



Pharmacophore modelling and molecular dynamics simulation to identify novel molecules targeting catechol-O-methyltransferase and dopamine D3 receptor to combat Parkinson's disease

Amitha Joy¹ · Sidharth Menon^{1,2} · Neethu Mariam Thomas^{1,2} · Meha Christy^{1,2} · Aiswarya D. Menon^{1,2} · Arun John^{1,2}

Received: 9 February 2023 / Revised: 10 September 2023 / Accepted: 11 November 2023 /
Published online: 2 December 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Parkinson's disease is a neurological illness that slowly impairs a small number of neurons in the substantia nigra, a part of the brain. Dopamine, a substance (neurotransmitter) that disseminates signals to different regions of the brain and, when it functions correctly, coordinates smooth and balanced muscular activity, is typically produced by these cells. One-hand tremor is frequently the first sign of Parkinson's disease. Loss of balance, stiffness, and delayed mobility are further symptoms. Proteins including catechol-O-methyltransferase and dopamine D3 receptors were taken into consideration as prospective therapeutic targets in this study. Two ligand-based pharmacophore models were generated with the help of compounds used for Parkinson's disease which have structural similarity, screened from the first 16 compounds found in the drug bank. In the second case, 9 compounds that have similar structure to the compound istradefylline were selected. Based on docking score, intermolecular interactions, ADME (absorption, distribution, metabolism, and excretion) features, pharmacophore, and toxicity investigations, the inhibitors among the chosen compounds were found. Additionally, the chosen inhibitor underwent a 100 nano-second molecular dynamics simulation with the two protein targets to determine its stability and binding affinity. The compound 3,4-Bis(1,3,5,6-heptatetraenyloxy) benzaldehyde was identified to be the most promising lead molecule in this analysis due to its better binding affinity, better pharmacophore properties, and greater stability. Hence, by targeting both specified proteins, the compound 3,4-Bis(1,3,5,6-heptatetraenyloxy) benzaldehyde should be beneficial against Parkinson's disease.

Keywords Parkinson's · Pharmacophore · Docking · Molecular dynamics simulation

Introduction

Parkinson's disease (PD), which affects 1–3% of people over the age of 60 worldwide, is classified as a progressive neurological condition which is fatal [1]. James Parkinson's famous book "Essay on the Shaking Palsy" (AD) published in 1817 defines the essential clinical symptoms of the second most common age-related neurodegenerative disorder after Alzheimer's disease [2]. The symptoms of both the motor and non-motor system are common in this idiopathic disorder of the nervous system [3]. The degeneration of dopaminergic (dopamine-producing) neurons in the substantia nigra, as well as the development of Lewy bodies in dopaminergic neurons, is pathological manifestations of Parkinson's disease (PD) [3]. Pathological alterations may occur approximately two decades or even much before visible symptoms appear. This selective loss of dopamine-producing neurons has a significant impact on the motor function [3]. Lewy bodies, or aberrant intracellular aggregates, include proteins such as alphasynuclein and ubiquitin, which affect neuron function [3]. The mechanism behind the PD is depicted in Fig. 1 [4]. Dopaminergic replacement therapies are the mainstay of PD therapy, which aims to reestablish dopamine functioning. These dopamine-targeted medications only treat the symptoms of PD-related motor impairments; they have no effect on the disease process [5]. These medications also exhibit diminished efficacy over time and have negative side effects such as hallucinations, dyskinesia, and on–off effects [5]. Additionally, the dopaminergic medications that are now on the market have extremely limited effects on the treatment of non-motor symptoms that typically accompany Parkinson's disease (PD), such as mood, postural instability, and cognitive difficulties. Therefore, the

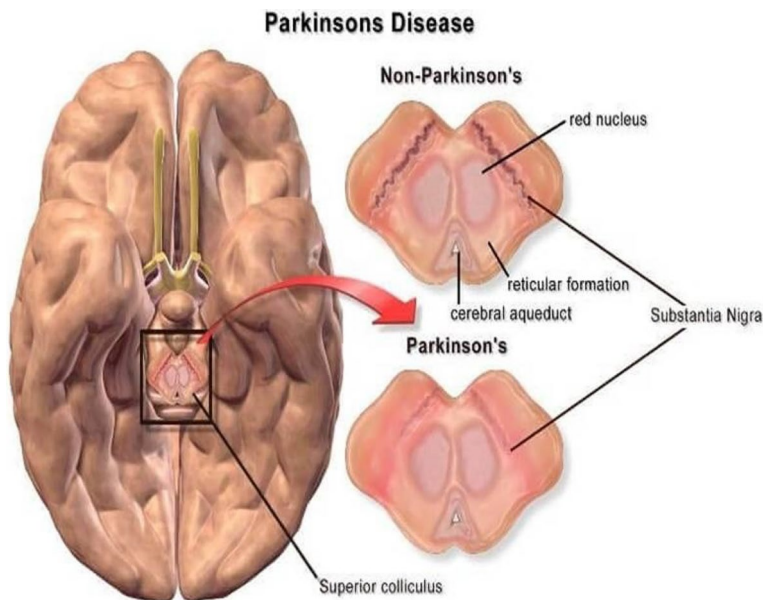


Fig. 1 Brain with and without Parkinson's disease

development of efficient and long-lasting symptomatic and disease-modifying medications is urgently needed for the long-term management of PD [5].

Even though the development of levodopa revolutionized the treatment of Parkinson's disease, it was quickly discovered that after numerous years of therapy, most patients acquire involuntary movements known as "dyskinesias," which are challenging to control and can greatly damage quality of life [2]. The current focus of study is on the mitigation of dopaminergic neuron loss. Despite progress towards this aim, all existing therapies are symptomatic; none stop or slow dopaminergic neuron loss [2]. The greatest impediment to the design of neuroprotective medications is a lack of understanding of the particular biochemical mechanisms that cause neurodegeneration in Parkinson's disease [2].

Given the active role that catechol-O-methyltransferase (COMT) plays in L-DOPA metabolism, both in the peripheral and in the CNS, where it converts more than 90% of administered L-DOPA before reaching the brain, the identification of a novel component known as COMT inhibitor with low toxicity, strong inhibitory potency, and selectivity for the CNS is of great interest [6]. S-adenosyl-L-methionine (SAM) is transformed into O-methylated products and S-adenosylhomocysteine by the magnesium-dependent enzyme COMT (AdoHcy) [6]. In Alzheimer's disease, potential COMT inhibitors have been discovered by pharmacophore-based inhibitor screening [7]. The dopamine transporter (DAT) also plays an important role in the termination of DA neurotransmission by absorbing DA released into the synapse. There are 5 types of DA receptors which are D1, D2, D3, D4, and D5. DA receptors are classified into two subfamilies which are D1 subfamily and D2 subfamily coupled with G-proteins. The DA receptors, D2, D3, and D4, are members of the D2 subfamily and bind to inhibitory G-proteins [8], whereas DA receptors D1 and D5 belong to the D1 subfamily [9]. Dopamine D1 receptor (D1R) is a significant therapeutic target implicated in a variety of mental and neurological illnesses. Selective D1R agonism is being pursued as a treatment option for various illnesses [10]. Most selective D1R agonists contain a dopamine-like catechol moiety in their chemical structure, hence limiting their therapeutic potential in vivo [10]. Unlike dopamine D1 receptors, the anatomical location of D5 receptors in the CNS is unknown. Most pharmacological and physiological studies do not identify the individual role of D1 and D5 receptors due to the lack of distinct agonists and antagonists [11]. The only significant difference that has been extensively recognized is that the D5 receptor has a higher affinity for dopamine than the D1 receptor [11]. The D4 receptor has been used for the study of many disorders like attention deficit disorder, metastatic progression, and erectile dysfunction. G-proteins, also known as guanine nucleotide-binding proteins, are a protein family that acts as a molecular switch inside cells, sending signals from various stimuli outside the cell to its inside [12]. The D2 dopamine receptor is the major target for both conventional and atypical antipsychotic medications, as well as for the pharmaceuticals used to treat Parkinson's disease [13]. D2 and D3 DA receptors are pharmacologically similar. Therefore, D3 receptors can also be considered as a possible target for antipsychotic and anti-Parkinson's medications [8]. Given the necessity to incorporate new technology in order to aid in the development of innovative medications, bioinformatics is gaining prominence in practically all therapeutic domains. Many researchers have

used some of these technologies to find and/or help in the creation of novel compounds since the 1990s [6]. Bioinformatic techniques have been widely employed to investigate the conformational space of a receptor in the binding domain of the chosen target protein due to the difficulties in acquiring the 3D structure of protein complexes [6]. A creative approach for finding molecules with increased potency is pharmacophore modelling. It has been a significant and effective tool for drug discovery over the recent decades, and it has been recommended that docking studies (binding affinity studies) be performed after pharmacophore screening [14]. In the absence of a macromolecular target structure, ligand-based pharmacophore modelling has emerged as a critical computational technique for aiding drug development. It is frequently performed by identifying shared chemical properties from the three-dimensional (3D) structures of a collection of well-known ligands that are indicative of the ligands' important interactions with a particular macromolecular target [15]. There are several 3D pharmacophore modelling applications available, some of which are free for academic users. The fundamental idea behind 3D pharmacophores does not change even though the exact meaning and application of pharmacophore properties and their attributes may vary between various 3D pharmacophore modelling applications [16]. In pharmacophore-based virtual screening, virtual libraries of compounds are screened against 3D pharmacophores created from a collection of active ligands, a target–ligand combination, or the entire target [16]. The libraries are queried for molecules with pharmacophore matches that match the query. Comparing virtual screening to *in vitro* high-throughput screens, the hit rate may be considerably increased for compounds, lowering the number of compounds that need to be tested experimentally [16].

Methodology

Database creation

Databases are the most important in the drug discovery context since they serve as the foundation for drug repositioning strategies. Thus, it is essential that they be properly developed and incorporate features that aid researchers in getting better results [17]. For setting up the database, the 3D chemical structures of drug compounds used in Parkinson's disease were downloaded in the ".sdf" format. Initially the list of compounds was derived from the DrugBank Database and the compounds were downloaded from the PubChem database. The DrugBank database is an all-inclusive, freely accessible online database that contains information on medications and drug targets that was created and is maintained by the University of Alberta and The Metabolomics Innovation Centre in Alberta, Canada. DrugBank amalgamates accurate drug data (chemical, pharmacological, and pharmaceutical) with extensive drug target data (sequence, structure, and route) along with their pathways and indications. PubChem is a database of chemical compounds and their activity against biological entities. The system is managed by the National Centre for Biotechnology Information, which is part of the National Library of Medicine, which is part of the National Institutes of Health in the United States. Millions of compound structures

and informative datasets are freely available for download via FTP. PubChem includes a variety of substances with descriptions as well as small molecules with less than 100 atoms and 1000 bonds. Over 80 database providers contribute to the ever-expanding PubChem database. A total of 16 compounds were selected and downloaded for the above-mentioned databases.

These compounds are: apomorphine, cycrimine, droxidopa, istradefylline, lisuride, mephensine, moxonidine, pimavanserine, pramipexole, profenamine, rivastigmine, ropinirole, rotigotine, safinamide, selegiline and tolcapone. In order to identify more number of candidate lead molecules, two sets of databases were prepared for model building. The first set consists of the 16 compounds listed above, which have similar pharmacological features. The second set consisted of compounds with structural similarity to the compound istradefylline. This set contains a total of nine compounds. As a supplement to levodopa/carbidopa in individuals with motor fluctuations, istradefylline is an alternative to dopaminergic medications [18].

Alignment checking

Furthermore, the alignment perspective tool found in the software LigandScout's (v.4.4) alignment perspective tool was used to align the ligands and their major structural characteristics were integrated [6]. Each of the 16 compounds was taken as a reference compound, and alignment perspective was completed. The RMSD (root mean square deviation) value was calculated. The compound set with the lowest RMSD value was selected for generating the database for the model building. The 16-compound database was used for screening purposes. An active set database was created with all these compounds. After excluding compounds that were not aligned properly, the first set then had 10 compounds and the second set remained the same with 9. The average difference between the corresponding atoms of two small molecules is indicated by the RMSD value; the lower the RMSD, the more similar the two structures are [19].

Inactive dataset

An inactive set database was created with the help of the DUD-E site [4]. A total of 49 compounds for the first case and 51 for the second case were found in the inactive database. This inactive set was used for screening purposes. The database was saved and kept aside. This was done for validation of the model. So, while screening, the software should predict 49 inactive compounds and 51 inactive compounds for the first and second sets, respectively.

Model building

For model building, we selected three compounds, which were the reference compound, droxidopa, and its similar compounds, istradefylline and apomorphine. Their multi-conformer was used to construct a number of different pharmacophore models [6]. Ten models were generated for both cases, and they were screened.

Model validation and screening

The models were validated using an active set of (10 compounds) and an inactive set of (49 compounds) compounds. Furthermore, the model was screened against the PubChem database to identify compounds with similar pharmacophore features. After the screening, 243 hits for the first case and 43 hits for the second case were observed. In the selection phase, only the ligands that shared all of the necessary structural moieties with the training set were considered hits [6].

Pharmacokinetic properties prediction

A powerful molecule needs to be concentrated sufficiently to reach its target in the body and remain there in a bioactive form long enough for the anticipated biological reactions to take place for it to be effective as a medication [20]. In order to produce new drugs, it is necessary to evaluate absorption, distribution, metabolism, and excretion (ADME) ever earlier in the discovery phase, when the number of potential compounds is high but access to the physical sample is scarce [14]. In that situation, computer models are appropriate substitutes for experimentation. Here, physicochemical characteristics, pharmacokinetics, drug-likeness, and medicinal chemistry are predicted using the web application SwissADME. A total of 286 compounds were subjected to pharmacokinetic properties prediction by adding the compounds of both cases. After the ADME analysis, 30 compounds were screened out of 243 from the first case and 3 compounds out of 40 compounds from the second case. The following criteria were used to filter the compounds: •The molecular weight is an important factor since as the MW increases, the absorption decreases. The drug must pass through the skin barrier for absorption to take place, so it is suggested that the MW should be less than 500 g/mol [21]. •According to Lipinski's Rule of Five, an orally active medicine should typically have a molecular weight under 500 g/mol, a partition coefficient, log P, of less than 5, not more than 5 hydrogen bond donors (OH and NH groups), and not more than 10 hydrogen bond acceptors [21]. •The blood–brain barrier (BBB) is a specialized network of brain microvascular endothelial cells that protects the brain from blood toxins, nourishes brain tissue, and removes dangerous substances from the brain and returns them to the bloodstream [22]. BBB permeability was also a criterion for screening.

Toxicity prediction

The ability to forecast toxicity has a significant impact on public health. Because it allows for the avoidance of various pharmacological research, toxicity prediction is crucial for lowering the cost and labour of a medicine's preclinical and clinical trials, in addition to its many other applications (clinical, animal, and cellular) [23]. Using the Toxtree V3.1 program, toxicity prediction was carried out. The compounds with high toxicity were removed, and the compounds with low

and medium toxicity were taken. While 10 compounds were screened in the first case, no compounds were successful in the second.

Molecular docking

The molecular docking problem's primary goal is to identify the best possible pairing of ligand and receptor that consumes the least amount of energy and attaches to a specific protein of interest [24]. The type and strength of the signal that will be generated can be predicted using docking. As a result of its ability to anticipate how small molecule ligands would bind to the proper target binding site, it is one of the most frequently employed methods in structure-based drug design [24]. Protein targets were identified and prepared from the Protein Data Bank (PDB) server. The Protein Data Bank is a database that contains three-dimensional structural data for big biological entities including proteins and nucleic acids. The data are collected by X-ray crystallography, NMR spectroscopy, or, increasingly, cryo-electron microscopy, and provided by biologists and biochemists from around the world. These data are publicly accessible on the Internet through the websites of its member organizations. The PDB is managed by a group called the Worldwide Protein Data Bank. The targets chosen for this investigation were dopamine D3 receptors (PDB ID: 7CMV) and catechol-O-methyltransferase (PDB ID: 1H1D). The three-dimensional structure of both the proteins is shown in Figs. 2 and 3, respectively. The 10 compounds from the first case were downloaded

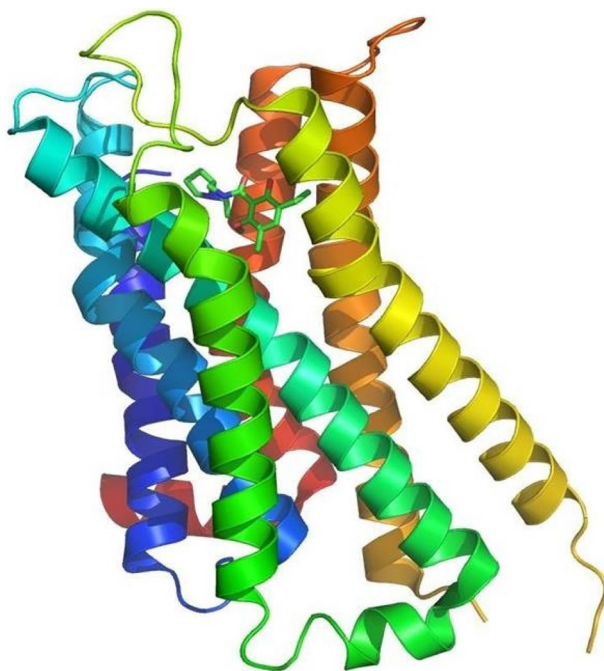


Fig. 2 PDB structure of dopamine receptor

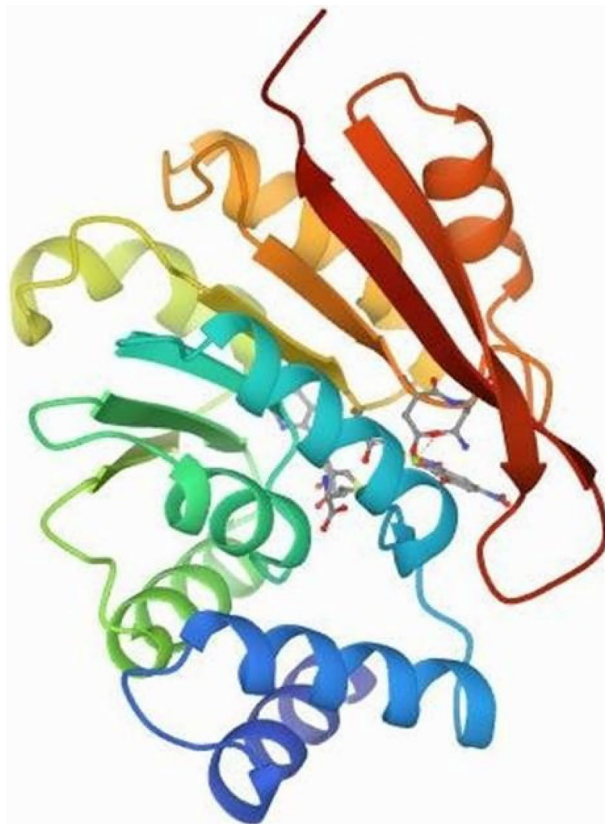


Fig. 3 PDB structure of catechol methyl transferase

from PubChem and converted to PDB format. The docking was done using the POAP server. Open Babel and the AutoDock package are designed to run with highly efficient parallelization using the parallel-based pipeline—POAP [25]. A special feature of POAP is the ligand preparation module, which provides a wide range of choices for geometry optimization, conformer creation, parallelization, and also quarantines incorrect datasets for smooth operation [25]. Additionally, POAP has multi-receptor docking, which may be used for virtual comparison screening and drug repurposing research [25].

The software Discovery Studio Version 16 was used to further examine the pose with the lowest binding energy, which was expressed in kcal/mol by the POAP server, and their atomic interactions in order to better visualize and recognize the most significant atomic interactions.

Molecular dynamics simulations

Molecular dynamics can be used to explore conformational space and is often the method of choice for large molecules such as proteins [26]. Due to the significant advancements in the field of technology, the application of molecular dynamics simulation is turning out to be arduous. Molecular simulation, in simple terms, is a systematic computer simulation of real experimental molecules. By applying Newton's equations of motion to the system, molecular dynamics explores the energy surface [27]. MD modelling mainly determines how a biomolecular system will react to a perturbation [26]. To find consistent differences in the outcomes in each of these situations, one should often run multiple simulations of both the disturbed and unperturbed systems [26]. With the rapid growth of molecular simulation technology, the world's major firms have created a variety of molecular simulation calculation tools, such as TINKER, Gromacs, Materials Studio, and LAMMPS, to suit the research needs of many fields [27]. This method has been used for various research pertaining to biomolecules and also to determine the relationship between various units of measurement. In a recent study [27] conducted in order to determine the properties of PAMAM dendrimer-based macromolecules, this method was used to determine the stability of the PAMAM dendrimers and the complex it forms with different drugs [27]. MD simulation was also used to study the dependence of temperature on specific volume of Lennard-Jones potential [28]. In a study conducted in 2012, MD simulation was also used to determine the affinity of plasticizers to nylon where the cohesive energy and the solubility parameters were taken into consideration [29]. Desmond, MD simulation software package was utilized for the molecular dynamics simulation. Both the protein–ligand complexes and the target protein utilized for docking have undergone MD simulation research. A series of steps were needed to be done in the software before engaging into molecular dynamics which were such as protein preparation, system build, and energy minimization. A protein trimer was positioned in the centre of the box during protein preparation, and 1.0 nm was chosen as the minimal distance between the solute surface and the box [30]. Further, a 100-nanosecond (ns) MD simulation was run for each system at 1 bar and 300 K [30]. Finally, for further examination, the atomic coordinates were saved to the trajectory file every 0.5 picoseconds (ps) [30]. In the subject of computer modelling evaluations, structure comparison approaches have been actively developed and employed for quantitative evaluation of anticipated model accuracy [27]. They are presently used to identify, evaluate, comprehend, and predict protein conformational changes, which are the essential foundation of their biological activity [31]. Any technique that depends on comparing the distances between points of reference in the model and their corresponding counterparts in the source template necessitates superimposing the model on the template beforehand, with the comparison's outcomes being blatantly dependent on the superimposition [31]. All superimposition-dependent approaches are hampered by the ambiguity that arises from the ambiguous goal of finding an optimum superimposition, which has several solutions that each optimize a different set of parameters [31]. The most commonly used quantitative metric for evaluating the degree of similarity between two stacked atomic coordinates is the root mean square deviation (RMSD) [31]. The RMSD of atomic locations in bioinformatics is a measurement of the typical separation between superposed protein atoms, which are typically backbone

atoms [32]. Non-protein molecules, such as small organic chemical compounds, can also be calculated using the RMSD method [32]. To compute the RMSD, first, carry out a least-squares fit to minimize the variation between two overlaid structures [33]. This presupposes that the structures are stiff and that the global minimum is quickly discovered using the least-squares method [33]. When the two structures are superimposed, the RMSD (in length units) is estimated based on

$$\text{RMSD} = \sqrt{\sum_{i=1}^N d_i^2 / N} \quad (1)$$

where N is the total number of equivalent atoms, and d_i is the distance that exists between atom i in each of the configurations [33]. The unit of RMSD is Å [31]. Only when the RMSD is as minimal as it is for nearly related proteins (< 3 Å), is success evident in structural comparison [34]. A difference in structure is shown by an RMSD of 3 Å between two tripeptides, but a comparable difference is indicated by an RMSD of 100 between two chains of 100 residues [34]. The closer the two structures are to one another, the lower the RMSD between them [34].

RMSD statistics were provided when the results were obtained from the Desmond software. The below equation is what the software uses to calculate RMSD. It is calculated across the trajectory's whole frame set [35]. The RMSD for frame x is:

$$\text{RMSD}_x = \sqrt{1/N \sum_{i=1}^N (r_i(t_x) - r_i(t_{ref}))^2} \quad (2)$$

where N is the number of atoms in the atom selection, t_{ref} is the reference time, usually the first frame, which is taken to be time $t = 0$, frame x is captured at time t_x , and r' is the position of the chosen atoms in frame x after superimposing on the reference frame. The procedure is repeated for every frame in the simulation trajectory [35]. A protein's function is influenced by both its structure and dynamics. The divergence in the evolution of protein movements can be studied to get a further understanding of the dynamics-function link [36]. The correlation between the root mean square fluctuations (RMSFs) of aligned residues is the most commonly used dynamical similarity score [36]. Any array of protein conformations, such as those generated through MD simulations or Monte Carlo (MC) simulations, can be used to determine RMSFs [36]. RMSF merely takes into account the total size of each C-atom's variation [36]. When analysing a structure, RMSF looks at the areas that deviate most (or least) from its mean structure [37].

A structure's RMSF is the time average of its RMSD [38]. It is computed using the following equation:

$$\rho_i = \sqrt{\langle (x_i - \langle x_i \rangle)^2 \rangle} \quad (3)$$

where x_i represents the coordinates of particle i and $\langle x_i \rangle$ represents the ensemble average location of i [38]

The RMSF can identify which portions of a system are the most mobile, whereas the RMSD measures the amount that a structure deviates from a reference over time [38]. The RMSF ought to be calculated based on the simulation's average structure rather than the initial state, where the RMSD is commonly calculated [38]. High mobility is indicated by a region of the structure where the RMSF values regularly depart from the average [38]. When RMSF analysis is performed on proteins, it is usually limited to alpha-carbon atoms, which are more indicative of changes in conformation than the more flexible side chains [38]. RMSD and RMSF both express the divergence of the particle's target state from the reference state in terms of their physical meaning. The RMS analysis combines the RMSD and RMSF evaluations [38].

RMSF statistics were also provided when the results were obtained from the Desmond software. The below equation is what the software uses to calculate RMSF. The RMSF for residue is:

$$\text{RMSF}_i = \sqrt{1/T \sum_{t=1}^T \langle (r_i(t)) - r_i(t_{\text{ref}})_2 \rangle} \quad (4)$$

where T is the trajectory time over which the RMSF is calculated, t_{ref} is the reference time, r_i is the position of residue i ; r' is the position of atoms in residue i after superposition on the reference, and the angle brackets indicate that the average of the square distance taken over the selection of atoms in the residue [35].

Results

Pharmacophore model

Using the structures acquired from the DrugBank and PubChem, pharmacophore models were created. The developed model should have the ability to distinguish between "active" and "inactive" ligands [39]. If an inactive molecule outperforms an active molecule, the hypothesis may be inaccurate since it does not distinguish between actives and inactive [40]. Greater selectivity is preferred since it suggests that the hypothesis is more likely to be specific to the ligands in the active set [39]. A ligand-based pharmacophore model was generated with a fitness score of 72.77. The obtained model structures along with the bond angles and bond lengths are given Figs. 4 and 5.

After screening each model from each case (i.e. first case and second case) with the respective active and inactive datasets, it was observed that model 8 in the first case had 9 hits consisting of 9 true positives and 0 false positives, and model 5 in the second case had 11 hits with 10 true positives and 1 false positive.

Further, the model was screened against the PubChem database to identify compounds with similar pharmacophore features. The effective implementation of a virtual database screening approach employing pharmacophore models as 3D queries has allowed the retrieval of prospective compounds that may be reliably employed

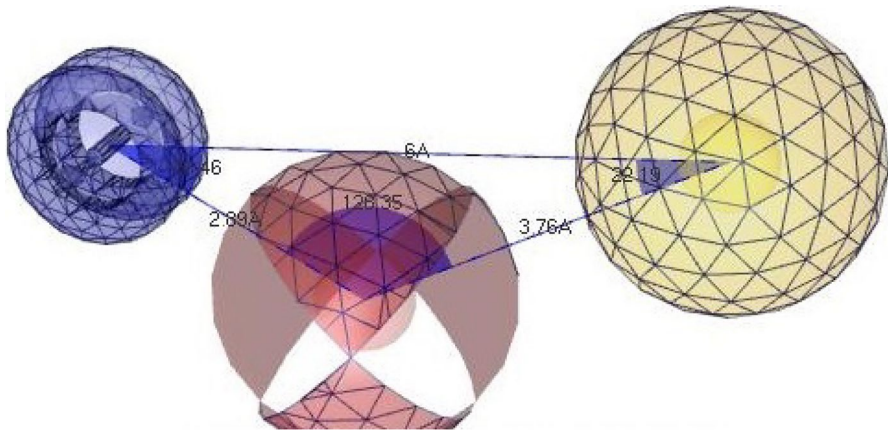


Fig. 4 Pharmacophore model obtained using 9 actives

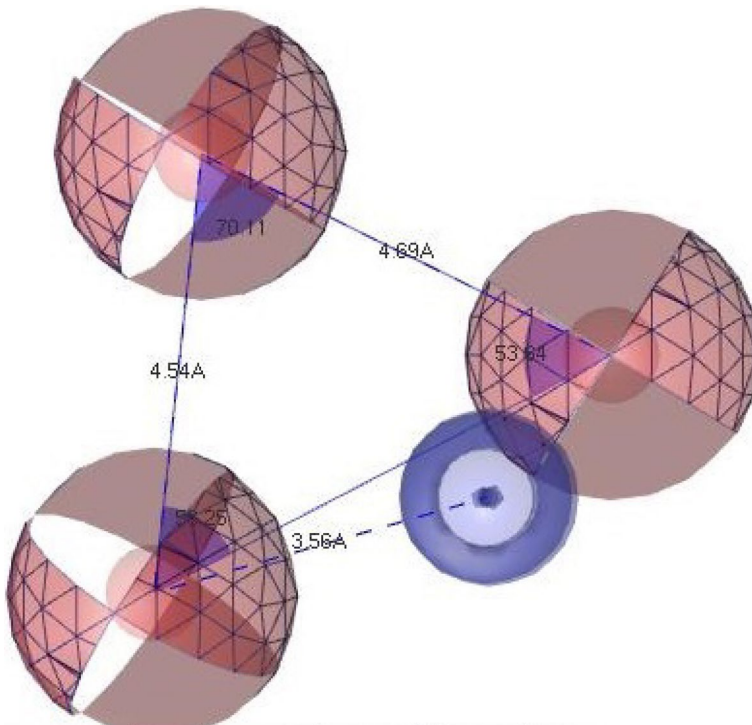


Fig. 5 Pharmacophore model obtained using 11 actives

in the discovery and development of innovative drugs [41]. In total, 243 compounds were obtained as hits in the first case, and 43 compounds were obtained in the second case after screening. Using docking studies, these compounds were tested for

binding affinity in protein targets such as dopamine receptors, adrenergic receptors, and adenosine receptors. The compounds that scored above -5 KJ/mol were chosen and subjected to pharmacokinetic property prediction, including toxicity studies.

ADME analysis

The compounds with drug-like properties and those that could cross the blood–brain barrier were filtered. All the 243 compounds in the first case and 43 compounds in the second case underwent ADME screening. Thirty compounds from the first case and three from the second case were BBB (blood–brain barrier) permeant and possessed drug-like properties. The detailed results of the ADME analysis are given in Table 1 in which the last three compounds belong to the second case.

Toxicity analysis

The toxicity of the chemicals assessed using the program Toxtree is shown in Table 2 in which the last three compounds belong to second case. Compounds of minimal toxicity were obtained and used in subsequent processes. In the first case, 10 compounds were obtained, whereas no compounds were retrieved in the second case.

Molecular docking studies

The ten chemicals obtained in the preceding stages were fetched from the PubChem repository. These chemicals were converted to the PDB format and submitted to the system. Docked results are given in Table 3. Figure 6 depicts the 2D docked structure obtained from the complex 1H1D-7471813. Figure 7 depicts the 2D docked structure obtained from the complex 7CMV-7471813. The dark green colour represents conventional hydrogen bonds. The light green colour represents van der Waals interaction. The orange colour represents Pi-anion. The dark yellow colour represents Pi-sulphur. The dark pink colour represents Pi-Pi stacked. The light pink colour represents Pi-alkyl. The coloured circles represent proteins (amino acids), and the cyclic structure represents the ligand molecule. Surface-accessible atoms that may potentially carry lone pairs are defined as hydrogen bond acceptor sites [39]. A hydrogen bond donor site is centred on a polar hydrogen atom, and a single vector feature is oriented along the hydrogen bond axis [39].

Molecular dynamics simulation studies

Initially, the target molecules, i.e. catechol-O-methyltransferase (PDB ID: 1H1D) and dopamine D3 receptors (PDB ID: 7CMV) were subjected to molecular dynamics simulation, and the RMSD values that we got for the targets were within the limits, i.e. between 1 and 3 for each of these targets, which are given in Table 4. And for each of these compounds, the corresponding graphs were obtained.

Table 1 ADME Results

Data set	LIGAND PubChem ID	Molecular Weight (g/ mol)	Hydrogen Bond Acceptor	Hydrogen Bond Donor	BBB Per- meant	Lipinski Viola- tions	Bio- avail- ability Score
First Case	20291156	246.3	2	0	YES	0	0.55
	20361792	209.25	3	0	YES	0	0.55
	20578428	494.62	4	1	YES	1	0.55
	21623110	334.45	4	1	YES	0	0.85
	23114952	286.32	4	0	YES	0	0.55
	104564	297.35	3	0	YES	0	0.55
	1100329	412.48	4	0	YES	0	0.55
	3759504	371.23	10	1	YES	0	0.55
	3797600	375.43	6	2	YES	0	0.55
	10939801	482.62	4	0	YES	1	0.55
	11059845	476.64	1	1	YES	1	0.55
	11067466	243.28	4	0	YES	0	0.55
	11370993	332.44	2	0	YES	1	0.55
	11401426	202.21	3	0	YES	0	0.55
	12173638	562.65	6	0	YES	1	0.55
	5262128	297.39	5	0	YES	0	0.55
	6138700	337.46	3	1	YES	0	0.55
	57527585	476.4	1	0	YES	1	0.55
	67123	358.39	6	0	YES	0	0.55
	96874	342.44	4	0	YES	0	0.55
	1754908	424.62	2	2	YES	0	0.55
	1754911	422.6	4	2	YES	0	0.55
	2297311	278.39	4	2	YES	0	0.55
3027735	425.56	5	1	YES	0	0.55	
3146121	490.63	6	0	YES	0	0.55	
3146122	490.63	6	0	YES	0	0.55	
7471813	318.37	3	0	YES	0	0.55	
17357779	326.39	4	1	YES	0	0.55	
17780758	334.45	4	0	YES	1	0.55	
17844478	211.26	3	1	YES	0	0.55	
Second case	11471102	266.12	8	0	YES	0	0.55
	33824	251.37	3	2	YES	0	0.55
	316301	293.45	3	2	YES	0	0.55

First Case: Compounds with structural similarity; Second Case: Compounds with similar structure to that of istradefylline

Then, molecular dynamics simulation was performed to each of the target–ligand complexes that were obtained after docking. The complexes of both targets produced with the ligand with PubChem ID 7471813, or

Table 2 Toxicity Results

Data set	LIGAND PubChem ID	Toxicity Level
First case	20291156	HIGH
	20361792	HIGH
	20578428	HIGH
	21623110	LOW
	23114952	LOW
	104564	LOW
	1100329	HIGH
	3759504	HIGH
	3797600	HIGH
	10939801	HIGH
	11059845	HIGH
	11067466	HIGH
	11370993	HIGH
	11401426	LOW
	12173638	HIGH
	5262128	LOW
	6138700	HIGH
	57527585	HIGH
	67123	LOW
	96874	HIGH
	1754908	HIGH
	1754911	HIGH
	2297311	LOW
	3027735	HIGH
	3146121	HIGH
3146122	HIGH	
7471813	LOW	
17357779	HIGH	
17780758	LOW	
17844478	LOW	
Second case	11471102	HIGH
	33824	HIGH
	316301	HIGH

3,4-Bis(1,3,5,6-heptatetraenyloxy) benzaldehyde (name provided in PubChem), were determined to be stable since their values were between the specified limits, i.e. 1 and 3. Table 4 includes the relevant information. The graph for the RMSD plot for the 1H1D and 1H1D-7471813 complex and 7CMV and 7CMV-7471813 complex is shown in Figs. 8 and 9, respectively. The blue indicates the selected target and the red the target–ligand complex. The time frames picosecond is displayed on the X-axis, while the corresponding RMSD values are displayed on the Y-axis (in Angstrom).

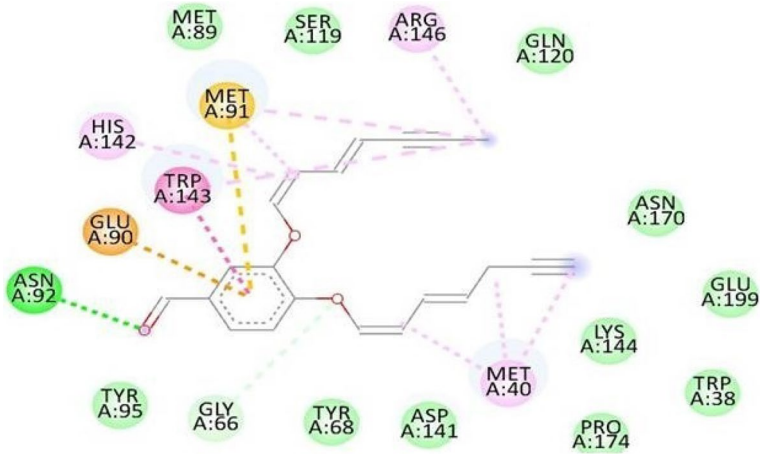


Fig. 6 2D Interaction of 1H1D-7471813 complex

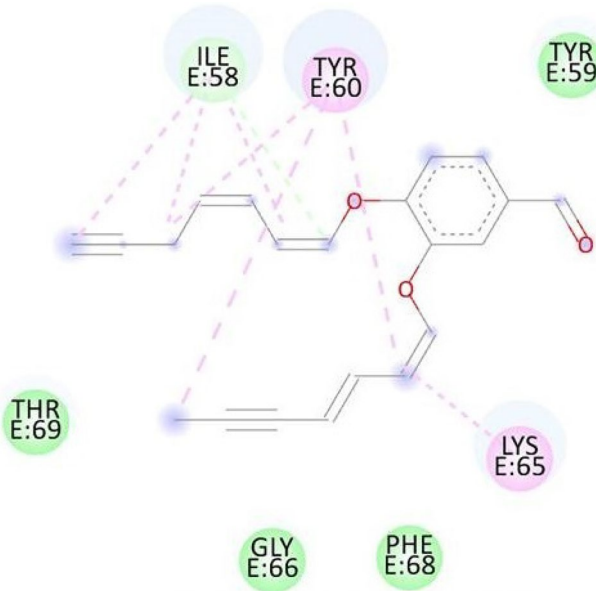


Fig. 7 2D Interaction of 7CMV-7471813 complex

The RMSF plots for 1H1D and 1H1D-7471813 complexes are shown in Fig. 10. The graph's Y-axis displays the RMSF value, while the X-axis displays the residue count. The protein regions with the highest fluctuations during the simulation are shown by the peaks. The protein's N- and C-terminal tails often change more than any other region of it. Secondary structural components like

Table 3 Docking Results

Ligand Molecule PubChem CID	Energy values for respective target molecules (kCal/mol)	
	1H1D	7CMV
CID7471813	– 7.9	– 5.0
CID104564	– 7.8	– 4.6
CID21623110	– 7.6	– 4.4
CID67123	– 7.7	– 4.0
CID23114952	– 6.6	– 4.9
CID11401426	– 6.8	– 4.4
CID17780758	– 6.7	– 4.4
CID2297311	– 6.3	– 4.6
CID17844478	– 6.6	– 4.1
CID5262128	– 6.3	– 3.4

Table 4 Statistics of RMSD and radius of gyration

Statistics	Apo and Holo forms	Mean	Median	SD
RMSD	7CMV	1.952	1.962	0.153
	1H1D	1.83	1.84	0.227
	7CMV-7471813 complex	2.702	2.557	1.019
	1H1D-7471813	1.880	2.069	0.345
Radius of gyration	1H1D-7471813	16.725	16.730	0.079
	7CMV-7471813	17.788	17.787	0.117

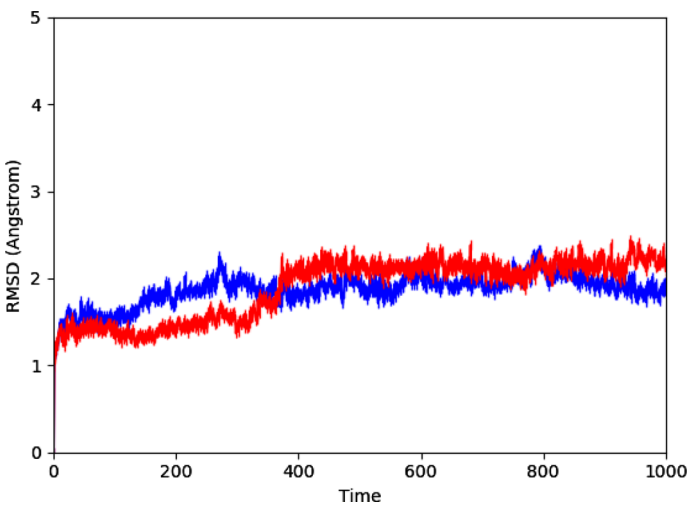


Fig. 8 RMSD plot for 1H1D and 1H1D-7471813

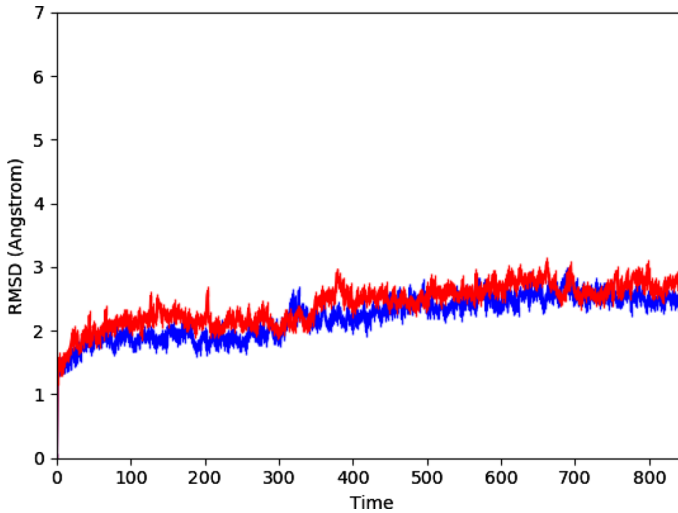


Fig. 9 RMSD plot for 7CMV and 7CMV-7471813

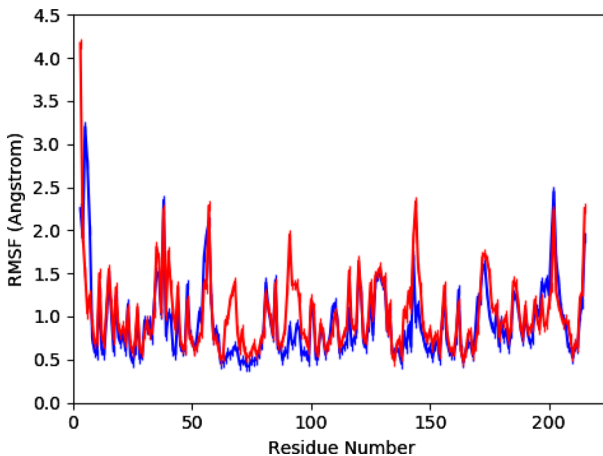


Fig. 10 RMSF plot of 1H1D and 7471813

alpha helices and beta strands change less than loop sections because they are often more rigid than the protein's unstructured portion.

The radius of gyration of the target–ligand complexes was then determined. The arrangement of atoms of a protein along its axis is known as the radius of gyration (Rg). Rg is the length that corresponds to the separation between the rotating point and the location where the energy transfer has the greatest impact. The Rg offers details about the size and compaction of the protein molecules. As illustrated in Table 4, their matching values were acquired in tabular form and

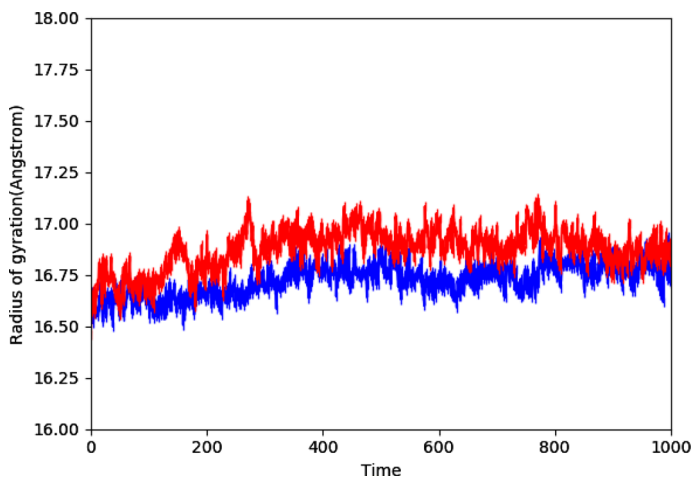


Fig. 11 Radius of gyration plot of 1H1D-74718133

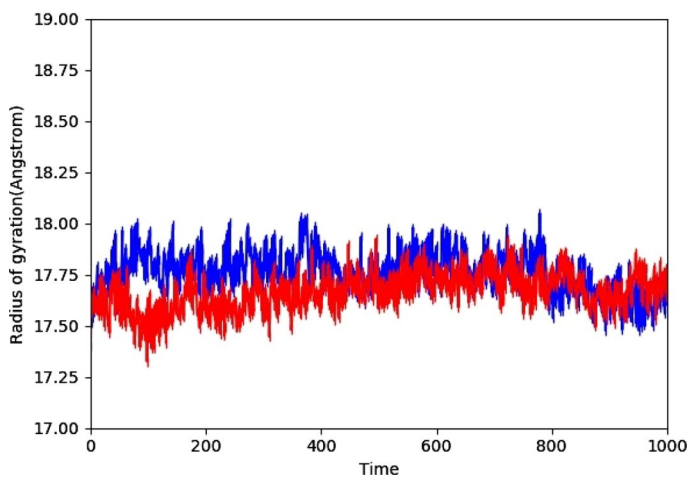


Fig. 12 Radius of gyration plot of 7CMV and 7CMV-7471813 complex

Figs. 11 and 12 show the graphical version of the data. The X-axis shows the time in picoseconds (ps) while the Y-axis shows the radius of the gyration value.

Discussion

The final pharmacophore model was reached through many trials and errors with different sets of databases. The selected pharmacophores were finalized based on their number of hits, or score, and their ability to distinguish between active sets and inactive sets. The pharmacophore models were rated based on their similarity to the active chemicals; a scoring function was used; and ten pharmacophore hypotheses

were retained [40]. Alignment of site points and vectors, volume overlap, selectivity, number of ligands matched, relative conformational energy, and activity all contribute to the scoring method [40]. Individual compounds or chemical libraries can be tested as “licence-in” opportunities using ADME in silico models, evaluating their eligibility as prospective therapeutic molecules [42]. The capacity to forecast the effect of a suggested structural adjustment in silico prior to compound synthesis would result in fewer redesign–synthesize–test cycles during lead optimization. This is most effective if a model can give suggestions on structural changes that can improve a property [42]. In silico approaches for estimating the toxicity of possible medicinal drugs are complicated systems. They are upgraded and evolved on an annual basis [43]. Their advancements are followed by the growth in the number of huge databases and the number of different computer applications [43]. Molecular docking was utilized to further understand and evaluate the preliminary data acquired in pharmacokinetic investigations, removing potential false negatives and examining the atomic interactions created with the active site in the proteins in great detail [6]. The docking approach regarded the conformationally sampled ligands and target proteins as hard structures. For each ligand, the ten docking postures with the lowest energy were preserved [44]. By comparing the atomic coordinates of each docked ligand to the X-ray coordinates of the reference ligands, the produced ligand–protein complexes were visually evaluated [44]. Asp141, His142, Trp143, Lys44, Asn170, Pro174, and Glu199 involved in the interaction profile of COMT [7] are present here in this docking study also. During the molecular dynamics simulation, the dynamic properties of the target and the target–ligand complex were determined using their respective RMSD and RMSF values. If the simulation’s fluctuations at the end are centred on a thermal average structure, RMSD analysis can tell whether the simulation has been equilibrated. For tiny, globular proteins, changes of the order of 1–3 are quite acceptable. However, changes that are considerably exceeding the given values suggest that the protein is undergoing a significant structural shift throughout the simulation. It is crucial that the RMSD values in the simulation are stable at a particular value. If the protein’s average RMSD is still rising or falling at the conclusion of the simulation, the system has not yet reached equilibrium, and your simulation may not have lasted long enough for a thorough study. None of the three MD simulations produced particularly unstable structures, which validated the convergence tendency of the systems throughout the simulation. The results are therefore credible for further research.

As future prospective of the work, the simulated compound first needs to be formulated. Primarily, drug compounds are the focus of early formulations, which are administered to animals via a variety of methods, including intravenous and oral administration, during the discovery and preclinical stages of drug development [45]. Once formulation is done, in vivo testing is the first step in the preclinical stage of the development of drugs to ascertain the medicine’s effectiveness and safety [46]. They are used to assess these substances against a variety of pharmacological objectives, particularly pharmacology (activity or efficacy), pharmacokinetics (PK), and toxic effects [45]. Pharmacology, PK, and toxicology studies are commonly conducted on experimental mice. Now, there are chips that mimic all the metabolic activities of a real cell or animal tissue. Ex vivo, in vitro, and in vivo tests

are performed on whole, live beings or cells, such as those from animals or people, or with the use of non-living organisms or tissue extract [46]. Examples of in vivo preclinical research include the generation of novel medications utilizing mouse, rat, and dog models [46]. In vitro research is a study done in a lab [46]. Ex vivo employs animal tissues or cells that are taken from dead animals [46]. Throughout the pre-clinical and clinical stages, formulation optimization should be carried out. It makes sure that the right amount of medication, at the right time, is administered to the correct location [46]. Then, a range of dosage forms and solubility-improving strategies, such as different solutions, suspensions, lipid-based formulations, and solid dispersions, are evaluated [45]. Beyond this are the clinical trials and approval.

Conclusion

Hence, by targeting both specified proteins, the compound 3,4-Bis(1,3,5,6-hepta-tetraenyloxy) benzaldehydes is found to have favourable pharmacokinetic properties and forms a stable complex with the target 1H1D. This interaction should be beneficial against Parkinson's disease. While conducting the molecular docking studies, it was inferred that the above-mentioned compound had a significantly high binding affinity with the catechol-O-methyltransferase(COMT) (1H1D), which is a component that is responsible for inactivation of L-3,4-dihydroxyphenylalanine (L-DOPA) [47] which is a precursor of dopamine which has a significant role in Parkinson's disease. This happens by the process of O-methylation [47] where more than 90% of administered L-DOPA is deactivated before reaching the brain. So, the compound 3,4-Bis(1,3,5,6-hepta tetra enyloxy) benzaldehyde acts as a COMT inhibitor to prevent the degradation of L-DOPA which in turn can prevent the rigorous symptoms of Parkinson's disease from taking place. As mentioned earlier, the D3 receptor can be considered as a possible target for antipsychotic and anti-Parkinson's medications [8], as the D2 dopamine receptor is the major target for both conventional and atypical antipsychotic medications, as well as for the pharmaceuticals used to treat Parkinson's disease [13] and D2 and D3 DA receptors are pharmacologically similar. The discovered compound also had a high binding affinity with the D3 receptor also. The ROC curve graphs demonstrate the relevance of the pharmacophore models, and the RMSD, RMSF, and radius of gyration plots are all stable. The 2D interaction figures reveal that there is good target–ligand interaction. The graphs of RMSD, RMSF, and radius of gyration then demonstrate stability during the interaction of the target and ligand. The screened compounds were subjected to ADME investigations, and only those that passed all of the screening criteria were advanced to the next stage. In the toxicity studies conducted only non-toxic substances were chosen. It can be inferred from all of these tests and steps that the resulting molecule, 3,4-Bis(1,3,5,6-heptatetraenyloxy) benzaldehyde, has no adverse effects according to the computational studies.

Author Contributions All authors contributed to the study conception and design. Material preparation, and data collection and analysis were performed by AJ, SM, NMT, MC, ADM, and AJ. The first draft

of the manuscript was written by NMT and SM, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding The authors did not receive support from any organization for the submitted work.

Availability of data and materials Not applicable.

Declarations

Conflict of interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication All the authors informed the consent for publication.

Code availability Not applicable.

References

1. Ball N, Teo W-P, Chandra S, Chapman J (2019) Parkinson's disease and the environment. *Front Neurol* 10:218
2. Dauer W, Przedborski S (2003) Parkinson's disease: mechanisms and models. *Neuron* 39(6):889–909
3. de Lau LML, Breteler MMB (2006) Epidemiology of Parkinson's disease. *Lancet Neurol* 5(6):525–535
4. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK (2012) Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J Med Chem* 55(14):6582–6594
5. Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkman J, Schrag A-E, Lang AE (2017) Parkinson disease. *Nat Rev Dis Primers* 3:17013
6. Cruz-Vicente P, Gonçalves AM, Ferreira O, Queiroz JA, Silvestre S, Passarinha LA, Gallardo E (2021) Discovery of small molecules as membrane-bound catechol-o-methyltransferase inhibitors with interest in parkinson's disease: Pharmacophore modeling, molecular docking and in vitro experimental validation studies. *Pharmaceuticals (Basel)* 15(1):51
7. Patel CN, George JJ, Modi KM, Narechania MB, Patel DP, Gonzalez FJ, Pandya HA (2017) Pharmacophore-based virtual screening of catechol-o-methyltransferase (COMT) inhibitors to combat Alzheimer's disease. *J Biomol Struct Dyn* 36(15):3938–3957
8. Joyce JN (2001) Dopamine D3 receptor as a therapeutic target for antipsychotic and antiparkinsonian drugs. *Pharmacol Ther* 90(2–3):231–259
9. Wang S, Wacker D, Levit A, Che T, Betz RM, McCorvy JD, Venkatakrishnan AJ, Huang X-P, Dror RO, Shoichet BK, Roth BL (2017) D(4) dopamine receptor high-resolution structures enable the discovery of selective agonists. *Science* 358(6361):381–386
10. Sun B, Feng D, Chu ML-H, Fish I, Lovera S, Sands ZA, Kelm S, Valade A, Wood M, Ceska T, Kobilka TS, Lebon F, Kobilka BK (2021) Crystal structure of dopamine D1 receptor in complex with G protein and a non-catechol agonist. *Nat Commun* 12(1):3305
11. Khan ZU, Gutiérrez A, Martín R, Peñafiel A, Rivera A, de la Calle A (2000) Dopamine D5 receptors of rat and human brain. *Neuroscience* 100(4):689–699
12. Milligan G, Kostenis E (2006) Heterotrimeric G-proteins: a short history. *Br J Pharmacol* 147(S1):46–55
13. Wang S, Che T, Levit A, Shoichet BK, Wacker D, Roth BL (2018) Structure of the D2 dopamine receptor bound to the atypical antipsychotic drug risperidone. *Nature* 555(7695):269–273


14. Ahmed SSSJ, Ahameethunisa A, Santosh W (2010) QSAR and pharmacophore modeling of 4-arylthieno [3, 2-d] pyrimidine derivatives against adenosine receptor of Parkinson's disease. *J Theor Comput Chem* 09(06):975–991
15. Yang S-Y (2010) Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discov Today* 15(11–12):444–450
16. Schaller D, Šribar D, Noonan T, Deng L, Nguyen TN, Pach S, Machalz D, Bermudez M, Wolber G (2020) Next generation 3D pharmacophore modeling. *WIREs Comput Mole Sci* 10(4):1468
17. Kaserer T, Beck KR, Akram M, Odermatt A, Schuster D (2015) Pharmacophore models and pharmacophore-based virtual screening: concepts and applications exemplified on hydroxysteroid dehydrogenases. *Molecules* 20(12):22799–22832
18. Cummins L, Cates ME (2022) Istradefylline: a novel agent in the treatment of “off” episodes associated with levodopa/carbidopa use in parkinson disease. *Ment Health Clin* 12(1):32–36
19. Reva BA, Finkelstein AV, Skolnick J (1998) What is the probability of a chance prediction of a protein structure with an RMSD of 6 Å? *Fold Des* 3(2):141–147
20. Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep* 7:42717
21. Benet LZ, Hosey CM, Ursu O, Oprea TI (2016) BDDCS, the rule of 5 and drugability. *Adv Drug Deliv Rev* 101:89–98
22. Persidsky Y, Ramirez SH, Haorah J, Kanmogne GD (2006) Blood-brain barrier: structural components and function under physiologic and pathologic conditions. *J Neuroimmune Pharmacol* 1(3):223–236
23. Krewski D, Acosta D Jr, Andersen M, Anderson H, Bailar JC 3rd, Boekelheide K, Brent R, Charnley G, Cheung VG, Green S Jr, Kelsey KT, Kerkvliet NI, Li AA, McCray L, Meyer O, Patterson RD, Pennie W, Scala RA, Solomon GM, Stephens M, Yager J, Zeise L (2010) Toxicity testing in the 21st century: a vision and a strategy. *J Toxicol Environ Health B Crit Rev* 13(2–4):51–138
24. Meng X-Y, Zhang H-X, Mezei M, Cui M (2011) Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des* 7(2):146–157
25. Samdani A, Vetrivel U (2018) POAP: a GNU parallel based multithreaded pipeline of open babel and AutoDock suite for boosted high throughput virtual screening. *Comput Biol Chem* 74:39–48
26. Adcock SA, McCammon JA (2006) Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem Rev* 106(5):1589–1615
27. Fatemi SM, Fatemi SJ, Abbasi Z (2020) Pamam dendrimer-based macromolecules and their potential applications: recent advances in theoretical studies. *Polym Bull* 77(12):6671–6691. <https://doi.org/10.1007/s00289-019-03076-4>
28. Al-Raei M, El-Daher MS (2021) Temperature dependence of the specific volume of Lennard-Jones potential and applying in case of polymers and other materials. *Polym Bull* 78(3):1453–1463. <https://doi.org/10.1007/s00289-020-03166-8>
29. Alperstein D, Knani D, Goichman A, Narkis M (2012) Determination of plasticizers efficiency for nylon by molecular modeling. *Polym Bull* 68(7):1977–1988. <https://doi.org/10.1007/s00289-012-0705-2>
30. Sahihi M, Gaci F, Navizet I (2021) Identification of new alpha-synuclein fibrillogenesis inhibitor using in silico structure-based virtual screening. *J Mol Graph Model* 108(108010):108010
31. Kufareva I, Abagyan R (2012) Methods of protein structure comparison. *Methods Mol Biol* 857:231–257
32. Services CB (2013) CD ComputaBio. Accessed on May 16th, 2023. <https://www.computabio.com/rmsd-rmsf-analysis-service.html>
33. Sargsyan K, Grauffel C, Lim C (2017) How molecular size impacts RMSD applications in molecular dynamics simulations. *J Chem Theory Comput* 13(4):1518–1524
34. Reva BA, Finkelstein AV, Skolnick J (1998) What is the probability of a chance prediction of a protein structure with an RMSD of 6 Å? *Fold Des* 3(2):141–147
35. Martínez L (2015) Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis. *PLoS ONE* 10(3):0119264
36. Fuglebakk E, Echave J, Reuter N (2012) Measuring and comparing structural fluctuation patterns in large protein datasets. *Bioinformatics* 28(19):2431–2440
37. Jolla L (2018) BioChemCoRe 2018. Accessed on June 3rd, 2023. <https://ctlee.github.io/BioChemCoRe-2018/rmsd>

38. Gowers R, Linke M, Barnoud J, Reddy T, Melo M, Seyler SL, Dotson D, Domanski J, Buchoux S, Kenney I (2016) Mdanalysis: a python package for the rapid analysis of molecular dynamics simulations
39. Durdagi S, Duff HJ, Noskov SY (2011) Combined receptor and ligand-based approach to the universal pharmacophore model development for studies of drug blockade to the hERG1 pore domain. *J Chem Inf Model* 51(2):463–474
40. Sen D, Chatterjee T (2013) Pharmacophore modeling and 3d quantitative structure-activity relationship analysis of febrifugine analogues as potent antimalarial agent. *J Adv Pharm Technol Res* 4:50–60. <https://doi.org/10.4103/2231-4040.107501>
41. Thangapandian S, John S, Lee Y, Kim S, Lee KW (2011) Dynamic structure-based pharmacophore model development: a new and effective addition in the histone deacetylase 8 (HDAC8) inhibitor discovery. *Int J Mol Sci* 12(12):9440–9462
42. Butina D, Segall MD, Frankcombe K (2002) Predicting ADME properties in silico: methods and models. *Drug Discov Today* 7(11):83–8
43. Toropov AA, Toropova AP, Raska I Jr, Leszczynska D, Leszczynski J (2014) Comprehension of drug toxicity: software and databases. *Comput Biol Med* 45:20–25
44. Crisan L, Istrate D, Bora A, Pacureanu L (2021) Virtual screening and drug repurposing experiments to identify potential novel selective MAO-B inhibitors for parkinson's disease treatment. *Mol Divers* 25(3):1775–1794
45. Li P, Zhao L (2007) Developing early formulations: practice and perspective. *Int J Pharm* 341(1–2):1–19
46. Pandey A (2003) North East BioLab. Accessed on May 22nd, 2023. <https://www.nebiolab.com/drug-discovery-and-development-process/>
47. Bonifácio MJ, Archer M, Rodrigues ML, Matias PM, Learmonth DA, Carrondo MA, Soares-Da-Silva P (2002) Kinetics and crystal structure of catechol-o-methyltransferase complex with co-substrate and a novel inhibitor with potential therapeutic application. *Mol Pharmacol* 62(4):795–805

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Amitha Joy¹  · Sidharth Menon^{1,2} · Neethu Mariam Thomas^{1,2} · Meha Christy^{1,2} · Aiswarya D. Menon^{1,2} · Arun John^{1,2}

✉ Amitha Joy
amitha@sahrdaya.ac.in

Sidharth Menon
sidharthmenon326@gmail.com

Neethu Mariam Thomas
mariamthomasneethu@gmail.com

Meha Christy
mehathottan@gmail.com

Aiswarya D. Menon
aiswaryadmenon@gmail.com

Arun John
arun.john023@gmail.com

- ¹ Biotechnology Department, Sahridaya College of Engineering and Technology, Kodakara, Thrissur, Kerala 680684, India
- ² Department of Molecular Analytics, Saveetha School of Engineering, Chennai, Tamilnadu 602105, India