



# Predictive interactive associate of Monoamine Oxidase-B inhibitors and role of PARK7 in Parkinson's disease.

Nimra Mehmood Malik , and Maria Zeb

<sup>1</sup> Biomedical Engineering and Sciences Department, School of Mechanical and Manufacturing Engineering (SMME), National University of Sciences and Technology (NUST), Islamabad, Pakistan.<sup>2</sup> Biomedical Engineering and Sciences Department, School of Mechanical and Manufacturing Engineering (SMME), National University of Sciences and Technology (NUST), Islamabad, Pakistan.  
Email: Nimra64malik@gmail.com

## ABSTRACT

Background: Parkinson's disease (PD) is an idiopathic neurodegenerative disorder targeting seven to ten million of the elderly population. The PARK7 mutation triggers the loss of DJ-1 protein, leading to the early onset of PD. Objective: Our study aims to identify the PARK7 protein interactions and its potent interacting inhibitor by implying various bioinformatics tools. Methodology: In the present study, expression analysis was performed to determine the PARK7 gene expression levels in the CSF of PD patients. Then a detailed spectrum analysis was performed for the quantification of the PARK7 protein fragment. Further, the ligand interactions of the PARK7 protein and  $\alpha$ -synuclein were studied through LIGPLOT. Several bioinformatic tools were employed to study the protein-protein interactions along with post-translational modifications. Lastly, the technique of molecular docking was implied to detect the potent MAO-B inhibitor of the PARK7 protein. Results: Our study demonstrated modified expression levels of the PARK7 protein. The PARK7 expression analysis displayed a P value of 0.0106 for healthy versus sporadic PD. While healthy versus genetic PD exhibited a P value of 0.0013. Our detailed spectrum analysis illustrated 13 percent sequence coverage for PARK7\_HUMAN across the MS sample. Further, this research introduced three biological replicates: (K) VTVAGLAGK(D);(K) VTTHPLAK(D)& (K) DGLILTSR(G) with 100%; 99%& 99% probabilities, respectively. Moreover, the docking results suggested Safinamide with a -5.4492 best energy score, to be a more potent interacting inhibitor as compared to Selegiline and Rasagiline. Conclusion: The outcome of this study offers early characterization and targeting of proteins, enabling effective diagnosis of PD.

**Keywords :** Parkinson's disease (PD); PARK7 protein, protein expression analysis;Protein-protein interactions (PPI); Monoamine oxidase enzymes;MAO-B inhibitors.

## 1 INTRODUCTION

**P**ARKINSON'S DISEASE (PD) is a progressive neurodegenerative disorder that is linked with the depletion of neuronal function in the substantia nigra, loss of dopaminergic neurons and alpha synuclein aggregation.

PD is characterized by motor traits like bradykinesia with cardinal indications as well as some non-motor traits such as dementia and dysautonomia [1]. Initially PD was viewed as spasmodic disease, however, on several occasion gene mutation was observed to fluctuate the rate of PD. Mutated PARK7 gene trigger the loss of DJ-1 protein activity leading towards early onset of autosomal recessive PD. So far, the major factor of the PARK7 mutation which initiates this pathway is yet to discover [2].

### 1.1 PARK7 protein

The PARK7 gene encodes DJ-1 protein which plays a neuroprotective role against oxidative stress. The PARK7 gene has been reported to act as a chaperone molecule, assists in folding and refolding of damaged proteins, allows transportation of specific proteins to proteasomes, also involves in RNA processing. Recent studies had discovered 25 PARK7 gene mutations associated with PD. The PARK7 gene mutation initiates the formation of unstable DJ-1 proteins, disruption of protein assembly, loss of chaperone activity, and eventually the cell death [3].

The PARK7 protein works by the upregulation of uncoupling mitochondrial proteins named UCP4 and UCP5. These proteins slowdown the process of mitochondrial membrane development through repression of ROS (Reactive Oxidative Species), therefore, plays a neuropro-

tective role in the cell survival. Similarly, when the PARK7 protein works with the regulator of IP<sub>3</sub>R (inositol triphosphate receptor) named Bcl-xL protein. It blocks the release of cytochrome c together with its apoptosis activity. In stressed environment the PARK7 protein dissociates from the Keaps (Kelch-like ECH-associated protein 1) Nrf2 inhibitor. Additionally, it turns on the transcription factor Nrf2 to defend the cell, by triggering the ATP and the NADPH production. The PARK7 prevents phosphatase activity by controlling the PI3K/PKB signalling through PTEN transnitrosylation. The PARK7 protein on one hand regulates the transcriptional process of the p53 by attaching. While on the other hand it prevents the p53 transcription through diacylation [4].

## 1.2 MAO-B inhibitors

Monoamine oxidase (MAO) enzymes are central to the normal functioning of brain. MAO plays a crucial role in neurotransmitter metabolism thus involved in some neurodegenerative diseases including Parkinson's disease. Existence of the isoforms monoamine oxidase A and B was established after separate gene sequences encode two enzymes. These isoforms are differentially expressed with MAO-A expressed in human placental mitochondria whereas MAO-B in human platelet mitochondria [5]. MAO plays an absolute role in the deamination of neurotransmitters thus controlling their quantities in the nervous system. Which consequently linked the MAO to the basic pathological signs of neurological disorders like PD. With the know-how of the crystal structure of MAO enzyme, it becomes a major drug target and enables ligand designing [6].

In this research, three MAO-B inhibitors were selected to study their interaction with the PARK7 protein. These MAO-B inhibitors includes Rasagiline [7], Safinamide [8], and Selegiline [9] as they are being used for the treatment of PD these days. Rasagiline is an irreversible MAO-B inhibitor which is free from amphetamine metabolites unlike selegiline. Rasagiline, is a neuroprotective agent with motor restoration and antiparkinsonian traits. It has this one critical metabolite named 1-aminoindan which provides neuroprotection against damaging. While Safinamide is a reversible MAO-B inhibitor and active derivative of alpha-aminoamide. It is used in an add-on therapy as an agonist of dopamine for the treatment of PD symptoms and in final stage of PD in combination with Levodopa [10]. Lastly, Selegiline reinforces the dopaminergic functioning of brain, therefore it is applied as a monotherapy to treat the early PD symptoms combined with levodopa. It is an irreversible MAO-B inhibitor, which offers protection against neuronal damaging in the patients of PD. Further clinical trials suggested that selegiline not only reduced the amount of levodopa but also alter the progression of PD [11].

Despite of all the genetic advancement not a single applicable diagnostic biomarker for PD has been discovered. Here we investigate the role of the PARK7 protein as a genetic biomarker of PD. In this study, the PARK7 protein expression levels in the cerebrospinal fluid (CSF) were analyzed for both genetic and sporadic PD. Furthermore, a detailed spectrum analysis of the PARK7 protein was carried out to detect the percentage of sequence coverage across the MS sample. In this research, various bioinformatic tools are employed to predict the protein-protein interactions (PPI), then protein-ligand interactions, along with their phosphorylated sites. To further detect a potent interacting inhibitor of MAO-B for the PARK7 protein, an approach of molecular docking was adopted. This discovery would be beneficial for treating PD caused by mutant PARK7 protein. These investigations will aid in the discovery of alpha synuclein and MAO-B related inhibitors which reduced the effects of PD triggered by mutant PARK7 protein. Outcome of this study, offers early characterization and targeting of proteins enabling effective diagnosis of PD.

## 2 METHODS

### 2.1 Sampling

The informed consent of all participated members was sought prior to CSF collection considering the international ethical guidelines for biomedical study. For specimen collection, technique of spinal tap was acquired to obtain CSF samples. The collected CSF samples were stored at -80°C after centrifugation at 2000Xg at 4 °C for 10 minutes. Then samples were distilled using SpeedVac and resuspended in lysis buffer of 40µl. Further, samples were loaded, and gel staining was achieved through Coomassie staining to extract protein bands. Moreover, In-Gel tryptic digestion was executed, in which protein bands went through mincing, reduction, alkylation succeeded by overnight trypsin digestion [12].

### 2.2 SWATH Quantitation

To create a spectral library, equal amounts of proteins from each sample were pooled to the final concentration of 80 µg. Then Pierce High pH Reversed-Phase Peptide Fractionation Kit, (Thermo Fisher Scientific) was utilized for separation into eight fractions. Further, a loading buffer was prepared to dissolve digested peptides at final concentration of 0.3 µg/µl. Total 1.5 µg of sample was enriched for each analysis on self-packed precolumn followed by its separation through analytical RP-C18 column [13]. Furthermore, the sample analysis was achieved through nanoflow chromatography system (Eksigent nanoLC425) hyphenated to a hybrid triple quadrupole-TOF mass spectrometer (TripleTOF 5600+) supplied with a Nanospray III ion source (Ionspray Voltage 2400 V, Interface Heater Temperature 150 centigrade, Sheath Gas Setting 12) which was controlled by Analyst TF 1.7.1 software build 1163 (all AB Sciex) [14].

### 2.3 Data dependent Analysis

The LC-MS quantitation combines the LC with MS/MS enabling analysis of complex mixtures. It allows separation of our analyte from mixture, providing confidence that measured component is correct. A Top30 data-dependent acquisition method along with an MS survey scan of *m/z* 380-1250 was performed at 35000 full widths at half maximum resolution for round about 250ms to study qualitative LC-MS/MS data.

Meanwhile, the quantitative SWATH analysis MS/MS data was accomplished in 100 variable size windows across the 400-1200  $m/z$  range. Then Peak View Software version 2.1 build 11041(AB Sciex) together with the SWATH quantitation microapp version 2.0 build 2003 were employed to create spectral library and to obtain the SWATH peaks [15].

#### 2.4 Statistical Analysis

The study performed column analyses on the PARK7 gene and its phosphor form, which were analysed through unpaired experiment design using the GraphPad Prism 9 software. Then parametric test was selected by assuming gaussian distribution. In these columns same SD were assumed as unpaired t-test was applied. Further, Two-tailed analysis was executed selecting the confidence level of 95% [16]. The statistical significance was achieved by applying parametric unpaired two tail t-test during the statistical analysis.

#### 2.5 Phosphorylated sites prediction analysis

NetPhos program (<http://www.cbs.dtu.dk/services/NetPhos/>) predicts the phosphorylated sites in the eukaryotic proteins with the low stringency by employing a combination of neuronal networks. In addition, both the kinase specific predictions together with the generic were executed for the PARK7 protein [17]. Further, the PhosphoSitePlus (<http://www.phosphosite.org/>) webtool which investigates experimentally discovered post-translational modifications of humans and mouse proteins was employed to study the PARK7 protein [18].

#### 2.6 Protein-ligand interaction analysis

LIGPLOT analysis was employed to study multiple in-depth interactions like hydrogen bonding, non-covalent hydrophobic bonding, main-chain, and side chain. LIGPLOT is a bioinformatic tool which forms 2D image of protein-ligand interactions through their 3D coordinates [19]. A mutant human PARK7 gene (5SYA receptor) along with the 4R0U9 (Tgvtava, amyloid forming portion of alpha synuclein) as a ligand was employed for this study.

#### 2.7 Protein-protein network analysis and data visualization

The STRING (<https://string-db.org>) gathers all the interacting genes including the known as well as the unknown protein-protein interactions (PPIs) [20]. To locate a functional protein association for the PARK7 gene, the STRING was employed. After the recovery of the PARK7 gene data from the STRING, that data was then mapped on network nodes through the Cytoscape. The Cytoscape software version 3.6.0 was selected to visualize the Protein-protein interactions (PPI) network retrieved from STRING database [21].

#### 2.8 Molecular Docking

Molecular docking is a structure-based drug designing process in which molecular compounds are docked into the receptor's binding site. It predicts the binding affinity along with the activity of protein-ligand complex [22]. Monoamine oxidase inhibition offers potential therapeutic target for the management of depression and Parkinson's disease. Thus, this present study was designed to analyse protein-ligand interaction by using monoamine oxidase b inhibitors (MAO-B) as ligands against PARK7 protein target. The commonly known three MAO-B inhibitors: Rasagiline, Safinamide and Selegiline were selected for this study.

##### 2.8.1 Ligand retrieval and its preparation

All the ligands namely Rasagiline, Safinamide and Selegiline used in our study were gathered from the existing literature. The three-dimensional molecular structures of all three were imported from PubChem in SDF format. This format was used to open these files in MOE and then perform the energy minimization using the default parameters. Then, the MMFF94x force field was used with no periodicity and the limitations were retained at the rigid water molecule level.

##### 2.8.2 Receptor's retrieval

To study the binding affinity of protein-ligand complexes, technique of molecular docking was adopted. In which the PARK7 protein was used as a target protein for the respective ligand inhibitors of MAO-B. The 5SYA retrieved from protein data bank (PDB) in PDB format, is a part of an atomic structure of a mutant DJ-1 D24N which is used as a PARK7 protein receptor (represented in **Figure 1. A**).

##### 2.8.3 Preparation of protein and Molecular Docking

Protein was prepared for molecular docking by energy minimization and prediction of active site for all ligands, while maintaining the default parameters. After which the ligands (Rasagiline, Safinamide and Selegiline) were docked with the target protein (PARK7) one by one using MOE software. For molecular docking, each ligand atom (Rasagiline, Safinamide and Selegiline) were selected. Then rescoring1 was set at London dG followed by rescoring2 at GBVI/WSA dG, which was run to determine the ligand interaction with the PARK7 protein. And as a result, the protein-ligand docking score along with its properties, 2D and 3D structures have been stored.

##### 2.8.4 Active site prediction

The PARK7 protein has three chains containing protein sequence of 192 amino acids residues, a ligand with two amino acids and a water molecule chain. The active site of the target protein PARK7 is the ligand molecule namely Ethylene glycol or 1,2-Ethanediol ( $C_2H_6O_2$ ) (mentioned in **Figure 1 B**) & **C**) respectively).

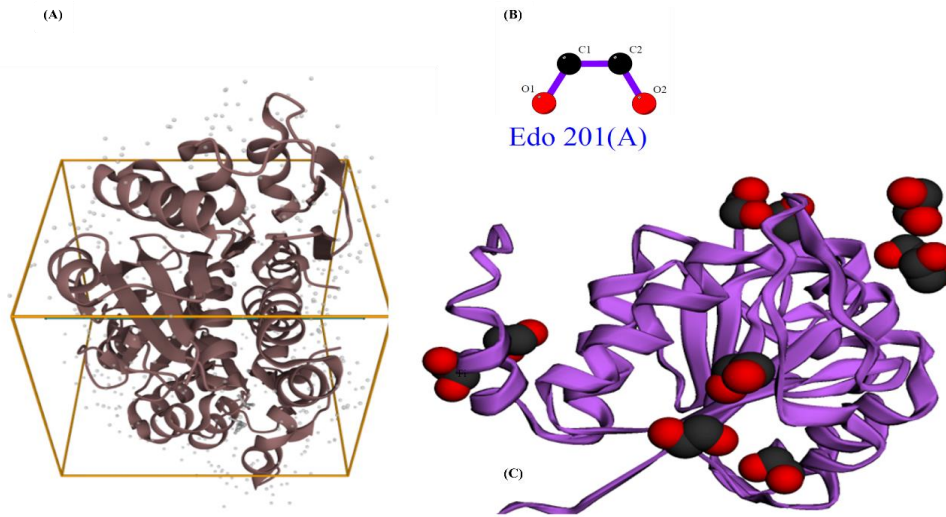


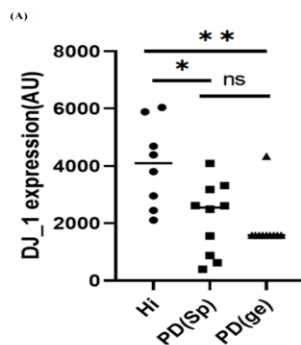
Fig. 1. (a) Atomic structure of mutant DJ-1 D24N which is used as a PARK7 protein receptor. (b) Ligand Edo 201(A). (c) 3D mol.js view of

### 3 RESULTS

#### 3.1 Expression analysis of PARK7 gene

The PARK7 gene expression analysis displays the P-value of 0.0013 in the healthy versus genetic PD, which was less than the significant difference of  $p < 0.01$ . Similarly, the P-value of 0.0106 was observed in the healthy versus sporadic PD, which was less than the significant difference of  $p < 0.05$ . Further investigating these findings in detail for the PARK7 gene, by analysing the column A Healthy (Hi) of sample size 8 and column B Sporadic PD (PD(Sp)) of sample size 10 through unpaired, two tail T-TEST, the P value obtained was 0.0106 which was less than our significant difference ( $*P < 0.05$ ). The difference between means (B - A)  $\pm$  SEM was  $-1866 \pm 644.8$  providing the 95% confidence interval of  $-3233$  to  $-498.8$  and the R squared (eta squared is the data variation explained by group means) value of 0.3435. Additionally, the F test was performed to compare the significant variance, but it does not display any significant difference. Similarly, when the analysis of column A Healthy (Hi) of sample size 8 and the column C genetic PD (PD (ge)) of the sample size 10 through unpaired, two tail T-TEST was examined, the P value obtained was 0.0013 which was less than our significant difference ( $*P < 0.05$ ,  $**p < 0.01$ ). The difference between means (C - A)  $\pm$  SEM was  $-2180 \pm 559.5$  providing the 95% confidence interval of  $-3366$  to  $-993.5$  and the R squared value of 0.4868. Then, again the F test was performed but no significant results were observed. Lastly, the PARK7 gene analysis of the column B PD (Sp) of the sample size 10 and the column C (PD (ge)) of the sample size 10 through unpaired, two tail T-TEST was determined, the P value obtained was 0.5231 which was greater than our significant difference. Hence, PD(Sp) vs PD (ge) was shown by ns on the graph means not significant as shown in **Figure 2. A**). In addition, the difference between means (C - B)  $\pm$  SEM is  $-313.9 \pm 482.05$  providing the 95% confidence interval of  $-1327$  to  $698.8$  and the R squared (eta squared) value of 0.02302. Like the previous results this F test when to compare significant variance, does not display any significant difference.

Fig. (2)



(B)

PARK7\_HUMAN (100%), 19,891.2 Da  
Protein deglycase DJ-1 OS=Homo sapiens GN=PARK7 PE=1 SV=2  
3 exclusive unique peptides, 3 exclusive unique spectra, 3 total spectra, 25/189 amino acids (13% coverage)

```

MASKRALVIL AKGAEEMETV IPVDVMRRAG IKVTVAGLAG KDPVQC SRDV
VICPDASLED AKKEGPDYDV VLPGGNLGAQ NLSESAAVKE ILKEQENRKG
LIAAICAGPT ALLAAHEIGFG SKVTTHPLAK DKMMNGGHY YSENRVEKDG
LILTSRGPPT SFEFALAIVE ALNGKEVAAQ VKAPLV LKD

```

Fig. 2 (a) The PARK7 gene expression analysis results were examined through Graph Pad Prism9 software. As illustrated in graph above, Hi vs PD (ge) and Hi vs PD(Sp) displays significant difference (P value less than 0.05 or 0.01) which is denoted by \*p & \*\*p. While PD(Sp) vs PD (ge) displays no meaningful results denoted by ns. Statistical significance: \* $p < 0.05$  & \*\* $p < 0.01$ . (b) These MS/MS spectra are produced through mass spectrometry by matching peptides with the given database which accurately paired with fragmentation pattern.

An independent LC-MS data acquisition analysis detected the T25 tryptic fragment of the PARK7 gene. Outcomes of the study introduced three unique peptides along with their three spectra. The scaffold analysis displayed thirteen percent sequence coverage for the PARK7\_HUMAN gene across the provided MS sample. In addition, three biological replicates with their sequences: (K)VTVAGLAGK(D); (K)VTTHPLAK(D) & (K)DGLILTSR(G) were obtained as mentioned in **Figure.2. B**). According to Scaffold about 98% of probability of sequence is required to be accurate. The above three sequences have shown: (K)VTVAGLAGK(D) 100%; (K)VTTHPLAK(D) 99% & (K)DGLILTSR(G) 99% probabilities, respectively. The mascot ion score of these three were: (K)VTVAGLAGK(D) 38.5; (K)VTTHPLAK(D) 30.5 & (K)DGLILTSR(G) 30.8. The resulting protein sequence includes 25 out of 189 amino acids residues. The first spectrum of (K)VTVAGLAGK(D) peptide displays the peak values of 201.1 & 428.3 for b2 and b5 ions together with B-H<sub>2</sub>O. Similarly, the y-ion values determined by the spectrum from y2 to y8 were represented in **Figure 3. A**). The second spectrum of (K)VTTHPLAK(D) peptide displays peak values of 201.1 for b2. Like wisely, the y-ion values examined by the spectrum from y2 to y5, y7.y8 along with Y-NH<sub>3</sub> and Y+2H values were shown in **Figure 3. B**). While the third spectrum of (K)DGLILTSR(G) peptides displays peak values of 173.1, 286.1 and 399.2 for b2, b3 & b4. Whereas the y-ion values determined by the spectrum from y3 to y8 along with Y-NH<sub>3</sub> and Y-H<sub>2</sub>O values were mentioned in **Figure 3. C**).

# IEEESEM

Fig. (3)

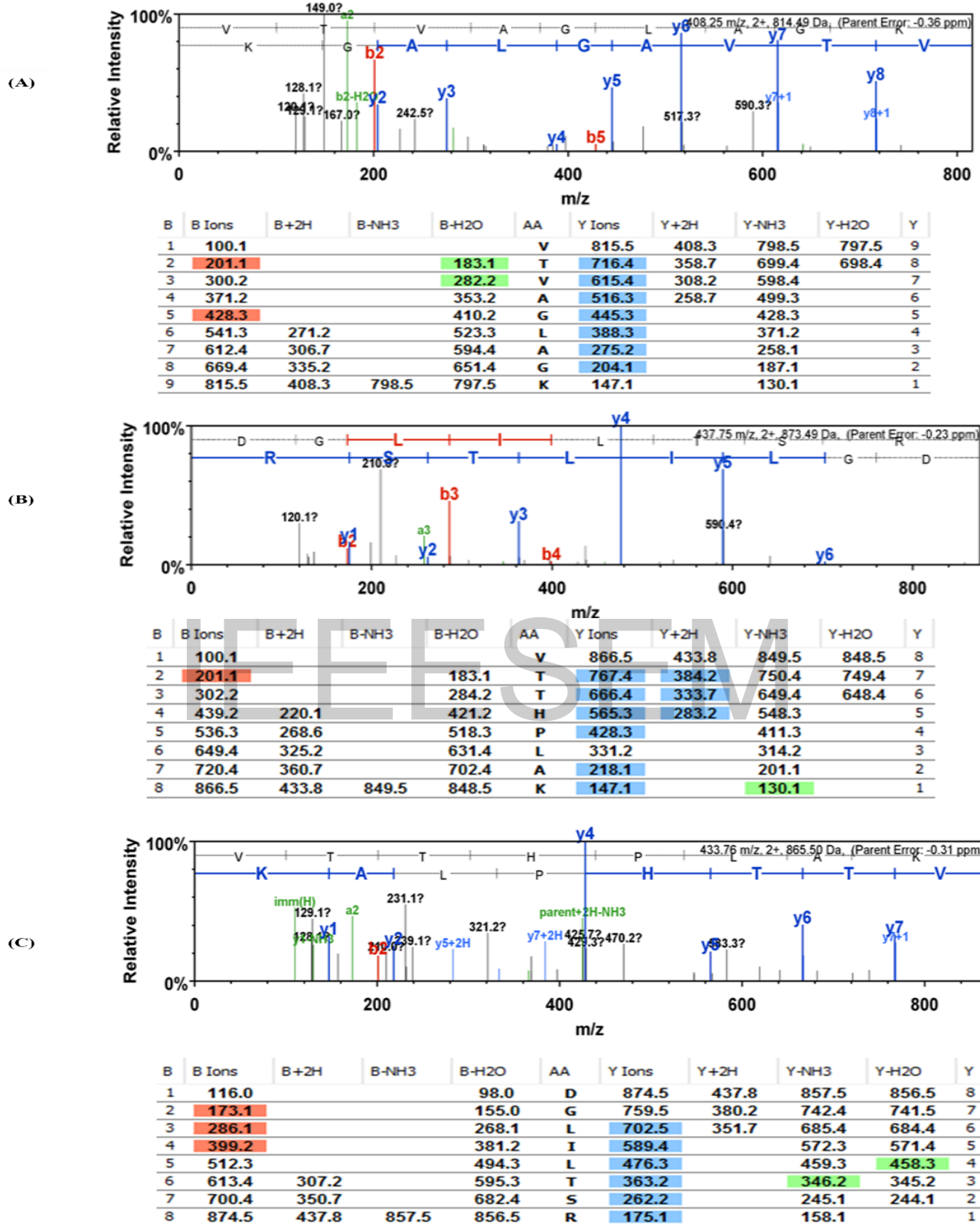


Fig. 3 (a) This is the graphical illustration of first MS/MS spectra with MS/MS peaks. In this spectrum, red-b ions represent fragment ions whereas blue ion indicates y ion. While green represents water, grey ones are not discovered. In this spectrum observed mass is 408.25 m/z than the actual mass which was 814.49 Da. And the fragment table displays the peaks value for b-ions which is 201.1 & 428.3 for b2 and b5 together with B-H2O. Like wisely, the y-ion values determined by the spectrum from y2 to y8 are also provided. (b) This is the graphical illustration of first MS/MS spectra with MS/MS peaks. In this spectrum, red-b ions represent fragment ions whereas blue ion indicates y ion. While green represents ammonium ion or neutral loss of NH3 or water, grey ones are not discovered. In this figure observed mass is 433.76 m/z than the actual mass which was 865.50 Da. While the fragment table displays the peaks value for b-ions which is 201.1 for b2. Like wisely, the y-ion values determined by the spectrum from y2 to y5, y7, y8 along with Y-NH3 and Y+2H values are also provided. (c) This is the graphical illustration of first MS/MS spectra with MS/MS peaks. In this spectrum, red-b ions represent fragment ions whereas blue ion indicates y ion. While green represents ammonium ions, grey ones are not discovered. In this figure observed mass is 437.75 m/z than the actual mass which was 873.49 Da. And the fragment table displays the peaks value for b-ions which is 173.1, 286.1 and 399.2 for b2, b3 & b4. Like wisely, the y-ion values determined by the spectrum from y3 to y8 along with Y-NH3 and Y-H2O values are also provided.



### 3.3 Phosphorylated sites prediction

The NetPhos output for the PARK7 gene was illustrated with graph of predicted post-translational glycosylation and phosphorylated proteins from 189 amino acid sequence. The graph also represents S (serine), T (threonine), and Y (tyrosine) residues in the PARK7 sequence with red, green, and blue lines (as shown in **Figure 4. A**). All these residues with the value less than 0.5 were considered as a phosphorylated site. Whereas if the residue crosses the threshold of 0.5 (any value above) it was predicted to be more secured. All the kinase specific predictions together with unspecified generic predictions are enlisted in the **Table 1**. All the tyrosine, threonine, and serine predictions with their context, score along with their kinase specifications (if present, otherwise unspecified generic predictions) are mentioned in the **Table 1**.

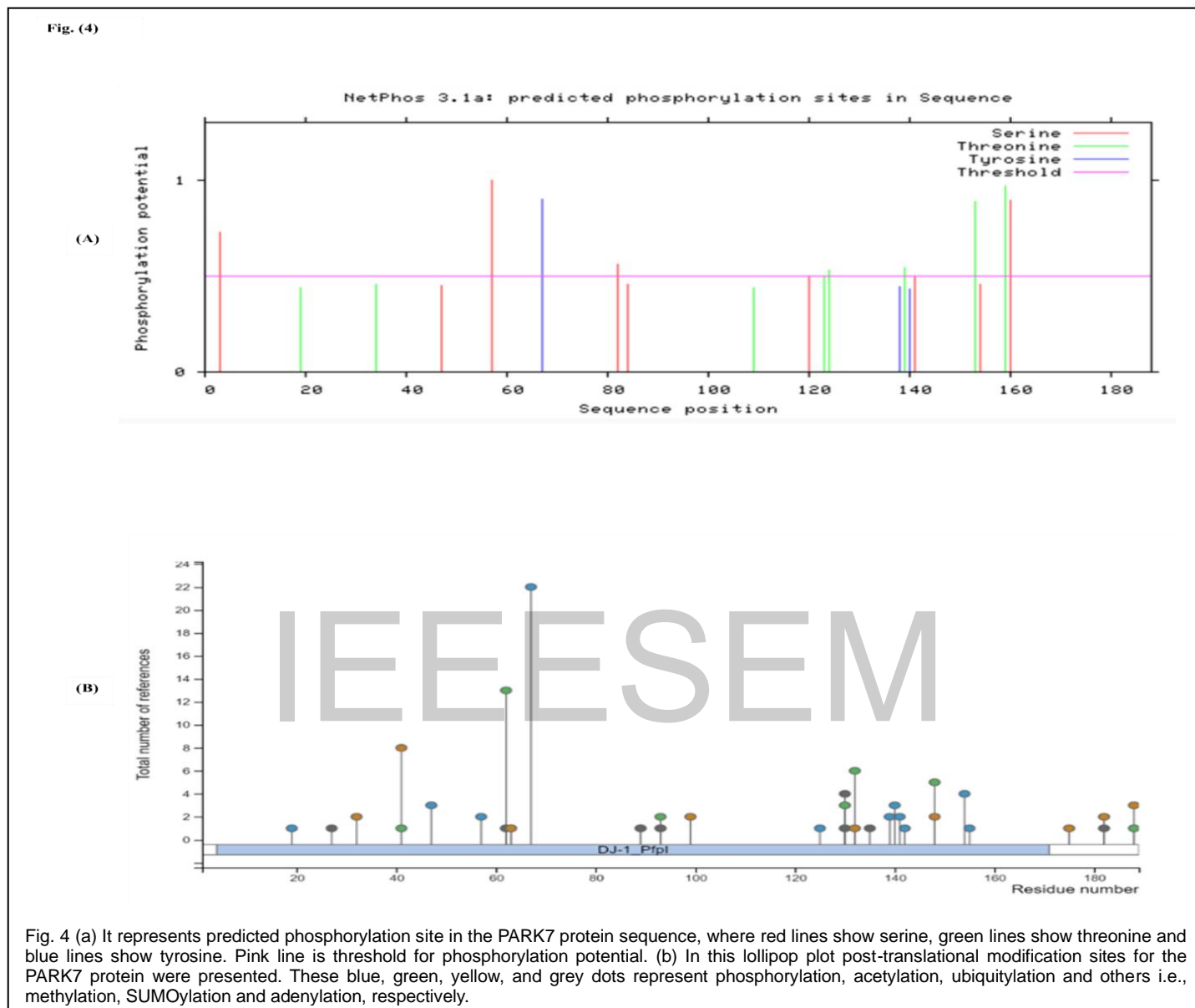
TABLE 1

NETPHOS RESULTS FOR ONLY THOSE SEQUENCES WHICH HAVE THE VALUES ABOVE THE THRESHOLD (0.500). THESE ENLISTED VALUES ARE PREDICTED TO BE MORE SECURED AND REPRESENTED WITH A YES SYMBOL.

Sequences	# x	Context	Score	Kinase	Answer
1.Sequence	3 S	--MASKRAL	0.728	PKC	YES
2. Sequence	57 S	CPDASLEDA	0.997	unsp	YES
3.Sequence	57 S	CPDASLEDA	0.558	DNAPK	YES
4. Sequence	67 Y	KEGPYDVVV	0.899	unsp	YES
5.Sequence	82 S	AQNLSSESAA	0.562	cdc2	YES
6.Sequence	82 S	AQNLSSESAA	0.520	unsp	YES
7.Sequence	82 S	AQNLSSESAA	0.502	CKII	YES
8.Sequence	124 T	SKVTTHPLA	0.528	PKC	YES
9. Sequence	139 T	GGHYTYSEN	0.544	CKII	YES
10. Sequence	153 T	GLILTSRGP	0.887	unsp	YES
11. Sequence	153 T	GLILTSRGP	0.807	PKC	YES
12.Sequence	159 T	RGPGTSFEF	0.969	unsp	YES
13. Sequence	159 T	RGPGTSFEF	0.549	CKII	YES

Furthermore, the PhosphoSitePlus was employed to examine the phosphorylation, acetylation, ubiquitylation and others i.e., methylation, SUMOylation, adenylation etc. in more detail than NetPhos. Moreover, the lollipop plot illustrates the linear amino acids sequence of the protein-protein domain DJ-1 Pfpl X-axis. While the Y-axis represents the HTP (high throughput papers includes modification sites from proteomic mass spectrometry) and the LTP (low throughput papers includes modification sites from other than mass spectrometry methods) ref-

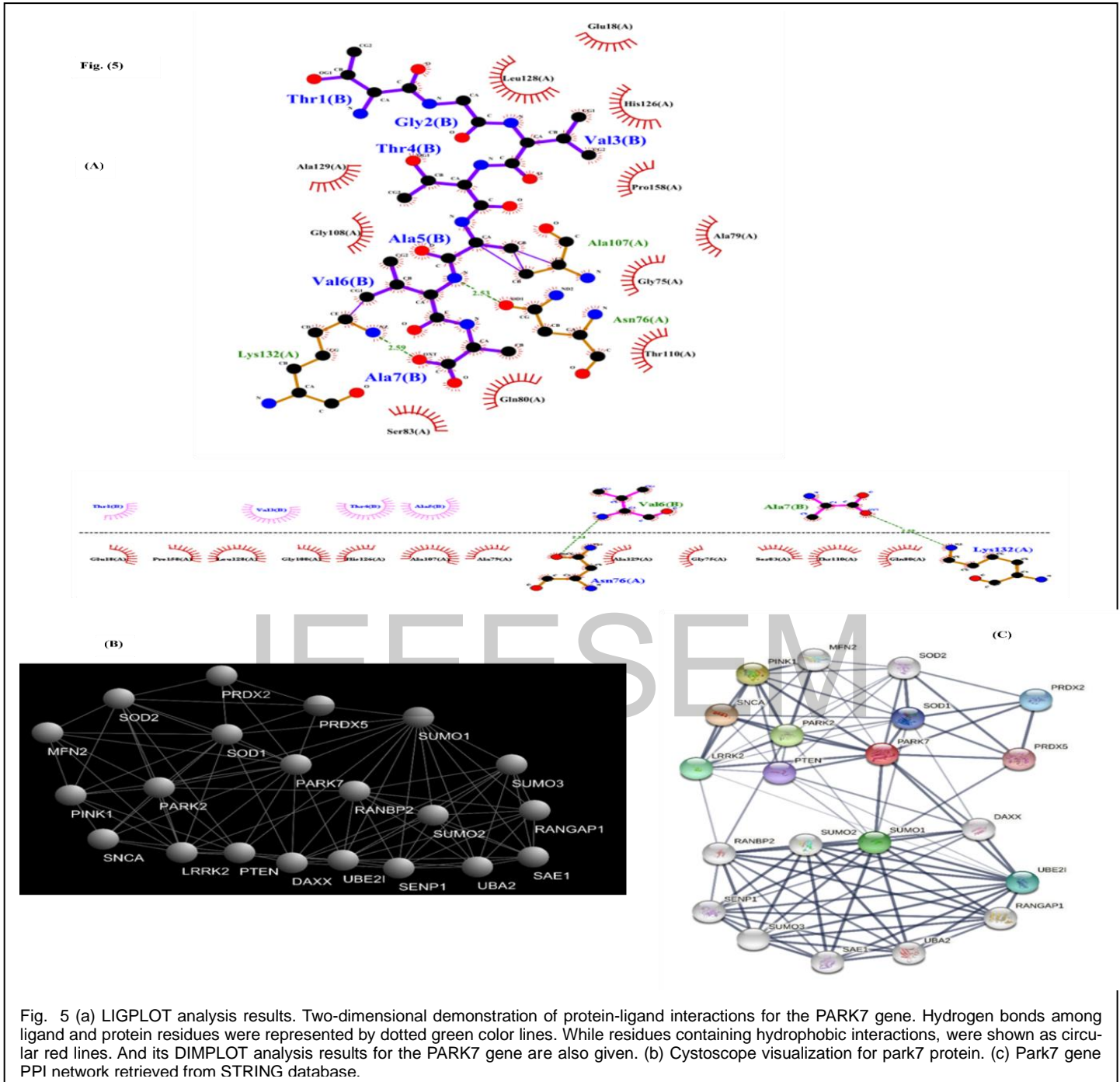
ferences as it is represented in **Figure 4.B**).



### 3.4 Protein-ligand interactions

The alpha synuclein ligand bond between the ligand residues were shown in purple, while the PARK7 protein amino acid residues were represented in brown. The total two hydrogen bonds were formed between the PARK7 protein and the alpha synuclein ligand. The first hydrogen bond was formed between the Asn76(A) [Asparagine 76 amino acid residue of chain A of PARK7 protein] and the Val6(B) [Valine 6 amino acid residue of chain B of  $\alpha$ -synuclein ligand] with the bond length of 2.53. While the second hydrogen bond of length 2.59 was formed between the Lys132(A) [Lysine 132 amino acid residue of chain A of PARK7 protein] and the Ala7(B) [Alanine 7 amino acid residue of chain B of  $\alpha$ -synuclein ligand]. While the Glu18(A), Leu128(A), His126(A), Pro158(A), Ala79(A), Gly75(A), Ala129(A), Gly108(A), Thr110(A), Gln80(A), and Ser83(A) were the eleven non-ligand amino acid residues involved in hydrophobic interactions and represented by red circular lines. The covalent/elastic bonds formed between the Lys132(A) and the Val6(B) together with the Ala107(A) and the Ala5(B) were represented as thin purple lines. The DIMPLOT was applied on the LIGPLOT results to study chain A and B interactions across the protein-protein interfaces. Where the horizontal dashed black line represents the interface as shown in **Figure 5. A**).





### 3.5 PPI network analysis and data visualization

The protein-protein network analysis predicted, eleven functional partners of the PARK7 gene which were alpha synuclein (SNCA), PINK1, PARK2, SUMO1, LRRK2, UBE21, PTEN, SOD1, PRDX2, and PRDX5, respectively. This network interaction analysis exhibited 11 nodes, 34 edges (greater than 15, commonly expected nodes), average node degree value of 6.18, average local clustering coefficient 0.775, and PPI enrichment p-value of  $1.01 \times 10^{-5}$ . These results were obtained by applying interaction score of 0.400 confidence to create protein-protein associations. This full network type displays edges indicating both functional and physical protein associations. These thicker lines between network edges indicates strength of the data as shown in **Figure 5. C**). The PPI enrichment analysis results stated in **Table 2**, encodes the information of nodes of the PARK7 gene network, which were retrieved from STRING with their p-values. This table states a few PPI enrichment results displayed through gene ontology, KEGG pathway, references from PubMed and Uniport.

TABLE 2. PPI ENRICHMENT ANALYSIS RESULTS INDICATES ENCODED INFORMATION OF NODES OF PARK7 GENE NETWORK RETRIEVED FROM STRING WITH THEIR P-VALUES.

PARK7 gene network data retrieved from STRING	P- value
1.Gene ontology (Biological process)	-----
cellular response to dopamine	0.00021
regulation of dopamine metabolic process	3.94e-06
positive regulation of autophagy of mitochondrion	3.35e-06
regulation of oxidative stress-induced cell death	3.60e-08
positive regulation of oxidative stress-induced intrinsic apoptotic signalling pathway	6.29e-05
dopamine metabolic process	0.00057
2.Reference publication (PubMed)	-----
Oxidative stress factors in Parkinson's disease.	3.85e-12
Back and to the Future: From Neurotoxin-Induced to Human Parkinson	3.85e-12
3.KEGG Pathways	-----
Parkinson's disease	3.68e-07
4.Annotated keywords (Uniport)	-----
Parkinsonism	7.23e-10

Then these STRING database outcomes were mapped on nodes and visualized through the Cytoscape as displayed in **Figure 5. B)**. This Cytoscape visualization exhibited clustering coefficient of 0.773, 21 nodes, and it took analysis time of 0.151. Further, the shortest paths observed was 420(100%), with the 8.857 of average number of neighbors and characteristic path length of 1.705. Meanwhile, the network analysis deduced network density of 0.443, and network heterogeneity of 0.271. Together with the network diameter of 3, network radius of 2, and network centralization of 0.284. In the **Table 3**, top fifteen differential genes including the PARK7 gene and with their characteristics (like betweenness centrality, closeness centrality, degree, and number of directed edges) were mentioned.

TABLE 3. THIS TABLE CONTAINS TOP FIFTEEN DIFFERENTIAL GENES INCLUDING PARK7 GENE AND THEIR CHARACTERISTICS RECOVERED FROM CYTOSCOPE.

Genes	Betweenness Centrality	Closeness Centrality	Degree	Number of Directed Edges
DAXX	0.040705	0.625	8	8
LRRK2	0.056283	0.666667	10	10
MFN2	0	0.5	6	6
PARK2	0.052335	0.666667	10	10
PARK7	0.117409	0.740741	13	13
PINK1	0.00364	0.526316	8	8
PRDX2	0	0.465116	4	4
PRDX5	0.012706	0.571429	5	5
PTEN	0.041363	0.645161	9	9
RANBP2	0.059397	0.645161	10	10
RANGAP1	0	0.540541	8	8
SAE1	0	0.540541	8	8
SENP1	0.004211	0.555556	9	9
SNCA	6.58E-04	0.5	6	6

### 3.6 Molecular Docking results

The docking results suggested that Safinamide with -5.4492 best energy score was more potent interacting inhibitor as compared to Selegiline and Rasagiline. Whereas Rasagiline, has shown favourable results with the -4.7925-energy score. Whereas Selegiline displayed the energy score of -5.1325 but it is a least interacting inhibitor as no hydrogen bond interaction was observed during protein-ligand interactions in their surface map. The docking complex images for Rasagiline, Safinamide and Selegiline demonstrating their interaction and surface maps are presented. All the details like the protein receptor interacting with the three ligands, type of the binding inside complexes, the length of hydrogen bonding, and the energies for docking complexes were stated.

#### (1) SAFINAMIDE

The most potent ligand, Safinamide, formed two hydrogen interactions with the asparagine (Asn) 76 and the sulfinoalanine (CSD) 106 active amino acid residues and one arene-hydrogen interaction. The Asn 76 was observed to make a polar hydrogen bond of bond length 2.91Å, with the oxygen atom double bonding with piperidine moiety of the ligand. And the CSD 106 formed greasy hydrogen interaction with the bond length of 3.03Å amino group of the ligand. Further, Proline (Pro)158 formed greasy arene-hydrogen interaction with the bond length of 3.86Å with benzene ring of the inhibitor. The Safinamide with the -5.4492 best energy score was more potent interacting inhibitor as compared to Selegiline and Rasagiline. All these interactions were represented in (in **Figure 6. A**) the surface map along with its docking complex image.

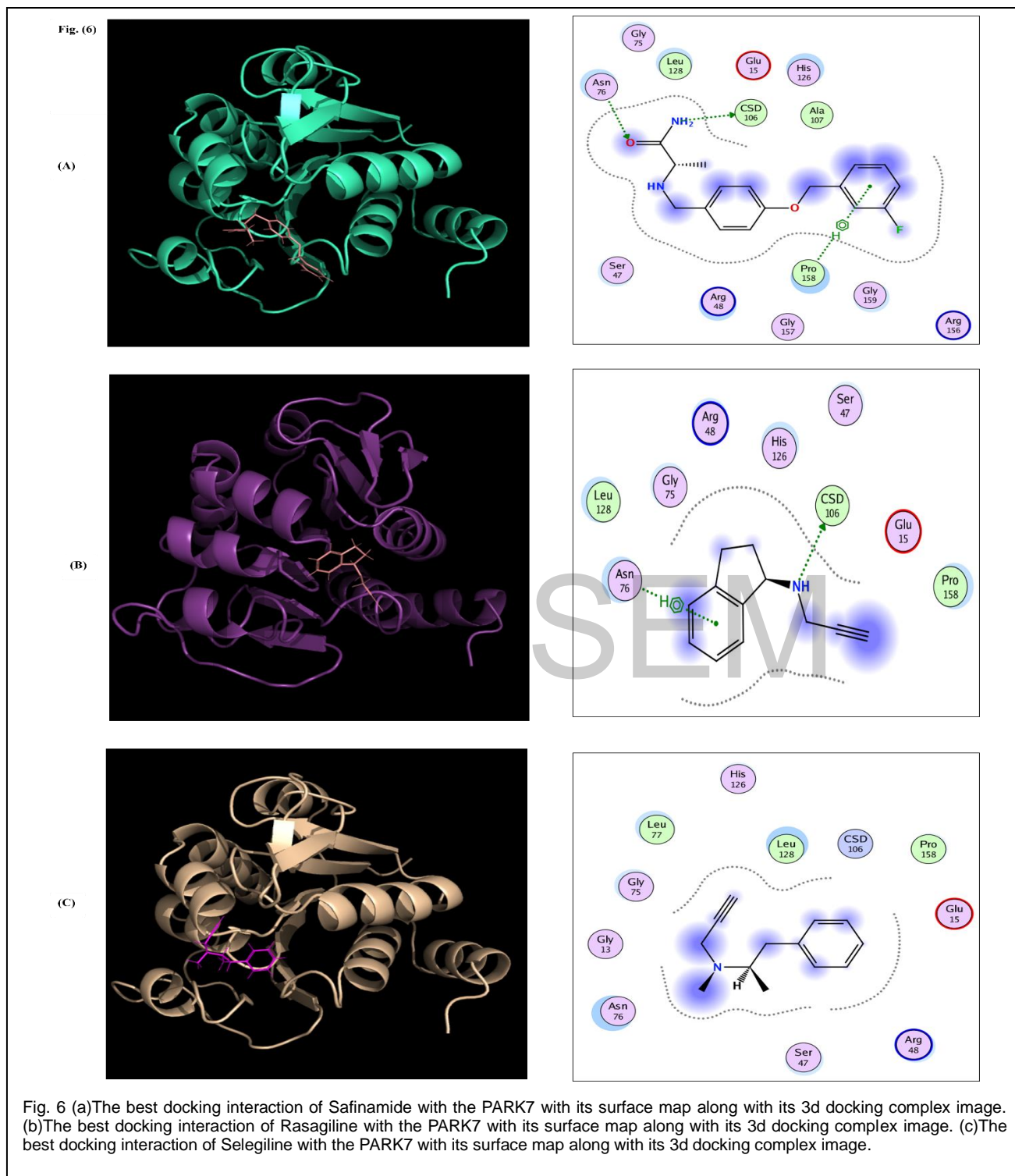
#### (2) RASAGILINE

Rasagiline, formed one hydrogen bond with the sulfinoalanine (CSD) 106 active amino acid residue and one arene-hydrogen interaction. The CSD 106 was observed to form a greasy hydrogen interaction, with the nitrogen atom of the ligand, with the bond length of 2.95Å. While the Asn 76 exhibited an arene-hydrogen bond with the bond length of 3.95Å with the benzene ring of the inhibitor. Rasagiline, has shown favourable results with the -4.7925-energy score. All these interactions were shown in (in **Figure 6. B**) the surface map along with its docking complex image.

#### (3) SELEGILINE

In Selegiline, no noticeable hydrogen bonding was observed between the polar Leu 128, Leu77, Gly 75, Asn 76, Ser 47, basic Arg 48 amino acid residues and the ligand. Selegiline has shown the energy score of -5.1325 but it is a least interacting inhibitor as no hydrogen bond interaction was observed during protein-ligand interaction in their surface map. All these observations were illustrated in (in **Figure 6. C**) the surface map along with its docking complex image.

IEEESEM



#### 4 DISCUSSION

Despite the quiet advancement in the field of neurological science, Parkinson's disease, its pathology is not clearly understood. Several studies have shown more than 25 PAK7 gene mutations associated with early onset of PD which occurs at an early age before 50. This research was conducted to define the role of these genetic biomarkers especially the PAK7 gene in PD. In this research we analyze the ex-

pression levels of PARK7 gene in the cerebrospinal fluid of the genetic and sporadic PD patients along with control (healthy) ones. The proteomic findings of this research indicate the aberrant expression levels of the PARK7 gene in the CSF of PD patients. These altered expression levels are the pathological hallmarks of PD [23].

We observed altered PARK7 expression levels in the cerebrospinal fluid of the selected patients. The PARK7 gene expression analysis results show the p-value of 0.0013 in healthy versus genetic PD, which was less than the significant difference of  $p < 0.01$  represented by  $**p$ . Similarly, the P-value of 0.0106 was observed in healthy versus sporadic PD, which was less than the significant difference of  $p < 0.05$  denoted by  $*p$ . Meanwhile, the column analysis performed on genetic and sporadic PD showed no significant difference as the p-value obtained was 0.5231 greater than  $p < 0.05$  and  $p < 0.01$ . The expression analysis of the PARK7 gene demonstrated that the genetic PD was statistically more significant as compared to sporadic PD. The results of this study describe the genetic mutation associated with PD, specifically in the PARK7 gene.

When independent LC-MS data acquisition analysis was performed to detect the sequence coverage in T25 tryptic fragment of the PARK7 gene. Outcomes of this analysis introduced three unique peptides along with their three spectra. It showed thirteen percent of sequence coverage for the PARK7\_HUMAN gene across the provided MS sample. The detailed spectrum analysis of the PARK7 protein showed that these three peptide sequences have (K)VTVAGLAGK(D) 100%; (K)VTTHPLAK(D) 99% & (K)DGLILTSR(G) 99% probabilities, respectively. In scaffolding about 98% probability of sequence is required for it to be accurate. Hence, these results demonstrate the detection and accuracy of PARK7 gene sequence in MS sample.

To study the PARK7 protein-ligand interactions LIGPLOT analysis was introduced. A mutant human PARK7 gene (5SYA receptor) along with the 4R0U9 (Tgvtava, amyloid forming portion of alpha synuclein) as a ligand was employed for this study. Previous studies had shown that an alpha-synuclein was a major contributing factor of PD, whose abnormal accumulation in the form of Lewy bodies lead towards the dysfunctional neuronal activity in the brain. The SNCA gene encodes an  $\alpha$ -synuclein which promotes the non-motor symptoms of PD [24].

So, we select an amyloid forming segment of an  $\alpha$ -synuclein as a ligand to investigate the protein-ligand interactions of the PARK7 protein. The two major hydrogen bonds were formed between the PARK7 protein and an  $\alpha$ -synuclein ligand residues of the length 2.53 and 2.59. While several hydrophobic interactions were also observed among the PARK7 protein's amino acid residues and an alpha-synuclein residues. This study enlisted Glu18(A), Leu128(A), His126(A), Pro158(A), Ala79(A), Gly75(A), Ala129(A), Gly108(A), Thr110(A), Gln80(A), and Ser83(A) as the non-ligand amino acid residues of the PARK7 protein. We observe the association between the PARK7 gene and an  $\alpha$ -synuclein protein so we could understand their pathological role in PD.

In addition to this different bioinformatic tools were employed to examine protein-protein interactions and post translation modifications of PARK7 gene. In-depth protein-proteins interactions were studied through the STRING and visualized by Cytoscape. We retrieved PTEN, SNCA, PINK1, PARK2, SUMO1, LRRK2, UBE21, SOD1, PRDX2, and PRDX5 proteins which interacts with the PARK7 protein. When this PPI network of the PARK7 protein was visualized top fifteen differential genes including the PARK7 gene were recovered through Cytoscape. These findings highlight the characteristics of these fifteen differential genes named DAXX, LRRK2, MFN2, PARK2, PINK1, PRDX2, PRDXS, PTEN, RANBP2, RANGAP1, SAE1, SENP1, SNCA, and SOD1 including the PARK7. Further, the number of directed edges, closeness and betweenness centrality of these genes was observed. Earlier studies illustrate the indirect role of PARK7 with PINK1 and parkin, along with its associated mitochondrial dysfunction which led towards PD [25]. The PARK7 gene works as an antioxidant protein during stress environment, which lead towards the transcription of genes that are antioxidant in nature these genes include SOD, NQO1 and Trx1. The SOD1 gene plays a major role in the glyoxalase activation during oxidative stress. So, this SOD1 mutation was associated with the PARK7 dysfunction and PD development [26]. In addition, the relation of the PINK1 and the SOD1 genes with PD and its interrelation with the PARK7 protein was observed through this PPI network analysis.

In this study, we used NetPhos tool to predict the phosphorylated sites of the PARK7 protein. The resulted graph predicts the post-translational glycosylation and phosphorylated sites for 189 amino acid sequence. It also represents S(serine), T(threonine), and Y(tyrosine) residues in the sequence. Then PhosphoSitePlus was employed to examine phosphorylation, acetylation, ubiquitylation and others i.e., methylation, SUMOylation, adenylation etc through high throughput and low throughput articles in more detail than NetPhos.

The molecular study examined the interaction of these MAO-B inhibitors (Rasagiline, Safinamide and Selegiline) with our target receptor protein (PARK7). The docking complex images for Rasagiline, Safinamide and Selegiline representing the interactions along with the surface maps were obtained. The docking results suggested that Safinamide with -5.4492 best energy score was more potent interacting inhibitor as compared to Selegiline and Rasagiline. Additionally, Rasagiline, has shown favorable results with -4.7925 energy score. Whereas Selegiline has shown the energy score of -5.1325 but it was a least interacting inhibitor as no hydrogen bond interaction was observed in their surface map.

Molecular docking results suggested variation from the best interaction to moderate binding interaction, which were based upon following observations (mentioned in **Table 4**):

- The energy scores (S) represent the range of best binding interactions to moderate ones which range from -5.4492 to -4.7925.
- The ligation mode was mostly covering N, O, and 6-ring sites.
- The main binding active amino acid residues of receptor were sulfinoalanine, asparagine and Proline, respectively.
- The hydrogen bond interactions include almost all types of bonds, such as H-donor, H-acceptor, and  $\pi$ -H except from  $\pi$ - $\pi$ .

The above discussion concludes that Safinamide was the best inhibitor as compared to Rasagiline and Selegiline. This study also exhibited interacting images of docking complexes explaining the overall docking results.

TABLE 4. THE ENERGY SCORE AND BINDING DATA FOR ALL THE THREE MAO-B INHIBITORS AGAINST SELECTED PROTEIN.

Compound	Protein	Ligand	Receptor		Interaction	Distance	E(Kcal/mol)	S (Energy score)
Safinamide	5SYA (PARK7)	N5	OD2 106	CSD (A)	H- donor	3.03	-5.3	-5.4492
		O3	ND2 76	ASN (A)	H-acceptor	2.91	-3.5	
		6-ring	CB 158	PRO (A)	Pi-H	3.86	-0.8	
Rasagiline	5SYA (PARK7)	N1	OD1 106	CSD (A)	H- donor	2.95	-1.9	-4.7925
		6-ring	N 76	ASN (A)	Pi-H	3.95	-2.1	
Selegiline	5SYA (PARK7)	-----	-----	-----	-----	-----	-----	-5.1325

All these findings collectively indicate that the PARK7 gene is a genetic biomarker of PD which is interrelated with alpha-synuclein, PINK1, and SOD1 as discussed earlier. Our research demonstrated the PARK7 gene as a potential biomarker of Parkinson's disease with the help of expression analysis and bioinformatic study. The altered PARK7 expression levels were reported in CSF samples of genetic and sporadic PD patients. Aside from expression analysis, bioinformatic tools were employed to study protein-ligand and protein-protein interactions. The predicted protein-ligand contacts through LIGPLOT study shows the relation of the PARK7 protein with alpha-synuclein which co-relates both with PD. The PINK1 and SOD1 genes were observed in protein-protein network analysis. The PARK7 gene was found to be a transcriptional regulator of SOD1 and co-relates with PINK1. Hence, the research outcomes illustrate that the PARK7 gene regulates some genes and involves with other to promote the early onset of Parkinson's disease. Molecular docking studies demonstrated the protein-ligand interaction and exhibited Safinamide as a potent inhibitor of the PARK7 protein. This outcome suggests Safinamide as a potent inhibitor, which can be employed to treat the PARK7 induced PD symptoms in future.

In summary, our research demonstrated the PARK7 levels in cerebrospinal fluid (CSF) of Parkinson's patient which were differently modulated. Then further research transcribes that some prominent biomarkers of Parkinson's (like alpha synuclein and other PD genes) were also associated with the PARK7 gene. The protein identification was achieved through implication of traditional mass spectrometry quantitation method known as SWATH quantitation analysis. To further examine the statistical significance in the PARK7 gene expression analysis was performed. In addition, the confidence of the identified PARK7 proteins fragments was enhanced through scaffolding and to increase the accuracy of obtained results. Furthermore, the PARK-7 gene and an alpha-synuclein ligand interactions were studied through LIGPLOT. Moreover, the PARK7 gene PPI network was retrieved from the STRING database and visualized in Cytoscape. Then, NetPhos tool predicted the specific and non-specific phosphorylated sites. Meanwhile, PhosphoSitePlus was used to study phosphorylation, acetylation, ubiquitylation and others i.e., methylation, SUMOylation, adenylation sites etc of the PARK7 gene. Lastly, the staggery of molecular docking was adopted which determined Safinamide as potent inhibitor of the PARK7 protein. These findings will be used to treat the PARK7 mutation related PD in future.



## 5 CONCLUSION

From the outcome of our study, we can speculate that the mutated PARK7 protein plays a significant role in the early onset of PD. The modified expression levels of the PARK7 protein which displayed the P-value of 0.0106 and 0.0013. Though no significant difference was observed between sporadic and genetic PD. These experimental investigations of the PARK7 displayed 13 percent sequence coverage for the PARK7\_HUMAN gene across the provided MS sample. Further, this research introduced three biological replicates along with their spectra. Our research reports Safinamide with -5.4492 best energy score to be more potent interacting inhibitor as compared to Selegiline and Rasagiline. These investigations will aid in the discovery of alpha synuclein and MAO-B related inhibitors which reduced the effects of PD triggered by the mutant PARK7 gene. Furthermore, targeted proteins and selective biomarkers enables more effective and applicable treatment, eliminating the further growth of the disease. Safinamide could be a promising inhibitor of the PARK7 protein, which can be used to treat PARK7 related PD in future. These findings may lead to the discovery of other novel inhibitors against the PARK7 protein and all age-related neurodegenerative disorders after more advanced research. Thus, Pave way for a better perception of Parkinson's pathophysiology and entirely focusing on all therapeutic approaches for PD.

## ACKNOWLEDGMENT

Not applicable.

## References

- [1] Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkman J, Schrag AE, Lang AE. Parkinson disease. Nature reviews Disease primers. 2017 Mar 23;3(1):1-21.
- [2] Musco VJ, Annesley SJ, Fisher PR. The DJ-1 gene and protein: links with Parkinson's disease. In Genetics, Neurology, Behavior, and Diet in Parkinson's Disease 2020 Jan 1 (pp. 35-49). Academic Press.
- [3] National Library of Medicine: MedlinePlus Genetics, PARK7 gene: Parkinsonism associated deglycase, [cited 2020 August 18]. Available from: <https://medlineplus.gov/genetics/gene/park7/#:~:text=Researchers%20have%20identified%20more%20than,which%20begins%20before%20age%2050>.
- [4] Dolgacheva LP, Berezhnov AV, Fedotova EI, Zinchenko VP, Abramov AY. Role of DJ-1 in the mechanism of pathogenesis of Parkinson's disease. Journal of bioenergetics and biomembranes. 2019 Jun;51(3):175-88.
- [5] Maggiorani D, Manzella N, Edmondson DE, Mattevi A, Parini A, Binda C, Mialet-Perez J. Monoamine oxidases, oxidative stress, and altered mitochondrial dynamics in cardiac ageing. Oxidative medicine and cellular longevity. 2017 May 4;2017.
- [6] Kumar B, Prakash Gupta V, Kumar V. A perspective on monoamine oxidase enzyme as drug target: challenges and opportunities. Current drug targets. 2017 Jan 1;18(1):87-97.
- [7] National Center for Biotechnology Information, PubChem Compound Summary for CID 3052776, Rasagiline; [Retrieved April 20, 2021], Available from <https://pubchem.ncbi.nlm.nih.gov/compound/Rasagiline>
- [8] National Center for Biotechnology Information, PubChem Compound Summary for CID 131682, Safinamide; [Retrieved April 20, 2021], Available from <https://pubchem.ncbi.nlm.nih.gov/compound/Safinamide>.
- [9] National Center for Biotechnology Information, PubChem Compound Summary for CID 26757, Selegiline; [Retrieved April 20, 2021]. Available from <https://pubchem.ncbi.nlm.nih.gov/compound/Selegiline>.
- [10] Finberg JP, Rabey JM. Inhibitors of MAO-A and MAO-B in psychiatry and neurology. Frontiers in pharmacology. 2016 Oct 18; 7:340.
- [11] Fabbri G, Abbruzzese G, Marconi S, Zappia M. Selegiline: a reappraisal of its role in Parkinson disease. Clinical neuropharmacology. 2012 May 1;35(3):134-40.
- [12] Schilde LM, Steinbach S, Serschnitzki B, Maass F, Bähr M, Lingor P, Marcus K, May C. Human cerebrospinal fluid data for use as spectral library, for biomarker research. Data in brief. 2020 Oct 1; 32:106048.
- [13] Gundry RL, White MY, Murray CI, Kane LA, Fu Q, Stanley BA, Van Eyk JE. Preparation of proteins and peptides for mass spectrometry analysis in a bottom-up proteomics workflow. Current protocols in molecular biology. 2010 Apr;90(1):10-25.
- [14] Schubert OT, Gillet LC, Collins BC, Navarro P, Rosenberger G, Wolski WE, Lam H, Amodei D, Mallick P, MacLean B, Aebersold R. Building high-quality assay libraries for targeted analysis of SWATH MS data. Nature protocols. 2015 Mar;10(3):426-41.
- [15] Govaert E, Van Steendam K, Willems S, Vossaert L, Dhaenens M, Deforce D. Comparison of fractionation proteomics for local SWATH library building. Proteomics. 2017 Aug;17(15-16):1700052.
- [16] Searle BC. Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies. Proteomics. 2010 Mar;10(6):1265-9.
- [17] Vlachakis D, Bencurova E, Papageorgiou L, Bhide M, Kossida S. Protein phosphorylation prediction: limitations, merits and pitfalls. Journal of Molecular Biochemistry. 2015 Jul 19;4(2).
- [18] Hornbeck PV, Kornhauser JM, Tkachev S, Zhang B, Skrzypek E, Murray B, Latham V, Sullivan M. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. Nucleic acids research. 2012 Jan 1;40(D1):D261-70.
- [19] Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery.
- [20] Karimizadeh E, Sharifi-Zarchi A, Nikaein H, Salehi S, Salamatian B, Elmi N, Gharibdoost F, Mahmoudi M. Analysis of gene expression profiles and protein-protein interaction networks in multiple tissues of systemic sclerosis. BMC medical genomics. 2019 Dec;12(1):1-2.

- [21] Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: network analysis and visualization of proteomics data. *Journal of proteome research*. 2018 Nov 19;18(2):623-32.
- [22] Seeliger D, de Groot BL. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *Journal of computer-aided molecular design*. 2010 May;24(5):417-22.
- [23] Riboldi GM, Di Fonzo AB. GBA, Gaucher disease, and Parkinson's disease: from genetic to clinic to new therapeutic approaches. *Cells*. 2019 Apr;8(4):364.
- [24] Gómez-Benito M, Granado N, García-Sanz P, Michel A, Dumoulin M, Moratalla R. Modeling Parkinson's disease with the alpha-synuclein protein. *Frontiers in pharmacology*. 2020;11.
- [25] McCoy MK, Cookson MR. DJ-1 regulation of mitochondrial function and autophagy through oxidative stress. *Autophagy*. 2011 May 1;7(5):531-2.
- [26] Biosa A, Sandrelli F, Beltramini M, Greggio E, Bubacco L, Bisaglia M. Recent findings on the physiological function of DJ-1: Beyond Parkinson's disease. *Neurobiology of disease*. 2017 Dec 1; 108:65

IEEESEM