-synuclein, and tau (Cookson, 2010; Dachsel and Farrer, 2010; Greggio and Cookson, 2009; Tsika and Moore, 2012; Gandhi *et al.*, 2009). α-Synuclein and tau also aggregate in the neurodegenerative disorders.

As an increase in kinase activity leads to cytotoxicity, inhibiting the enzyme activity might be therapeutically beneficial strategy. This can be accomplished by designing ATP competitive LRRK2 kinase inhibitor (Rudenko *et al.*, 2012). Compounds **LRRK2-IN1**, **CZC-25146**, **TEA684**, **GSK2578215A,** and **HG-10-102-01** are reported as potent inhibitors of the LRRK2 kinase domain (supplementary file, Fig. S1) (Reith *et al.*, 2012; Zhang *et al.*, 2012; Deng *et al.*, 2011; Estrada *et al.*, 2012; Choi *et al.*, 2012; Lee *et al.*, 2012). All of them showed promising activity in both wild and mutated (A2016T and G2019S) types of LRRK2. Different structural features of these compounds are responsible to impart inhibitory activity by occupying receptor cavity (Supplementary file, Fig. S2). Although **LRRK2-IN1** and **CZC-25146** have high selectivity toward LRRK2, their ability to penetrate the blood–brain barrier (BBB) is limited. **TEA684** shows selectivity against mutant A2016T, as it avoids steric clashes with A2016T residue. The compounds **LRRK2-IN1**, **CZC-25146**, and **TEA684** contain aminopyrimidine scaffold, while **GSK2578215A** is a benzamide derivative. Selectivity among kinome and BBB penetration are two major challenges to design an ATP competitive LRRK2 inhibitor. In this study, emphasis is given to explore the structural requirement of LRRK2 inhibitor using ligand-based approach, and to design potent inhibitors through de novo approach and homology modeling of the kinase domain of LRRK2 to possess sufficient selectivity as well as BBB penetration property.

Materials and methods

Pharmacophore modeling

The dataset consisting of 736 LRRK2 inhibitors (Supplementary file, Table S1) based on the biological assay method (Baker-Glenn *et al.*, 2011a, b; Chan *et al.*, 2011, 2012) was used for the generation of the pharmacophore models. The inhibitory activity (IC_{50} value) of the molecules is spanned across the range of 0.3–858 nM. During pharmacophore model generation, selection of a suitable training set is an important step to assist in determining the quality of the generated model. A set of 28 compounds was selected based on the principles of structural diversity and activity range. The remaining 708 compounds of the original dataset were used as a test set for validation of the generated pharmacophore model.

For evaluation purposes, the activity values were classified on the scale of highly active (+++, $IC_{50} < 50$ nM), moderately active (++, 300 nM $< IC_{50} >$ 50 nM), and least active (+, $IC_{50} > 300$ nM). Numerous conformations of all the inhibitors were generated using the BEST method, which provides a complete and improved coverage of the conformational space by performing a rigorous energy minimization, and optimizing the conformations in both torsional and cartesian space by the polling algorithm (Smellie *et al.*, 1995). During this process, the maximum number of conformers was set to 255 with 4.0 kcal/mol as the energy cutoff, and all other parameters kept default (Dhoke *et al.*, 2012). Pharmacophore models were generated by the hypogen algorithm implemented in Accelrys Discovery Studio (DS2.5) (Discovery studio 2.5, 2009).

The structural features considered for model generation include hydrogen-bond acceptor (HBA), hydrogen-bond donor (HBD), hydrophobic aliphatic (HYA), hydrophobic (HY), and ring aromatic (RAr). The minimum and maximum numbers of features to be included during

pharmacophore generation were set to 0 and 5, respectively. The uncertainty value was set to 2. Activity data rescaled to 4 orders of magnitude as difference between most and least active compounds is less than 4 orders of magnitude on logarithmic scale (Kristam *et al.*, 2005). The quality of the generated pharmacophore hypothesis can be judged in terms of several statistical parameters, such as fixed cost, null cost, and total cost. Additionally, three other cost values that play a vital role in determining the quality of a generated hypothesis are weight cost, configuration cost, and error cost. The best model was selected based on high-correlation coefficient (r), lowest total cost, highest cost difference, and lowest root mean squared deviation (RMSD) values (Singh *et al.*, 2013a). The top ten pharmacophore hypotheses with significant statistical parameters were generated by the 3D QSAR Pharmacophore Generation module in the DS2.5.

The best pharmacophore hypothesis was validated to determine its capability for differentiating between active, moderate, and least active compounds, and for predicting their activities accurately. Three different methods (Fischer randomization test, test set prediction, and decoy test) were employed for validation of the developed pharmacophore hypothesis. A Fischer randomization (Cat-Scramble) test was utilized with the goal to check whether there is a substantial correlation between the chemical structural features and the biological activity of the training set compounds. During this validation, 19 random pharmacophore hypotheses were generated to achieve 95 % confidence level. The test set compounds with a wide range of inhibitory activity were used to validate the best hypogen hypothesis.

On further validation of the pharmacophore hypothesis, decoy set was generated by DecoyFinder1.1 (Cereto-Massagué *et al.*, 2012). Decoys are resembled to active ligands physically, while chemically distinct to avoid bias in the enrichment factor calculation. Decoys were selected based on the five physicochemical descriptors (molecular weight, number of rotational bonds, hydrogen-bond donor count, hydrogen-bond acceptor count, and the octanol–water partition coefficient) of the active ligands. MACCS fingerprints were calculated according to the maximum Tanimoto coefficient values to discriminate decoys and active ligands chemically. In the decoy set, 15 LRRK2 inhibitors from test set were included to calculate various statistical parameters, such as accuracy, precision, sensitivity, specificity, goodness of hit score (GH), and enrichment factor (E value). Decoys molecules are supposed to be inactive against a target and used to validate the performance of the virtual screening workflow. The screening of databases of test and decoy sets was performed using the hypothesis 1 as a 3D structural query. The accuracy, precision, sensitivity, and specificity of the best pharmacophore model were estimated. The two major parameters, GH and E value play important role for judging capability of the generated pharmacophore hypothesis (Singh *et al.*, 2013b), are calculated as per the Eqs. 1 and 2, respectively.

$$\mathrm{GH} = (\mathrm{TP}/4\mathrm{Ht}\ A)(3A + \mathrm{Ht}) \times (1 - ((\mathrm{Ht} - \mathrm{TP})/(D - A))) \quad (1)$$

$$E = (\mathrm{TP} \times D)/(\mathrm{Ht} \times A) \quad (2)$$

where D, A, Ht, and TP represent the total number of compounds of the database, total number of actives, total number of compounds screened by a pharmacophore model, and total number of active compounds screened, respectively. To calculate the E value and GH score of the test set, highly and moderately active compounds (IC_{50} <300 nM) were consider as active compounds.

Homology modeling

Crystal structure of the kinase domain of LRRK2 is not reported yet; therefore, building of a homology model using related crystal structure as a template is initiated. Choice of template varies from structural closeness, functionality, identity, and quality of crystal structure available. On PSI-BLAST search, none of the kinase showed sequence identity more than JAK3. In this study, JAK3 (PDB ID: 3LXL) was selected as a template to perform a homology modeling, because it shares a similar inhibitory profile as LRRK2 inhibitors (Chrencik *et al.*, 2010; Chen *et al.*, 2012). The selected template crystal structure has a resolution of 1.74 Å. Secondary structure prediction of target sequence was done using three different online servers, Jpred3 (http://www.compbio.dundee.ac.uk/www-jpred/), PROF (http://www.aber.ac.uk/~phiwww/prof/), and PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) (Cole *et al.*, 2008). Initially, pairwise sequence alignment of the template and target sequence was done in ClustalW and align 2d python script of Modeller 9v11 (Larkin *et al.*, 2007; Eswar *et al.*, 2008). The alignment was manually edited on the basis of secondary structure prediction, considering that active site residues were properly aligned and gaps present in loop region. Modeller 9v11 was used to build model of LRRK2 kinase domain using sequence alignment. The models were built with ligand modeling by changing default optimization and refinement protocol in molecular dynamic of simulated annealing technique. In the process, ligand in the template transferred to model keeping orientation and restrained identical. The best model was selected on the basis of molpdf and dope score, and further subjected to loop refinement using python script in Modeller. After subsequent steps of loop refinement, final model was validated using Ramchandran and Errat plots (Laskowski *et al.*, 1993; Colovos and Yeates,

1993; Ramachandran *et al.*, 1963), which were assessed through PROCHECK program of SAVES online server (http://nihser-ver.mbi.ucla.edu/SAVES). Active site was validated by docking ATP and reference compounds by using Glide 5.5, and visual analysis was done through PyMoL and VMD 1.9 (PyMOL 1.3, 2010; Humphrey *et al.*, 1996).

Structure-based drug design

A validated model of the kinase domain was used for de novo design of ATP competitive inhibitor of LRRK2 using Ligbuilder 2.0 (Yuan *et al.*, 2011; Kare *et al.*, 2013). It follows the genetic algorithm to design ligand site iteratively, based on active site cavity of a target protein by using organic fragments. The growing and linking strategies for constructions of ligands were processed. An empirical function was used to evaluate binding affinities between ligands and receptor. Design of ligands involved the use of POCKET, GROW, and PROCESS modules step by step. Pocket module was used for analyzing binding site and generating grid with pharmacophoric features essential to build ligand. Information of grid file was used in the grow step. Two seed structures (Fig. S3 of supplementary material) were provided to GROW module on which new ligands were grown using genetic algorithm. First seed pyrrolopyridine ring, which can occupy a hinge region of receptor pocket attached to pyridine ring, acts as aromatic linker. Aminopyromidine class possesses sufficient brain penetration ability was used as second seed (Deng *et al.*, 2012). All the hydrogens were selected as growing site. Seven fragments were extracted from these seeds (supplementary file, Fig. S4). A fragment library containing ring structures, such as benzene, pyrimidine, indole, amino pyrimidine, etc., was developed. Molecular input parameters selected for growing ligands were molecular weight: 160–500, Hbond donor: 1–5, Hbond acceptor: 2–10, log*P*: 2–5, and PKD: 6–12 to follow the Lipinski's rule of 5. Designed ligands from the ligand collection file were explored by PROCESS module to extract desired molecules. A total number of 495 molecules were obtained, which were energy minimized in SYBYL (SYBYL 7.1 2005).

Pharmacophore-based virtual screening

Pharmacophore-based virtual screening was performed with the database of compounds, obtained through de novo design. After applying the Lipinski's filter, newly designed compounds were mined for the virtual screening process, which helps in identification of novel and potential hits suitable for further development. The validated pharmacophore hypothesis (Hypo1) as a 3D query with the Best/Flexible search option was used for virtual screening of compounds.

The maximum omitted features were set to "−1." Hit compounds were screened for their predicted biological activity values. Hits, obtained from the pharmacophore screening (estimated IC_{50} ≤50 nM), were selected and subsequently submitted for evaluation of drug-likeness properties, such as MW, QPlogPo/w, QPlogHERG (potential hERG channel blockage leads to QT syndrome), QPlogBB (the ability to cross the BBB), #metab (predicting the number of metabolic reactions likely to occur), and human oral absorption. Further, the toxicity profile of designed compounds was predicted using Toxicity Prediction through Komputer Assisted Technology (TOPKAT) (Enslein *et al.*, 1994), which measured the molecular-based potential toxicity end points. The optimum prediction space (OPS) of the query molecules was also checked. The predicted toxicity value considered acceptable only if the query molecule is within the OPS limit, otherwise rejected. Human Ether-à-go–go-Related Gene (hERG) toxicity of the accepted compounds was predicted using QikProp (QikProp 3.2, 2009). The compounds passed through ADME and toxicity screenings were subjected to molecular docking by Glide (Glide 5.5, 2009).

Molecular docking

To investigate the binding affinity and interaction pattern with residues, molecular docking study was carried out on the kinase domain of LRRK2, which developed through homology modeling. After screening for ADMET, the designed molecules were prepared for docking in LigPrep (LigPrep 2.3, 2009), using OPLS2005 force field. Grid size of 20 Å was generated keeping ATP at the center. GlideSP mode was used for docking purpose. The binding affinity of potent known inhibitors was compared with that of the screened hits and ranked the molecules on the basis of interactions with amino acid residues at the active site. Homology models of the kinase domain of wild and mutated (G2019S, I2020T, and A2016T) LRRK2 were processed using a protein preparation module of Schrödinger (Protein Preparation Wizard, 2009). Docking studies of final hits were carried out on wild as well as the mutated kinase domain of LRRK2 for checking selectivity.

Results and discussion

Pharmacophore modeling

In ligand-based pharmacophore modeling studies, top ten hypotheses were generated using diverse training set

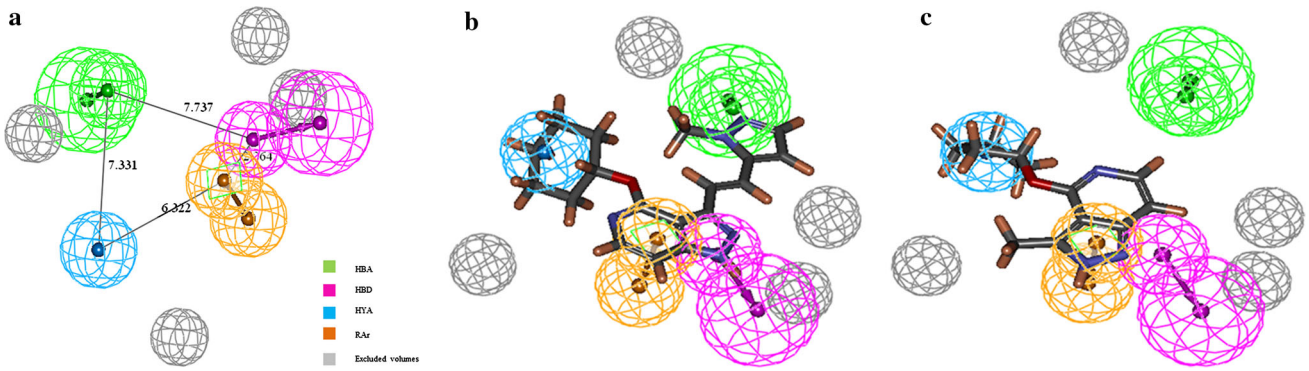

Fig. 2 **a** Chemical features of best pharmacophore hypothesis (Hypo1) with their inter-feature distance constraints in Å. **b** Alignment of most active (compound **1**) and **c** Alignment of least active (compound **26**) LRRK2 inhibitors in Hypo1

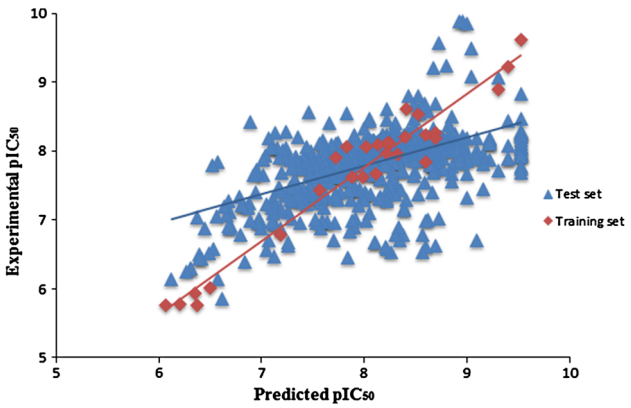

Fig. 3 Scatter plot of predicted versus experimental pIC_{50} of LRRK2 inhibitors

compounds. The results obtained during the pharmacophore hypothesis generation are summarized in supplementary file (Table S2). The best pharmacophore hypothesis (Hypo1) is characterized by the best correlation coefficient, lowest root mean square error, highest cost difference, and consisted of spatial arrangement of the four chemical features (HBA, HBD, HYA, and RAr) and four excluded volumes (Fig. 2a). Among the 28 training set compounds, all the active, moderately and least active compounds predicted accurately (Table S1) in the model. Interestingly, most of the highly active compounds mapped to all the pharmacophoric features (Fig. 2b). But in case of moderately active and least active compounds, one or two features were missing or mapping partly (Fig. 2c). The estimated versus predicted activity values of the LRRK2 training set and test molecules are depicted in Fig. 3. Fit value is a measure of the overlap between the features in the pharmacophore and chemical features in the molecule, which assists in understanding the chemical meaning of the pharmacophore hypothesis. The most active compound in the training set has shown the fitness score of 8.97. However, the least active compound has shown the fitness score 5.28, when mapped to Hypo1.

Pharmacophore model validation

Several methods and parameters, such as cost analysis, test set prediction, Fischer's randomization method, goodness of fit, and enrichment factor were used to judge the quality of the generated pharmacophore model. Initially, various cost values calculated during the pharmacophore model generation were evaluated. The total cost value for the hypothesis was 111.381, while the null cost value was 259.481. The cost difference between the null cost and the total cost values (Δcost = 148.10) characterizes that the pharmacophore hypothesis can significantly correlate the data by more than 90 %. Pharmacophore hypothesis (Hypo1) has shown the highest correlation coefficient value of 0.963. In addition, the RMSD value of 0.95 and the configuration cost of 13.782 justify the model acceptability. Most of the compounds in the test set were predicted correctly in respect to their experimental biological activity with coefficient of determination (r^2) of 0.60. The test set validation was also assessed by calculating Enrichment factor and GHs (Table 1). In the test set validation, it was found that 583 of 627 highly active, 32 of 71 moderately active, and 8 of 10 least active compounds were predicted correctly. The 44 active compounds were underestimated as moderately active, the 29 moderately active compounds were overestimated as actives, and 3 were underestimated as least actives, and the 2 least active compounds were overestimated as moderately active. Fischer randomization test at 95 % confidence level showed (supplementary file, Fig. S5) that none of the randomly generated pharmacophore hypotheses has scored better statistical results than Hypo1, indicating that the model has not been generated by any chance correlation.

A small database containing 390 compounds including 15 actives and 375 decoys was generated using

Table 1 The statistical parameters obtained from Test set and Decoy test

Sr. no.	Parameters	Test set	Decoy set
1	Total compounds in database (*D*)	708	390
2	Total number of actives in database (*A*)	698	15
3	Total hits (Ht)	695	15
4	Active hits (TP)	693	15
5	True negative (TN)	8	375
6	Enrichment factor or enhancement (*E*)	1.01	26
7	False negatives (FN = *A* − TP)	5	0
8	False positives (FP = Ht − TP)	2	0
9	GH score (goodness of hit list)	0.79	1
10	Accuracy = (TP + TN)/(TP + TN + FP + FN)	0.99	1
11	Precision = TP/(TP + FP)	0.99	1
12	Sensitivity = TP/(TP + FN)	0.99	1
13	Specificity = TN/(TN + FP)	0.80	1

DecoyFinder1.1. The database was used to verify the capability of hypo1 model for discriminating the active and inactive compounds. Hypo1 has shown an *E* value of 1.01 and 26 for test set and decoy set, respectively. The calculated GH score for both test and decoy sets is greater than 0.5, which indicated that the developed pharmacophore model is superior than any other hypothesis. From the overall validation results, it assured that the model (Hypo1) is able to discriminate between the active and decoys compounds.

Homology modeling

PSI BLAST results shown 44 % identities, 42 % positives, and 8 % gaps in template and target sequence. Alignment between template and target conserved residues was done correctly (supplementary file, Fig. S6). Topology of the model was checked in different protein prediction servers (Jpred3, PROF & PSIPRED) while building the model. Stereochemistry of modeled protein was assessed by Ramchandran plot. The Ramchandran and ERRAT plot of the model are provided in supplementary file (Figs. S7, S8; Table S3). Statistical parameters indicate that 99.1 % residues are present in the core and additional allowed region of Ramchandran plot, while 96.74 % residues are below error value in Errat plot. The RMSD and alignment score of 1.52 and 0.09, respectively, between the template and the model are obtained through Protein Structure Alignment tool of Schrödinger suite. The model is well aligned with the template (supplementary file, Fig. S9). Docked pose for ATP in the active site of kinase domain (supplementary file, Fig. S10a) shows hydrogen-bond interactions of ATP with residues LYS 1906, GLU 1948, ALA 1950, and ASN 1999. Catalytic residues, GLU 1948, and ALA 1950 form H-bond with heterocyclic adenine ring, while LYS 1906 and ASN 1999 with oxygen of phosphate group. The docked ATP in the model was further aligned with crystal structure of the ROCO4 kinase (PDB ID: 4FOF) (Gilsbach *et al.*, 2012), which bound ligand phosphomethylphosphonic acid adenylate ester to check the right conformation and sharing structural similarity with ATP. The superimposed active site residues of ROCO4 and modeled kinase domain confirm that docked ATP is not only showing interaction with key residues of the kinase domain but also occupying the active site in right conformation (Supplementary file, Fig. S10b). The model suggests that the ATP binding site is present between the interfaces of two lobes. The α-helix is present in both N- and C-terminals, whereas β-sheet only present in the N-terminal of the kinase domain. The ATP binding pocket is partially closed by residues LYS 1885 and ARG 1957, covering sugar moiety of ATP (supplementary file, Fig. S10c). Pocket remains open toward the adenine ring and the phosphate moiety occupied region. Residues GLU 1948 and ALA 1950 form the hinge region to behave either hydrogen-bond donor or acceptor. Amino acids, LYS 1906, and GLU 1920 are present in the back of the side wall of the pocket.

The docking scores of ATP and other reported compounds with their model energy (Supplementary file, Table S4) proved that the G2019S mutation affects binding of compounds, indicating by decrease in G-score and E-model values of ATP and reference compounds. Decrease in binding affinity toward G20119S mutant may be due to the side chains of serine, which creates slight steric hindrance for ligands (supplementary file, Fig. S11). Although this hindrance is too small to affect activity, it is good enough to reduce the concrete binding of LRRK2 inhibitors.

Structure-based drug design

Initial grid file and pharmacophore file were generated from cavity module. The key site features, generated from pocket structure-based pharmacophore model (Supplementary file, Fig. S12), consist of 5 hydrogen-bond donor (blue), 3 hydrogen-bond acceptor, and 2 hydrophobic sites (green). Total 495 molecules were obtained through the GROW module after subjecting to PROCESS module.

Pharmacophore model-based virtual screening

Initially, the validated pharmacophore model (Hypo1) of LRRK2 inhibitors was used as a query tool for screening database of newly designed compounds, which were prefiltered by the Lipinski's rule of five. The model retrieved

368 hit compounds of which some compounds are structurally similar to that of the existing LRRK2 inhibitors. A set of 89 hit compounds shown estimated IC_{50} value below 50 nM. These selected screened hits were subsequently submitted for ADMET analysis. Molecules were screened for properties, such as QPlogPo/w, QPlogHERG, QPlogBB, #metab, human oral absorption, and Jorgensen's rule of three that includes aqueous solubility >−5, Caco-2 cell permeability >22 nm/s, and primary metabolites less than 7. Compounds that possess less violation of these properties are considered to be effective on oral absorption (QikProp 3.2, 2009). The calculated pharmaceutically important properties and structural information are provided in the supplementary file (Table S5). All reported compounds were violating the QPlogHERG properties. The BBB penetration ability was found to be low except compound **HG-10-102-01**. Designed molecules are within the recommended range and showed drug-likeness properties. While assessing toxicity in TOPKAT, two filters were applied: (i) molecular carcinogenicity or mutagenicity, and (ii) molecular OPS limit or applicability domain. Total 12 molecules were obtained after ADMET screening.

Molecular docking

Molecular docking studies showed that most of the designed molecules have good affinity toward LRRK2 in terms of docking score compared to standard compounds (Table 2). Top hits have shown a comparatively less affinity toward G2019S mutant among wild and other mutant (A2016T and I2020T) types. All the molecules have shown H-bond interactions with GLU 1948, ALA 1950, and ASP 1999. Docked poses of compound **CT_5** in wild and mutated (A2016T) proteins are shown in Fig. 4a and b respectively. There is no steric clash between the side chain of mutated THR 2016 with **CT_5**. All the hits have

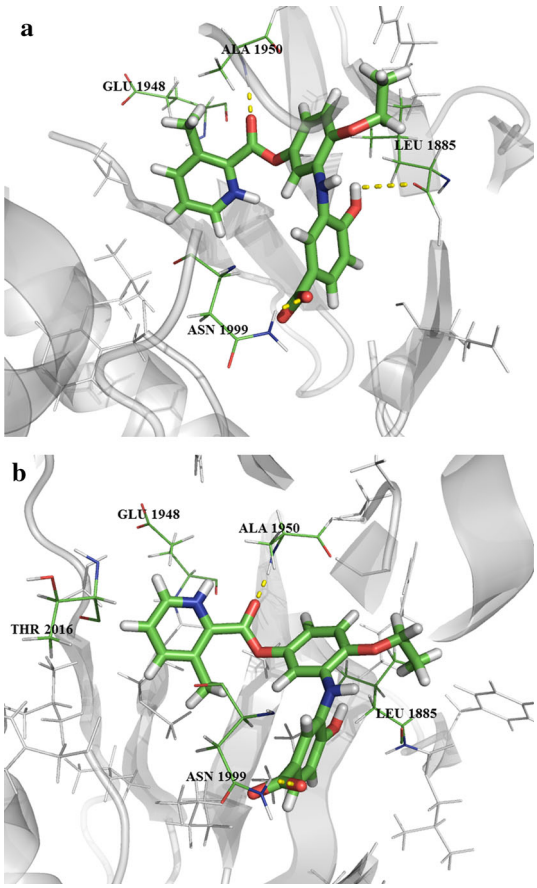

Fig. 4 The docking pose of designed molecule **a CT_5** in wild type LRRK2, **b CT_5** in A2016T mutated LRRK2. Side chain THR 2016 not clash with **CT_5**. H-bonding interaction shown in *dotted lines*

Table 2 Docking score, fit value, predicted LRRK2 (Wild) inhibitors activity of top virtual hits

Compound	Docking score				Hypo1 fit value	Estimated IC_{50} (nM) (wild type)
	Wild	I2020T	G2019S	A2016T		
CT_1	−7.04	−7.69	−5.96	−7.65	8.91	6.83
CT_2	−3.73	−3.32	−4.42	−4.03	8.03	9.88
CT_3	−9.57	−9.22	−6.86	−8.71	8.25	6.03
CT_4	−8.87	−9.36	−6.60	−8.78	7.60	26.64
CT_5	−9.33	−9.46	−6.41	−9.49	7.57	28.42
CT_6	−7.35	−5.75	−7.63	−6.93	7.45	37.91
CT_7	−6.96	−6.10	−3.14	−6.48	7.92	12.72
CT_8	−6.67	−6.26	−8.07	−5.91	7.34	49.10
CT_9	−7.53	−7.24	−3.92	−6.92	7.60	26.79
CT_10	−6.06	−5.84	−6.22	−6.27	8.41	4.11
CT_11	−7.31	−7.87	−5.10	−7.72	7.75	18.81
CT_12	−5.48	−6.00	−5.83	−6.96	7.60	26.89

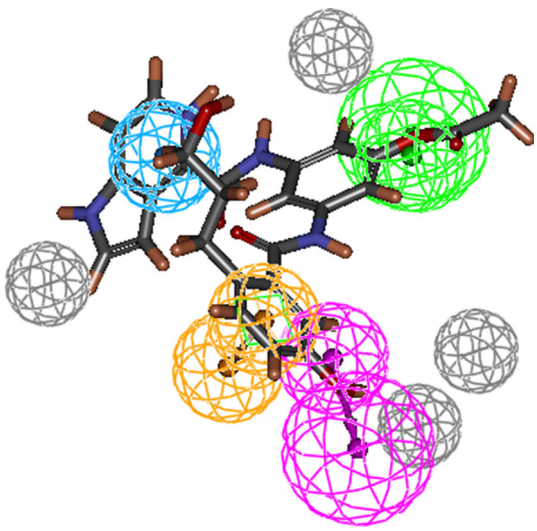

Fig. 5 Alignment of **CT_8** with best pharmacophore hypothesis (Hypo1)

shown similar binding orientation as that of the known LRRK2 inhibitors in docking studies. Compound **CT_2** shows docking score >−5 in mutated as well as the wild forms of LRRK2 protein, suggesting that molecule do not have a sufficient affinity toward LRRK2 in terms of dock score. The compound **CT_8** is perfectly mapped in pharmacophore model (Fig. 5), indicating potent inhibitory activity for LRRK2.

Conclusions

A mutation in LRRK2 leads to increase in kinase activity which is a genetic cause of PD. In an effort to design ATP competitive inhibitor using ligand-based pharmacophore modeling, structure-based drug design and homology modeling approaches were used. Structural studies of LRRK2 inhibitors shown that inhibitors must contain small hydrophobic moiety to occupy a distinct region near to hinge region, heterocyclic ring to exchange H-bond with GLU 1948 and ALA 1950 the aromatic linker and electronegative H-bond acceptor group attached to aromatic linker. The best pharmacophore model also highlighted the key structural features that include hydrogen-bond acceptor, donor, HYA, and RAr. Considering these features, de novo approach was used to design new molecules. Virtual hits were subjected to pharmacophore mapping and ADMET properties based screening including BBB penetration property. Molecules which possess drug-likeness were studied for affinity and interaction with the ATP binding site of kinase domain of LRRK2, which was developed through homology modeling. These virtual hits might prove promising lead molecules to be tested as potential and selective LRRK2 inhibitors.

Acknowledgments The authors would like to thank the University Grant Commission (UGC), New Delhi, India for financial assistance of the project.

References

Anand VS, Reichling LJ, Lipinski K, Stochaj W, Duan W, Kelleher K, Pungaliya P, Brown EL, Reinhart PH, Somberg R (2009) Investigation of leucine rich repeat kinase 2. FEBS J 276(2):466–478. doi:10.1111/j.1742-4658.2008.06789.x

Baker-Glenn C, Burdick DJ, Chambers M, Chan BK, Chen H, Estrada A, Gunzner JL, Shore D, Sweeney ZK, Wang S (2011a) Aminopyrimidine derivatives as LRRK2 modulators. WO Patent App. PCT/EP2011/059009

Baker-Glenn C, Burdick DJ, Chambers M, Chan BK, Chen H, Estrada A, Sweeney ZK (2011b) Pyrazole aminopyrimidine derivatives as LRRK2 modulators. WO Patent App. PCT/EP2011/069696

Cereto-Massagué A, Guasch L, Valls C, Mulero M, Pujadas G, Garcia-Vallvé S (2012) DecoyFinder: an easy-to-use python GUI application for building target-specific decoy sets. Bioinformatics 28(12):1661–1662. doi:10.1093/bioinformatics/bts249

Chan B, Estrada A, Sweeney Z, Mciver EG (2011) Pyrazolopyridines as inhibitors of the kinase LRRK2. WO Patent App. PCT/GB2011/050937

Chan BK, Estrada AA, Chen H, Atherall J, Baker-Glenn C, Beresford A, Burdick DJ, Chambers M, Dominguez SL, Drummond J (2012) Discovery of a highly selective, brain-penetrant aminopyrazole LRRK2 inhibitor. ACS Med Chem Lett 4(1):85–90. doi:10.1021/ml3003007

Chen H, Chan BK, Drummond J, Estrada AA, Gunzner Toste J, Liu X, Liu Y, Moffat J, Shore D, Sweeney ZK (2012) Discovery of selective LRRK2 inhibitors guided by computational analysis and molecular modeling. J Med Chem 55(11):5536–5545. doi:10.1021/jm300452p

Choi HG, Zhang J, Deng X, Hatcher JM, Patricelli MP, Zhao Z, Alessi DR, Gray NS (2012) Brain penetrant LRRK2 inhibitor. ACS Med Chem Lett 3(8):658–662. doi:10.1021/ml300123a

Chrencik JE, Patny A, Leung IK, Korniski B, Emmons TL, Hall T, Weinberg RA, Gormley JA, Williams JM, Day JE (2010) Structural and thermodynamic characterization of the TYK2 and JAK3 kinase domains in complex with CP-690550 and CMP-6. J Mol Biol 400(3):413–433. doi:10.1016/j.jmb.2010.05.020

Cole C, Barber JD, Barton GJ (2008) The Jpred 3 secondary structure prediction server. Nucleic Acids Res 36(suppl 2):W197–W201. doi:10.1093/nar/gkn238

Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 2(9):1511–1519. doi:10.1002/pro.5560020916

Cookson MR (2010) The role of leucine-rich repeat kinase 2 (LRRK2) in Parkinson's disease. Nat Rev Neurosci 11(12):791–797. doi:10.1038/nrn2935

Dachsel JC, Farrer MJ (2010) LRRK2 and Parkinson disease. Arch Neurol 67(5):542. doi:10.1001/archneurol.2010.79

Deng X, Dzamko N, Prescott A, Davies P, Liu Q, Yang Q, Lee JD, Patricelli MP, Nomanbhoy TK, Alessi DR (2011) Characterization of a selective inhibitor of the Parkinson's disease kinase LRRK2. Nat Chem Biol 7(4):203–205. doi:10.1038/nchembio.538

Deng X, Choi HG, Buhrlage SJ, Gray NS (2012) Leucine-rich repeat kinase 2 inhibitors: a patent review (2006–2011). Expert Opin Ther Pat 22(12):1415–1426. doi:10.1517/13543776.2012.729041

Dexter DT, Jenner P (2013) Parkinson's disease: from pathology to molecular disease mechanisms. Free Radic Biol Med. 62:132–144. doi:10.1016/j.freeradbiomed.2013.01.018

Dhoke GV, Gangwal RP, Sangamwar AT (2012) A combined ligand and structure based approach to design potent PPAR-alpha agonists. J Mol Struct 1028:22–30. doi:10.1016/j.molstruc.2012.06.032

Discovery studio 2.5 (2009) Accelrys Inc., San Diego

Enslein K, Gombar VK, Blake BW (1994) Use of SAR in computer-assisted prediction of carcinogenicity and mutagenicity of chemicals by the TOPKAT program. Mutat Res 305(1):47–61. doi:10.1016/0027-5107(94)90125-2

Estrada AA, Liu X, Baker-Glenn C, Beresford A, Burdick DJ, Chambers M, Chan BK, Chen H, Ding X, DiPasquale AG (2012) Discovery of highly potent, selective, and brain-penetrable leucine-rich repeat kinase 2 (LRRK2) small molecule inhibitors. J Med Chem 55(22):9416–9433. doi:10.1021/jm301020q

Eswar N, Eramian D, Webb B, Shen MY, Sali A (2008) Protein structure modeling with MODELLER. Structural Proteomics. Springer, London, pp 145–159. doi:10.1007/978-1-60327-058-8_8

Gandhi PN, Chen SG, Wilson-Delfosse AL (2009) Leucine rich repeat kinase 2 (LRRK2): a key player in the pathogenesis of Parkinson's disease. J Neurosci Res 87(6):1283–1295. doi:10.1002/jnr.21949

Gilsbach BK, Ho FY, Vetter IR, van Haastert PJ, Wittinghofer A, Kortholt A (2012) Roco kinase structures give insights into the mechanism of Parkinson disease-related leucine-rich-repeat kinase 2 mutations. Proc Natl Acad Sci USA 109(26):10322–10327. doi:10.1073/pnas.1203223109

Glide 5.5 (2009) Schrödinger. LLC, New York

Gloeckner CJ, Kinkl N, Schumacher A, Braun RJ, O'Neill E, Meitinger T, Kolch W, Prokisch H, Ueffing M (2006) The Parkinson disease causing LRRK2 mutation I2020T is associated with increased kinase activity. Hum Mol Genet 15(2):223–232. doi:10.1093/hmg/ddi439

Greggio E, Cookson MR (2009) Leucine-rich repeat kinase 2 mutations and Parkinson's disease: three questions. ASN Neuro 1(1):13–24. doi:10.1042/AN20090007

Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14(1):33–38. doi:10.1016/0263-7855(96)00018-5

Kare P, Bhat J, Sobhia ME (2013) Structure-based design and analysis of MAO-B inhibitors for Parkinson's disease: using in silico approaches. Mol Divers 17(1):111–122. doi:10.1007/s11030-012-9420-z

Kristam R, Gillet VJ, Lewis RA, Thorner D (2005) Comparison of conformational analysis techniques to generate pharmacophore hypotheses using catalyst. J Chem Inf Model 45(2):461–476. doi:10.1021/ci049731z

Larkin M, Blackshields G, Brown N, Chenna R, McGettigan P, McWilliam H, Valentin F, Wallace I, Wilm A, Lopez R (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23(21):2947–2948. doi:10.1093/bioinformatics/btm404

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26(2):283–291. doi:10.1107/S0021889892009944

Lee BD, Dawson VL, Dawson TM (2012) Leucine-rich repeat kinase 2 (LRRK2) as a potential therapeutic target in Parkinson's disease. Trends Pharmacol Sci 33(7):365–373. doi:10.1016/j.tips.2012.04.001

Lewis PA (2009) The function of ROCO proteins in health and disease. Biol Cell 101(3):183–191. doi:10.1042/BC20080053

LigPrep 2.3 (2009) Schrödinger. LLC, New York

Protein Preparation Wizard (2009) Schrödinger. LLC, New York

PyMOL 1.3 (2010) Schrödinger. LLC, New York

QikProp 3.2 (2009) Schrödinger. LLC, New York

Ramachandran G, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. J Mol Biol 7:95. doi:10.1016/S0022-2836(63)80023-6

Reith AD, Bamborough P, Jandu K, Andreotti D, Mensah L, Dossang P, Choi HG, Deng X, Zhang J, Alessi DR (2012) GSK2578215A; a potent and highly selective 2-arylmethyloxy-5-substituent-*N*-arylbenzamide LRRK2 kinase inhibitor. Bioorg Med Chem Lett 22(17):5625–5629. doi:10.1016/j.bmcl.2012.06.104

Rudenko IN, Chia R, Cookson MR (2012) Is inhibition of kinase activity the only therapeutic strategy for LRRK2-associated Parkinson's disease? BMC Med 10(1):20. doi:10.1186/1741-7015-10-20

Singh R, Balupuri A, Sobhia ME (2013a) Development of 3D-pharmacophore model followed by successive virtual screening, molecular docking and ADME studies for the design of potent CCR2 antagonists for inflammation-driven diseases. Mol Simul 39(1):49–58. doi:10.1080/08927022.2012.701743

Singh U, Gangwal RP, Prajapati R, Dhoke GV, Sangamwar AT (2013b) 3D QSAR pharmacophore-based virtual screening and molecular docking studies to identify novel matrix metalloproteinase 12 inhibitors. Mol Simul 39(5):385–396. doi:10.1080/08927022.2012.731506

Smellie A, Teig SL, Towbin P (1995) Poling: promoting conformational variation. J Comput Chem 16(2):171–187. doi:10.1002/jcc.540160205

SYBYL 7.1 (2005) Tripose Inc., St. Louis

Tsika E, Moore DJ (2012) Mechanisms of LRRK2-mediated neurodegeneration. Curr Neurol Neurosci Rep 12(3):251–260. doi:10.1007/s11910-012-0265-8

Yuan Y, Pei J, Lai L (2011) LigBuilder 2: a practical de novo drug design approach. J Chem Inf Model 51(5):1083–1091. doi:10.1021/ci100350u

Zhang J, Deng X, Choi HG, Alessi DR, Gray NS (2012) Characterization of TAE684 as a potent LRRK2 kinase inhibitor. Bioorg Med Chem Lett 22(5):1864–1869. doi:10.1016/j.bmcl.2012.01.084