Universidade de Aveiro Departamento de Química
Ano 2012/2013

Liliana Raquel Rodrigues Loureiro

Interactores de DJ-1 mutadas – Elucidações sobre a Doença de Parkinson

Universidade de Aveiro Departamento de Química
Ano 2012/2013

Liliana Raquel Rodrigues Loureiro

Interactores de DJ-1 mutadas – Elucidações sobre a Doença de Parkinson

DJ-1 mutants binding partners – Insights into Parkinson´s Disease

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biotecnologia, Ramo de Biotecnologia Molecular, realizada sob a orientação científica do Doutor Bruno Manadas, Investigador Auxiliar do Centro de Neurociências e Biologia Celular (CNC) da Universidade de Coimbra do Doutor António Correia, Professor Catedrático do Departamento de Biologia da Universidade de Aveiro.

Financiado por Fundos FEDER através do Programa Operacional Factores de Competitividade – COMPETE e por Fundos Nacionais através da FCT – Fundação para a Ciência e a Tecnologia no âmbito dos projectos PEst-C/SAU/LA0001/2013-2014 e PTDC/NEU-NMC/0205/2012.

o júri

presidente

Prof.ª Doutora Etelvina Maria de Almeida Paula Figueira
Professora auxiliar do Departamento de Biologia da Universidade de Aveiro

Prof.ª Doutora Isaura Isabel Gonçalves Simões
Investigadora Auxiliar do Centro de Neurociências e Biologia Celular (CNC) da Universidade de Coimbra

Prof. Doutor Bruno José Fernandes Oliveira Manadas
Investigador Auxiliar do Centro de Neurociências e Biologia Celular (CNC) da Universidade de Coimbra

agradecimentos

Outro capítulo da minha vida está encerrado e na conquista de mais um triunfo é de salientar a importância do apoio, confiança e colaboração das inúmeras pessoas que me rodearam e que actualmente me rodeiam. Assim começo por agradecer ao Doutor Bruno Manadas e Doutor Mário Grãos pela oportunidade concedida na realização deste projecto, mas mais importante pelo apoio, partilha de conhecimentos, compreensão e disponibilidade ao longo de todo o trabalho. Agradeço também ao Doutor António Correia pelo apoio e disponibilidade prestados.

Um agradecimento ainda para o Doutor Pedro Castanheira que me ensinou e auxiliou ao longo do processo de produção e purificação das proteínas. Na realização das análises de dicroísmo circular quero agradecer ao Doutor Tiago Faria e ao Doutor Rui Brito pela colaboração e disponibilidade.

Passo agora aos agradecimentos a todos os amigos e colegas do Biocant, em particular, um muito obrigada…

A todas as pessoas do grupo de Proteómica e Metabolómica que me acolheu e com as quais desenvolvi uma grande amizade. À Sandra e à Matilde um obrigada especial por todos os ensinamentos, orientação, apoio, disponibilidade e paciência, tendo sido dois pilares importantes na realização deste projecto. À Lúcia, com quem desenvolvi parte do projecto em conjunto, um grande obrigada pelo companheirismo, amizade e momentos de partilha. À Cátia, que se mostrou sempre disponível em partilhar os seus ensinamentos e ajuda. À Margarida por ouvir tantos desabafos da companheira de secretária. À Vera pelos conhecimentos e apoio partilhados. À Joana, à Alexandra, à Mariana, ao Zé e à Susana, que apesar de não terem estado sempre presentes, contribuíram com a sua amizade e bom humor.

A toda a unidade de Biotecnologia Molecular, por me terem recebido de braços abertos e partilhado os seus conhecimentos, incentivo e apoio estendendo essa partilha aos momentos de divertimento e convívio.

À unidade de Biologia Celular que se mostraram sempre disponíveis para ensinar e ajudar em todos os imprevistos e problemas que foram surgindo.

A todos os meus amigos que de uma forma directa ou indirecta contribuíram para a realização desta tese. Aqui tenho que mencionar a minha família basquetebolística, que me acompanha desde a infância, em especial à equipa de basquetebol do CPE, que me proporcionou momentos essenciais de descontracção e relaxamento durante este ano de tese.

Ao meu namorado, com quem diariamente partilhei as minhas inquietações, desânimos, alegrias e conquistas, “roubando-lhe” sempre algum tempo em cada dia que passava. Um muito obrigada pela paciência, apoio e carinho que sempre me deste durante todo este tempo que partilhámos.

À família Cunha que me “adotou” e tratou como uma verdadeira filha, acompanhando-me durante todo o meu percurso académico e transmitindo constantemente sentimentos de partilha, apoio e paciência. À “irmã” queria agradecer pela contínua disponibilidade em ouvir os meus desabafos e pelo apoio incansável.

Por último, mas detentores de maior importância…a minha família! Ao meu irmão pela sua incansável curiosidade que me leva a enriquecer diariamente as minhas capacidades em “ensinar”. Aos meus Pais, que com os seus ensinamentos, amor incondicional, incentivos, ajuda e alegria me acompanharam na conquista de todos os meus objectivos, fazendo com que nunca duvidasse de mim e mais importante, com que nunca me sentisse desamparada.

Muito Obrigada a Todos!

palavras-chave

DJ-1, Doença de Parkinson, mutações na DJ-1, stress oxidativo, neuroproteção, interactoma dinâmico, produção e purificação de proteínas, LC-MS/MS, dicroísmo circular, estabilidade térmica, HPLC de exclusão molecular, análise SWATH.

resumo

A Doença de Parkinson, a segunda doença neurodegenerativa progressiva mais comum, é uma doença multifatorial causada conjuntamente por fatores genéticos e fatores ambientais. De entre os inúmeros genes associados à Doença de Parkinson, a DJ-1 é uma proteína multifuncional envolvida na resposta ao stress oxidativo e neuroproteção. Mutações na DJ-1, tais como L166P, M26I e E136K levam à perda de função da proteína causando a forma de Parkinson autossomal recessiva com desenvolvimento precoce. De salientar que o resíduo C106 é considerado crucial na função de sensor de stress oxidativo que a DJ-1 desempenha. Neste estudo, foram produzidas e caracterizadas duas mutações sintéticas no resíduo C106 (C106A e C106DD) e uma mutação natural (E163K) de modo a avaliar o efeito neuroprotetor de cada mutação bem como caracterizar o seu interactoma dinâmico.

Análises estruturais confirmaram a produção de todos os mutantes na forma dimérica, apresentando um peso molecular de aproximadamente 43kDa. A estabilidade térmica das proteínas foi ainda avaliada por *thermal shift* e os resultados revelaram que o mutante E163K foi o menos estável enquanto que o mutante C106A foi o mais estável. Análise da estrutura secundária foi realizada por dicroísmo circular revelando elevada semelhança entre as estruturas secundárias da DJ-1 nativa e mutantes.

Por fim, foi realizada uma análise de LC-MS/MS de modo a determinar os contaminantes das proteínas produzidas e verificou-se que a maioria dos contaminantes era proveniente do sistema de expressão e meio de cultura utilizados na produção das proteínas.

Seguidamente, ensaios de neuroproteção revelaram que a DJ-1 nativa não exercia um efeito neuroprotetor nas células SH-SY5Y em condições de stress oxidativo.

O interactoma dinâmico da DJ-1 nativa e mutantes C106DD, E163K e C106A foi caracterizado sob condições de stress oxidativo. Um elevado número de interactores foram identificados e para alguns deles foi possível obter uma quantificação nas diferentes condições. Os referidos interactores apresentam uma enorme variedade de funções, contudo a grande maioria está associada à resposta celular ao stress oxidativo.

O estudo das mutações na DJ-1 é considerado muito relevante visto que fornece importantes elucidações relativamente às funções e mecanismos da DJ-1 nativa associados à doença. Neste sentido, os supostos interactores da DJ-1 nativa identificados ainda carecem de validação, mas da caracterização dos interactomas dinâmicos, elucidações podem ser obtidas sobre a patologia da Doença de Parkinson e identificação de novos potenciais alvos para prevenção e terapia desta doença, tal como a RNA helicase DDX3X dependente de ATP aqui identificada como novo interactor dinâmico da DJ-1.

keywords

DJ-1, Parkinson´s Disease, DJ-1 mutations, oxidative stress, neuroprotection, dynamic interactome, protein production and purification, LC-MS/MS, circular dichorism, thermal shift, size exclusion-HPLC, SWATH analysis.

abstract

Parkinson´s disease (PD), the second most common progressive neurodegenerative disorder, is a multifactorial disease caused by both genetic and environmental factors. Among the genes associated with PD, DJ-1 is a multifunctional protein involved in oxidative stress response and neuroprotection. DJ-1 mutations, such as L166P, M26I and E163K lead to loss of protein function causing early onset autosomal recessive PD. Moreover, the residue C106 is considered crucial in DJ-1 function as a sensor of oxidative stress. In this study, one missense mutations (E163K) and two engineered mutations in the residue C106 (C106A and C106DD) were produced and characterized in order to evaluate the neuroprotective effect of each mutation and also characterize their dynamic interactome.

Structural analysis confirmed the production of all the mutants in the dimeric form, with a molecular weight of approximately 43kDa. Moreover, protein´s thermal stability was assessed by thermal shift and the results showed that the mutant E163K was the less stable and the C106A the most stable. Secondary structure analysis was performed by circular dichroism and revealed similar secondary structures between DJ-1 WT and mutants.

In addition, a LC-MS/MS was performed to determine proteins´ contaminants and the majority of the protein contaminants were coming from the expression system and culture medium used in proteins´ production.

Moreover, neuroprotection assays revealed that DJ-1 WT did not protect SH-SY5Y cells under oxidative stress conditions.

The dynamic interactome of DJ-1 WT and mutants C106DD, E163K and C106A was characterized under oxidative stress conditions. A wide number of binding partners were identified and for some of them quantification in the different conditions was also determined. These interactors have a broad range of functions but the majority are associated with cellular response to oxidative stress.

The study of DJ-1 mutations is very important, since it gives elucidations into DJ-1 WT functions and related disease mechanisms. In this way, the putative DJ-1 WT interactors identified still lack validation, but from these characterized dynamic interactomes further elucidations can be obtained into Parkinson's Disease pathology and potential new targets for PD prevention and therapy, like ATP-dependent RNA helicase DDX3X herein identified as new dynamic interactor of DJ-1.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

6-OHDA	6-hydroxydopamine

A

ACN	acetonitrile
AD	activation domain
ANOVA	analysis of variance
AP–MS/MS	affinity purification coupled to tandem mass spectrometry
AR-PD	autosomal recessive juvenile parkinsonism
ASK1	apoptosis signal regulating kinase 1

B

BCA	bicinchoninic acid
BSA	bovine serum albumin

C

CCD	charge-coupled device
CD	circular dichroism
CID	collision-induced dissociation
Cys-SO_2H	cysteine-sulfinic acid
Cys-SO_3H	cysteine-sulfonic acid
Cys-SOH	cysteine-sulfenic acid

D

DAQ	dopamine quinones
DBD	DNA binding domain
DMEM	dulbecco's modified eagle medium
DPBS	dulbecco's phosphate buffered saline

E

ER	endoplasmic reticulum
ESI	electrospray ionization

F

FA	formic acid
FBS	fetal bovine serum
FDR	false discovery rate

G

GBA	glucocerebrosidase gene
GO	gene ontology
GOrilla	Gene Ontology enRIchment anaLysis and visuaLizAtion tool
GSH	glutathione

GWAS	genome-wide association study

H

H_2O_2	hydrogen peroxide
HTRA2	high temperature requirement A2

I

IDA	information dependent acquisition
iRT	indexed retention times

K

KRS	Kufor-Rakeb syndrome

L

LB	Lewy bodies
LC	liquid chromatography
LC-MS/MS	liquid chromatography - tandem mass spectrometry
LD50	lethal dose 50
LRRK2	leucine-rich repeat kinase 2

M

MALDI	matrix-assisted laser desorption ionization
MAPT	microtubule-associated protein tau
MS	mass spectrometry
mtDNA	mitochondrial DNA
MW	molecular weight

N

Nrf2	nuclear factor erythroid 2-related factor

O

OD	optical density

P

PBS	phosphate buffered saline
PD	Parkinson´s disease
PINK1	PTEN-induced putative kinase 1
PPI	protein-protein interactions

Q

q-AP-MS	quantitative AP-MS

R

RNA	ribonucleic acid
ROS	reactive oxygen species

rpm	revolutions per minute
rPQ	relative peptide query
RS	regulatory subunit

S

SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SNpc	*substantia nigra pars compacta*
SE-HPLC	size exclusion – high-performance chromatography
SWATH	Sequential Windowed data independent Acquisition of the Total High-resolution mass spectra

T

TAP	tandem affinity purification
T_m	melting temperature

U

UBL	amino terminal ubiquitin-like
UCHL1	ubiquitin c-terminal hydrolase 1
UPS	ubiquitin-proteasome system
UV	ultraviolet

W

WT	wild-type

X

XIC	eXtracted-Ion Chromatogram

Y

Y2H	yeast two-hybrid

1. INTRODUCTION

1.1. PARKINSON´S DISEASE

Parkinson's disease (PD) was first described in 1817 by James Parkinson in the classic ''Essay on the Shaking Palsy'' [1], and is the second most common neurodegenerative disorder, after Alzheimer's disease [2].

Parkinson´s disease prevalence increases with age, affecting 1-2% of the population over the age of 65 years and reaching a prevalence of almost 4% in those aged above 85 years [3, 4]. Due to demographic changes and increased life expectancy the prevalence of PD is expected to double by the year 2030 as populations age, forecasting an impending burden on the healthcare systems of many countries [3, 5].

From the clinical point of view, PD is characterized by cardinal symptoms, usually described as "parkinsonism", including resting tremor, rigidity, bradykinesia, and postural instability [2, 4]. Furthermore, many non-motor symptoms frequently appear in PD, such as cognitive impairment, behavioral abnormalities, autonomic insufficiency and sleep disorders [4].

The pathological hallmark of classical PD includes loss of nigrostriatal dopaminergic neurons in the *substantia nigra pars compacta* (SNpc) leading to loss of dopamine in the striatum, and the presence of Lewy bodies (LB) and Lewy neurites in the surviving neurons (Figure 1.1) [2-4]. LB are dense cytosolic inclusions that contain a variety of insoluble proteins being the main component fibrillar α-synuclein [6, 7]. PD diagnoses are typically based on the presence of LB, however some monogenic forms of PD lack this characteristic (LB) pathology [3]. In addition, the symptoms of PD can vary extensively between patients and many neurological insults can cause parkinsonism. Therefore, a post-mortem examination of the neural tissue must be done in order to obtain a definitive diagnosis of PD [5]. Symptoms appear when about 50–60% of the nigral neurons and about 80–85% of the dopamine content of the striatum are lost [2]. Unfortunately, only some improvements of the symptoms are obtainable by current treatments, which are based on levodopa and dopaminergic therapy, however currently there is no available treatment to detain the progression of the disease [4].

PD can be considered mostly as sporadic (90-95%) and to a lesser extent as familial (5-10%) [8, 9]. The etiology of PD is considered complex and multifactorial,

resulting probably from an elaborated interplay of several factors, like genes, modifying effects by susceptibility alleles, environmental exposures and gene-environment interactions (e.g., influence of environmental agents on gene expression), and their direct impact on the developing and aging brain [10]. Although the majority of cases appear sporadic in nature, the discovery of genes linked to rare familial forms of PD have confirmed the role of genetics in development of PD, providing vital clues in understanding molecular pathogenesis of this disease [9, 11]. Even though knowledge has grown rapidly, the pathogenesis of PD is still unknown [2].

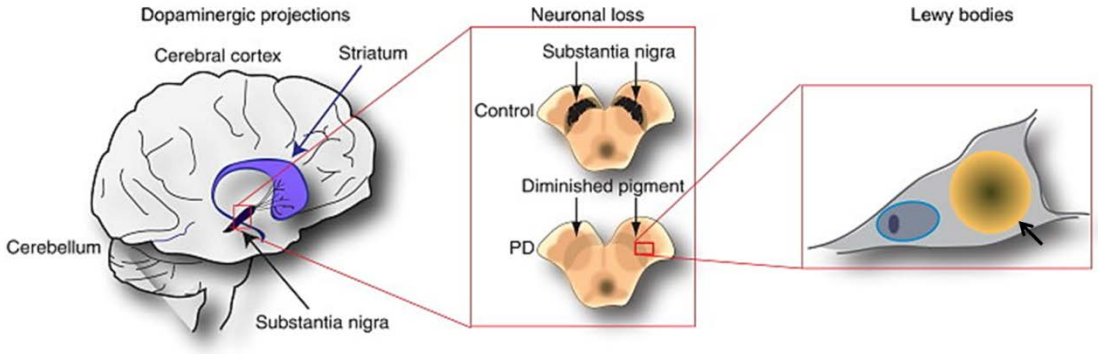


Figure 1.1 | PATHOLOGY OF PARKINSON'S DISEASE. A simplified view of the main neuropathological events in PD is shown at three levels from left to right. In the brain, a major pathway is degeneration of the dopaminergic projections from the substantia nigra (in black) to the striatum (in purple). At the level of substantia nigra, the neurons are normally melanized and are easily identified by this pigment in control brains (upper panel). In contrast, the loss of neurons in this region is substantial leading to depigmentation in PD cases (lower panel). Of the few remaining cells, many show pathological changes, including the accumulation of proteins and lipids in LB (arrow, right panel) [12]**.**

1.1.1. ETIOLOGY OF PD

Environmental factors were long thought to be the predominant cause of Parkinson's disease, but currently PD is thought to arise from a combination of environmental exposures, and complex gene-environment interactions superimposed on slow and sustained neuronal dysfunction due to aging (Figure 1.2) [4, 13].

The environmental hypothesis was supported by the identification, in the early 1980s, of an exogenic toxin that caused Parkinsonian symptoms. After intravenous application of a synthetic drug that contained 1-methyl-4-tetrahydropyridine (MPTP),

some of the drug addicts developed an akinetic-rigidic syndrome that responded to levodopa therapy [14]. MPTP causes selective degeneration of the nigrostriatal pathway by inhibiting mitochondrial complex I leading to a parkinsonian syndrome in rodents, primates, and humans (Figure 1.3) [15]. Since the discovery that MPTP causes parkinsonism in humans and non-human primates, as well as in various other mammalian species, it has been extensively used as a model of PD [16].

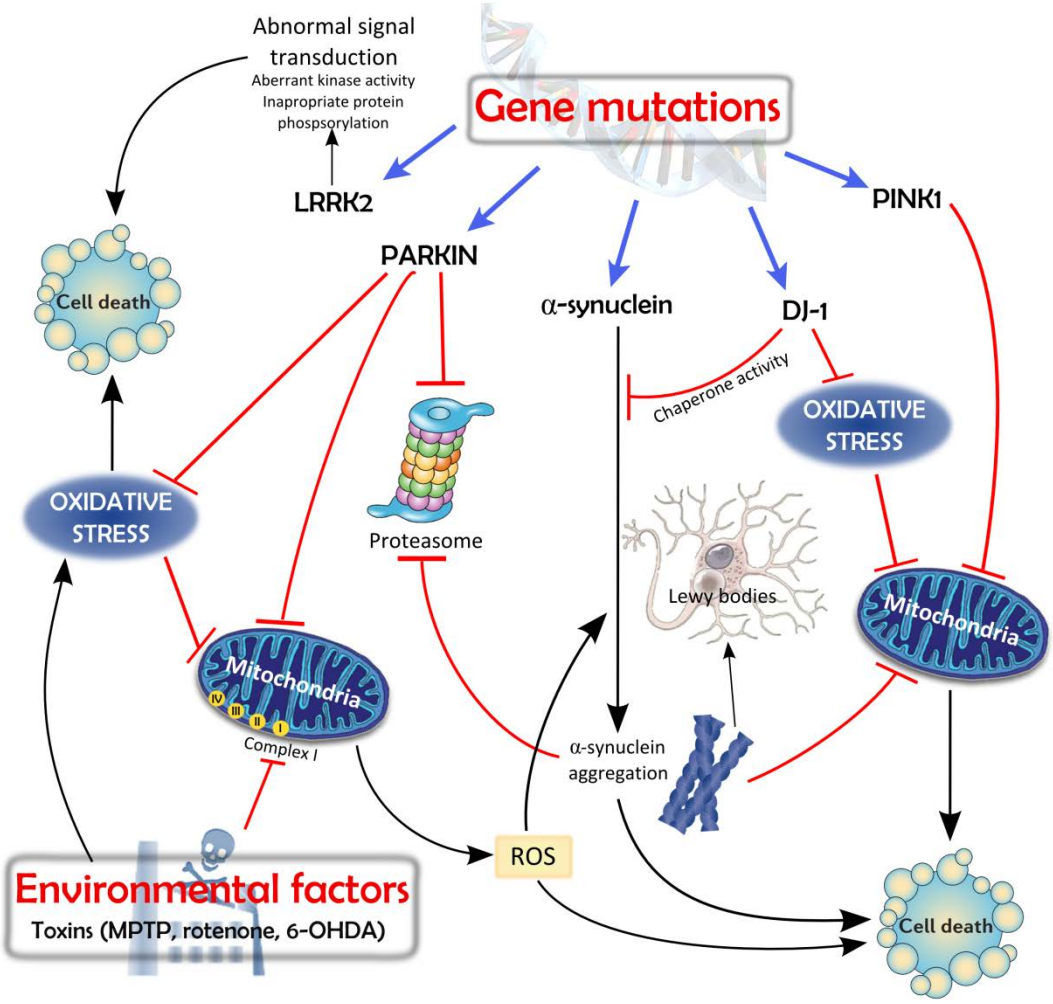


Figure 1.2 | KEY ENVIRONMENTAL AND GENETIC FACTORS THAT CONTRIBUTE TO THE DEVELOPMENT OF PARKINSONISM. The etiology of PD is thought to be multifactorial with contributions from environmental and genetic factors. Both factors are involved in several pathways, such as mitochondrial dysfunction, oxidative stress, proteasome inhibition, LB and ROS formation, which ultimately lead to cell death and neurodegeneration [17-21].

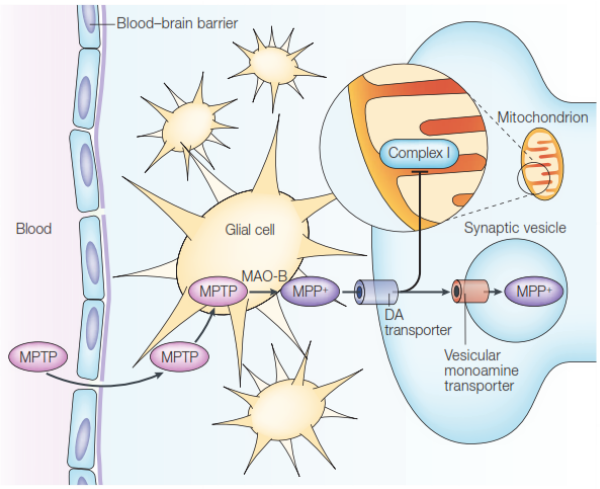


Figure 1.3 | **REPRESENTATION OF MPTP METABOLISM**. MPTP can cross the blood–brain barrier and is metabolized to MPP^+ with the assistance of the enzyme monoamine oxidase B (MAO-B) in glial cells. MPP^+ is then taken up into neuronal cells by dopamine transporter for which it has high affinity. Once inside neurons, MPP^+ is concentrated within the mitochondria, where it impairs mitochondrial oxidative phosphorylation by inhibiting complex I of the respiratory chain [18].

Toxicological studies showed that subacute exposure to other toxins, like rotenone, paraquat and 6-hydroxydopamine (6-OHDA) also induces behavioral and pathological changes characteristic of PD in animal models [19]. Several epidemiological studies showed that additional environmental factors, like rural living, well-water drinking, prolonged pesticide/insecticide exposure, and metal exposure, also increase the risk of developing PD [22]. Furthermore, some studies showed that cigarette smoking and coffee or tea drinking could be protective against PD, but the physiologic mechanisms for these relations are poorly understood [23].

Although environmental factors may play an important role in PD etiology, in the past 15 years it has been witnessed a seed change in the etiologic basis of this disease. This change has been largely driven by genetics, progressing through the early recognition of familial forms of the disease, to include the identification of gene mutations that cause rare familial forms of PD, finding moderate risk variants, and most recently the mapping of multiple low-risk conferring loci by genome-wide association studies (GWAS) [24].

The current view of PD etiology is that it is multifactorial in which environmental insults and genetic predisposition interact to cause the disease. Still, knowledge about the roles of individual pathogenic factors may lead to the useful insights into the disease process and further identification of targets responsive to therapy.

1.1.2. MONOGENIC FORMS OF PD

It is currently believed that PD is largely sporadic, meaning that the disease arises in individuals without a family history of PD. However, some patients have monogenic PD, where the cause of the disease can be ascribed to mutations in single genes that have been convincingly demonstrated to be pathogenic [13]. Monogenic PD probably accounts for about 5–10% of the total population of PD patients and exhibit an autosomal dominant or recessive mode of inheritance [3].

During the past few years, 18 *loci*, also called PARK *locus*, and a few other genes have been assigned through linkage analysis or GWAS (Table 1.1) [10, 25].

Table 1.1 | GENE/*LOCI* UNDERLYING MONOGENIC PARKINSONISM [10, 25-27].

PARK *Locus*	Gene	Map Position	Inheritance	Mutations Identified (Pathogenic)	Clinical Phenotype	Pathology
PARK 1/4	***SNCA***	**4q21**	**AD;Rarely sporadic**	**27 (25)**	**EOPD**	**LB+**
PARK 2	***Parkin***	**6q25-q27**	**AR; Sporadic**	**214 (128)**	**Juvenile and EOPD**	**LB- (most brains)**
PARK 3	*Unknown*	2p13	AD	NA	LOPD	LB+
PARK 5	*UCHL1*	4p14	AD	NA	LOPD	Unknown
PARK 6	***PINK1***	**1p35-p36**	**AR**	**138 (28)**	**EOPD**	**LB+ (only 1 brain**
PARK 7	***DJ-1***	**1p36**	**AR**	**28 (6)**	**EOPD**	**Unknown**
PARK 8	***LRRK2***	**12q12**	**AD; Sporadic**	**128 (6)**	**LOPD**	**LB+ (most brains)**
PARK 9	***ATP13A2***	**1p36**	**AR**	**NA**	**Kuffor-Rakeb syndrome**	**Unknown**
PARK 10	*Unknown*	1p32	Unclear	NA	LOPD	Unknown
PARK 11	*GIGYF2*	2q36-q37	AD	NA	LOPD	Unknown
PARK 12	*Unknown*	Xq21-q25	X-linked	NA	LOPD	Unknown
PARK 13	*HTRA2*	2p13	AD	NA	Not clear	Unknown
PARK 14	*PLA2G6*	22q12-q13	AR	NA	Juvenile levodopa-responsive dystonia-parkinsonism	LB+
PARK 15	*FBXO7*	22q12-q13	AR	NA	EOPD pyramidal syndrome	Unknown
PARK 16	*Unknown*	1q32	Unclear	NA	LOPD	Unknown
PARK 17	*VPS35*	16q11.2	AD	NA	LOPD	
PARK 18	*EIF4G1*	3q27.1	AD	NA	LOPD	
FTDP-17	*MAPT*	17q21.1	-	-	Dementia (sometimes parkinsonism)	Neurofibrillary tangles
Gaucher's locus	*GBA*	1q21	Recessive for Gaucher´s disease	-	LOPD	LB+
SCA2	*Ataxin 2*	12q24.1	Ataxia (sometimes parkinsonism)	-	Not clear	Unknown
SCA3	*Ataxin 3*	14q21	Ataxia (sometimes parkinsonism)	-	Not clear	Unknown

Abbreviations: AD, Autosomal dominant; AR, autosomal recessive; EOPD, Early-onset Parkinson´s Disease; LOPD, Late-onset Parkinson Disease; LB+, Presence of LB pathology; NA, not available.
Note: Well-validated monogenic forms are listed in bold.

Some follow-up genetic studies have been inconsistent for certain genes, nonetheless there is evidence that mutations in five of those genes (*SNCA*, *parkin*, *PINK1*,*DJ-1* and *LRRK2*) cause typical PD [28], where *SNCA* and *LRRK2* are conclusively associated with autosomal dominant PD, and *parkin*, *PINK1* and *DJ-1* with early-onset autosomal recessive PD [29]. Moreover, mutations in *ATP13A2* gene cause Kufor-Rakeb disease, a neurodegenerative disease characterized by parkinsonism [30].

Polymorphic variants in *SNCA* and *LRRK2* and in candidate genes such as microtubule-associated protein tau (*MAPT*), as well as rare mutations in the glucocerebrosidase gene (*GBA*) have also emerged as susceptibility factors in several populations [25].

1.1.2.1. SNCA

α-synuclein is a 140 amino-acid protein encoded by the SNCA gene that is highly abundant in the mammalian brain particularly in presynaptic nerve terminals, and also present in other tissues, like red blood cells[4]. It was originally identified in association with synaptic vesicles in the presynaptic nerve terminal [31] and has been revealed to interact with membranes both *in vitro* and *in vivo* [10]. Fibrillar forms of α-synuclein are the major structural component of LB, the pathogenomic hallmark of all forms of PD [32]. Currently, studies show that mutations in α-synuclein lead to an increased tendency of the protein to form oligomers and fibrillar aggregates critical to LB formation, suggesting a link between α-synuclein aggregation and PD [33]. From the information given before, there are evidences that α-synuclein could lead to PD pathology through both toxic-gain of function of mutant protein and over-production of wild-type (WT) protein.

In 1997 *SNCA* mutations were the first to be identified in monogenic PD [24], and patients with variants in this gene are usually associated to early-onset PD with a rapid progression (sometimes with atypical clinical features) [10].

Three missense mutations (A53T, A30P and E64K), and duplications or triplications of the locus containing *SNCA* have been identified in families with autosomal dominant forms PD (Figure 1.4) [34]. Furthermore, clinical characteristics of patients with duplications presented typical late-onset PD phenotype while patients

with triplications had lower age at onset, rapid progression, and more often dementia [29]. Hence, the severity of clinical phenotype in *SNCA* multiplication seems to be gene dose dependent [34], where an increased number of *SNCA* copies, and therefore higher levels of the protein encoded by this gene, α-synuclein, is associated with a more severe phenotype, and faster disease progression [31]. Importantly, GWAS linked single nucleotide polymorphisms in the SNCA gene with increased susceptibility to sporadic PD [13].

The biological function of α-synuclein is still not fully understood, however several studies suggest that it is involved in modulating neuronal plasticity and synaptic transmission, as well as providing a stabilizing effect on complexes of SNARE proteins critical for neurotransmitter release, vesicle recycling, and synaptic integrity [21]. Furthermore, several studies associate the *SNCA*-expression levels with the appearance and possible mechanisms underlying PD [26, 32].

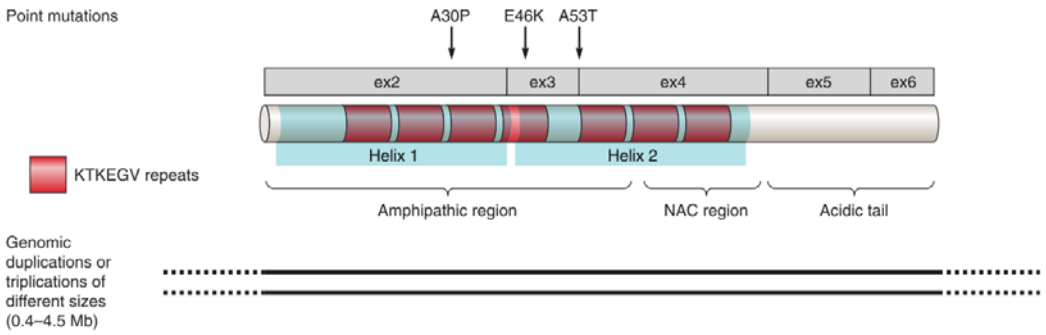


Figure 1.4 | SCHEMATIC REPRESENTATION OF A-SYNUCLEIN PROTEIN. There are three pathogenic missense mutations represented above the protein organization and duplications or triplications of different genomic sizes below the protein (red lines). α-synuclein has seven imperfect repeats (KTKEGV) in the amino terminal, a central hydrophobic NAC (nonamyloid component) region, and an acidic carboxilic-terminal region [16].

1.1.2.2. LRRK2

The leucine-rich repeat kinase 2 (LRRK2) gene encodes for a 2527 amino acid protein named dardarin, which has several predicted functional domains (e.g. Roc GTPase, Roc-COR tandem domain and protein kinase domains) (Figure 1.5) suggesting that this protein has multiple functions, including scaffold protein function, substrate binding, and protein phosphorylation [21]. However, how these domains interact in

order to influence *LRRK2* function under physiological and pathophysiological conditions remains to be fully characterized [33].

In cellular models, evidences show that overexpression of disease-causing *LRRK2* mutants is toxic and that toxicity is associated with kinase activity, GTP-binding, and mitochondria-dependent apoptosis [21].

Mutations in the LRRK2 gene, the most common known cause of PD, are associated with late-onset autosomal dominant (accounting for an estimated 5–10% of occurrences) and sporadic forms of PD (an estimated 1–5% of occurrences) [13]. So far, over 100 variations in *LRRK2* have been described (http://www.molgen.ua.ac.be/PDmutDB) [28] but only seven mutations have been proven to be pathogenic (N1437H, R1441G/C/H, Y1699C, G2019S, and I2020T) [3]. The G2019S mutation is the most common, found in a wide range of ethnic groups, and in 1– 3% of sporadic and 4–8% of familial cases (Figure 1.5) [26, 29]. In addition to these mutations that alone are sufficient to cause the disease, two common non-synonymous SNPs in *LRRK2* (G2385R and R1628P) appear to be true risk variants for PD in the Asian population [16].

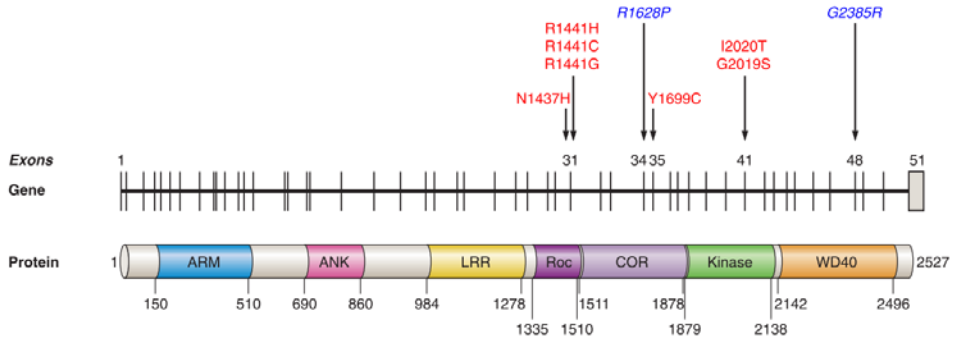


Figure 1.5 | SCHEMATIC REPRESENTATION OF LRRK2 GENE AND PROTEIN. This gene has seven proven pathogenic mutations (in red) and the two Asian-specific risk factors (in blue) above the protein organization. *LRRK2* contains several conserved domains: ARM (Armadillo), ANK (ankyrin repeat), LRR (leucine-rich repeat), Roc (Ras of complex proteins: GTPase), COR (carboxilic terminal of Roc), kinase domain, and WD40 [16].

One significant feature of pathogenic *LRRK2* mutations is that they typically result in clinical phenotypes that resemble late-onset sporadic PD even though LB pathology is sometimes absent or lacking [26].

PD-causative mutations have been localized in several different domains, and it is possible that these mutations could interfere with LRRK2 protein functions in different ways [13].

1.1.2.3. PARKIN

PARK2 gene codes for a 465 amino acid protein (Figure 1.6)- parkin which is expressed widely, but most prominently in muscle and brain [35, 36]. Parkin contains an amino terminal ubiquitin-like (UBL) domain followed by three RING finger domains (RING 0–2) separated by a 51-residue IBR domain in the carboxylic terminal part [16]. Functionally, parkin protein is a member of a family of E3 ubiquitin ligases responsible for tagging proteins for proteasomal degradation mediated by the transfer of activated ubiquitin molecules to a protein substrate [16, 35]. Moreover, under physiological conditions, parkin is also involved in mitochondrial maintenance and clearance of dysfunctional mitochondria by autophagy (mitophagy) [21]. Most PD-causing mutations in Parkin abolish its E3 ubiquitin ligase activity causing a loss of parkin catalytic competency [16]. Furthermore, several studies refer that dopaminergic, oxidative, and nitrosative stress, key players in PD, can inactivate Parkin indicating that

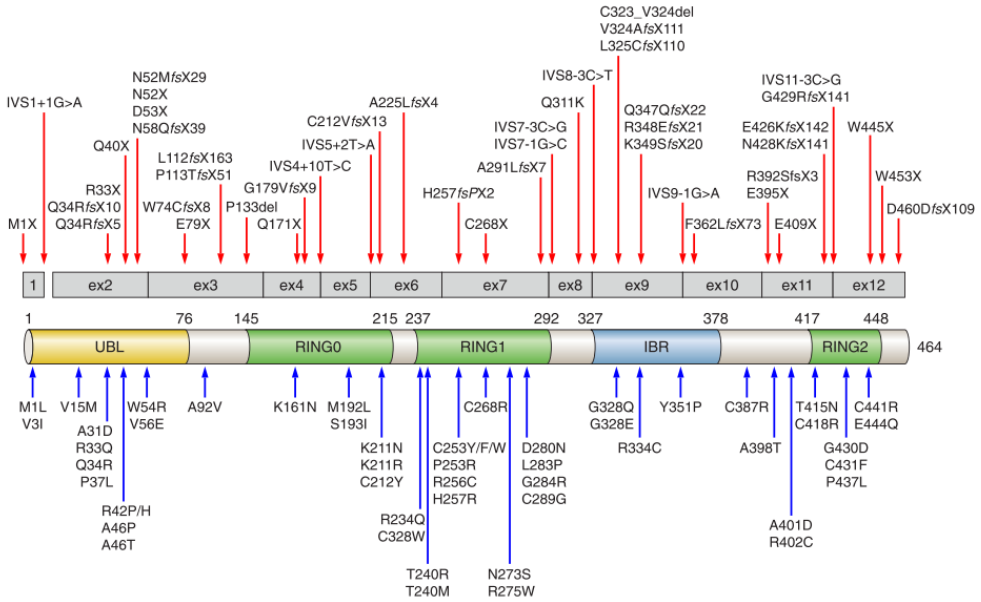


Figure 1.6 | SCHEMATIC REPRESENTATION OF *PARKIN* ON TRANSCRIPT LEVEL AND FUNCTIONAL DOMAINS. Exonic deletions above the transcript (red lines) and exonic duplications (green lines) or triplications (blue lines) below the transcript. Only homozygous or compound heterozygous mutations are listed [16].

parkin dysfunction plays a role in both autosomal recessive parkinsonism (owing to *parkin* mutations) and sporadic PD [37]. Thus, accumulation of parkin substrates is thought to contribute to DA neurodegeneration owing to parkin inactivation in the pathogenesis of PD but there is the need of further clarification [26].

Mutations in PARK2 gene have been identified as the most common known cause of autosomal recessive juvenile parkinsonism (AR-JP) [36]. Currently it is thought that up to half of hereditary parkinsonism and 10% of all early-onset PD cases are associated with mutations in this gene [38]. Besides this, more than 100 putatively pathogenic mutations have been observed in PD families (Figure 1.6) [28]. Clinically, the disease is usually indistinguishable from early onset idiopathic PD, characterized by early onset and a marked response to levodopa treatment but generally there is no LB formation [24].

1.1.2.4. PINK1

PINK1 protein is a kinase with 581 amino acids, containing an N-terminal mitochondrial targeting motif and a highly conserved serine/threonine kinase domain (Figure 1.7) [39]. It is widely expressed and mainly localizes in mitochondria, especially in the outer membrane [32]. Studies suggest that PINK1 may regulate mitochondrial response to cellular and oxidative stress, protecting cells from mitochondrial dysfunction [21, 39]. This protein has been reported to participate in multiple mitochondrial functions, such as calcium dynamics, mitochondrial trafficking and respiration efficacy, reactive oxygen species (ROS) formation, and opening of the mitochondrial permeability transition pore [21]. Interestingly, recent studies provided evidence that PINK1 and Parkin function in a common pathway in the maintenance of mitochondrial quality via autophagy (mitophagy) [10, 26, 40].

Mutations in the *PINK1* (PTEN-induced putative kinase 1) gene are so far the second most common cause of autosomal recessive early-onset PD after mutations in the parkin gene [39]. To date, more than 50 pathogenic mutations have been reported in *PINK1* (http://www.molgen.ua.ac.be/PDmutDB/) (Figure 1.7) [41]. It is thought that pathogenic mutations in *PINK1* influences its kinase activity, stability and localization [21, 39]. Most of the mutations are observed in or near the serine/threonine

kinase domain, suggesting that the disruption of kinase activity plays a crucial part in the pathogenesis of *PINK1*-linked PD [40]. Clinically, patients usually present a later age at onset in between 30 and 50 years but otherwise are similar to *parkin* related disease [24].

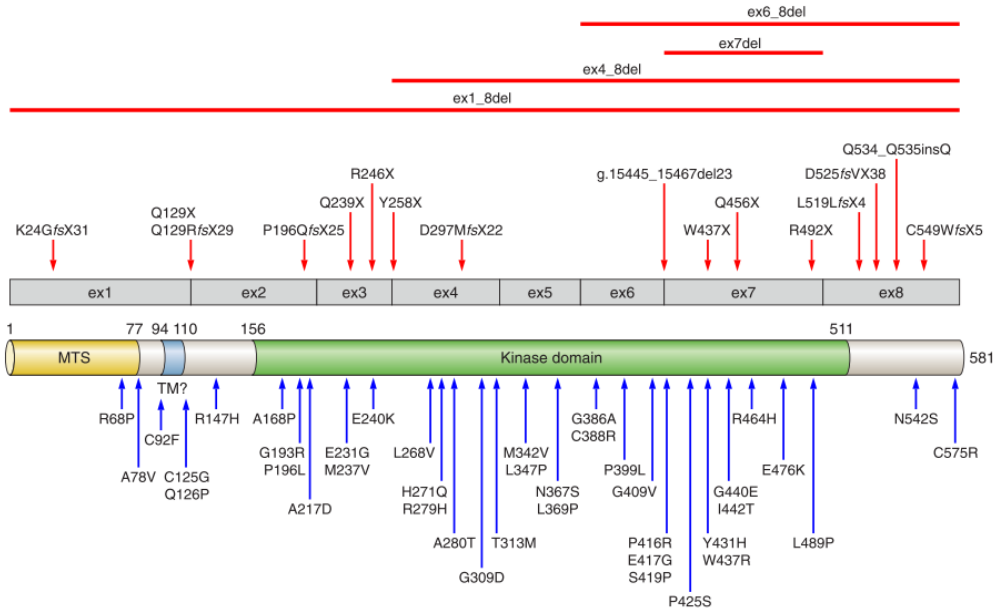


Figure 1.7 | SCHEMATIC REPRESENTATION OF *PINK1* ON TRANSCRIPT LEVEL AND FUNCTIONAL DOMAINS. Rare deletions (above, red lines). Only homozygous or compound heterozygous mutations are listed. *PINK1* has an amino terminal mitochondrial targeting signal (MTS) motif, a putative transmembrane (TM) region, and a serine-threonine kinase domain [16].

1.1.2.5. DJ-1

PARK7/DJ-1 was the third gene associated with AR-PD, is located on chromosome 1p36.23, contains eight exons, the first two of which are noncoding and subject to alternative splicing, and codify a 189-amino acid protein of approximately 20 kDa, the DJ-1 protein (Figure 1.8) [42, 43]. DJ-1 is a multifunctional protein that has been widely recognized as involved in the oxidative stress response pathways and this is believed to have an important role in the preservation of neuronal viability [2, 44]. However, the exact function and regulation of DJ-1 remains to be elucidated and this has been an active field of study for over a decade [45].

Several mutations have been associated to this gene (Figure 1.8) leading to a form of parkinsonism characterized by an age of onset in the 20s or 30s with resting tremor, postural tremor, bradykinesia, loss of postural reflexes, an asymmetric onset

of symptoms, and slow progression with response to levodopa [24, 46]. A more detailed description of all the aspects related to DJ-1 and its mutations will be given in following topics of this chapter.

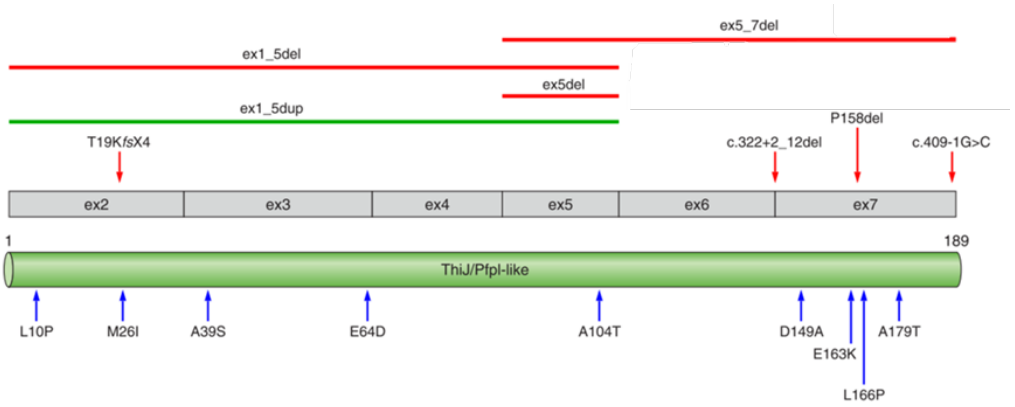


Figure 1.8 | SCHEMATIC REPRESENTATION OF DJ-1 ON TRANSCRIPT LEVEL AND FUNCTIONAL DOMAINS. Pathogenic frameshift mutations above the transcript and protein organization and missense mutations below the protein. Rare deletions (red lines) and duplications (green lines) (above) are represented. Only homozygous or compound heterozygous mutations are listed [16].

1.1.2.6. ATP13A2

In recent years, another gene has been associated to autosomal recessive forms of parkinsonism, the ATP13A2 gene (PARK9) [32]. *ATP13A2* encodes a P-type adenosine triphosphatase (ATPase) [47], which is localized to the lysosomal membrane with ten transmembrane domains [32]. Interestingly, truncating mutations promote the mislocalization of ATP13A2 to the endoplasmic reticulum (ER) where they are subsequently degraded by the proteasome [47]. This might indicate that lysosomal dysfunction could play an important role in the pathogenesis of parkinsionism and may provide crucial insights into the pathogenic mechanisms of PD [30]. The function of ATP13A2 protein is still largely unknown, although it seems to participate in protein degradation, like protecting against α-synuclein misfolding and toxicity, and lysosomal pathway, where loss of ATP13A2 function might lead to insufficient lysosomal protein degradation [30]. Thus, ATP13A2 may represent the first genetic link between lysosomal impairment and PD [32].

Mutations in this gene have been identified as causative in Kufor-Rakeb Syndrome (KRS), a rare form of autosomal recessive juvenile or early-onset, levodopa-responsive type of parkinsonism [47, 48]. Numerous mutations (homozygous or

compound heterozygous and truncating mutations) in ATP13A2 gene have been reported and considered pathogenic, leading to a loss-of-function of the ATP13A2 protein [48].

1.1.3. PATHWAYS RELATED TO PD

The discovery of PD genes, together with the pathological and epidemiological investigation of sporadic PD cases, has provided some insights into the main pathways of PD pathogenesis [49].

Even though the molecular mechanisms that can trigger degeneration of SNpc neurons are not fully understood, there are evidences indicating mitochondrial impairment, oxidative stress, accumulation of misfolded and aggregated proteins, and impairment of clearance systems such as ubiquitin-proteasome system (UPS) and the autophagy pathway as the main factors involved in the neurodegenerative process (Figure 1.9) [50, 51].

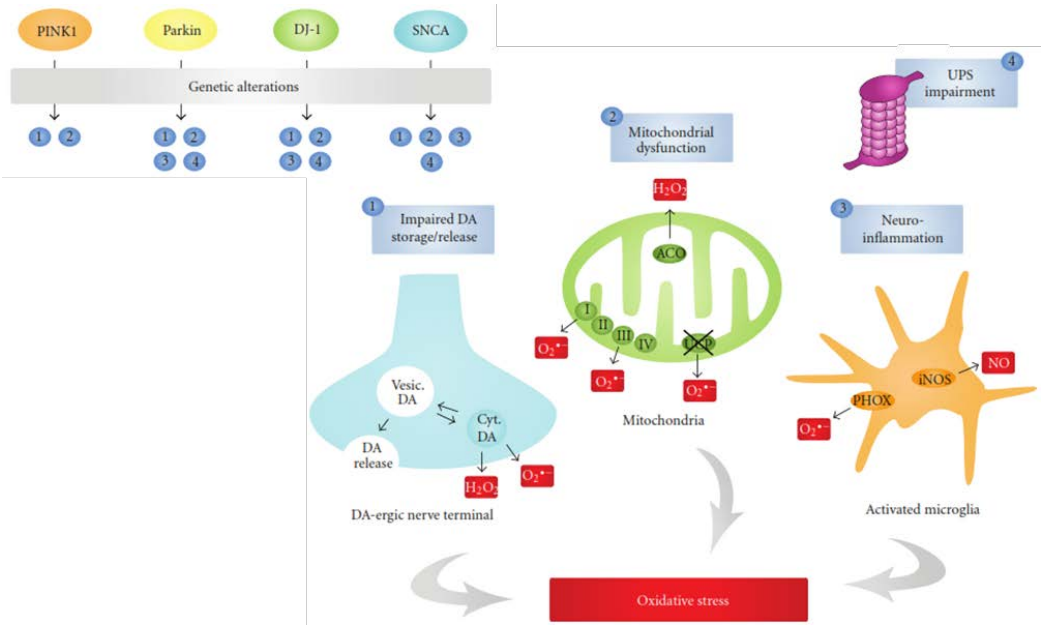


Figure 1.9 | PATHWAYS LEADING TO OXIDATIVE STRESS IN PD AND THE MODULATION BY PD-RELATED GENES. Different pathways contribute to high levels of oxidative stress in dopaminergic neurons, including impaired DA handling (1), mitochondrial dysfunction (2), and neuroinflammation (3). As further discussed in the text, alterations in PD genes can potentiate all these impairments and, therefore, lead to enhanced levels of oxidative stress. UPS dysfunction (4), another important pathogenic pathway in PD, can contribute to the damaging effects of α-syn, due to improper degradation of α-syn mutants or oligomers [52].

These cellular pathways are believed to be interconnected, since mitochondrial dysfunction leads to increased free radical generation, which further evokes deficits in the respiratory chain. Importantly, the UPS is dependent on oxidative phosphorylation for energy production and oxidatively damaged proteins increase the bulk of substrates to be degraded by the UPS. Moreover, this leads to increased cell dysfunction and a lowered threshold to apoptosis [53]. However, the exact sequence in which these events act and which points of interaction between events are key to the demise of SNpc dopaminergic neurons are still unknown [19].

Moreover, another unclear issue is whether the multiple cell death-related molecular pathways activated during PD neurodegeneration ultimately engage common downstream machinery, such as apoptosis, or remain highly divergent [19].

1.1.4. MITOCHONDRIAL DYSFUNCTION

It has been suggested that mitochondrial dysfunction may be an important player in loss of neurons and development of PD, and could be triggered by environmental factors alongside with certain gene variants and typical age-related loss of mitochondrial function [26]. This assumption is supported by several lines of evidence such as: (i) post mortem biochemical studies have demonstrated decreased complex I activity, a major component of the electron transport chain, in the SNpc of PD brains compared to controls; (ii) inhibition of complex I with rotenone, paraquat or MPTP reproduces the pathological degeneration of the SNpc in mammalian models; and (iii) the protein products of genes known to cause familial forms of parkinsonism have clear roles in the maintenance of functional mitochondria [44]. Furthermore, it was observed a 20-35% decrease in complex I activity in several cases of PD [5, 54].

Inhibition of complex I contribute to two main consequences to neuronal degeneration in PD through the decrease of ATP synthesis and consequently impairment of all ATP dependent cellular processes, and excess production of free radicals (ROS) that cause oxidative stress [55]. Inversely, higher oxidative stress can induce anomalies in mitochondrial function [54]. Moreover, mitochondria have also an important role in regulation and mediating cell death pathways and are involved in calcium homeostasis in cells [56].

Interestingly, many of the genes associated with familial forms of PD are also associated to mitochondria in disease pathogenesis. From the several genes involved in the pathogenesis of PD, *SNCA*, *parkin*, *DJ-1*, *PINK1*, *LRRK2* and high temperature requirement A2 (*HTRA2*) directly or indirectly impact mitochondrial function (Figure 1.10) [57].

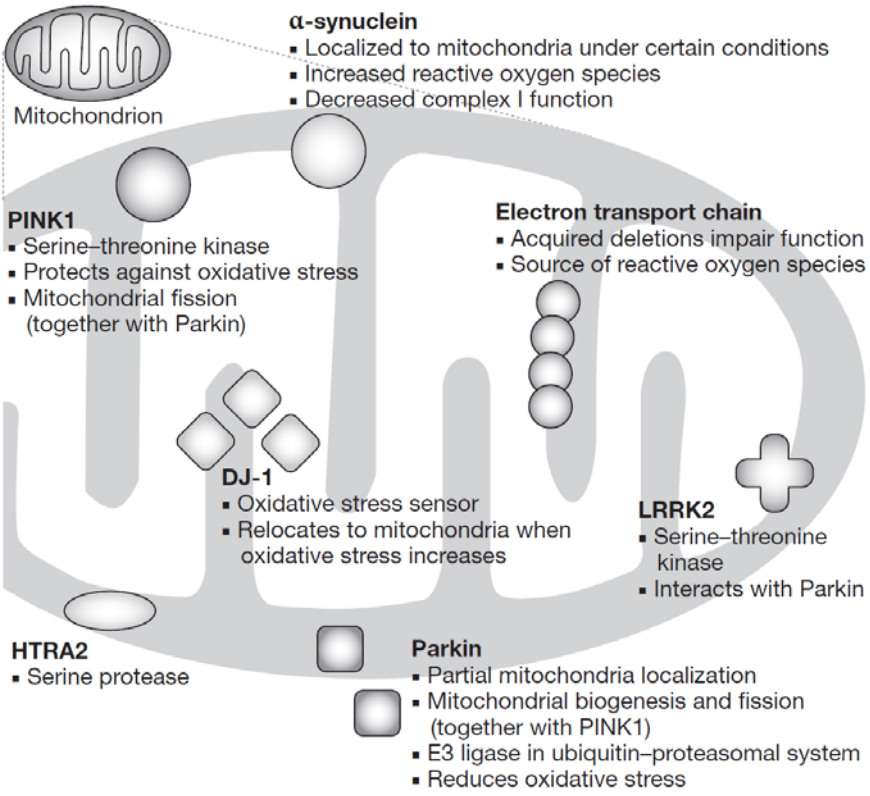


Figure 1.10 | PRODUCTS OF PD-ASSOCIATED GENES THAT AFFECT MITOCHONDRIAL FUNCTION AND OXIDATIVE STRESS. Acquired somatic mutations affect mitochondrial electron transport chain function, and such mutations are increased in the substantia nigra in patients with PD. Rare inherited mutations in genes encoding electron transport chain components have been associated with parkinsonism. Parkin, α-synuclein, PINK1, DJ-1, LRRK2 and HTRA2, are all encoded by nuclear genes, mutations in which can lead to PD, and all show a degree of localization to the mitochondria. Parkin is partially localized to the outer mitochondrial membrane, protects against oxidative stress, and has a hypothesized role in mitochondrial biogenesis. LRRK2 associates, at least in part, with the outer mitochondrial membrane; its precise function in that location is unclear, but it is thought to interact with Parkin. HTRA2 is a mitochondrial serine protease, the release of which might be involved in apoptotic cell death. PINK1 is a mitochondrial serine–threonine kinase that affords protection against oxidative stress and acts with Parkin to regulate the balance of mitochondrial fission and fusion. DJ-1 is relocated to mitochondria under conditions of oxidative stress and is thought to be neuroprotective under such conditions. The α-synuclein protein has an amino-terminal mitochondrial targeting sequence and, when overexpressed or under conditions of acidification, is at least partially associated with the inner mitochondrial membrane, where it might cause direct damage [56].

The accumulation of point mutations and deletions in mitochondrial DNA (mtDNA) is another interesting possibility that could cause the mitochondrial defects and dysfunction observed in PD. Several studies have shown increased mtDNA deletions in brains of PD patients when compared to controls [7, 50]. The proteins encoded by mtDNA include subunits from all parts of the electron transport chain, therefore, point mutations or deletions in any of the genes that encode these proteins would alter complex I activity [5].

1.1.4.1. MITOPHAGY

Recently, mitochondrial physiology and autophagy in the central nervous system has also captured additional attention, since it could give new insights in PD pathogenesis [58]. Mitophagy, a pathway for selective degradation of mitochondria by autophagy, is evolving as an important pathway in PD and has been linked to this disorder through the study of its recessively inherited forms, involving *PINK1*, *parkin* and *DJ-1* [59]. Particularly, mitophagy is important in neurons because they are post-mitotic and rely heavily and continuously on high levels of ATP production [59]. Therefore, neuronal activity and survival is believe to be severely dependent on mitochondrial integrity and functionality [58, 60].

It has been known that the process of mitophagy is greatly regulated and conserved from yeast to mammals, nonetheless much remains to be learned about this degradation process [60]. It appears that mitophagy shares key regulatory steps with the macroautophagy pathway, while displaying distinct steps specific for mitochondrial removal [60]. Recent findings indicate that parkin and PINK1 are involved in regulation of mitochondrial dynamics, and quality through the modulation of mitophagy [50, 61]. In the absence of PINK1 or Parkin, cells often develop fragmented mitochondria. Parkin is important in the repair of damaged and mutated mtDNA and mitochondrial biogenesis, among others, whereas PINK1 regulate the mitochondrial membrane potential, cristae structure and calcium homeostasis in order to maintain mitochondrial integrity. In the mitophagy pathway, parkin is specifically recruited by PINK1 that selectively targets the mitochondria promoting the degradation and removal of dysfunctional mitochondria [50, 61]. This has particularly

interest for cells that lack PINK1, where mitochondrial defects are rescued by parkin overexpression [61]. As expected, mutations in *PINK1*, parkin and also DJ-1 have been associated with mitochondrial dysfunction, likely leading to dopaminergic neuron degeneration in PD [59, 60].

Concluding, excessive mitochondrial stress derived from exposure to environmental toxins, increased oxidative stress or defects in mtDNA, together with the incapability of the cell to remove damaged mitochondria through mitophagy, may contribute to PD pathogenesis [60].

1.1.5. OXIDATIVE STRESS

Oxidative stress is characterized by a redox imbalance between the generation of free radicals or other reactive species and antioxidant defenses [62]. Many cellular reactions use molecular oxygen for catalysis and energy production. These reactions in turn produce ROS such as superoxide anions and hydrogen peroxide (H_2O_2) which in excessive production is detrimental to the cell and can lead to cell death. Cells have antioxidants such as glutathione (GSH) and vitamin E to manage this situation [63]. Moreover, dopamine metabolism can also be a source of ROS in dopaminergic neurons since dopamine autoxidation produces dopamine quinones (DAQ) that support ROS formation [57].

The detection of lipid peroxidation, increased iron, and proteins oxidation in autopsy tissue from the brains of individuals with PD support the evidence of an oxidative stress state in the SNpc. In addition, reduced levels of GSH in the midbrain are consistent with an increase of free radical levels [55, 57]. Oxidative stress in PD has been suggested to be a consequence of complex I dysfunction [62] and on the other hand the formation of ROS seems to impair mitochondrial proteins, further impairing mitochondrial function [64]. The linking of mitochondrial dysfunction and increased oxidative stress is also predicted to interfere with the UPS function, leading to the accumulation of misfolded or damaged proteins and formation of LB [56].

The increase in oxidative stress together with the decline in endogenous antioxidants is an important underlying risk factors for older people to develop neurodegenerative diseases including Alzheimer's and PD [63].

1.1.6. MISFOLDING AND PROTEIN AGGREGATION

The abnormal accumulation of misfolded or aggregated proteins within neurons has long been recognized a central pathogenetic mechanism associated with neurodegenerative disorders, including PD [50]. These aberrantly folded proteins were first thought to be pathogenic, affecting several cell functions. However, recent evidence has suggested that these aggregates may be a defensive measure by the cell to separate these aberrant proteins, therefore reducing the amount of toxic soluble misfolded proteins in cells. Indeed, reduction of toxicity is associated with increased accumulation of protein aggregates in healthy neurons, indicating that the appearance of such structures is a response to pathological stress [55, 65].

Specifically in PD, significant evidence came from the finding that LB are mainly constituted by aggregated α-synuclein, which in its native state is a soluble and unfolded protein [7, 50]. This protein is very sensitive to environmental stimuli and is a target of oxidative modifications induced by its interaction with ROS and reactive nitrogen species leading to a high propensity to aggregate [7].

Usually misfolded proteins are cytotoxic and there is a need to remove them from cells through cellular mechanisms for quality control, namely protein degradation systems, such as molecular chaperones, the UPS, and autophagy/lysosomal degradation [65]. As expected, under pathological conditions an impairment of these degradation systems can occur, resulting in the accumulation of aberrant and toxic proteins within or outside cells [55, 65]. Regarding the UPS system, parkin (an ubiquitin E3-ligase that ubiquitinates protein targets for UPS-mediated degradation) contributes to the removal of misfolded and damaging proteins, and the UCHL1 (a brain-specific ubiquitin C-terminal hydrolase) is involved precisely in α-synuclein degradation [7, 50]. The proteins LRRK2 and PINK1 are related to autophagy, the latter being known to directly activate autophagy and to interact with other autophagic proteins, and also ATP13A2 and GBA gene encode for lysosomal enzymes [50]. Several mutations (in molecular chaperones or UPS-associated enzymes), oxidative stress, mitochondrial dysfunction and ageing have been associated with the protein degradation systems dysfunction. Subsequently, dysfunction of protein degradation has arisen as an key contributor to nigral neuronal death and pathophysiology of PD [19, 64, 65].

1.2. **DJ-1**

DJ-1 is highly conserved throughout diverse species, and was initially identified as an oncogene whose transforming ability is enhanced by H-ras, with possible roles in sperm maturation and fertilization [46]. Furthermore, it was observed an increased DJ-1 expression in several types of cancer and also DJ-1 as a circulating tumor antigen in breast cancer [46, 66].

DJ-1 belongs to the Thi/PfpI/GAT protein superfamily and is ubiquitously expressed in many cell types, including the brain. It mainly localizes in the cytosol, but it can also be found in the nucleus and in some specific conditions in the mitochondria [42, 66]. In the human brain DJ-1 is predominantly expressed in astrocytes and astrocytic processes, and in a much less extent in neurons, which can be explained by the restrict antibody access within neurons due to the presence of DJ-1 binding partners [67]. These evidences suggest that DJ-1 might be involved in the cross-talk between glia and neurons [67].

1.2.1. STRUCTURAL BIOLOGY

The study and determination of DJ-1 crystal structure [68, 69] provided significant molecular insights regarding the function of this redox-reactive protein [46]. DJ-1 protein is a homodimer both in crystal and in solution that comprises 8 α-helices (α1-α8) and 11 β-strands (β1-β11) [45, 66]. This protein has a flavodoxin-like fold adopting a helix-strand-helix sandwich structure. The dimer interface is formed by close contact of several α-helices of each monomer. These interactions occur mainly through hydrophobic contacts, but hydrogen bonds and ionic interactions can also take part, forming a hydrophobic core between the two monomers (Figure 1.11) [66].

The superfamily to which DJ-1 belongs is characterized by the archetypical bacterial ThiJ and PfpI structures, but despite the remarkable structural similarities between DJ-1 and the bacterial PfpI/PH1704 proteases, there are two features that are not conserved: (i) DJ-1 possesses a distorted catalytical triad (Cys106-His126-Glu18) which is vital for protease activity and (ii) the additional C-terminal helix that DJ-1 has blocks the putative active site of this protein.

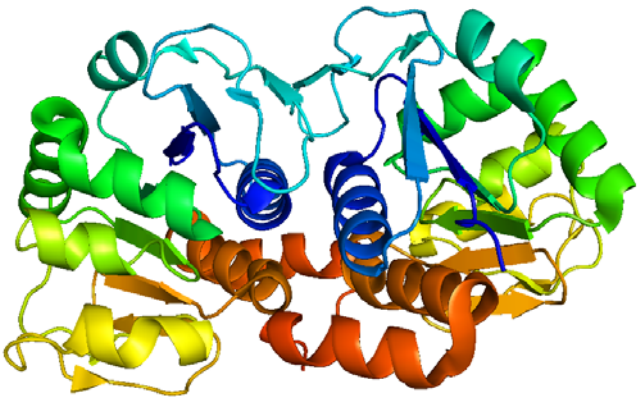


Figure 1.11 | **RIBBON DIAGRAM OF HUMAN DJ-1 DIMER.** The structure is colored from blue (N-terminal) to red (C-terminal) (PDB entry 1UCF, [68]). Image was created using PyMOL (DeLano Scientific; http://www.pymol.org).

Nonetheless, upon removal of the C-terminal helix DJ-1 shows proteolytic activity [70]. Furthermore, DJ-1 quaternary structure differs from that of other members of this superfamily therefore DJ-1 lacks protease activity [46].

Despite the differences in the configuration of the catalytical triad residues E18, C106, and H126, DJ-1 retains the highly conserved cysteine residue 106 (C106) and in the close proximity there are also the side chains of Glu-18 and His-126 (Figure 1.12) [46, 66]. The C106 is the best conserved among the three cysteine residues that human DJ-1 contains (C46, C53 and C106) and resides at the sharp turn between a β-strand and an α-helix called the “nucleophile elbow” [45].

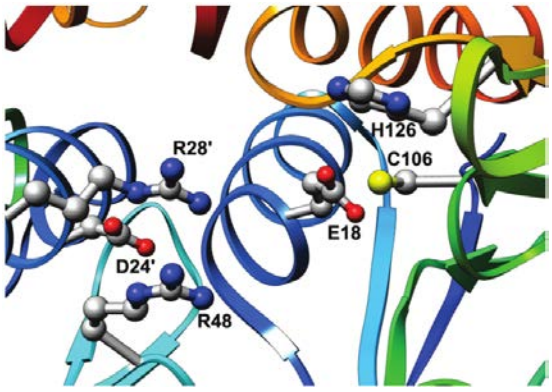


Figure 1.12 | RIBBON DIAGRAM OF HUMAN DJ-1 SHOWING THE LOCATION AND CONFIGURATION OF THE CATALYTIC TRIAD. The side chain of highly conserved C106 is located near the side chains of His-126 and buried Glu-18. Residues Asp-24’ and Arg-28’ from the opposite subunit contribute to stabilizing interactions across the subunit-subunit interface in the dimer [66].

This residue is high oxidizable and is believed to be involved in DJ-1 activation. In this process the side chain of Glu18 is thought to stabilize the oxidized C106 sulfinic acid, which is considered the active form of DJ-1 [46]. Nonetheless, the orientation that these residues display is not the appropriated for proton transfer which is characteristic of the cysteine protease catalysis [69].

It seems that most of the PD-causing mutations (L166P, M26I, E163K, D149A) form heterodimers with WT DJ-1. Furthermore, DJ-1 mutants are unstable, frequently not correctly folded and rapidly degraded in cells [10]. For instance, the L166P mutant displays severe structural disruption leading to an intense destabilization and thus loss of DJ-1 neuroprotective function [46]. Commonly, the others DJ-1 mutants also lead to loss of DJ-1 neuroprotective function but most of them present only slight structural perturbations compared to WT DJ-1. This suggest that even the slight structural defects in DJ-1 structure can cause DJ-1 loss of function, but the biophysical basis of the pathogenicity of these mutants and how this conformational changes affect the C106 oxidation and its cytoprotective effect remains unclear [45, 46].

1.2.2. CELLULAR FUNCTION

DJ-1 is a multifunctional protein with several proposed biochemical and cellular activities. It is involved in cellular processes like oxidative stress, cellular transformation, androgen-receptor signaling, spermatogenesis and fertilization [71]. Among the multiple functions ascribed for this protein, it is important to highlight the function as transcriptional co-regulator, molecular chaperone, antioxidant, RNA-binding, mitochondria protection against oxidative stress [67, 72] and recently discover glyoxalase activity [73]. A connection between DJ-1 and astrocytes has been reported that could lead to increase of neuronal susceptibility to oxidative stress. Studies indicate that astrocyte dysfunction can be triggered by DJ-1 deficiency thereby neuronal protection and secretion of soluble protective factors by astrocytes is impaired [74]. Furthermore, DJ-1 is involved in cell signaling pathways like the PTEN/Akt pathway and interacts with the apoptosis-implicated proteins Daxx, apoptosis signal regulating kinase 1 (ASK1) and p53 (Figure 1.13) [45]. Due to the

complexity and diverse roles that DJ-1 plays in cells the exact function of this protein is still to be determined despite the known neuronal cytoprotective effect.

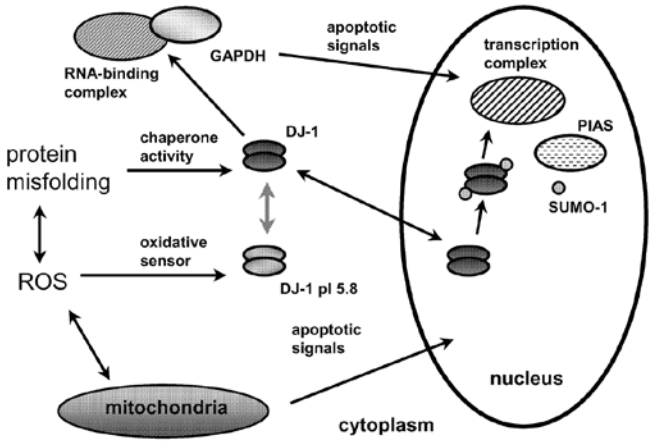


Figure 1.13 | MODEL OF DJ-1 AS A STRESS-RESPONSIVE, MULTIFUNCTIONAL PROTEIN. DJ-1 has multiple functions and some of the most important are chaperone activity, oxidative stress sensor, modulation of expression of genes (interacting with PIAS and cytosolic RNA-binding protein complexes, and other nuclear cofactor) and an association with GAPDH [43].

DJ-1 is a member of the ThiJ/PfpI family, as mentioned above, which is associated with functions like protease and chaperone activity. Additionally, the analogy between DJ-1 and Hsp31 (a molecular chaperone in *Escherichia coli*), and several observations that DJ-1 prevents aggregation of citrate synthase, luciferase and the PD-linked protein α-synuclein led to the identification of the DJ-1 chaperone activity [66, 71]. DJ-1 chaperone activity is activated in an oxidative environment and is dependent on oxidation state of DJ-1 at residue C106. Furthermore, for robust chaperone activity the mildly oxidized form of DJ-1 (Cys106-SO_2^-) is required [45, 75].

In addition, several studies suggest that DJ-1 protects neurons through an antioxidant function [6, 9]. Upon oxidative stress conditions DJ-1 shifts its pI from 6.2 to 5.8 and is converted to more acidic isoforms, specifically through the oxidation of C106 to cysteine sulfinic acid (Cys-SO_2H). Also, DJ-1 has the capability to undergo self-oxidation in order to eliminate H_2O_2. These evidences and the fact that C106 oxidation necessarily consumes ROS lead to the proposed role of DJ-1 acting as scavenger of ROS [6, 9, 45]. The DJ-1 antioxidant function can be dependent on its translocation to the mitochondria and the C106 oxidation seems to be part of an oxidative stress sensing mechanism that promotes the DJ-1 localization to mitochondria [6, 9].

Some studies also show that DJ-1 antioxidant action can be due to the regulation of antioxidant gene expression rather than by simply scavenging H_2O_2. In fact, it has been shown that DJ-1 stabilize the antioxidant transcriptional regulator Nrf2 (nuclear factor erythroid 2-related factor) by preventing association with its inhibitor. DJ-1 also serves as scaffold protein attenuating the inflammatory responses of astrocytes and microglia [76]; modulates the expression of genes such as glutamate–cysteine ligase, the rate-limiting enzyme of GSH biosynthesis; and regulates the expression of superoxide dismutase-1 and 3 [42, 72]. It was also reported that DJ-1 can modulate expression of several other genes such as androgen receptor reporter activity by interacting with PIASx; tyrosine hydroxylase and Bax [46, 72]. These studies and results propose that DJ-1 participates in numerous pathways in order to enhance cell survival in response to oxidative stress conditions.

Another area in which DJ-1 is involved is the mitochondria and its functions. Several lines of evidence suggest that DJ-1 is important in maintaining mitochondrial dynamics, integrity and fusion [26, 77]. Indeed, there are many studies that report partial localization of DJ-1 towards mitochondria and this re-localization is enhanced under oxidative stress stimuli [77]. The oxidative-induced modification of DJ-1 (at the C106) leads to DJ-1 mitochondrial translocation and it is believed that this step is necessary for the DJ-1 neuroprotective function which is abolished by certain mutations in the protein [26, 77]. Recent evidences suggest that DJ-1 is involved in the *PINK1/parkin* pathway in order to help maintain mitochondrial function, dynamics and integrity when an oxidative stimulus occurs [4].

Originally DJ-1 was described as a regulatory subunit (RS) which is a regulatory component of an RNA binding complex. Van der Brug and colleagues exposed the RNA binding function of DJ-1 that can associate with RNA targets in cells and the brain, such as mitochondrial genes, genes involved in GSH metabolism and members of the PTEN/PI3K cascade. They also discovered that this protein binds in an oxidation-dependent manner to several RNA targets since after oxidative stress the interaction between DJ-1 and RNA is disrupted. Thus, the hypothesis was that the postulated pleiotropic effects of DJ-1 could be related to the single function of binding to multiple mRNA transcripts [78].

1.2.3. ROLE IN NEUROPROTECTION

As mentioned before, in normal human brain DJ-1 is expressed both in astrocytes, and neurons although in the later in lower proportions [46]. Moreover, it was found that DJ-1 does not accumulate in α-synuclein-positive LBs yet oxidized and reduced DJ-1 protein levels were observed in PD brains [42, 46].

Several studies report the translocation of DJ-1 towards mitochondria after an oxidative stimulus, and this re-localization is believed to play a neuroprotective role in cells through mitochondrial autophagy regulation [51, 79]. Importantly, the C106 residue is also involved in neuroprotection mediated by DJ-1, since it plays a crucial role in DJ-1 acidic pI shift in oxidative stress conditions [80].

In model systems, DJ-1 confers protection to dopaminergic neurons against various insults, such as rotenone, mutant α-synuclein, H_2O_2 and 6-OHDA [42]. Several studies, including in *Drosophila, C. elegans* and mice that cells lacking DJ-1 are more sensitive to stress insults, which can be rescued by overexpression of DJ-1 WT but not by DJ-1 mutants [51, 81]. Interestingly, DJ-1 protects cells against low rotenone concentrations but at higher levels or prolonged stress the protective function of DJ-1 is abolished and apoptotic pathways are activated [51]. These evidences seem to mimic the corresponding ones in PD pathogenesis, thus provide a very helpful way to study pathogenesis PD associated to DJ-1.

Among the several functions attributed for DJ-1, the putative role in neuroprotection is the most relevant and it has been mainly linked to oxidative stress response. However, the mechanisms by which DJ-1 plays its neuroprotective roles are far from being elucidated and understood [81].

1.2.4. DJ-1 MUTATIONS

DJ-1 has been receiving more attention in the last years due to its involvement in familial early onset PD and also due to a large number of mutations identified in this gene [82, 83]. Mutations in DJ-1 are rare, varying between populations and ethnic groups and account for up to 1-2% of the sporadic cases with early-onset [83]. However, the study of these DJ-1 variations provides insights into the molecular and cellular mechanisms involved in the etiology of this disorder [46, 75]. PD-linked

mutations associated to DJ-1/PARK7 gene comprise CNVs (exonic deletions and truncations), homozygous (L10P, M26I, E64D, E163K and L166P) and heterozygous (A39S, A104T and D149A) missense mutations and also rare polymorphisms (R98Q, A171S), among others (Table 1.2) [84].

Table 1.2 | **DJ-1 GENETIC VARIANTS**.

Mutation	Inheritance	Population	Effect	Ref.
14-kbdeletion	Homozygous	Dutch	Loss of protein	[46]
L166P	Homozygous	Italian	Protein instability	[46]
M26I	Homozygous	Ashkenazi Jewish	Decreased stability	[46]
D149A	Heterozygous	Afro Caribbean	Unknown	[46]
IVS6-1G-C	Heterozygous	Hispanic	Altered transcript	[46]
c.56delC	Heterozygous	Hispanic	Frame shift	[46]
c.57G→A	Both	Global	Polymorphism	[46]
A104T	Heterozygous	Latino	Unknown	[46]
Ex5-7del	Heterozygous	Northern Italian	Altered transcript	[46]
IVS5+2-12del	Heterozygous	Russian	Altered transcript	[46]
R98Q	Heterozygous	Global	Polymorphism	[46]
E64D	Homozygous	Turkish	Unknown	[46]
E163K	Compound	Italian	Altered activity	[46]
g.168_185dup	Homozygous		Unknown	[46]
P158del	Homozygous	Dutch	Unknown	[46]
A179T	Heterozygous	Dutch	Unknown	[46]
Ex1–5dup	Heterozygous	Dutch	Unknown	[46]
A39S	Heterozygous	Chinese	Protein instability	[85]
A167A	Both	Global	Unknown	[86]
L10P	Homozygous	Chinese	Unknown	[87]
A171S	Heterozygous	American	Polymorphism	[88]

Both – homozygous and heterozygous; **Compound –**comprises two different mutations in the paternal and maternal alleles

Due to the recessive inheritance of DJ-1 associated PD, it is assumed that disease development is due to loss of function and therefore point mutations are expected to trigger loss of function of the protein, the presumed cause of their pathogenicity [89]. Identification of novel disease-linked missense mutations would help the discovery and analysis of functionally important domains in the DJ-1 protein and furthermore in understanding the role and mechanisms of DJ-1 in PD [43].

In recent years, remarkable developments have been accomplished in the characterization of some pathological DJ-1 mutations, however structural explanations for the observed deleterious effects of these missense mutations is still missing [90].

In the following sections three missense mutations that are believed to be the more studied and prominent will be analyzed and described: the L166P, M26I and

E163K mutations (Figure 1.14). Moreover, a topic related to the C106 engineered mutations will also be discussed.

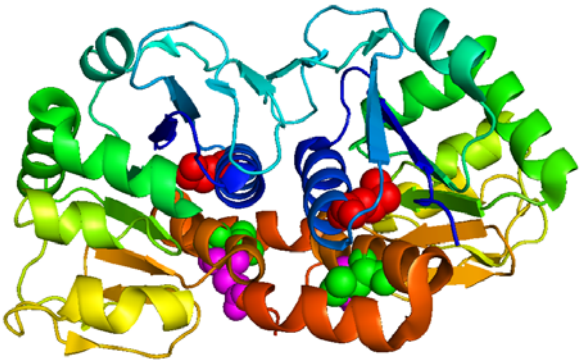


Figure 1.14 | **RIBBON DIAGRAM OF HUMAN DJ-1 DIMER.** The structure is colored from blue (N-terminal) to red (C-terminal) (PDB entry 1UCF, [68]). The L166 residue is represented by green spheres, the M26 residue by red spheres and the E163 residue by magenta spheres. The L166P substitution is located between two helices that form part of the dimer interface. Image was created using PyMOL (DeLano Scientific; http://www.pymol.org).

1.2.4.1. L166P

The substitution of the highly conserved leucine at position 166 for a proline (L166P) is one of the most extensively investigated and deleterious missense mutation in DJ-1 [66, 90]. This mutation was first reported for an Italian family and is located in the center of α-helix 7 (the second to last C-terminal helix), near the dimer interface, which together with the α-helix 8 mediates dimerization of DJ-1 (Figure 1.15) [91, 92].

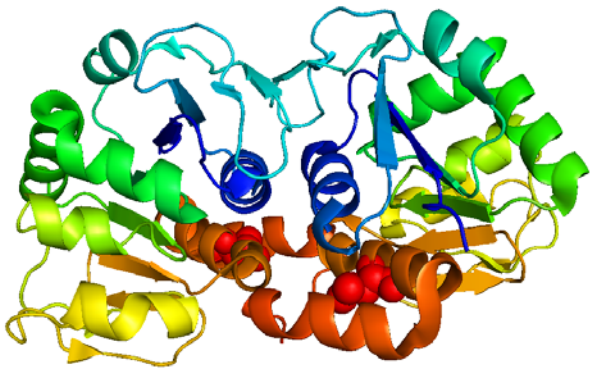


Figure 1.15 | **RIBBON DIAGRAM OF HUMAN DJ-1 DIMER WITH THE L166 POSITION HIGHLIGHTED.** The structure is colored from blue (N-terminal) to red (C-terminal) (PDB entry 1UCF, [68]). The L166 residue is represented by red spheres. The L166P substitution is located between two helices that form part of the dimer interface. Image was created using PyMOL (DeLano Scientific; http://www.pymol.org).

It was proposed that this substitution to proline acts as a helix breaker, disturbing the packing interactions at the dimer interface and consequently, leading to the disruption of the dimer which is associated with an increase in the total hydrophobic surface area exposed to solvent, probably explaining the propensity of this mutant to aggregate (Figure 1.16) [66, 92, 93].

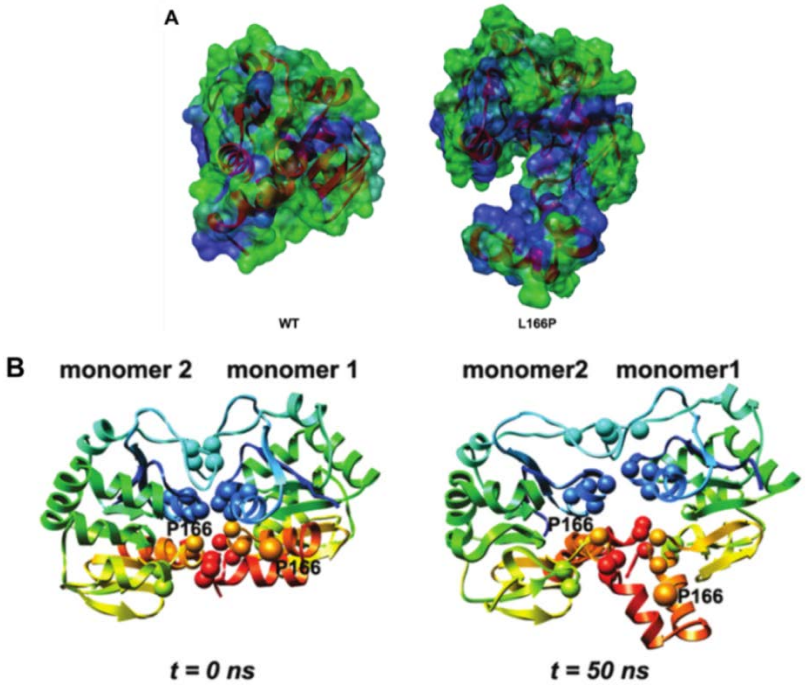


Figure 1.16 | STRUCTURAL PERTURBATIONS CAUSED BY L166P MUTANT REGARDING THE DIMERIC STRUCTURE AND HYDROPHOBIC SURFACE. A) Increase in hydrophobic solvent accessible surface areas (SASA) of L166P DJ-1 (right) relative to WT DJ-1 (left). Molecular surface is colored by residue from least hydrophobic (green) to most hydrophobic (blue). B) Perturbations in the DJ-1 dimer interface following L166P substitution in each subunit. Loss of intermolecular contacts that mediate DJ-1 dimerization after 50 ns of L166P molecular dynamics simulations. The key hydrophobic interactions across the interface are represented by spheres in both subunits. Residue 166 positions are indicated by orange spheres[66].

The unfolded structure of L166P exhibits an impaired ability to form stable homodimer with itself or a heterodimer with DJ-1 WT however is capable of forming high molecular-weight complexes containing DJ-1 oligomers or aggregates with other proteins [66]. Due to the structural perturbations, the L166P mutant protein is significantly unstable in mammalian cells [93], probably due to direct proteasomal endoproteolytic cleavage not requiring prior ubiquitination for its degradation [84]. This lead to expression of L166P at lower steady-state protein levels compared to the WT protein and an accelerated turnover rate [92].

As a consequence of impaired dimerization and reduced stability the L166P mutation abolish DJ-1 chaperone and protease activity of DJ-1 protein [90, 92]. Furthermore, this substitution also led to major perturbations in the region of the C106 residue, highly conserved, that is involved in dimerization and critical for the putative chaperone function of DJ-1, suggesting that the structural effects from this mutation are not confined to the vicinity of the substitution, but spread quickly throughout the entire protein [66].

Additionally, this mutant fail to protect cells from death induced by H_2O_2 and in removing it when in cells [75], and under stress conditions the L166P protein is only partially mislocalizes or sequestered to mitochondria [82, 91, 94].

1.2.4.2. M26I

The second most studied mutation in DJ-1 is the conservative amino acid substitution of the methionine in position 26 for an isoleucine (M26I)[95, 96]. This methionine residue is highly conserved throughout evolution, suggesting that it could be critical for the structure and function of DJ-1 and is located on the C-terminal end of α-helix A (or 1), causing packing defects in the core of the dimer (Figure 1.17) [6, 90]. This mutation was first identified in an Ashkenazi Jewish patient with early-onset PD [95].

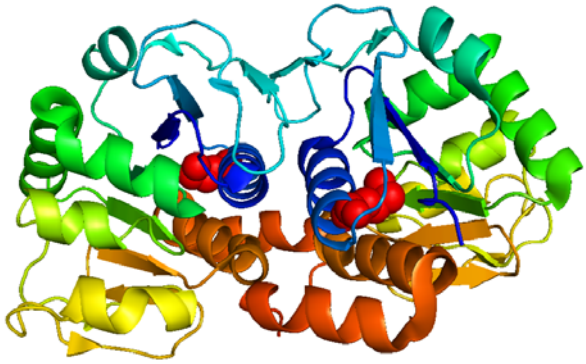


Figure 1.17 | RIBBON DIAGRAM OF HUMAN DJ-1 DIMER WITH THE M26 POSITION HIGHLIGHTED. The structure is colored from blue (N-terminal) to red (C-terminal) (PDB entry 1UCF, [68]). The M26 residue is represented by red spheres. The M26I substitution is located in the core of the DJ-1 dimer. Image was created using PyMOL (DeLano Scientific; http://www.pymol.org).

Contradictory data is reported regarding the stability of this mutant. Some studies refer that M26I is stable in cells and not subjected to proteasomal degradation

[84, 89] where others claim that this mutant form of DJ-1 is unstable and produced in low steady state levels in cell cultures [6, 93]. On the other hand, several studies refer the ability of M26I mutant to adopt the dimeric DJ-1 WT fold as well as to form heterodimers with the WT protein [75, 93]. Furthermore, the M26I substitution leads to decreased thermal stability and an enhanced propensity to form high molecular weight oligomers [6]. This pathogenic mutation can localize in the cytoplasm or mitochondria in slightly varying degrees compared to the DJ-1 WT, but the localization in the latter compartment is enhanced by oxidative conditions [93].

Moreover, studies show that M26I mutant aberrantly bound to ASK1, failing to suppress ASK1 activity and nuclear export of the death domain-associated protein Daxx, thereby leading to loss of cytoprotective activity [97]. Also, some reported approaches can help to stabilize the M26I mutant protein. A previous study revealed that heterodimer formation between M26I mutant protein and DJ-1 WT could help to stabilize the mutant protein and this could play a central role in preserving DJ-1 function in heterozygous individuals [6]. Another study shows that introduction of a disulfide linkage between residue Val 51 and Cys 53 stabilized the mutant M26I against thermal denaturation, improved their ability to scavenge ROS, and restored a chaperone-like function of blocking α-synuclein aggregation [79].

1.2.4.3. E163K

The E163K missense mutation in exon 7 and is characterized by the substitution of the glutamic acid residue in the position 163 for a lysine residue. This is a very non-conservative exchange of a strictly conserved acidic amino acid residue within the structurally essential helix G (or 7) (Figure 1.18) [46, 98]. This mutation was first reported for a family in southern Italy, resulting in a severe phenotype as early as 24 years of age. Subjects homozygous for this E163K mutation develop symptoms that include parkinsonism, dementia, and amyotrophic lateral sclerosis [99].

The E163K pathogenic mutant retains similar properties to DJ-1 WT protein, under normal conditions, in terms of solubility and dimerization, and only subtle perturbations in the DJ-1 structure [90, 99]. Nonetheless, this mutant compromises the ability of DJ-1 to protect against oxidative stress while retaining the ability to protect

the cells against mitochondrial stress and proteasome inhibition [99]. This means that the effect of this mutation on impairing DJ-1 function is specific to oxidative stress, in contrast to the other mutations in DJ-1 [99]. Another important particularity of this mutant is that it causes the disruption of a structurally critical salt bridge with the arginine residue 145 (R145), which leads to an increased mobility of R145 and consequently interferes with a network of hydrogen bonds that involves L186 and spans the dimer interface [90].

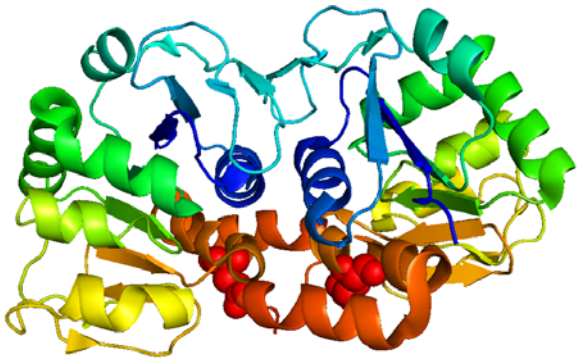


Figure 1.18 | RIBBON DIAGRAM OF HUMAN DJ-1 DIMER WITH THE E163 POSITION HIGHLIGHTED. The structure is colored from blue (N-terminal) to red (C-terminal) (PDB entry 1UCF, [68]). The residue E163 is represented by red spheres. The E163K substitution is located in the DJ-1 surface. Image was created using PyMOL (DeLano Scientific; http://www.pymol.org).

Under oxidative stress conditions the E163K mutant shows a lack of localization towards mitochondria. It is though that the change from a negatively to positively charged residue in addition to the subtle structural effects may prevent the interaction of DJ-1 with other proteins that may be involved in the redistribution of DJ-1 under conditions of oxidative stress [99].

The features described above reveal that the loss of DJ-1 protective function can occur without evident biochemical changes on the protein and demonstrate that alterations to specific residues can affect precisely individual functions indicating that probably DJ-1 is involved in multiple cellular pathways [99].

1.2.4.4. CYS 106 ENGINEERED MUTANTS

There are specific residues that are very important and highly conserved and C106 is one of those residues. Cysteine 106 is crucial for DJ-1 activity in

neuroprotection and replacement of this residue by other amino acids blocks the neuroprotective DJ-1 function [100].

Through direct oxygen addition, C106 forms 3 different species, specifically, cysteine-sulfenic acid (Cys-SOH), cysteine-sulfinic acid (Cys-SO_2H), and cysteine-sulfonic acid (Cys-SO_3H). In cells exposed to oxidative stress the pI of DJ-1 shifts to a more acidic value and C106 is oxidized to Cys-SO_2H or Cys-SO_3H [101, 102]. Currently, existing data suggests that mild oxidation to Cys106-SO_2^- is associated with DJ-1 cytoprotection whereas extensive oxidation of DJ-1 is associated with aging and neurodegenerative disease (Figure 1.19) [45].

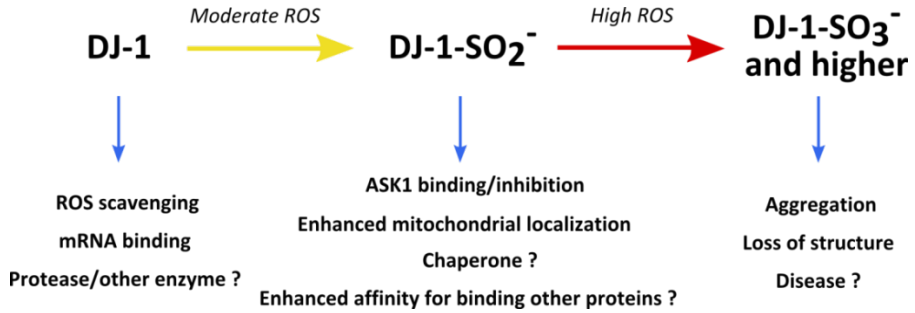


Figure 1.19 | ROLE OF MAJOR OXIDIZED DJ-1 FORMS. The question mark is used to indicate unresolved aspects of the current set of proposals for DJ-1 function [45].

Moreover, the partially oxidized DJ-1 isoform where C106 is converted to the sulfinic acid form is believed to trigger DJ-1 chaperone activity and suppress α-synuclein fibrillization, while further oxidation of the protein is thought to lead to partial loss of protein structure and to abolish this chaperone function [6, 103]. Additionally, oxidation of C106 to Cys106-SO_2^- necessarily consumes ROS and therefore it has been suggested that DJ-1 acts as direct scavenger of ROS [45]. To validate this statement the engineered mutant C106DD (substitution of C106 for two aspartic acids) should be nonfunctional for a putative C106 dependent ROS scavenging role in DJ-1, but that is not the case. Hence, current data suggest that this effect is not the main contributor to the protective function of DJ-1, despite oxidation of C106 indeed results in some reduction of cellular ROS [45].

Cysteine residue 106 was also identified as the functionally critical amino acid that promotes direct binding of DJ-1 to ASK1 and thus cytoprotection in oxidatively stressed cells [97]. DJ-1 forms a mixed disulfide bonds involving the central C106 with the N-terminus of ASK1, similar to a known negative regulator, Trx1 and this could be

recapitulated by the C106DD mutation [45, 97]. On the other hand, the substitution of C106 for an alanine specifically abolished binding to ASK1 and the cytoprotective activity of DJ-1 [97].

Unlike C106A that impairs DJ-1 neuroprotective functions, C106DD mimics and activates the cytoprotective protein DJ-1 and consequently provides protection to cells against oxidative stress [45]. However, it is unknown whether the loss of DJ-1 protective function in these mutants is due to the absence of C106 oxidation or the absence of the cysteine residue itself [104].

Some residues around C106 are responsible for the modulation of C106 reactivity toward ROS. Among these residues the glutamate in position 18 (E18) is essential to the stabilization of Cys-106-SO_2^- residue through the formation of a hydrogen bond with it [104]. Thereby mutations in Glu-18 to other residues affected the ability of C106 to oxidize to sulfinic acid without changing the cysteine residue itself [104]. Furthermore, under oxidative conditions, DJ-1 WT translocates to mitochondria and this recruitment is closely related with its susceptibility for oxidation, thus C106A abrogate this feature of DJ-1 and is not relocated in this way [80].

1.2.5. ***DJ-1*** ***INTERACTOME***

Protein interaction networks are crucial to acquire knowledge about proteins´ functions and characterization, disease mechanisms, and in some cases it could lead to the discovery of novel targets for therapeutic drugs [105]. Proteins can interact with several types of molecules (other proteins, nucleic acids, carbohydrates, and lipids among others) in order to modulate biochemical reactions and/or mediate numerous stimuli towards signaling pathways [106]. Interactome can be referred as a network of all types of molecular interactions, but protein-protein interactions (PPI) have a fundamental role in cellular systems. In order to function properly these interactions need to occur at the right time and location, and proteins must bind specifically with appropriate interacting partners [106].

Following this perspective, several studies have been performed with the purpose of discovering DJ-1 interacting partners (Table 1.3).

Table 1.3 | EXAMPLES OF DJ-1 INTERACTING PARTNERS REFERRED IN THE LITERATURE AND DATABASES.

Molecular Function Classification	Protein	Reference
Ubiquitin-protein ligase activity	Parkin	[11, 107]
	TRAF6	[108]
Dopamine release and transport	α-synuclein	[109, 110]
SUMO ligase activity	SUMO-1	[111, 112]
	SUMO- conjugating enzyme UBC9 (UBE2I)	[112]
	PIAS1	[111]
Protein binding	polyubiquitin-C	[113]
Phosphatase activity	PTEN	[114]
	SHP-1	[76]
Chromatin binding protein	MTA1	[115]
	MTA2	[115]
Kinase activity	PINK1	[107]
	MAP3K5	[97]
	ERK1/2	[116]
Chaperone Activity	HSPA4	[11]
	HSP70 and 90	[117]
	BAG1	[118]
	Mortalin	[119]
	grp94	[119]
	calnexin	[119]
Apoptosis regulator activity	BCL2L1	[77]
Tumor necrosis factor	TRAF and TNF receptor-associated protein	[108]
Ubiquitin thiolesterase activity	Cezanne/Za20d1	[120]
Histone deacetylase activity	HDAC1	[115]
	Histone-binding protein RBBP4	[115]
	Mi-2/NuRD complex	[115]
RNA binding	Nucleolin	[119]
Monoamine transmembrane transporter activity	VMAT2	[121]
Structural molecule activity	MAP1B	[122]
	clathrin	[119]
Superoxide dismutase activity	SOD1	[123]
Protein Kinase	ASK1	[124]
Transcriptional factor	p53	[125]
	DAXX	[112]
	C-myc	[126]
	DJBP	[127]
Dehydrogenase and nitrosylase activities (phosphorylating activity)	GAPDH	[126]
Proline biosynthesis activity	PYCR 1	[128]

Nowadays, there are already databases that aim to map protein-protein interaction networks, namely PPI networks associated with the human interactome. BioGRID, DIP, IntAct are some examples of PPI databases [129], where a list of

interactors of a target protein could be found and in some database a list of interactors of the target protein could be accessed.

As mentioned before, the function and protein interactors vary with the time, conditions and location in cells. In this way, some studies are focused on studying the interactome changes when the time and conditions are modified. An example of that is the work done by Sandra Anjo during her master thesis project (unpublished data) in which the endogenous DJ-1 interactome was studied under normal and stress conditions for different time points of oxidative stress exposure.

Genetic mutations associated with genetically inherited disorders doesn't cause the disease only due to the malfunction of the mutated gene product, but rather through the perturbations in the interaction network in which that altered protein is inserted [129]. Thus, understanding and revealing the complex interaction network in which the disease-related proteins are linked is essential to understand the network perturbations associated with the disease phenotype [129, 130]. From this standpoint the study of DJ-1 mutations dynamic interactome may generate significant information about the modifications in the DJ-1 interaction network caused by these mutations and furthermore in understanding the PD molecular mechanisms and phenotype. For some mutations interesting interacting partners have been identified [11, 82, 96, 97, 108], but a full interactome analysis under different conditions is still lacking.

1.3. INTERACTOMICS

For several years, the genomic field has been generating a huge amount of data concerning genetic sequences with special highlight to the human genome project. However, currently the challenge has shifted to the study of the proteins´ functions encoded by genes identified in the human genome project [131]. The proteome has dynamic properties compared to the static nature of the genome. Therefore, the focus of proteomics is the study of protein structures, modifications and localization, and protein–protein interactions among them [131].

Frequently, in several cellular processes proteins fulfill their functions within a complex network of interactions that could be stable or dynamic and where a single component can affect a wide range of other components [132, 133]. Thus, PPI are

crucial to biological processes and the characterization of PPI networks, called interactomes, are important to translate molecular mechanisms into cells functions and physiology [132, 134]. PPIs can also help to elucidate the functions of newly discovered interacting proteins or new functions of proteins already identified [132, 133]. Importantly, PPIs are frequently regulated in response to a specific stimulus or cell condition, where different interactomes vary according to tissue, cell type, and biological state [132, 135].

It was observed that a single protein, under different conditions, can interact with several different partners leading to different biological outcomes. Therefore, PPI mapping is particularly important for many diseases helping to understand their molecular basis and pathogenesis that could lead to the identification of new disease markers [129, 136, 137]. Disruption of PPIs is also associated with disease development since many cellular processes are regulated by multimeric protein complexes. Hence, a new type of drugs that influence PPIs regulation by blocking the formation of functional protein complexes instead of affecting individual target proteins could be developed [132]. In fact, to date, some interaction studies related to PD-associated proteins have been made, particularly in α-synuclein and LRRK2 proteins [137, 138].

Several methodologies can be used to map PPI, with major use of two approaches: the yeast two-hybrid system (Y2H) and the affinity purification coupled to tandem mass spectrometry (AP–MS/MS) [130, 139]. The Y2H system was used to obtain the first interactome maps and is based in a binary assay that captures direct PPI, whereas AP-MS identifies and maps direct and indirect protein associations within protein complexes [129, 134]. In the first approach a transcription factor can be separated into two parts, a DNA binding domain (DBD) that is fused to a bait-protein (known protein used to capture other proteins) and an activation domain (AD) that is fused to a prey-protein (one or several proteins that interact with the bait-protein). The transcription factor is functional only when the bait and prey proteins interact physically which lead to the junction of DBD and AD. The functional transcription factor will induce transcriptional activation of a reporter gene and subsequently to a detectable and measurable outcome, which could be growth on selective conditions media or generation of a colorimetric or fluorescent signal [140, 141].

The advantages of Y2H system are the fact that is a scalable method and the speed in PPIs identification. However, this method has high false-positive and false-negative rates due to numerous aspects, including the fact that yeast is a heterologous system. Thus, AP-MS/MS has been an alternative method widely used [131, 135].

1.3.1. ***AP-MS/MS***

AP-MS/MS is a high-throughput and powerful technology, thus currently is the most common approach used to characterize protein complexes and analyze interactomes [142, 143] (Figure 1.20). In this methodology, the protein of interest (bait) is made by the fusion of the protein with a specific tag (recombinant tagged proteins) that is immobilized in a solid support (usually agarose) or it is purified using antibodies bound to agarose or magnetic beads [131]. The next step is separation of the protein (bait) and its binding partners (preys) that is accomplished through cleavage in specific sites (between the bait protein and support) by a specific enzyme or using a soluble substrate that will compete for binding sites [144]. Then, the purified protein complexes can be analyzed in a gel based approach using 1D or 2D gels and the desired bands can be proteolyticly digested into small peptides that are subsequently identified by mass spectrometry (MS). Alternatively, a gel free based approach using sample digestion and direct analysis by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) can also be employed (Figure 1.20) [131, 145]. MS data generate a list of proteins that comprises the protein of interest (bait), its cellular binding partners and also some background contaminant proteins [135].

The biggest challenge of AP-MS method is to distinguish between true interaction partners and co-purifying contaminants (unspecific background of proteins) [143]. Two approaches to overcome this problem are tandem affinity purification (TAP) and quantitative AP-MS (q-AP-MS). In the former tagged bait proteins are purified by two-step purification procedure, reducing the amount of contaminants, while the later uses quantitative proteomics providing evidence to differentiate real interactions from false positives [135, 143].

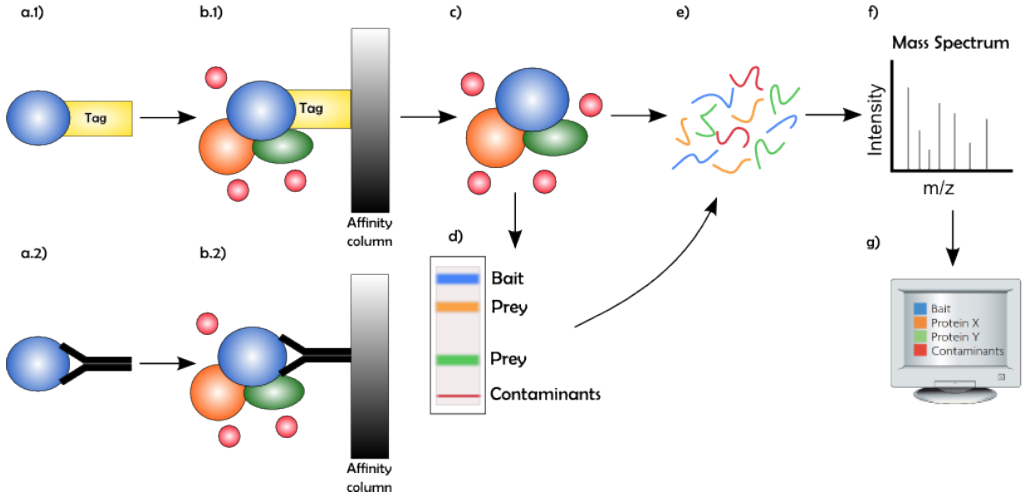


Figure 1.20 | OVERVIEW OF AN AFFINITY PURIFICATION COUPLED TO MS APPROACH. (a) Specific bait protein (blue) is tagged by specific methods (a.1) or has affinity to a specific antibody (a.2).(b) The expressed bait proteins with the affinity tag are purified from a cell lysate of the bait protein and binding partners (orange and green) using affinity chromatography (b.1) or beads with the specific antibody for the bait protein are used to purify the bait protein and binding partners (orange and green) from a cell lysate (b.2). (c) Purified protein complex and contaminants (red) that can also be present. (d) In an optional step, proteins in the complex can be separated by SDS–PAGE. (e) Proteins are subject to proteolysis (usually with trypsin). (f) MS analysis of peptides, involving in most cases LC-MS/MS that yields a tandem MS (MS/MS) spectrum (simplified MS/MS scan is shown). (g) Interpretation of the MS data and protein identification using database searching and statistical software which generate a list of the tagged protein, its interacting partners and contaminant [139, 146].

The TAP method involves the expression of a dual tag fused to the bait protein and the original TAP tag was a sequence stretch containing a calmodulin-binding peptide (CBP), a tobacco etch virus (TEV) protease cleavage site and the immunoglobulin G (IgG)-binding domain of protein A from *Staphylococcus aureus* (SpA), but through the years many variations of this strategy have been developed [143, 147]. This strategy has been commonly and successfully used in numerous protein interaction studies, increasing the efficiency of isolation of protein complexes [143]. The main advantages are the mild washing conditions that prevent the destruction of protein complexes and the near physiological conditions of the method [132]. However it has limitations in identifying very weak or transient interactions due to the long purification protocol that preserves only stable interactions being too extensive for this type of associations [142, 147].

The other approach developed to eliminate nonspecific interactors without losing abundant and transiently interacting proteins is q-AP-MS [129]. In this method two or more affinity purifications are made in parallel. In one of the affinity purifications, the bait protein is captured with an antibody or a specific/tagged ligand and the other one is planned in order to capture only background proteins or a different condition from the first one, therefore working as a negative control or as different experimental condition, respectively. After that the relative amount of each protein in those conditions must be determined and for that quantitative proteomics is employed. True interactors will appear only in one of the pull-downs and will have peaks with higher abundance compared to the false positives that should be equally present and appear in both pull-downs [143]. To help elucidating and validating the direct protein interactions bioinformatics tools will correlate MS data with other methods like Y2H, or interactive MS measurements in conjunction with chemical crosslinking [148].

Despite the limitations of AP-MS/MS, this technique has also many advantages such as the native cellular environment where the expression occurs, which allows the accurate regulation and post-translational modifications of the protein of interest in its cellular compartment [136, 139]. Moreover, when used in combination with quantitative proteomics techniques, this method can be used to prove dynamic changes in PPIs of protein complexes, providing information associated with cell signaling [135, 143].

1.3.2. ***MS***

Mass spectrometry has revolutionized the proteomics field becoming the method of choice to analyze and identify large numbers of proteins in complex protein samples [131, 148]. This technique measures the mass to charge ratios (m/z) of gas-phase ions from the sample [149]. In the last years, MS has played an increasingly significant role in the biological sciences and nowadays is used to characterize structurally several types of biomolecules [149]. It is characterized by extreme sensitive, speed, allows the identification of peptides present at femtomole levels, is

compatible with high-throughput strategies, is able to characterize peptide modifications, and can also quantitatively measure peptide abundance [139].

A mass spectrometer consists of an inlet device to introduce the analyte into the ion source (usually a separation technique such as liquid chromatograph or just a direct injection), an ion source to produce ions from the sample, one or more mass analyzers, to separate the ions according to their m/z ratios, a detector that registers the number of ions at each m/z value and a computer, to process the data and produce the mass spectrum [148, 149].

Analysis of proteins and peptides by MS was only possible with the development of soft ionization methods, like electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) that allow the transfer of large, polar, thermally labile biomolecules (like peptides and proteins) into the gaseous phase for mass analysis [131]. ESI ionizes the analytes out of a solution phase sample generating gaseous ions and therefore can be easily coupled to liquid-based separation technologies such as LC and capillary electrophoresis. MALDI sublimates and ionizes the samples out of a dry, crystalline matrix via laser pulses and is normally used to analyze simple peptide mixtures, while for the analysis of complex samples liquid-chromatography ESI-MS methods (LC-MS) are preferred [148, 149].

These advances led to the development of new mass analyzers and complex multistage devices in order to resolve the challenges of protein and proteome analysis. In some cases when additional information is needed, after the initial mass determination, specific ions are selected and subjected to fragmentation through collision in a methodology called tandem mass spectrometry (MS/MS) [150]. In the past few years, tandem MS has been widely used for high-throughput analysis of complex protein samples [151]. This method combines sequence and mass information in order to identify proteins (MS/MS). The first step is to measure the peptides masses, then isolate single peptides in the mass spectrometer and promote their collision with inert gas molecules (collision-induced dissociation (CID)), which will lead to fragmentation of the peptide backbone. Each MS–MS spectrum is associated with an amino acid sequence and in a second phase the data obtained is compared against one or more databases to infer the identity of the proteins in the sample [129, 151].

Files with thousands of MS and MS/MS spectra are generated when tandem MS is performed and there is the need to analyze all the information for further use in protein identification and characterization [151]. Usually, peptides are identified by the submission of the data to selected protein sequence databases [151]. In this methodology, each acquired MS/MS spectrum is compared with theoretical spectra obtained from a sequence database and each acquired MS/MS spectrum is thereby assigned the best matching database peptide [151]. According to the degree of similarity between the experimental measured data and the theoretical data there is a score that indicates the reliability of peptide identification [151].

Protein identification and characterization by itself does not give information regarding proteome dynamics that include significant cellular events such as protein synthesis and degradation, or the formation of protein assemblies [152]. Quantitative proteomics is based in MS quantification in the changes in protein abundances and is an important tool used to study protein interactions´ dynamics [153]. Thus, protein quantification in LC-MS methods is centered on the peak height or area of the proteolytic peptide peaks in the mass spectrum and/or chromatogram [152]. Different approaches have been developed to perform quantitative proteomics experiments and can be broadly divided into labeling with stable isotopes and label-free methods [153]. The former can be divided into metabolic labeling (SILAC) and chemical/enzymatic (ICAT, ICPL, and iTRAQ) whereas the latter in spectral counting and peptide ion chromatogram extraction [153, 154]. Labeling methods involve the incorporation of stable isotopes (artificial labeling of peptides or proteins) into one or more samples being studied while in label free methods the samples native isotope composition is maintained and identical peptides between samples can be compared [154].

In general terms, an MS-based proteomics analysis uses site-specific enzymatic proteases to digest proteins into peptides that are subsequently fragmented in the mass spectrometer, giving MS/MS spectra which are used to identify the peptides sequences through databases and moreover quantify through quantitative MS (Figure 1.21) [155].

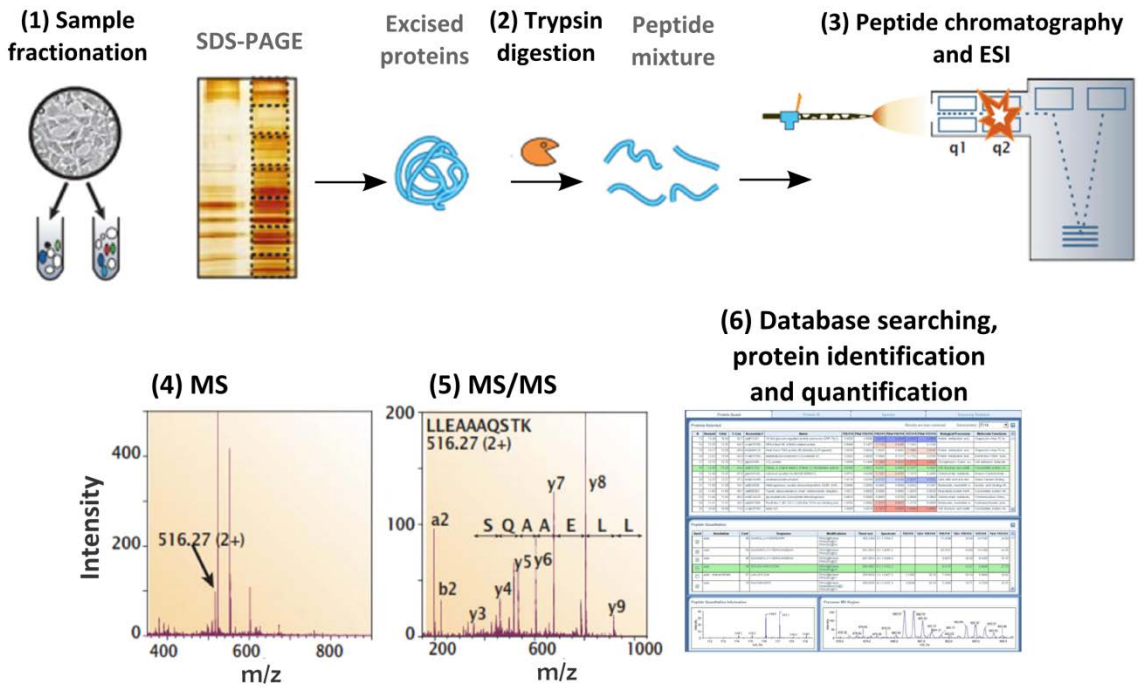


Figure 1.21 | GENERIC MS-BASED PROTEOMICS APPROACH. A classic proteomics approach comprises five stages. In stage 1, the proteins to be analyzed are isolated from cell lysate or tissues. This often includes a final step of one-dimensional gel electrophoresis. Proteins are degraded enzymatically to peptides in stage 2, usually by trypsin. In stage 3, the peptides are separated by high-pressure liquid chromatography and eluted into an electrospray ion source. In stage 4, a mass spectrum of the peptides eluting at this time point is taken (MS1 spectrum). The computer generates a prioritized list of these peptides for fragmentation and a series of tandem mass spectrometric or "MS/MS" experiments ensues (stage 5). The MS and MS/MS spectra are acquired and stored for matching against protein sequence databases. The outcome of the experiment is the identity of the peptides and therefore the proteins making up the purified protein population (stage 6) [148].

1.4. OBJECTIVES

The exact functions and mechanisms of DJ-1 remain largely unknown, however the involvement of this protein in PD and its fundamental role in neuroprotection is indisputable. DJ-1 neuroprotection in cells is believed to be associated with an oxidative stress response, with DJ-1 playing diverse functions.

Mutations associated to this gene lead to structural defects and loss of DJ-1 function which will cause an inherited form of PD. Identifying and clarifying the mechanisms of DJ-1 related disease has already provided insight into pathophysiology of PD and may promote the understanding of many other mechanisms of brain neuronal maintenance. Additionally, the identification and study of PD associated genes has a fundamental contribution in understanding and unravel the pathophysiologic mechanisms of common sporadic forms of this disease. Following these perceptions, this study aims to characterize the dynamic interactome of WT DJ-1 with the interactome of certain DJ-1 mutants, particularly E163K, C106A and C106DD mutants, under normal and stress conditions. These approaches will help to identify the changes in protein-protein interactions that DJ-1 undergoes when mutated and later on in understanding how this changes compromise DJ-1 functions and neuroprotection.

2. EXPERIMENTAL PROCEDURES

2.1. DJ-1 CONSTRUCTS AND SITE-DIRECTED MUTAGENESIS

A previously designed pET28a plasmid construct was used, the DJ-1_pSKB-3 (Figure 2.1) that contains the cDNA sequence of wild-type DJ-1 fused to an N-terminal hexahistidine tag.

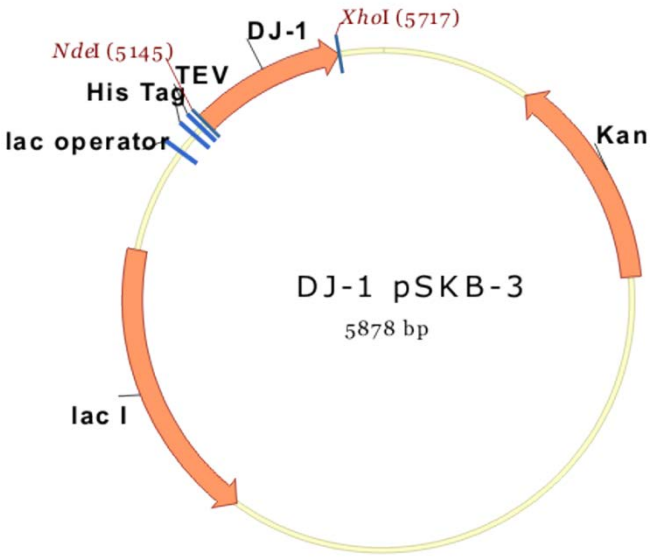


Figure 2.1 | MAP OF DJ-1_PSKB-3 CONSTRUCT. Representation of the DJ-1_pSKB-3 construct that contains: lac I – lac repressor gene; His Tag – hexahistidine tag; TEV – tobaco etch virus protease recognition site; kan – kanamycin resistance gene; the DJ-1 cDNA (optimized for *E. coli* expression) cloned between Nhe I and Xho I restriction endonuclease sites. The vector pSKB-3 corresponds to pET28a where thrombin recognition site was modified to TEV recognition site.

The QuickChange® Site Directed Mutagenesis Kit from Stratagene was used in the DJ-1_pSKB-3 construct in order to generate constructs with the E163K, C106A and C106DD mutants´ forms of human DJ-1. The primers containing the pathogenic mutations were designed according to the manufacturer's recommendations, with codons optimized for expression in *Escherichia coli* (Table 1.1).

DJ-1 mutants´ constructs were used to transform *E. coli* XL1 blue competent cells (Agilent), grown at 37^{o}C in Luria Broth Base (Miller's LB Broth Base) (Invitrogen/agar plates supplemented with kanamycin (50µg/mL)). Colonies were picked and grown in liquid LB supplemented with kanamycin (50µg/mL), the DNA was isolated using PureLink® Quick Plasmid Miniprep Kit (Invitrogen) and the DNA

sequence of all DJ-1 mutants´ constructs were confirmed by Sanger DNA Sequencing at STAB Vida.

Table 1.1 | DJ-1 MUTATIONS WITH RESPECTIVE CODON SUBSTITUTIONS AND PRIMERS.

Mutation	*Codon substitution*	*Primer*
E163K	GAA → AAA	Forward: 5´- GTC CGG GTA CAA GCT TTA AAT TTG CAC TGG CAA TTG -3´ Reverse: 5´- CAA TTG CCA GTG CAA ATT TAA AGC TTG TAC CCG GAC -3´
C106A	TGT → GCG	Forward: 5´- CTG ATT GCA GCA ATT GCG GCA GGT CCG ACC GCA CTG -3´ Reverse: 5´- CAG TGC GGT CGG ACC TGC CGC AAT TGC TGC AAT CAG -3´
C106DD	TGT → GAT GAT	Forward: 5´- CGC AAA GGT CTG ATT GCA GCA ATT GAT GAT GCA GGT CCG ACC GCA CTG C -3´ Reverse: 5´- GCA GTG CGG TCG GAC CTG CAT CAT CAA TTG CTG CAA TCA GAC CTT TGC G -3´

2.2. DJ-1 MUTANTS PRODUCTION AND PURIFICATION

The DJ-1 mutants´ constructs were transformed into competent *E. coli* BL21star (DE3) strain and plated into LB/agar supplemented with 50µg/mL kanamycin. One colony was used to inoculate 50mL of LB supplemented with 50µg/mL kanamycin that was grown overnight at 37°C. To inoculate 1L of LB, supplemented with 50µg/mL kanamycin, 25mL of the previous cell culture were used. The cells were allowed to grow at 37°C with shaking. When the optical density at 600nm reached 0.5, the temperature of the incubator was decreased to 18°C and after one hour IPTG was added to a final concentration of 1mM to induce protein expression. This was allowed to occur for 16 hours.

The cell suspension was centrifuged (20 min, 4 000*g*, 4°C), the cellular pellet was resuspended in 60mL of 20mM sodium phosphate, 500mM NaCl, 20mM Imidazole, pH 7.5, and disrupted through a high pressure homogenizer EmulsiFlex-C3 (Avestin), by 3 passages at 1 000 bar. The cellular extract obtained was clarified by centrifugation (20 min, 18 900*g*, 4°C) and the supernatant was loaded into a 5mL HisTrap HP column (GE Healthcare) using the peristaltic pump EP-1 Econo Pump (Bio-

Rad). After column loading, the column was extensively washed with loading buffer (20mM sodium phosphate, 500mM NaCl, 20mM Imidazole, pH 7.2) and subsequent protein elution was obtained by stepwise increasing of imidazole concentration up to 500mM (50, 100, 300 and 500mM). The fractions containing the DJ-1 mutants were combined, loaded into a HiLoad 26/60 Superdex 200 prep grade and protein elution was obtained using a solution of PBS (8mM K_2HPO_4, 2mM $NaH_2PO_4 \cdot H_2O$, 150mM NaCl). The fraction containing the dimeric form of the DJ-1 mutants was concentrated using 10kDa molecular weight cut-off (MWCO) centrifugal concentration (Millipore) and then protein concentration was determined by Nanodrop ND-1000 Spectrophotometer (Thermo FisherScientific) measuring the absorption at 280nm, and stored in PBS with 10% glycerol at -80^{o}C. These fractions were also analyzed by silver stained SDS-PAGE [4-20% polyacrylamide gel (Bio-Rad)]. The other fractions recovered were quantified using the BCA protein determination assay (Thermo Fisher Scientific) and analyzed by Coomassie Brilliant Blue G-250.

2.3. GEL SEPARATION, GEL BAND PROCESSING AND PEPTIDE EXTRACTION

Samples from DJ-1 WT and mutants were electrophoretically separated in a pre-cast 4-20% SDS-polyacrylamide gel (Bio-Rad) using a Mini-PROTEAN Tetra Electrophoresis System (Bio-Rad). Sample preparation involved protein denaturation with Laemmli Sample Buffer 2x (65.8mM Tris-HCl, pH 6.8, 2.1% SDS, 26.3% (w/v) glycerol, 0.01% bromophenol blue) (Bio-Rad) and boiling at 95^{o}C for 5 minutes. After cooling at room temperature, 4µL of acrilamide/bis-acrilamide solution for each 30µL of sample was added to promote the cysteine alkylation. After protein separation, the gel was stained with Colloidal Coomassie Blue. The entire gel lanes were sliced into small pieces and transfered to microcentrifuge tubes with 1mL of deionized water to prevent gel bands dehydration. To destain the gel bands the water was remove, 1mL of destaining solution (50mM ammonium bicarbonate and 30% acetonitrile) was added and the tubes were placed in a thermomixer (Comfort, Eppendorf) at 850 rpm for 15 minutes. The previous procedure was repeated if the gel pieces that remained blue, otherwise 1mL of water was added and tubes were placed in a thermomixer for 10 minutes at 850 rpm. After this, the water was removed and 100µL of acetonitrile

(ACN) was added to promote faster gel dehydration. The ACN was removed and the gel bands were dehydrated on the Concentrador Plus (Eppendorf) for 1 hour at 60°C. Adequate volume of trypsin (15ng/µL in 10mM ammonium bicarbonate) was added to cover the dried gel bands and tubes were incubated for 10 minutes on ice to rehydrate the gel. After this time, sufficient volume of 10mM ammonium bicarbonate (final concentration of trypsin 10ng/µL) was added to cover gel bands and samples were incubated overnight at room temperature in the dark to perform the in-gel digestion. The tryptic solution (containing trypsin and peptides) was collected to low binding microcentrifuge tubes (Eppendorf) and the remaining peptides were extracted by sequential addition of 50µL of 30%, 50%, and 98% of ACN in 1% formic acid (FA). After the addition of each solution, the tubes were agitated in a thermomixer at 1 050 rpm for 15 minutes and the solution was collected to the tube containing the initial tryptic solution. Samples containing the peptides were concentrated on the Concentrador Plus (Eppendorf) at 60°C, resuspended to a final volume of 100µL in a solution of 2% ACN and 1% FA, and sonicated on a Sonics 750W using a cup-horn (2 min with 1 sec on 1 sec off cycles at 20% of amplitude). The next step was desalting the peptides using C18 Bond Elut OMIX solid phase extraction pipette tips (Agilent technology). Briefly, tip columns were hydrated with 100µL of 50% ACN and equilibrated with 100µL of 2% ACN and 1% FA. Peptides were loaded to the columns and this step was repeated five times, followed by a washing step with 100µL of 2% ACN and 1% FA solution and elution to new protein LoBind tubes with 100µL of 70% ACN and 0.1% FA. Eluates were concentrated using the Concentrator Plus at 60°C until approximately 10µL of solution remained. Sample volume was adjusted to 15 or 30µL in a solution with final concentration of 2% ACN and 0.1 % FA. Finally, samples were sonicated on a Sonics 750W using a cup-horn (2 min with 1 sec on 1 sec off cycles at 20% of amplitude), centrifuged 5 minutes at 14 100*g* and the supernatant was transferred into vials for subsequent LC-MS/MS analysis.

2.3.1. POLYACRYLAMIDE GEL STAINING METHODS

To evaluate the SDS-PAGE profile of the fractions recovered from the proteins´ production and purification homemade gels with 12.5% of polyacrylamide and 1.5mm thickness were used. For gel band digestion and protein identification, pre-cast gels

with 4 to 20% polyacrylamide and 0.75mm (Bio-Rad) were used. Samples were electrophoretically separated using a Mini-PROTEAN Tetra Electrophoresis System (BioRad) and gels were stained with Coomasie Brilliant Blue G-250 (Bio-Rad) or Silver (Sigma-Aldrich).

2.3.1.1. COLLOIDAL COOMASSIE STAINING

The Coomassie staining was adapted from the work described by Candiano and co-workers [156] with some adjustments. Briefly, after electrophoresis, the gel was washed with distilled water and immersed in staining solution [10% (v/v) of 85% solution of phosphoric acid, 10% (w/v) ammonium sulphate, 20% (v/v) methanol]. The Coomassie powder was added to the solution with a filter to allow the formation of colloidal particles under agitation and allowed to stain overnight. After that, successive washes with distilled water were made and the gel was maintained in water to use in further analysis.

2.3.1.2. SILVER STAINING

Silver staining was used to allow the detection of low level proteins separated by electrophoresis. Therefore, the method used by O'Connell and Stults [157] was adapted in which the gel was washed with distilled water and immersed in the fixation solution [25% (v/v) methanol and 5% (v/v) acetic acid] for 30 minutes with agitation (all steps were performed under agitation). Subsequently, the gel was dehydrated by consecutive immersion in two alcoholic solutions of gradually decreasing concentrations [50% (v/v) ethanol followed by 30% (v/v) ethanol] for 10 minutes each. Then, the gel was submersed in a solution with 0.2g/L of sodium thiosulfate for 1 minute followed by 3 washes with de-ionized water and immersion in the solution of silver nitrate with a final concentration of 2g/L for 20 minutes. Finally, it was added the developing solution [0.7mL/L of a solution of 37% of formaldehyde, 30g/L of sodium carbonate anhydrous and 10mg/L of sodium thiosulfate]. Once the desired staining was reached, a stop solution [50g/L of Tris and 2.5% of acetic acid] was added for 1 minute. Until further use the gel was stored in distilled water.

2.4. PROTEIN IDENTIFICATION BY LC-MS/MS

Protein identification experiments were carried out on a hybrid quadrupole time-of-flight mass spectrometer (Triple TOF™ 5600 System; ABSciex) after both gel and liquid digestions. Peptides separation was performed using liquid chromatography (nanoLC Ultra 2D, Eksigent) on a ChromXP™ C18 reverse phase column (300µm ID × 15cm length, 3µm particles, 120Å pore size, Eksigent) at 5µL/min. Peptides were eluted into the mass spectrometer with an acetonitrile gradient in 0.1% FA (2% to 35% ACN, in a multiple step gradient for 25 min for gel digested samples and 5% to 35% ACN, in a multiple step gradient for 25 min for liquid digested samples), using an electrospray ionization source (DuoSpray™ Source, ABSciex). The mass spectrometer was set for information dependent acquisition (IDA) scanning full spectra (350-1250 m/z), followed by 20 MS/MS on multiple charged ions (+2 to +5) and performed one MS/MS before adding those ions to the exclusion list for 15s (mass spectrometer operated by Analyst® TF 1.6, ABSciex).

Protein identification was accomplished using Protein Pilot software (v4.5, ABSciex) with the following search parameters: SwissProt database, against *E. coli* and all species, SwissProt database with DJ-1 mutants in study, against all species; trypsin digestion; acrylamide as cysteine alkylating reagent; thorough ID search effort, and 0.05 unused ProtScore (10% confidence score) as detected protein threshold. Data analysis was based on an independent False Discovery Rate analysis (FDR) using the target-decoy approach (5% critical FDR). When the analysis based on FDR was not possible to perform, the criteria used in positive identification were proteins presenting at least 95% confidence and unused score of 1.3 or more. Proteins with a single peptide hit were considered if the individual confidence was above 95% and a minimum sequence tag of 3 amino acids (4 consecutive peaks in the MS/MS spectrum).

2.5. INTACT PROTEIN ANALYSIS BY LC-MS

To analyze the intact DJ-1 WT and mutants, 400pmols of each protein were mixed with 6 volumes of acetone at -20^{o}C and incubated at -20^{o}C for 20 minutes. After

that time the samples were centrifuged at 20 000*g* (Eppendorf Centrifuge 5417R), 4^{o}C for 20 minutes. The pellet was resuspended in 15µl of 2% ACN and 0.1% FA and transferred into vials for further LC-MS analysis. Protein analysis was carried out on a hybrid quadrupole time-of-flight mass spectrometer (Triple TOF™ 5600 System; ABSciex). Proteins separation was performed using liquid chromatography (nanoLC Ultra 2D, Eksigent) on a ChromXP™ C18 reverse phase trap-column (300 µm ID × 0.5mm length, 3µm particles, 120Å pore size, Eksigent) at 7µL/min. Proteins were eluted into the mass spectrometer with an acetonitrile gradient in 0.1% FA (2% to 50% ACN, in a multiple step gradient for 10min), using an electrospray ionization source (DuoSpray™ Source, ABSciex) with the Intact Protein mode activated. The mass spectrometer was set for full spectra scanning (600-1250 m/z, operated by Analyst® TF 1.6, ABSciex). Intact mass was determined using BioAnalyst™ Software (ABSciex). Briefly, the deconvolution of the spectra was obtained by Baesian Protein Reconstruct of the average of the spectra, with the following parameters: mass range from 15 to 30kDa (defined according with the expected protein molecular weight), spectrum limit range from 620 to 1260m/z (defined according to the protein charge envelope profile), signal to noise threshold around 10 (determined by a script present in Analyst software that is based in the ratio from the less intense peak used in calculation and the spectrum noise). A peak from the isotopic distribution of the deconvoluted spectrum was chosen to modeling the data and subsequently calculate the molecular weight of the protein.

2.6. SIZE EXCLUSION-HPLC

Size exclusion (SE) - High-performance liquid chromatography (HPLC) was performed using a Prominence Shimadszu system (Shimadzu Scientific Instruments) and a Superdex 200 5/150 GL column (GE Healthcare), and data were collected and analyzed by LC Solution Software (Shimadzu Scientific Instruments). The proteins were eluted using PBS with 10% glycerol (the same buffer in which the proteins were stored) at a flow rate of 0.4mL/min and monitoring absorbance at 214nm. Apparent molecular weights of eluted proteins were determined using a calibration curve established with the following standards: aprotinin (6.5kDa) (AppliChem), Ribonuclease A (13.7kDa),

Carbonic anhydrase (29kDa), Ovalbumin (44kDa), Conalbumin (75kDa), Aldolase (158kDa), Ferritin (440kDa), and Blue Dextran (≈2 000kDa) (all from GE Healthcare).

2.7. CIRCULAR DICHROISM

The secondary structure of DJ-1 wild-type and mutant proteins was evaluated by Circular Dichroism (CD) spectroscopy. The proteins were dissolved in PBS buffer containing 10% glycerol at concentrations around 2.0mg/mL as determined by ultraviolet (UV) absorbance measurements. Far UV CD spectra were acquired on an Olis DSM 20 circular dichroism spectropolarimeter continuously purged with nitrogen, equipped with a Quantum Northwest CD 150 temperature-controlled cuvette and controlled by the GlobalWorks software. Spectra were recorded at 37^{o}C between 190 and 260nm at 1 nm intervals using a 0.05mm pathlength cuvette. Two scans obtained with an integration time of 6 seconds were averaged and corrected by subtracting a baseline spectrum acquired under the same experimental conditions. The results are expressed as the mean residue molar ellipticity $[\Theta]_{MRW}$ (deg cm^2 dmol^{-1}) defined as $[\Theta]_{MRW} = \Theta_{obs} \times 100 \times MW / (l \times c \times N)$, where Θ_{obs} is the observed ellipticity (mdeg), MW is the protein molecular weight (g mol^{-1}), l is the cuvette pathlength (cm), c is the protein concentration (mg mL^{-1}) and N is the number of aminoacid residues of the protein. The secondary structure contents were calculated with the GlobalWorks software using the CONTILL, CDSSTR and Selcon3 algorithms against the CLSTR reference basis set which contains soluble and denatured proteins with known secondary structure.

2.8. THERMAL SHIFT

Fluorescent dye Sypro orange (Sigma-Aldrich), an environmentally sensitive dye, was used to monitor protein unfolding. The unfolding process exposes the hydrophobic region of proteins and results in a large increase in fluorescence that is used to monitor the protein-unfolding transition. The thermal shift assay was conducted in the 7500 Fast Real-time PCR system (Applied Biosystems), originally designed for PCR. The system contains a heating/cooling device for accurate temperature control and a charge-coupled device (CCD) detector for simultaneous

imaging of the fluorescence changes in the wells of the microplate. To each well of the MicroAmp® Optical 96-well reaction plate (Applied Biosystems) 5µL of 50x Sypro orange was added, the correspondent volume of protein to a final concentration of 0.5mg/mL and 0.2mg/mL, and buffer (PBS with 10% glycerol) to the final volume of 50µL. The plate was heated from 25 to 95^{o}C with a heating rate of 1^{o}C/min. The fluorescence intensity was measured with excitation/emission wavelengths of 490/580nm.

2.9. DJ-1 AND DJ-1 MUTANTS MEDIATED NEUROPROTECTION

2.9.1. CELL CULTURE

Human neuroblastoma SH-SY5Y cells were cultured in Dulbecco's modified Eagle medium (DMEM) with Glutamax™ and low glucose (1g/L) (Gibco) supplemented with 10% fetal bovine serum (FBS) (Gibco), 1.25µg/mL of amphotericin B solution (Invitrogen) and 1% penicilin-streptomycin solution (Pen-Strep) (Cambrex). Dulbecco's phosphate buffered saline (DPBS) (Cambrex, Charles City, IA) was used to wash the cells and trypsin-EDTA (0.05% solution in phosphate buffered saline (PBS)) (Invitrogen) to detach for cell passage. Cells were maintained at 37^{o}C and 95% air/ 5% CO_2 atmosphere in a humidified incubator (Shel Lab 3517-2) (Sheldon Manufacturing, Inc.).

For the neuroprotection assays, SH-SY5Y cells were seeded at 30x10^{3}cells/cm^{2} in DMEM with 10% FBS in 96 well plates (Corning) at 37^{o}C, 5% CO_2/95% air in a humidified incubator (Sheldon Manufacturing, Inc.) for 4h.

2.9.2. OXIDATIVE STRESS STIMULI

Four hours after plating, cells were stimulated with different concentrations (0, 100 and 200µM) of hydrogen peroxide (Sigma-Aldrich) in DMEM with 0.1% of FBS in the presence or absence of recombinant His-tagged human DJ-1 protein (1µM), recombinant His-tagged human DJ-1 mutants (E163K, C106A or C106DD) or the corresponding vehicle (PBS with 10% of glycerol) for 24h. Change of the culture medium to DMEM with 0.1% of FBS was used as control condition.

2.9.3. CELL VIABILITY ASSESSMENT

Cell viability was performed in white opaque 96-well plates (Corning) using the Cell Titer-Glo® Luminescent assay (Promega). The luminescent signal was detected by a LUMIstar Galaxy automated microplate luminescence reader (BMG Labtech), according to the manufacturer's instructions. Briefly, this method determine the viable cell number based on bioluminescent quantitation of the ATP present in the sample, which indicates the presence of metabolically active cells.

2.10. PULL-DOWN ASSAYS

2.10.1. OXIDATIVE STIMULI

SH-SY5Y cells were seeded at 75 000 cells/cm^2 in DMEM medium with 10% FBS. After 24h, serum starvation was induced by changing the culture medium to DMEM medium with 0.1% FBS. The stimulus was performed 16h afterwards with the addition of 1mM of H_2O_2 in DMEM medium with 0.1% FBS to the cells for different time points: 0, 15 and 40 minutes. Once stimulation time was over and to finish the stimulation, the cell culture medium was removed and PBS at room temperature was added to the cells.

2.10.2. PROTEIN EXTRACTS

After cell washing with PBS, cells were placed on ice and scraped with ice-cold Extraction Solution (ES) (IP (from Dynabeads co-immunoprecipitation kit, Invitrogen), 100mM of NaCl, Complete Mini protease inhibitor mixture and Complete Mini phosphatase inhibitor mixture (Roche Diagnostics)), harvested to centrifuge tubes and incubated on ice for 15 minutes. After that time cells were centrifuged at 2 600*g* for 5 minutes at 4^{o}C. Later, the supernatant (protein extract) was collected to new tubes. A sample of the protein extract was collected to calculate the protein concentration using the BCA protein determination assay (Thermo Fisher Scientific) and the remaining extracts were frozen at -80°C until further use.

2.10.3. PROTEIN QUANTIFICATION – BCA ASSAY

The bicinchoninic acid (BCA) assay was used to determine the protein concentration in protein extracts. Bovine serum albumin (BSA) was used as standard and the assay was performed following the manufacturer protocol (Pierce as part of Thermo Fisher Scientific, Rockford, IL) for microwell plates. Shortly, the standards were prepared in IP buffer (IP and 100 mM of NaCl), then samples and standards were mixed with BCA Working Reagent in a 1:8 proportion and incubated for 30 minutes at 37^{o}C. After that time, samples´ absorvance was measured at 562nm on a Microplate Spectrophotometer (PowerWave XS, BioTek, Bad Friedrichshall, Germany). Using the results, the calibration curve was obtained and the protein extract concentration was interpolated from it.

2.10.4. PULL-DOWN

Initially the HisTrap resin (Ni Sepharose™ High Performance GE Healthcare) was resuspended in approximately 50 times the resin volume of PBS with 10% glycerol and 40mM imidazole and centrifuged at 1 000*g* for 3 minutes. This process was repeated twice. After supernatant removal, enough volume of PBS and 40mM imidazole was added to the resin to allow distribution of 60µL of resin (Ni Sepharose™ High Performance HisTrap™ HP - GE Healthcare) into each tube. The resin was incubated with 5µg of recombinant DJ-1 wild-type or mutants (amount determine for resin saturation) for 2h at 1 100 rpm and 4^{o}C in a thermomixer (Eppendorf). Following centrifugation at 1 000*g* for 3 minutes, the supernatant containing proteins that did not bind to HisTrap resin was collected to a clean tube. The resin was washed with 500µL of PBS and 40mM imidazole, centrifuged at 1 000*g* for 3 minutes and the supernatant was recovered to a new tube. The previous procedure was repeated three times. Then, the resin was incubated with protein extracts (obtained under normal and oxidative stress conditions – section 2.10.2) for 16h at 4^{o}C and 900 rpm. After that time, the tubes were centrifuged at 1 000*g* for 3 minutes and the supernatant was recovered. One milliliter of IP buffer was added to the resin, the tubes were centrifuged at 1 000*g* for 3 minutes and the supernatant recovered. This process was

repeated three times. Finally, the resin with the proteins attached was denatured with 30µL of Laemmli Sample Buffer 2x (65.8mM Tris-HCl, pH 6.8, 2.1% SDS, 26.3% (w/v) glycerol, 0.01% bromophenol blue) (Bio-Rad) at 95^{o}C for 15 minutes at 14 000 rpm in a thermomixer. The fractions of interest were electrophoretically separated using a polyacrylamide gel electrophoresis and other fractions and supernatants were stored at -20^{o}C until further use.

2.10.5. GEL SEPARATION, GEL BAND PROCESSING AND PEPTIDE EXTRACTION

Pull-down samples, containing resin with DJ-1 WT, DJ-1 mutants and their interactors, were electrophoretically separated in a pre-cast 4-20% SDS-polyacrylamide gel (Bio-Rad) using a Mini-PROTEAN Tetra Electrophoresis System (Bio-Rad) for 15min at 110V. Gel staining and processing was done as described in section 2.3 with modifications after sample evaporation and before trypsin addition. In this case, 50µL of DTT was added and incubated with agitation at 25^{o}C, for 30min at 850rpm. Then the DTT was removed and the samples were evaporated for 30min to allow the addition of 50µL of acrylamide and incubation with agitation at 25^{o}C, for 30min at 850rpm. After that, the acrylamide was removed and samples were evaporated for 30min. Finally, 200µL of water was added and removed to allow the hydration of the gel pieces and samples were dried on a Concentrador Plus (Eppendorf).

2.10.6. PROTEIN IDENTIFICATION AND QUANTIFICATION

Samples were analyzed on an AB Sciex 5600 TripleTOF in two phases: information-dependent acquisition (IDA) followed by SWATH (Sequential Windowed data independent Acquisition of the Total High-resolution Mass Spectra) acquisition, on the same sample. Peptides separation was performed using liquid chromatography (nanoLC Ultra 2D, Eksigent) on a Halo Fused-Core C18 reverse phase column (300µm ID × 15cm length, 2.7µm particles, 90Å pore size, Eksigent) at 5µL/min. Peptides were eluted into the mass spectrometer with an acetonitrile gradient in 0.1% FA (2% to 30%, in a linear gradient for 45 min), using an electrospray ionization source (DuoSpray™ Source, ABSciex). Rolling collision was used with a collision energy spread of 5. One third of the volume of each sample was analyzed using classical shotgun data

acquisition. The mass spectrometer was set for IDA scanning full spectra (350-1250m/z) for 250ms, followed by up to 20 MS/MS scans (100–1500 m/z for 100ms each). Candidate ions with a charge state between +2 and +5 and counts above a minimum threshold of 70 counts per second were isolated for fragmentation and one MS/MS spectra was collected before adding those ions to the exclusion list for 15 seconds (mass spectrometer operated by Analyst® TF 1.6, ABSciex). Two thirds of the volume of each sample were set for quantitative analysis by acquisition in SWATH mode. The SWATH setup was essentially performed as in Gillet et al [158], with the same chromatographic conditions used as in the IDA run described above. For SWATH-MS–based experiments, the mass spectrometer was operated in a looped product ion mode. The instrument was specifically tuned to allow a quadrupole resolution of 25-m/z mass selection. Using an isolation width of 26 m/z (containing 1 m/z for the window overlap), a set of 29 overlapping windows was constructed covering the precursor mass range of 350–1100 m/z. A 50 ms survey scan (350-1500 m/z) was acquired at the beginning of each cycle for instrument calibration and SWATH MS/MS spectra were collected from 100–1500 m/z for 100 ms resulting in a cycle time of 3.25s from the precursors ranging from 350 to 1100 m/z. The collision energy for each window was determined according to the calculation for a charge 2+ ion centered upon the window with a collision energy spread of 15.

2.10.6.1. LIBRARY GENERATION AND SWATH DATA FILE PROCESSING

A specific library of precursor masses and fragment ions was created for each recombinant protein used in the interactomic study – by combining all files from IDA experiments - and used for subsequent SWATH processing. Libraries were obtained using Protein Pilot software (v4.5, ABSciex) with the following search parameters: canonical SwissProt complete proteome database for human with DJ-1 mutants in study and iRTs peptides; trypsin digestion; acrylamide as cysteine alkylating reagent; thorough ID search effort, and 0.05 unused ProtScore (10% confidence score) as detected protein threshold. Data analysis was based on an independent False Discovery Rate (FDR) analysis using the target-decoy approach.

Prior to data processing, peptides were selected automatically from the library using the following criteria: (i) the unique peptides for a specific targeted protein were ranked by the intensity of the precursor ion from the IDA analysis as estimated by the ProteinPilot software, and (ii) Peptides that contained biological modifications and/or were shared between different protein entries/isoforms were excluded from selection. Up to 15 peptides were chosen per protein, and SWATH quantitation was attempted for all proteins in library files that were identified below 5% local FDR from ProteinPilot searches. In SWATH™ Acquisition data, peptides are confirmed by finding and scoring peak groups, which are a set of fragment ions for the peptide.

Target fragment ions, up to 7, were automatically selected following the criteria described in Lambert et al [159]: (i) Fragment ions for a selected peptide were ranked according to ion intensity; (ii) ions higher in m/z than the y4 fragment ion for each selected peptide were ranked highest; (iii) ions within the SWATH isolation window were excluded from selection; (iv) if insufficient target ions were found, ions lower than y4 but outside of the SWATH window were chosen; (v) if there were still insufficient ions, then fragment ions from within the SWATH window region were chosen.

Peak groups were scored following the criteria described in Lambert et al [159], briefly, each target fragment ion was scored based on the retention time adjustment, peak width overlap, peak intensity ratio, correct isotopic state, m/z error and MS/MS score. The individual peak group scores were combination of all subscores (scores from fragment ions) and the peak group with the best score was taken forward. Peak group confidence threshold was determined based on a FDR analysis using the target-decoy approach and 1% extraction FDR threshold was used for all the analyses.

Peptide features (i.e., peptides in a given charge state) that met the 1% FDR threshold in all the experiments were retained, and the peak areas of the target fragment ions of those peptides were extracted across the experiments using a extracted-ion chromatogram (XIC) window of 5 minutes.

Protein levels were estimated by summing all the transitions from all the peptides for a given protein (an adaptation of [160]). Protein levels were normalized for the levels of a peptide used as internal standard (the iRT peptide that demonstrated more stability (low %CV) across 4 analytical replicates). To identify the

true protein interactions, normalized data were compared between DJ-1 pull-down assays and resin (negative control). Only proteins which have twice or more the intensity in DJ-1 pull-down assays compared to resin were considered true interactions.

For comparisons between DJ-1 purifications across the time course, abundances were normalized with respect to the four most abundant tryptic peptides from DJ-1. All comparisons were made to the control condition (0 min of hydrogen peroxide stimulus).

2.11. BINDING PARTNERS QUALITATIVE ASSESSMENT

Venn diagrams were used to compare the number of binding partners that are shared or unique among two or more conditions. The Venn diagram tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) from Bioinformatics & Evolutionary Genomics was used. The lists containing the proteins from the desired conditions to be compared were uploaded and the selected output was symmetric Venn diagrams. The tool generates a Venn diagram of the inserted conditions and a ".txt" file containing the accession numbers of the analyzed proteins and the distribution of these proteins within the different conditions compared. Both diagrams and ".txt" files were downloaded and presented as results or used in further analysis.

2.12. GENE ONTOLOGY (GO) ENRICHMENT ANALYSIS

GO enrichment analysis was performed for each binding partners list, using the web-based application Gene Ontology enRIchment anaLysis and visuaLizAtion tool – GOrilla (http://cbl-gorilla.cs.technion.ac.il/) [161].

GOrilla helps to identify and visualize enriched GO terms in genes or proteins lists. The output of this analysis is visualized as a hierarchical structure, using a DAG (directed acyclic graph) graphical representation with color-coding that reflects the enrichment degree based on the *p*-value. Additionally, the nodes are clickable giving further information displayed on a table with information of the enriched GO terms

analyzed, their enrichment *p*-values and the relevant annotated genes/proteins used in the analysis. This data can be exported as an Excel file.

All the analyses were done for the Homo sapiens species, in the single ranked list of genes running mode and based on biological process. Both diagram and excel file were exported and presented as results or used in further analysis.

2.13. STATISTICAL ANALYSIS

Statistical analysis of results from cell viability was performed using SPSS (Statistical Package for the Social Sciences) version 20.0 (IBM®). Statistical significance was considered relevant for $*p < 0.05$, $**p < 0.01$ and $***p < 0.001$ using one-way ANOVA (Analysis Of Variance). Parametric assumptions (data normality and homoscedasticity) were tested using Shapiro-Wilk Test and Levene's Test, respectively. Data presented as mean ± standard error of the mean (S.E.M.).

3. RESULTS

DJ-1 has been unequivocally linked to familial early onset Parkinson´s Disease [43, 81] and although its exact function has not been determined, this protein is widely recognized as involved in oxidative stress response, a main feature in preservation of neurons viability [2, 67]. Importantly, cysteine in position 106 is crucial in neuroprotection mediated by DJ-1 due to its essential role in DJ-1 acidic pI shift in oxidative stress conditions [80, 100].

DJ-1 mutations trigger loss of the protein neuroprotective function leading to the development of PD [75, 89]. PD-linked mutations associated to DJ-1 comprise natural and also engineered mutations developed to mimic and study the oxidative states of this protein [45, 101].

Understanding the role and mechanisms of this protein can give insight into the pathogenesis of PD [94]. Therefore, in order to study the effect of DJ-1 mutations in neuronal cells under normal and stress conditions the project comprises the production, purification, structural characterization, viability and interactomic assays of DJ-1 wild-type and mutants E163K, C106A and C106DD. Specifically and following the project plan, the first step was confirmation of sequences mutation followed by production and purification of the proteins for further characterization by size exclusion chromatography, circular dichroism, thermal shift and mass spectrometry, study of the neuroprotective effect in a neuronal cell line and finally characterize the dynamic interactome of each protein.

It is worth mentioning that DJ-1 WT was produced and purified by Matilde and kindly given to be used and studied in the project assays. Furthermore, in this project major emphasis will be given to the study of the mutants E163K, C106A and C106DD.

3.1. NUCLEOTIDE SEQUENCING

In order to produce the DJ-1 mutants of interest, the first step was to construct plasmids containing each mutated DJ-1 gene, in this case the plasmid previously used to produce DJ-1 WT (see Experimental Procedures – 2.1). Before plasmids incorporation into competent cells, the sequence of DJ-1 gene needed to be confirmed for the presence of the correct mutated nucleotides. Thus, all plasmid DNA obtained was sequenced by Sanger DNA Sequencing method and the results from C106DD mutant showed the correct substitution (Figure 3.1).

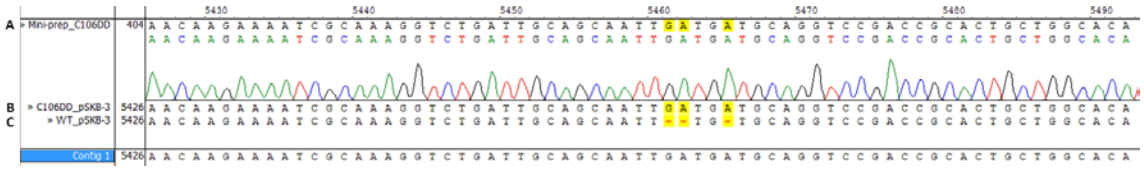


Figure 3.1 | DNA SEQUENCING RESULTS FOR THE DJ-1 MUTANT C106DD CODIFIYING SEQUENCE. Nucleotide sequence alignment of small fragments of DJ-1 mutant C106DD gene sequenced (A), its corresponding sequence in the C106DD designed plasmid (B) and in the DJ-1 WT designed plasmid (C). The codon containing the mutation is highlighted. The codon substitution was TGT to GAT GAT.

Comparing to the DJ-1 WT gene it was verified the insertion of nucleotides as expected since this mutation implicates the substitution of one codon by two codons. Given this, it was confirmed the design of a plasmid DNA with the DJ-1 gene incorporating the C106DD mutation and it was possible to proceed to the production and purification of the mutated proteins.

Similar results were obtained for the sequencing of mutants C106A and E163K (Supplementary Figure 7.1).

3.2. DJ-1 MUTANTS PRODUCTION AND PURIFICATION

3.2.1. PRODUCTION AND PURIFICATION – AFFINITY AND SIZE EXCLUSION CHROMATOGRAPHIES

The plasmids previously constructed and sequences verified were used for protein production followed by protein purification. The recombinant mutants E163K, C106A and C106DD, containing a His-tag, were expressed in *E. coli* XL1 blue competent

cells and purified by affinity chromatography using a HisTrap column followed by a size-exclusion chromatography using a Superdex 200 column. The HisTrap column was used to purify the recombinant DJ-1 and mutants from the culture medium, and to separate the homodimeric DJ-1 from DJ-1 oligomers the Superdex 200 column was used.

The affinity chromatography (Figure 3.2) was performed using an imidazole step gradient (20mM sodium phosphate, 500mM NaCl, 20mM Imidazole, pH 7.2) and from the fractions recovered, only fraction 4 did not contained DJ-1. The highest value of optical density (OD) was observed in fraction 3, which suggests that this fraction contains a higher protein content. Therefore, and since in this elution step 300mM of imidazole were used, most of the proteins which do not contain the hexahistidine peptide have already been eluted, and this was the fraction chosen to load into the Superdex 200 column to perform a size exclusion chromatography.

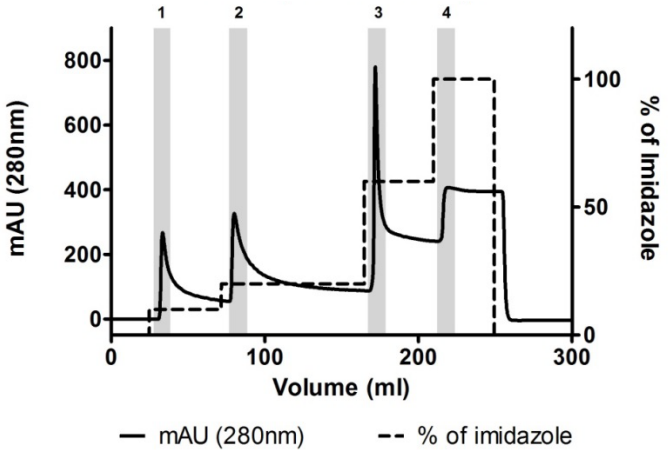


Figure 3.2 | AFFINITY CHROMATOGRAPHY PURIFICATION OF DJ-1 C106DD. The affinity chromatography was performed using a HisTrap column. Imidazole step gradient (50mM, 100mM, 300mM and 500mM) is shown in dashed line. The grey bars correspond to the fractions that were collected from each elution step - fraction 1, 2, 3 and 4 eluted with 50, 100, 300 and 500mM of imidazole, respectively.

The size exclusion chromatography allowed the recovery of two protein fractions (Figure 3.3) with the second one containing a higher protein content as shown by the highest OD at 280nm. Both fractions were collected and further analyzed with major interest in fraction 2.

For the other mutants, similar results were obtained from affinity and size exclusion chromatography (Supplementary Figure 7.2 and Supplementary Figure 7.3).

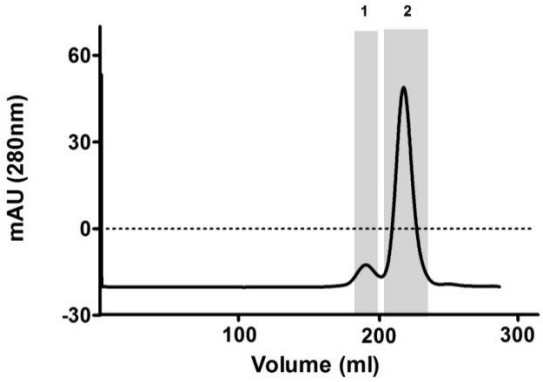


Figure 3.3 | SIZE EXCLUSION CHROMATOGRAPHY PURIFICATION OF DJ-1 C106DD. Fraction 3 recovered from the affinity chromatography (Figure 3.2) was applied in an HiLoad 26-60 Superdex 200 column, using PBS as the mobile phase. The grey bars represent the fractions that were collected (fraction 1 and 2).

3.2.2. PROTEIN CONCENTRATION

After the fractions of interest for each protein have been recovered, the protein concentration was determined by measuring the absorbance at 280nm (NanoDrop). The blank used in this procedure was the buffer in which the protein was and would be stored (PBS with 10% glycerol) and the result obtained was 2.19µg/µL for fraction 2 of the C106DD mutant. The protein concentration of the remaining mutants is shown in Supplementary Table 7.1.

3.2.3. PROTEIN PURIFICATION ASSESSMENT

The fractions recovered during the production and purification steps were resolved in a SDS-PAGE in order to evaluate the protein profile (Figure 3.4 and Figure 3.5). In the case of protein expression the fractions recovered and analyzed were the BI (cellular extract obtained before induction of protein expression), AI I (insoluble protein fraction obtained from centrifugation after induction of protein expression) and AI S (soluble protein fraction obtained after induction of protein expression), and in protein purification using affinity chromatography the BL (cellular extract loaded on the HisTrap column), FT (flow-through obtained from HisTrap column), HT 1 (fraction 1 eluted from the HisTrap column), HT 2 (fraction 2 eluted from the HisTrap column), HT 3 (fraction 3 eluted from the HisTrap column) and HT 4 (fraction 4 eluted from the

HisTrap column (Figure 3.2)) (Figure 3.4). In the size exclusion chromatography the fractions were the Sx1 and Sx2, in protein concentration the BC Sx2, FT Sx2 and AC Sx2 and the stored protein sample was the C106DD (Figure 3.5).

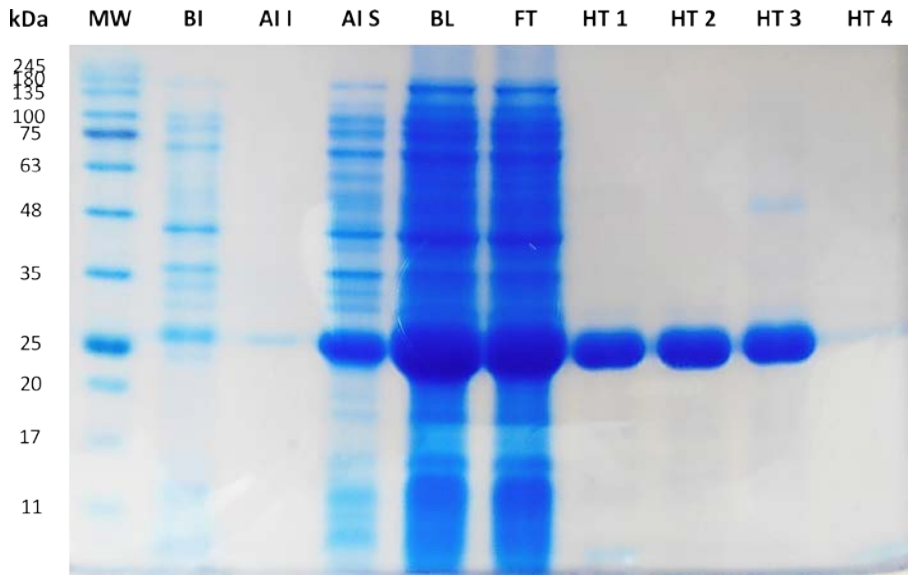


Figure 3.4 | ANALYSIS OF DJ-1 C106DD EXPRESSION AND PURIFICATION BY SDS-PAGE. SDS-PAGE followed by Coomassie staining. MW – molecular weight marker: NZY Colour Protein Marker II (NZYTech); BI – cellular extract obtained before induction of protein expression; AI I – insoluble protein fraction obtained from centrifugation after induction of protein expression; AI S – soluble protein fraction obtained after induction of protein expression; BL – cellular extract loaded on the HisTrap column; FT – flow-through obtained from HisTrap column; HT 1 –fraction 1 eluted from the HisTrap column; HT 2 - fraction 2 eluted from the HisTrap column; HT 3 – fraction 3 eluted from the HisTrap column; HT 4 - fraction 4 eluted from the HisTrap column (Figure 3.2). The sample volume added to each well was 10µL.

The results show that there isn´t any intense band at approximately 22 kDa, which is the molecular weight of DJ-1 monomer in the sample obtained before induction (BI). The same was observed in the insoluble sample recovered after induction of protein expression (AI I). On the other hand, in the soluble sample recovered after induction of protein expression (AI S) an intense band at approximately DJ-1 monomer molecular weight was observed that indicated that DJ-1 was produced in a soluble form and not in inclusion bodies (AI I). In addition, from samples BI to AI it was observed enrichment in DJ-1, which means that after IPTG induction, the cells produced recombinant DJ-1. When a HisTrap column is used to purify the protein it can be seen the presence of the DJ-1 monomer band with high intensity in the sample loaded into the column (BL) that means that DJ-1 was loaded

into the column but there was also an intense band in the flow through (FT) recovered after sample loading into the column, which indicate that a large amount of DJ-1 was also lost. From all the fractions recovered from the protein elution of the HisTrap column, only fraction 4 did not contained an intense band around 22kDa.

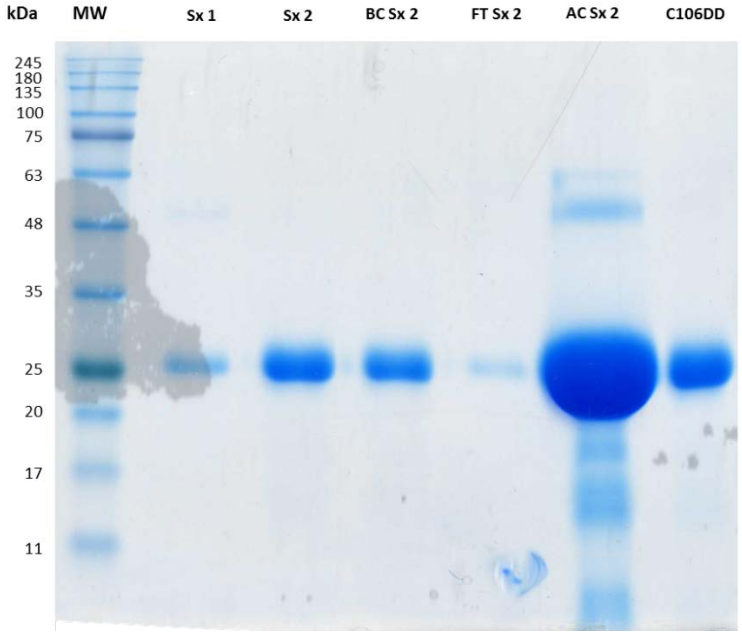


Figure 3.5 | ANALYSIS OF DJ-1 C106DD SIZE EXCLUSION CHROMATOGRAPHY BY SDS-PAGE. SDS-PAGE followed by Coomassie staining. MW – molecular weight marker: NZY Colour Protein Marker II (NZYTech); After loading the fraction HT 3 from the HisTrap elution on a HiLoad 26/60 Superdex 200 prep grade fractions 1 (Sx 1) and 2 (Sx 2) were recovered; BC Sx 2 – sample containing fraction 2 obtained from Superdex 200 before concentration; FT Sx 2 – flow-through obtained from concentration of fraction 2; AC Sx 2 - purified DJ-1 C106DD obtained after concentration of fraction 2; C106DD – sample AC Sx 2 stored in PBS with 10% glycerol at -80°C. The sample volume added to each well was 10µL.

The results obtained from the SDS-PAGE showed an intense band corresponding to DJ-1 (≈22kDa) in fraction 2 obtained from elution of the Superdex 200 column (Sx 2), in the fraction that was concentrated (BC Sx 2), in the fraction obtained after protein concentration (AC Sx 2) and in sample stored (C106DD) (Figure 3.5). In fraction 1 acquired after elution from the Superdex 200 column (Sx 1) this band is much less intense and in the flow through recovered after protein concentration (FT Sx 2) the band is almost undetectable.

For the other proteins being used in this study (DJ-1 WT, C106A and E163K) similar results were obtained (Supplementary Figure 7.4 to 7.6).

3.3. SEQUENCE ANALYSIS BY LC-MS/MS

The protein sequence of DJ-1 WT and mutants was confirmed by LC-MS/MS after protein digestion (see Experimental Procedures – 2.4). The C106DD mutant was identified with a Protein Pilot unused score of 141.28, a sequence coverage of 99.5% (the initial methionine was not identified) (Figure 3.6) and 115 peptides were identified with 95% or more confidence. The loss of the initial methionine is very common since a significant number of intracellular proteins produced in *E. coli* are post- or co-translational modified and the amino-terminal is removed enzymatically by the methionine aminopeptidase [162].

MGSSHHHHHHDYDIPTTENLYFQGHMASKRALVILAKGAEEMETVIPVDVMRRAGIKVTVAGLAGKDPVQC
SRDVVICPDASLEDAKKEGPYDVVVLPGGNLGAQNLSESAAVKEILKEQENRKGLIAAIDDAGPTALLAHEIGFG
SKVTTHPLAKDKMMNGGHYTYSENRVEKDGLILTSRGPGTSFEFALAIVEALNGKEVAAQVKAPLVLKD

Figure 3.6 | SEQUENCE COVERAGE OF DJ-1 MUTANT C106DD. Green letters represent the residues identified in peptides with at least 95% confidence; grey letters represent the unidentified residues. Mutated residues are highlight in yellow.

Similar results were obtained for DJ-1 WT and mutants C106A and E163K (Supplementary Figure 7.7 to 7.9).

3.4. PROTEIN CHARACTERIZATION

After proteins production and purification, they were characterized to ensure protein structure and stability to use in further assays, including neuroprotection and interactomics. Thus, the secondary and quaternary structures of each protein were evaluated by circular dichroism and size exclusion (SE) - high-performance liquid chromatography (HPLC), respectively. The thermal stability was also assessed by thermal shift and the protein´s molecular weight was confirmed by LC-MS. Moreover, protein´s identification and contaminants identification were performed using LC-MS/MS.

3.4.1. SIZE EXCLUSION (SE) – HPLC

Many studies report that the functional structure of DJ-1 is its homodimeric form [66, 69]. Moreover, it has been reported that the DJ-1 mutants in study (C106A,

C106DD and E163K) retain the homodimeric form despite the loss of protein function [90, 97, 99]. Given this, assessment of the quaternary structure of these proteins is essential and thus a SE - HPLC was performed in which the absorbance was measured at 214 nm over elution volume (Figure 3.7).

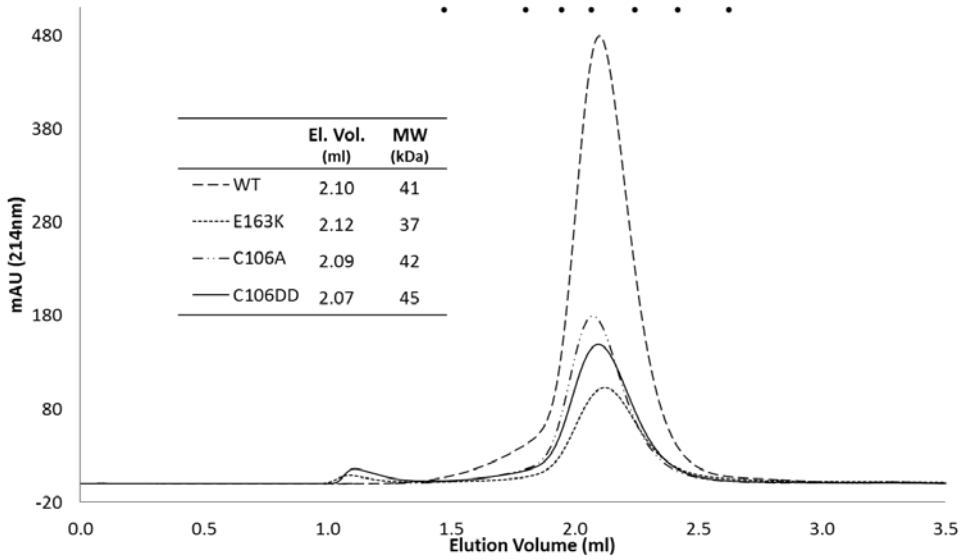


Figure 3.7 | HPLC-SIZE EXCLUSION CHROMATOGRAMS OF PURIFIED DJ-1 WT AND MUTANTS C106A, C106DD AND E163K. Standards molecular weight (black dots) from left to right: Ferritin (440kDa); Aldolase (158kDa); Conalbumin (75kDa); Ovalbumin (44kDa); Carbonic Anhydrase (29kDa); Ribonuclease A (13.7kDa); Aprotinin (6.5kDa). The table shows the proteins´ legend as well as the elution volume at which the protein was retrieved and the corresponding molecular weight. El. Vol. – Elution volume; MW – molecular weight.

With the obtained elution volume (2.07mL), and using the elution volumes of standards, it is possible to calculate the approximate molecular weight of the produced proteins (Supplementary Figure 7.10). The molecular weights obtained were around 41kDa for the WT and C106A, approximately 37 and 45kDa for E163K and C106DD mutants, respectively. These values correspond to the molecular weight reported in the literature for the DJ-1 dimer [68]. The differences observed were considered small taking into account the wide range of the calibration curve (from 6.5kDa to 440kDa) and therefore even the E163K mutant was also considered to be present in the dimeric form. In addition, the results show that in the case of DJ-1 mutants E163K and C106DD there is another peak around 1.1mL and based on the value of molecular weight obtained for these peaks – 1 801kDa and 1 664kDa, respectively – the corresponding proteins seemed to be forming oligomers. This means that the purified proteins DJ-1

E163K and C106DD have not only the protein in the homodimeric form, considered the functional form, but also a small amount of protein in the oligomeric form, a nonfunctional form [69, 92, 163].

3.4.2. LC-MS OF INTACT PROTEINS

The LC-MS/MS analysis of the digested proteins shows that the protein sequence obtained was according with the expected from the vector, and that the initial methionine was not identified. However, this analysis does not indicate if the protein is present on its intact form, or if is a mixture of non-complete sequences. In order to elucidate if the protein has its correct molecular weight, an intact protein analysis was performed by LC-MS (Figure 3.8).

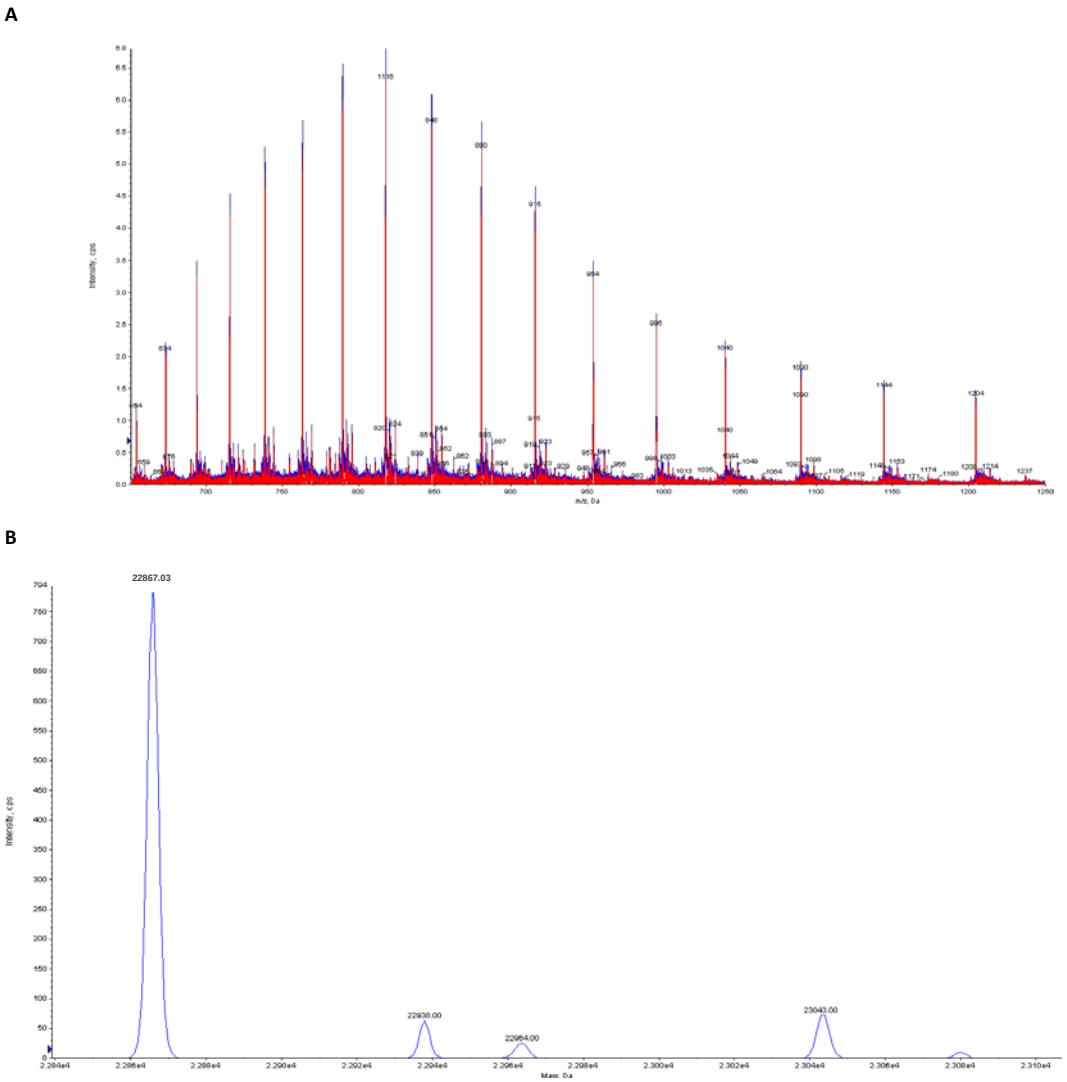


Figure 3.8 | LC-MS ANALYSIS OF INTACT DJ-1 MUTANT C106DD. A - Average spectra of protein charge envelop of intact C106DD (in blue). Deconvoluted mass spectrum of protein charge envelop of intact C106DD (in red). B - Graphic of deconvoluted masses using BioAnalyst™.

From the graphic of the deconvoluted masses and using the BioAnalyst™ Software the value of the predicted molecular weight of the intact protein can be retrieved (Figure 3.8 - B). For the mutant C106DD in study, the value obtained was 22 867.037 Da whereas the value calculated using the Mass Calculator tool, without the initial methionine, was 22 865.953 Da. The values are similar presenting a shift of 1.085 Da.

The same analysis was applied to the other mutants and DJ-1 WT and the values of average, predicted masses, and the mass shift observed were similar (Supplementary Table 7.2).

3.4.3. THERMAL SHIFT

The thermal shift concept is based on protein unfolding due to the increase in the temperature, which allows the exposure of its hydrophobic regions. During this process, the dye binds to these regions and the fluorescence signal increases with the exposure of more hydrophobic regions. A spectrum is obtained and the melting temperature can be determined and used to assess the thermal stability of the purified proteins. To assess if the mutants in study affect the stability of DJ-1, thermal denaturation experiments were performed where the fluorescence intensity was monitored while temperature increased 1°C/min from 25°C to 95°C in a thermal shift assay (Figure 3.9).

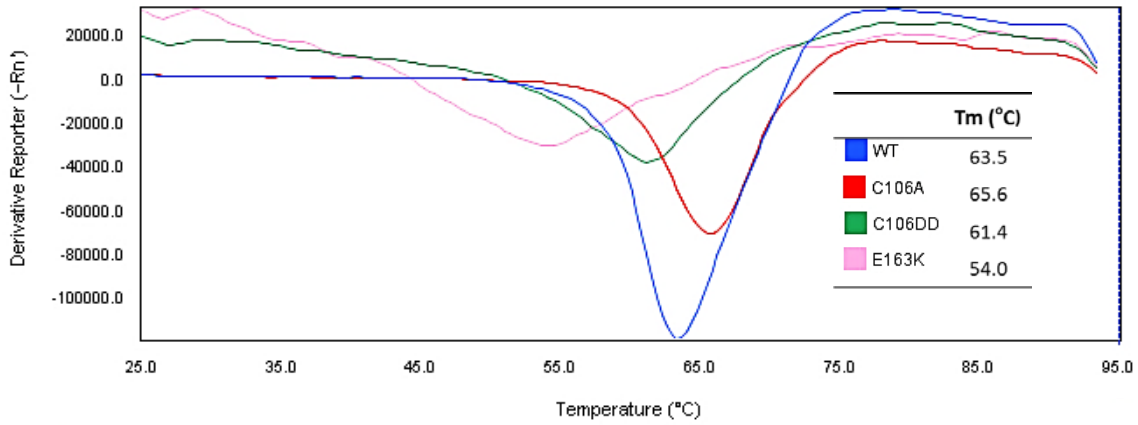


Figure 3.9 | FIRST DERIVATIVE CURVES OF DJ-1 WT AND MUTANTS C106A, C106DD AND E163K OBTAINED FROM THERMAL SHIFT ASSAY. The concentration of the recombinant DJ-1 WT and its mutants was 0.5mg/mL. The temperature at the peak is the melting temperature (T_m). The first derivative curves and melting temperatures were obtained by the normalized curves analysis and transformation as explained in Supplementary data 7.6. The table displays proteins´ legend as well as the T_m obtained for each protein.

The concentration of all the protein samples was adjusted to 0.5mg/mL and the results revealed that the folded structure of DJ-1 WT was relatively stable and protein unfolding was achieved for a melting temperature (T_m – temperature at which 50% of the protein is denatured) of 63.5°C. The mutant presenting the higher thermal stability was the mutation C106A with a T_m of 65.6°C. The most thermally unstable mutant was E163K exhibiting a T_m of 54.0°C and the mutant C106DD presented a T_m similar to the DJ-1 WT displaying a value of 61.4°C. Of note, the reduced signal amplitude of mutants E163K and C106DD compared to DJ-1 WT and mutant C106A. This fact could be associated with the protein concentration used in this assay, since only two different concentrations were tested (0.2 and 0.5mg/mL) and the results showed that the most adequate curves were the ones corresponding to 0.5mg/mL (data not shown). However, no other concentrations were tested due to the reduced amount of mutant proteins available for further studies.

3.4.4. CIRCULAR DICHROISM

In addition to the assessment of the primary and quaternary structure of a protein, the secondary structure was also evaluated. Thus, DJ-1 and DJ-1 mutants secondary structure was determined by circular dichroism (Figure 3.10).

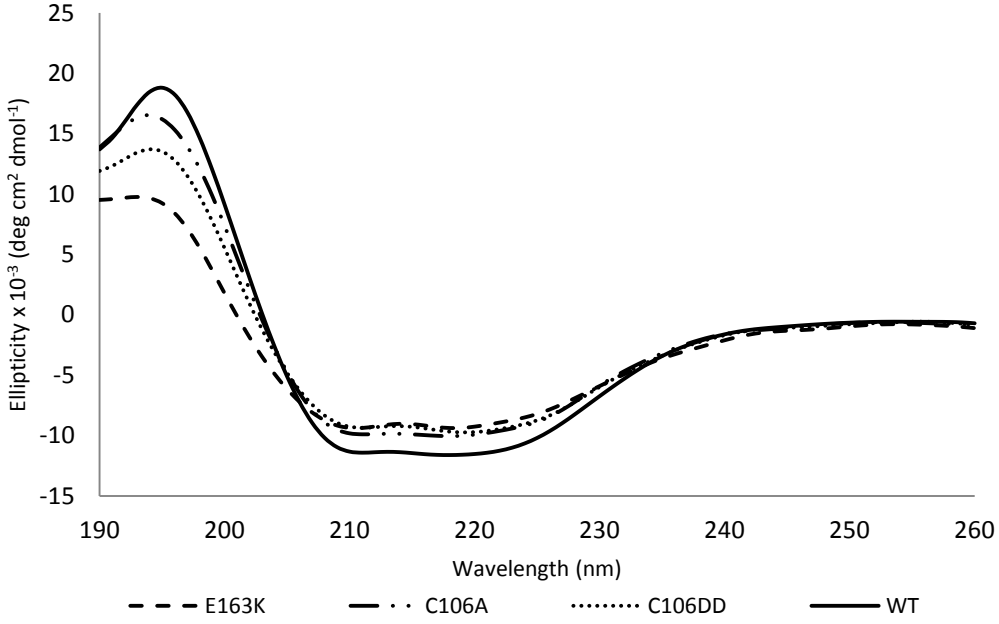


Figure 3.10 | FAR-ULTRAVIOLET CIRCULAR DICHROISM SPECTRA OF DJ-1 WT AND MUTANTS C106A, C106DD AND E163K. CD spectra were obtained under normal conditions at 37°C using a 0.05mm cell. The protein concentrations used were approximately 2.0mg/mL.

At physiologic pH and 37°C the far-UV CD spectra of the DJ-1 mutants showed high similarities in shape to the DJ-1 WT, but exhibited a decreased signal amplitude, with this effect being more pronounced for the E163K mutant. Despite this, all CD spectra displayed a broad negative ellipticity from 208 to 222 nm, consistent with a mixed α/β structure. The secondary structure content was similar for all the mutants, containing 20 to 30% of structure in α-helix, 20 to 30% of structure in β-sheet and approximately 50% without regular secondary structure. For DJ-1 WT, the secondary structure content was higher for the structure in α-helix, with approximately 40%, 20 to 25% of structure in β-sheet and around 40 to 50% without regular secondary structure. These values were calculated by the GlobalWorks software (described in Experimental Procedures – 2.7).

3.4.5. LC-MS/MS ANALYSIS OF DIGESTED SAMPLES

An important feature that needs to be evaluated in the production of a recombinant protein before its use in further biological assays is the determination of contaminants in the proteins samples. In this way, an analysis by LC-MS/MS to the digested samples was performed to identify all the proteins present in each sample (Table 3.1). For this analysis both liquid and gel digestions were made and the results from both were combined.

Table 3.1 | PROTEIN IDENTIFICATION BY LC-MS/MS ANALYSIS OF EACH PURIFIED PROTEIN.

					WT		**C106A**		**C106DD**		**E163K**	
Name	**Accession Number**	**Species**	**MW (Da)**	**Function**	**UP**	US	**UP**	US	**UP**	US	**UP**	US
DJ-1	Q99497	HUMAN	19,891	Stress Response	**113**	165.4	**96**	183.2	**129**	193.5	**112**	165.3
Superoxide dismutase [Cu-Zn]	Q96VL0	CLAPU	15,839	Antioxidant; Oxidoreductase			**2**	3.64	**2**	4	**2**	1.16
Alpha-S1-casein	P02662	BOVIN	24,529	Antioxidant					**2**	4	**2**	3.72
Beta-casein	P02666	BOVIN	25,107	Antioxidant (protease activity)					**1**	2	**1**	2
Glutaredoxin-1	P68688	ECOLI	9,685	Oxidoreductase; cell redox homeostasis	**2**	3.04	**2**	2.67	**2**	2.35	**2**	3.19
Thioredoxin-1	P0AA25	ECOLI	11,807	Oxidoreductase; cell redox homeostasis	**5**	8.64	**5**	8.47	**4**	8.01	**4**	8.06
Glutathione reductase	P06715	ECOLI	48,773	Oxidoreductase					**3**	5.74	**2**	2.65
FMN-dependent NADH-azoreductase	P41407	ECOLI	21,658	Oxidoreductase					**2**	4.09		
Cytochrome b	Q34945	LUMTE	42,882	Oxidoreductase			**2**	1.72				
Hydrogenase-1 small chain	P69739	ECOLI	40,681	Oxidoreductase					**1**	2		

Probable hypoxanthine oxidase XdhD	Q46814	ECOLI	103,519	Oxidoreductase						**1**	2
4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	Q5GRK4	WOLTR	47,381	Oxidoreductase						**1**	2
Aldehyde dehydrogenase 5, mitochondrial	P40047	YEAST	56,693	Oxidoreductase (catabolic process)			**1**	1.54		**1**	2
Ketol-acid reductoisomerase	A4IRH9	GEOTN	37,666	Oxidoreductase					**1**	2	
Protochlorophyllide reductase B, chloroplastic	Q42850	HORVU	42,148	Oxidoreductase	**1**	2	**1**	0.74			
Ferric uptake regulation protein	P0A9A9	ECOLI	16,795	Transcription regulation (repressor)			**2**	4	**2**	4	**5**	10
Transcriptional regulatory protein CusR	P0ACZ8	ECOLI	25,395	Transcription regulator (activator)	**1**	2					
50S ribosomal protein L28	P0A7M2	ECOLI	9,006	Translation (Ribonucleo-protein)			**1**	2			
50S ribosomal protein L11P	A3CSJ5	METMJ	16,317	Translation (RNA binding)	**1**	2	**1**	2			
Beta-lactoglobulin	P02754	BOVIN	19,883	Transporter (retinol binding)					**2**	4.01	**1**	1.32
Fatty acid-binding protein, epidermal	Q01469	HUMAN	15,164	Transporter							**2**	4
Ferrous iron transport protein B	P33650	ECOLI	84,474	Transporter			**2**	4			
Farnesyl diphosphate synthase	P22939	ECOLI	32,160	Transferase			**2**	3.28	**2**	5.08	**1**	2
Cyclomaltodextrin glucanotransferase	P31797	GEOSE	78,923	Transferase (carbohydrate metabolic process)	**1**	0.22	**2**	3.77			
Glycerol-3-phosphate acyltransferase	A6Q218	NITSB	22,430	Transferase (Lipid biosynthesis and metabolism)	**1**						
Serum albumin	P02769	BOVIN	69,293	DNA binding (negative regulation of apoptotic process)							**3**	3.29
Putative two-component response regulator-like APRR6	Q9C9F6	ARATH	86,182	DNA binding	**1**	0.24					
Uncharacterized protein yffS	P76550	ECOLI	29,751	DNA binding							**1**	2
Minor capsid protein 10B	P19727	BPT7	41,830	Caspsid Protein			**1**	2	**1**	2	
Filaggrin-2	Q5D862	HUMAN	248,073	Structural molecule (calcium binding)					**1**	2	
Dermcidin	P81605	HUMAN	11,284	Peptidase			**1**	0.61	**2**	3.92	**2**	3.66
Lysozyme C	P61626	BOVIN	16,537	Hydrolase (lysozyme activity)					**2**	4	**1**	2
High-molecular weight cobalt-containing nitrile hydratase subunit alpha	P21219	RHORH	22,835	Lyase			**1**	0.52			**1**	0.91
Putative lipocalin 1-like protein 1	Q5VSP4	HUMAN	17,918	Binding to lipids					**1**	2.01	**1**	0.78
Uncharacterized protein yaiL	P51024	ECOLI	19,923	Unknown	**1**	2					
Putative molybdate metabolism regulator	P33345	ECOLI	140,939	Unknown							**1**	2

UP – Unique peptides; **US –** Unused score; **CLAPU -** *Claviceps purpurea*; **ECOLI -** *Escherichia coli*; **LUMTE -** *Lumbricus terrestris*; **WOLTR –** *Wolbachia* subsp. *Brugia malayi*; **GEOTN -** *Geobacillus thermodenitrificans*; **HORVU -** *Hordeum vulgare*; **METMJ -** *Methanoculleus marisnigri*; **GEOSE -** *Geobacillus stearothermophilus*; **NITSB –** *Nitratiruptor* sp.; **ARATH -** *Arabidopsis thaliana*; **BPT7 -** *Enterobacteria phage T7*; **RHORH -** *Rhodococcus rhodochrous*.

Proteins were grouped by function where the considered most relevant are presented on the top of the table (Table 3.1). The most significant proteins in the DJ-1 WT sample are the ones that have similar DJ-1 functions, such as antioxidant and oxidoreductase activity. On the other hand, for the DJ-1 mutants the proteins of interest are the ones that have functions that could lead to cell dysfunction and neurodegeneration, like proteins that play a role in transcription and transport. Worth mentioning that most of the protein contaminants are bovin and *E. coli* proteins. The presence of bovin proteins is explained by the use of LB medium in bacterial culture and *E. coli* was the host selected for cloning and expression of the proteins in study therefore proteins from this species could be commonly identified.

3.5. DJ-1 WT MEDIATED NEUROPROTECTION

Previous reports have shown a protective effect of exogenous DJ-1 addition after cells exposure to hydrogen peroxide [164, 165]. To address this issue, the neuroprotective effect of DJ-1 WT and mutants was evaluated in the SH-SY5Y cell line using hydrogen peroxide as oxidative stress agent (Figure 3.11).

The hydrogen peroxide concentrations used were chosen based on the dose-response curve previously performed in SH-SY5Y cells, where a LD50 value of 247.4µM was obtained (data not shown). The results show no significant increase in cell viability in cells treated with hydrogen peroxide in the presence of DJ-1 WT (Figure 3.11).
This indicates that DJ-1 WT was not protecting cells against H_2O_2-induced cell death. However, it is noteworthy that high values of cell viability were observed in the treatment with 100µM and 200µM of H_2O_2 (Figure 3.11) when comparing with previous results obtained in similar assays, where cell viability at 100µM and 200µM of H_2O_2 were around 50% and 30%, respectively (data not shown). Furthermore, the results indicate that cells treated with hydrogen peroxide and in the presence of the protein´s vehicle have an increase in cell viability similar to the effect of the protein, however this increase is not statistically significant (Figure 3.11). Additionally, the protection conferred to cells by DJ-1 WT could not be observed in a significant manner.

Considering that in the current conditions there is a higher cell viability in the presence of H_2O_2, and the neuroprotection previously observed by DJ-1 was not

confirmed, the assays with the mutants were not performed.

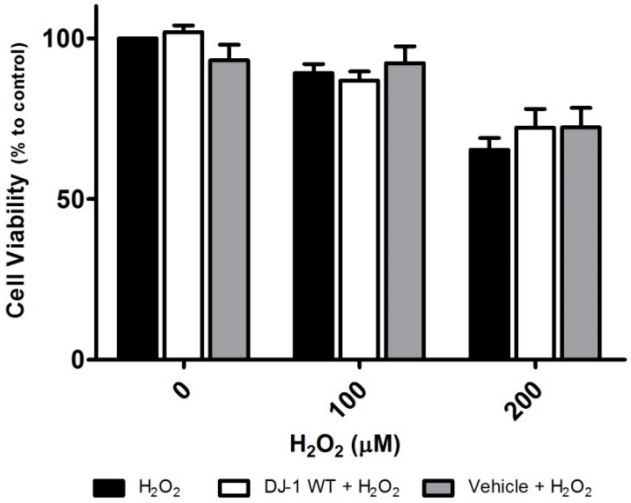


Figure 3.11 | PROTECTIVE EFFECT OF DJ-1 WT AGAINST HYDROGEN PEROXIDE-INDUCED OXIDATIVE STRESS. SH-SY5Y cells were plated and after 4h were treated with different concentrations of H_2O_2 (0, 100 and 200µM) with the addition or not of DJ-1 WT (1µM) or vehicle. Cell viability was measured after 24h and represented to the control (0µM of H_2O_2). Data represent the mean±SEM of three experiments, based on untreated cultures as 100% of viability. No significant results were obtained after statistical analysis using one-way ANOVA.

3.6. DJ-1 WT AND MUTANTS INTERACTOME

The mechanisms by which DJ-1 potentiate cell neuroprotection against oxidative stress are largely unknown. Given this, DJ-1 WT and mutants binding partners were identified after cell exposure to H_2O_2 for different periods (0, 15 and 40min). This approach gives an overview of the dynamic interactome of DJ-1 WT and its mutants under different oxidative stress moments.

A pull-down assay was performed to identify DJ-1 WT and mutants binding partners. For that, each recombinant protein was attached to the resin and the protein extracts recovered from cells exposed to H_2O_2 for different time points were placed in contact with the resin attached proteins. Recombinant DJ-1 proteins and their binding partners were recovered and protein identification was performed by LC-MS/MS. A list of proteins and peptides used in the identification of these proteins was obtained for each recombinant DJ-1 protein. Due to the extent list of results, the tables are shown in a supplementary file ("Proteins Identified by LC-MS_MS.pdf").

Due to the nature of this assay, a condition to assess the proteins that bind unspecifically to the resin (non-specific interactors) needed to be performed. This

comprised the execution of the experiment using the resin without any recombinant protein attached. In order to eliminate this non-specific interactors from further analysis, proteins were selected based on the relative peptide query (rPQ) [166], which is the number of peptides identified in each condition divided by the number of peptides identified in the control condition (resin). Proteins with rPQ of 2 or greater were considered as specifically co-purifying with DJ-1 WT and mutants.

To analyze the extensive lists of recombinant DJ-1 proteins´ binding partners obtained several tools were used, such as Venn Diagrams and Gene Ontology enRIchment anaLysis and visualization tool (GOrilla).

3.6.1. BINDING PARTNERS QUALITATIVE ASSESSMENT

Venn diagrams were used to help visualize the number of shared and unique proteins in each condition over the time and among different conditions (qualitative assessment) (see Experimental Procedures 2.11). With this intention, two main analyses were performed comparing the recombinant DJ-1 proteins in all and in each time point (Figure 3.12). Moreover, the number of shared and unique DJ-1 C106DD binding partners was evaluated over the different time points (0,15 and 40 minutes) (Figure 3.13). Similar analyses were performed for the other DJ-1 mutants (Supplementary Figure 7.18).

The comparison between DJ-1 WT and mutants C106A, C106DD and E163K among all the time points revealed that 213 proteins were shared by all the recombinant proteins. Worth to mention that an overall decrease of the proteins identified for the C106A mutant was observed in almost every analyses. This fact is explained by the decreased number of proteins identified by LC-MS/MS (“Proteins Identified by LC-MS_MS.pdf”) in the analysis performed for this mutant. Regarding the analysis of all the recombinant proteins in each time point (0, 15 and 40min) it was shown that at 0min, 8 proteins were shared among all the recombinant proteins, at 15min that number increased to 54 and at 40min, 25 proteins were shared. Considering the number of unique proteins identified, it should be noted the decrease in the unique proteins of the C106DD mutant (170, 9 and 38 at 0, 15 and 40min, respectively) and on the other hand, the increase in the unique proteins of DJ-1 WT (33, 74 and 126 at 0, 15 and 40min, respectively).

Moreover, the C106DD mutant was reported to mimic DJ-1 WT in its oxidized form that commonly occurs under stress conditions. Therefore, a qualitative assessment of the binding partners identified for this mutant in each time point was performed (Figure 3.13).

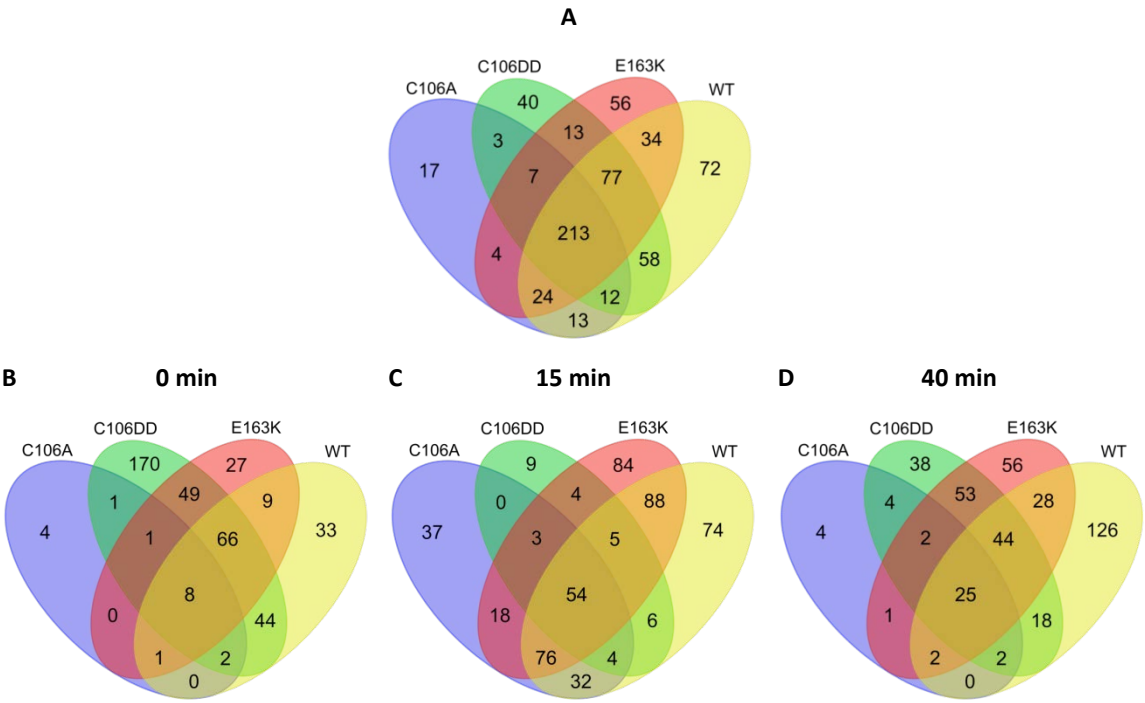


Figure 3.12| VENN DIAGRAMS OF THE IDENTIFIED BINDING PARTNERS IN DJ-1 WT AND MUTANTS FOR ALL AND EACH TIME POINT. A- Venn diagram showing the distribution of shared and unique binding partners in all time points (0, 15 and 40min) of H_2O_2 stimulation in DJ-1 WT and mutants C106A, C106DD and E163K; with a combined protein number of 643. B- Venn diagram showing the distribution of shared and unique binding partners in the 0min condition. The combined protein number is 415. C - Venn diagram showing the distribution of shared and unique binding partners in the 15min of H_2O_2 stimulation. The combined protein number is 494. D - Venn diagram showing the distribution of shared and unique binding partners in the 40min of H_2O_2 stimulation. The combined protein number is 403.

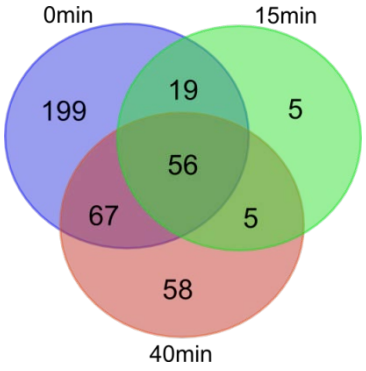


Figure 3.13| VENN DIAGRAM ILLUSTRATING THE NUMBER OF SHARED OR UNIQUE C106DD MUTANT BINDING PARTNERS AMONG THE DIFFERENT TIME POINTS. There are 56 common binding partners of C106DD in all time points of H_2O_2 stimulation and the combined protein number is 409.

The results obtained over the different conditions revealed that 56 proteins were commonly identified and a large variation in the unique proteins identified for the different time points was seen, with 199, 5 and 58 proteins identified for the 0, 15 and 40min, respectively.

Further combinations, with more targeted analyses, were performed using Venn diagrams to facilitate the analysis of results obtained for each recombinant protein (Supplementary Figure 7.14 to 7.18). Additionally, the proteins identified for DJ-1 WT in the described pull-down assay were compared to the proteins identified in the immunoprecipitation performed in previous work (Supplementary Figure 7.19).

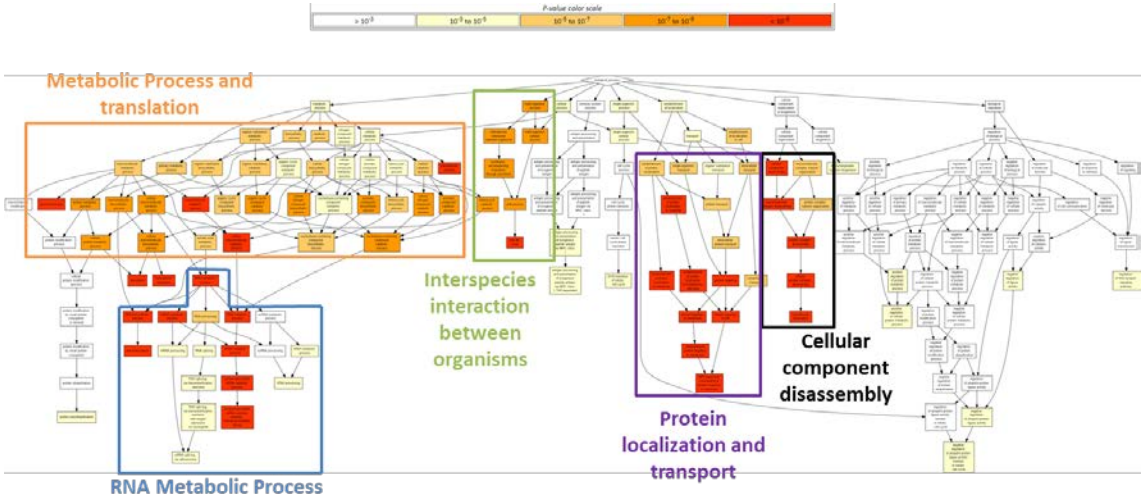
3.6.2. GO ENRICHMENT ANALYSIS

An important analysis was to evaluate whether the overall lists of binding partners obtained were significantly enriched in any particular GO terms based on biological processes. For that Gene Ontology enRIchment anaLysis and visuaLizAtion tool (GOrilla) was used (see Experimental Procedures 2.12), which provides a clear view of the relations between the enriched GO terms in the form of tree diagram.

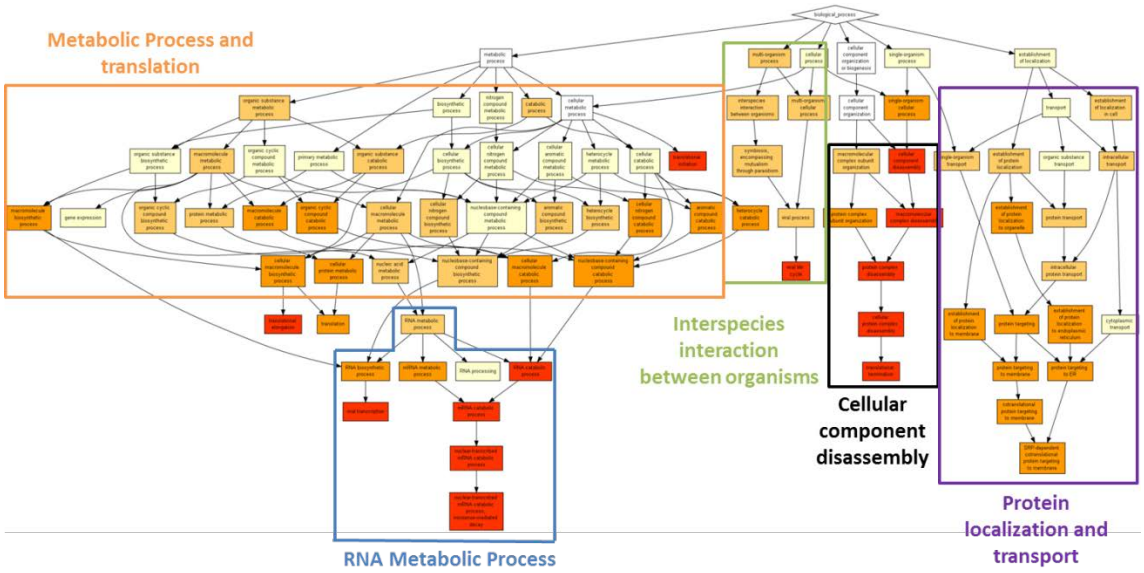
The GO enrichment analysis of the binding partners identified in all time points for DJ-1 WT, and C106DD, C106A and E163K mutants identified 100, 80, 59 and 78 significant enriched GO terms, respectively (Figure 3.14). These GO terms were sorted into 6 to 8 major biological processes, usually common to all the recombinant proteins, and further into a larger number of subgroups. Several subgroups were considered relevant in this analysis, considering the enrichment values and variations among conditions, and thus 5 main groups of subgroups were formed for a simplified visualization of the results - “metabolic process and translation”, “protein localization and transport”, “RNA metabolic process”, “cellular component disassembly” and “interspecies interaction between organisms” (Figure 3.14 – colored boxes).

As observed this type of analysis generates large amounts of results and information to be evaluated. So, the major groups were assessed taking into account the number of subgroups with higher GO terms enrichment and with the greater variations among recombinant proteins and time points.

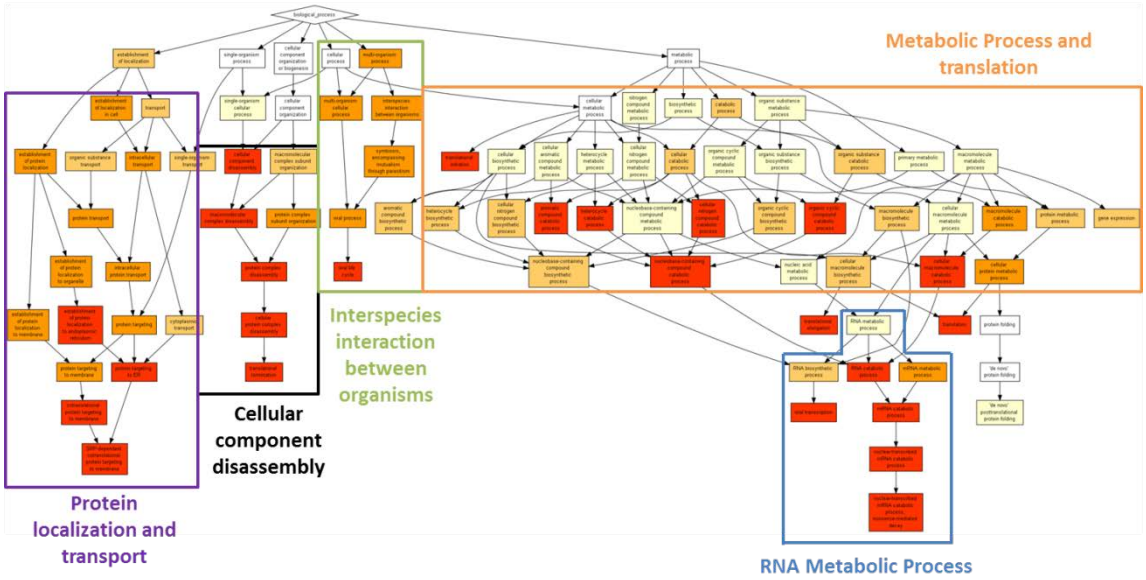
A


Metabolic Process and translation
Interspecies interaction between organisms
Cellular component disassembly
Protein localization and transport
RNA Metabolic Process

B


Metabolic Process and translation
Interspecies interaction between organisms
Cellular component disassembly
Protein localization and transport
RNA Metabolic Process

C



D

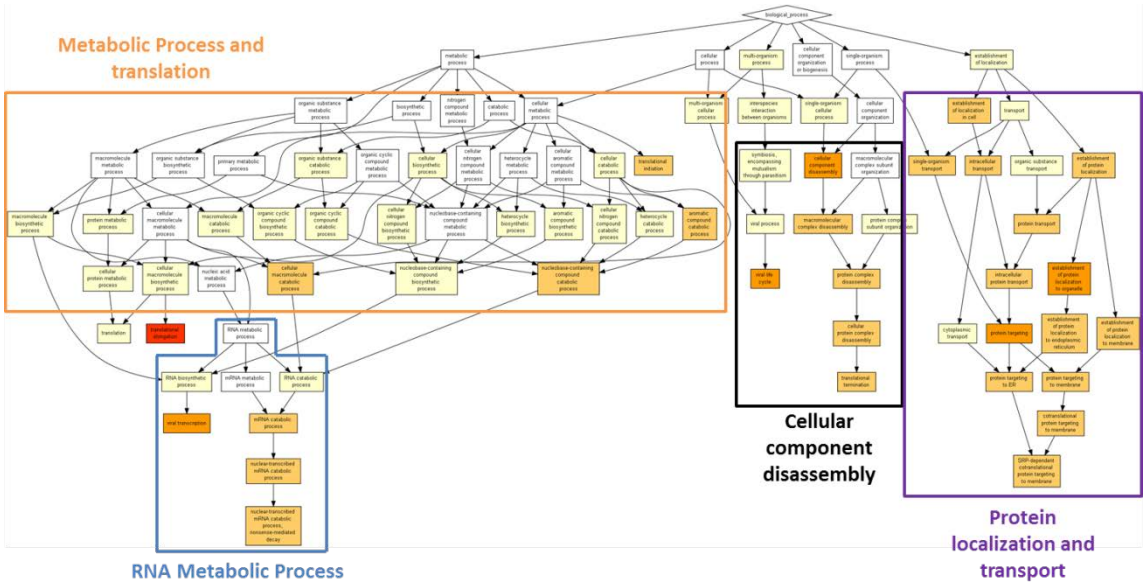


Figure 3.14| GENE ONTOLOGY (GO) ENRICHMENT OF THE IDENTIFIED BINDING PARTNERS FOR EACH RECOMBINANT PROTEIN. The binding partners analyzed were those corresponding to all time points performed for each recombinant protein. The main groups of GO terms considered relevant were highlighted using colored boxes. A to D - GO enrichment of the identified binding partners for DJ-1 WT and C106DD, E163K and C106A mutants, respectively.

In this way, the group of "RNA metabolic processes" was selected for further analysis and discussion. Comparisons of this group between DJ- WT and mutants revealed a larger number of subgroups with a high degree of GO terms enrichment identified for DJ-1 WT (Figure 3.15). Therefore, a targeted GO enrichment analysis was performed for DJ-1 WT in the different time points (Figure 3.16).

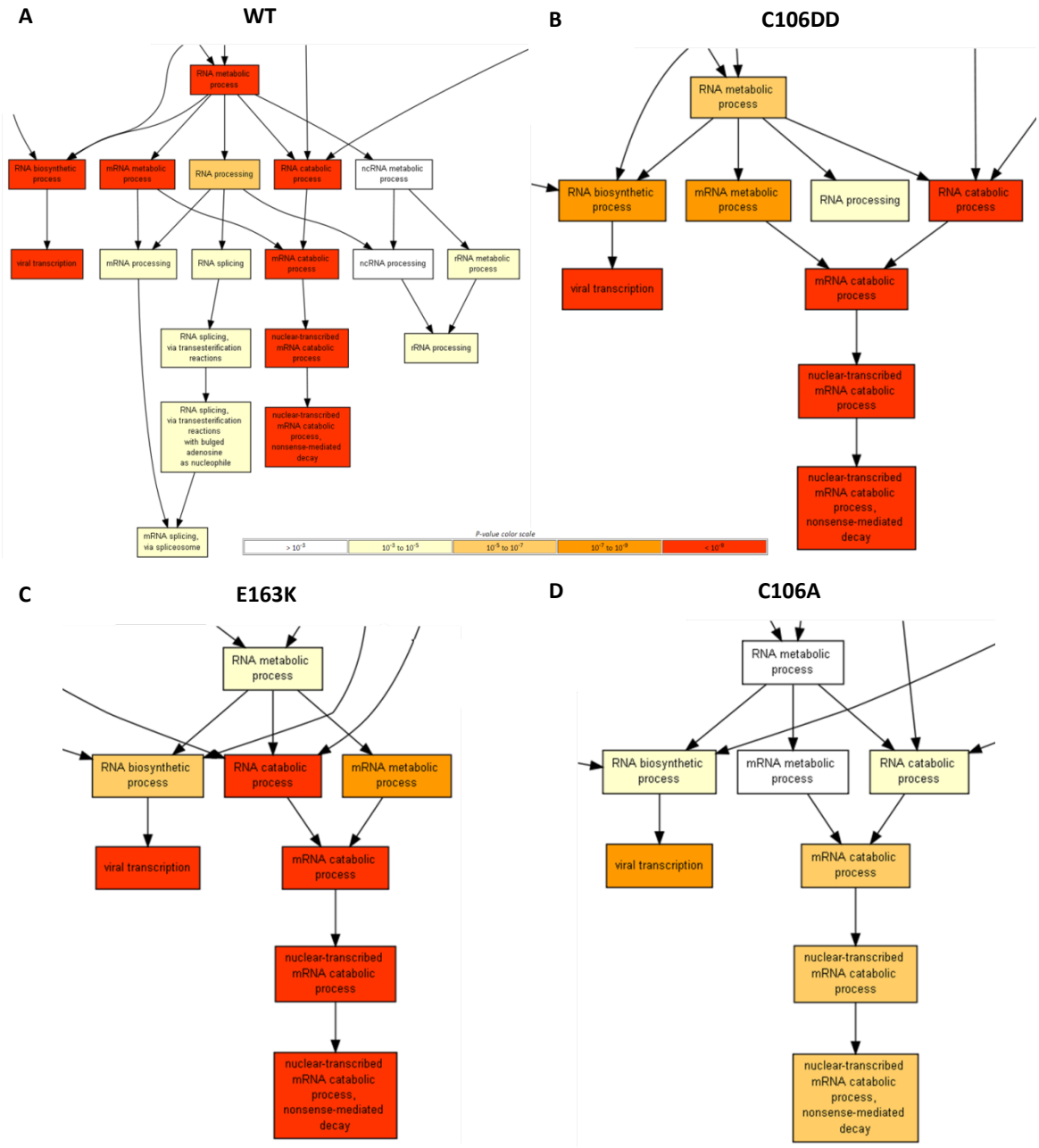


Figure 3.15| GENE ONTOLOGY (GO) ENRICHMENT OF THE RNA METABOLIC PROCESS GROUP FOR EACH RECOMBINANT PROTEIN. A to D - Targeted view of the RNA metabolic process group highlighted by blue boxes in Figure 3.14 for DJ-1 WT, C106DD, E163K and C106A mutants, respectively.

Examining the "RNA metabolic process" enrichment, it was observed an increase in the enrichment degree at 15 minutes of H_2O_2 stimulation when comparing with the control condition (0 min), followed by a decrease to 40 minutes after stimulation.

Similar analyses among the different time points were performed for the DJ-1 mutants (Supplementary Figure 7.20 to 7.22).

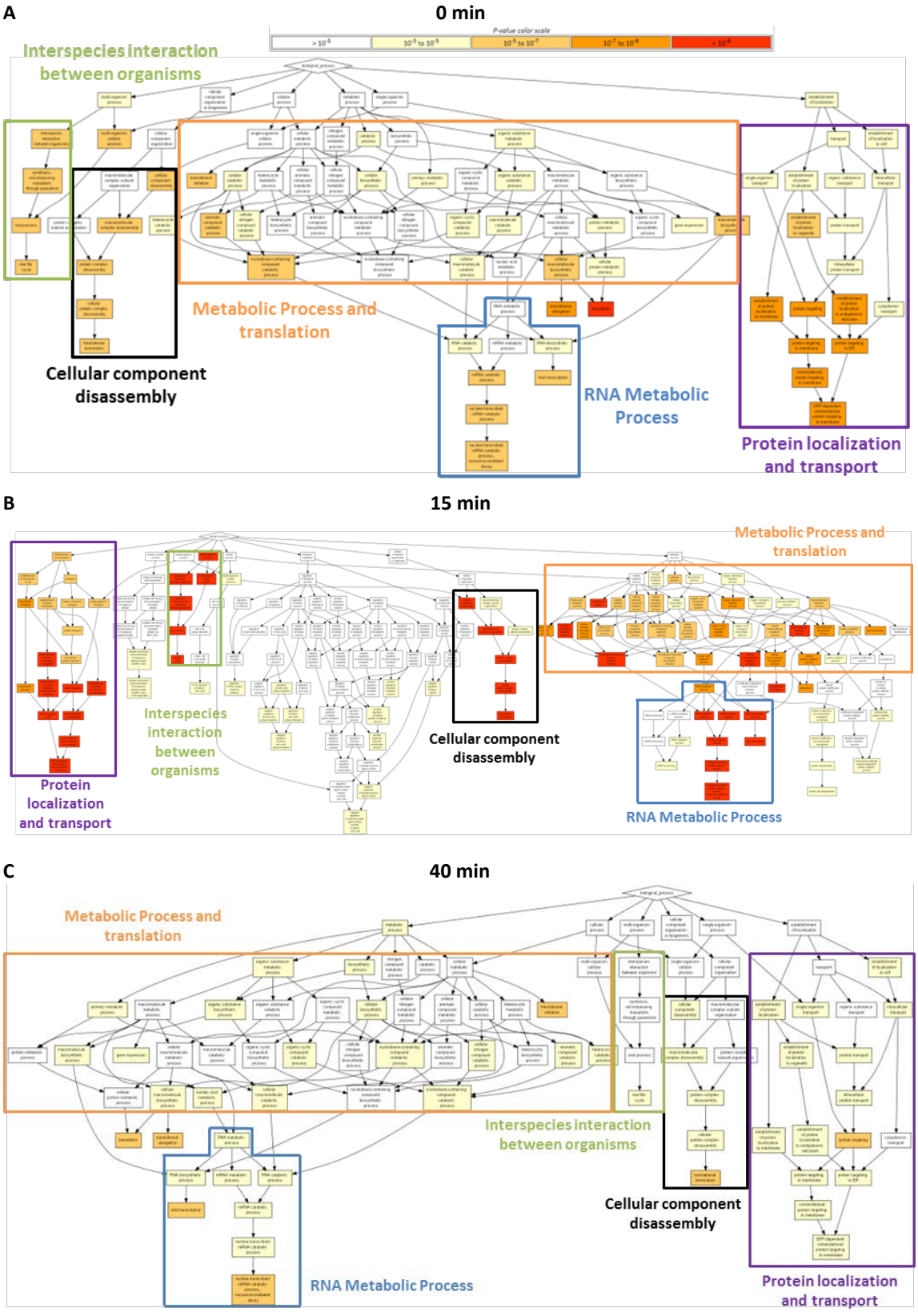


Figure 3.16| GENE ONTOLOGY (GO) ENRICHMENT OF THE IDENTIFIED BINDING PARTNERS FOR DJ-1 WT IN EACH TIME POINT. The main groups of GO terms considered relevant were highlighted using colored boxes. A to C - GO enrichment of the identified binding partners for DJ-1 WT at 0, 15 and 40 min, respectively.

3.6.3. QUANTITATIVE ANALYSIS BY SWATH

Focus was given to DJ-1 WT binding partners and a list of identified interactors in the RNA metabolic process group was obtained (supplementary file available in digital support – "DJ-1 WT binding partners _RNA Metabolic Process.pdf"). In this list 83 proteins were identified and an analysis was performed to determine which proteins from this list could be quantified by SWATH–MS. Briefly, this approach generates a complete recording of the fragment ion spectra of all the analytes in a sample within a determined m/z window and then acquired data from the chromatographic peak of each peptide is further used to determine the relative quantity of a certain protein.

The 42 DJ-1 WT binding partners that were quantifiable by SWATH are shown in Table 3.2 along with the fold increase values represented in log scale and normalized for the control condition (normal condition – 0 minutes). Additionally, a visual representation of these results is presented in Figure 3.17.

Table 3.2 | QUANTITATIVE ANALYSIS OF DJ-1 BINDING PARTNERS BY SWATH-MS.

Protein name	Accession	15 min (log10)	40 min (log10)
U5 small nuclear ribonucleoprotein 200 kDa helicase	O75643|U520_HUMAN	-0.16	0.11
ATP-dependent RNA helicase DDX3X	O00571|DDX3X_HUMAN	-0.15	0.01
Splicing factor U2AF 65 kDa subunit	P26368|U2AF2_HUMAN	-0.08	-0.11
40S ribosomal protein S4, X isoform	P62701|RS4X_HUMAN	0.05	-0.05
Asparagine--tRNA ligase, cytoplasmic	O43776|SYNC_HUMAN	0.05	0.02
Small nuclear ribonucleoprotein Sm D2	P62316|SMD2_HUMAN	0.06	-0.02
Histone-binding protein RBBP4	Q09028|RBBP4_HUMAN	0.10	0.08
40S ribosomal protein S3	P23396|RS3_HUMAN	0.12	0.09
Small nuclear ribonucleoprotein-associated protein N	P63162|RSMN_HUMAN	0.12	-0.01
Heterogeneous nuclear ribonucleoprotein H	P31943|HNRH1_HUMAN	0.15	0.03
40S ribosomal protein S3a	P61247|RS3A_HUMAN	0.16	0.04
General transcription factor II-I	P78347|GTF2I_HUMAN	0.17	0.28
Heterogeneous nuclear ribonucleoprotein D0	Q14103|HNRPD_HUMAN	0.19	0.08
40S ribosomal protein S27	P42677|RS27_HUMAN	0.19	0.06
Polypyrimidine tract-binding protein 1	P26599|PTBP1_HUMAN	0.19	0.13
40S ribosomal protein SA	P08865|RSSA_HUMAN	0.21	0.12
Aspartate--tRNA ligase, cytoplasmic	P14868|SYDC_HUMAN	0.22	0.03
40S ribosomal protein S2	P15880|RS2_HUMAN	0.22	0.09
Heat shock cognate 71 kDa protein	P11142|HSP7C_HUMAN	0.23	0.30
60S ribosomal protein L9	P32969|RL9_HUMAN	0.23	0.04
40S ribosomal protein S15a	P62244|RS15A_HUMAN	0.23	0.08
Poly(rC)-binding protein 2	Q15366|PCBP2_HUMAN	0.24	0.10
Arginine--tRNA ligase, cytoplasmic	P54136|SYRC_HUMAN	0.24	0.14
Poly(rC)-binding protein 1	Q15365|PCBP1_HUMAN	0.25	0.11
40S ribosomal protein S6	P62753|RS6_HUMAN	0.26	0.02
40S ribosomal protein S24	P62847|RS24_HUMAN	0.28	0.15
Heterogeneous nuclear ribonucleoproteins A2/B1	P22626|ROA2_HUMAN	0.29	0.11
40S ribosomal protein S5	P46782|RS5_HUMAN	0.29	0.26
60S ribosomal protein L23a	P62750|RL23A_HUMAN	0.30	0.07
40S ribosomal protein S10	P46783|RS10_HUMAN	0.31	0.22

60S ribosomal protein L7	P18124|RL7_HUMAN	0.32	0.06
Heterogeneous nuclear ribonucleoprotein F	P52597|HNRPF_HUMAN	0.32	0.25
60S ribosomal protein L7a	P62424|RL7A_HUMAN	0.33	0.08
40S ribosomal protein S16	P62249|RS16_HUMAN	0.34	0.27
40S ribosomal protein S13	P62277|RS13_HUMAN	0.34	0.12
40S ribosomal protein S7	P62081|RS7_HUMAN	0.34	0.24
60S ribosomal protein L4	P36578|RL4_HUMAN	0.36	0.18
60S ribosomal protein L14	P50914|RL14_HUMAN	0.36	0.08
60S ribosomal protein L11	P62913|RL11_HUMAN	0.36	0.19
60S ribosomal protein L6	Q02878|RL6_HUMAN	0.37	0.12
40S ribosomal protein S23	P62266|RS23_HUMAN	0.38	0.20
40S ribosomal protein S18	P62269|RS18_HUMAN	0.42	0.26

Analyzing the results obtained in Figure 3.17, two main groups of proteins are very distinctive. One of them is represented by three proteins - the U5 small nuclear ribonucleoprotein 200 kDa helicase, the ATP-dependent RNA helicase DDX3X and the splicing factor U2AF 65 kDa subunit – where it was seen lower quantities at 15 minutes of stimulus compared to the control condition (0min of stimulus). Another characteristic group comprises the large set of proteins mainly containing 40S and 60S ribosomal proteins, which were presented in a significant amount at 15min with more than two-fold increase in comparison to the control.

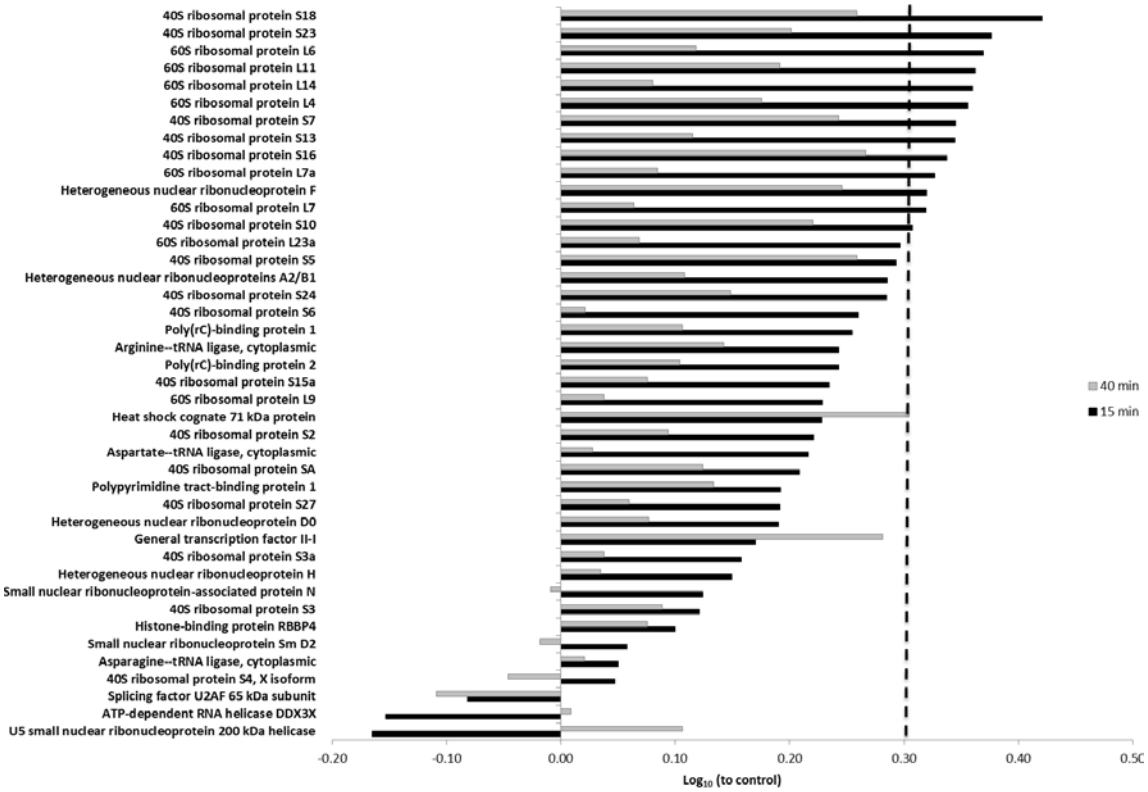


Figure 3.17 | FOLD-INCREASE OF DJ-1 WT BINDING PARTNERS IN OXIDATIVE STRESS CONDITIONS. The relative quantitative binding at 15 and 40 minutes of oxidative stress stimulus are presented in a log scale and were normalized for the control (normal conditions – 0 min). The dashed line represents the 2 fold-increase (0.301 in the log scale).

4. DISCUSSION

Parkinson´s disease is a progressive neurodegenerative disorder that can be sporadic or familial [8, 9]. However, the etiology and pathology of this disorder is largely unknown [2, 10]. The study of the mechanisms of monogenic forms of PD has provided insight into the pathophysiology of this disease. Among these genes, DJ-1 has been associated with PD due to the variety of functions related to neuroprotection [44]. This protein participates in the response to oxidative stress acting as an oxidative stress sensor in which the cysteine residue 106 plays an essential role [80]. Mutations associated to DJ-1 lead to loss of protein function and ultimately development of PD [94]. Given all of the previous facts, the main goals of this study were to evaluate the neuroprotective effect of the DJ-1 mutations E163K (natural mutation), C106A and C106DD (engineered mutations) and to characterize the dynamic interactome of these mutants and DJ-1 WT.

DJ-1 WT and mutants were produced and characterized to be used in further studies of neuroprotection against hydrogen peroxide-induced cell death and dynamic interactome. The coding DNA sequence of DJ-1 and its mutants was sequenced by Sanger DNA sequencing and all the mutations were confirmed. The molecular weight of all the mutants was approximately 43kDa (Figure 3.4, Figure 3.5, Figure 3.7, Supplementary Figure 7.4 to 7.6) corresponding to the dimeric form, which is an important feature of these proteins [69, 90, 97, 99] and the molecular weight of the monomers of all the proteins (Figure 3.8 and Supplementary Table 7.2) was in accordance with the described (approximately 22kDa) [6]. Also, the results from the intact protein analysis showed the presence of other peaks in addition to the peak corresponding to the intact protein (Figure 3.8 - B and Supplementary Figure 7.11 - B). Some of these peaks can be explained by modifications in the intact protein, like oxidations or even addition of amino acid residues, such as methionine. For instance, for the C106DD mutant the peak with a MW of 23 043.54Da (mass shift of 178 Da comparing to the MW of the intact protein) could be due to the occurrence of three oxidations in the intact protein in addition to the initial non-cleaved residue of methionine. A similar peak with a mass shift of 178 Da comparing to the MW of the intact protein (at 22 885.32Da) was obtained for the C106A mutant (Supplementary Figure 7.11 - B) and the same explanation could be applied. For the E163K mutant, the presence of three additional peaks in the graphic of deconvoluted masses at 22

770.90Da, 22 786.45Da and 22 915.82Da (mass shift of 33, 49 and 178Da comparing to the MW of the intact protein, respectively) (Supplementary Figure 7.11 - D) could be explained by two and three oxidations in the intact protein, and three oxidations in addition to one more residue of methionine, respectively. Furthermore, the amino-acid sequence of the purified proteins was confirmed, in which only the initial methionine was not identified due to a natural post-translational modification that occurs in *E. coli*, the host used in protein production (Figure 3.6 and Supplementary Figure 7.7 to 7.9).

The CD spectrum of DJ-1 WT revealed a well folded polypeptide with a considerable amount in α-helix content of approximately 40%, 20 to 25% structure in β-sheet and around 40 to 50% without regular secondary structure (Figure 3.10), results that are in accordance with previous reports from CD analysis and also from crystallographic analysis [69, 90, 92, 103, 163]. The secondary structure content of DJ-1 mutants in study was similar and in accordance with previous reports [90, 92, 167], with 20 to 30% of structure in α-helix, 20 to 30% of structure in β-sheet and approximately 50% without regular secondary structure (Figure 3.10). Furthermore, the decrease in signal amplitude observed for the mutants could result from a minor perturbation in the structure of each one of the mutants. Given this, the results suggest that none of the DJ-1 mutants in study significantly alters the secondary structure of DJ-1 in solution.

The thermal stability of DJ-1 mutants was not significantly altered, except for the E163K mutant, compared to the DJ-1 WT with all the purified proteins presenting similar melting temperatures (Figure 3.9), which means that the mutations impact on DJ-1 structure are not very significant, as expected [45, 99], at least for the analyzed parameters. Noteworthy the overall reduced signal amplitude for the E163K mutant curve that could be due to the concentration used, since this parameter is considered a fundamental variable to improve data quality [168, 169]. Therefore, other concentration values, preferably higher, should be tested to improve the signal amplitude and thereby obtain a more accurate melting temperature value.

The final purified proteins samples have some contaminants in their composition, mainly derived from the LB medium and *E. coli* used in bacterial culture and as a host, respectively, during protein production (Table 3.1). Moreover, some

proteins were identified as human proteins. These proteins could be identified as human, by the software (ProteinPilot), based on the peptides confidence and sequence coverage, but in this case, it is more likely that they are derived from a similar organism, such as bovin, since bovin and not human proteins were used in proteins´ production and purification experiments.

Protein contaminants with similar DJ-1 WT functions were identified that could influence DJ-1 function in cells and its effect in neuroprotection. However, it should be considered the low level of identification on these proteins. Given this, the proteins with similar antioxidant function and with similar molecular weight (Alpha-S1-casein and Beta-casein) were identified with only two and one peptide, respectively, in the C106DD and E163K mutants (Table 3.1). In contrast, DJ-1 mutants were identified with 129 and 112 peptides, clearly showing a large enrichment in comparison to these contaminants. Moreover, the contaminants are barely seen on the SDS-PAGE gel (Figure 3.5 – lane C106DD, Supplementary Figure 7.5 – lane C106A and Supplementary Figure 7.6 – lane E163K). In DJ-1 WT, two proteins with similar oxidoreductase and cell redox homeostasis functions (Glutaredoxin-1 and Thioredoxin-1) were identified, with 2 and 5 peptides, respectively. However, bands with small molecular weight that could correspond to these contaminants did not appear on the SDS-PAGE gel (data not shown) and DJ-1 WT was identified with 113 peptides, showing a strong enrichment of DJ-1 compared to the contaminants.

Previous work was done by our research group to evaluate DJ-1-mediated neuroprotection in SH-SY5Y cells treated with different concentrations of hydrogen peroxide (0, 25, 50, 100 and 200µM). In these experiments, exogenous addition of recombinant DJ-1 had a protective effect on H_2O_2-induced cell death. Taking these results into account, one of the goals of this project was to assess DJ-1 WT and mutants C106A, C106DD and E163K neuroprotection in SH-SY5Y treated with H_2O_2. However, in DJ-1 mediated neuroprotection assays it was not observed a protective effect by DJ-1 in cells stimulated with hydrogen peroxide. This result could be associated with the low level of cell death that makes difficult to observe a significant increase in cell viability when DJ-1 is present. Moreover, an increment in cell viability was observed in cells in the presence of the vehicle (PBS with 10% glycerol), which could be explained by the already reported hydroxyl radical scavenger function of

glycerol in cells [170, 171]. This fact could be also affecting the correct assessment of DJ-1 neuroprotective effect, since the neuroprotection that was observed in some cases could be resulting from the protein´s buffer and not from DJ-1 itself. Besides protein´s buffer, DJ-1 WT was produced in its function form (dimer) (Figure 3.7), with the desired thermal stability (Figure 3.9) and secondary structure (Figure 3.10) and by the analysis of the contaminants´ function these should not negatively influence DJ-1 neuroprotective function (Table 3.1). A proposed solution could be to produce and store the proteins in a buffer that does not affect cell viability and research is being made to this end. Furthermore, in order to understand the high values of cell viability observed, a new hydrogen peroxide was tested but the results remained very similar. Therefore, another explanation for these values was the cell line quality and development that could be impaired somehow. Nonetheless, further work is being done involving the research for a new and suitable buffer to store the protein and improvements in cell line development and quality, in which the acquisition of a new cell line could be a possibility.

DJ-1 has been reported as a multifunctional protein [43, 45, 67] and it has been demonstrated that interactions between proteins are essential in the biological processes in cells [134]. Thus, the study of the interactome, defined as a group of protein-to-protein interactions that occurs in a cell under a specified condition and time point, is considered important in the translation of molecular mechanisms into cellular functions [132, 134]. Given this, DJ-1 WT and mutants interactome under normal and stress conditions in different time points - dynamic interactome – was evaluated through pull-down assays of each one of the recombinant proteins with protein extracts recovered from cells not subjected to oxidative stimulus (0 minutes condition), and upon H_2O_2 stimulus after 15 and 40 minutes.

The qualitative assessment of the binding partners identified for DJ-1 WT and mutants C106DD, E163K and C106A performed by Veen diagrams revealed that the majority of binding partners identified are shared among DJ-1 WT and mutants demonstrating that these binding partners are constant in DJ-1 WT and mutants, not suffering possible interferences caused by DJ-1 mutations. This also reveals that the analysis of primary, secondary and quaternary structure, which indicated minor differences, is in accordance with a high degree of shared binding partners. They do

not suffer interference from the mutations in the interaction with DJ-1 (Figure 3.12 - A). Analyzing each time point separately, an increase in proteins binding to DJ-1 WT, C106A and E163K mutants was observed after 15min of stimulus, suggesting that this state of oxidation promotes unique protein interactions in these recombinant proteins (Figure 3.12 – B to D). For the C106DD mutant a decrease in unique interactors was seen for 15 and 40min (oxidative stress conditions) in comparison to the 0min (normal conditions). Since C106DD mutant seems to mimic DJ-1 WT active form in cells, this could mean that when DJ-1 WT is in its active form promotes interactions with proteins similar to the ones present in cells under normal conditions rather than the ones present under stress conditions (Figure 3.12 – B to D). This fact was confirmed by the time course evaluation of C106DD binding partners (Figure 3.13), showing that a higher number of interactors was identified under normal conditions. Additionally, this analysis showed that the lowest number of unique interactors of C106DD was at 15 minutes of oxidative stress. This suggests that for this mutant an interaction with proteins under increased oxidative stress conditions was promoted over lower oxidative stress conditions. It is worth to mention that the number of total proteins identified for each recombinant protein influences this type of analysis, as seen for the C106A mutant analysis.

A GO term enrichment analysis is a common tool used in large lists of genes or proteins, allowing the visualization and identification of the major biological processes to which the binding partners identified are associated. In this respect, from the major groups identified, the RNA metabolic process emerged as an interesting target in this study. This was because the enrichment GO terms values among DJ-1 proteins was higher for DJ-1 WT, which revealed that this group in DJ-1 WT would have unique proteins associated with this processes in comparison to the other mutants (Figure 3.15). In addition, RNA metabolic processes comprise cellular chemical reactions and pathways associated with RNA, namely transcription and translation processes and DJ-1 have been associated with the regulation of RNA metabolism [172]. Evaluating the enrichment degree in this group for DJ-1 WT among the different times, it was seen a significant increase in enrichment after 15min of oxidative stimulus in comparison with normal (0 min) and higher oxidative stress (40 min) conditions (Figure 3.16). This suggests that after 15min of stimulus more proteins involved in RNA metabolic

processes bound to DJ-1 WT and after 40min of stimulus this number decreased in comparison to the 15min condition, presenting a similar GO enrichment pattern to the normal conditions (0min).

Using a SWATH analysis it was possible to quantify some of the DJ-1 WT binding partners identified for the RNA metabolic processes in the different oxidative stress conditions. Among these proteins there are several 40 and 60S ribosomal proteins and specifically the U5 small nuclear ribonucleoprotein 200 kDa helicase that have already been identified in the IP previously performed by our research group, and are present in the shared proteins identified both in the IP and PD procedures (Supplementary Figure 7.19). Interestingly, the SWATH analysis revealed a large set of 40S and 60S ribosomal proteins being identified with a two-fold increase or higher at 15 minutes of oxidative stimulus (Figure 3.17). Worth mention that the 15 and 40 minutes of stimulus were chosen in order to represent an initial and a more extended state of oxidation, respectively. This means that DJ-1 WT in initial oxidative stress conditions greatly binds to these ribosomal proteins, which are involved in translational processes. A previous study demonstrated that under oxidative stress conditions the formation of aggregate between proteins and RNA in ribosomes was promoted [173]. Being these proteins structural elements of the ribosome, DJ-1 WT binding could directly stabilize the ribosome structure preventing the formation of protein aggregates with RNA in the ribosomes and promote the interaction with the RNA leading to a proper translation process. This ribosome stabilization could be associated with a chaperone like function previously reported that helps other proteins or complexes to maintain their structure under oxidative stress conditions. For these set of ribosomal proteins after 40 minutes of stimulus the amount of binding protein to DJ-1 WT was reduced in comparison to the 15 minutes condition. This suggests that the proposed function for DJ-1 WT is more important in an initial state of oxidation probably because it directly disturbs translation with immediately consequences in protein production as response to oxidative stress. Therefore, due to the importance of this process DJ-1 WT needs to act in early states of oxidative stress conditions.

Another protein identified in the IP assay and quantified in this study was the U5 small nuclear ribonucleoprotein 200 kDa helicase, also known as BRR2. In contrast to the observed for the 40 and 60S ribosomal proteins, a lower quantity of this

interactor was observed after 15 minutes of stimulus (Figure 3.17). This protein is a member of the DExH/Dbox helicase family and plays a key role in the activation of the spliceosome, which is involved in pre-mRNA splicing [174]. BRR2 unwinds the U4/U6 snRNA duplices promoting the formation of the spliceosome catalytic center [175]. In this way, the observed decrease in DJ-1 WT interaction with U5 small nuclear ribonucleoprotein 200 kDa helicase at an initial state of oxidative stress conditions can be explained as DJ-1 WT releasing the BRR2 protein, and promoting its location at the spliceosome to interact with the U4 snRNA to displace it from the U6 snRNA, a essential step to create the active site for splicing catalysis [176]. In a more extended state of oxidative stress conditions (40 min) the amount of BRR2 protein binding to DJ-1 WT increases in comparison to the control condition. It has been reported that DExH/Dbox helicases, like BRR2, are also involved in remodeling RNA–protein complexes formed in the spliceosome and ribosome, for instance [177, 178]. Given this, at this level of oxidation DJ-1 WT could promote the interaction with BRR2 protein in order to modulate its function and act in remodeling the interactions between RNA and proteins for a long term effect in oxidative stress regulation.

In addition, a new and interesting DJ-1 WT interactor, not identified in the IP assay previously mentioned, and involved in apoptotic processes was identified and quantified in this study - ATP-dependent RNA helicase DDX3X. It was seen that after 15 minutes of stimulus a lower amount of this protein was interacting with DJ-1 WT (Figure 3.17). As a member of the DEAD-box RNA helicase family, functions correlated with RNA processing, transport and also translation processes have been reported for this protein [179]. More interesting was the described function as antiapoptotic protein, negatively regulating the apoptotic processes via death domain receptors [180]. It was shown that DDX3X protein form a death antagonizing signaling complex with other antiapoptotic proteins at death receptors to inhibit apoptotic signals and therefore ensuring the maintenance of unstimulated death receptors in a quiescent state. In this line of thought, the observed decrease in DJ-1 WT interaction with ATP-dependent RNA helicase DDX3X at an initial state of oxidative stress conditions can be explained as DJ-1 WT release of DDX3X protein promoting its interaction with death receptors and as a result inhibit the activation of apoptotic signaling pathways. In a more extended state of oxidative stress conditions (40 min) the amount of DDX3X

protein binding to DJ-1 WT returns to values near the ones observed under normal conditions. At this point, DJ-1 WT may adjust its form of action in response to the oxidative stress events that are occurring in cells promoted by this extended level of oxidation. In this way, DJ-1 WT could redirect its functions like chaperone, antioxidant and ROS scavenger activities, which represent more directed forms of response to oxidative stress.

In summary, two proteins were confirmed to be DJ-1 WT interactors by two different assays with further quantification of these proteins accomplished in this study. A new DJ-1 WT interactor was identified, the ATP-dependent RNA helicase DDX3X. For this protein, the quantification by SWATH was important and provided a more accurate binding rate profile for this protein, since based on the number of peptides identified no modifications seemed to occur among the different time points when in fact a decrease in the binding rate to DJ-1 WT at 15 minutes of oxidative stress was observed (supplementary file - “Proteins Identified by LC-MS_MS.pdf”). Nonetheless, a higher number of SWATH analysis should be performed in order to increase the confidence in the results obtained and improve the number of proteins that could be quantified using this method. In addition, the identified interactors should be validated and further research should be performed to confirm and obtain more insights into these suggested DJ-1 molecular mechanisms of action against oxidative stress.

Noteworthy that one of the main goals of this project was to characterize the dynamic interactome and quantify the binding partners of DJ-1 WT and mutants E163K, C106A and C106D. Through the identification of the differences between DJ-1 WT and mutants binding partners obtain insight into the molecular mechanisms by which DJ-1 exerts its neuroprotective function and which functions are impaired in DJ-1 mutants. However, due to some problems in the samples processing this goal was somehow compromised and these analyses will be repeated. Furthermore, to take into account the losses through the sample processing steps, an internal standard will be added to the samples, such as GFP. In this way, a thorough analysis will be made giving a more accurate understanding on the loss of DJ-1 mutants binding partners and furthermore elucidations into DJ-1 WT functions that were impaired by these mutations would be obtained.

5. CONCLUSIONS

The main goal of this study was to obtain insight into of DJ-1 WT function in cells through the characterization of the dynamic interactome of DJ-1 WT and mutants E163K, C106A and C106DD, under normal and stress conditions. A focus was given to a natural mutation and two engineered mutations, specifically at the C106 residue due to its importance in DJ-1 functions against oxidative stress. To accomplish this, DJ-1 WT and mutants were produced for further use in SH-SY5Y cells to assess their neuroprotective effect and evaluate the changes in their interactome under stress conditions.

The previously mentioned DJ-1 mutations were successfully accomplished and proteins´ characterization revealed that these proteins were in the dimeric form, had similar secondary structures contents and identical thermal stabilities, as previously described. Furthermore, no significant contaminants were identified in the purified proteins that could greatly influence these proteins´ functions in cells. The characterization steps are particularly relevant when the protein will be used for biological assays, such as survival and interactomics assays, where not only the primary sequence is important, but mainly the maintenance of the structure that will be recognized by the biological system.

Contrarily to what was previously observed, DJ-1 WT-mediated neuroprotection was not confirmed in cells stimulated with hydrogen peroxide.

Through the dynamic interactomic analysis, several new putative binding partners for DJ-1 WT and mutants E163K, C106DD and C106A were identified, which can contribute with insight in revealing DJ-1 WT functions against oxidative stress. However, validation of these interactors is required as well as further assessment of these interactors' function and influence in the regulation of DJ-1 functions.

Furthermore, a recent developed label-free method to quantify proteins allowed a better characterization of the continuously changing protein-protein interactions that occurs in the dynamic interactome, significantly contributing to the insight previously obtained. Thus, further elucidations can be obtained into Parkinson's Disease pathology and potential new targets for PD prevention and therapy could have been identified.

6. REFERENCES

1. Parkinson, J., *An Essay on the Shaking Palsy. 1817.* The Journal of Neuropsychiatry and Clinical Neurosciences, 2002. **14**(2): p. 223-236.
2. Wirdefeldt, K., et al., *Epidemiology and etiology of Parkinson's disease: a review of the evidence.* European journal of epidemiology, 2011. **26**: p. 1-58.
3. Crosiers, D., et al., *Parkinson disease: insights in clinical, genetic and pathological features of monogenic disease subtypes.* Journal of chemical neuroanatomy, 2011. **42**(2): p. 131-141.
4. Coppedè, F., *Genetics and Epigenetics of Parkinson's Disease.* The Scientific World Journal, 2012. **2012**.
5. Hauser, D.N. and T.G. Hastings, *Mitochondrial dysfunction and oxidative stress in Parkinson's disease and monogenic parkinsonism.* Neurobiology of Disease, 2012.
6. Hulleman, J.D., et al., *Destabilization of DJ-1 by familial substitution and oxidative modifications: implications for Parkinson's disease.* Biochemistry, 2007. **46**(19): p. 5776-5789.
7. Ferrer, I., *Early involvement of the cerebral cortex in Parkinson's disease: convergence of multiple metabolic defects.* Progress in neurobiology, 2009. **88**(2): p. 89-103.
8. Sassi, C., *Genetics of Parkinson Disease, Etiology and Pathophysiology of Parkinson's Disease*. 2011: InTech.
9. Thomas, B. and M.F. Beal, *Parkinson's disease.* Human molecular genetics, 2007. **16**(R2): p. R183-R194.
10. Klein, C. and A. Westenberger, *Genetics of Parkinson's Disease.* Cold Spring Harbor Perspectives in Medicine, 2012. **2**(1).
11. Moore, D.J., et al., *Association of DJ-1 and parkin mediated by pathogenic DJ-1 mutations and oxidative stress.* Human molecular genetics, 2005. **14**(1): p. 71-84.
12. Cookson, M.R., *Evolution of Neurodegeneration.* Current Biology, 2012. **22**(17): p. R753-R761.
13. Horowitz, M. and J. Greenamyre, *Gene–environment interactions in Parkinson's disease: the importance of animal modeling.* Clinical Pharmacology & Therapeutics, 2010. **88**(4): p. 467-474.
14. Burbulla, L.F. and R. Krüger, *Converging environmental and genetic pathways in the pathogenesis of Parkinson's disease.* Journal of the neurological sciences, 2011. **306**(1): p. 1-8.
15. Singer, T.P. and R.R. Ramsay, *Mechanism of the neurotoxicity of MPTP: an update.* FEBS letters, 1990. **274**(1): p. 1-8.
16. Corti, O., S. Lesage, and A. Brice, *What genetics tells us about the causes and mechanisms of Parkinson's disease.* Physiological Reviews, 2011. **91**(4): p. 1161-1218.
17. Farrer, M.J., *Genetics of Parkinson disease: paradigm shifts and future prospects.* Nature Reviews Genetics, 2006. **7**(4): p. 306-318.
18. Vila, M. and S. Przedborski, *Targeting programmed cell death in neurodegenerative diseases.* Nature Reviews Neuroscience, 2003. **4**(5): p. 365-375.
19. Dauer, W. and S. Przedborski, *Parkinson's disease: mechanisms and models.* Neuron, 2003. **39**(6): p. 889-909.

20. Abou-Sleiman, P.M., M.M.K. Muqit, and N.W. Wood, *Expanding insights of mitochondrial dysfunction in Parkinson's disease.* Nature Reviews Neuroscience, 2006. **7**(3): p. 207-219.
21. Gao, H.M. and J.S. Hong, *Gene–environment interactions: key to unraveling the mystery of Parkinson's disease.* Progress in neurobiology, 2011. **94**(1): p. 1-19.
22. Di Monte, D.A., M. Lavasani, and A.B. Manning-Bog, *Environmental factors in Parkinson's disease.* Neurotoxicology, 2002. **23**(4): p. 487-502.
23. Liu, R., et al., *Caffeine Intake, Smoking, and Risk of Parkinson Disease in Men and Women.* American journal of epidemiology, 2012. **175**(11): p. 1200-1207.
24. Houlden, H. and A.B. Singleton, *The genetics and neuropathology of Parkinson's disease.* Acta neuropathologica, 2012: p. 1-14.
25. Lesage, S. and A. Brice, *Role of Mendelian genes in "sporadic" Parkinson's disease.* Parkinsonism & Related Disorders, 2012. **18**: p. S66-S70.
26. Martin, I., V.L. Dawson, and T.M. Dawson, *Recent advances in the genetics of Parkinson's disease.* Annual review of genomics and human genetics, 2011. **12**: p. 301-325.
27. Kumar, K.R., A. Djarmati-Westenberger, and A. Grünewald, *Genetics of Parkinson's Disease.* Semin Neurol, 2011. **31**(05): p. 433-440.
28. Nuytemans, K., et al., *Genetic etiology of Parkinson disease associated with mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 genes: a mutation update.* Human mutation, 2010. **31**(7): p. 763-780.
29. Lesage, S. and A. Brice, *Parkinson's disease: from monogenic forms to genetic susceptibility factors.* Human molecular genetics, 2009. **18**(R1): p. R48-R59.
30. Dehay, B., et al., *Loss of P-type ATPase ATP13A2/PARK9 function induces general lysosomal deficiency and leads to Parkinson disease neurodegeneration.* Proceedings of the National Academy of Sciences, 2012. **109**(24): p. 9611-9616.
31. Ross, O.A., et al., *Genomic investigation of α-synuclein multiplication and parkinsonism.* Annals of Neurology, 2008. **63**(6): p. 743-750.
32. Brockmann, K. and T. Gasser, *Genetic Basis of Monogenic Forms of Parkinson Disease.* eLS - John Wiley & Sons, Ltd, 2010.
33. Lee, B.D., V.L. Dawson, and T.M. Dawson, *Leucine-rich repeat kinase 2 (LRRK2) as a potential therapeutic target in Parkinson's disease.* Trends in pharmacological sciences, 2012.
34. Pihlstrøm, L. and M. Toft, *Genetic variability in SNCA and Parkinson's disease.* neurogenetics, 2011. **12**(4): p. 283-293.
35. Shimura, H., et al., *Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase.* Nature genetics, 2000. **25**(3): p. 302-305.
36. Kitada, T., et al., *Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism.* nature, 1998. **392**(6676): p. 605-608.
37. Dawson, T.M. and V.L. Dawson, *The role of parkin in familial and sporadic Parkinson's disease.* Movement disorders, 2010. **25**(S1): p. S32-S39.
38. Klein, C. and K. Lohmann-Hedrich, *Impact of recent genetic findings in Parkinson's disease.* Current opinion in neurology, 2007. **20**(4): p. 453.
39. Valente, E.M., et al., *Hereditary early-onset Parkinson's disease caused by mutations in PINK1.* science, 2004. **304**(5674): p. 1158-1160.

40. Kawajiri, S., et al., *Genetic mutations and functions of PINK1.* Trends in pharmacological sciences, 2011. **32**(10): p. 573-580.
41. Mills, R.D., et al., *Biochemical aspects of the neuroprotective mechanism of PTEN-induced kinase-1 (PINK1).* Journal of neurochemistry, 2008. **105**(1): p. 18-33.
42. Wilhelmus, M.M.M., et al., *Involvement and interplay of Parkin, PINK1 and DJ1 in neurodegenerative and neuroinflammatory disorders.* Free Radical Biology and Medicine, 2012.
43. Bonifati, V., B.A. Oostra, and P. Heutink, *Linking DJ-1 to neurodegeneration offers novel insights for understanding the pathogenesis of Parkinson's disease.* Journal of molecular medicine, 2004. **82**(3): p. 163-174.
44. Hauser, D.N. and M.R. Cookson, *Astrocytes in Parkinson's disease and DJ-1.* Journal of neurochemistry, 2011. **117**(3): p. 357-358.
45. Wilson, M.A., *The role of cysteine oxidation in DJ-1 function and dysfunction.* Antioxidants & redox signaling, 2011. **15**(1): p. 111-122.
46. Kahle, P.J., J. Waak, and T. Gasser, *DJ-1 and prevention of oxidative stress in Parkinson's disease and other age-related disorders.* Free radical biology & medicine, 2009. **47**(10): p. 1354.
47. Ramirez, A., et al., *Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase.* Nature genetics, 2006. **38**(10): p. 1184-1191.
48. Park, J.S., et al., *Pathogenic effects of novel mutations in the P-type ATPase ATP13A2 (PARK9) causing Kufor-Rakeb syndrome, a form of early-onset parkinsonism.* Human mutation, 2011. **32**(8): p. 956-964.
49. Yao, Z. and N.W. Wood, *Cell death pathways in Parkinson's disease: role of mitochondria.* Antioxidants & redox signaling, 2009. **11**(9): p. 2135-2149.
50. Valente, E.M., et al., *Molecular pathways in sporadic PD.* Parkinsonism & Related Disorders, 2012. **18**: p. S71-S73.
51. Alberio, T. and M. Fasano, *Proteomics in Parkinson's disease: An unbiased approach towards peripheral biomarkers and new therapies.* Journal of biotechnology, 2011. **156**(4): p. 325-337.
52. Varçin, M., et al., *Oxidative stress in genetic mouse models of Parkinson's disease.* Oxidative Medicine and Cellular Longevity, 2012. **2012**.
53. Rego, A.C. and C.R. Oliveira, *Mitochondrial dysfunction and reactive oxygen species in excitotoxicity and apoptosis: implications for the pathogenesis of neurodegenerative diseases.* Neurochemical research, 2003. **28**(10): p. 1563-1574.
54. Schapira, A.H.V., *Mitochondria in the aetiology and pathogenesis of Parkinson's disease.* The Lancet Neurology, 2008. **7**(1): p. 97-109.
55. Gandhi, S. and N.W. Wood, *Molecular pathogenesis of Parkinson's disease.* Human molecular genetics, 2005. **14**(18): p. 2749-2755.
56. Henchcliffe, C. and M.F. Beal, *Mitochondrial biology and oxidative stress in Parkinson disease pathogenesis.* Nature Clinical Practice Neurology, 2008. **4**(11): p. 600-609.
57. Onyango, I.G., *Mitochondrial dysfunction and oxidative stress in Parkinson's disease.* Neurochemical research, 2008. **33**(3): p. 589-597.

58. de Vries, R.L.A. and S. Przedborski, *Mitophagy and Parkinson's disease: Be eaten to stay healthy.* Molecular and Cellular Neuroscience, 2012.
59. Batlevi, Y. and A.R. La Spada, *Mitochondrial autophagy in neural function, neurodegenerative disease, neuron cell death, and aging.* Neurobiology of disease, 2011. **43**(1): p. 46-51.
60. Palikaras, K. and N. Tavernarakis, *Mitophagy in neurodegeneration and aging.* Frontiers in Genetics, 2012. **3**.
61. Büeler, H., *Mitochondrial dynamics, cell death and the pathogenesis of Parkinson's disease.* Apoptosis, 2010. **15**(11): p. 1336-1353.
62. Perfeito, R., T. Cunha-Oliveira, and A. Cristina Rego, *Revisiting oxidative stress and mitochondrial dysfunction in the pathogenesis of Parkinson's disease-resemblance to the effect of amphetamine drugs of abuse.* Free Radical Biology and Medicine, 2012.
63. Miller, R.L., et al., *Oxidative and inflammatory pathways in Parkinson's disease.* Neurochemical research, 2009. **34**(1): p. 55-65.
64. Mizuno, Y., et al., *Progress in the pathogenesis and genetics of Parkinson's disease.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2008. **363**(1500): p. 2215-2227.
65. Nakamura, T., D.H. Cho, and S.A. Lipton, *Redox regulation of protein misfolding, mitochondrial dysfunction, synaptic damage, and cell death in neurodegenerative diseases.* Experimental Neurology, 2012.
66. Anderson, P.C. and V. Daggett, *Molecular Basis for the Structural Instability of Human DJ-1 Induced by the L166P Mutation Associated with Parkinson's Disease†.* Biochemistry, 2008. **47**(36): p. 9380-9393.
67. Hatano, T., et al., *Pathogenesis of familial Parkinson's disease: new insights based on monogenic forms of Parkinson's disease.* Journal of neurochemistry, 2009. **111**(5): p. 1075-1093.
68. Honbou, K., et al., *The crystal structure of DJ-1, a protein related to male fertility and Parkinson's disease.* Journal of Biological Chemistry, 2003. **278**(33): p. 31380-31384.
69. Wilson, M.A., et al., *The 1.1-Å resolution crystal structure of DJ-1, the protein mutated in autosomal recessive early onset Parkinson's disease.* Proceedings of the National Academy of Sciences, 2003. **100**(16): p. 9256-9261.
70. Chen, J., L. Li, and L.-S. Chin, *Parkinson disease protein DJ-1 converts from a zymogen to a protease by carboxyl-terminal cleavage.* Human molecular genetics, 2010. **19**(12): p. 2395-2408.
71. Lev, N., et al., *Role of DJ-1 in Parkinson's disease.* Journal of molecular neuroscience, 2006. **29**(3): p. 215-225.
72. Kim, S.J., et al., *Nuclear translocation of DJ-1 during oxidative stress-induced neuronal cell death.* Free Radical Biology and Medicine, 2012.
73. Lee, J.-y., et al., *Human DJ-1 and its homologs are novel glyoxalases.* Human molecular genetics, 2012. **21**(14): p. 3215-3225.
74. Lev, N., et al., *Knocking Out DJ-1 Attenuates Astrocytes Neuroprotection Against 6-Hydroxydopamine Toxicity.* Journal of Molecular Neuroscience, 2013: p. 1-9.
75. Malgieri, G. and D. Eliezer, *Structural effects of Parkinson's disease linked DJ-1 mutations.* Protein Science, 2008. **17**(5): p. 855-868.

76. Kim, J.-h., et al., *DJ-1 facilitates the interaction between STAT1 and its phosphatase, SHP-1, in brain microglia and astrocytes: A novel anti-inflammatory function of DJ-1.* Neurobiology of Disease, 2013. **60**(0): p. 1-10.
77. Ren, H., et al., *Oxidized DJ-1 interacts with the mitochondrial protein BCL-XL.* Journal of Biological Chemistry, 2011. **286**(40): p. 35308-35317.
78. Van Der Brug, M.P., et al., *RNA binding activity of the recessive parkinsonism protein DJ-1 supports involvement in multiple cellular pathways.* Proceedings of the National Academy of Sciences, 2008. **105**(29): p. 10244-10249.
79. Logan, T., L. Clark, and S.S. Ray, *Engineered disulfide bonds restore chaperone-like function of DJ-1 mutants linked to familial Parkinson's disease.* Biochemistry, 2010. **49**(27): p. 5624-5633.
80. Canet-Avilés, R.M., et al., *The Parkinson's disease protein DJ-1 is neuroprotective due to cysteine-sulfinic acid-driven mitochondrial localization.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(24): p. 9103-9108.
81. Abou-Sleiman, P.M., D.G. Healy, and N.W. Wood, *Causes of Parkinson's disease: genetics of DJ-1.* Cell and tissue research, 2004. **318**(1): p. 185-188.
82. Ren, H., et al., *L166P mutant DJ-1 promotes cell death by dissociating Bax from mitochondrial Bcl-XL.* Molecular Neurodegeneration, 2012. **7**(1): p. 40.
83. Bonifati, V., *Autosomal recessive parkinsonism.* Parkinsonism & Related Disorders, 2012. **18**: p. S4-S6.
84. Alvarez-Castelao, B., et al., *Reduced protein stability of human DJ-1/PARK7 L166P, linked to autosomal recessive Parkinson disease, is due to direct endoproteolytic cleavage by the proteasome.* Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2011.
85. Tang, B., et al., *Association of PINK1 and DJ-1 confers digenic inheritance of early-onset Parkinson's disease.* Human molecular genetics, 2006. **15**(11): p. 1816-1825.
86. Anvret, A., et al., *DJ-1 Mutations are Rare in a Swedish Parkinson Cohort.* The open neurology journal, 2011. **5**: p. 8.
87. Guo, J.F., et al., *Mutation analysis of Parkin, PINK1, DJ-1 and ATP13A2 genes in Chinese patients with autosomal recessive early-onset Parkinsonism.* Movement Disorders, 2008. **23**(14): p. 2074-2079.
88. Clark, L.N., et al., *Analysis of an early-onset Parkinson's disease cohort for DJ-1 mutations.* Movement disorders, 2004. **19**(7): p. 796-800.
89. Baulac, S., et al., *Dimerization of Parkinson's disease-causing DJ-1 and formation of high molecular weight complexes in human brain.* Molecular and Cellular Neuroscience, 2004. **27**(3): p. 236-246.
90. Lakshminarasimhan, M., et al., *Structural impact of three Parkinsonism-associated missense mutations on human DJ-1.* Biochemistry, 2008. **47**(5): p. 1381-1392.
91. Moore, D.J., et al., *A missense mutation (L166P) in DJ-1, linked to familial Parkinson's disease, confers reduced protein stability and impairs homo-oligomerization.* Journal of neurochemistry, 2003. **87**(6): p. 1558-1567.
92. Olzmann, J.A., et al., *Familial Parkinson's disease-associated L166P mutation disrupts DJ-1 protein folding and function.* Journal of Biological Chemistry, 2004. **279**(9): p. 8506-8515.

93. Blackinton, J., et al., *Effects of DJ-1 mutations and polymorphisms on protein stability and subcellular localization.* Molecular brain research, 2005. **134**(1): p. 76-83.
94. Bonifati, V., et al., *Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism.* Science, 2003. **299**(5604): p. 256-259.
95. Abou-Sleiman, P.M., et al., *The role of pathogenic DJ-1 mutations in Parkinson's disease.* Annals of neurology, 2003. **54**(3): p. 283-286.
96. Zucchelli, S., et al., *Aggresome-forming TTRAP mediates pro-apoptotic properties of Parkinson's disease-associated DJ-1 missense mutations.* Cell Death & Differentiation, 2008. **16**(3): p. 428-438.
97. Waak, J., et al., *Oxidizable residues mediating protein stability and cytoprotective interaction of DJ-1 with apoptosis signal-regulating kinase 1.* Journal of Biological Chemistry, 2009. **284**(21): p. 14245-14257.
98. Annesi, G., et al., *DJ-1 mutations and parkinsonism-dementia-amyotrophic lateral sclerosis complex.* Annals of neurology, 2005. **58**(5): p. 803-807.
99. Ramsey, C.P. and B.I. Giasson, *The E163K DJ-1 mutant shows specific antioxidant deficiency.* Brain research, 2008. **1239**: p. 1-11.
100. Cookson, M.R., *Parkinsonism Due to Mutations in PINK1, Parkin, and DJ-1 and Oxidative Stress and Mitochondrial Pathways.* Cold Spring Harbor Perspectives in Medicine, 2012. **2**(9).
101. Kinumi, T., et al., *Cysteine-106 of DJ-1 is the most sensitive cysteine residue to hydrogen peroxide-mediated oxidation in vivo in human umbilical vein endothelial cells.* Biochemical and biophysical research communications, 2004. **317**(3): p. 722-728.
102. Miyama, A., et al., *Oxidation of DJ-1 induced by 6-hydroxydopamine decreasing intracellular glutathione.* PLoS One, 2011. **6**(11): p. e27883.
103. Girotto, S., et al., *Dopamine-derived Quinones Affect the Structure of the Redox Sensor DJ-1 through Modifications at Cys-106 and Cys-53.* Journal of Biological Chemistry, 2012. **287**(22): p. 18738-18749.
104. Blackinton, J., et al., *Formation of a stabilized cysteine sulfinic acid is critical for the mitochondrial function of the parkinsonism protein DJ-1.* Journal of Biological Chemistry, 2009. **284**(10): p. 6476-6485.
105. Hase, T. and Y. Niimura, *Protein-Protein Interaction Networks: Structures, Evolution, and Application to Drug Design.* Protein-Protein Interactions - Computational and Experimental Tools, 2012.
106. Jung, W., H.H. Jeong, and K.Y. Lee, *Protein Interactome and Its Application to Protein Function Prediction.* Protein-Protein Interactions - Computational and Experimental Tools, 2012.
107. Xiong, H., et al., *Parkin, PINK1, and DJ-1 form a ubiquitin E3 ligase complex promoting unfolded protein degradation.* The Journal of clinical investigation, 2009. **119**(3): p. 650.
108. Zucchelli, S., et al., *TRAF6 promotes atypical ubiquitination of mutant DJ-1 and alpha-synuclein and is localized to Lewy bodies in sporadic Parkinson's disease brains.* Human Molecular Genetics, 2010. **19**(19): p. 3759-3770.
109. Meulener, M.C., et al., *DJ-1 is present in a large molecular complex in human brain tissue and interacts with α-synuclein.* Journal of Neurochemistry, 2005. **93**(6): p. 1524-1532.

110. Shendelman, S., et al., *DJ-1 Is a Redox-Dependent Molecular Chaperone That Inhibits α-Synuclein Aggregate Formation.* PLoS Biology, 2004. **2**(11).
111. Takahashi, K., et al., *DJ-1 Positively Regulates the Androgen Receptor by Impairing the Binding of PIASxα to the Receptor.* Journal of Biological Chemistry, 2001. **276**(40): p. 37556-37563.
112. Junn, E., et al., *Interaction of DJ-1 with Daxx inhibits apoptosis signal-regulating kinase 1 activity and cell death.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(27): p. 9691-9696.
113. Wagner, S.A., et al., *A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles.* Molecular & Cellular Proteomics, 2011. **10**(10).
114. Kim, Y.C., et al., *Oxidation of DJ-1-dependent cell transformation through direct binding of DJ-1 to PTEN.* International journal of oncology, 2009. **35**(6): p. 1331-1341.
115. Opsahl, J.A., et al., *Increased interaction between DJ-1 and the Mi-2/nucleosome remodelling and deacetylase complex during cellular stress.* Proteomics, 2010. **10**(7): p. 1494-1504.
116. Lu, L., et al., *DJ-1 upregulates tyrosine hydroxylase gene expression by activating its transcriptional factor Nurr1 via the ERK1/2 pathway.* The International Journal of Biochemistry & Cell Biology, 2011.
117. Knobbe, C.B., et al., *Choice of biological source material supersedes oxidative stress in its influence on DJ-1 in vivo interactions with Hsp90.* Journal of proteome research, 2011. **10**(10): p. 4388-4404.
118. Deeg, S., et al., *BAG1 restores formation of functional DJ-1 L166P dimers and DJ-1 chaperone activity.* The Journal of cell biology, 2010. **188**(4): p. 505-513.
119. Jin, J., et al., *Identification of Novel Proteins Associated with Both α-Synuclein and DJ-1.* Molecular & Cellular Proteomics, 2007. **6**(5): p. 845-859.
120. McNally, R.S., et al., *DJ-1 Enhances Cell Survival through the Binding of Cezanne, a Negative Regulator of NF-κB.* Journal of Biological Chemistry, 2011. **286**(6): p. 4098-4106.
121. Ishikawa, S., et al., *Stimulation of vesicular monoamine transporter 2 activity by DJ-1 in SH-SY5Y cells.* Biochemical and Biophysical Research Communications, 2012.
122. Wang, Z., et al., *DJ-1 can inhibit microtubule associated protein 1 B formed aggregates.* Molecular Neurodegeneration, 2011. **6**(1): p. 38.
123. Yamashita, S., et al., *DJ-1 forms complexes with mutant SOD1 and ameliorates its toxicity.* Journal of neurochemistry, 2010. **113**(4): p. 860-870.
124. Mo, J.S., et al., *DJ-1 modulates the p38 mitogen-activated protein kinase pathway through physical interaction with apoptosis signal-regulating kinase 1.* Journal of cellular biochemistry, 2010. **110**(1): p. 229-237.
125. Baulac, S., et al., *Increased DJ-1 expression under oxidative stress and in Alzheimer's disease brains.* Molecular neurodegeneration, 2009. **4**(1): p. 12.
126. Cookson, M.R., *Pathways to parkinsonism.* Neuron, 2003. **37**(1): p. 7-10.
127. Niki, T., et al., *DJBP: A Novel DJ-1-Binding Protein, Negatively Regulates the Androgen Receptor by Recruiting Histone Deacetylase Complex, and DJ-1 Antagonizes This Inhibition by Abrogation of This Complex1 1 Ministry of Education, Science, Sport, Culture and Technology of Japan. 2 2 The nucleotide*

sequence reported in this paper has been submitted to the DDBJ/GenBank/EBI Data Bank with accession number AB073862. Molecular cancer research, 2003. **1**(4): p. 247-261.

128. Yasuda, T., et al., *DJ-1 cooperates with PYCR1 in cell protection against oxidative stress.* Biochemical and biophysical research communications, 2013.
129. Arndt, V. and I. Vorberg, *Defining the Cellular Interactome of Disease-Linked Proteins in Neurodegeneration*. 2012: Dr. Jianfeng Cai.
130. Vidal, M., M.E. Cusick, and A.L. Barabasi, *Interactome networks and human disease.* Cell, 2011. **144**(6): p. 986-998.
131. Abu-Farha, M., F. Elisma, and D. Figeys, *Identification of Protein–Protein Interactions by Mass Spectrometry Coupled Techniques.* Protein–Protein Interaction, 2008: p. 67-80.
132. Ivanov, A., V. Zgoda, and A. Archakov, *Technologies of protein interactomics: A review.* Russian Journal of Bioorganic Chemistry, 2011. **37**(1): p. 4-16.
133. Chautard, E., N. Thierry-Mieg, and S. Ricard-Blum, *Interaction networks: from protein functions to drug discovery. A review.* Pathologie Biologie, 2009. **57**(4): p. 324-333.
134. Cusick, M.E., et al., *Interactome: gateway into systems biology.* Human molecular genetics, 2005. **14**(suppl 2): p. R171-R181.
135. Paul, F.E., F. Hosp, and M. Selbach, *Analyzing protein–protein interactions by quantitative mass spectrometry.* Methods, 2011. **54**(4): p. 387-395.
136. Cho, S., et al., *Protein-protein interaction networks: from interactions to networks.* Journal of biochemistry and molecular biology, 2004. **37**(1): p. 45.
137. Suzuki, H., *Protein–protein interactions in the mammalian brain.* The Journal of physiology, 2006. **575**(2): p. 373-377.
138. Meixner, A., et al., *A QUICK screen for Lrrk2 interaction partners–leucine-rich repeat kinase 2 is involved in actin cytoskeleton dynamics.* Molecular & Cellular Proteomics, 2011. **10**(1).
139. Gingras, A.C., et al., *Analysis of protein complexes using mass spectrometry.* Nature reviews Molecular cell biology, 2007. **8**(8): p. 645-654.
140. Suter, B., S. Kittanakom, and I. Stagljar, *Two-hybrid technologies in proteomics research.* Current opinion in biotechnology, 2008. **19**(4): p. 316-323.
141. Lemmens, I., S. Lievens, and J. Tavernier, *Strategies towards high-quality binary protein interactome maps.* Journal of proteomics, 2010. **73**(8): p. 1415.
142. Pardo, M. and J.S. Choudhary, *Assignment of protein interactions from affinity purification/mass spectrometry data.* Journal of Proteome Research, 2012. **11**(3): p. 1462-1474.
143. Marcilla, M. and J.P. Albar, *Quantitative proteomics: A strategic ally to map protein interaction networks.* IUBMB life, 2013. **65**(1): p. 9-16.
144. Banks, C.A.S., S.E. Kong, and M.P. Washburn, *Affinity purification of protein complexes for analysis by multidimensional protein identification technology.* Protein Expression and Purification, 2012.
145. Dunham, W.H., M. Mullin, and A.C. Gingras, *Affinity-purification coupled to mass spectrometry: Basic principles and strategies.* Proteomics, 2012. **12**(10): p. 1576-1590.
146. Auerbach, D., M. Fetchko, and I. Stagljar, *Proteomic approaches for generating comprehensive protein interaction maps.* Targets, 2003. **2**(3): p. 85-92.

147. Sardiu, M.E. and M.P. Washburn, *Building protein-protein interaction networks with proteomics and informatics tools.* Journal of Biological Chemistry, 2011. **286**(27): p. 23645-23651.
148. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.
149. Lane, C., *Mass spectrometry-based proteomics in the life sciences.* Cellular and molecular life sciences, 2005. **62**(7): p. 848-869.
150. Domon, B. and R. Aebersold, *Mass spectrometry and protein analysis.* Science Signalling, 2006. **312**(5771): p. 212.
151. Nesvizhskii, A.I. and R. Aebersold, *Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS.* Drug Discovery Today, 2004. **9**(4): p. 173-181.
152. Sap, K.A. and J.A.A. Demmers, *Labeling Methods in Mass Spectrometry Based Quantitative Proteomics*. Integrative Proteomics. 2012: InTech.
153. Ramisetty, S.R. and M.P. Washburn, *Unraveling the dynamics of protein interactions with quantitative mass spectrometry.* Critical reviews in biochemistry and molecular biology, 2011. **46**(3): p. 216-228.
154. Wong, J. and G. Cagney, *An overview of label-free quantitation methods in proteomics by mass spectrometry.* Methods Mol. Biol, 2010. **604**: p. 273-283.
155. Ong, S.E. and M. Mann, *Mass spectrometry–based proteomics turns quantitative.* Nature chemical biology, 2005. **1**(5): p. 252-262.
156. Candiano, G., et al., *Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis.* Electrophoresis, 2004. **25**(9): p. 1327-1333.
157. O'Connell, K.L. and J.T. Stults, *Identification of mouse liver proteins on two-dimensional electrophoresis gels by matrix-assisted laser desorption/ionization mass spectrometry of in situ enzymatic digests.* Electrophoresis, 1997. **18**(3-4): p. 349-359.
158. Gillet, L.C., et al., *Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis.* Molecular & Cellular Proteomics, 2012. **11**(6).
159. Lambert, J.-P., et al., *Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition.* Nature methods, 2013.
160. Collins, B.C., et al., *Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-3 system.* Nature methods, 2013.
161. Eden, E., et al., *GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists.* BMC bioinformatics, 2009. **10**(1): p. 48.
162. Sherman, F., J.W. Stewart, and S. Tsunasawa, *Methionine or not methionine at the beginning of a protein.* BioEssays, 1985. **3**(1): p. 27-31.
163. Tao, X. and L. Tong, *Crystal structure of human DJ-1, a protein associated with early onset Parkinson's disease.* Journal of Biological Chemistry, 2003. **278**(33): p. 31372-31379.
164. Yanagisawa, D., et al., *DJ-1 protects against neurodegeneration caused by focal cerebral ischemia and reperfusion in rats.* Journal of Cerebral Blood Flow & Metabolism, 2007. **28**(3): p. 563-578.

165. Yanagida, T., et al., *Oxidative stress induction of DJ-1 protein in reactive astrocytes scavenges free radicals and reduces cell injury.* Oxidative medicine and cellular longevity, 2009. **2**(1): p. 36-42.
166. Schwenk, J., et al., *Functional proteomics identify cornichon proteins as auxiliary subunits of AMPA receptors.* Science, 2009. **323**(5919): p. 1313-1319.
167. Zhou, W., et al., *The oxidation state of DJ-1 regulates its chaperone activity toward α-synuclein.* Journal of molecular biology, 2006. **356**(4): p. 1036-1048.
168. Zhang, R. and F. Monsma, *Fluorescence-based thermal shift assays.* Current opinion in drug discovery & development, 2010. **13**(4): p. 389.
169. Phillips, K. and A.H. de la Peña, *The combined use of the Thermofluor assay and ThermoQ analytical software for the determination of protein stability and buffer optimization as an aid in protein crystallization.* Current Protocols in Molecular Biology, 2011: p. 10.28. 1-10.28. 15.
170. Limoli, C.L., et al., *Attenuation of radiation-induced genomic instability by free radical scavengers and cellular proliferation.* Free Radical Biology and Medicine, 2001. **31**(1): p. 10-19.
171. Rosenblum, W.I. and F. El-Sabban, *Dimethyl sulfoxide (DMSO) and glycerol, hydroxyl radical scavengers, impair platelet aggregation within and eliminate the accompanying vasodilation of, injured mouse pial arterioles.* Stroke, 1982. **13**(1): p. 35-39.
172. Xu, J., et al., *The Parkinson's disease-associated DJ-1 protein is a transcriptional co-activator that protects against neuronal apoptosis.* Human molecular genetics, 2005. **14**(9): p. 1231-1241.
173. Mirzaei, H. and F. Regnier, *Protein-RNA cross-linking in the ribosomes of yeast under oxidative stress.* Journal of proteome research, 2006. **5**(12): p. 3249-3259.
174. Liu, S., et al., *The network of protein–protein interactions within the human U4/U6. U5 tri-snRNP.* Rna, 2006. **12**(7): p. 1418-1430.
175. Nguyen, T.H.D., et al., *Structural Basis of Brr2-Prp8 Interactions and Implications for U5 snRNP Biogenesis and the Spliceosome Active Site.* Structure, 2013.
176. Raghunathan, P.L. and C. Guthrie, *RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2.* Current biology, 1998. **8**(15): p. 847-855.
177. Jankowsky, E. and H. Bowers, *Remodeling of ribonucleoprotein complexes with DExH/D RNA helicases.* Nucleic acids research, 2006. **34**(15): p. 4181-4188.
178. Schwer, B., *A new twist on RNA helicases: DExH/D box proteins as RNPases.* Nature Structural & Molecular Biology, 2001. **8**(2): p. 113-116.
179. Chang, P., et al., *DDX3, a DEAD box RNA helicase, is deregulated in hepatitis virus-associated hepatocellular carcinoma and is involved in cell growth control.* Oncogene, 2005. **25**(14): p. 1991-2003.
180. Sun, M., et al., *Identification of an antiapoptotic protein complex at death receptors.* Cell Death & Differentiation, 2008. **15**(12): p. 1887-1900.

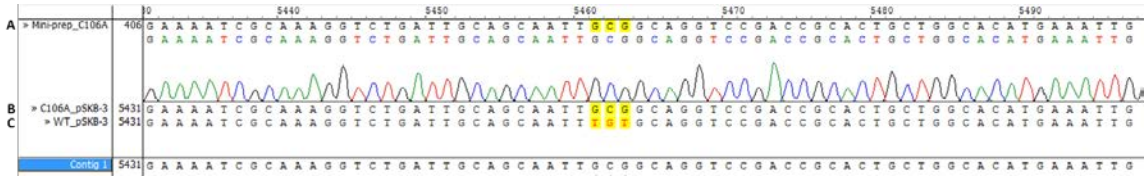
Notes:
Some figures were adapted from their original publications for simplicity.

7. SUPPLEMENTARY DATA

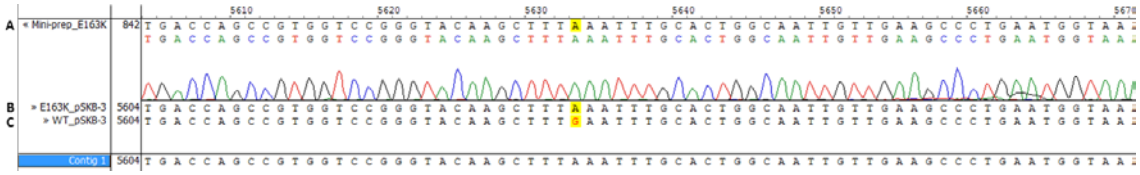
7.1. SEQUENCING

As shown for the results obtained for the C106DD mutant, the sequencing confirmed the acquisition of plasmid DNA with the DJ-1 C106A and E163K mutated genes to use for further production and purification of the respective mutated proteins. This was confirmed by comparison of the different plasmids constructs of these mutants (Supplementary Figure 7.1) where the codon substitution was confirmed (highlighted nucleotides in the figures). Codon substitutions for all the mutants are described in Experimental Procedures - 2.1.

I



II

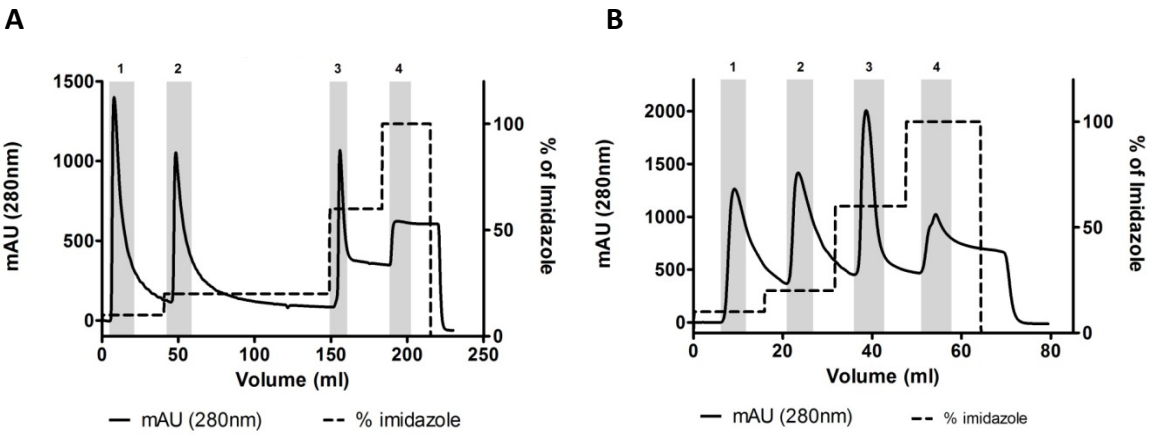


Supplementary Figure 7.1| DNA SEQUENCING RESULTS OF DJ-1 MUTANTS C106A (I) AND E163K (II) CODIFIYING SEQUENCES. Nucleotide sequence alignment of small fragments of DJ-1 mutant C106A (or E163K) gene sequenced (A), its corresponding sequence in the C106A (or E163K) designed plasmid (B) and in the DJ-1 WT designed plasmid (C). The codon containing the mutation is highlighted. The codon substitution was TGT to GCG for the C106A mutation and GAA to AAA in the E163K mutation.

7.2. DJ-1 WT AND MUTANTS PRODUCTION AND PURIFICATION

7.2.1. PRODUCTION AND PURIFICATION – AFFINITY AND SIZE EXCLUSION CHROMATOGRAPHIES

As mentioned above, after sequencing the mutated genes, the proteins were produced and purified. In the purification step two columns were used, the HisTrap and the Superdex 200, and the chromatograms obtained for the mutants C106A and E163K are shown below (Supplementary Figure 7.2).

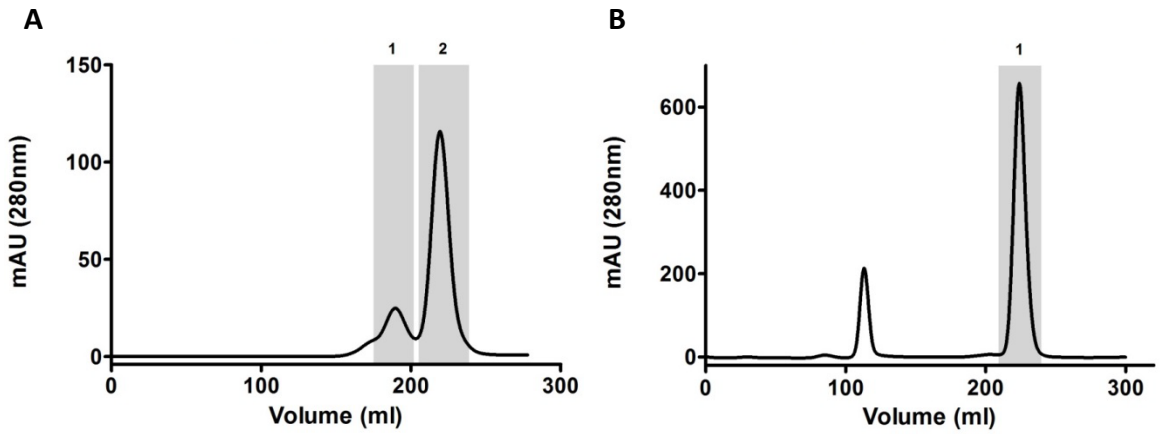


Supplementary Figure 7.2 | AFFINITY CHROMATOGRAPHY PURIFICATION OF DJ-1 C106A (A) AND E163K (B). The affinity chromatography was performed using a HisTrap column. Imidazole step gradient (50mM, 100mM, 300mM and 500mM) is shown in dashed line. The grey bars correspond to the fractions that were collected from each elution step - fraction 1, 2, 3 and 4 eluted with 50, 100, 300 and 500mM of imidazole, respectively.

From each affinity purification four fractions were recovered and in both cases, the fraction 3 eluted with 300mM of imidazole was collected and loaded into the Superdex 200 column by the same reasons mentioned in section 3.2. Fraction 4 from the C106A chromatogram (Supplementary Figure 7.2 – A) did not contain protein but the same fraction from E163K chromatogram contains a low amount of protein as it can be observed by the small peak obtained.

From the purification of mutants C106A and E163K two and one fraction were recovered, respectively (Supplementary Figure 7.3). In the purification of E163K a problem occurred and the first peak could not be recovered. However, based in the low values of volume in which this peak appeared the protein present in this fraction would not be in the homodimeric form, but instead in oligomeric forms. Therefore, this would not be the fraction of interest to use. The fractions of interest were the ones recovered around 210mL, which were fraction 2 from C106A chromatogram and fraction 1 from E163K chromatogram.

Furthermore, analysis of fraction 1 from C106A mutant purification by HPLC showed the presence of protein in the form of oligomers (data not shown). Given the above mentioned, the fractions stored for further analysis were fraction 2 from C106A purification and fraction 1 from E163K size exclusion chromatographies.

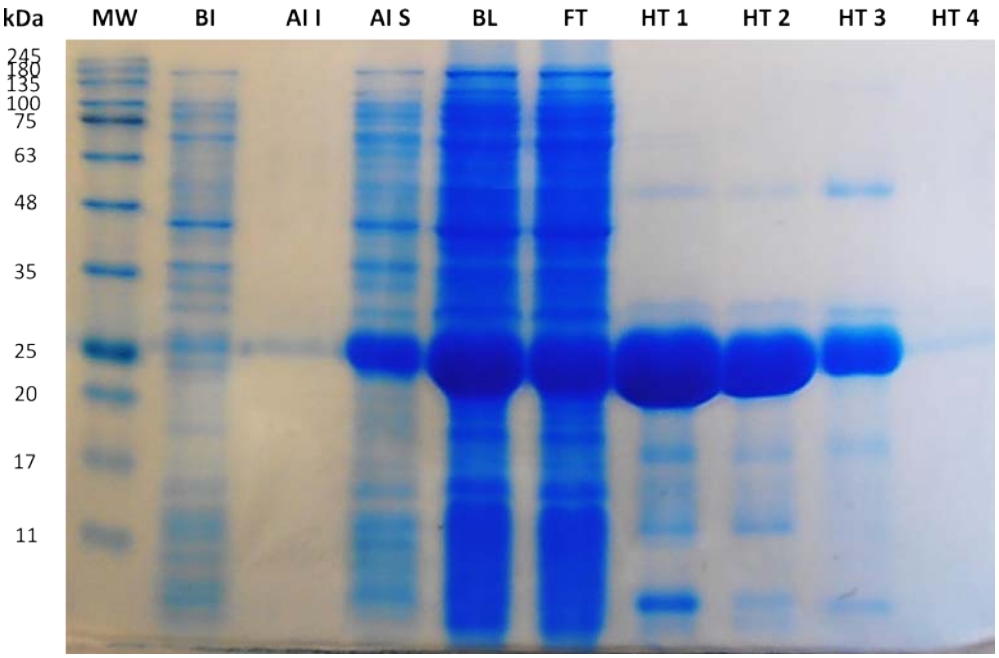


Supplementary Figure 7.3 | SIZE EXCLUSION CHROMATOGRAPHY PURIFICATION OF DJ-1 C106A (A) AND E163K (B). Fraction 3 recovered from the affinity chromatography of C106A and E163K (Supplementary Figure 7.2 – A and B, respectively) was applied in an HiLoad 26-60 Superdex 200 column. The mobile phase used was PBS. The grey bars represent the fractions that were collected (fraction 1 and 2 (A) and fraction 1 (B)).

7.2.2. PROTEIN PURIFICATION ASSESSMENT

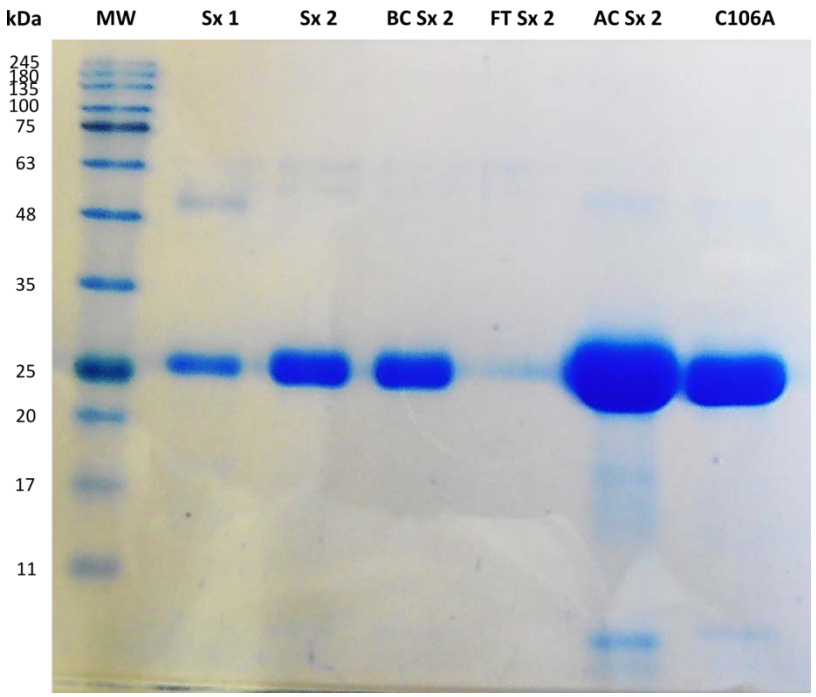
The following figures show the protein profile obtained for samples recovered from protein expression (BI, AI I and AI S) and protein purification using affinity chromatography (BL, FT, HT 1, HT 2, HT 3, HT 4) – Supplementary Figure 7.4 – and samples recovered from size exclusion chromatography (Sx1 and Sx2), from protein concentration (BC Sx2, FT Sx2 and AC Sx2) and the stored protein sample (C106A) – Supplementary Figure 7.5.

Analyzing the results it can be seen that the samples obtained before induction (BI) and insoluble protein fraction obtained after induction of protein expression (AI I) have a band around the molecular weight of DJ-1 with very low intensity (Supplementary Figure 7.4). This suggests that DJ-1 was not present in the medium before protein expression induced with IPTG and neither in the insoluble protein fraction recovered after induction of protein expression. On the contrary, an intense band around 24kDa was observed in the soluble protein fraction after induction of protein expression (AI S) which indicates that DJ-1 was produced in a soluble form and present in the soluble proteins fraction.



Supplementary Figure 7.4 | ANALYSIS OF DJ-1 C106A EXPRESSION AND PURIFICATION BY SDS-PAGE. SDS-PAGE followed by Coomassie staining. MW – molecular weight marker: NZY Colour Protein Marker II (NZYTech); BI – cellular extract obtained before induction of protein expression; AI I – insoluble protein fraction obtained from centrifugation after induction of protein expression; AI S – soluble protein fraction obtained after induction of protein expression; BL – cellular extract loaded on the HisTrap column; FT – flow-through obtained from the HisTrap column; HT 1 –fraction 1 eluted from the HisTrap column; HT 2 - fraction 2 eluted from the HisTrap column; HT 3 - fraction 3 eluted from the HisTrap column; HT 4 - fraction 4 eluted from the HisTrap column (Supplementary Figure 7.2 - A). The sample volume added to each well was 10µL.

Regarding protein purification using an HisTrap column it could be observed the presence of a DJ-1 band in all fractions except fraction 4. This means, in the case of fraction recovered before column loading (BI) and fraction recovered from the flow through after column loading (FT), that DJ-1 was effectively loaded into the column but also that a large quantity of DJ-1 did not bind to the resin and therefore was lost. The presence of DJ-1 in fractions 1 to 3 (HT 1 to 3) and absence from fraction 4 (HT 4) was expected through the analysis of the affinity purification chromatogram. Noteworthy that in fractions 1 to 3 more bands are present besides DJ-1´s band and that the intensity of these bands is lower in fraction 3, which suggests that this fraction has a lower amount of contaminants. Therefore, this fraction was chosen to be purified by size exclusion chromatography.

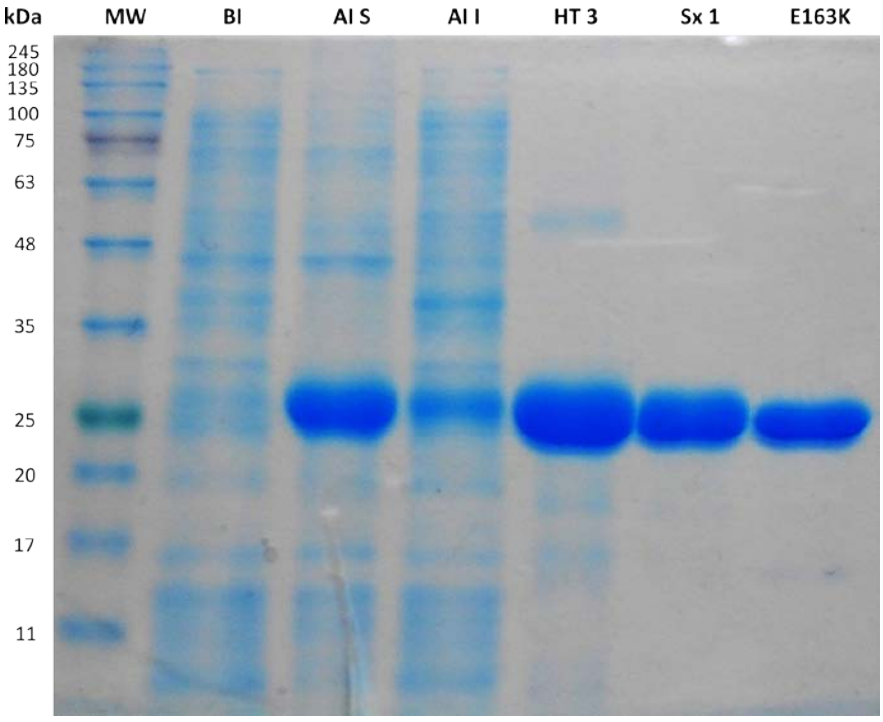


Supplementary Figure 7.5 | ANALYSIS OF DJ-1 C106A SIZE EXCLUSION CHROMATOGRAPHY BY SDS-PAGE. SDS-PAGE followed by Coomassie staining. MW – molecular weight marker: NZY Colour Protein Marker II (NZYTech); After loading the fraction HT 3 from the HisTrap elution on a HiLoad 26/60 Superdex 200 prep grade fractions 1 (Sx 1) and 2 (Sx 2) were recovered; BC Sx 2 – sample containing fraction 2 obtained from Superdex 200 before concentration; FT Sx 2 – flow-through obtained from concentration of fraction 2; AC Sx 2 - purified DJ-1 C106A obtained after concentration of fraction 2; C106A – sample AC Sx 2 stored in PBS with 10% glycerol at -80°C. The sample volume added to each well was 10µL.

The two fractions recovered from size exclusion chromatography (Sx 1 and Sx 2), after loading fraction 3 (HT 3), both contained a band corresponding to DJ-1 (≈22kDa) being the Sx 2 band more intense (Supplementary Figure 7.5). Additionally, the sample containing fraction 1 also presented another band around 48 kDa that could correspond to oligomers, also supported by the size exclusion chromatogram analysis. Thereby, fraction 2 (Sx 2) was selected to storage and use in further assays. Before storage, this fraction was concentrated and in the sample retrieved before concentration (BC Sx 2) the presence of DJ-1 in the sample was confirmed by the band shown at approximately 22 kDa. Due to the very low intensity of DJ-1 band in the sample obtained from the flow through of the concentration (FT Sx 2) it can be concluded that almost no protein was lost during this process. This is confirmed by the very intense DJ-1 band in the sample recovered after protein concentration (AC Sx 2).

In this sample, there are other bands present, which suggested that these contaminants were also concentrated. The concentrated sample was stored in PBS with 10% glycerol (C106A) and the SDS-PAGE showed only an intense band at DJ-1 molecular weight compared to the sample obtained after concentration. This fact could be explained by the dilution performed in PBS with 10% glycerol to store the protein.

Samples recovered during protein expression (BI, AI I, AI S), protein purification using a HisTrap column (HT 3) followed by Superdex 200 column (Sx 1) from E163K were resolved by SDS-PAGE (Supplementary Figure 7.6). To evaluate the protein profile of this protein only one SDS-PAGE gel was performed and some fractions were not recovered.



Supplementary Figure 7.6 | ANALYSIS OF DJ-1 E163K EXPRESSION AND PURIFICATION BY SDS-PAGE. SDS-PAGE followed by Coomassie staining. MW – molecular weight marker: NZY Colour Protein Marker II (NZYTech); BI – cellular extract obtained before induction of protein expression; AI I – insoluble protein fraction obtained from centrifugation after induction of protein expression; AI S – soluble protein fraction obtained after induction of protein expression; HT 3 – fraction 3 eluted from the HisTrap column; Sx 1 – fraction 1 eluted after loading sample HT 3 on a HiLoad 26/60 Superdex 200 prep grade; E163K – sample Sx F1 stored in PBS with 10% glycerol at -80°C. The sample volume added to each well was 10µL.

The results show that there was no band at DJ-1 molecular weight before protein expression induction (BI), which means that DJ-1 was not present in the medium before protein expression induced with IPTG. After protein expression induction DJ-1 was present in large amount in the soluble fraction (AI S) and in a smaller amount in the insoluble fraction (AI I). This means that DJ-1 was mostly produced in a soluble form and present in the soluble protein fraction. Fraction 3 (HT 3) recovered from the HisTrap column, corresponds to protein elution at 300mM of imidazole with an intense DJ-1 band along with other bands with very low intensity. After protein elution from the HisTrap, fraction 3 was loaded into the Superdex 200 column and one fraction was recovered from this elution (Sx 1). In this fraction an intense band for DJ-1 was present again which suggests the presence of DJ-1 in large amounts in this sample. Therefore, the sample Sx 1 was diluted in PBS with 10% glycerol and stored at -80^{o}C (E163K).

7.2.3. PROTEIN CONCENTRATION

Protein samples obtained for each mutant were analyzed by absorbance at 280nm (NanoDrop) to determine the protein concentration before storage. As mentioned above, the blank used in this procedure was the same buffer in which the protein was stored (PBS with 10% glycerol) and the results are shown in Supplementary Table 7.1.

Supplementary Table 7.1 | PROTEIN CONCENTRATION OF FINAL SAMPLES CONTAINING THE MUTANTS IN STUDY.

Protein	Concentration (µg/µL)
Wild-type	1.97
C106A	2.46
C106DD	2.19
E163K	2.05

The desired concentration in the sample to store was around 2.0mg/mL. As observed in Supplementary Table 7.1 all the concentrations are around this value with the exception of mutant C106A with a higher concentration.

7.3. SEQUENCE ANALYSIS BY LC-MS/MS

The C106A mutant was identified with an unused protein score of 134.34, a sequence coverage of 99.5% and 107 peptides were identified with 95% or more confidence. The results indicated that the initial methionine was not identified which is represented by a sequence coverage of 99.5% (Supplementary Figure 7.7).

MGSSHHHHHHDYDIPTTENLYFQGHMASKRALVILAKGAEEMETVIPVDVMRRAGIKVTVAGLAGKDPVQC
SRDVVICPDASLEDAKKEGPYDVVVLPGGNLGAQNLSESAAVKEILKEQENRKGLIAAIAAGPTALLAHEIGFGS
KVTTHPLAKDKMMNGGHYTYSENRVEKDGLILTSRGPGTSFEFALAIVEALNGKEVAAQVKAPLVLKD

Supplementary Figure 7.7 | SEQUENCE COVERAGE OF DJ-1 MUTANT C106A. Green letters represent the residues identified in peptides with at least 95% confidence; grey letters represent the unidentified residues. Mutated residue is highlight in yellow.

The E163K mutant was identified with an unused protein score of 127.55, a sequence coverage of 100% and 96 peptides were identified with 95% or more confidence. Unlike the results obtained for the other proteins, for this mutant the initial methionine was identified but with low confidence (Supplementary Figure 7.8).

MGSSHHHHHHDYDIPTTENLYFQGHMASKRALVILAKGAEEMETVIPVDVMRRAGIKVTVAGLAGKDPVQC
SRDVVICPDASLEDAKKEGPYDVVVLPGGNLGAQNLSESAAVKEILKEQENRKGLIAAICAGPTALLAHEIGFGS
KVTTHPLAKDKMMNGGHYTYSENRVEKDGLILTSRGPGTSFKFALAIVEALNGKEVAAQVKAPLVLKD

Supplementary Figure 7.8 | SEQUENCE COVERAGE OF DJ-1 MUTANT E163K. Green letters represent the residues identified in peptides with at least 95% confidence; grey letters represent the unidentified residues; red letters represent the residues identified in peptides with low confidence. Mutated residue is highlight in yellow.

This fact justifies the sequence coverage value of 100%. Probably for this protein the cleavage of the initial methionine was inefficient and that allowed the identification of this residue even though with low confidence**.**

DJ-1 WT was identified with an unused protein score of 152.33, a sequence coverage of 99.5% and 116 peptides were identified with 95% or more confidence. The results indicated that the initial methionine was not identified contributing to a value of 99.5% for sequence coverage (Supplementary Figure 7.9).

MGSSHHHHHHDYDIPTTENLYFQGHMASKRALVILAKGAEEMETVIPVDVMRRAGIKVTVAGLAGKDPVQC
SRDVVICPDASLEDAKKEGPYDVVVLPGGNLGAQNLSESAAVKEILKEQENRKGLIAAICAGPTALLAHEIGFGS
KVTTHPLAKDKMMNGGHYTYSENRVEKDGLILTSRGPGTSFEFALAIVEALNGKEVAAQVKAPLVLKD

Supplementary Figure 7.9 | SEQUENCE COVERAGE OF DJ-1 WT. Green letters represent the residues identified in peptides with at least 95% confidence; grey letters represent the unidentified residues.

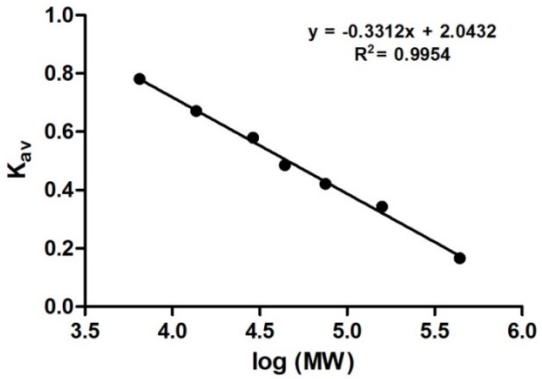
7.4. SIZE EXCLUSION– HPLC

To evaluate if the proteins were in a homodimeric form, a size exclusion – HPLC was performed. In this technique, the molecular weight of each protein was determined using a calibration curve obtained by the retention time of certain standards – Supplementary Figure 7.10.

The molecular weight of DJ-1 and its mutants was calculated using the retention time of each sample. Firstly, the retention time was converted to elution volume (V_e) using the column flow rate value. With elution volume and the values of column void volume (V_0) – blue dextran elution volume - and geometric column volume (V_c), the partition coefficient (K_{av}) was calculated using the following formula:

$$\boldsymbol{K_{av}} = \frac{\boldsymbol{V_e - V_0}}{\boldsymbol{V_c - V_0}} \,.$$

The calibration curve was represented by the partition coefficient (K_{av}) versus the logarithm of molecular weight (log (MW)) as shown in Supplementary Figure 7.10.

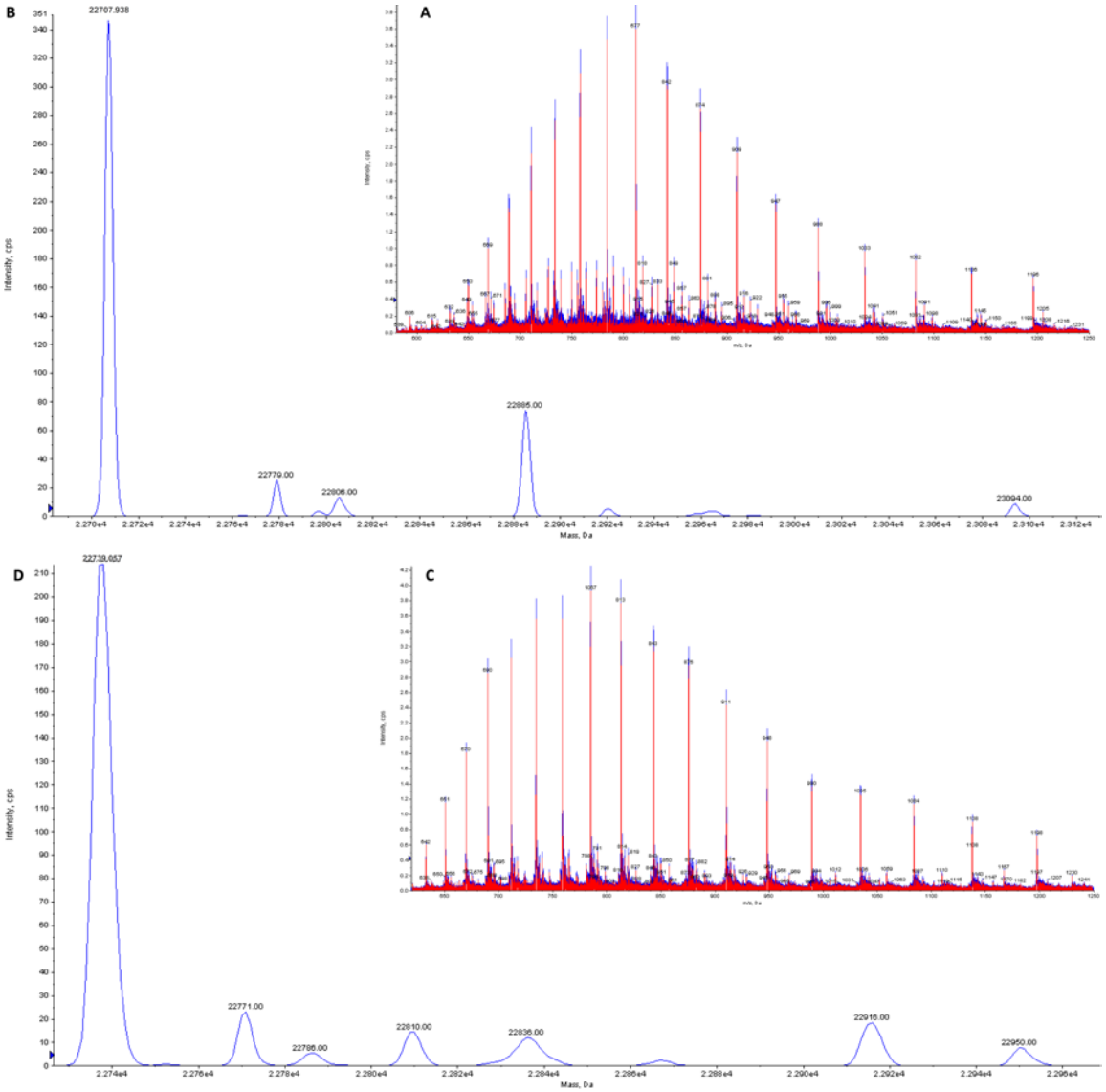


Supplementary Figure 7.10 | HPLC-SIZE EXCLUSION CHROMATOGRAPHY CALIBRATION CURVE. Molecular weight of standards (dots) from right to left: Ferritin (440 kDa); Aldolase (158 kDa); Conalbumin (75 kDa); Ovalbumin (44 kDa); Carbonic Anhydrase (29 kDa); Ribonuclease A (13.7 kDa); Aprotinin (6.5 kDa). The mobile phase was PBS with 10% glycerol (buffer used to store the proteins). K_{av} – partition coefficient, log (MW) - logarithm of molecular weight.

Therefore, using the calculated value of K_{av} of the protein, the log (MW) can be retrieved and finally the molecular weight can be inferred.

7.5. LC-MS OF INTACT PROTEINS

The following results represent the deconvoluted mass spectrum and graphic of deconvoluted masses for the mutants C106A and E163K (Supplementary Figure 7.11).



Supplementary Figure 7.11 | LC-MS ANALYSIS OF INTACT DJ-1 MUTANTS C106A AND E163K. A and C - Average spectra of protein charge envelop of intact C106A and E163K, respectively (in blue). Deconvoluted mass spectrum of protein charge envelop of intact C106A and E163K, respectively (in red). B and D - Graphic of deconvoluted masses of C106A and E163K, respectively, using BioAnalyst™.

In the same way as explained in the section 3.4.2 from results, the values from the average and predicted masses were obtained for DJ-1 WT and mutants C106A and E163K – Supplementary Table 7.2.

Supplementary Table 7.2 | PROTEINS´ MASSES OBTAINED WITH BIOANALYST AND PREDICTED FROM THE SEQUENCE. Average and predicted masses were obtained using BioAnalyst™ and Mass Calculator tool, respectively. The predicted masses were calculated without the initial methionine. The shift between the two mass values is also shown.

Protein	Average Mass (Da)	Predicted Mass (Da)	Shift (Da)
WT	22739.998	22738.899	1.099
E163K	22739.057	22737.908	1.149
C106A	22707.938	22707.407	0.531
C106DD	22867.037	22865.953	1.085

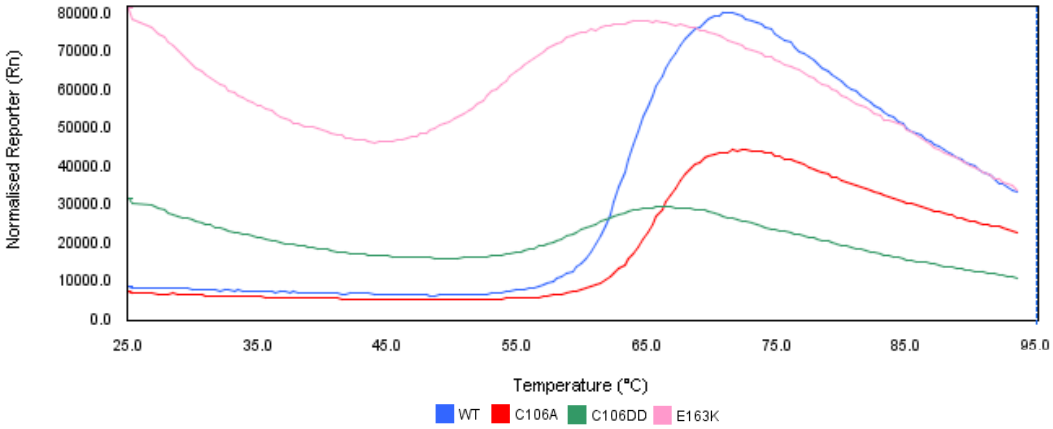
In a general way, similar results for predicted and average mass were obtained giving low shift values. All the values were very close to 22.7 kDa as expected since this is the approximate molecular weight of DJ-1 WT monomer.

7.6. THERMAL SHIFT – NORMALIZED CURVES

The results obtained from the thermal shift assay were represented by normalized curves (normalized fluorescence intensity versus temperature) (Supplementary Figure 7.12). To simplify the acquisition of the melting temperature of each protein – temperature at the inflection point of the normalized curves slope – derivative curves were obtained from the normalized curves, where the T_m is the temperature value of the peak obtained in these curves (Figure 3.9).

The normalized curves indicated that mutants E163K and C106DD presented higher fluorescence for initial values of temperature. Based on the concept of thermal shift this means that the structure of these proteins have more hydrophobic regions initially exposed allowing the dye to bind to more protein regions and promoting the increase in fluorescence signal. Due to this fact inferior T_m values were obtained for these proteins compared to the DJ-1 WT and mutant C106A. Since these proteins had

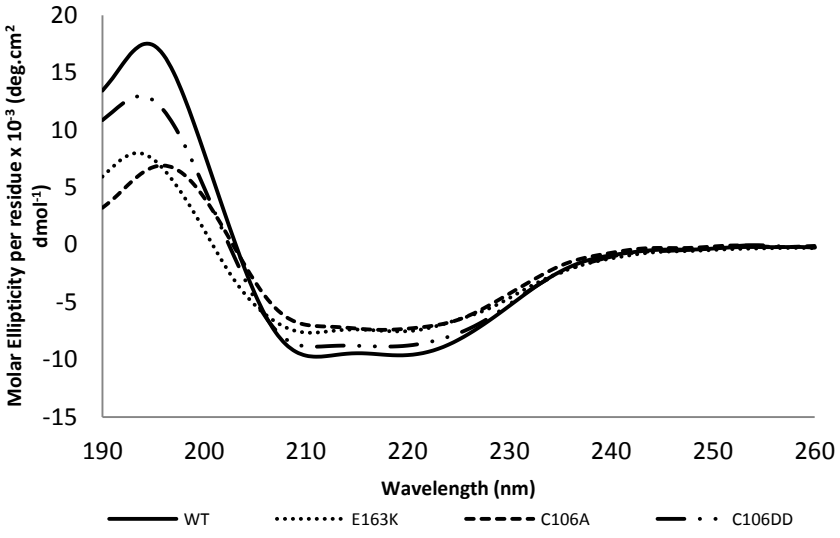
already more hydrophobic regions exposed, the temperature necessary to denature 50% of the protein structure was lower.



Supplementary Figure 7.12 | NORMALISED CURVES OF DJ-1 WT AND MUTANTS C106A, C106DD AND E163K OBTAINED FROM THERMAL SHIFT ASSAY. The concentration of the recombinant DJ-1 WT and its mutants was 0.5mg/mL. The melting temperature (T_m) was determined using the first derivative of the normalized curve for each protein (Figure 3.9).

7.7. CIRCULAR DICHROISM

Initially, the analysis of CD was performed at 25°C (Supplementary Figure 7.13), a commonly used temperature to determine the secondary structure content of proteins. However, the assays in which the proteins would be used were performed at



Supplementary Figure 7.13 | FAR-ULTRAVIOLET CIRCULAR DICHROISM SPECTRA OF DJ-1 WT AND MUTANTS C106A, C106DD AND E163K. CD spectra were obtained under normal conditions at 25°C using a 0.05mm cell. The protein concentrations used were approximately 2.0mg/mL.

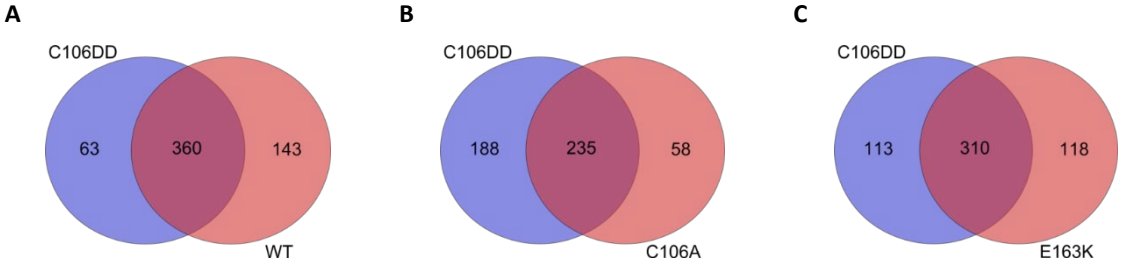
37°C and thereby the CD analysis was repeated and performed at 37°C (Figure 3.10).

At physiologic pH and 25°C the circular dichroism (CD) spectral analysis showed that DJ-1 mutants C106A, C106DD and E163K have essentially the same structure as DJ-1 WT with a decrease in signal amplitude for mutants spectra. However, all CD spectra displayed a broad negative ellipticity from 208 to 222 nm, consistent with a mixed α/β structure. Despite these differences, the secondary structure content was similar for all proteins containing 20 to 30% of structure in α-helix, 20 to 30% of structure in β-sheet and approximately 50% without regular secondary structure, as calculated by the GlobalWorks (described in Experimental Procedures – 2.7).

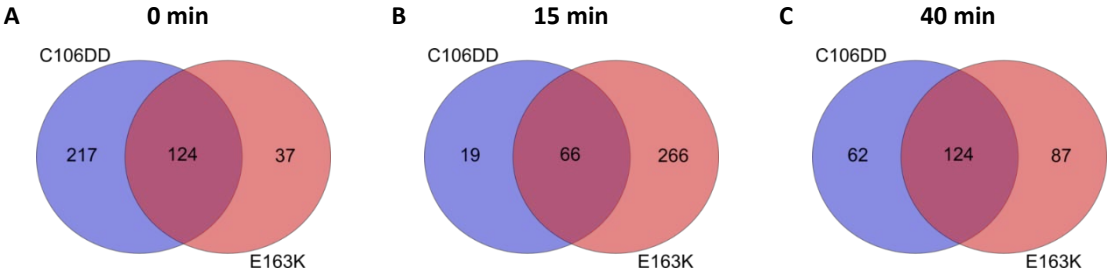
Overall, these results are similar to the results obtained for 37°C, with the exception of the DJ-1 WT that showed a higher content of structure in α-helix, approximately 40%, compared to the 20 to 30% obtained at 25°C.

7.8. BINDING PARTNERS QUALITATIVE ASSESSMENT

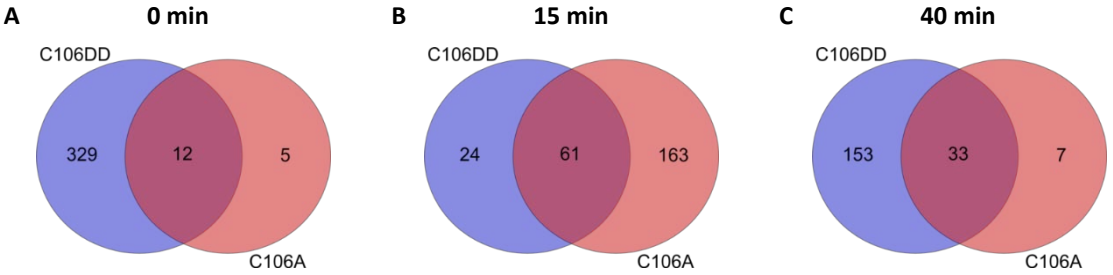
A general Venn diagram previously shown illustrates the shared and unique proteins among DJ-1 WT and mutants (Figure 3.12). However, due to the elevated number of conditions to be compared, the comparison between only two conditions can be time consuming and visually confusing. In this way, the following figures (Supplementary Figure 7.14 to 7.18) provide a more targeted comparison between each recombinant protein and C106DD mutant, as an example. More and similar combinations were performed for each recombinant protein (data not shown).



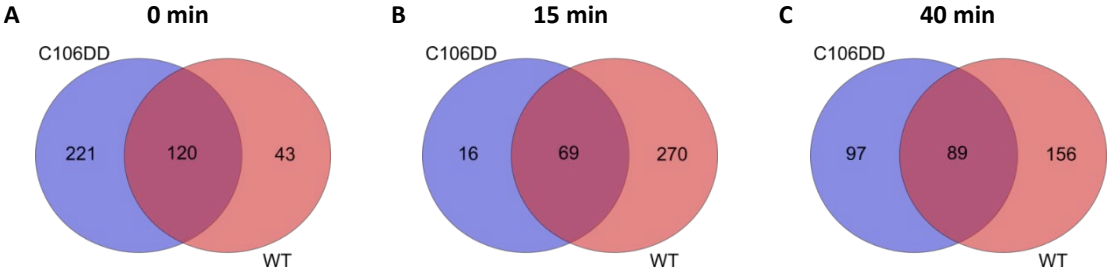
Supplementary Figure 7.14| VENN DIAGRAMS OF THE IDENTIFIED BINDING PARTNERS IN DJ-1 WT AND MUTANTS C106A AND E163K. A- Venn diagram showing the distribution of shared binding partners for all time points among DJ-1 C106DD and DJ-1 WT. The combined protein number is 566. B - Venn diagram showing the distribution of shared binding partners for all time points among DJ-1 C106A and C106DD. The combined protein number is 481. C - Venn diagram showing the distribution of shared binding partners for all time points among DJ-1 C106DD and E163K. The combined protein number is 541.



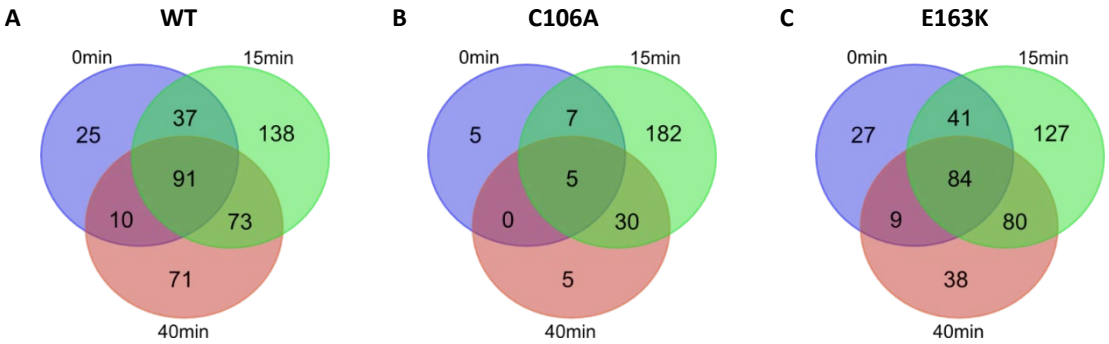
Supplementary Figure 7.15| VENN DIAGRAMS OF THE IDENTIFIED BINDING PARTNERS IN DJ-1 MUTANTS C106DD AND E163K FOR EACH TIME POINT. A- Venn diagram showing the distribution of shared and unique binding partners in the 0min condition. The combined protein number is 378. B - Venn diagram showing the distribution of shared and unique binding partners in the 15min condition. The combined protein number is 351. C - Venn diagram showing the distribution of shared and unique binding partners in the 40min condition. The combined protein number is 273.



Supplementary Figure 7.16| VENN DIAGRAMS OF THE IDENTIFIED BINDING PARTNERS IN DJ-1 MUTANTS C106DD AND C106A FOR EACH TIME POINT. A- Venn diagram showing the distribution of shared and unique binding partners in the 0min condition. The combined protein number is 346. B - Venn diagram showing the distribution of shared and unique binding partners in the 15min condition. The combined protein number is 248. C - Venn diagram showing the distribution of shared and unique binding partners in the 40min condition. The combined protein number is 193.

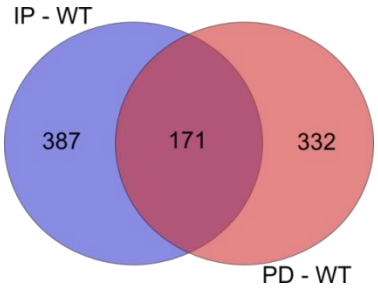


Supplementary Figure 7.17| VENN DIAGRAMS OF THE IDENTIFIED BINDING PARTNERS IN DJ-1 WT AND C106DD MUTANT FOR EACH TIME POINT. A- Venn diagram showing the distribution of shared and unique binding partners in the 0min condition. The combined protein number is 384. B - Venn diagram showing the distribution of shared and unique binding partners in the 15min condition. The combined protein number is 355. C - Venn diagram showing the distribution of shared and unique binding partners in the 40min condition. The combined protein number is 442.



Supplementary Figure 7.18| VENN DIAGRAMS ILLUSTRATING THE NUMBER OF SHARED OR UNIQUE BINDING PARTNERS OF DJ-1 WT AND MUTANTS C106A AND E163K AMONG THE DIFFERENT TIME POINTS. A- There are 91 common binding partners in all time points for DJ-1 WT and the combined protein number is 445. B- There are 5 common binding partners in all time points for DJ-1 C106A and the combined protein number is 234. C- There are 84 common binding partners in all time points for DJ-1 E163K and the combined protein number is 406.

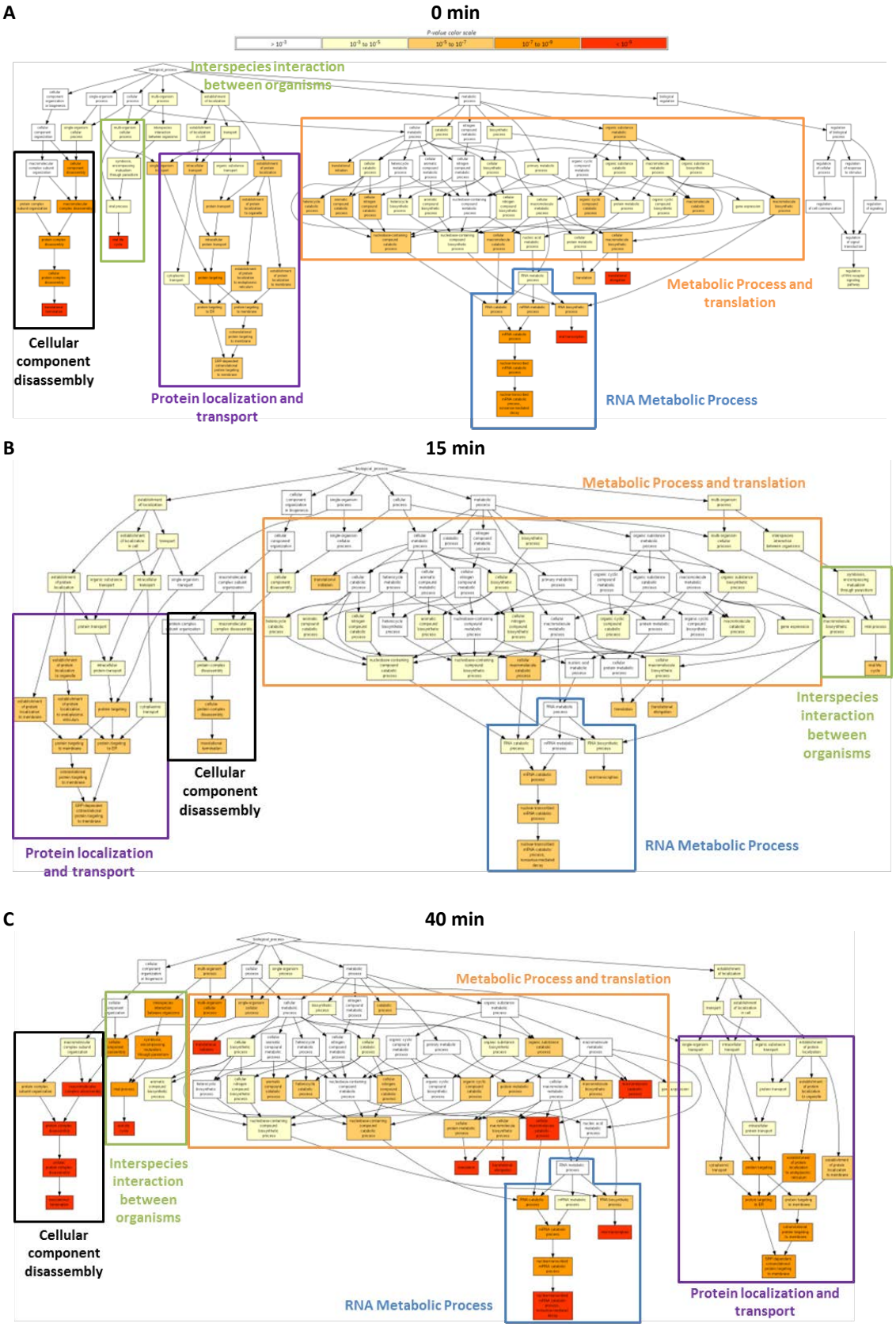
As mentioned above, a comparison was made between the binding partners identified in the pull-down assay (PD) performed in this project and in the immunoprecipitation (IP) performed in previous work (Supplementary Figure 7.19). The results show that 171 binding partners were identified in both assays, and 332 and 387 are unique binding partners for the PD and IP assay, respectively.



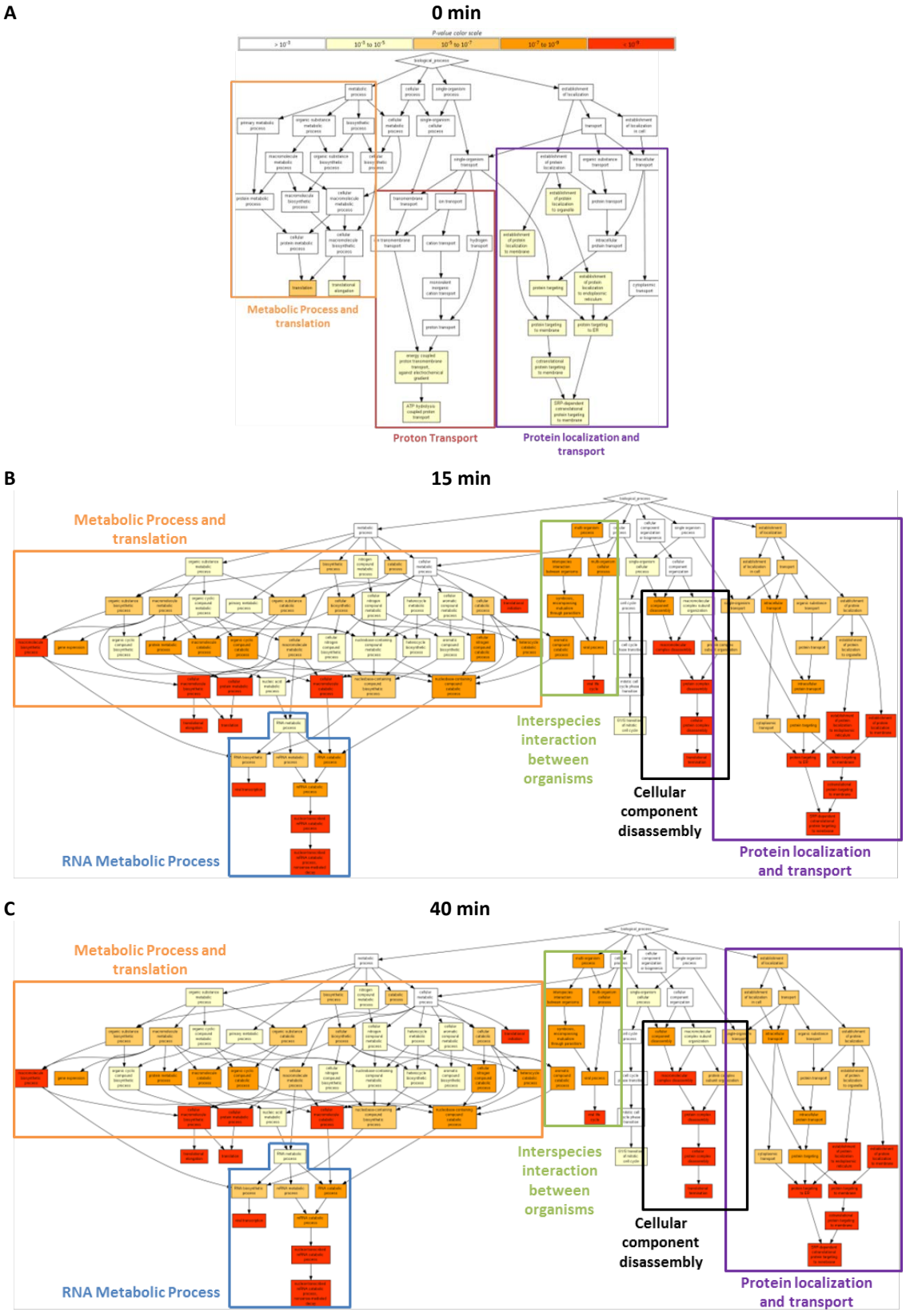
Supplementary Figure 7.19| VENN DIAGRAM ILLUSTRATING THE NUMBER OF SHARED OR UNIQUE BINDING PARTNERS AMONG THE INTERACTOMIC ASSAYS FOR DJ-1 WT PERFORMED BY IMUNNOPRECIPITATION AND PULL-DOWN. There are 171 common binding partners between the IP assay previously performed (data not shown) and the PD assay performed in this project. The combined protein number is 890. IP – Immunoprecipitation; PD – Pull-down.

7.9. GO ENRICHMENT ANALYSIS

Similarly to the analysis presented for DJ-1 WT, a GO term enrichment analysis was performed for each DJ-1 mutant in the different time points (Supplementary Figure 7.20 to 7.22).

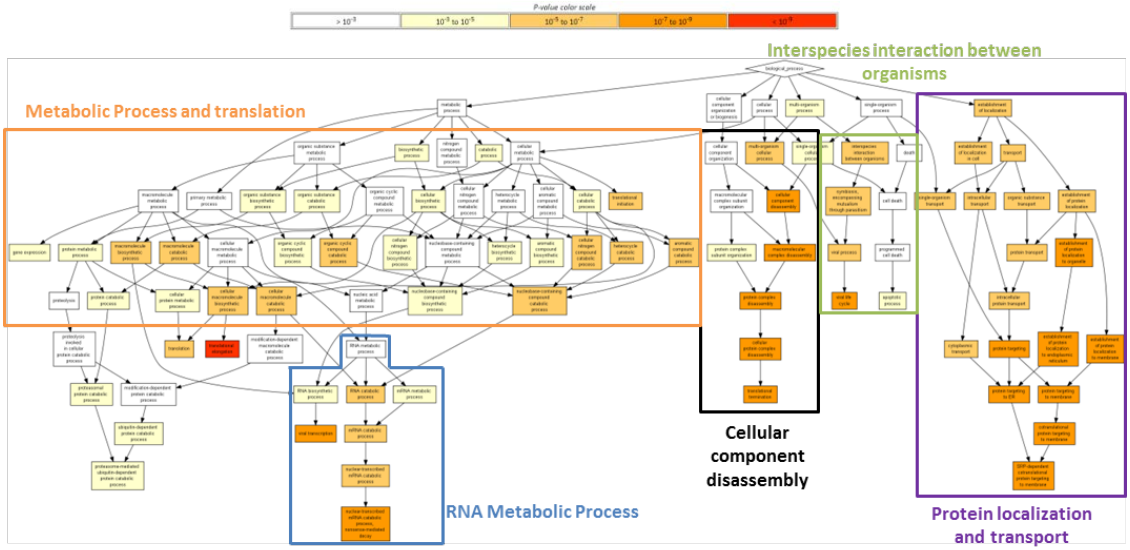


Supplementary Figure 7.20| GENE ONTOLOGY (GO) ENRICHMENT OF THE IDENTIFIED BINDING PARTNERS FOR C106DD MUTANT IN EACH TIME POINT. The main groups of GO terms considered relevant were highlighted using colored boxes. A to C - GO enrichment of the identified binding partners for C106DD mutant at 0, 15 and 40 min, respectively.

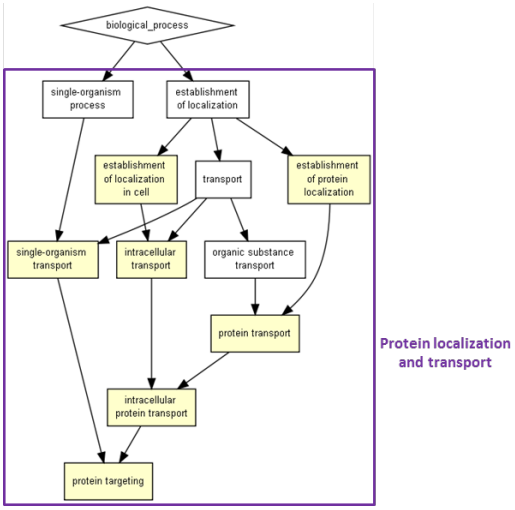


Supplementary Figure 7.21| GENE ONTOLOGY (GO) ENRICHMENT OF THE IDENTIFIED BINDING PARTNERS FOR E163K MUTANT IN EACH TIME POINT. The main groups of GO terms considered relevant were highlighted using colored boxes. A to C - GO enrichment of the identified binding partners for E163K mutant at 0, 15 and 40 min, respectively.

A **15 min**



B **40 min**



Supplementary Figure 7.22| GENE ONTOLOGY (GO) ENRICHMENT OF THE IDENTIFIED BINDING PARTNERS FOR C106A MUTANT IN EACH TIME POINT. The main groups of GO terms considered relevant were highlighted using colored boxes. A and B - GO enrichment of the identified binding partners for C106A mutant at 15 and 40 min, respectively.

Due to the low number of identified proteins for the C106A mutant at 0 minutes condition it was not possible to perform a GO term enrichment analysis. This was only possible for the 15 and 40 minutes conditions and in the later also a reduced number of proteins were identified, which conditioned the results obtained in this analysis leading to the establishment of fewer groups and with smaller GO term enrichment (Supplementary Figure 7.22).

NOTE: The tree diagrams obtained from GO term enrichment analysis are available at high resolution in PNG files (folder “Gorilla - Tree Diagrams”).