CONSTRUCTION OF LANDSCAPE OF DISEASE CAUSING MUTATION IN NEURODEGENERATIVE DISEASE

AND

DOCKING STUDIES TO IDENTIFY INHIBITORS TO PREVENT GTP BINDING AND INCREASE IN KINASE ACTIVITY

MUSUKHA MALA BRAHMA

A Dissertation Submitted to

Indian Institute of Technology Hyderabad

In Partial Fulfilment of the Requirements for

The Degree of Master of Technology

भारतीय प्रौद्योगिकी संस्थान हैदराबाद

Indian Institute of Technology Hyderabad

Department of Biotechnology Engineering

June, 2018

Declaration

I declare that this written submission represents my ideas in my own words, and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited, or from whom proper permission has not been taken when needed.

(Signature)

MUSUKHA MALA BRAHMA

BO16MTECH11005

Approval Sheet

This thesis entitled "Construction of Landscape of Disease Causing Mutation in Neurodegenerative Disease and Docking Studies to Identifying Inhibitors to Prevent GTP Binding and Increase in Kinase Activity" by Musukha Mala Brahma is approved for the degree of Master of Technology from IIT Hyderabad.

Dr. Sandeep Kumar Singh
Assistant Professor
Department of Biotechnology

Dr. Falguni Pati
Assistant Professor
Department of Biomedical Engineering

21/6/18

Dr. Ashish Misra
Assistant Professor
Department of Biotechnology

Acknowledgements

It is my privilege to express my sincerest regards to my project supervisor Dr. Ashish Misra, for encouraging me to do the project on the topic “Construction of Landscape of Disease Causing Mutation in Neurodegenerative Disease and Docking Studies to Find Inhibitors to Prevent GTP Binding and Increase in Kinase Activity” at my department for the partial fulfilment of the requirements leading to the award of degree of Master of Technology in Biotechnology Department. I take this opportunity to give my sincere thanks to Arivarasan Sampath for initially helping me to make the Schrodinger software familiar. And, also thank all my friends and family for their cooperation and support.

Table of Contents: **Page No.**

Abstract

Leucine-rich repeat kinase 2(LRRK2) is the greatest known genetic contributor to Parkinson's disease (PD). LRRK2 encodes a 2527 amino-acid multi-domain protein. It has several regions predicted to be involved in protein-protein interactions, which include an ankyrin repeat domain, a leucine-rich repeat domain and a WD40 domain. It also has two catalytic domains including a kinase domain of the tyrosine kinase-like family and a GTPase domain of the Ras of complex proteins family (ROC). Tyrosine kinase and Ras family GTPases like kinases are often associated as elements of the same intracellular signalling pathway, that suggest a functional interaction between both of these catalytic functions within LRRK2. Roc GTPase domain represents hotspot for mutations it has three PD mutations at a single R1441 residue (i.e., R1441C, R1441G, and R1441H), two PD mutations at a single T1343 residue (i.e., T1343G and T1343V), N1437H, I1371V, K1347A, R1398H, and T1348N. Y1699C mutation is the only known disease-causing variant in COR domain. G2019S and I2020T, located in adjacent residues of the kinase activation loop, are known to cause familial PD.

Mutation in LRRK2 causes impairment in GTPase activity and therefore alters kinase activity. Kinase activity of LRRK2 requires an intact of ROC and GTPase domain. The activity of LRRK2 kinase increases when GTP binds to the GTPase (ROC) domain. The increase in Kinase activity on GTP binding is the cause of PD as known for now. There is a number of mutations in the domains present in LRRK2 that is responsible for changes in GTPase activity and Kinase activity. Mutations in the catalytic ROC-COR and kinase domains of LRRK2 are considered as the most common causes of PD. Mutations located in the kinase domain influence only kinase activity whereas mutations in ROC domain may influence both GTP-binding and kinase activity. The GTPase domain plays a role in regulating kinase activity, and in mediating the neurotoxic effects of LRRK2. Therefore, the GTPase domain and Kinase domain is studied as a therapeutic target for inhibiting the pathogenic effects of LRRK2 mutations.

LRRK2 Roc domain structure from UniProt: Q5S007 (LRRK2_HUMAN) and Homology modelling of Kinase domain are done in Swiss-Model to study the inhibition efficiency of GTP binding and Kinase activity.

Abbreviations:

ALS: amyotrophic lateral sclerosis.

DM: Myotonic dystrophy.

FTD: Frontotemporal dementia

FXS: Fragile X syndrome.

FXTAS: Fragile X-associated tremor/ataxia syndrome.

OPMD: Oculopharyngeal muscular dystrophy.

PD: Parkinson's disease.

POMA: Paraneoplastic opsoclonus-myoclonus ataxia.

SCA2: Spinocerebellar ataxia type 2.

SMA: spinal muscular atrophy.

LRR: Leucine-rich repeat.

ROC: Renin-angiotensin system in complex proteins.

COR: C-terminal domain of ROC.

MAPKKK: Mitogen-activated protein kinase kinase kinase

List of Figures:

1. INTRODUCTION:

1.1 Neurodegenerative Diseases:

Neurodegenerative disease is a term for a range of neuronal conditions. It is the progressive loss of function of neurons that leads to the death of neurons. Neurons are the building blocks of the nervous system which includes the brain and spinal cord. When neurons become damaged or die they cannot be replaced by the body as they don't reproduce or replace themselves. Neurodegenerative diseases are incurable conditions that result in progressive degeneration and death of nerve cells, which causes ataxias or dementias.

Example of neurodegenerative diseases are:

- Alzheimer's disease (AD)
- Parkinson's disease (PD)
- Amyotrophic lateral sclerosis (ALS)
- Prion disease
- Huntington's disease (HD)
- Spinocerebellar ataxia (SCA)
- Spinal muscular atrophy (SMA)

In these study, I have taken Amyotrophic lateral sclerosis (ALS) and Parkinson's disease (PD) to construct the landscape of disease-causing mutation in neurodegenerative disease.
ALS and PD are the second and third most common human adult-onset neurodegenerative diseases, respectively, after Alzheimer's disease (1).

They are characterized by prominent age-related neurodegeneration in selectively vulnerable neural systems. Some forms of PD and ALS are inherited and some are sporadic and genes causing these diseases have been identified.

Both ALS and PD are neurodegenerative diseases, and are characterized by the presence of intra-neuronal inclusions; however, different classes of neurons are affected and the primary protein in the inclusions differs between the diseases, and in some cases is different in distinct forms of the same disease.

The major symptoms of both diseases arise from the degeneration of particular classes of the neuron: the dopaminergic neurons of the substantia nigra pars compacta in the case of PD and the spinal motor neurons in the case of ALS.

1.2 RNA Binding Proteins:

RBPs are involved in all aspects of RNA processing, controlling the life cycle of RNAs from synthesis to degradation. Hallmark features of RBPs in neuron dysfunction include misregulation of RNA processing, mislocalization of RBPs to the cytoplasm, and abnormal aggregation of RBPs (2). Recent advances in neurodegenerative diseases point to novel mechanisms of protein aggregation that revolve around the unique biology of RNA binding proteins (3). RNA binding proteins normally are present in the nucleus (3). Under conditions of cell stress, these RNA binding proteins translocate to the cytoplasm where they form stress granules, which function in part to sequester specialized transcript and promote translation of protective proteins (3). Studies in humans show that pathological aggregates occurring in ALS, Alzheimer's disease and other dementias co-localize with stress granules. Mutations in RNA binding proteins or prolonged periods of stress causes the formation of very stable, pathological stress granules (3). The consolidation of RNA binding proteins away from the nucleus and neuronal arbours into pathological stress granules might impair the normal physiological activities of these RNA binding proteins causing the neurodegeneration associated with these diseases.

RBPs TDP-43 and FUS for ALS and SNCA and LRRK2 for PD are focused in this study. That is, how the mutations in these proteins cause the disease is studied.

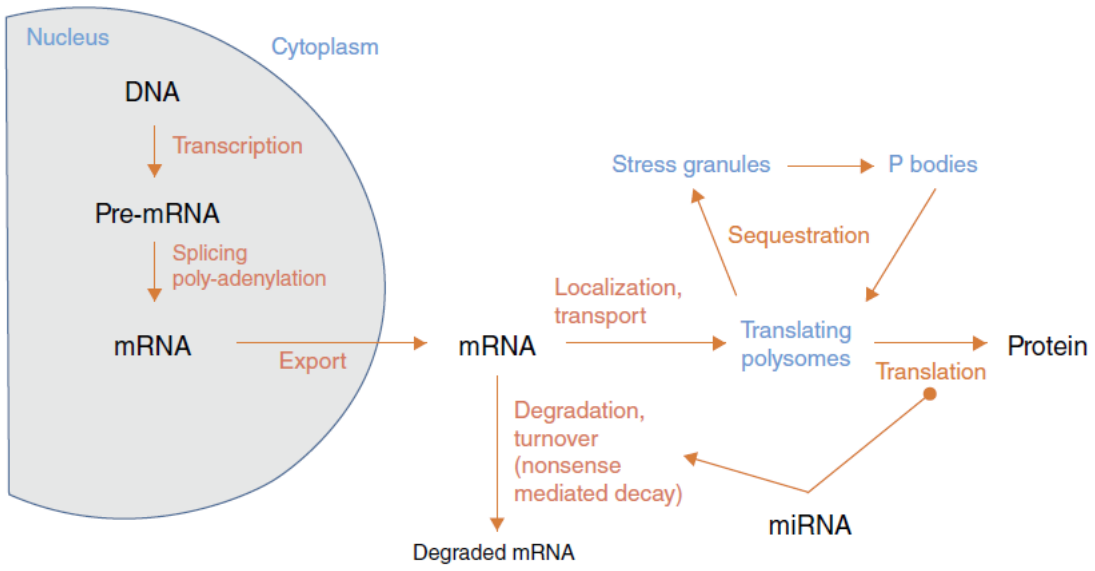


Figure from RNA-binding proteins implicated in neurodegenerative diseases; Mark R Cookson

Figure 1: RNA-binding proteins (RBPs) implicated in several neurological diseases. RBPs implicated in neurological diseases have a role in several steps of RNA processing, including transcription, alternative splicing, alternative polyadenylation, localization of transcripts, including sequestration into inclusions and stress granules, translation, RNA degradation, and turnover (4). The physiological functions of these RBPs often lend significant insights into the mechanisms of pathogenesis.

RNA processing step, RNA binding proteins and the diseases:

Transcription:

Taf15, EWS, FUS **(ALS/FTD)**

Alternative splicing:

MBNL1/2, CuGBP **(DM)**

TDP43, FUS, TAF15, hnRNPA1, hnRNPA2/B1, EWS **(FTD/ALS)**

NOVA **(POMA)**

RBFOX **(Ataxia)**

Alternative polyadenylation:

NOVA (POMA)

MBNL1/2 (DM)

PABPN1 (OPMD)

Localization transport and sequestration:

hnRNPA2/B1, TDP43, FUS, TAF15, EWS **(ALS/FTD)**

FMRP **(FXS, FXTAS)**

ATXN2 **(SCA2, ALS)**

SMN **(SMA)**

Degradation and turnover:

CELF4, HuR/ELAVLI **(PD)**

MATR3 **(ALS)**

Translation:

FMRP, DGCR8, DROSHA **(FXS, FXTAS)**

hnRNPA2/B1, hnRNPA1 **(ALS/FTD)**

PARK7 **(PD)**

2. Scope of study:

Since these RNA binding proteins are found to be aggregated and toxic in ALS and PD condition, it could be hypothesized to find some small molecules/phytochemical to inhibit the kinase activity and GTP binding.

Objectives:

1. To find the most common gene mutation of severely damaging protein in the pathogenesis**.**
2. To find small molecules to inhibit the GTP binding and Kinase activity.
3. Docking study with the protein and small molecule.

3. Objective1:

To find the most common gene mutation of severely damaging protein in the pathogenesis.

The literature survey is done in order to find the epidemiology of ALS and PD, from the statistics of the gene which severely affects the population and which is prone to mutation

(hotspot for mutation) is found. The epidemiology statistics is done in order to narrow down the search to find the most common and severely affecting gene responsible for the disease pathogenesis. From several searches in database and literature survey.

Table 1: Selected the gene targets for ALS and PD.

Disease	Genes selected
PD	LRRK2, SNCA
ALS	TDP-43, FUS

3.1 Amyotrophic lateral sclerosis (ALS):

Amyotrophic lateral sclerosis (ALS or Lou Gehrig's disease) is a progressive neurodegenerative disease specifically affecting cortical and spinal motor neurons. It is 100% fatal. ALS usually strikes people between the ages of 40-70 and approximately 20,000 Americans can have the disease at any given time (5).

There are two different types of ALS, sporadic and familial.

a) Sporadic ALS (sALS), accounts for 90 to 95 percent of all cases (5). The most common form of the disease in the U.S.
b) Familial ALS (fALS) accounts for 5 to 10 percent of all cases in the U.S. In those families, there is 50% probability of each offspring to inherit the gene mutation and disease development.

Cytoplasmic inclusions containing hyperphosphorylated and ubiquitinated TDP-43 are a pathological hallmark of ALS, and mutations in the gene encoding TDP-43 have been directly linked to the development of the disease (6). TDP-43 is a ubiquitous DNA/RNA-binding protein with a nuclear role in pre-mRNA splicing.

TDP-43 protein and FUS protein aggregates have been implicated in several cases of the disease, and the most common known cause of sporadic ALS is thought to be due to mutation in chromosome 9 (C9orf72)

3.1.1 TDP-43 Is an ALS Disease Protein:

TDP-43 was found in cytoplasmic inclusions in the diseased brain, which helped to understand the causes of ALS and methods to study its pathogenesis which was severely limited by an inability to identify ALS genes. After which many mutations in TDP43 were found in sALS and fALS cases, but not in control cases.

TDP-43 has various roles in RNA metabolism, it is highly conserved, ubiquitously expressed protein (7). TDP-43 has two RNA Recognition Motifs (RRMs) and a C-terminal Glycine-rich domain. It is predominantly a nuclear protein, in ALS motor neurons which were found to mislocalize to the cytoplasm, with some affected neurons (8). TDP-43 was also found to be modified and to be misfolded in an aggregated state. Its modifications include ubiquitination, phosphorylation, and cleavage. The implications of these modifications as well as their origins and roles in ALS pathology is unclear.

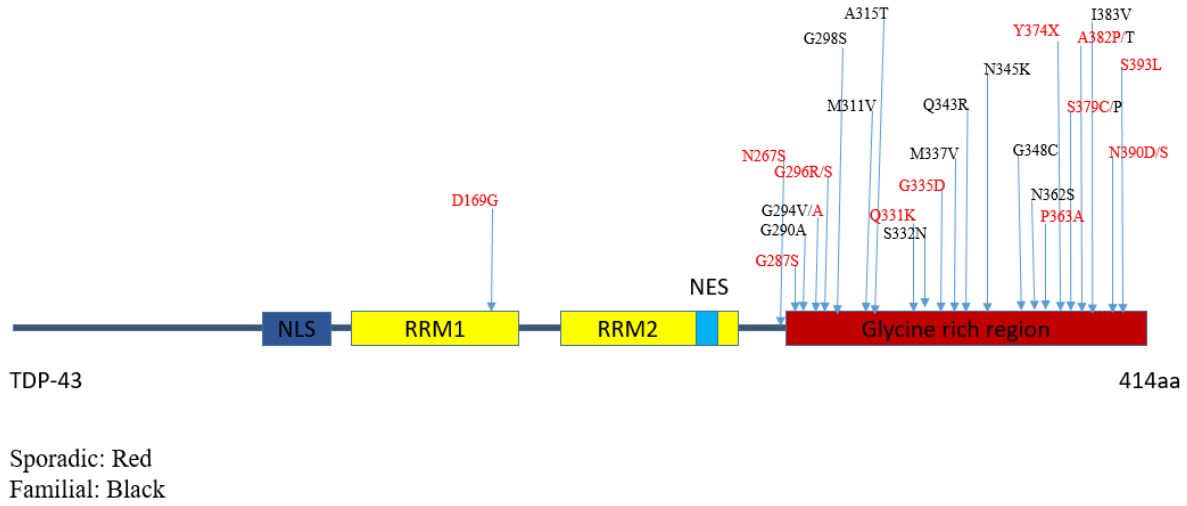


Figure 2: TDP-43 mutations in ALS: Mutations in TDP-43 identified in sALS and fALS patients. All are missense mutations. RRM1 and RRM2: RNA recognition motifs; NLS, nuclear localization signal; (NES), predicted nuclear export signal.

3.1.2 FUS (fused in sarcoma) Protein in ALS Disease:

Like TDP-43, FUS is a ubiquitously expressed protein that is normally nuclear but is found in cytoplasmic inclusions in ALS brain (9). High levels of *FUS* (fused in sarcoma), either the normal or mutant forms, kill motor neurons, and certain mutations change the protein's location, which might impair the normal functioning of neurons.

Mutations in FUS account for ~35% of FALS in patients younger than 40 years old. Meta-analyses of 154 ALS cases with FUS mutations (including FALS and SALS with *de novo* FUS mutations) show an average disease onset of 43.8±17.4 years. More than 60% of cases with FUS mutations show disease onset before 45 years of age, with many juvenile ALS cases presenting with disease onset in late teens and early 20's

Mutation and proteinopathies in RNA binding protein FUS, which is closely related to TDP-43, raises the intriguing possibility that perturbations to the RNA homeostasis and metabolism in neurons contribute to the pathogenesis of these diseases. They converge on the same mechanisms leading to neurodegeneration, unlike TDP-43 there is an evidence that FUS mutations target distinct mechanisms to cause early disease onset and aggressive progression of the disease. FUS mutations perturb the maintenance of dendrites through fundamental processes in RNA splicing, RNA transport and DNA damage response/repair (10).

FUS contains an N-terminal QGSY-rich domain, followed by a Glycine-rich domain, an RRM, and two RGG-rich domains. Like TDP-43, FUS has roles in a number of different RNA-processing activities, including splicing and mRNA export to the cytoplasm (9). About 50 ALS-linked FUS mutations have been identified.

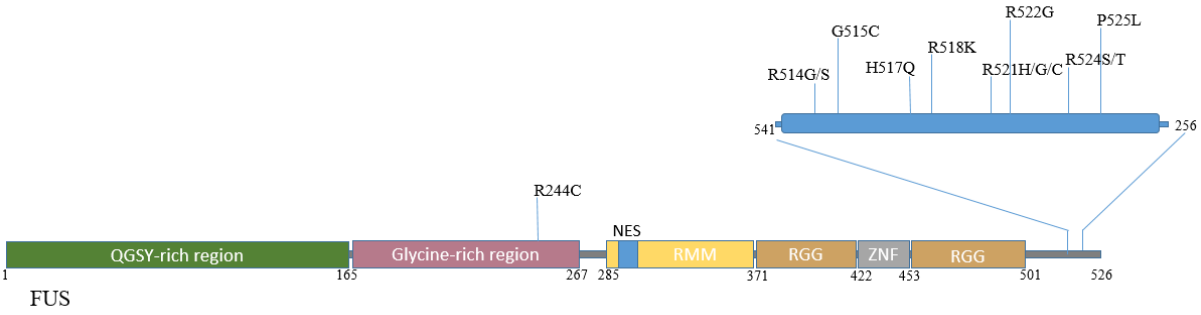


Figure 3: Genomic organization of the human FUS Gene, FUS mutations identified in ALS, and functional domains in FUS proteins.

3.2 Parkinson's disease

Parkinson's disease is the second most common neurodegenerative disease and manifests as bradykinesia, rigidity, resting tremor and posture instability. Parkinson's disease is a degenerative disorder of the central nervous system. It results from the death of dopamine-generating cells in the substantia nigra, a region of the midbrain; the cause of cell-death is unknown. PD is characterized by the core motor symptoms collectively called parkinsonisms, including resting tremor, bradykinesia, muscle rigidity, postural instability and gait impairment, and it is also accompanied by a range of non-motor symptoms, which include constipation, urinary symptoms, sleep disorder, and dementia. The mechanism by which the brain cells in Parkinson's are lost may consist of an abnormal accumulation of the protein alpha-synuclein bound to ubiquitin in the damaged cells (11). The protein accumulation forms cytoplasmic inclusions called Lewy bodies.

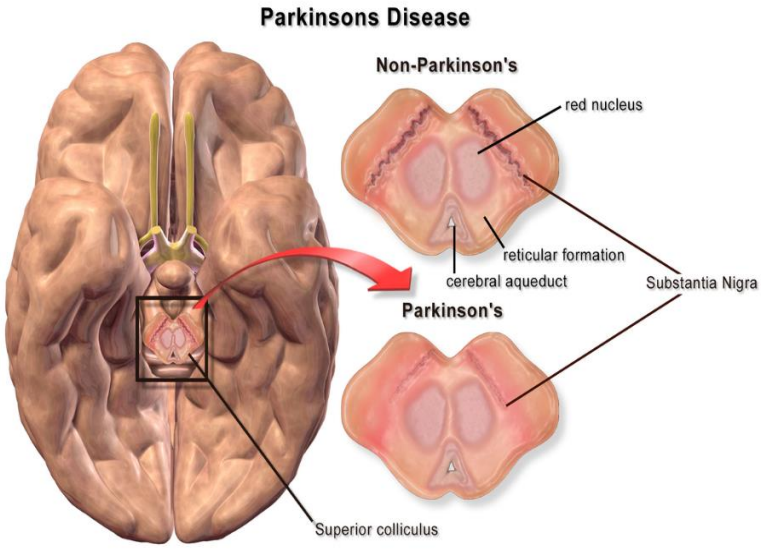


Figure from: Blausen Gallery 2014/wikiversity journal of medicine, doi:10.15347/wjm/2014.010.

Figure 4: A brain without and with Parkinson's Disease compared in Substantia Nigra

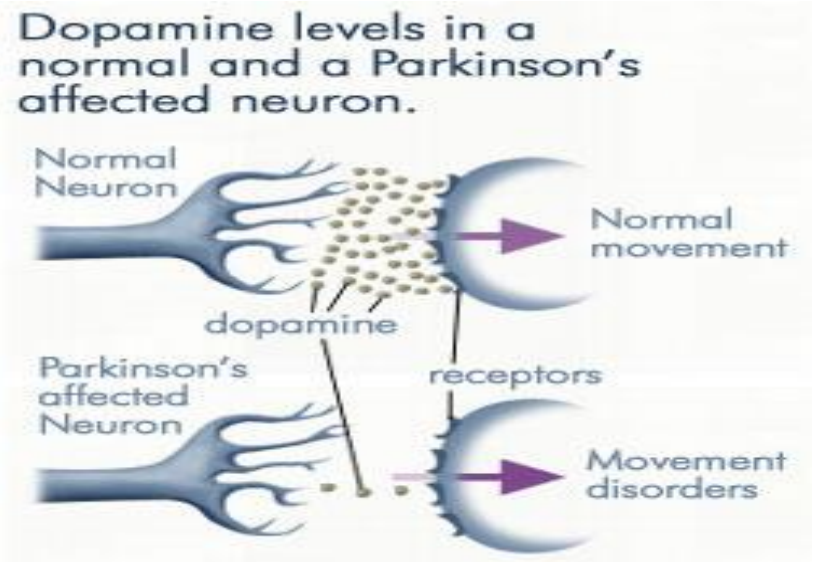


Figure from: anti- agingfirewalls.com

Figure 5: Decreased in dopamine level produced by the neurons affected by Parkinson's disease as a result less dopamine is available to bind to the dopamine receptor on the post-synaptic membrane.

An estimated seven million to 10 million people worldwide have Parkinson's disease. 10 percent of family history. The prevalence of the disease ranges from 41 people per 100,000 in the fourth decade of life to more than 1,900 people per 100,000 among those 80 and older. An estimated 4 percent of people with Parkinson's are diagnosed before the age of 50.

The crude prevalence rate of PD has been reported to range from 15 per 100,000 to 12,500 per 100,000, and the incidence of PD from 15 per 100,000 to 328 per 100,000, with the disease being less common in Asian countries.

Men are 1^1/2 times more likely to have Parkinson's than women.

Parkinson's statistics in selected countries

United States

About 1million Americans are thought to have Parkinson's.

Every year, about 60,000 Americans are diagnosed with Parkinson's.

Canada

There are over 100,000 Canadians living with Parkinson's today and approximately 6,600 new cases of PD are diagnosed each year in Canada.

About 56 percent of patients receive formal or informal assistance due to their condition. Of those who receive it, 84 percent rely on family, friends or neighbours, while 56 percent obtain other assistance.

United Kingdom

The prevalence of Parkinson's in the United Kingdom is one in 500 people, with about a total of about 127,000 people having the disease. Reports show that every hour, someone is diagnosed with Parkinson's in Britain. Most are 50 or older.

The main known risk factor is age. Genes implicated in the development of PD include SNCA, LRRK2, GBA, PRNK, PINK1, PARK7, VPS35, EIF4G1, DNAJC13 AND CHCHD2.

Mutations in some genes, including SNCA, LRRK2 and GBA, have been found to be risk factors for "sporadic" (non-familial) PD. LRRK2 is one of the 20,000 genes which are passed from parents to children. Mutations in the gene LRRK2 are the most common known cause of familial and sporadic PD, accounting for approx. 5% of individuals with a family history of the disease and 3% of sporadic cases.

3.2.1 α-Synuclein and LRRK2:

α-synuclein expressed highly in the brain, it is a 14 kDa neuron protein. Normally, α-synuclein is typically associated with synaptic vesicles at axon terminals. The physiological role of α-synuclein is unclear (12). The increase in striatal alpha-synuclein levels induces increased Lrrk2 mRNA levels which suggested that Lrrk2 and alpha-synuclein mRNA levels are possibly co-regulated (12). G2019S mutation in Lrrk2 has a significantly greater capacity than wild-type Lrrk2 to phosphorylate alpha-synuclein. Loss of LRRK2, on the other hand, seems to cause both accumulation and aggregation of SNCA. Co-immunoprecipitation showed that endogenous LRRK2 and α-synuclein have interaction in cells, mouse and human brain tissue.

3.2.2 SNCA; α-Synuclein:

α-Synuclein is a neuronal protein that is linked genetically and neuropathologically to Parkinson's disease (PD). α-Synuclein contributes to PD pathogenesis in a number of ways, but it is generally thought that its aberrant soluble oligomeric conformations, termed protofibrils, are the toxic species that mediate disruption of cellular homeostasis and neuronal death, through effects on various intracellular targets, including synaptic function (12).

α-Synuclein is a member of the synuclein family of proteins, which also include β- and γ-synuclein. The only difference of α-synuclein from the other members structurally is the NAC region. All three are neuronal proteins that localize preferentially to presynaptic terminals. Point mutations in β-synuclein in rare cases causes DLB and neurodegeneration.

But wild-type β-synuclein is protective in various settings against α-synuclein-mediated neurodegeneration

The SNCA gene encodes for a 140 amino acid protein, which in aqueous solutions does not have a defined structure, hence the term "natively unfolded protein." α-Synuclein does form α-helical structures on binding to negatively charged lipids, such as phospholipids present on cellular membranes, and β-sheet-rich structures on prolonged periods of incubation. The protein is composed of three distinct regions:

(1) an amino terminus (residues 1–60), containing apolipoprotein lipid-binding motifs, which are predicted to form amphiphilic helices conferring the propensity to form α-helical structures on membrane binding,

(2) a central hydrophobic region (61–95), so-called NAC (non-Aβ component), which confers the β-sheet potential, and

(3) a carboxyl terminus that is highly negatively charged, and is prone to be unstructured. There are at least two shorter alternatively spliced variants of the SNCA gene transcript, but their physiological and pathological roles have not been well characterized.

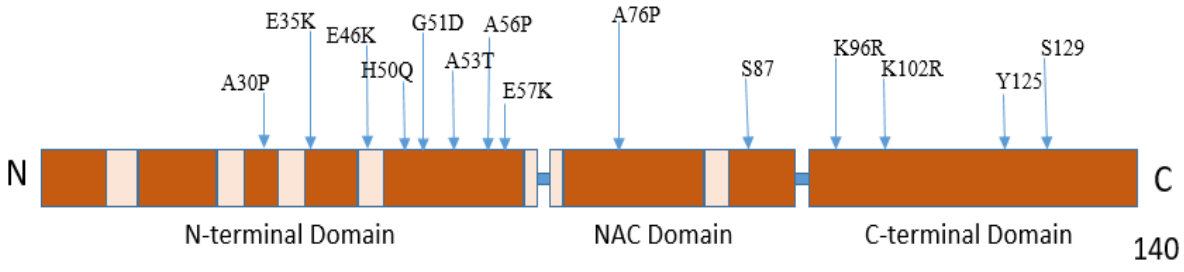


Figure 6: Scheme representing the structure of human ASYN with the three distinct domains (N-terminal, NAC and C-terminal).

3.2.3 Effect of mutations on ASYN oligomerization and aggregation:

From Epifluorescence microscopy known that, all the ASYN variants formed oligomers in HEK cells.

The aggregation paradigm shows that, the seven ASYN mutations (A30P, E46K, A53T, E35K, E57K, TP and Y125F) had the most pronounced effects for subsequent analysis. ASYN mutants could be examined using different assays, including toxicity measurements, biochemical analysis of ASYN, ASYN secretion, degradation pathways, in order to obtain detailed information on the cellular effects of specific types of ASYN accumulations.

3.2.4 LRRK2 (Leucine-rich repeat kinase 2):

LRRK2 is a member of the leucine-rich repeat kinase family. Variants of this gene are associated with an increased risk of Parkinson's disease. A large number of novel *LRRK2* mutations have been described as putative causes of PD.

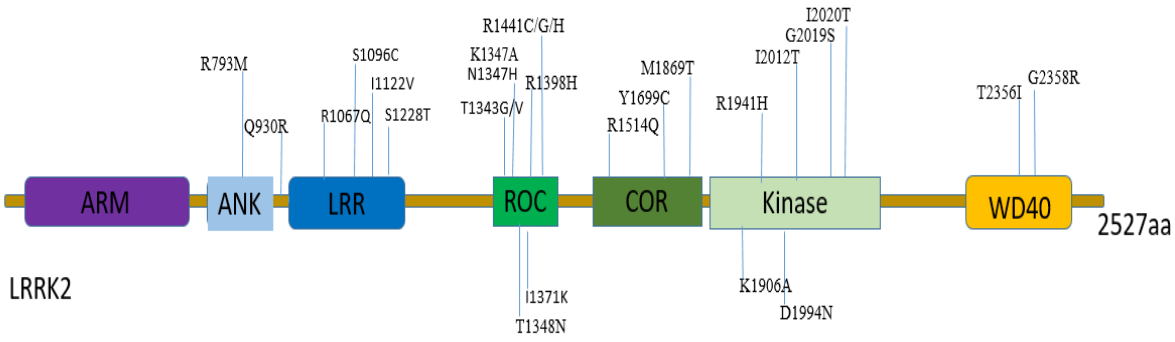


Figure 7: Structure of the leucine-rich repeat kinase 2 protein and functional domains and PD-linked point mutations

3.2.5 Normal function of LRRK2:

The LRRK2 gene provides instructions for making a protein called dardarin. The LRRK2 gene is active in the brain and other tissues throughout the body.

Dardarin has an enzyme function known as kinase activity. Proteins with kinase activity assist in the transfer of a phosphate group (a cluster of oxygen and phosphorus atoms) from the energy molecule ATP to amino acids in certain proteins. Dardarin also has a second enzyme function referred to as a GTPase activity. This activity is associated with a region of the protein called the ROC domain. The ROC domain may help control the overall shape of the dardarin protein.

3.2.6 Mutation in LRRK2 causes PD:

Mutations in Leucine-rich repeat kinase-2 (*LRRK2*) cause a genetic form of PD and have implications in sporadic PD. LRRK2 is a cytoplasmic protein and contains both GTPase and kinase domains. Mutations in the catalytic Roc-COR and kinase domains of LRRK2 are considered common causes of familial PD. LRRK2 mutations causes PD with age-related penetrance and clinical features identical to late-onset sporadic PD. There is an increase in LRRK2 kinase activity for kinase domain and a decrease in GTPase activity for Roc-COR mutations.

Mutations in the LRRK2 gene and especially the common mutation G2019S may account for 4–8% of familial and 1–3% of sporadic cases of Parkinson's disease, including those of early and late onset. The LRRK2 gene is situated on chromosome 12p11.2–q13.1. Dardarin contains several functional domains, including a leucine-rich repeat domain, WD40, renin-angiotensin system/guanosine triphosphatases and kinase domains. The presence of the leucine-rich repeat and WD40 domains suggests a role in protein-protein interaction. Frequencies of G2019S LRRK2 mutations vary significantly across different ethnic groups:

Caucasian: 5% of familial PD cases and approximately 2% of apparently sporadic PD cases.

Ashkenazi Jewish: 40% of familial and 13% of sporadic

North African Berber Arabs: 39% of familial cases and 40% of sporadic cases.

Asian population shows that 2 protein-coding variants, G2385R and R1628P, increase the risk for PD approximately by 2-fold; these mutations are present at a frequency of approximately 6% in cases.

4. Objective 2:

To find small molecules to inhibit GTP binding and kinase activity.

To find the small molecule, several chemical databases like PubChem, drug bank etc. along with the literature study is done and the suitable physio chemical properties of the phytochemicals are studied. The Roc-COR domain LRRK2 gene responsible for PD is taken to study the GTP binding and kinase activity.

For inhibiting kinase activity several kinase inhibitors are studied of which I have selected Nilotinib and Sunitinib based on the idea of their interactions and pharmacology. Also as it has been studied to some extent for other Neurodegenerative diseases such as Alzheimer's. The GTP binding inhibitor compound 68 and FX2149 which can reduce LRRK2 GTP binding activity and can reduce LRRK2 kinase activity and protect against mutant LRRK2 toxicity.

The structure of small molecules is drawn in 2D sketcher from the Schrodinger software which is then converted to 3D for the docking studies.

4.1: 2D structure of small molecules:

i. **Nilotinib:** Kinase inhibitor
   Molecular Formula: $C_{28}H_{22}F_3N_7O$
   Molecular weight: 592.527

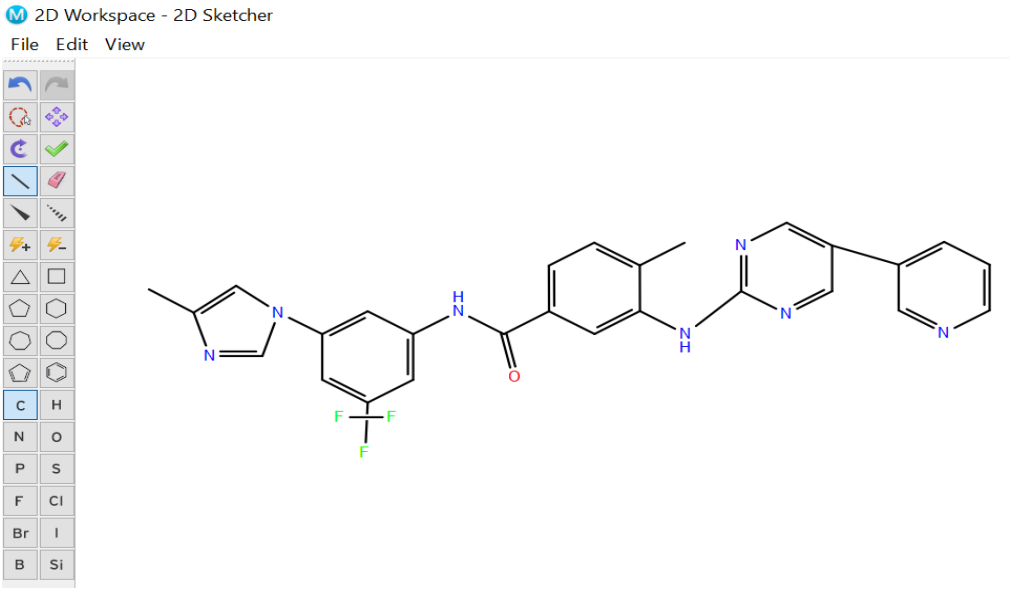


Figure 8: Structure of Nilotinib.

ii. **Sunitinib:** Kinase inhibitor

Molecular Formula: $C_{22}H_{27}FN_4O_2$

Molecular weight: 398.482

Structure:

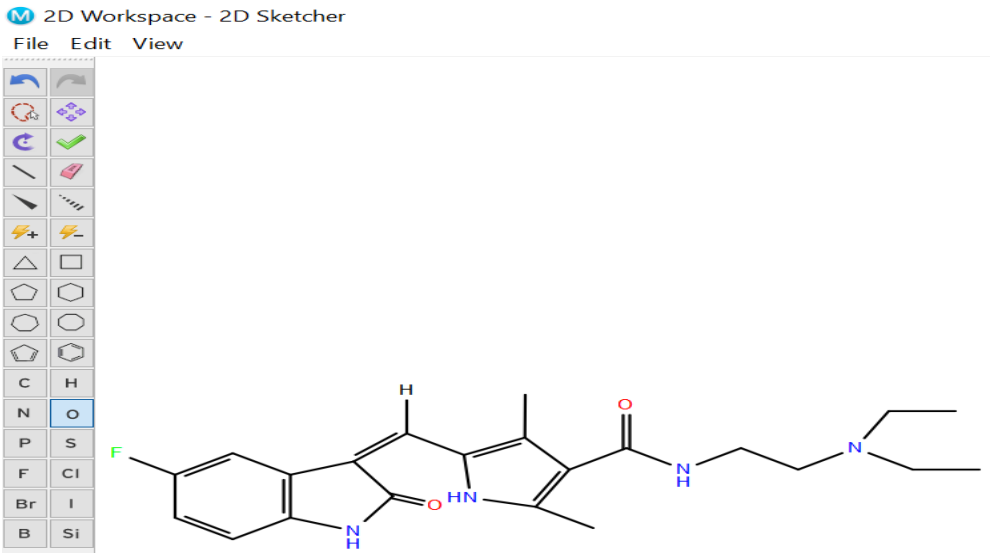


Figure 9: Structure of Sunitinib.

iii. **Compound 68:** GTP binding inhibitor

Molecular Formula: $C_{16}H_{19}N_2O_4S$

Molecular weight:334

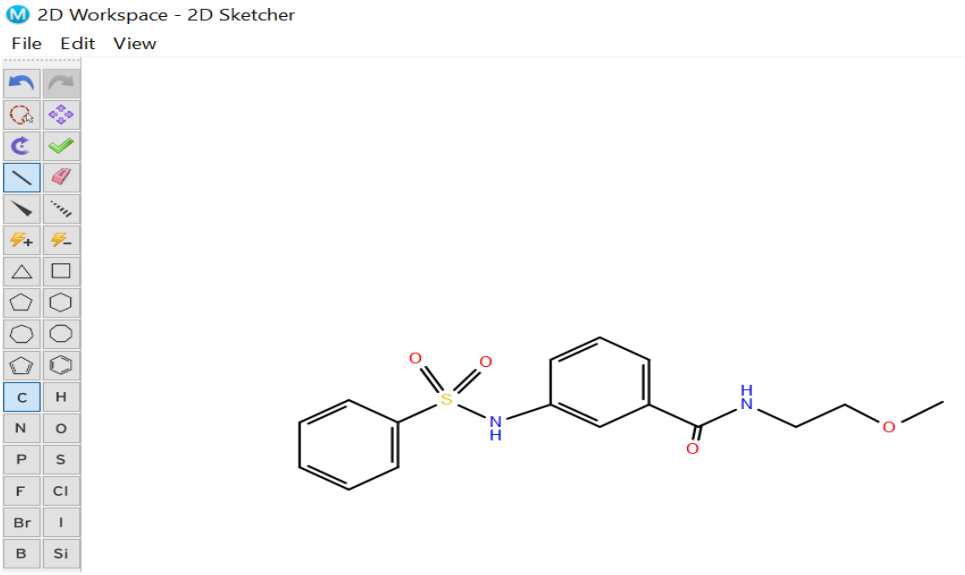


Figure 10: Structure of compound 68.

iv. **FX2149:** GTP binding inhibitor

Molecular Formula: $C_{15}H_{17}N_3O_3S$

Molecular weight:319

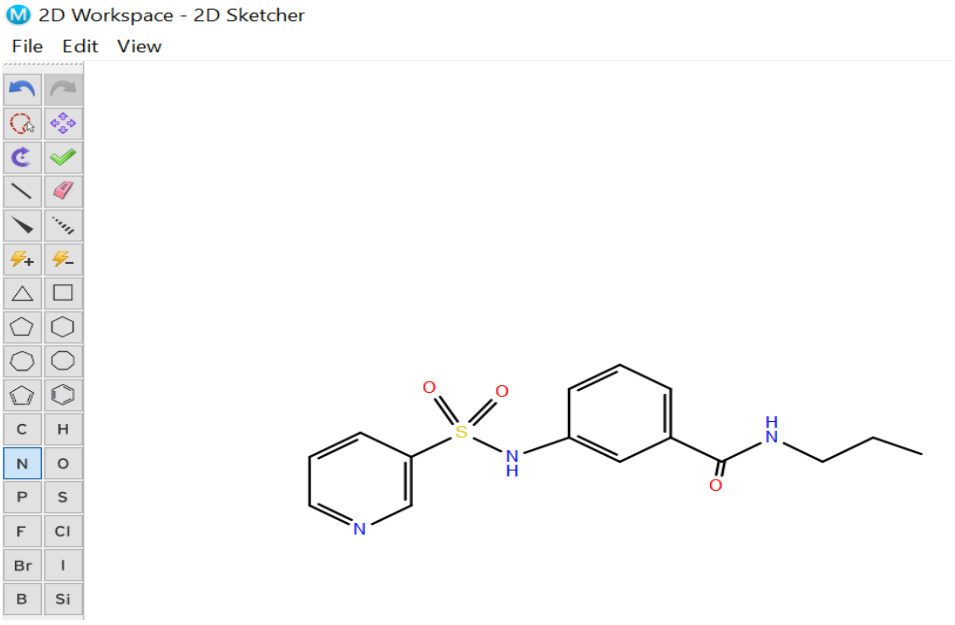


Figure 11: Structure of FX2149.

5. Objective 3:

Docking study with the protein and small molecule.

The docking and scoring are done with help of Schrodinger software. The Protein (LRRK2) structure is taken from UniProt. Which covers the Roc domain of LRRK2 that is residues from amino acids 1333–1516. PD-linked mutations within the GTPase domain alter either GTP binding or GTPase activity. GTP binding by the K1347A mutation alters LRRK2 kinase activity. As the structure of COR region of GTPase domain is not available Homology modelling is done to build the structure in Swiss Model.

5.1 Introduction to software:

The following gives a brief introduction of the software that is used to derive the results of Docking and Scoring.

Schrodinger:

A software that is considered as the scientific leader in developing the state-of-the-art chemical simulation for use in pharmaceutical, biotechnology, and materials science research.

Maestro:

Maestro is the portal to all of Schrodinger's computational technology. The powerful, unified portal to modelling small-molecules, biologics, and materials with Schrodinger technology**.** Far more than just a user interface, Maestro also helps researchers organize and analyze data.

Specific buttons, menus, and other controls that can be found in the Maestro user interface. The image below shows the maestro portal. The interface consists of three main regions: the selection toolbar, with its selection tools and primary actions; the workspace navigator with its two panes for managing and reviewing content.

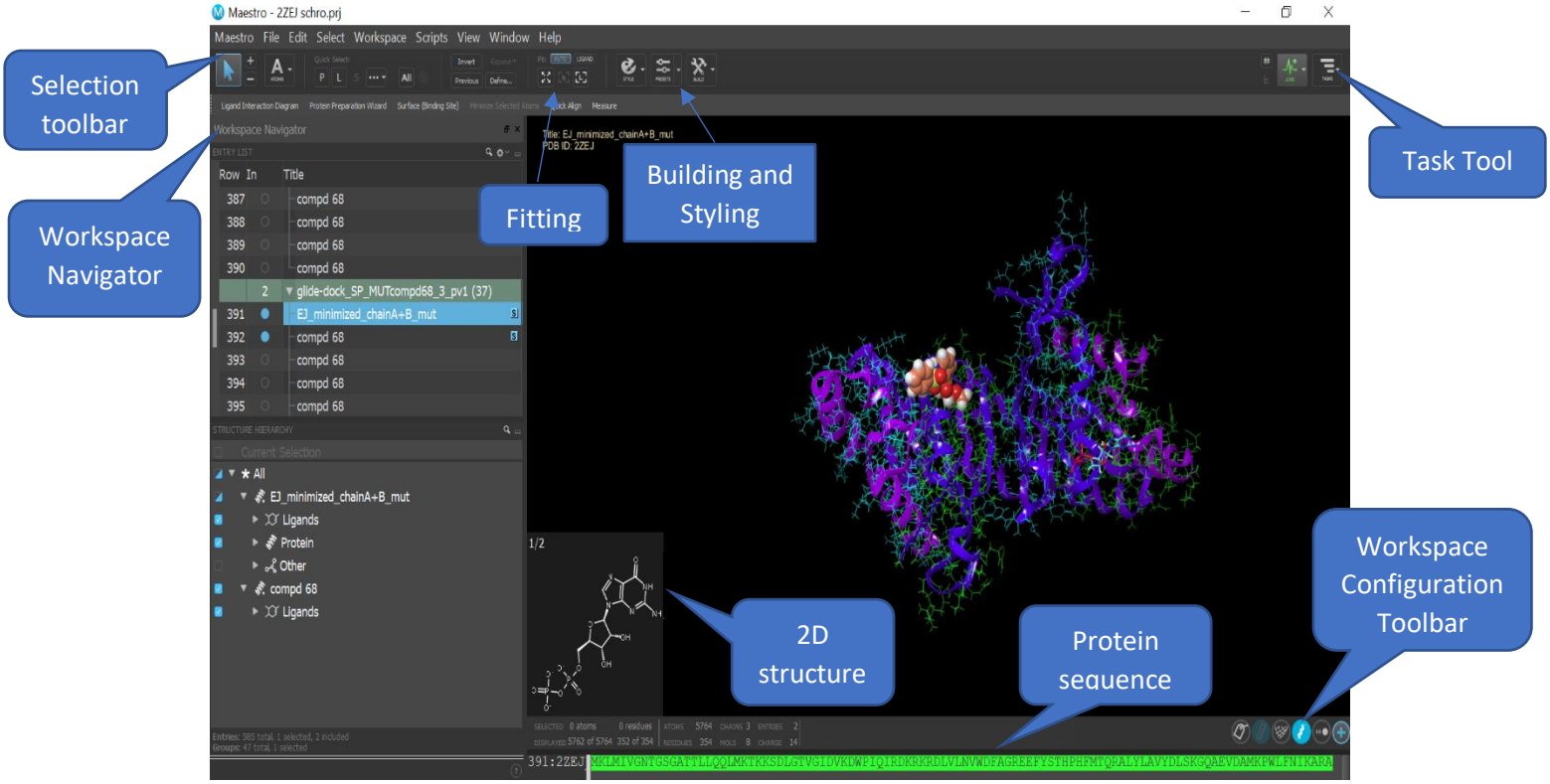


Figure 12: Layout of Maestro.

5.2 Docking:

Also referred to as small molecular docking is a method that predicts the best orientation of one molecule to the other when bound to each other to form a stable complex. Molecular docking is one of the most frequently used methods in structure-based drug design, due to its ability to predict the binding-conformation of small molecule ligands to the appropriate target binding site. Molecular docking is a study of how two or more molecular structures, for instance, drug and catalyst or macromolecule receptor, match along to be a perfect fit. Binding orientation of small-molecule drug candidates to their macromolecular targets predicts the affinity and activity of a given small molecule.

5.3 Methodology:

5.3.1 Protein Preparation:

The typical protein structure file taken from the PDB is not suitable for immediate use in molecular modelling calculation. They usually consist of heavy atoms and may include a co-crystallized ligand, water molecules, metal ion and cofactors. Some structures are multimeric and need to be reduced to a single unit and because of the limited resolution of X-ray experiments, it can be difficult to distinguish between the carboxyl oxygen and secondary amine nitrogen of amides in the crystal structure. So, the placement of these groups must be checked. PDB structure may also present with missing atoms and connectivity information which must be assigned along with bond orders and charges. Therefore, protein preparation is a must before performing molecular modelling.

To do the protein preparation, in the maestro portal click the task toolbar and take protein preparation wizard. The protein preparation wizard consists of a series of tools to prepare proteins in a form that is suitable for modelling and calculations.

There are three tabs in the protein preparation wizard. These wizards consist of all the tools for stages of the protein preparation.

1.Import and process tab: Properties structure can be imported and a basic task for fixing the structure can be performed.

2.Review and modify tab: Unwanted waters and chains can be deleted and fix or delete het groups.

3.Refined tab: It can be used to optimize the orientation of H-bonded groups and minimize the structure.

All these 3-stages should be performed for an unprocessed protein, such as the one from the PDB. When the protein is ready to use, the grid points are set using the Glide Grid Generation.

5.3.2 Ligand Preparation:

To prepare the ligands to use in computational studies such as Schrodinger modeling application for docking. Go to the task and click on LigPrep. The LigPrep panel opens which shows several options used for preparing the ligand. The ligand structure imported in the

workspace is used. So, the option “use structure from project table” is used in the procedure. The other criteria such as ionization state, choosing stereoisomers, desalt, choosing the output format are selected according to the requirement. After completion of which the name of the job is renamed as preferred and the job is run. It takes 10-15 minutes of running time to complete the job. The structure of some ligands appears multiple times represented by different ionization or tautomeric states. In the project table click the property tree and show only the Epik results. The one which has the lowest Epik penalty among the structures represents the more energetically favorable state than the one with more penalty value. Therefore, the structures are fully prepared to be used in further computational studies within the Schrodinger suite.

5.3.3 Protein–ligand docking:

It is the most commonly used docking technique. It predicts the position of a ligand when it is bound to the protein. The ligand might act as an inhibitor or a promoter. In the maestro portal go to the task then click to Glide Docking, when the ligand docking window is open, in the option receptor grid, browse the protein whose grid point is set in the Receptor Grid Generation. Next, in use ligand from option select project table selected entries. Keep the scaling factor 0.08. Rename the job name and run the job. The running time of the job is about 10-20 minutes. Once the job is done the docked product will appear in the workspace, which is further used for a binding affinity study using Prime MM-GBSA.

5.3.4 MM-GBSA:

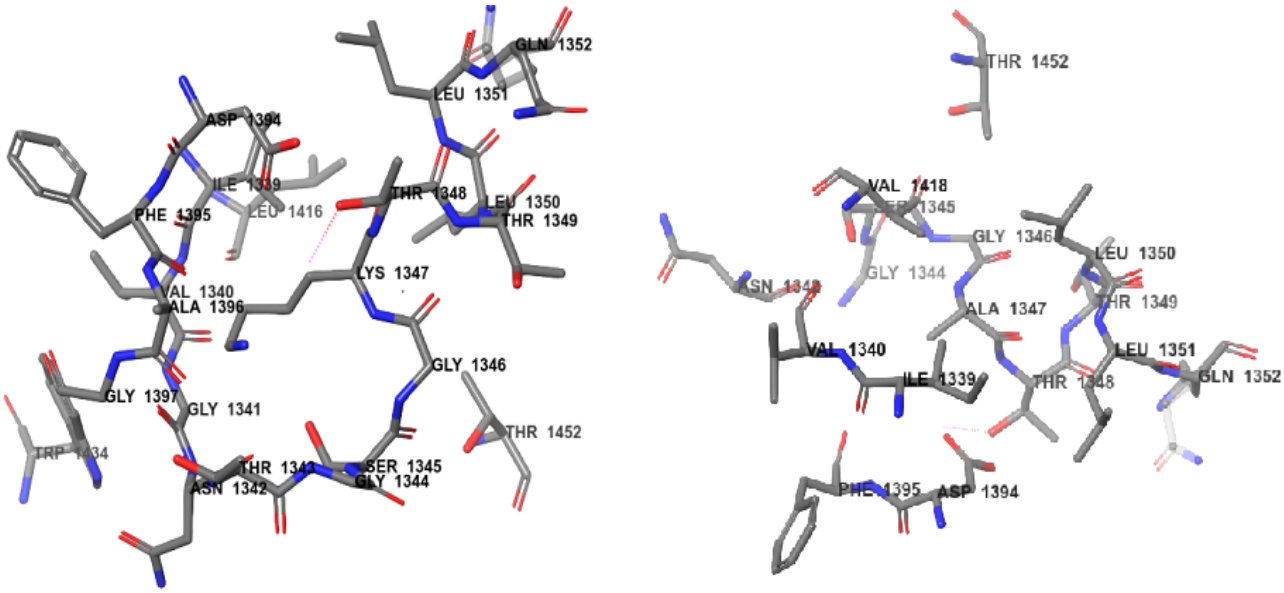
Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) calculations. It is a method to calculate relative binding free energy for molecules by combining molecular mechanics calculations and continuum (implicit) solvation models. The absolute values calculated are not necessarily in agreement with experimental binding affinities. The ranking of the ligands based on the calculated binding energies (MMGBSA DG Bind) can be expected to agree reasonably well with ranking based on experimental binding affinity, particularly in the case of congeneric series.

The job running time of MM-GBSA takes about 9-10 hours. A more negative value indicates stronger binding as the MM-GBSA binding energies are approximate free energies of binding.

6. Results:

6.1 Roc domain study:

The Roc domain of the LRRK2 protein is studied and one of the mutations (K1347A), that is playing a role in causing the PD is considered in this study. K1347A is said to impair the GDP/GTP binding. The most common PD-linked mutation, G2019S, has abnormally elevated kinase activity. The GTP binding by the K1347A mutation increases LRRK2 kinase activity, which is the reason that causes PD.



(A) (B)

Figure 13: ROC domain of LRRK2 protein. (A) Wild Type Lys1437 residue and (B) Mutation of LYS to ALA1347.

6.2 Docking of small molecule and Binding affinity analysis:

The kinase inhibitor Nilotinib and the GTP binding inhibitor FX2149 and Compound 68 taken as a small molecule are prepared in the Maestro portal's LigPrep. Likewise, the Roc domain of the LRRK2 protein is prepared by the Protein preparation wizard. Through these, the small molecules' and the protein's unwanted particles are removed and are made compatible for Docking. The already present ligand of the protein of interest is been removed to dock with the selected inhibitors. The Grid point where the molecule is to be docked is set by using the Receptor Grid generation. The results of which are saved as glid_grid_4 for Wild Type and glid_grid_7 for Mutant protein. Finally, these two completely prepared proteins are further used in docking process. The docking is done by incorporating the small molecule in the place of the removed ligand that was already present in the raw protein and whose grid has been set for the area of interest in the protein. Thereafter, the binding affinity of the small molecules to the protein is calculated by the method MMGBSA. Figure 14, figure 18, and figure 22 shows the ΔG value of wild type protein with the Nilotinib, FX2149 and Compound 68 complex respectively. Similarly, figure 16, figure 20, figure 24 shows the ΔG value of mutant protein with the Nilotinib, FX2149 and Compound 68 complex respectively.

Row	In	Title		Prime MMGBSA Input Structure Hash	MMGBSA moved residues	MMGBSA dG Bind
		prime_mmgbsa_WTnilotinib_2-out1 (14)				
267		644241	0.142	76d428d7173c6fb2a13a85c25c1a05a66783bc11e0bff05bba...	A:1343,A:1344,A:1353,A:1396,A:2,B:...	-31.916
268		644241	9.604	f79c5e16a3f619aa0b14f1c8048518089a4a9c0f3c1412e8cd8...	A:1343,A:1344,A:1346,A:1349,A:135...	-54.575
269		644241	0.341	0425b275d8eb9f0a1d9e075e5418a65d447754a577422912b...	A:1349,A:1368,B:1455,B:1490,B:1491	-58.870
270		644241	0.077	4b4c3f36c0e24f7e659de1c5d8de1b79e8e95c5acdd68fea934...	A:1346,A:1349,A:1368,A:2,B:1455,B:...	-0.733
271		644241	0.021	00ca0dd198e590364bfba3b5a622d71179f6ba724f1c88425a...	A:1343,A:1344,A:1368,A:1369,A:2,B:...	-52.335
272		644241	0.198	93cabc3f22ea09872e589630ae6b11d5b4c250e691181684e3...	A:1349,A:1368,A:2,B:1490	-27.055
273		644241	9.733	6082768c7555e2f40031079e95426f5270a12b5b1d1ac1368c...	A:1348,A:1349,A:1368,A:1396,A:2,B:...	-83.683
274		644241	9.865	96bbbf2fb0bc5f62570c918f19b1eca8f609c64397f38afc22aef...	A:1343,A:1344,A:1348,A:2,B:1455	-4.137
275		644241	0.091	0f2539fa94f4e21ac086b44e2d6d1182cc8499c20a463420d7...	A:1346,A:1349,A:1368,A:2,B:1490,B:...	-4.491
276		644241	9.553	e4ec69f78025d57c9e15db2ba9ebfa19370bdfc60bf68a08d8a...	A:1343,A:1344,A:1368,A:1396,A:2,B:...	-78.762
277		644241	9.937	47948dd4b00dc2ade1d7037eaf5a697fc68504523c9c6782e9...	A:1344,A:1396,A:2,B:1455,B:1491	-50.069
278		644241	0.048	2b69af3e9e3495e13149102aa199b9ead4de81217f826c258a...	A:1347,A:1348,A:1396,A:2,B:1455,B:...	-89.673
279		644241	9.743	e94060d3b57931330b63365d1948f8b712bc9ee7cc08217b9...	A:1343,A:1344,A:1396,A:2,B:1490,B:...	-46.788
280		644241	0.086	bd2630481e81e3f1574c81692a4b98604a6f272e9bdf23c6f7...	A:1343,A:1344,A:1348,A:1396,B:1491	-27.526

Figure 14: ΔG value of WT Roc domain and Kinase inhibitor Nilotinib complex calculated by MMGBSA method.

From the MMGBSA study, the two highest negative value is selected as it indicates stronger binding since the MM-GBSA binding energies are approximate free energies of binding. The orientation with the highest negative value in the binding of these small molecules is considered to be the most ideal. Then the two structures are studied further, and their interaction between

the wild-type domain, and small molecules as well as how they differ with the mutant domain, and small molecule interaction is compared. Figure 15, Figure 19, Figure 23 shows the interaction between wild-type Roc domain of LRRK2 protein with the kinase inhibitor Nilotinib, and GTP binding inhibitors FX2149, and Compound 68 respectively. The interaction between mutant Roc domain of LRRK2 protein with the Kinase inhibitor Nilotinib, and GTP binding inhibitors FX2149 and Compound 68 are shown in the Figure 17, Figure 21 and Figure 23 respectively. The binding affinity is good enough to say that the small molecule inhibitors would work as a good inhibitor and could be used to inhibit the GTP binding to stop the kinase activity that is said to causes the PD.

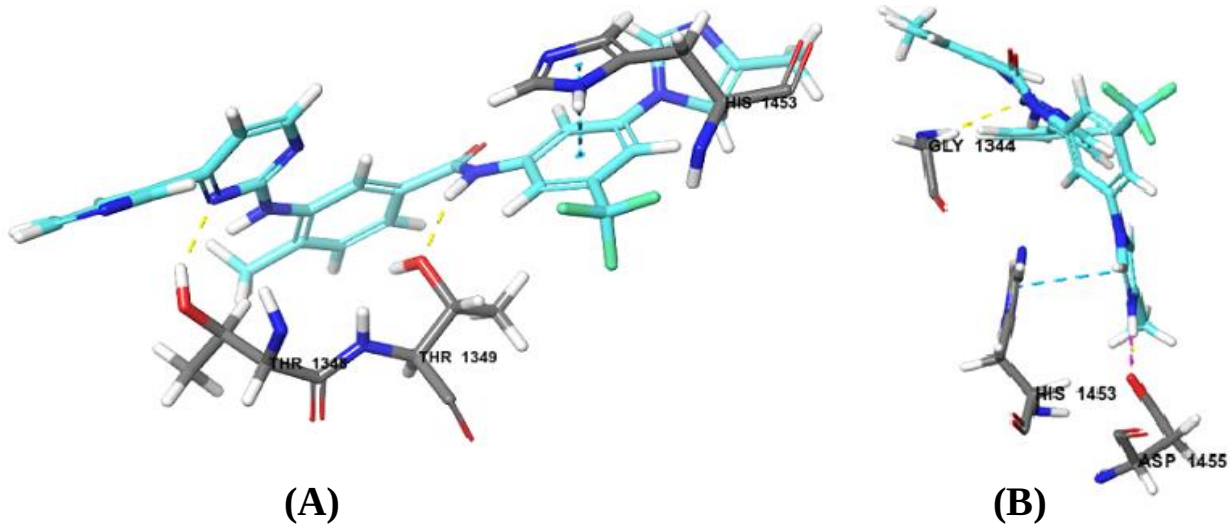


Figure 15: Interaction between wild type Roc domain of LRRK2 protein with the Kinase inhibitor Nilotinib.

Row	In	Title		Prime MMGBSA Input Structure Hash	MMGBSA moved residues	MMGBSA dG Bind
		prime_mmgbsa_MUTnilotinib_3-out1 (19)				
283		644241	9.298	7ccc87c09d076eef536e6fd34f681d5cbb98d164bce26a0f1ce...	A:1348,A:1352,A:1368,A:1369,A:139...	8.574
284		644241	9.238	1afb4afd4ccf5b6494336771fbfcded7f384d6dfc722f4fb95c9d...	A:1344,A:1368,A:1396,B:1489	-42.961
285		644241	9.354	89adca9f1e1f8d357c0f456ac8baa5ed5319bcb3627e44493ab...	A:1348,A:1352,A:1368,A:1369,A:139...	7.791
286		644241	8.904	3bf84d31b7ad15cfc6720815f1a7ddc2b4013a2890472faf59b...	A:1343,A:1349,A:1352,A:1353,A:1396	-36.679
287		644241	8.839	3bcab8208eb8b56edee28618464b71a9db9576fad5a4b39fe2...	A:1368,A:1369,A:1396,A:2	-45.770
288		644241	8.845	c63ee22406ce3f2ee419df8eb9bd160998e75bb15b767e143f...	A:1344,A:1348,A:1352,A:1368,A:136...	-51.472
289		644241	9.062	909d9a605c4db29fd19d4107a71da77e86fbd267b4e2bbe9b...	A:1343,A:1368,A:1369,A:1396,A:2,B:...	-74.319
290		644241	8.913	5ef4e732af09224931312761926ee96e65b2325da743b5b07ee...	A:1349,A:1368,A:1396,B:1455,B:149...	-11.713
291		644241	8.963	714ea446b57e40cf86134b829c38cc1e2a1c4e7385835cd7f0...	A:1343,A:1344,A:2,B:1455	-38.220
292		644241	9.266	c15510b8ea798bfd433565867d2d94ae0beff116ad854a812f...	A:1344,B:1490	4.154
293		644241	9.250	76bcbaa2664d1ef43b40593db5101d753e602c2a6cf614d1cd...	A:1344,B:1455,B:1490	-1.317
294		644241	9.016	555da5853ae3028109ea899ace8a5c3c58ee06baaaf5a53d09...	A:1344,B:1490	-35.642
295		644241	9.259	ed3ecc40eaf4285c4d36c4fb060c17b7597630571e322c58dd...	A:1344,A:1396,A:2,B:1455	-52.573
296		644241	9.427	7d1d5e060db6e3657328d086b402cf88fbddda535f21ae269e...	A:1344,A:1396,A:2,B:1490	-12.778
297		644241	9.184	76637e74fdfd6f22bfd0a3171c529e988e3acfa063fc81d4f8f5...	A:1343,A:1368,A:1369,A:1396,A:2,B:...	-58.875
298		644241	9.203	17902d747b5a85f4b39ea5a1d523b7598b13bd1527d4eaaab...	A:1368,A:1396,B:1490	-26.049
299		644241	8.881	ef806938fa08933c6d292a231afc9227637c4b0cff40fb1e8cd7...	A:1344,A:1396,B:1490	-40.091
300		644241	9.453	4b3344958637ccbfe47c597fd747509f3644ac9422be5da58e...	A:1344,A:1347,A:1396	30.428
301		644241	9.161	7b8aee14d4345db20665babf2815f088cb347270f474de233f...	A:1344,A:1347,A:1396,B:1490	-38.161

Figure 16: ΔG value of Mutant Roc domain and Kinase inhibitor Nilotinib complex calculated by MMGBSA method.

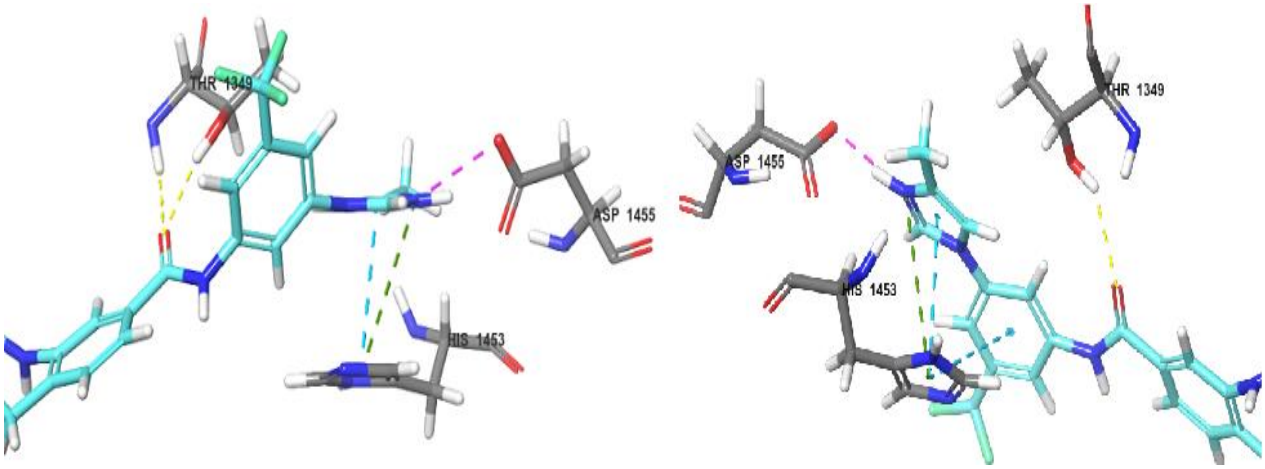


Figure 17: Interaction between mutant Roc domain of LRRK2 protein with the Kinase inhibitor Nilotinib.

Row	In	Title		Prime MMGBSA Input Structure Hash	MMGBSA moved residues	MMGBSA dG Bind
		prime_mmgbsa_WTfx2149_13-out1 (16)				
332		FX2149_str	9.703	7ffa026c54961ee1963648a55715d8ef86655e148350b8fe20...	A:1343,A:1349,A:1368,A:2	-58.404
333		FX2149_str	9.637	765cf95e88788135626e7229bb0e2ecd72170c2ed0f48f40c9...	A:1343,A:1349,A:2	-61.954
334		FX2149_str	9.476	751bf91e5ff0145f5f6bae1c6bc7937b9e77f89279139cbd451...	A:1343,A:1349,A:2	-56.006
335		FX2149_str	9.609	bdadbe8fef8d3f03989e1be7bb17911133fb4ca0c66295fddd9...	A:1347,A:1349,A:1368,A:1396,A:2	-47.879
336		FX2149_str	9.822	09dbb1cb0459908b578cd08cc59f6505d920d54946f7f88234...	A:1348,A:1368,A:1369,A:2	-46.352
337		FX2149_str	9.733	ce186ad56fbbc0225f7e137e5602c44c2c985377d5b071daa7...	A:1344,A:1368,A:2	-32.925
338		FX2149_str	9.726	2e9e83cacf3bf9ceda8cad039faa9bf76fc2a499827cb93edde9...	A:1343,A:1349,A:1368,A:2	-24.856
339		FX2149_str	0.075	00e6e60daf17e53a6cda4eb89de17f644b8a59d3c0efd5dbda1...	A:1368,A:1369,A:2	-17.718
340		FX2149_str	9.908	472ac4679e2577ad08fbb89ab00b2b3909b250929008f14fa4...	A:1343,A:1368,A:1369,A:2	-43.358
341		FX2149_str	0.110	331bf5221828bb982cc2fbe6e98f1bdbab240fe42bb22d395c...	A:1368,A:1369,A:2	-1.775
342		FX2149_str	0.274	2372401d2ca1c6695140fb89c63bc6fe4f07e3246189d54012...	A:1368,A:1369,A:2	-21.096
343		FX2149_str	9.877	b35e202cbe3f5b6e6f6f823fb39a42aa3f8d57c20bf44da6808...	A:1348,A:1368,A:1369,A:2	-24.340
344		FX2149_str	9.532	bca4a6dcdfa38d5ea35221e064e8c196312c036c935eb3c647f...	A:1349,B:1490	-84.338
345		FX2149_str	0.246	6cb55fd389234968921885 85bdf8a8cddc8d29e890136a10c2...	A:1344,A:1348,A:1369,A:2	-0.195
346		FX2149_str	9.742	9de9a8f2e6d7c0bf681c1d1887dee34ed77cb1f43a8b752bc49...	A:1344,A:1368,A:1369,A:2	9.996
347		FX2149_str	9.793	49532851f6f044815dd2879b6dedc284bcd3fe4f75ad54722f5...	A:1343,A:1344,A:1347,A:1348,A:1368...	-61.332

Figure 18: ΔG value of wild type Roc domain and GTP binding inhibitor FX2149 complex calculated by MMGBSA method.

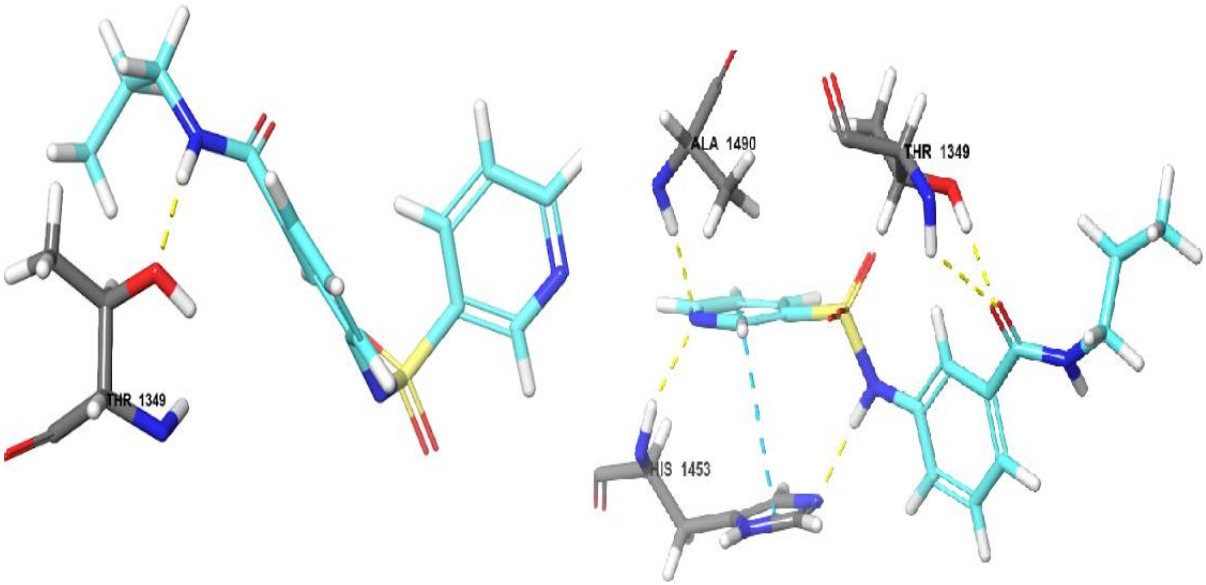


Figure 19: Interaction between Wild Type Roc domain of LRRK2 protein with the GTP binding inhibitor FX2149.

Row	In	Title		Prime MMGBSA Input Structure Hash	MMGBSA moved residues	MMGBSA dG Bind
		prime_mmgbsa_MUTfx2149_12-out1 (17)				
315		FX2149_str	27.615	a616bbf623f493f0cf474260e2f7fd0dcf8c3f41188708c79b92...	A:1343,A:1344,A:1368,A:1369,A:2	-6.207
316		FX2149_str	29.084	6a8c49a3b94c904e5190931c6e644221a0862a02ed54d728f4...	A:1343,A:1344,A:1368,A:1369	-61.626
317		FX2149_str	29.041	08e62a0989f58ef97056098fcc338f0a14f13973ea4673bed13...	A:1343,A:1344,A:1396,A:2	-6.905
318		FX2149_str	29.304	f1139e8011fda562c5e64b65e24e2d0adffd6c5c3f51cb4d228...	A:1344,A:1349,A:1368,A:2	-12.518
319		FX2149_str	29.421	7248a46c64f195bca00dbb7583aa550f69f2a1f0b24ef7de33d...	A:1344,A:1368,A:1369,A:2	-64.798
320		FX2149_str	29.104	fa5aded15874f6471d64a26f90d802ef6d9e1428e60daf2cb8b...	A:1343,A:1344,A:1345,A:1368,A:136...	-47.373
321		FX2149_str	29.296	a6c042b4cc264eb8c44eccef3013191605c5efbc836cd7f9271...	A:1343,A:1344,A:1368,A:1369,A:1396	-45.060
322		FX2149_str	29.483	dbf3ca14605e0dbbc4ef3301007d3e53cc11d930b5b559dbaff...	A:1343,A:1344,A:1368,A:1369,A:2	-16.170
323		FX2149_str	29.359	0b07feb3f7c3d8f0a776388041144becf3560ea198c50e77120...	A:1343,A:1344,A:1368,A:1369,A:1396	-56.855
324		FX2149_str	28.615	87d92fafcccf98cd875d6aafa487153f5a2f37bbbd85bf89ca64f...	A:1343,A:1348,A:1368,A:1369	-12.868
325		FX2149_str	29.203	2742c930d6b6426c23c999faa1df62ad498c8730b7ec7ab1e0...	A:1344,A:1349,A:1368,B:1490	-60.871
326		FX2149_str	28.597	518c28c221196444524e2a9378a9b12aeb45cbd06c69bfcec0...	A:1343,A:1348,A:1368,A:1369	-11.105
327		FX2149_str	28.898	1f9715fba9fd15cfb08415cb7434a885e395bdcba2063b403f4...		-12.035
328		FX2149_str	29.323	4099a3ef320d919d7478fae3b34bd2198edbbcec74be6f9a12...	A:1344,B:1490	-10.697
329		FX2149_str	29.255	985ad0f2195c7354146ca2870de6262865c8b91559b3e831b...	A:1344,A:1348,A:1368,A:2	33.314
330		FX2149_str	25.676	d8b2f7e1fc023f0f98eeeb8a09de444e5f123c2c19cbc25d0feb...	A:1343,A:1344,A:1345,A:1368,A:136...	-10.553
331		FX2149_str	28.924	286238d5569e8d4fbcfe9abf1dfad79e1dc0693e46753fb5c56...	A:1343,A:1344,A:1349,A:1396	-23.342

Figure 20: ΔG value of Mutant Roc domain of LRRK2 protein and GTP binding inhibitor FX2149 complex calculated by MMGBSA method.

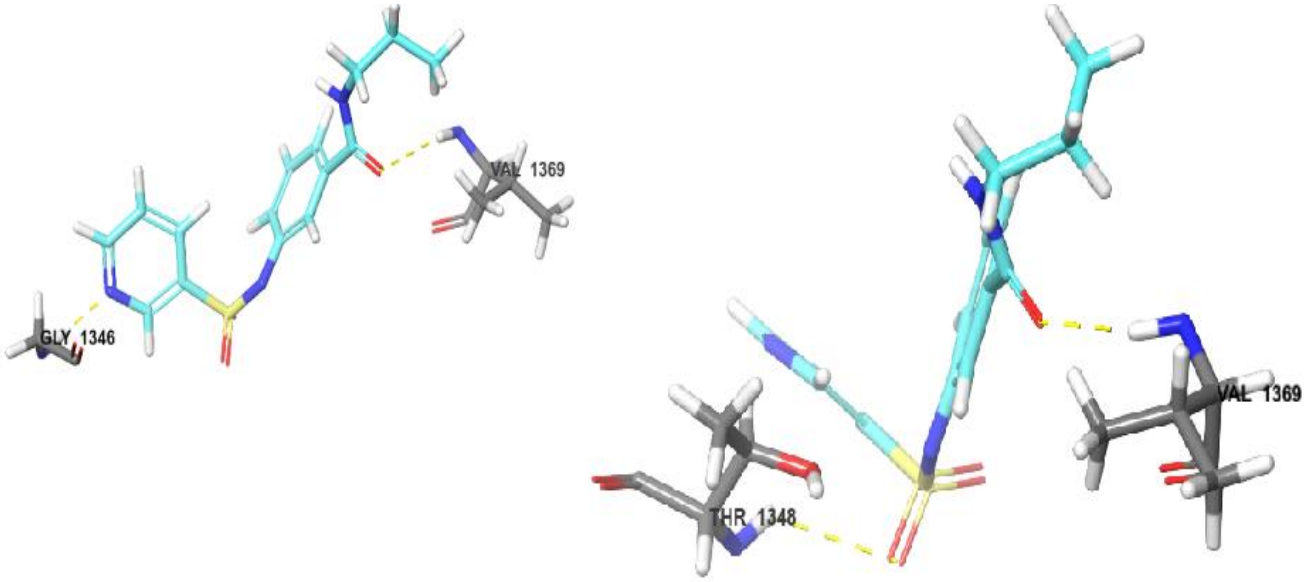


Figure 21: Interaction between Mutant Roc domain of LRRK2 protein with the GTP binding inhibitor FX2149.

Row	In	Title		Prime MMGBSA Input Structure Hash	MMGBSA moved residues	MMGBSA dG Bind
		prime_mmgbsa_1_WTcompd68_2-out1 (38)				
427		compd 68	9.593	76059c083897de4441fe92e4f7cc251162f1bf14ed014b0d2f7...	A:1343,A:1347,A:1396,A:2	-61.866
428		compd 68	9.375	77f41e4eaffdcba7fef75305b3e830ed1085ccac02620dad31cd...	A:1343,A:1348,A:1368,A:1369,A:2	-85.088
429		compd 68	9.288	c1f94e4bceba7b27a218f580677974d15630d308178e5b095f...	A:1349,A:1368,A:1396,A:2,B:1490	-25.115
430		compd 68	9.990	65642c59d54be4e95be9c789bb3b6a9c751aab7305d461e1e...	A:1348,A:1349,A:1369,A:2	-0.470
431		compd 68	9.692	5b8a744be17f09cb34aab8dc89f353f91da3eeb66b4f27021f8...	A:1343,A:1344,A:1348,A:1368,A:136...	-32.411
432		compd 68	9.944	ab2079a10b3c4aa55364848e8610bb0c1700a39ec33a775a9...	A:1368,A:1369,A:1396,A:2	-29.382
433		compd 68	9.906	768c47c4e559dc9d4b1eeca5c7040fa57b7d10fccae11ddcb91...	A:1348,A:1349,A:1368,A:1369,A:139...	-19.877
434		compd 68	0.371	73aa57562e60e7bfe1f9531d89063d5aa493f41ebc2f8724a13...	A:1347,A:1348,A:1368,A:1369,A:2	-59.812
435		compd 68	9.480	2f1e5cfc0ab0533a6c2fc5d3af4538ab5a4234a1d06f7496fdd5...	A:1344,A:1349,A:1368,A:2	-21.613
436		compd 68	9.935	41ffb6cdb1bf2caf05a9a1edb7b245ffb9681aaad8e44c98b5ff...	A:1344,A:1368,A:1369,A:1396,A:2,B:...	-26.428
437		compd 68	9.301	d351c037f3431ca3db76d47355de392c00a504c0c94c395d15...	A:1349,A:1368,A:2	-20.518
438		compd 68	0.016	bc65dd754edf14bf039c5551d9487a88aa0b1c5ebc9e323bf7...	A:1343,A:1347,A:1348,A:1349,A:1368...	-45.431
439		compd 68	9.954	5d82162a35bd9a417d448a4a78ed4c2cf747d1ba000b6fa6d4...	A:1348,A:1349,A:1368,A:1369,A:2,B:...	-49.027
440		compd 68	9.871	4c922286402164f08266e54a01a395477a74f9d58a61cf34b6...	A:1348,A:1368,A:1369,A:2	-50.775
441		compd 68	0.033	6ad7f8417f5adcee8372947e4e8b7ef8f9476e9f1316486b4dc...	A:1343,A:1344,A:1368,A:1369,A:139...	-29.054
442		compd 68	9.815	d637ab60c9e1a88650cb1f57d5555629ebed717fc16f078534...	A:1368,A:1369,A:1396,A:2,B:1491	-45.655
443		compd 68	9.610	81500189a5ec4fcca74b008181029443c97f4452a4c2e7a418...	A:1344,A:1368,A:1369,A:1396,A:2	-13.413
444		compd 68	9.767	1ba6b421360ba7acfc50436e5e16e8f9796563bb4672ce63a5...	A:1344,A:1368,A:1369,A:1396,A:2,B:...	-25.688
445		compd 68	9.895	f603f579cb464dc1e79f42db3b91f275152e89a1386dbaf15c5...	A:1346,A:1368,A:1369,A:2	-61.950
446		compd 68	9.931	4ab3eb751bb09594275972efada53ee0cc63c3503e22bbfa79f...	A:1343,A:1348,A:1368,A:1369,A:1396	-53.873
447		compd 68	9.825	2139b3d3628f485b9bceccde622b35ebd5d527e470263756a...	A:1343,A:1347,A:1348,A:1368,A:1369...	-45.005
448		compd 68	9.575	c76ac92fca8883216250f9799f554131a998ce86a689747f58d...	A:1343,A:1347,A:1368,A:1369,A:1396...	-72.468
449		compd 68	0.031	b127fd56bd9cb241e467b157f945b7ccac1737d0c5a4d8e79f...	A:1368,A:1372,A:2	-29.538
450		compd 68	9.959	8a6592c7b267de638e316e1ede50f1067f4ce437f92e37e5c4f...	A:1344,A:1347,A:1349,A:1368,A:2,B:...	-83.635

Figure 22: ΔG value of wild type Roc domain and GTP binding inhibitor compound 68 complex calculated by MMGBSA method.

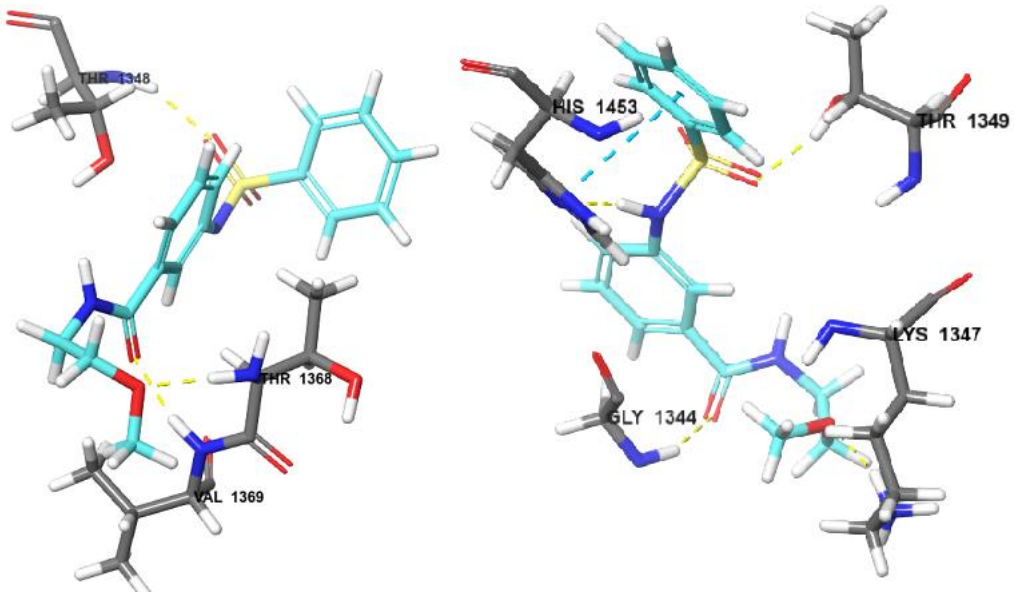


Figure 23: Interaction between WT Roc domain of LRRK2 protein with the GTP binding inhibitor compound 68.

Row	In	Title		Prime MMGBSA Input Structure Hash	MMGBSA moved residues	MMGBSA dG Bind
		prime_mmgbsa_1_MUTcompd68_3-out1 (36)				
465		compd 68	9.540	9a489e74d42d655329a36e1a59d47de3b6b15e7e36f305130...	A:1344,A:1349,A:1395,A:1396,A:2,B:...	-11.697
466		compd 68	8.552	e3f1f86fed50aa83f9f054519b6d16436e7def2b410b7c0ba56f...	A:1343,A:1395,A:1396,A:2,B:1453	-48.394
467		compd 68	9.456	2b565e0fc5505cdb30c5200fd3da1783b71ad67f7ce1e70953...	A:1395,A:1396,A:2	17.278
468		compd 68	9.276	a6de0570ae93fb3b2579d8af3b240ada86acc95456afe9c6054...	A:1344,A:1349,A:1368,A:1395,A:2,B:...	-27.229
469		compd 68	9.244	ead0a2afa09bf85619862a3ac60a8efd9db7fdfa0216d8feb52d...	A:1343,A:1395,A:1396,A:2	-4.931
470		compd 68	9.104	e0b7c4c52d5a234b273b83d0c135074c807491ce171608dab...	A:1343,A:1395,A:1396	-9.376
471		compd 68	9.785	d1e96977e6cc41a963e85b1162224153010e52f0a4d6ffca9a...	A:1344,A:1368,A:1369	12.034
472		compd 68	8.693	702bb3b3216fe91b59203e02df24254fc47e13365cf613f4254...	A:1343,A:1368,A:1369,A:1396,A:2	-17.515
473		compd 68	9.181	da802ce590b832e1a7fad22dbdcc6eb229df6d68fd994823cc3...	A:1368,A:1396,A:2	6.052
474		compd 68	9.426	e786e5b297bf7ae62eb40bbcd5fa0708b5bb18c4c4a622ab3a...	A:1343,A:1348,A:1349,A:1368,A:139...	-2.585
475		compd 68	9.616	9debd5d831b33b160ef80d637f6138e4843004aeae850738cb...	A:1349,A:1368,A:2	-0.442
476		compd 68	8.838	06decddee2829aae56f133ec1594ccaca10a5ce074a7f684e82...	A:1343,A:1344,A:1395,A:1396,A:2,B:...	-27.663
477		compd 68	5.647	b7c482a145bf5008e7b6b0ac4e1e0ae4acbb4a87457321a10b...	A:1368,A:1395,A:1396,A:2,B:1453	-63.163
478		compd 68	9.506	03dc8738fac451f9ffc975abd4cb31883fdbc41b094e4d192ae...	A:1368,A:1395,A:1396,A:2	-7.616
479		compd 68	9.064	d916ff72e1f808b0f2524de184db027e5f446796f7b25aa0efd...	A:1349,A:1395,A:1396,B:1491	-24.231
480		compd 68	9.199	ab42cdbd51184d621a2cc89f35d5a58d2775c3b40c36f5a7df...	A:1344,A:1368,A:1369,A:2,B:1453	-22.033

Figure 24: ΔG value of Mutant Roc domain and GTP binding inhibitor compound 68 complex calculated by MMGBSA method.

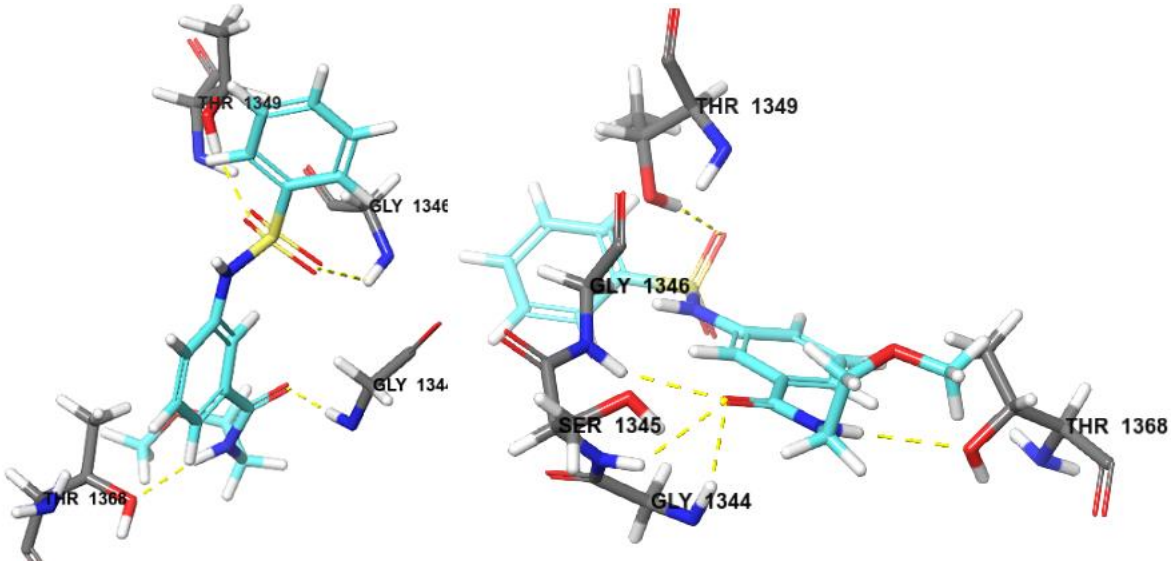


Figure 25: Interaction between Mutant Roc domain of LRRK2 protein with the GTP binding inhibitor compound 68.

Row	In	Title	MMGBSA moved residues	MMGBSA dG B
		prime_mmgbsa_GDPmut7_2-out1 (31)		
814		EJ_minimized_ligand1	A:2,B:1455,B:1490	
815		EJ_minimized_ligand1	A:1343,A:1346,A:1349,A:2,B:1453,B:...	
816		EJ_minimized_ligand1	A:1348,A:1349,B:1453,B:1455,B:1490	
817		EJ_minimized_ligand1	A:1344,A:1349,A:2	
818		EJ_minimized_ligand1	A:1368,A:2	
819		EJ_minimized_ligand1	A:1343,A:1344,A:1348,A:1349,A:135...	
820		EJ_minimized_ligand1	B:1455,B:1490	
821		EJ_minimized_ligand1	A:1348,A:2,B:1453	
822		EJ_minimized_ligand1	A:1348,A:2,B:1453	
823		EJ_minimized_ligand1	A:1343,A:1348,A:1349,A:2	
824		EJ_minimized_ligand1	A:1343,A:1348,A:1368,A:2,B:1452	
825		EJ_minimized_ligand1	A:1344,B:1455,B:1490	
826		EJ_minimized_ligand1	A:1343,A:1344,A:1349,A:1368,B:145...	
827		EJ_minimized_ligand1	A:1344,A:2,B:1455	
828		EJ_minimized_ligand1	A:2,B:1491	
829		EJ_minimized_ligand1	A:1368,A:1369,A:1372,A:1396,A:2	

Figure 26: ΔG value of WT Roc domain and GDP complex calculated by MMGBSA method.

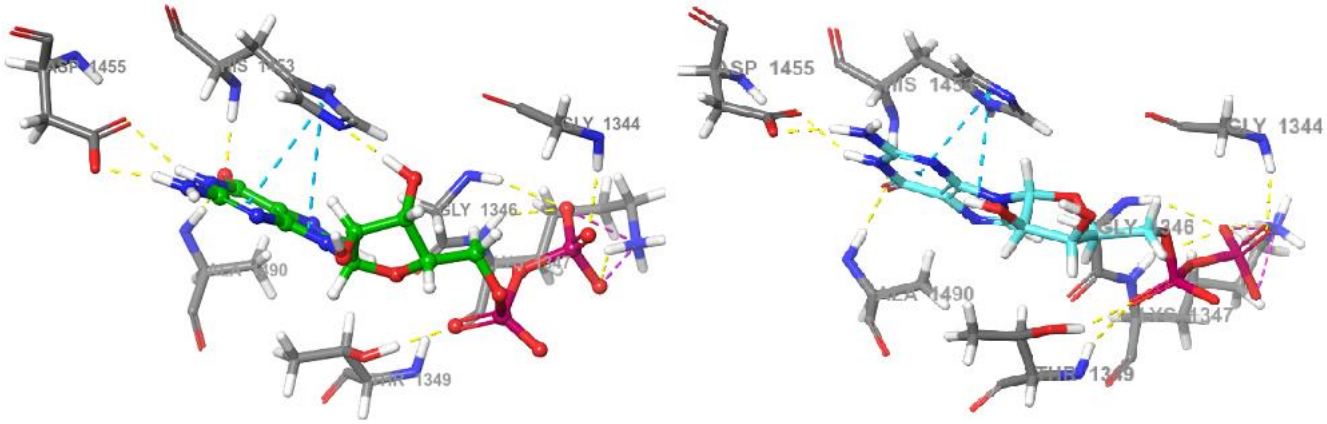


Figure 27: Interaction between WT Roc domain of LRRK2 protein with the GDP.

Row	In	Title	MMGBSA moved residues	MMGBSA dG Bind
		prime_mmgbsa_GDPwt7-out1 (32)		
782		EJ_minimized_ligand1	A:2,B:1455,B:1490	-41.447
783		EJ_minimized_ligand1	A:1346,A:2	-73.952
784		EJ_minimized_ligand1	A:1347,A:1348,B:1455	-28.308
785		EJ_minimized_ligand1	A:1343,A:1344,A:1348,A:1368,B:145...	-18.734
786		EJ_minimized_ligand1	A:1344,A:1349,B:1453,B:1455,B:1490	-53.380
787		EJ_minimized_ligand1	A:1368,A:2,B:1455	-15.166
788		EJ_minimized_ligand1	A:1343,A:1346,A:1347,A:1368,B:1453...	-22.008
789		EJ_minimized_ligand1	A:1349,A:2,B:1455	-74.698
790		EJ_minimized_ligand1	A:1347,A:1348,A:1368	-33.329
791		EJ_minimized_ligand1	A:1343,A:1347,A:1349,A:1368	-9.151
792		EJ_minimized_ligand1	A:1348,A:2,B:1453	8.538
793		EJ_minimized_ligand1	A:1348,A:2,B:1453	3.608
794		EJ_minimized_ligand1	A:1349,A:2,B:1453	22.594
795		EJ_minimized_ligand1	A:1348,A:2,B:1491	-29.917

Figure 28: ΔG value of Mutant Roc domain and GDP complex calculated by MMGBSA method.

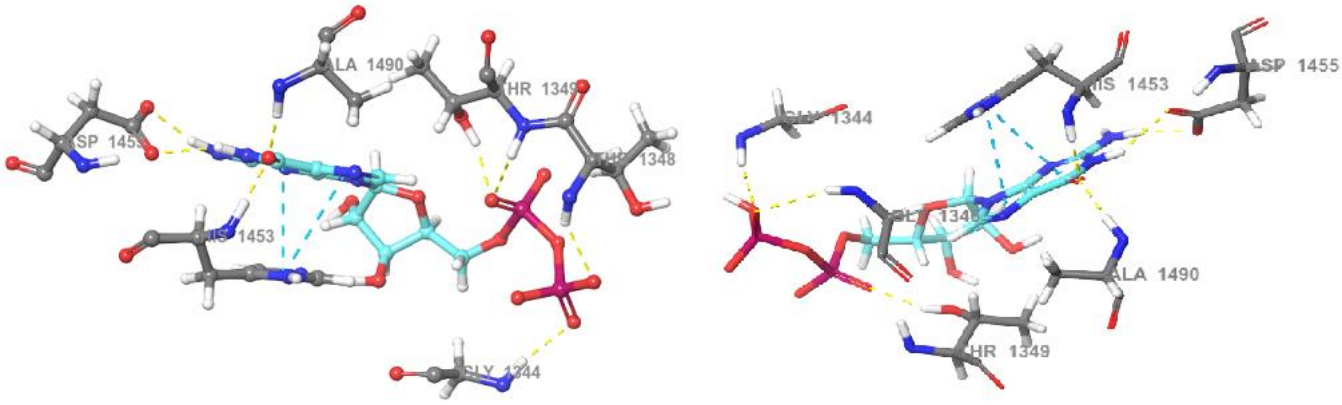


Figure 29: Interaction between Mutant Roc domain of LRRK2 protein with the GDP.

7. Conclusion:

LRRK2 is the main causative protein in PD and an attractive target for therapeutic development. Though the LRRK2-related mechanisms leading to neurotoxicity remain incompletely understood, recent studies have brought to an idea that the kinase activity of

LRRK2 is a key therapeutic target since the most common G2019S mutation increases kinase activity and induces neuronal toxicity. LRRK2 contains an evolutionarily conserved Roc-COR tandem domain, and many familial mutations are clustered within the Roc and COR domains and impair GTPase activity. Therefore, GTPase activity is important for LRRK2 function, for regulating kinase activity, and for the development of PD. The mutation K1347A impairs GTP binding activity and is responsible to increase the kinase activity of G2019S. If the GTP binding could be prohibited by GTP binding inhibitor there might not be the elevation of kinase activity. Potential therapeutic strategies for targeting the Roc-COR tandem domain might include inhibition of GTP binding and kinase inhibition. So, taking these into an account a small molecule inhibitors are studied and the docking study is done.

The molecular docking approach can be used to model the interaction between small molecule and a protein at the atomic level, which allows to characterize the behaviour of small molecules in the binding site of target proteins as well as to explain fundamental biochemical processes. The docking process involves two basic steps: prediction of the ligand conformation as well as its position and orientation within these sites and assessment of the binding affinity. These two steps are related to sampling methods and scoring schemes, respectively. Knowing the location of the binding site before docking processes significantly increase the docking efficiency. Computationally the inhibitors taken for this study is found to be compatible. Although, the validation of each of these strategies is now required in disease-relevant models and may rely upon the future development of small-molecule compounds.

References:

1. Mitochondrial pathobiology in Parkinson's disease and amyotrophic lateral sclerosis; Martin LJ. 2010; doi: 10.3233/JAD-2010-100348.

2. Genetic mutations in RNA-binding proteins and their roles in ALS; Katannya Kapeli, Fernando J. Martinez, and Gene W. Yeo; 2017 Jul 31. doi: 10.1007/s00439-017-1830-7.

3. RNA binding proteins and the genesis of neurodegenerative diseases; Benjamin Wolozin, M.D., Ph.D. and Daniel Apicco; 2015 Dec 29; doi: 10.1007/978-3-319-08927-0_3.

4. RNA-binding proteins in neurodegeneration: Seq and you shall receive; Julia K. Nussbacher1, Ranjan Batra1, Clotilde Lagier-Tourenne2,3, and Gene W. Ye; Trends Neurosci. 2015 Apr;38(4):226-36. doi: 10.1016/j.tins.2015.02.003.

5. Genetics of familial and sporadic amyotrophic lateral sclerosis; Francois Gros-Louis Claudia Gaspar Guy A.Rouleau; https://doi.org/10.1016/j.bbadis.2006.01.004.

6. The ALS disease protein TDP-43 is actively transported in motor neuron axons and regulates axon outgrowth; Fallini C, Bassell GJ, Rossoll W; 2012 Aug 15;21(16):3703-18. doi: 10.1093/hmg/dds205. Epub 2012 May 28.

7. Rethinking ALS: the FUS about TDP-43; Lagier-Tourenne C, Cleveland DW; Cell. 2009 Mar 20;136(6):1001-4. doi: 10.1016/j.cell.2009.03.006.

8. ALS-associated mutant FUS induces selective motor neuron degeneration through toxic gain of function; Aarti Sharma, Lei Lu, Sara Ebrahimi Nasrabady, Margot Elmaleh, Monica Mendelsohn, Adriana Nemes, Juan Carlos Tapia, George Z. Mentis, and Neil A. Shneider; 2016 Feb 4. doi: 10.1038/ncomms10465.
9. RNA-Binding Proteins in Amyotrophic Lateral Sclerosis and Neurodegeneration Scott E. Ugras and James Shorter; Volume 2012 (2012), Article ID 432780, 5 pages http://dx.doi.org/10.1155/2012/432780.
10. Mechanisms of FUS mutations in familial amyotrophic lateral sclerosis YuleiShang, Eric J.Huang https://doi.org/10.1016/j.brainres.2016.03.036.
11. Computational identification and analysis of neurodegenerative disease associated protein kinases in hominid genomes; Saranya Jayapalan, Devika Subramanian, JeyakumarNatarajan; https://doi.org/10.1016/j.gendis.2016.04.004.
12. The role of LRRK2 in Parkinsonism; Jie-Qiong Li, Lan Tan, and Jin-Tai Yu; 2014 Nov 12. doi: 10.1186/1750-1326-9-47.
13. The role of TDP-43 in amyotrophic lateral sclerosis and frontotemporal dementia; Ian R.A. Mackenzie, MD and Rosa Rademakers, PhD; doi: 10.1097/WCO. 0b013e3283168d1d.
14. Motor neuron intrinsic and extrinsic mechanisms contribute to the pathogenesis of *FUS*-associated amyotrophic lateral sclerosis; Jelena Scekic-Zahirovic, Hajer El Oussini, Sina Mersmann, Kevin Drenner, Marina Wagner, Ying Sun, Kira Allmeroth, Stéphane Dieterlé, Jérôme Sinniger, Sylvie Dirrig-Grosch, Frédérique René, Dorothee Dormann, Christian Haass, Albert C. Ludolph, Clotilde Lagier-Tourenne, Erik Storkebaum, and Luc Dupuis; doi: 10.1007/s00401-017-1687-9, 2017 Feb 28.
15. Parkinson Disease and LRRK2; Andrew Singleton, PhD; Chief Editor: Keith K Vaux, MD; Updated: Feb 17, 2016
16. α-Synuclein in PD; Leonidas Stefanis; doi:10.1101/cshperspect.a009399
17. RNA-binding proteins with prion-like domains in ALS and FTLD-U Aaron D Gitler and James Shorter; 2011 Jul 1. doi: 10.4161/pri.5.3.17230
18. Systematic Comparison of the Effects of Alpha-synuclein Mutations on Its Oligomerization and Aggregation; Diana F. Lázaro, Eva F. Rodrigues, Ramona Langohr, Hedieh Shahpasandzadeh, Thales Ribeiro, Patrícia Guerreiro, Ellen Gerhardt, Katharina Kröhnert, Jochen Klucken, Marcos D. Pereira, Blagovesta Popova, Niels Kruse,Published; November 13, 2014; https://doi.org/10.1371/journal.pgen.1004741

19. Proteostasis and Movement Disorders: Parkinson's Disease and Amyotrophic Lateral Sclerosis; Daryl A. Bosco, Matthew J. LaVoie, Gregory A. Petsko, and Dagmar Ringe; doi: 10.1101/cshperspect.a007500
20. GTPase activity plays a key role in the pathobiology of LRRK2; Xiong Y[1], Coombes CE, Kilaru A, Li X, Gitler AD, Bowers WJ, Dawson VL, Dawson TM, Moore DJ.
21. Practical Considerations in Virtual Screening and Molecular Docking; Michael Berry, Junaid Gamieldien, in Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology, 2015.
22. GTP binding is essential to the protein kinase activity of LRRK2, a causative gene product for familial Parkinson's disease; Ito G, Okai T, Fujino G, Takeda K, Ichijo H, Katada T, Iwatsubo T; Biochemistry. 2007 Feb 6;46(5):1380-8.
23. LRRK2 Kinase Activity Is Dependent on LRRK2 GTP Binding Capacity but Independent of LRRK2 GTP Binding; Jean-Marc Taymans, Renee Vancraenenbroeck, Petri Ollikainen, Alexandra Beilina, Evy Lobbestael, Marc De Maeyer, Veerle Baekelandt, and Mark R. Cookson: Philipp J. Kahle, Editor; Aug 12. doi: 10.1371/journal.pone.0023207.
24. A Novel GTP-Binding Inhibitor, FX2149, Attenuates LRRK2 Toxicity in Parkinson's Disease Models; Tianxia Li, Xinhua He, Joseph M. Thomas, Dejun Yang, Shijun Zhong, Fengtian Xue, and Wanli W. Smith; 2015 Mar 27 doi:10.1371/journal.pone.0122461.
25. Novel LRRK2 GTP-binding inhibitors reduced degeneration in PD cell and mouse models; Tianxia Li Dejun Yang Shijun Zhong Joseph M. Thomas Fengtian XueJingnan Liu Lingbo Kong Pamela Voulalas Hazem E. Hassan Jae-Sung Park; https://doi.org/10.1093/ hmg/ ddu341.
26. Molecular Docking: A powerful approach for structure-based drug discovery; Xuan-Yu Meng, Hong-Xing Zhang, Mihaly Mezei, and Meng Cui; 2011 Jun 1; 7(2): 146–157.
27. Practical Considerations in Virtual Screening and Molecular Docking; Michael Berry, Junaid Gamieldien, in Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology, 2015