

Computational Identification and Characterization of MAO-B Inhibitors for Parkinson's Disease: An Integrated Docking and Statistical Analysis Approach



By

Asad Zahid Malik

CIIT/SP21-BSI-011/ISB

&

Muhammad Ali Ejaz

CIIT/SP21-BSI-033/ISB

BS Thesis

In

Bioinformatics

COMSATS University Islamabad, Pakistan

Spring 2024



COMSATS University Islamabad

**Computational Identification and
Characterization of MAO-B Inhibitors for
Parkinson's Disease: An Integrated Docking
and Statistical Analysis Approach**

A Thesis Presented to

COMSATS University Islamabad

In Partial Fulfilment
Of the requirement for the Degree of

BS Bioinformatics

By

Asad Zahid Malik

CIIT/SP21-BSI-011/ISB

&

Muhammad Ali Ejaz
CIIT/SP21-BSI-033/ISB

Spring 2024

Computational Identification and Characterization of MAO-B Inhibitors for Parkinson's Disease: An Integrated Docking and Statistical Analysis Approach

This thesis is submitted to the Department of Biosciences in partial fulfilment
of the requirements for the award of Degree of BS Bioinformatics.

Name	Registration Number
Asad Zahid Maik	CIIT/SP21-BSI-011/ISB
Muhammad Ali Ejaz	CIIT/SP21-BSI-033/ISB

Supervisor

Dr. Saira Amir

Associate Professor

Department of Biosciences

COMSATS University Islamabad

Final Approval

This thesis titled

Computational Identification and
Characterization of MAO-B Inhibitors for
Parkinson's Disease: An
Integrated Docking and Statistical Analysis
Approach

By

Asad Zahid Malik
CIIT/SP21-BSI-011/ISB

&

Muhammad Ali Ejaz
CIIT/SP21-BSI-033/ISB

Has been approved
For COMSATS University Islamabad

External Examiner: _____

Supervisor: _____

Dr. Saira Amir
Department of Biosciences
COMSATS University Islamabad

Head of the Department: _____

Prof. Dr. Ijaz Ali
Department of Biosciences
COMSATS University Islamabad

Declaration

I, Asad Zahid Malik (Registration No: CIIT/SP21-BSI-011/ISB) and Muhammad Ali Ejaz (Registration No: CIIT/SP21-BSI-033/ISB) hereby declare that we have produced the work presented in this thesis, during the scheduled period of study. We also declare

that we have not taken any material from any source except referred to wherever due that amount of plagiarism is within acceptable range. If a violation of HEC rules on research has occurred in this thesis, we shall be liable to punishable action under the plagiarism rule of HEC.

Date: 17th December, 2024

Signature of the Student

Asad Zahid Malik
(CIIT/SP21-BSI-011/ISB)

Signature of the Student

Muhammad Ali Ejaz
(CIIT/SP21-BSI-033/ISB)

Certificate

It is certified that Asad Zahid Malik (CIIT/SP21-BSI-011/ISB) & Muhammad Ali Ejaz (CIIT/SP21-BSI-033/ISB), has carried out all the work related to this thesis under my supervision at the

Department of Biosciences, COMSATS University Islamabad, and the work fulfills the requirements for the award of the degree of BS in Bioinformatics.

Date: 17th December 2024

Supervisor:

Dr. Saira Amir
Associate Professor
Department of Biosciences
COMSATS University Islamabad

Head of the Department:

Prof. Dr. Ijaz Ali
Department of Biosciences
COMSATS University Islamabad

DEDICATION

Dedicated to
Asad's Parents, Siblings and Friends
&
Ali's Parents and Brothers
Whose prayers, affection, encouragement
and support made all the difference!
&
The Greatest Muhammad Ali 
Whose motivation kept us going!

ACKNOWLEDGEMENT

All praise be to **ALLAH** Almighty, the nourisher and the cherisher of the whole world, who has bestowed knowledge and wisdom endowed to all mankind. Countless salutations on the noble personage of **Hazrat Muhammad ﷺ**, and peace and mercy upon His noble companions' eminent members of family and true followers till the Day of Judgment.

We would like to express our sincere gratitude to our supervisor **Dr. Saira Amir** for her consistent support and cooperation, throughout the project and to **Khizra Jabeen** for providing guidance and feedback. We are thankful to our fellow students in the BS Bioinformatics program, especially **Baseerat Fatima, Muskan Raja, Nimra Bibi, Qurat Tul Ain, Sajad Ahmed, Sidra Kiani** and **Tooba Zahra** for their moral support and for the cherished time we spent together.

A special thanks to our family for their unconditional love, encouragement and belief in us. This thesis would not have been possible without their unwavering support and sacrifices. To conclude, we would like to thank everyone who has contributed to this thesis in any way.

Asad Zahid Malik

CIIT/SP21-BSI-011/ISB

&

Muhammad Ali Ejaz

CIIT/SP21-BSI-033/ISB

Abstract

Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized by motor and cognitive impairments. Monoamine oxidase B (MAO-B) inhibitors have emerged as effective therapeutic agents for managing PD symptoms by reducing dopamine degradation. This study employs a computational approach to identify and evaluate potential MAO-B inhibitors. Molecular docking was utilized to analyze the binding interactions of various compounds with MAO-B, focusing on key parameters such as binding affinity, van der Waals forces, and electrostatic interactions.

In addition, pIC50 predictions were performed to assess compound potency, complemented by statistical analyses of molecular descriptors such as LogP, molecular weight, and hydrogen bond acceptors. Frequency distribution and box plot visualizations highlighted significant differences between active and inactive compounds, with p -values ≤ 0.05 confirming statistical significance. Results revealed strong docking affinities for compounds like selegiline, along with clear trends correlating molecular properties to bioactivity.

This research provides insights into the structural and physicochemical characteristics that influence MAO-B inhibition, aligning with Lipinski's Rule of Five for drug-likeness. The findings emphasize the importance of computational methods in early drug discovery, offering a foundation for experimental validation and further development of PD therapeutics. Ultimately, this study underscores the potential of integrating molecular docking, statistical analysis, and bioactivity predictions to accelerate the identification of promising drug candidates.

Table of Contents

COMSATS University Islamabad.....	ii
BS Bioinformatics.....	ii
Final Approval.....	iii
Declaration	iv
Certificate	v
DEDICATION.....	vii
Abstract	ix
Table of Contents.....	x
List of Figures.....	xii
1.Introduction.....	15
2.1 Idiopathic Parkinson's Disease (IPD).....	16
2.2 Genetic Parkinson's Disease	18
• SNCA (Alpha-synuclein)	18
• LRRK2 (Leucine-rich repeat kinase 2)	18
• GBA (Glucocerebrosidase)	19
• VPS35 (Vacuolar protein sorting 35).....	20
• UCHL1 (Ubiquitin C-terminal hydrolase L1)	20
2.3 Autosomal Recessive (AR) Genes	21
• PRKN (Parkin).....	21
• PINK1 (PTEN-induced kinase 1)	22
• DJ1 (Parkinsonism-associated deglycase DJ1)	22
• ATP13A2.....	23
• PLA2G6 (Phospholipase A2 group VI)	23
1. RAB39B	24
3.Treatment of Parkinson's Disease	25
3.1 Pharmacological Treatments	25
3.2 Dopamine agonists	26

3.3 Monoamine oxidase-B (MAO-B) inhibitors,.....	26
3.4 COMT inhibitors.....	27
3.5 Emerging Therapies and Experimental Treatments.....	27
3.6 Surgical Treatment.....	28
3.7 Future Treatments for PD	29
4.1 Challenges in Diagnosing Parkinson's Disease	29
4.2 International Diagnostic Criteria for Parkinson's Disease	30
4.3 Diagnosis in Pre-Diagnostic Stages	30
4.4 Imaging Techniques in Parkinson's Diagnosis	31
4.5 Fluid and Tissue Biomarkers	31
Chapter 2.....	32
7. Role of Monoamine Oxidase B in Drug Discovery	35
7.2 Novel MAOB Inhibitors in Development.....	36
8.1 Introduction to Computational Drug Discovery.....	37
9.1 Definition and Importance of Molecular Descriptors.....	39
9.2 Types of Molecular Descriptors	39
9.3 Role of Molecular Descriptors in MAOB Inhibition.....	39
10.Machine Learning in Drug Discovery	41
10.1 Machine Learning Models for Predicting Drug Activity	41
12. Dataset Acquisition and Preprocessing	47
14.2 Model Validation	53
Calculated Molecular Descriptors – After running the PaDEL-	56
Why It Stands Out.....	57
1. Selegiline	70
2. Apigenin.....	73
3. 1-Phenylethylamine.....	75
Conclusion	80

List of Figures

Figure 1: Neurons in the substantia nigra.....	32
Figure 2: Interface of the app.....	52
Figure 3: Calling chembl database through websource client.....	57
Figure 4: Selecting relevant Chemb1 ID and importing bioactivity data of our target MOAB.....	57
Figure 5: Classifying them as active and inactive (≥ 10000 =inactive) (≤ 1000 =active)	58
Figure 6: Calculating Lipinski descriptors through canonical smiles.....	60
Figure 7: IC50 to pIC50	60
Figure 8: removing intermediate class	61
Figure 9: Converting our bioactivity data to Pubchem fingerprints by PaDel descriptors.....	61
Figure 10: Removing Low variance features before training the dataset	60
Figure 11: Random Forest Model Performance	61
Figure 12: Frequency of ac8ve and inac8ve compounds	62

Figure 13: pIC50 Value vs Bioac8vity Class (Box Plot)	
.....	62
Figure 14 logP vs Bioactivity class(Box Plot)	65
Figure 15 MW vs Bioactivity Class (Box Plot)	
.....	66
Figure 16 interpreting p-values	
.....	67
Figure 17: Docking results of selegiline and MOAB	
.....	69
Figure 18: Binding of MAOB & Selegiline	
.....	69
Figure 19: Selegiline pIC50 value of 7.9214	
.....	71
Figure 20: Apigenin Docking results	
.....	72
Figure 21: Apigenin & MOAB binding	
.....	72
Figure 22: Apigenin pIC50 value of 5.2	
.....	73
Figure 23: 1-Phenylethylamine docking results.....	75
Figure 24: Docking with MOAB.....	75
Figure 25: 1-Phenylethylamine pIC50 value of 4.8.....	77

Chapter 1

Introduction

1. Introduction

The hallmark features of pathology of Parkinson's disease (PD) include progressive neurodegenerative disorder characterized by the profound loss of dopaminergic cells in the basal ganglia, a part of the brain involved in movement control. The reduction in dopamine impairs the basal ganglia's ability to coordinate inhibitory and excitatory neural motor signals within cortical-subcortical circuits. The motor consequences of this dysfunction include rigidity, tremors, and dyskinesia. (Kwan & Whitehill, 2011)

Parkinson's disease (PD) is the second most common neurodegenerative disorder after Alzheimer's disease and affects more than 5 million cases worldwide (Olanow et al., 2009)

The frequency of Parkinson's disease is highest among Caucasians in Europe and North America, moderately affects Asians, especially in China and Japan, and is least common in Black populations across Africa.

(Marras and Tanner, 2004) . The risk of developing Parkinson's disease increases with age, with the disease typically emerging in individuals during their fifth or sixth decade of life. (Dorsey et al. 2007) The prevalence of PD in industrialized countries is generally estimated at 0.3% of the entire population and about 1% in people over 60 years of age (De Lau and Breteler 2006). However, a smaller proportion of cases occur in younger individuals, including those under 40 years of age, and are classified as early-onset Parkinson's disease. In rare instances, the condition can even manifest in adults as young as 20 years old. These early-onset cases are often associated with a family history of Parkinson's disease, where individuals may have a parent, sibling, or child who also suffers from the condition.

Up to 80% of the dopaminergic cells that produce dopamine in the brain are atrophied or degenerated before the symptoms of the PD start to emerge (Chung et al. 2001). One of the first symptoms is bradykinesia, which means a slow start to voluntary movements, along with a gradual decrease in the speed and range of repetitive actions. Other symptoms involve muscular rigidity, resting tremor or postural instability, are a

prerequisite for the diagnosis (Hughes et al. 1992). The aetiology of Parkinson's disease (PD) is multifactorial, and currently, no treatment is available that can halt or reverse the progression of the disease. However, various methods exist to manage its symptoms. Treatment plans can differ from person to person, depending on their specific symptoms and the effectiveness of different therapies. Medications are the primary approach to treating PD. Levodopa, a central nervous system agent, is converted into dopamine in the brain. Initially, the response to levodopa is generally positive. As the disease progresses and the brain's ability to store dopamine decreases, most patients experience a reduced duration of response to each dose (wearing-off symptoms), as well as involuntary movements (dyskinesias) and other motor complications. To manage these fluctuations, additional dopaminergic medications are used, including monoamine oxidase type B inhibitors, catechol-O-methyltransferase inhibitors, the NMDA receptor antagonist amantadine, and dopamine receptor agonists (Jankovic & Stacy, 2007). The most common surgical treatment for Parkinson's disease is deep brain stimulation (DBS). This treatment strategy is typically reserved for bradykinesia, rigidity and tremor in patients who no longer respond to medication in a predictable manner (Lee et al., 2018)

2.Types of Parkinson's Disease

2.1 Idiopathic Parkinson's Disease (IPD)

“Idiopathic” means that a health condition has no identifiable cause. Idiopathic Parkinson is the most common form of Parkinson which means that there's no known cause. Parkinson's disease has a known genetic cause in only 10–15% of people. These individuals often develop Parkinson's disease at an earlier age than those with IPD. According to the Parkinson's Foundation, it's difficult to know early on if an individual has IPD or if there's a known cause of parkinsonism due to their similar symptoms. (Pennington et al., 2010) . In IPD, nerve cells in the brain are damaged and expire. The buildup of the abnormal clumps of a protein called alpha-synuclein in Lewy bodies inside nerve cells is believed to

disrupt normal cellular function of the disrupted neurons which eventually lead the death of neurons. What causes Lewy bodies to form isn't known. The substantia nigra is the most commonly affected area of the brain in IPD. This part of the brain is important for controlling movement and contains many dopamine producing nerve cells. Dopamine is a neurotransmitter that helps control muscle tone and movement. A loss of dopamine results in the main symptoms of Idiopathic Parkinson's disease.

The symptoms of Idiopathic disease usually are:

1-Bradykinesia (slow movement)

2-Tremors (rhythmic shaking movements)

3-Rigidity (stiffness of the arms or legs)

4-Balance and coordination problems

5-Changes in speech

6-Muscle spasms or cramps

7-Trouble chewing or swallowing

8-Changes in posture

9-sleeping patterns

10-Issues with urination

The diagnosis of Idiopathic Parkinson's Disease (IPD) is primarily clinical, as no specific test can definitively confirm the condition. This is because several symptoms of Parkinson's disease can resemble those of other medical conditions. To establish a diagnosis, physicians focus on excluding other potential causes of the symptoms, such as side effects of medications or other neurological disorders. A comprehensive evaluation typically includes a detailed medical history, thorough physical and neurological examinations, and additional tests like blood work or imaging studies to rule out alternative explanations for the patient's symptoms.

2.2 Genetic Parkinson's Disease

Genetic Parkinson's Disease refers to forms of Parkinson's disease (PD) that arise due to inherited mutations in specific genes. Although these cases make up a smaller percentage of Parkinson's disease cases (about 5-10%), they offer important insights into the mechanisms behind PD and its progression. (Lunati et al., 2018) The cause of Parkinson's disease (PD) remains unclear for the majority of patients, although research has made significant progress since the identification of the first genetic mutation linked to PD in the SNCA gene in 1997. (Franco et al., 2021). Since then, many other genes with Mendelian inheritance patterns have been discovered to be associated with PD. These genes are categorized into three inheritance patterns: autosomal dominant (AD), autosomal recessive (AR), and X-linked

2.2.1- Autosomal Dominant (AD) Genes

- ***SNCA (Alpha-synuclein)***

SNCA gene is responsible for the production of the Alpha-synuclein the role of the Alphasynuclein is regulating neurotransmitter release and maintaining synaptic function in neurons. (Magistrelli et al., 2021)

Mutation in the SNCA causes the improper folding of the Alphasynuclein which leads to the toxic clumps, known as Lewy bodies. (Kim et al., 2014) These clumps accumulate in neurons, disrupting their normal function and leading to neurodegeneration. The accumulation of alpha-synuclein in dopaminergic neurons, especially in the substantia nigra, causes a reduction in dopamine production which causes early onset pd

- ***LRRK2 (Leucine-rich repeat kinase 2)***

LRRK2 encodes a protein called Leucine-rich repeat kinase 2, which plays a crucial role in signal transduction, vesicle trafficking, and autophagy. (Greggio & Cookson, 2009) Mutations in the LRRK2 gene are among the most common genetic causes of Parkinson's disease (PD), particularly in familial cases, though

they can also contribute to sporadic forms of the disease. The most common mutation, G2019S, leads to an increase in the kinase activity of the LRRK2 protein, disrupting its normal cellular functions. Abnormal LRRK2 can impair the lysosomal system and autophagy, both of which are responsible for degrading and recycling cellular waste and removing damaged or dead cells, respectively. This impairment results in the accumulation of toxic substances in neurons, contributing to neurodegeneration, especially in dopaminergic neurons of the substantia nigra. Additionally, LRRK2 mutations can affect mitochondrial function, leading to energy deficits in neurons and making them more vulnerable to damage. LRRK2 mutations are linked to both early-onset and late-onset Parkinson's disease.

- ***GBA (Glucocerebrosidase)***

GBA is a gene that encodes the enzyme glucocerebrosidase, which is crucial for breaking down a lipid called glucocerebroside in lysosomes. Lysosomes are cellular compartments responsible for degrading and recycling various substances, including lipids and proteins. Glucocerebrosidase specifically helps to degrade glucocerebroside into simpler components that can be reused by the cell. (Behl et al. 2021) When mutation occurs in the GBA gene, the enzyme becomes defective or inactive, impairing its ability to process glucocerebroside. As a result, lipids accumulate in neurons, particularly in the brain's substantia nigra, leading to cellular dysfunction and stress. Additionally, the defective GBA enzyme compromises the degradation of misfolded proteins, such as alpha-synuclein, a protein that tends to aggregate into toxic clumps known as Lewy bodies, a hallmark of Parkinson's disease (Alcalay et al. 2015)

- ***VPS35 (Vacuolar protein sorting 35)***

VPS35 was originally identified in yeast as a member of the retromer complex. This complex is involved in the intracellular trafficking of proteins (Seaman et al., 1998). It is part of the retromer complex, which is responsible for the recycling of proteins from endosomes back to the trans-Golgi network or the plasma membrane. (Seaman, 2012, Follett et al., 2014a, Trousdale and Kim, 2015). When mutations such as D620N occur in VPS35, it disrupts its ability to regulate endosomal trafficking and vesicular recycling. This leads to an accumulation of misfolded proteins within the cell and disrupts the sorting of proteins required for normal cellular function. In the context of Parkinson's disease (PD), this impairment in trafficking specifically affects dopaminergic neurons, contributing to their degeneration. As a result, the accumulation of damaged or improperly sorted proteins, combined with a failure to maintain cellular homeostasis, contributes to neurodegeneration, particularly in the substantia nigra, which is critical in the development of PD. (Vilarino-Güell et al. 2011)

- ***UCHL1 (Ubiquitin C-terminal hydrolase L1)***

UCHL1 (Ubiquitin C-terminal hydrolase L1) is essential for maintaining cellular balance by regulating the ubiquitin-proteasome system (UPS), which degrades damaged or misfolded proteins. It works by removing ubiquitin tags from proteins marked for degradation, allowing them to be processed by the proteasome. This process is crucial for clearing abnormal proteins, especially in neurons, which are vulnerable to protein accumulation. When mutations occur in the UCHL1 gene, the enzyme becomes dysfunctional, impairing its ability to process ubiquitinated proteins. This dysfunction hampers the ubiquitin-proteasome system, preventing it from efficiently degrading misfolded or damaged proteins. As a result, these abnormal proteins accumulate in neurons, particularly in the dopaminergic system, which is vulnerable in Parkinson's disease (PD).

The buildup of proteins like alpha-synuclein leads to toxic aggregates known as Lewy bodies, a hallmark of PD. These aggregates disrupt cell function, cause oxidative stress, and impair mitochondrial function, contributing to neuronal degeneration. In PD, defective UCHL1 accelerates the accumulation of damaged proteins, causing the death of dopaminergic neurons, which ultimately leads to motor symptoms like tremors, rigidity, and bradykinesia. (NG et al., 2020)

2.3 Autosomal Recessive (AR) Genes

- ***PRKN (Parkin)***

PRKN encodes a ubiquitin ligase, which is a crucial enzyme for maintaining mitochondrial quality control. This enzyme plays a critical role in a process called mitophagy, wherein damaged mitochondria are tagged for degradation and removal from the cell. Mitochondria are the powerhouses of the cell, and they are especially important for neurons due to their high energy demands. The process of mitophagy ensures that dysfunctional mitochondria, which can produce harmful reactive oxygen species (ROS) and contribute to oxidative stress, are removed. This removal is vital for cellular health, particularly in neurons, which are highly dependent on mitochondrial function to support neurotransmission and other cellular processes. However, when mutations occur in the *PRKN* gene, the enzyme's ability to mark damaged mitochondria for removal is impaired. As a result, dysfunctional mitochondria accumulate within the cell, causing an increase in oxidative stress and mitochondrial dysfunction. This leads to neuronal damage, and over time, neuronal death. The accumulation of dysfunctional mitochondria in neurons, particularly dopaminergic neurons in the brain's substantia nigra, is a major contributing factor to the pathophysiology of Parkinson's disease (PD). The progressive loss of dopaminergic neurons, responsible for dopamine production and motor control, results in the characteristic motor symptoms of PD, such as tremors, rigidity, bradykinesia, and postural instability (Arkinson et al., 2018).

- ***PINK1 (PTEN-induced kinase 1)***

PINK1 is a kinase that plays a critical role in maintaining mitochondrial health, particularly by regulating mitophagy. In a healthy cell, PINK1 is continuously imported into mitochondria, where it is rapidly degraded by the mitochondrial proteases. However, when mitochondria become damaged or dysfunctional, PINK1 accumulates on the outer mitochondrial membrane. This accumulation triggers the recruitment of Parkin, another protein responsible for tagging damaged mitochondria for degradation. PINK1's role in this process is crucial for protecting cells from mitochondrial dysfunction and oxidative damage. (Song et., 2013) When mutations occur in the PINK1 gene, it disrupts the kinase's ability to accumulate on the outer membrane of damaged mitochondria, impairing the recruitment of Parkin. Without this critical step, damaged mitochondria are not effectively removed, and they continue to accumulate in neurons. This buildup of dysfunctional mitochondria contributes to increased oxidative stress, which damages cellular components, particularly in dopaminergic neurons. As oxidative stress accumulates, it results in neuronal damage, and over time, this leads to the progressive degeneration of dopaminergic neurons, which is a hallmark of Parkinson's disease. This disruption in mitochondrial maintenance and clearance contributes significantly to the pathophysiology of PD.

- ***DJ1 (Parkinsonism-associated deglycase DJ1)***

DJ1 is a multifunctional protein that plays an important role in protecting cells from oxidative stress. It is known to help maintain mitochondrial function and is involved in various cellular processes, such as the stabilization of cellular proteins and the removal of harmful chemical modifications from proteins. DJ1 also functions as a deglycase, which means it helps to remove toxic chemical modifications from proteins, a process essential for maintaining protein homeostasis within the cell. Mutations in the DJ1 gene impair its protective functions, leaving cells, especially dopaminergic neurons, vulnerable to oxidative damage and mitochondrial dysfunction. These impairments lead to an accumulation of

harmful molecules and toxic proteins within the cell. Over time, this damage leads to neuronal dysfunction, and as the degeneration of dopaminergic neurons progresses, the hallmark symptoms of Parkinson's disease, such as tremors, bradykinesia, and rigidity, manifest. DJ1's involvement in oxidative stress response makes it an important protein in maintaining the health of neurons and preventing the development of neurodegenerative diseases like Parkinson's disease. (Richarme and Diaoru, 2017)

- ***ATP13A2***

ATP13A2 encodes a lysosomal membrane protein that plays a key role in the transport of various ions across the lysosomal membrane and the clearance of cellular waste products. It is involved in maintaining lysosomal function, which is essential for the degradation of toxic substances within the cell. The proper functioning of lysosomes is critical for neurons, as they rely on the degradation of cellular waste, such as misfolded proteins and damaged organelles, to maintain cellular health. Mutations in the ATP13A2 gene disrupt lysosomal function, impairing the degradation of cellular waste products, including toxic proteins. This leads to the accumulation of undigested substances within the cell, contributing to neuronal dysfunction.

In Parkinson's disease, the loss of ATP13A2 function is particularly harmful to dopaminergic neurons, which are especially vulnerable to the buildup of toxic substances. The accumulation of cellular waste, particularly toxic proteins like alpha-synuclein, contributes to neuronal degeneration, further accelerating the onset and progression of Parkinson's disease, particularly in early-onset forms of the disease. (Park et al. 2015)

- ***PLA2G6 (Phospholipase A2 group VI)***

PLA2G6 encodes an enzyme that plays a crucial role in lipid metabolism, particularly in the breakdown of phospholipids within cellular membranes.

This enzyme is involved in the production of bioactive lipids that participate in various cellular processes, including inflammation, membrane repair, and the maintenance of membrane integrity. PLA2G6 is essential for neuronal function and survival, as it helps regulate membrane fluidity and signal transduction in the brain. Mutations in the PLA2G6 gene impair the enzyme's activity, leading to a disruption in lipid metabolism and the accumulation of defective lipids in neuronal membranes. This disruption compromises membrane integrity, affecting cell signaling, synaptic function, and overall neuronal health. The accumulation of damaged lipids and impaired cellular membrane integrity results in neuronal dysfunction, particularly in dopaminergic neurons, which are especially vulnerable to these metabolic disturbances. Over time, this leads to neurodegeneration and the formation of toxic aggregates that disrupt normal cellular processes. PLA2G6 mutations are associated with a rare form of Parkinson's disease known as neurodegeneration with brain iron accumulation (NBIA), which typically presents with early-onset symptoms, including motor dysfunction, dystonia, and Parkinsonism. This form of PD is characterized by the accumulation of iron in the brain, which further exacerbates neuronal damage and contributes to the progression of the disease.

(Guo et al. 2018)

2.4- X-linked Genes:

1. RAB39B

RAB39B is a small GTPase involved in vesicle trafficking, which plays a key role in the transport of proteins and other cellular components within neurons. It regulates the movement of synaptic vesicles and cargo, ensuring proper communication and function within the cell, particularly in neurons that require efficient vesicle transport for neurotransmitter release and synaptic plasticity. (Gao et al. 2020) RAB39B is essential for maintaining the integrity of synaptic vesicles and supporting normal neuronal function.

Mutations in the RAB39B gene impair its trafficking function, leading to disrupted vesicle dynamics in dopaminergic neurons. This dysfunction causes the accumulation of mis localized or damaged proteins, contributing to cellular stress, oxidative damage, and mitochondrial dysfunction. These disruptions accelerate neuronal degeneration, particularly in the substantia nigra, the region responsible for dopamine production, which is heavily affected in Parkinson's disease (PD). (Gao et al. 2020)

In women, RAB39B mutations typically lead to a more typical PD phenotype, with symptoms resembling those of idiopathic Parkinson's disease, such as motor dysfunction. In men, however, the disease presentation is more variable, potentially leading to differences in disease progression or severity. This variability highlights the complex role of RAB39B in PD, with gender-specific factors influencing the severity of neuronal degeneration. (Mata et al. 2015)

3.Treatment of Parkinson's Disease

Parkinson's Disease (PD) is a progressive neurodegenerative disorder that significantly impacts patients' quality of life and their ability to perform daily activities. Despite advances in symptom management, there is currently no cure for PD, highlighting the need for novel neuroprotective drugs and therapeutic strategies. Treatments for PD are typically tailored to patients' symptoms and needs, encompassing pharmacological, rehabilitative, and surgical options. Additionally, the identification of distinct PD subtypes through cluster analysis has facilitated personalized therapeutic approaches based on motor and non-motor symptom presentations (Lauretani et al., 2014).

3.1 Pharmacological Treatments

The cornerstone of PD treatment is levodopa (L-dopa), the metabolic precursor of dopamine. Levodopa is converted to dopamine by dopa-decarboxylase in dopaminergic neurons. To prevent peripheral decarboxylation, levodopa is often coadministered with carbidopa or benserazide, which are peripheral dopa-decarboxylase inhibitors that do

not cross the blood-brain barrier (BBB). This combination enhances levodopa's bioavailability in the central nervous system (CNS) (Goldenberg, 2008). While levodopa is effective in managing PD symptoms, its long-term use is associated with complications such as dyskinesias. A notable limitation of levodopa therapy arises from its conversion by serotonergic neurons, which can result in reduced efficacy and increased side effects (De Duerwaerdère et al., 2017). Recent innovations include Rytary—a capsule containing levodopa, carbidopa, and entacapone—and Duopa, an intestinal gel formulation of levodopa and carbidopa designed to prolong therapeutic effects and reduce motor fluctuations. Additionally, Safinamide, a reversible MAO-B inhibitor with antiglutamatergic properties, has shown effectiveness in enhancing motor symptom control and mitigating the 'wearing-off' phenomenon (Olanow et al., 2014; Reiter, 1995).

3.2 Dopamine agonists

dopamine agonists mimic dopamine's effects by binding to post-synaptic dopamine receptors. Common agents in this category include pramipexole, ropinirole, apomorphine, rotigotine, and piribedil. These drugs are often employed as initial monotherapy to delay motor complications associated with levodopa (Jankovic and Aguilar, 2008). Rotigotine, available as a transdermal patch, provides continuous drug delivery, improving symptom control, while apomorphine, administered via subcutaneous infusion, is effective in reducing motor fluctuations. However, dopamine agonists can cause side effects such as nausea, daytime sleepiness, and impulse control disorders. These adverse effects are more frequent in elderly patients with cognitive impairments, making careful patient selection crucial.

3.3 Monoamine oxidase-B (MAO-B) inhibitors,

Monoamine oxidase-B (MAO-B) inhibitors, such as selegiline and rasagiline, play an essential role in managing PD by preventing dopamine degradation and reducing oxidative stress. Selegiline has demonstrated levodopa-sparing effects and may slow disease progression (Reiter, 1995;

MONOCOMB study). Rasagiline is effective as an add-on therapy in advanced PD, helping reduce motor fluctuations and levodopa dependency. Safinamide, a newer MAO-B inhibitor, offers enhanced motor control and improves quality of life by reducing ‘wearing-off’ symptoms (Reiter, 1995).

3.4 COMT inhibitors

COMT inhibitors including entacapone and tolcapone, work by blocking levodopa’s peripheral metabolism, thereby increasing its bioavailability and half-life. Recent advancements have introduced nebicapone, which has shown superior efficacy compared to entacapone with fewer side effects, and opicapone, administered as a once-daily oral dose. These options significantly reduce OFF periods and extend ON periods without troublesome dyskinésias, improving patient outcomes in advanced PD (Goldenberg, 2008).

3.5 Emerging Therapies and Experimental Treatments

The formation of protein aggregates, such as alpha-synuclein (α -syn), is a hallmark of PD pathology. Therapies targeting this process are under active investigation. Experimental drugs such as NPT200-11 inhibit α -syn oligomerization, protecting neurons from toxic aggregates (Oertel, 2017). Similarly, ANLE138b blocks the conversion of α -syn monomers to toxic oligomers and fibrils, while squalamine, a natural product, prevents α -syn aggregation by displacing it from lipid vesicles, reducing its toxicity (Perni et al., 2017).

Rapamycin, an mTOR inhibitor, induces autophagy and facilitates the clearance of toxic protein aggregates. This mechanism is particularly promising for halting PD progression (Hochfeld et al., 2013). Additionally, adenosine A2A receptor antagonists, such as caffeine, have shown neuroprotective effects, reducing the risk and progression of PD. Transgenic models with mutant α -syn and deleted adenosine A2A receptor genes demonstrate reduced PD-related neurodegeneration (Kachroo and Schwarzschild, 2012). Non-selective PDE inhibitors, such as

dipyridamole, have shown potential in preclinical trials by protecting dopaminergic neurons from α -syninduced degeneration (Höllerhage et al., 2017).

3.6 Surgical Treatment

Deep brain stimulation (DBS) of either the subthalamic nucleus (STN) or globus pallidus interna (GPi) is a well-known treatment for patients with motor complications. For treatment of tremors, thalamic DBS is a viable option. Surgical treatment is preferred when motor fluctuations and dyskinesias become disabling despite responsiveness of the motor symptoms to levodopa. The average time before DBS is performed is about 10–13 years after the diagnosis of Parkinson's disease has been established. Findings of the EARLYSTIM trial, a multicenter randomized control trial, showed that DBS in the early course of disease (mean disease duration 7.5 years, with motor fluctuations for less than three years) could improve the patient's quality of life and several secondary outcome measures more than the best medical therapy.

DBS is reversible and can be adjusted for disease progression. The presence of dementia, acute psychosis, and major depression are exclusion criteria for DBS.

Bilateral DBS of the STN improves the UPDRS II (activities of daily living) and UPDRS III (motor) scores, on average, by 50–60% compared with the preoperative medical OFF state. The total daily dopaminergic drug dosage is reduced by about 60% following the institution of DBS, and dyskinesias decrease by 60–70%. Subthalamic nucleus (STN) DBS was associated with decreased requirement of levodopa doses. The mortality of DBS is less than 0.5%, and the important adverse events include intracranial bleeding or device-related complications (such as infections and lead misplacements, among others). Non-pharmacological therapies available for PD include exercise, education, support groups, speech therapy, and nutrition. Evidence from the literature recommends their usage early in the course of the disease.

3.7 Future Treatments for PD

Accurate diagnosis of Parkinson's Disease still relies largely on clinical acumen due to the lack of definitive diagnostic laboratory tests. Early PD has a heterogeneous presentation and an insidious onset, which makes clinical diagnosis prone to errors. A valid diagnostic biomarker for PD will be critical in improving patient outcomes on several accounts: improving diagnostic accuracy, especially in early stages; enabling the use of neuroprotective strategies; and better defining diagnostic subtypes for tailored treatment approaches. Biomarker research holds the promise of advancing early detection and personalized interventions in PD, ultimately transforming the clinical management and prognosis of the disease.

4.1 Challenges in Diagnosing Parkinson's Disease

Diagnosing idiopathic Parkinson's disease (PD) is relatively straightforward in cases with a classic history and typical asymmetric motor signs. However, in clinical practice, diagnostic errors are common, with misclassification rates ranging from 15% to 24%. Even with stringent criteria, approximately 10% of PD diagnoses are incorrect, as alternative pathologies are often present. A significant challenge is distinguishing PD from atypical parkinsonian disorders early in the disease course.

Atypical parkinsonism syndromes, such as multiple system atrophy (MSA), progressive supranuclear palsy (PSP), and corticobasal degeneration (CBD), share overlapping features with PD. These syndromes differ in progression and prognosis but are difficult to distinguish clinically in early stages. Misclassification rates in these conditions range from 7% to 35%. The clinical overlap underscores the need for improved diagnostic tools.

4.2 International Diagnostic Criteria for Parkinson's Disease

To enhance diagnostic accuracy, the International Parkinson and Movement Disorder Society (MDS) developed criteria that refine the Queen Square Brain Bank (QSBB) guidelines. These criteria require evidence of bradykinesia combined with rigidity or asymmetric resting tremor. Supportive features and red flags help distinguish PD from atypical disorders.

The MDS criteria define two levels of diagnostic certainty: clinically established and clinically probable. While the former prioritizes specificity over sensitivity, the latter balances sensitivity and specificity. Validation studies show these criteria achieve up to 96% sensitivity and 98.5% specificity. Ancillary tests incorporated into these criteria aim to improve diagnostic precision, particularly in early disease stages.

4.3 Diagnosis in Pre-Diagnostic Stages

Parkinson's disease likely begins well before clinical symptoms appear. However, no reliable biomarkers currently exist to diagnose pre-diagnostic or prodromal stages with high sensitivity and specificity. Identifying individuals at risk remains critical for understanding disease progression and enabling early intervention.

Features associated with increased risk include a positive family history, genetic mutations, and non-motor symptoms such as REM sleep behavior disorder (RBD). RBD has been strongly linked to PD, with over 90% of cases progressing to neurodegenerative parkinsonism. Other indicators, such as olfactory dysfunction and abnormalities on DAT-SPECT imaging, can identify those nearing clinical conversion. Prospective studies and diagnostic algorithms, including online tools like the PREDICTPD framework, are being developed to estimate PD risk. However, these tools lack sufficient sensitivity and predictive value, highlighting the urgent need for robust biomarkers.

4.4 Imaging Techniques in Parkinson's Diagnosis

Structural magnetic resonance imaging (MRI) plays a key role in distinguishing PD from secondary or atypical parkinsonian syndromes. Advanced MRI techniques, such as neuromelanin imaging (NMI) and quantitative susceptibility mapping (QSM), allow visualization of nigral pathology. Studies show NMI achieves over 80% sensitivity and specificity in detecting PD.

Radionuclide imaging, particularly DAT-SPECT, is widely used to examine dopaminergic function and differentiate PD from non-degenerative parkinsonism. However, it is less effective in distinguishing PD from other neurodegenerative parkinsonisms. Emerging radiotracers targeting alpha-synuclein deposits may enhance early detection but are not yet clinically available.

4.5 Fluid and Tissue Biomarkers

Alpha-synuclein remains a key focus in biomarker research. Biopsy studies suggest that assessing phosphorylated alpha-synuclein in tissues like salivary glands or skin could help diagnose prodromal PD. Advanced in vitro assays, such as Real-Time Quaking Induced Conversion (RT-QuIC), detect alpha-synuclein seeding activity with over 90% accuracy. These assays also show promise in distinguishing different synucleinopathies, such as PD and MSA, based on distinct alpha-synuclein strains. However, their clinical application is still under investigation.

Chapter 2

Literature Review

5. Role of Monoamine Oxidase B in Parkinson's Disease

Monoamine Oxidase B (MAOB) is an enzyme that plays a crucial role in the metabolism of neurotransmitters in the brain. Specifically, MAOB is responsible for the oxidative deamination of biogenic amines, including dopamine, which is essential for proper motor control and cognitive function. Dopamine is a key neurotransmitter in the central nervous system, particularly within the dopaminergic pathways that regulate movement and reward processing. In healthy individuals, MAOB helps maintain a balance in dopamine levels, preventing excess accumulation, which could lead to neurotoxicity. However, in the context of Parkinson's disease (PD), the role of MAOB becomes more complex and detrimental. (Tan,2022)

Parkinson's disease is characterized by the progressive degeneration of dopaminergic neurons in the substantia nigra, a region of the brain critical for motor control. (Figure 1) As these neurons die, the brain's ability to produce dopamine is severely impaired, resulting in the hallmark symptoms of PD, such as tremors, rigidity, and bradykinesia

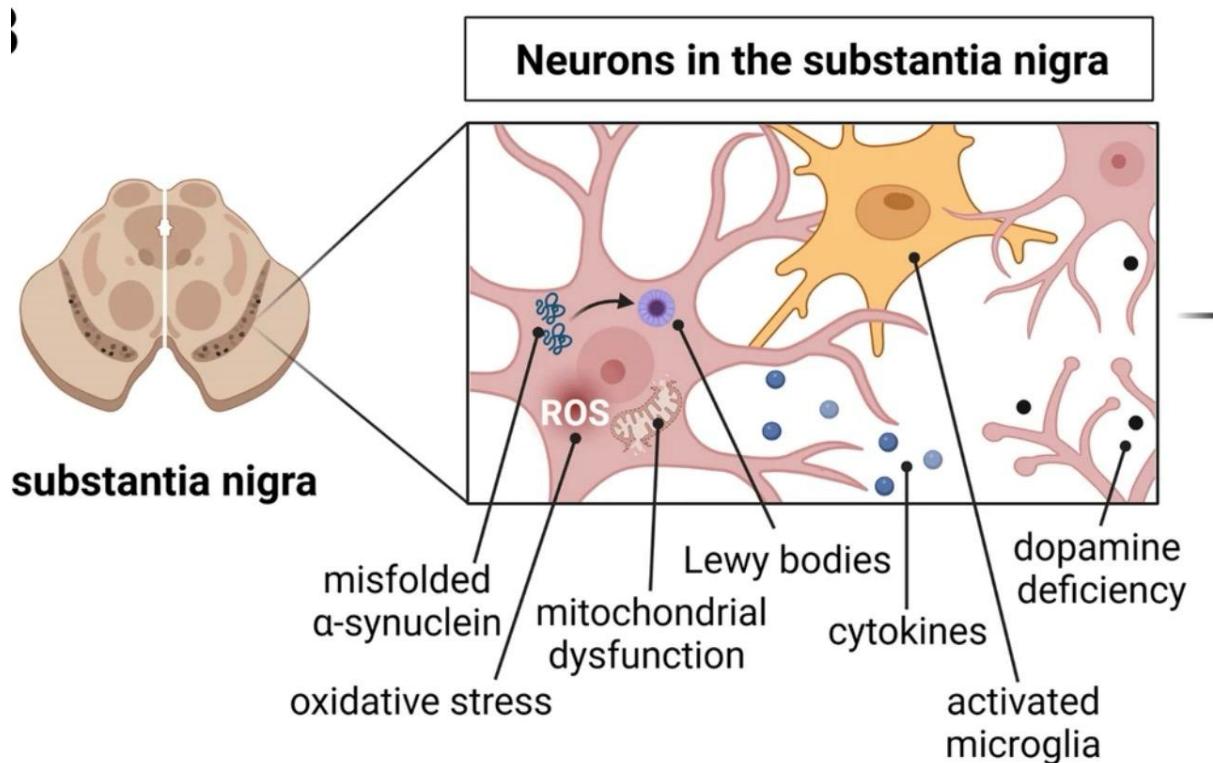


Figure 1: Neurons in the substantia nigra

One of the key contributors to the progression of the disease is the increased activity of MAOB in the remaining dopaminergic neurons. When MAOB activity is elevated, it leads to the accelerated breakdown of dopamine in the brain, further exacerbating the dopaminergic deficit. This process accelerates the depletion of available dopamine and contributes to the worsening of motor symptoms in PD patients.

The role of MAOB in the pathophysiology of Parkinson's disease has made it an important therapeutic target. By inhibiting the activity of MAOB, researchers have been able to slow the progression of the disease and alleviate some of the debilitating symptoms associated with dopamine depletion. MAOB inhibitors, such as selegiline and rasagiline, are currently used in the treatment of PD. These inhibitors work by preventing the breakdown of dopamine, thus preserving dopamine levels in the brain and helping to manage motor symptoms. Beyond simply preserving dopamine, some MAOB inhibitors may also offer neuroprotective effects, helping to reduce oxidative stress and mitigate the neurodegeneration that characterizes PD. (Robaskis,2015) Furthermore, there is growing evidence that MAOB inhibitors may have broader benefits beyond symptomatic relief. They may slow the rate of disease progression by preventing the accumulation of toxic byproducts from dopamine metabolism, such as hydrogen peroxide, which can lead to further neuronal damage. While MAOB inhibitors are not a cure for Parkinson's disease, they represent a promising therapeutic approach to managing the disease and improving the quality of life for patients. Ongoing research into the role of MAOB in Parkinson's disease continues to inform the development of new and more effective therapies aimed at slowing the progression of the disease and preserving neuronal function. (Chew,2021)

6. Challenges in Parkinson's Disease Drug Development

Despite the availability of these symptomatic treatments, PD remains an incurable disease. One of the main challenges in PD drug development is the limited efficacy of current therapies in addressing disease progression. Levodopa, while effective initially, does not prevent neuronal degeneration. Moreover, the side effects associated with long-term

levodopa use underscore the need for alternative therapies. Research is increasingly focusing on neuroprotective agents, gene therapies, and drugs that target underlying disease mechanisms, including oxidative stress and mitochondrial dysfunction. (Pires,2021)

7. Role of Monoamine Oxidase B in Drug Discovery

7.1 Mechanism of Action of MAOB Inhibitors

Monoamine Oxidase B (MAOB) inhibitors play a crucial role in the treatment of

Parkinson's disease (PD) by targeting the enzyme responsible for the breakdown of dopamine, a neurotransmitter that is critically depleted in PD patients. MAOB catalyzes the oxidative deamination of dopamine, converting it into inactive metabolites, including hydrogen peroxide, which can be neurotoxic and contribute to the degeneration of dopaminergic neurons. MAOB inhibitors work by blocking this enzyme, preventing the breakdown of dopamine and thus increasing its availability in the synaptic cleft. This increase in dopamine levels helps alleviate the motor symptoms of Parkinson's disease, such as tremors, rigidity, and bradykinesia.

In addition to their primary role in increasing dopamine levels, MAOB inhibitors may have potential neuroprotective effects. By inhibiting the activity of MAOB, these compounds help reduce the accumulation of harmful metabolites, such as hydrogen peroxide and aldehydes, which are byproducts of dopamine metabolism. The accumulation of these neurotoxic metabolites can lead to oxidative stress, exacerbating the degeneration of dopaminergic neurons. Therefore, by minimizing oxidative damage, MAOB inhibitors may not only provide symptomatic relief but also slow the progression of neurodegeneration, offering long-term benefits beyond just controlling the symptoms of PD. (Finberg, 2014)

Currently, several MAOB inhibitors are approved for clinical use, with selegiline and rasagiline being the most well-known. These drugs have been shown to improve motor symptoms in PD patients, often in

combination with levodopa, and provide modest neuroprotective effects. Selegiline, for instance, has been used for decades and has demonstrated efficacy in delaying the need for levodopa therapy, while rasagiline, a more selective MAOB inhibitor, has shown potential for slowing disease progression, making it an important tool in managing PD.

7.2 Novel MAOB Inhibitors in Development

Despite the success of current MAOB inhibitors, there remains a significant need for the development of more potent, selective, and safer compounds. Existing MAOB inhibitors, while effective, have limitations, including side effects and interactions with other medications, especially in long-term use. Therefore, research efforts continue to identify novel MAOB inhibitors that offer improved pharmacokinetic properties, better selectivity for the MAOB isoenzyme, and a more favorable side effect profile. Novel MAOB inhibitors are being explored through a combination of traditional drug discovery approaches and advanced computational techniques. High-throughput screening, virtual screening, and molecular docking are increasingly employed to identify potential compounds that can bind more effectively to MAOB, improving their potency and selectivity. These computational methods allow researchers to predict the binding affinity of new molecules, helping to streamline the drug discovery process. Moreover, by targeting specific regions of the MAOB enzyme, researchers aim to develop inhibitors that minimize off-target effects and improve the safety of these treatments. (Pires et al., 2017)

Furthermore, efforts are being made to design MAOB inhibitors that can cross the blood-brain barrier more efficiently, ensuring that they can act directly in the central nervous system. Some novel compounds under investigation also have the potential to combine MAOB inhibition with other mechanisms, such as anti-inflammatory or antioxidant activity, to provide a more comprehensive approach to treating Parkinson's disease. The development of these novel compounds is crucial for improving the treatment of PD, offering hope for more effective therapies that not only alleviate symptoms but also modify the disease course.

8. Computational Approaches in Drug Discovery

8.1 Introduction to Computational Drug Discovery

Computational drug discovery is an advanced and efficient approach that leverages computer-based techniques to predict and optimize drug candidates before experimental validation. This methodology uses simulations to model how drug molecules interact with biological targets, allowing researchers to explore their potential therapeutic effects in silico. The primary advantage of computational drug discovery is its ability to predict the activity, toxicity, and efficacy of compounds rapidly, thus, reducing the time and costs associated with traditional drug development processes. (Smith et al., 2018)

Several computational methods are commonly used in drug discovery, including virtual screening, molecular docking, and Quantitative Structure-Activity Relationship (QSAR) models. Virtual screening involves scanning large compound libraries to identify molecules that are likely to bind to a specific target, such as an enzyme or receptor. Molecular docking simulates the interaction between a drug molecule and its biological target, helping to determine how well a compound binds to the target site and assessing its binding affinity. QSAR models, on the other hand, use statistical techniques to relate the chemical structure of compounds to their biological activity, providing insights into how structural changes can improve the effectiveness of drug candidates.

These computational techniques not only aid in the identification of new drug candidates but also help in optimizing existing ones, making them more effective and safer before moving on to in vitro or in vivo testing. This pre-emptive evaluation allows researchers to focus their experimental resources on the most promising compounds. (Ou-Yang et al., 2012)

8.2 Applications of Computational Methods in Parkinson's Disease

In the context of Parkinson's disease (PD), computational approaches are particularly useful for identifying and developing inhibitors of Monoamine Oxidase B (MAOB), a key enzyme involved in the degradation of dopamine. Since PD is characterized by a deficiency of dopamine in the brain, inhibiting MAOB can help increase dopamine levels and alleviate symptoms. Computational drug discovery allows researchers to rapidly identify potential MAOB inhibitors from large compound libraries, including those found in databases like ChEMBL, which provides bioactivity data for small molecules.

By using computational techniques like virtual screening and molecular docking, researchers can simulate the binding of various compounds to MAOB, assessing them binding affinity and optimizing their structure to improve efficacy and selectivity. This helps identify the most promising candidates for further experimental validation. Additionally, computational methods can assist in the design of structure-based drug discovery by revealing key binding sites on the MAOB enzyme, guiding the creation of new inhibitors with improved potency and reduced side effects.

Machine learning algorithms are also playing an increasingly important role in PD drug discovery. These algorithms can be trained using data from previously successful MAOB inhibitors to predict the activity of new compounds, further accelerating the identification of potential drug candidates. Moreover, machine learning can help analyse large datasets from virtual screening and docking studies, automating the identification of patterns and optimizing the selection of compounds for further development.

Overall, computational drug discovery not only accelerates the process of identifying novel therapies for Parkinson's disease but also offers a more efficient and cost-effective approach to developing drugs that may slow or halt the progression of this debilitating condition.

9. Molecular Descriptors and Their Use in Drug Discovery

9.1 Definition and Importance of Molecular Descriptors

Molecular descriptors are quantitative representations of the chemical structure of a molecule. They describe various physicochemical properties, such as molecular size, polarity, and hydrophobicity, which are crucial for predicting the biological activity of compounds. In drug discovery, molecular descriptors are used to build models that can predict the activity of compounds against specific targets. (Moriwaki et al., 2018) Descriptors provide a bridge between chemical structure and biological activity, allowing researchers to prioritize compounds with the highest likelihood of success.

9.2 Types of Molecular Descriptors

Molecular descriptors can be categorized into two broad groups: global descriptors, which provide information about the overall structure of the molecule (e.g., molecular weight, logP), and topological descriptors, which describe the arrangement of atoms and bonds within the molecule. Descriptors like lipophilicity (logP), hydrogen bond donors and acceptors, and molecular weight are commonly used in predictive models to assess the drug-likeness of compounds. These properties are particularly important for assessing the pharmacokinetic behavior of drugs, such as their absorption, distribution, metabolism, and excretion (ADME) profiles. (Moriwaki et al., 2018)

9.3 Role of Molecular Descriptors in MAOB Inhibition

In the context of Monoamine Oxidase B (MAOB) inhibition, molecular descriptors play a crucial role in identifying compounds that are likely to effectively bind to the MAOB enzyme and inhibit its activity. These descriptors are numerical representations of the physicochemical properties of compounds, such as molecular shape, charge distribution, size, and hydrophobicity. These properties are critical as they influence how well a compound can interact with the MAOB active site, which is essential for inhibiting the enzyme's activity. (Murugan et al., 2020)

Computational techniques, such as molecular docking and Quantitative Structure Activity Relationship (QSAR) modeling, are frequently used to correlate these molecular descriptors with the inhibitory activity of compounds. Molecular docking simulates the binding of compounds to the MAOB enzyme, providing insight into the strength and orientation of the interaction. QSAR models, on the other hand, use statistical analysis to relate the molecular descriptors of compounds to their biological activity, allowing for the prediction of the inhibitory potential of new compounds. By analyzing the molecular descriptors in conjunction with experimental data, researchers can identify compounds with high potential for MAOB inhibition. This computational approach helps in the efficient screening of large compound libraries, ultimately accelerating the discovery of novel MAOB inhibitors for Parkinson's disease treatment.

9.4 Application of Lipinski's Rule in Parkinson's Disease Drug Discovery

Lipinski's Rule of Five is a vital tool in Parkinson's disease drug discovery, especially for developing orally bioavailable compounds. This rule outlines key physicochemical properties that increase the likelihood of a compound being an effective oral drug. Many existing MAOB inhibitors, such as selegiline and rasagiline, adhere to these guidelines, which helps ensure their suitability for oral administration. By screening for compounds that meet Lipinski's criteria—such as molecular weight, lipophilicity ($\log P$), and hydrogen bonding potential—researchers can eliminate compounds with poor pharmacokinetic properties early in the drug discovery process. This step is crucial in optimizing drug-likeness, ensuring that only those compounds with favorable absorption, distribution, metabolism, and excretion (ADME) profiles proceed to further stages of development. Ultimately, Lipinski's Rule aids in identifying promising MAOB inhibitors that are not only effective but also practical for clinical use in treating Parkinson's disease.

10. Machine Learning in Drug Discovery

Machine learning (ML) has emerged as a powerful tool in drug discovery due to its ability to analyze large volumes of complex data and identify patterns that may not be immediately apparent. At its core, machine learning involves the use of algorithms to create models that can predict outcomes based on input data. In drug discovery, these models can predict the biological activity of compounds, their potential toxicity, and even their pharmacokinetic properties. (Dara et al., 2022) The key advantage of machine learning is its ability to process and learn from vast datasets, allowing researchers to predict the efficacy and safety of compounds before experimental testing. By leveraging machine learning, researchers can identify promising drug candidates more efficiently, reducing both the time and cost involved in traditional drug development processes. (Vamathevan et al., 2019)

10.1 Machine Learning Models for Predicting Drug Activity

Machine learning algorithms, such as Random Forests, Support Vector Machines (SVM), and Neural Networks, have gained significant attention in drug discovery for predicting the biological activity of compounds.

These models are trained using datasets that consist of known compounds along with their molecular descriptors and corresponding biological activities. The training process enables the models to recognize patterns within the data and make predictions about new, untested compounds. For example, in predicting MAOB inhibition, molecular descriptors like molecular weight, logP, and hydrogen bonding properties serve as inputs for the machine learning model, which learns the relationship between these descriptors and the observed biological activity (e.g., IC₅₀ values).

(Moriwaki et al., 2018) Random Forests, a popular ensemble learning method, are particularly effective in handling high-dimensional data, such as molecular descriptors, and can provide valuable insights into which features are most important for predicting activity. Support Vector Machines (SVM) are another powerful tool in classification tasks, particularly for identifying whether a compound is likely to be an active

inhibitor or not. Neural Networks, on the other hand, can model complex relationships and handle large datasets, making them well-suited for predicting the activity of novel compounds. These machine learning models are not only able to predict the activity of compounds with high accuracy but also provide valuable insights into the underlying patterns that drive biological interactions.

10.2 Machine Learning Applications in Parkinson's Disease

In Parkinson's disease (PD) drug discovery, machine learning has become an essential tool for identifying new inhibitors of MAOB, a key enzyme involved in the disease's pathophysiology. Using large datasets that contain molecular descriptors and their corresponding biological activities, researchers can train machine learning models to predict the inhibitory potential of new compounds. By focusing on compounds that are likely to bind effectively to the MAOB active site, these models help streamline the discovery process, reducing the reliance on time-consuming and costly experimental screening. (Dara et al., 2022)

For Parkinson's disease, machine learning models are trained on datasets of known MAOB inhibitors, which include information such as SMILES notation, molecular weight, logP, and IC₅₀ values. Once trained, the models can predict the activity of novel compounds, helping researchers prioritize which compounds to move forward with. This is particularly useful when screening large compound libraries, where manually testing each compound would be prohibitively time-consuming and expensive. The integration of machine learning with computational techniques such as virtual screening and molecular docking further enhances the predictive power of these models, providing a comprehensive approach to drug discovery.

Machine learning is also used to optimize existing compounds by identifying structural modifications that could improve their activity or pharmacokinetic properties. By predicting the effects of these modifications on MAOB inhibition, researchers can create new compounds with enhanced efficacy and safety profiles. The ability to

rapidly predict and evaluate the potential of new drug candidates is revolutionizing the drug discovery process for Parkinson's disease, accelerating the identification of novel therapeutic agents. Ultimately, machine learning is helping to bridge the gap between computational predictions and experimental validation, speeding up the development of effective treatments for Parkinson's disease.

11. Molecular Docking: An Overview

Molecular docking is a computational technique used to predict the interaction between a small molecule (ligand) and a biological target, such as a protein or enzyme. It aims to simulate the binding process by determining the optimal orientation, conformation, and position of a ligand within the binding pocket of a target molecule. Docking evaluates the strength of these interactions by calculating binding affinity and other energy-based parameters, such as van der Waals (vdW) forces and electrostatic energies.

This method is particularly important in understanding molecular interactions at an atomic level, helping researchers identify key binding sites and stabilizing forces. Docking also provides insights into the strength and specificity of ligand-target interactions, making it an indispensable tool in modern drug discovery. By simulating these interactions in silico, docking reduces the need for expensive and timeconsuming experimental procedures during the initial stages of drug development.

11.1 Role of Docking in Drug Discovery

Docking plays a critical role in drug discovery by enabling researchers to identify and optimize potential drug candidates. It helps screen large libraries of compounds to identify those with strong binding affinities and favorable interaction profiles. This process saves considerable time and

resources by prioritizing compounds with a high likelihood of success in subsequent experimental and clinical stages.

In the context of drug discovery, docking evaluates how well a ligand fits into the binding pocket of a target protein. Key outputs include binding affinity, which reflects the strength of the interaction, and specific energy components such as vdW and electrostatic energies, which indicate the forces stabilizing the ligand-protein complex. This detailed analysis helps predict the efficacy of potential drugs and provides a molecular-level understanding of their mechanisms of action.

Docking is particularly valuable in the early stages of drug discovery, as it allows for the rapid identification of lead compounds. It also aids in structure-based drug design by revealing which molecular features enhance or hinder binding. By focusing on compounds with favorable docking scores, researchers can refine and optimize chemical structures to improve their pharmacological properties.

11.2 Affinity Score in Docking

The affinity score measures the binding energy between a ligand and its target protein. It is typically expressed in kcal/mol and indicates the strength of the binding interaction. A more negative affinity score reflects a stronger binding, suggesting that the ligand forms stable interactions within the binding site.

In drug discovery, affinity scores are critical for evaluating the potential effectiveness of a compound as a drug candidate. Compounds with highly negative scores are prioritized, as they are more likely to exhibit strong inhibitory effects on the target protein. This metric is a key determinant in the selection and optimization of lead compounds during the docking process.

11.3 Total Energy in Docking

Total energy represents the overall energy involved in the docking interaction, combining contributions from both attractive and repulsive forces. Positive total energy indicates that some energy is required to stabilize the binding, while negative values suggest that the binding process is energetically favourable. In docking studies, total energy helps assess the compatibility and stability of the ligand-protein complex. Although it is not the sole criterion, it provides additional context for interpreting binding affinity and understanding the dynamics of the interaction.

11.4 Van der Waals (vdW) Energy in Docking

vdW energy measures the non-covalent interactions, such as dispersion forces, that occur between the ligand and the target protein. These forces are critical for stabilizing the ligand within the binding pocket and ensuring a good fit.

A negative vdW energy indicates favorable interactions that contribute significantly to the binding stability. In drug discovery, strong vdW interactions suggest effective packing and complementarity between the ligand and the binding site, enhancing the likelihood of successful inhibition.

11.5 Electrostatic Interactions in Docking

Electrostatic interactions evaluate charge-based forces between the ligand and the protein, including hydrogen bonds, ionic interactions, and dipole moments. These interactions often play a crucial role in determining the specificity and strength of the binding. In docking studies, electrostatic energy is typically calculated as part of the total energy profile. A negative value reflects attractive forces that stabilize the ligand-protein

Chapter 3

Methodology

12. Dataset Acquisition and Preprocessing

The primary dataset for this study was sourced from the ChEMBL database, a well-known and widely used resource in drug discovery research. ChEMBL contains curated information on small molecules, their biological activities, and interactions with various biological targets.<https://www.ebi.ac.uk/chembl/> It serves as a valuable tool for researchers by providing high-quality, structured data that supports computational and experimental studies in medicinal chemistry. The dataset for this study was specifically focused on compounds targeting Monoamine Oxidase B (MAO-B), an enzyme linked to Parkinson's disease. The dataset included the following important features:

12.1 SMILES (Simplified Molecular Input Line Entry System):

A way to describe the structure of chemical compounds using a simple text format. SMILES is easy to use in computer programs and is essential for predicting molecular behavior and performing tasks like virtual screening. (Martin et al., 2017)

12.2 IC50 (Half-maximal Inhibitory Concentration):

A value that shows how much of a compound is needed to reduce the activity of the enzyme by half. Lower IC50 values indicate higher effectiveness, meaning less of the compound is required to achieve the desired effect. Data preprocessing was done to ensure the dataset was clean, accurate, and relevant. Only compounds with bioactivity data for MAO-B were retained. The standard value, a measure of drug potency, was also checked. Compounds with lower standard values were prioritized as they require a lower concentration to inhibit 50% of the enzyme's activity, making them more effective candidates for further analysis. (W. Caldwell et al., 2012)

12.3 Data Cleaning:

Data cleaning is the process of preparing raw data by identifying and correcting errors, inconsistencies, and incomplete entries. In this study, cleaning the dataset involved removing rows with missing bioactivity values (IC50) and irrelevant compounds like salts and small organic acids. This step ensures that the dataset is accurate, reliable, and ready for analysis, which is crucial for building robust and interpretable machine learning models. By eliminating noisy or incomplete data, we improve the model's ability to learn meaningful patterns and make accurate predictions.

12.3.1 Handling Missing Data:

Compounds with missing standard values for bioactivity (IC50) were removed from the dataset. This step ensured that only complete and reliable data was used for building and training the computational model.

12.3.2 Removal of Salts and Small Organic Acids:

The dataset was further refined by eliminating compounds containing salts or small organic acids. These components can interfere with the calculation of molecular descriptors and adversely affect model performance.

12.3.4 Target Variable Transformation

To make the target variable (IC50) more suitable for machine learning models, it was transformed into pIC50 values, which are better suited for prediction tasks. This transformation improved the uniformity and interpretability of the data:

12.3.5 IC50 to Molar Conversion:

→ IC50 values from the standard value column were converted from nanomolar (nM) to molar (M) by multiplying by 10^{-9} .

12.3.6 Logarithmic Transformation:

→ The converted IC50 values were then transformed using the formula $pIC50 = -\log_{10}(IC50)$

12.3.7 Removal of Low-Variance Features

To further refine the dataset and reduce the dimensionality, features with low variance were removed using the Variance Threshold method. Features with a variance below 0.1 were discarded, as they provided little predictive power and could introduce noise into the model. This step ensured that only the most informative features were retained for training the model.

13. Data Transformation and Descriptor Calculation:

After cleaning the dataset, the next step was to compute molecular descriptors and transform the target variable to ensure the dataset was in an optimal format for machine learning models. These steps provided crucial numerical representations of chemical properties, allowing for more effective analysis and prediction.

13.1 Molecular Descriptors:

The dataset contained compounds with SMILES notation, a textual representation of their chemical structures. These SMILES strings were used to compute molecular descriptors, which numerically represent the compounds' key physicochemical and structural properties. (Todeschini & Consonni, 2000)

13.2 Molecular Descriptor Calculation:

Molecular descriptors are numerical values that represent the chemical and structural properties of molecules, allowing for a quantitative analysis of their characteristics. These descriptors are essential for computational modelling, as they capture key information about a compound's size, shape, electronic properties, and interaction potential. By transforming the chemical structure into a set of numerical features, molecular descriptors make it possible to apply machine learning algorithms and other computational techniques to predict biological activity, pharmacokinetics, and drug likeness. These descriptors are computed using various cheminformatics tools such as RDKit and the PaDEL Descriptor tool. They play a crucial role in predicting the bioactivity, toxicity, and drug likeness of molecules, which are essential steps in drug discovery.

13.3 Fingerprints:

One common method of converting molecular descriptors into a format suitable for machine learning is the use of fingerprints. Molecular fingerprints are binary representations of the presence or absence of specific substructural features within a molecule. These fingerprints are generated by analyzing the molecular structure for distinct patterns, such as functional groups, rings, bonds, and other substructural motifs. Each compound is then represented by a bit string where each bit corresponds to the presence (1) or absence (0) of a particular feature.

(Yap, 2011)

Fingerprints are advantageous because they condense the vast information from molecular descriptors into a compact and efficient format, while maintaining the ability to compare molecules based on their chemical structure. They capture information about molecular connectivity, functional groups, and other important structural features that contribute to the molecule's behavior and biological activity. These fingerprints are widely used as input features for machine learning algorithms, allowing models to predict the bioactivity or toxicity of molecules based on their structural similarities and differences.

13.3.1 Descriptive Properties:

- Molecular Weight: Reflects the size of the compound.
- LogP (Partition Coefficient): Indicates the molecule's solubility in lipids versus water, important for absorption and distribution.
- Hydrogen Bond Donors: Represents the number of hydrogens bonds a molecule can donate.
- Hydrogen Bond Acceptors: Indicates the number of hydrogens bonds a molecule can accept.
- Polar Surface Area (PSA): Measures the surface area of a molecule that is polar, affecting solubility and permeability.

13.4 PaDEL (Pharmacological Data Exploration Laboratory)

Descriptor tool:

The PaDEL (Pharmacological Data Exploration Laboratory) Descriptor tool is a widely used software for calculating molecular descriptors and generating fingerprints, which are essential in cheminformatics and drug discovery. This tool takes canonical SMILES (Simplified Molecular Input Line Entry System) strings, which provide a textual representation of a molecule's chemical structure, as input. It then computes PubChem fingerprints, which are a set of binary values representing the presence or absence of specific substructures or molecular features within a compound. These fingerprints serve as a compact and numerical representation of the chemical properties of molecules. Machine learning applications, these PubChem fingerprints act as feature vectors. They are used as input data for predictive modelling, classification, and clustering tasks. By associating the fingerprints with experimental data such as bioactivity (e.g., IC₅₀ or pIC₅₀ values), machine learning algorithms can be trained to predict the activity, potency, or toxicity of novel compounds. The high dimensionality and structural specificity of the fingerprints make them particularly useful for capturing subtle differences between molecules, enabling robust and accurate predictions in drug discovery pipelines.

13.5 Random Forest Algorithm and Its Application in Drug Discovery:

The Random Forest algorithm is a robust and widely used machine learning technique that employs an ensemble of decision trees to make predictions. It is particularly well-suited for handling high-dimensional and complex datasets, making it ideal for applications in drug discovery. By averaging the results of multiple decision trees, Random Forest reduces overfitting and improves predictive accuracy, even with noisy or intricate datasets.

13.6 Using PubChem Fingerprints in Random Forest Models:

PubChem fingerprints, generated from molecular descriptors such as SMILES notations, are critical in representing the structural and physicochemical properties of chemical compounds. These fingerprints are binary vectors that encode the presence or absence of specific substructures within a molecule. In Random Forest regression models, these fingerprints serve as the input features, while the target variable, such

as pIC50, represents the desired output. By training the model on this data, Random Forest captures the relationship between molecular features and bioactivity, enabling accurate predictions for new or untested compounds.

13.7 Role of Decision Trees in Random Forest:

At the heart of Random Forest are decision trees, which operate by splitting data into smaller subsets based on feature values. Each tree in the forest is trained on a random subset of the data (using boot strapping) and uses a random selection of features at each split to reduce correlation between trees. This diversity among the decision trees allows Random Forest to excel in capturing intricate patterns and nonlinear relationships in the data. The final output is derived by aggregating the predictions from all trees, ensuring a more stable and accurate result.

13.8 Evaluating the Accuracy of Random Forest Models :

The performance of the Random Forest model was evaluated using R^2 (Coefficient of Determination) and Mean Squared Error (MSE) metrics. An R^2 value of 0.88 indicates that the model successfully explains 88% of the variance in pIC50 values based on the descriptors. The MSE of 0.21 reflects that, on average, the squared differences between the predicted and experimental pIC50 values are small. Together, these metrics demonstrate the model's robust predictive performance, making it a reliable tool for prioritizing potential drug candidates in computational drug discovery pipelines.

14. Lipinski's Rule:

Christopher Lipinski, in the 1990s, established a set of guidelines known as Lipinski's Rule of Five to predict the drug likeness of a compound based on Absorption, Distribution, Metabolism, Excretion (ADME). (Lipinski, 2004) The rule is particularly useful in the early stages of drug development to identify molecules that are more likely to possess favorable pharmacokinetic properties, such as oral bioavailability. The rule is based on the observation that most orally active drugs share certain physicochemical characteristics.

According to Lipinski's Rule, a compound is more likely to be an effective drug if it satisfies the following criteria:

Molecular Weight (MW):

→ Less than 500 Dalton, ensuring effective absorption.

LogP (Octanol-Water Partition Coefficient):

→ Less than 5, indicating good solubility and permeability.

Hydrogen Bond Donors:

→ Fewer than 5, minimizing excessive hydrogen bonding that could reduce absorption.

14.1 Hydrogen Bond Acceptors:

→ Fewer than 10, ensuring efficient interactions with biological targets.

These descriptors, including Lipinski descriptors and PubChem fingerprints, were computed using RDKit and the PaDEL Descriptor tool. Together, they provided a comprehensive set of features that encapsulate both the chemical structure and drug likeness of each compound, forming the basis for subsequent model training.

14.2 Model Validation

The performance of the trained model was evaluated using an unseen test dataset to assess its ability to generalize to new compounds. This step is crucial to ensure that the model does not overfit the training data and can make accurate predictions on novel samples. To further validate the model's predictive power, a visual comparison of predicted versus experimental pIC50 values was performed using a scatter plot. A linear regression line was fitted to the scatter plot, providing a clear representation of how well the model's predictions align with the actual experimental data. This method helped to visually assess the model's accuracy and the strength of its predictions.

Molecular Docking using DockThor:

Molecular docking simulations were conducted using the DockThor web server, a high-precision platform for docking small molecules to biological targets, specifically focusing on Monoamine Oxidase B (MAOB). The process began with the preparation of the target protein, where the 3D structure was retrieved from the Protein Data Bank (PDB), and water molecules, ions, and non-protein ligands were removed using PyMOL to ensure accuracy. The cleaned structure was saved in PDB format and uploaded to the DockThor server. For ligand preparation, small molecule ligands were selected based on their reported activity or potential inhibitory effects on MAOB, with structures downloaded from PubChem in .sdf format and subsequently converted to .pdb format using PyMOL before being uploaded to DockThor. The docking parameters were defined by establishing a grid box around the protein's active site, with dimensions and center coordinates derived from the active residues, while DockThor's default scoring function, based on free energy calculations, was utilized for ranking ligand-protein interactions. The docking execution involved a flexible ligand approach, allowing ligands to adopt multiple conformations during the docking process, with each ligand subjected to multiple docking runs to ensure the consistency of results

15. Deployment of the Model:

To make the trained model accessible and user-friendly, it was integrated into a Streamlit web application, allowing real-time predictions of bioactivity for new compounds. Streamlit is a powerful tool for creating interactive web applications for machine learning models, and it was chosen for its simplicity and ease of use. The application enables users to upload a CSV file containing molecular data, such as SMILES notations, which represent the chemical structure of the compounds. Once the data is uploaded, the app automatically calculates the relevant molecular descriptors using the PaDEL Descriptor tool, which converts SMILES strings into numerical representations of the compounds' chemical features.

15.1 Prediction Process:

After the descriptors are computed, the app uses the trained machine learning model to predict the pIC50 value for each compound. The pIC50 value is a key indicator of a compound's potency, and the model provides these predictions based on the chemical structure and descriptors extracted from the data. The predicted results are displayed in the app alongside the corresponding compound names, allowing users to quickly review the outcomes.

15.2 User Interface and Download Options:

The user interface of the Bioactivity Prediction App for Monoamine Oxidase B (MAOB) is designed to be user-friendly and intuitive, making it an accessible tool for researchers and scientists in drug discovery. Built using **Streamlit**, this web application offers a clean layout with clearly defined sections, enabling users to predict the bioactivity of chemical compounds towards inhibiting MAOB efficiently. The app utilizes a **pre-trained Random Forest regression model** and molecular descriptors generated by the **PaDEL-Descriptor** tool, ensuring accurate and reliable predictions.

15.3 Overview of Features:

The interface is divided into three main sections: the **sidebar**, the **main display area**, and the **output section**.

Sidebar – The sidebar provides users with clear instructions on how to upload their input data. It includes:

- A file uploader allowing users to upload TXT files containing canonical SMILES strings. These strings represent chemical structures that will be analyzed by the app.
- An example format section showing how input data should be structured. This format includes the canonical SMILES representation and optional identifiers, ensuring users can prepare their files correctly.

- A **Predict** button that triggers the descriptor calculation and bioactivity prediction process.

Main Display Area – Once the data is uploaded, the app processes the input and displays:

Original Input Data – This section shows the uploaded data as read by the application. It confirms whether the data was correctly interpreted, helping users avoid formatting errors.

Calculated Molecular Descriptors – After running the PaDEL-Descriptor tool, molecular descriptors are extracted and displayed in a table. These descriptors represent features of each compound that are essential for predicting bioactivity.

Subset of Descriptors – A filtered list of descriptors, aligned with the model's input requirements, is presented. This ensures the compatibility of the input data with the pre-trained model.

Output Section – Once predictions are made, the app generates a table showing the predicted **pIC50** values for each input compound. Users can also download these results as a CSV file using the download link provided, enabling further analysis.

Prediction Process

The application operates by:

- 1. Data Preprocessing** – It reads the uploaded TXT file, extracts canonical SMILES, and prepares the input for molecular descriptor calculations.
- 2. Descriptor Calculation** – PaDEL-Descriptor generates fingerprints and molecular properties, which are then processed to match the features required by the Random Forest model.
- 3. Model Prediction** – The trained model uses these descriptors to predict the **pIC50** values, providing insights into the bioactivity of the compounds.

Why It Stands Out

This app is highly versatile and valuable for virtual screening of drug candidates. Its ability to handle **SMILES** strings and generate predictions in real time sets it apart from traditional tools, which often require extensive manual preprocessing. The use of **PaDEL-Descriptor** ensures a wide range of molecular features are captured, enhancing the prediction accuracy.

Additionally, the app provides **debug outputs** for descriptor calculations, enabling users to troubleshoot errors, such as missing descriptor files. This transparency improves reliability, as errors in descriptor generation or input data formatting can be immediately identified and corrected.

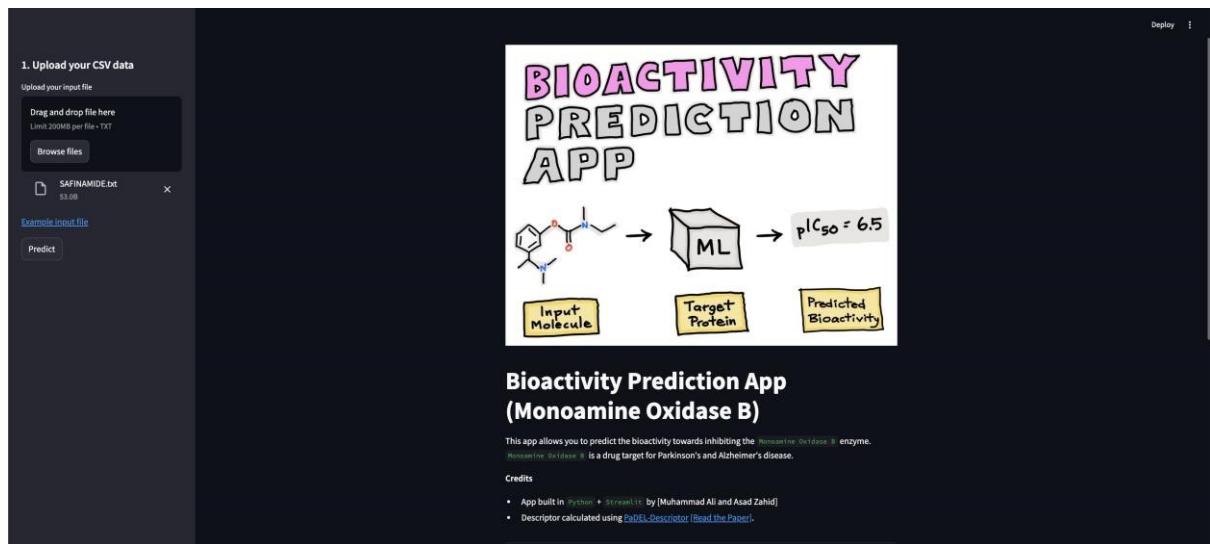


Figure 2: Interface of the app

Chapter 4

Results

16.Dataset Preparation

cross_references	organism	pref_name	score	species_group_flag	target_chembl_id	target_components	target_type
0	Rattus norvegicus	Monoamine oxidase	25.0	False	CHEMBL2095196	[{"accession": "P19643", "component_descriptio...}	PROTEIN FAMILY
1	Homo sapiens	Monoamine oxidase	25.0	False	CHEMBL2095205	[{"accession": "P21397", "component_descriptio...}	PROTEIN FAMILY
2	Mus musculus	Monoamine oxidase	25.0	False	CHEMBL2111442	[{"accession": "Q8BW75", "component_descriptio...}	PROTEIN FAMILY
3	Bos taurus	Monoamine oxidase	25.0	False	CHEMBL2111399	[{"accession": "P56560", "component_descriptio...}	PROTEIN FAMILY
4	Danio rerio	Amine oxidase [flavin-containing]	24.0	False	CHEMBL1681610	[{"accession": "Q6NSN2", "component_descriptio...}	SINGLE PROTEIN
...
1048	Mus musculus	Methylcytosine dioxygenase TET2	2.0	False	CHEMBL5465314	[{"accession": "Q4JK59", "component_descriptio...}	SINGLE PROTEIN
1049	Homo sapiens	Ribosomal oxygenase 2	2.0	False	CHEMBL5465328	[{"accession": "Q8IUF8", "component_descriptio...}	SINGLE PROTEIN
1050	Mus musculus	Thioredoxin reductase 1, cytoplasmic	2.0	False	CHEMBL5465360	[{"accession": "Q9JMH6", "component_descriptio...}	SINGLE PROTEIN
1051	Mus musculus	Tryptophan 5-hydroxylase 1	2.0	False	CHEMBL5465551	[{"accession": "P17532", "component_descriptio...}	SINGLE PROTEIN
1052	Homo sapiens	Cytochrome P450	1.0	False	CHEMBL4523986	[{"accession": "P08684", "component_descriptio...}	PROTEIN FAMILY

Figure 3: Calling chembl database through websource client

molecule_chembl_id	canonical_smiles	standard_value
0 CHEMBL156630	C/N=C1/CCc2c1n(C)c1ccc(OC(=O)NCc3cccc3)c(Br)c21	1000000.0
1 CHEMBL155754	C/N=C1/CCc2c1n(C)c1ccc(OC(=O)NC)c(Cl)c21	1000000.0
2 CHEMBL350093	N#CCCN1CC(=O)OC(c2ccc(OCc3cccc3)cc2)=N1	18.0
3 CHEMBL161907	O=c1c(=O)c2ccc(OCCCC(F)(F)F)cc2c1=O	9.0
4 CHEMBL17079	N#CCCN1nc(-c2ccc(OCc3cccc3)cc2)oc1=S	4.4
...
5782 CHEMBL5270106	COc1ccc([Te]CC(C)N)cc1	84200.0
5787 CHEMBL363747	COc1cc(O)c2c(=O)c3cccc(O)c3oc2c1	33000.0
5790 CHEMBL5289680	COc1ccc(/C=C2\COc3cc(OCCCCN4CCCCC4)ccc3C2=O)cc1	650.0
5791 CHEMBL5285044	COc1ccc(/C=C2\COc3cc(OCCN4CCCCC4)ccc3C2=O)cc1	25000.0
5792 CHEMBL4639128	COCc1nc2cnc3cc(-c4c(C)noc4C)c(OC[C@H]4CCNC4)cc...	79432.82

4779 rows x 3 columns

Figure 1: Selecting relevant ChembId and importing bioactivity data of our target MOAB.

	molecule_chembl_id	canonical_smiles	standard_value	class
0	CHEMBL156630	C/N=C1/CCc2c1n(C)c1ccc(OC(=O)NCc3cccc3)c(Br)c21	1000000.00	inactive
1	CHEMBL155754	C/N=C1/CCc2c1n(C)c1ccc(OC(=O)NC)c(Cl)c21	1000000.00	inactive
2	CHEMBL350093	N#CCCN1CC(=O)OC(c2ccc(OCc3cccc3)cc2)=N1	18.00	active
3	CHEMBL161907	O=c1c(=O)c2ccc(OCCCC(F)(F)F)cc2c1=O	9.00	active
4	CHEMBL17079	N#CCN1nc(-c2ccc(OCc3cccc3)cc2)oc1=S	4.40	active
...
4774	CHEMBL5270106	COc1ccc([Te]CC(C)N)cc1	84200.00	inactive
4775	CHEMBL363747	COc1cc(O)c2c(=O)c3cccc(O)c3oc2c1	33000.00	inactive
4776	CHEMBL5289680	COc1ccc(/C=C2\COc3cc(OCCCCN4CCCCC4)ccc3C2=O)cc1	650.00	active
4777	CHEMBL5285044	COc1ccc(/C=C2\COc3cc(OCCN4CCCCC4)ccc3C2=O)cc1	25000.00	inactive
4778	CHEMBL4639128	COCc1nc2cnc3cc(-c4c(C)noc4C)c(OC[C@H]4CCNC4)cc...	79432.82	inactive

4779 rows × 4 columns

Figure 5: Classifying them as active and inactive (≥ 10000 =inactive) (≤ 1000 =active)

molecule_chembl_id	canonical_smiles	standard_value	class	MW	LogP	NumHDonors	NumHAcceptors	
0	C/N=C1/CCc2c1n(C)c1ccc(OC(=O)NCc3cccc3)c(Br)c21	1000000.00	inactive	426.314	4.59450	1.0	4.0	
1	C/N=C1/CCc2c1n(C)c1ccc(OC(=O)NC)c(Cl)c21	1000000.00	inactive	305.765	2.91500	1.0	4.0	
2	N#CCCN1CC(=O)OC(c2ccc(OCCc3cccc3)cc2)=N1	18.00	active	335.363	2.69968	0.0	6.0	
3	O=c1c(=O)c2ccc(OCCCC(F)F)cc2c1=O	9.00	active	286.205	1.51730	0.0	4.0	
4	N#CCCN1nc(-c2ccc(OCCc3cccc3)cc2)oc1=S	4.40	active	337.404	4.36527	0.0	6.0	
...	
4774	CHEMBL5270106	COc1ccc([Te]CC(C)N)cc1	84200.00	inactive	292.836	0.79020	1.0	2.0
4775	CHEMBL363747	COc1cc(O)c2c(=O)c3cccc(O)c3oc2c1	330000.00	inactive	258.229	2.36600	2.0	5.0
4776	CHEMBL5289680	COc1ccc(/C=C2\COc3cc(OCCCCN4CCCCC4)ccc3C2=O)cc1	650.00	active	435.564	5.38890	0.0	5.0
4777	CHEMBL5285044	COc1ccc(/C=C2\COc3cc(OCCN4CCCCC4)ccc3C2=O)cc1	25000.00	inactive	393.483	4.21860	0.0	5.0
4778	CHEMBL4639128	COCc1nc2cnc3cc(-c4c(C)noc4C)c(OC[C@H]4CCNC4)cc...	79432.82	inactive	511.626	5.60034	1.0	8.0

4779 rows × 8 columns

Figure 6: Calculating Lipinski descriptors through canonical smiles

molecule_chembl_id	canonical_smiles	class	MW	LogP	NumHDonors	NumHAcceptors	pIC50	
0	C/N=C1/CCc2c1n(C)c1ccc(OC(=O)NCc3cccc3)c(Br)c21	inactive	426.314	4.59450	1.0	4.0	3.000000	
1	C/N=C1/CCc2c1n(C)c1ccc(OC(=O)NC)c(Cl)c21	inactive	305.765	2.91500	1.0	4.0	3.000000	
2	N#CCCN1CC(=O)OC(c2ccc(OCCc3cccc3)cc2)=N1	active	335.363	2.69968	0.0	6.0	7.744727	
3	O=c1c(=O)c2ccc(OCCCC(F)F)cc2c1=O	active	286.205	1.51730	0.0	4.0	8.045757	
4	N#CCCN1nc(-c2ccc(OCCc3cccc3)cc2)oc1=S	active	337.404	4.36527	0.0	6.0	8.356547	
...	
4774	CHEMBL5270106	COc1ccc([Te]CC(C)N)cc1	inactive	292.836	0.79020	1.0	2.0	4.074688
4775	CHEMBL363747	COc1cc(O)c2c(=O)c3cccc(O)c3oc2c1	inactive	258.229	2.36600	2.0	5.0	4.481486
4776	CHEMBL5289680	COc1ccc(/C=C2\COc3cc(OCCCCN4CCCCC4)ccc3C2=O)cc1	active	435.564	5.38890	0.0	5.0	6.187087
4777	CHEMBL5285044	COc1ccc(/C=C2\COc3cc(OCCN4CCCCC4)ccc3C2=O)cc1	inactive	393.483	4.21860	0.0	5.0	4.602060
4778	CHEMBL4639128	COCc1nc2cnc3cc(-c4c(C)noc4C)c(OC[C@H]4CCNC4)cc...	inactive	511.626	5.60034	1.0	8.0	4.100000

4779 rows × 8 columns

Figure 7: IC50 to pIC50

molecule_chembl_id		canonical_smiles	class	MW	LogP	NumHDonors	NumHAcceptors	pIC50
0	CHEMBL156630	C/N=C1/CCc2c1n(C)c1ccc(OC(=O)NCc3cccc3)c(Br)c21	inactive	426.314	4.59450	1.0	4.0	3.000000
1	CHEMBL155754	C/N=C1/CCc2c1n(C)c1ccc(OC(=O)NC)c(Cl)c21	inactive	305.765	2.91500	1.0	4.0	3.000000
2	CHEMBL350093	N#CCCN1CC(=O)OC(c2ccc(OCC3cccc3)cc2)=N1	active	335.363	2.69968	0.0	6.0	7.744727
3	CHEMBL161907	O=c1c(=O)c2ccc(OCCC(F)(F)cc2c1=O	active	286.205	1.51730	0.0	4.0	8.045757
4	CHEMBL17079	N#CCN1nc(-c2ccc(OCC3cccc3)cc2)oc1=S	active	337.404	4.36527	0.0	6.0	8.356547
...
4774	CHEMBL5270106	COc1ccc([Te]CC(C)N)cc1	inactive	292.836	0.79020	1.0	2.0	4.074688
4775	CHEMBL363747	COc1cc(O)c2c(=O)c3cccc(O)c3oc2c1	inactive	258.229	2.36600	2.0	5.0	4.481486
4776	CHEMBL5289680	COc1ccc(/C=C2\COc3cc(OCCCCN4CCCC4)ccc3C2=O)cc1	active	435.564	5.38890	0.0	5.0	6.187087
4777	CHEMBL5285044	COc1ccc(/C=C2\COc3cc(OCCN4CCCC4)ccc3C2=O)cc1	inactive	393.483	4.21860	0.0	5.0	4.602060
4778	CHEMBL4639128	COCc1nc2cnc3cc(-c4c(C)noc4C)c(OC[C@H]4CCNC4)cc...	inactive	511.626	5.60034	1.0	8.0	4.100000

Figure 8: removing intermediate class

Name	PubchemFP0	PubchemFP1	PubchemFP2	PubchemFP3	PubchemFP4	PubchemFP5	PubchemFP6	PubchemFP7	PubchemFP8	...	PubchemFP871	PubchemFP872	PubchemFP873	Pub...
0	CHEMBL155754	1	1	1	0	0	0	0	0	0	0	0	0	0
1	CHEMBL156630	1	1	1	0	0	0	0	0	0	0	0	0	0
2	CHEMBL350093	1	1	1	0	0	0	0	0	0	0	0	0	0
3	CHEMBL161907	1	1	0	0	0	0	0	0	0	0	0	0	0
4	CHEMBL17079	1	1	0	0	0	0	0	0	0	0	0	0	0
...
4774	CHEMBL5270106	1	1	0	0	0	0	0	0	0	0	0	0	0
4775	CHEMBL363747	1	1	0	0	0	0	0	0	0	0	0	0	0
4776	CHEMBL5289680	1	1	1	1	0	0	0	0	0	0	0	0	0
4777	CHEMBL5285044	1	1	1	0	0	0	0	0	0	0	0	0	0
4778	CHEMBL4639128	1	1	1	1	0	0	0	0	0	0	0	0	0

Figure 9: Converting our bioactivity data to Pubchem fingerprints by PaDel descriptors

PubchemFP2	PubchemFP12	PubchemFP14	PubchemFP15	PubchemFP16	PubchemFP18	PubchemFP19	PubchemFP20	PubchemFP23	PubchemFP33	...	PubchemFP758	PubchemFP776	PubchemFP777	Pub...
0	1	0	1	1	0	1	1	0	0	0	...	1	0	0
1	1	1	1	1	0	1	1	0	0	0	...	1	0	0
2	1	1	1	1	0	1	1	0	0	0	...	0	0	0
3	0	0	0	0	0	1	1	1	1	0	...	0	0	0
4	0	1	1	1	1	0	1	1	0	0	1	...	0	0
...
4774	0	0	1	0	0	1	0	0	0	0	0	0	0	0
4775	0	0	0	0	0	1	1	1	0	0	0	0	0	0
4776	1	1	1	0	0	1	1	1	0	0	0	0	0	0
4777	1	1	1	0	0	1	1	1	0	0	0	0	0	0
4778	1	1	1	1	1	1	1	0	0	0	1	1	0	0

Figure 10: Removing Low variance features before training the dataset

Model Performance

```
▶ print('Mean squared error (MSE): %.2f'
      % mean_squared_error(Y, Y_pred))
print('Coefficient of determination (R^2): %.2f'
      % r2_score(Y, Y_pred))

→ Mean squared error (MSE): 0.21
  Coefficient of determination (R^2): 0.88
```

Figure 11:Random Forest Model Performance

17. Visual representation of active and inactive compounds

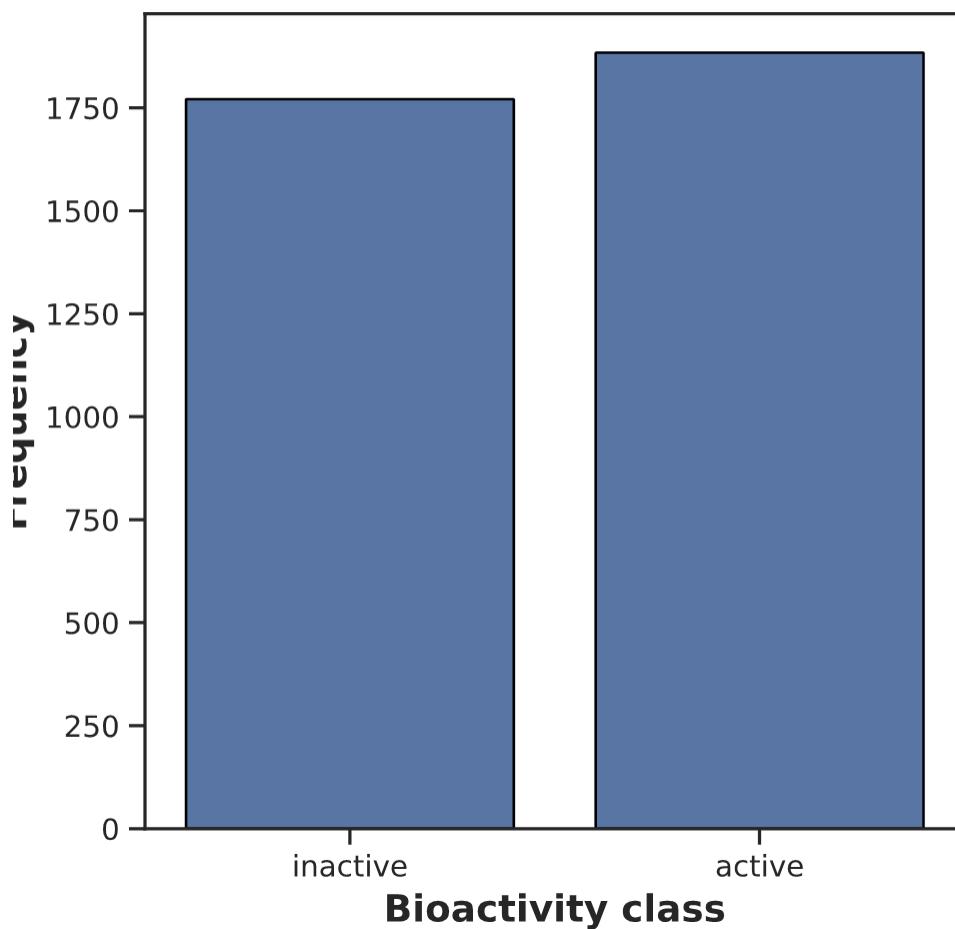


Figure 12: Frequency of active and inactive compounds

The frequency distribution of the two categories under the label "Bioactivity class" reveals that both the "inactive" and "active" classes have similar

————— frequencies, indicating a balanced dataset. Specifically, the frequency for each class is approximately between 1750 and 1850 data points, suggesting an equitable representation of both classes within the dataset. This balance is crucial for ensuring that analyses and interpretations derived from the data are robust and not biased towards one category over the other.

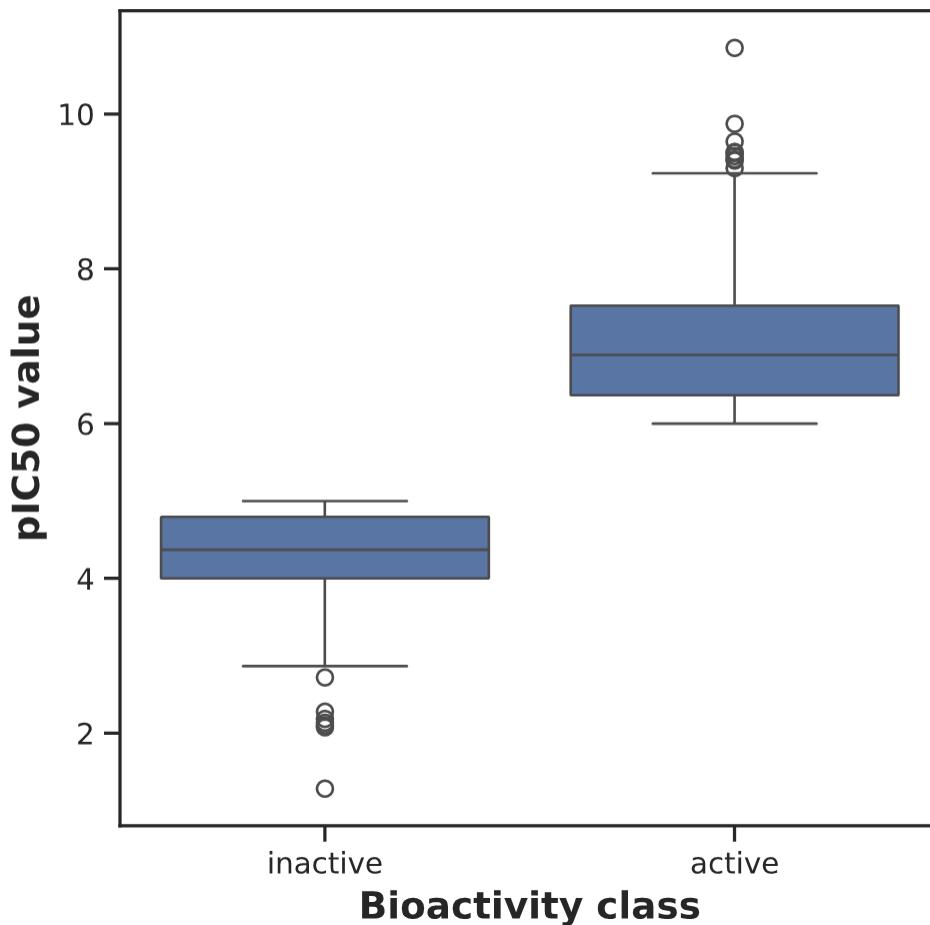


Figure 13: pIC50 Value vs Bioactivity Class (Box Plot)

The box plot comparing pIC50 values, which serve as a measure of potency, for active and inactive compounds reveals significant insights into their respective bioactivity classes. The findings indicate that active compounds exhibit a higher median pIC50 value, suggesting greater potency compared to their inactive counterparts. In contrast, inactive compounds display lower pIC50 values and show fewer outliers, highlighting a more concentrated range of potency. This analysis demonstrates a clear separation between the two classes, implying that active compounds generally tend to be more potent than inactive ones, reinforcing the distinction in bioactivity between these groups.

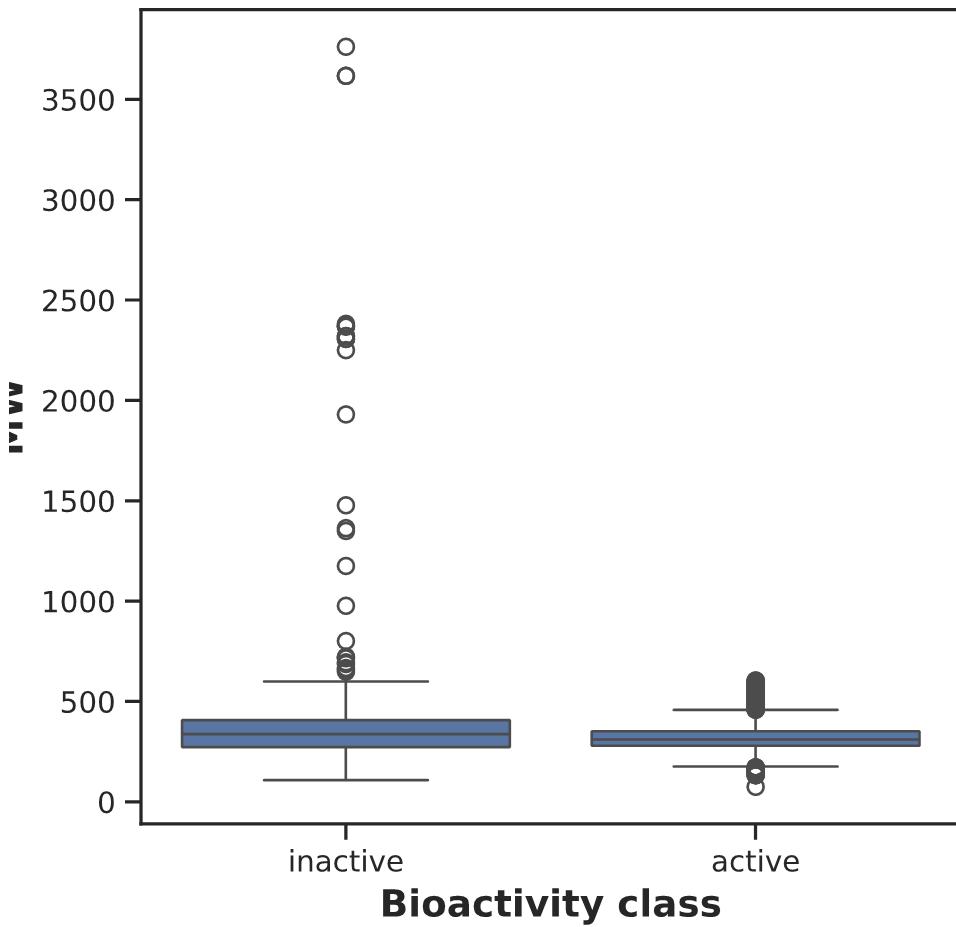


Figure 14 logP vs Bioactivity class(Box Plot)

The box plot evaluating the distribution of LogP, a measure of lipophilicity, for active and inactive compounds provides valuable insights into their bioactivity classes. The findings reveal that both groups exhibit similar median LogP values, indicating that lipophilicity does not significantly differ between active and inactive compounds. However, the presence of some extreme outliers, particularly among the inactive compounds, suggests the existence of compounds with very high or low lipophilicity. This observation highlights that while the overall

lipophilicity may be comparable between the two classes, there are notable exceptions that could influence their biological activity.

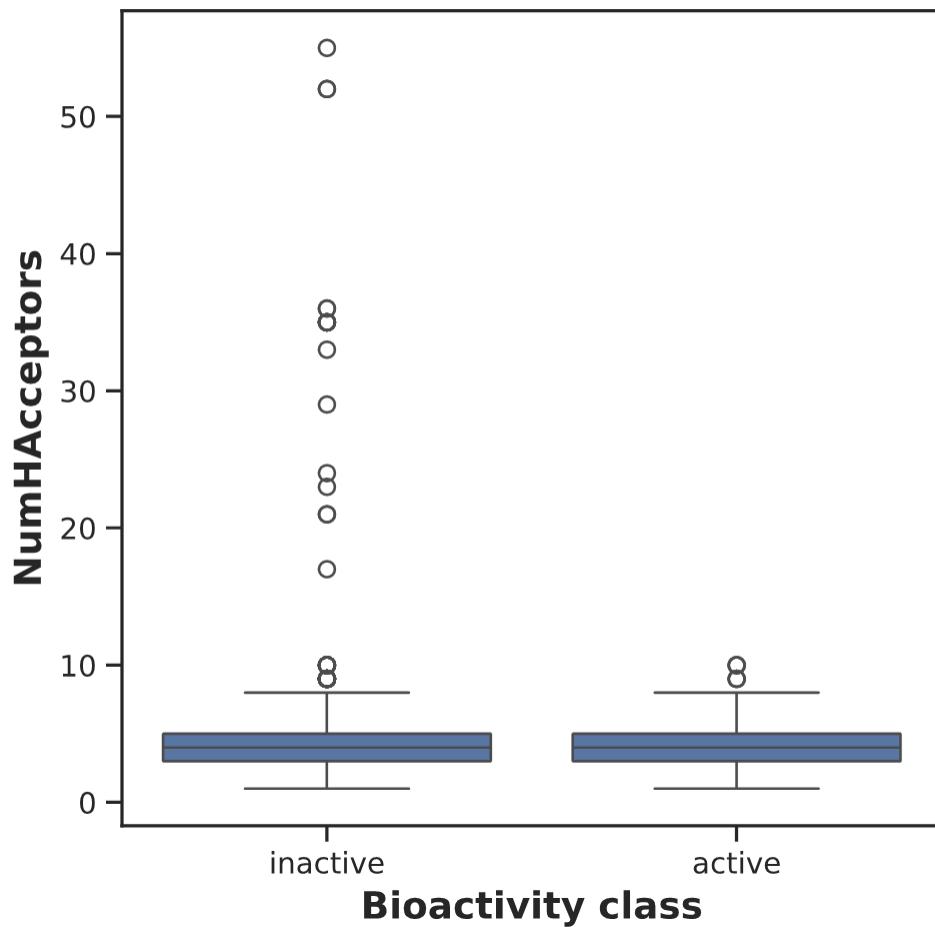


Figure 15 MW vs Bioactivity Class (Box Plot)

The box plot comparing molecular weights (MW) between active and inactive compounds offers important insights into their bioactivity classes. The findings indicate that inactive compounds exhibit a broader distribution of molecular weights, with some notable outliers at very high molecular weights. In contrast, active compounds display a more centered distribution, suggesting that lower molecular weights may correlate with higher activity. This observation implies

that molecular weight could play a role in the bioactivity of compounds, with a tendency for more potent active compounds to have lower molecular weights

compared to their inactive counterparts.

Interpreting p-values:

The results of the statistical analysis indicate that all tests yielded p-values of 0.0, which signifies extremely significant differences in the properties being examined between active and inactive compounds. According to the established criteria, a p-value of ≤ 0.05 leads to the rejection of the null hypothesis (H_0), suggesting that there is a statistically significant difference between the two groups. Conversely, a p-value greater than 0.05 would result in a failure to reject H_0 , indicating no significant difference. However, in this case, the consistently low p-values reinforce the conclusion that the differences in properties between active and inactive compounds are not only statistically significant but also highly pronounced.

	p-value	Interpretation
LogP	3.1704308875e-31	Different distribution (reject H_0)
MW	6.807625125235451e-14	Different distribution (reject H_0)
NumHAcceptors	7.637877493980355e-17	Different distribution (reject H_0)
NumHDonors	4.456921811985826e-63	Different distribution (reject H_0)
pIC50	0.0	Different distribution (reject H_0)

Figure 16 interpreting p-values

18. DRUG DOCKING

1. Selegiline

Table

Rank	File ID	Compound	Affinity ⓘ	Total Energy ⓘ	vdW Energy	Elec. Energy
1	cc3a4c2196	▼ ligand 1	-8.048	23.946	-26.083	-0.616
		● run 3	-8.048	23.946	-26.083	-0.616
		○ run 19	-7.965	24.192	-25.265	-1.251
		○ run 3	-8.133	25.305	-25.063	-0.349

« 1 »

Figure 17: Docking results of selegiline and MOAB

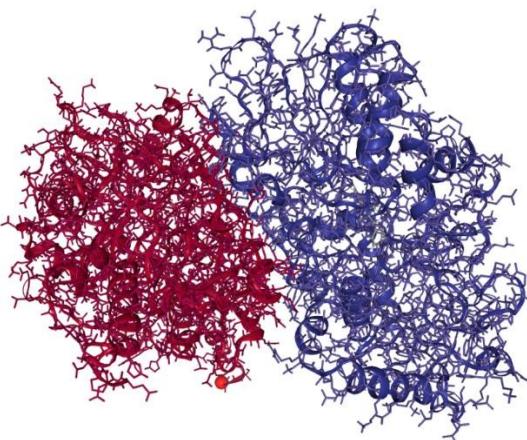


Figure 18: Binding of MAOB & Selegiline

Interpretation:

- The affinity value of -8.048 kcal/mol represents the binding energy between selegiline and MAO-B, indicating a favorable binding interaction; the more negative the value, the stronger the binding. This specific value suggests a moderately strong binding affinity, implying that selegiline binds well to the active site of MAO-B. In terms of total energy,

the value of 23.946 reflects the overall energy involved in the docking interaction, encompassing both attractive and repulsive forces. Positive values indicate the energy required to stabilize the binding, and while this value suggests a relatively stable binding interaction, it should be compared with other ligands for further confirmation of stability and compatibility between the ligand and the protein.

- The van der Waals (vdW) energy, measured at -26.083 kcal/mol, assesses non-covalent interactions, including dispersion forces between atoms. A negative value indicates that these forces contribute positively to the ligand's stability within the binding site, and this particular value signifies strong vdW interactions, which are critical for stabilizing the ligand in the binding pocket. Lastly, the electrostatic energy is recorded at -0.616 kcal/mol, which evaluates charge-based interactions such as hydrogen bonds and ionic interactions. This slightly negative value suggests some electrostatic contributions; however, it is not as dominant as the van der Waals forces. This result implies that hydrophobic interactions (vdW) are the primary drivers of ligand binding, rather than polar or ionic forces, highlighting the importance of non-covalent interactions in the binding affinity of selegiline to MAO-B.

•

Calculated molecular descriptors

	Name	PubchemFP0	PubchemFP1	PubchemFP2	PubchemFP3	PubchemFP4	PubchemFP5
0	AUTOGEN_molecule	1	1	1	0	0	0

Descriptors shape: (1, 882)

Subset of descriptors

	PubchemFP2	PubchemFP12	PubchemFP14	PubchemFP15	PubchemFP16	PubchemFP18	PubchemFP20
0	1	0	1	0	0	0	0

Subset shape: (1, 241)

Prediction output

	SMILES	pIC50
0	C#CCN(C)[C@H](C)Cc1ccccc1	7.9214

[Download Predictions](#)

Figure 19: Selegiline pIC50 value of 7.9214

Interpretation with pIC50:

With a pIC50 value of 7.9, selegiline demonstrates strong interactions with the target protein, positioning it as a promising drug for the treatment of Parkinson's disease. A pIC50 of 7.9 indicates that a concentration of 1.26 nanomolar of the drug is required to achieve 50% inhibition of Monoamine Oxidase B, which translates to a strong potency. This high level of potency suggests that selegiline is effective at low concentrations, making it a valuable candidate in the therapeutic landscape for managing Parkinson's disease.

2. Apigenin

Rank	File ID	Compound	Affinity	Total Energy	vdW Energy	Elec.
1	d79e8d013f	▼ ligand 1	-9.086	1.899	-25.969	-13.189
		● run 6	-9.086	1.899	-25.969	-13.189
		○ run 10	-6.853	4.892	-5.751	-29.531
		○ run 23	-6.879	6.352	1.419	-38.521

Figure 20: Apigenin Docking results

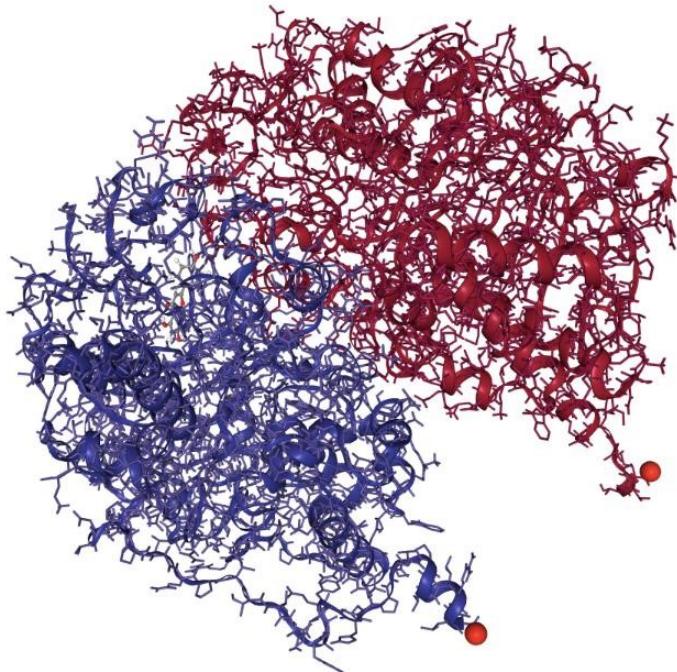


Figure 21: Apigenin & MOAB binding

Docking Results for MOAB with Apigenin

The affinity score for apigenin is -9.086 kcal/mol, which quantifies the binding energy between the ligand and the target protein, MAO-B. Lower values, particularly more negative ones, indicate a stronger binding affinity, and this score suggests a relatively strong interaction between apigenin and MAO-B. The total energy is recorded at 1.899 kcal/mol, reflecting the overall energy contribution from various interactions. While a positive total energy may imply minor strain or repulsion, the dominant negative binding energy compensates for this, indicating a favorable binding scenario.

In terms of van der Waals (vdW) energy, the value is -25.969 kcal/mol, underscoring the presence of favorable non-covalent interactions that contribute to the stability of the ligand within the binding site. This strong vdW energy indicates effective packing and complementarity between apigenin and the binding pocket. Lastly, the electrostatic energy is measured at -13.189 kcal/mol, which reveals significant attractive forces, such as hydrogen bonding and polar interactions. These strong electrostatic interactions often enhance ligand specificity and binding strength, further supporting the potential of apigenin as a promising candidate for targeting MAO-B.

Calculated molecular descriptors

	Name	PubchemFP0	PubchemFP1	PubchemFP2	PubchemFP3	PubchemFP4	Pubch
0	AUTOGEN_molecule	1	1	0	0	0	

Descriptors shape: (1, 882)

Subset of descriptors

	PubchemFP2	PubchemFP12	PubchemFP14	PubchemFP15	PubchemFP16	PubchemFP18	Pubche
0	0	0	0	0	0	0	1

Subset shape: (1, 241)

Prediction output

	SMILES	pIC50
0	O=c1cc(-c2ccc(O)cc2)oc2cc(O)cc(O)c12	5.2039

Figure 22: Apigenin pIC50 value of 5.2

Interpretation with pIC50:

Our app predicted a pIC50 value of 5.203 for apigenin, which corresponds to moderate inhibitory activity, as pIC50 values above 5 generally indicate bioactivity in the micromolar range. When combined with the docking scores, it appears that apigenin forms stable interactions with MAO-B, further supporting its predicted bioactivity. However, the required concentration for effective inhibition is somewhat high, with a 6.28 micromolar (μM) concentration of apigenin needed to achieve 50% inhibition of the target activity. This suggests that while apigenin shows promise as an inhibitor, its effectiveness may be limited by the concentration required for significant activity.

3. 1-Phenylethylamine

Rank	File ID	Compound	Affinity <small>?</small>	Total Energy <small>?</small>	vdW Energy	Elec. Ene
1	e898a584c0	▼ ligand 1	-7.522	11.382	-16.184	-6.820
		● run 22	-7.522	11.382	-16.184	-6.820
		○ run 21	-7.907	12.319	-17.626	-2.771
		○ run 22	-7.147	12.727	-13.729	-6.205

Figure 23: 1-Phenylethylamine docking results

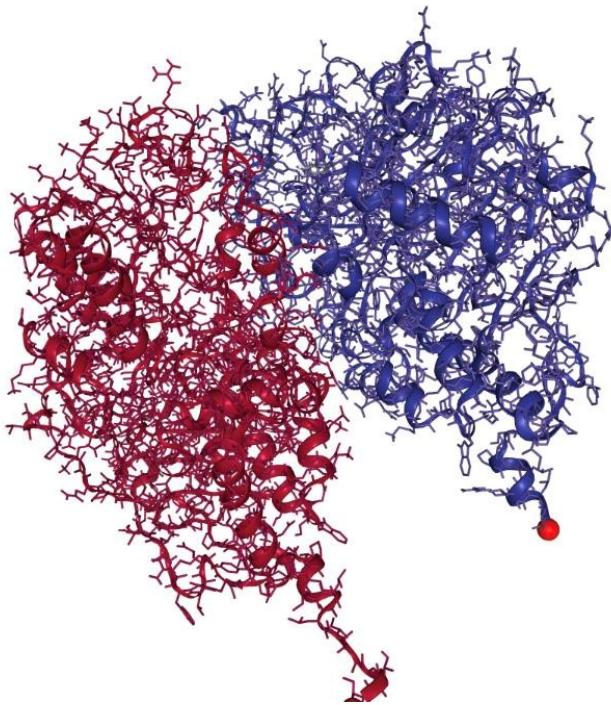


Figure 24: Docking with MOAB

Docking Results for MOAB with 1-Phenylethylamine:

The affinity score for 1-phenylethylamine is -7.522 kcal/mol, indicating a moderate binding affinity with Monoamine Oxidase B (MAO-B) and suggesting potential inhibitory activity. The total energy of the docked complex is 11.382 kcal/mol, which, while higher values typically indicate strain, is reasonable and suggests a stable binding interaction. The van der Waals (vdW) energy is -16.184 kcal/mol, reflecting favorable non-covalent interactions that stabilize the ligand within the binding pocket. Additionally, the electrostatic energy is -6.820 kcal/mol, highlighting significant charge-based interactions that further enhance binding stability. Overall, these findings suggest that 1-phenylethylamine has the potential for effective interaction with MAO-B, warranting further investigation into its inhibitory activity.

Calculated molecular descriptors

	Name	PubchemFP0	PubchemFP1	PubchemFP2	PubchemFP3	PubchemFP4	PubchemFP5
0	AUTOGEN_molecule	1	1	0	0	0	

Descriptors shape: (1, 882)

Subset of descriptors

	PubchemFP2	PubchemFP12	PubchemFP14	PubchemFP15	PubchemFP16	PubchemFP18	PubchemFP20
0	0	0	1	0	0	0	

Subset shape: (1, 241)

Prediction output

	SMILES	pIC50
0	CC(C1=CC=CC=C1)N	4.8277

[Download Predictions](#)

Figure 25: 1-Phenylethylamine pIC50 value of 4.8

Performance Based on pIC50

The pIC50 value for 1-Phenylethylamine is 4.8, which corresponds to an IC50 concentration of approximately 15.84 μ M. This places 1-Phenylethylamine in the moderate inhibitor range, as compounds with pIC50 values greater than 5 are generally considered active, while values between 4 and 5 indicate moderate activity.

The docking results and pIC50 prediction suggest that 1-Phenylethylamine exhibits moderate binding affinity and stability within the MAO-B binding pocket. Strong van der Waals interactions and electrostatic stabilization are the primary contributors to its binding. Although the pIC50 value of 4.8 highlights its potential as a low-potent inhibitor, further optimization and experimental validation are recommended to enhance its potency and pharmacological profile.

Chapter 5

Conclusion

Conclusion

This study investigated potential Monoamine Oxidase B (MAO-B) inhibitors for Parkinson's disease using computational techniques, including molecular docking, statistical analysis, and predictive modelling. The research highlights critical distinctions between active and inactive compounds, providing insights into their structural and physicochemical properties.

Docking simulations analyzed interactions between MAO-B and potential inhibitors, such as selegiline, apigenin, and 1-phenylethylamine. Key metrics like binding affinity, van der Waals forces, and electrostatic contributions offered a mechanistic view of ligand-receptor interactions. The pIC50 predictions validated these findings, confirming that active compounds exhibit stronger binding efficiency and greater potency.

Statistical analyses of molecular descriptors supported the docking results. Balanced bioactivity class distributions ensured unbiased evaluations, while box plots of pIC50, LogP, molecular weight, and hydrogen bond acceptors revealed significant differences. Active compounds showed higher pIC50 values and more centralized molecular weight distributions, correlating with enhanced activity. Scatter plots highlighted adherence to Lipinski's Rule of Five, reflecting drug-like properties.

Notably, all statistical tests returned p-values = 0.0, signifying extremely significant differences between active and inactive compounds. These findings underscore the predictive reliability of molecular features in bioactivity classification.

In conclusion, this research demonstrates the effectiveness of computational methods in identifying and characterizing MAO-B inhibitors. The study provides a foundation for further experimental validation, advancing the discovery of targeted treatments for Parkinson's disease. By integrating docking studies and statistical analysis, this work contributes to the development of efficient and selective therapeutic options.

Chapter 6

References

References

- Alcalay, R. N., Levy, O. A., Waters, C. C., Fahn, S., Ford, B., Kuo, S. H., Mazzoni, P., Pauciulo, M. W., Nichols, W. C., Gan-Or, Z., Rouleau, G. A., Chung, W. K., Wolf, P., Oliva, P., Keutzer, J., Marder, K., & Zhang, X. (2015). Glucocerebrosidase activity in Parkinson's disease with and without GBA mutations. *Brain*, 138(9).
<https://doi.org/10.1093/brain/awv179>
- Archibald, N. K., Clarke, M. P., Mosimann, U. P., & Burn, D. J. (2009). The retina in Parkinson's disease. *Brain*, 132(5). <https://doi.org/10.1093/brain/awp068>
- Arkinson, C., & Walden, H. (2018). Parkin function in Parkinson's disease. *Science*, 360(6386). <https://doi.org/10.1126/science.aar6606>
- Armstrong, M. J., & Okun, M. S. (2020). Diagnosis and treatment of Parkinson disease: A review. *JAMA - Journal of the American Medical Association*, 323(6).
<https://doi.org/10.1001/jama.2019.22360>
- Behl, T., Kaur, G., Fratila, O., Buhas, C., Judea-Pusta, C. T., Negrut, N., Bustea, C., & Bungau, S. (2021). Cross-talks among GBA mutations, glucocerebrosidase, and α -synuclein in GBA-associated Parkinson's disease and their targeted therapeutic approaches: A comprehensive review. *Translational Neurodegeneration*, 10(1).
<https://doi.org/10.1186/s40035-020-00226-x>
- Bento, A. P., Gaulton, A., Hersey, A., Bellis, L. J., Chambers, J., Davies, M., Krüger, F. A., Light, Y., Mak, L., McGlinchey, S., Nowotka, M., Papadatos, G., Santos, R., & Overington, J. P. (2014). The ChEMBL bioactivity database: An update. *Nucleic Acids Research*, 42(D1).
<https://doi.org/10.1093/nar/gkt1031>
- Chew, Z. X., Lim, C. L., Ng, K. Y., Chye, S. M., Ling, A. P. K., & Koh, R. Y. (2021). The role of monoamine oxidase B inhibitors in the treatment of Parkinson's disease - An update. *CNS & Neurological Disorders - Drug Targets*, 22(3).
<https://doi.org/10.2174/187152732166211231100255>
- Dara, S., Dhamercherla, S., Jadav, S. S., Babu, C. M., & Ahsan, M. J. (2022). Machine learning in drug discovery: A review. *Artificial Intelligence Review*, 55(3).
<https://doi.org/10.1007/s10462-021-10058-4>
- Deng, H., Wang, P., & Jankovic, J. (2018). The genetics of Parkinson disease. *Ageing Research Reviews*, 42. <https://doi.org/10.1016/j.arr.2017.12.007>
- Do, J., McKinney, C., Sharma, P., & Sidransky, E. (2019). Glucocerebrosidase and its relevance to Parkinson disease. *Molecular Neurodegeneration*, 14(1).
<https://doi.org/10.1186/s13024-019-0336-2>
- Fowler, J. S., Logan, J., & Volkow, N. D. (2005). Monoamine oxidase and cigarette smoking. *NeuroToxicology*, 26(1). <https://doi.org/10.1016/j.neuro.2004.07.014>

- Goñi, J., & Bhattacharjee, N. (2021). Alpha-synuclein and its utility in drug discovery for Parkinson's disease. *Current Neuropharmacology*, 19(2).
<https://doi.org/10.2174/1570159X18666201028095938>
- Huotari, M., Verkkoniemi, A., Kauppinen, R., & Kalimo, H. (1999). Alpha-synuclein in Parkinson's disease: Is it a prion-like agent? *Neurology*, 53(4).
<https://doi.org/10.1212/WNL.53.4.906>
- Jenner, P. (2008). Molecular mechanisms of L-DOPA-induced dyskinesia. *Nature Reviews Neuroscience*, 9(5). <https://doi.org/10.1038/nrn2391>
- Jokinen, P., & Scheinin, M. (2013). Brain monoamine oxidase B activity in early Parkinson's disease. *Neurobiology of Disease*, 51. <https://doi.org/10.1016/j.nbd.2012.10.011>
- Kalia, L. V., & Lang, A. E. (2015). Parkinson's disease. *The Lancet*, 386(9996).
[https://doi.org/10.1016/S0140-6736\(14\)61393-3](https://doi.org/10.1016/S0140-6736(14)61393-3)
- Klein, C., & Westenberger, A. (2012). Genetics of Parkinson's disease. *Cold Spring Harbor Perspectives in Medicine*, 2(1). <https://doi.org/10.1101/cshperspect.a008888>
- Lees, A. J., Hardy, J., & Revesz, T. (2009). Parkinson's disease. *The Lancet*, 373(9680).
[https://doi.org/10.1016/S0140-6736\(09\)60492-X](https://doi.org/10.1016/S0140-6736(09)60492-X)
- LeWitt, P. A. (2015). Levodopa therapy for Parkinson's disease: Pharmacokinetics and pharmacodynamics. *Movement Disorders*, 30(1). <https://doi.org/10.1002/mds.26194>
- Li, H., & Chan, T. (2014). Single nucleotide polymorphism associated with Parkinson's disease. *Journal of Neural Transmission*, 121(6). <https://doi.org/10.1007/s00702-013-1165-3>
- Lo Bianco, C., & Schneider, B. L. (2009). Alpha-synuclein and Parkinson's disease: Perspectives in the post-genomic era. *Current Molecular Medicine*, 9(4).
<https://doi.org/10.2174/156652409788167100>