

Electronic Thesis and Dissertation Repository

---

6-20-2013 12:00 AM

## Structure of C-terminal Domain of Parkin, IBR-RING2

Yeong Ju Noh, *The University of Western Ontario*

Supervisor: Dr. Shaw, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biochemistry

© Yeong Ju Noh 2013

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Biochemistry Commons](#)

---

### Recommended Citation

Noh, Yeong Ju, "Structure of C-terminal Domain of Parkin, IBR-RING2" (2013). *Electronic Thesis and Dissertation Repository*. 1324.

<https://ir.lib.uwo.ca/etd/1324>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

# Structure of C-terminal Domain of Parkin, IBR-RING2

(Thesis format: Monograph)

by

Yeong Ju Noh

Graduate Program in Biochemistry

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Master of Science

The School of Graduate and Postdoctoral Studies  
The University of Western Ontario  
London, Ontario, Canada

© Yeong Ju Noh, 2013

## **Abstract**

Parkin is an E3 ubiquitin ligase which degrades misfolded proteins and prevents the formation of abnormal protein aggregates often formed in Parkinson's disease. The main goal of this thesis was to perform structural analysis on the IBR(In-Between-RING)-RING2 (Really Interesting New Gene) domain of parkin. After determining the three-dimensional solution structure of the protein by NMR spectroscopy, the RING2 domain was identified to be similar to the IBR domain, showing that it is not a canonical RING domain. The catalytic cysteine on RING2 was also shown to be solvent exposed, supporting the recently proposed RING/HECT hybrid mechanism of parkin as an RBR E3 ligase. The structure also revealed that IBR and RING2 domains do not interact. This was confirmed with two dimensional NMR experiments and split GFP system. The 14 disease-state IBR-RING2 proteins were analyzed using NMR spectroscopy to monitor the structural impact of autosomal recessive juvenile parkinsonism (ARJP) related mutations.

### **Keywords:**

Parkinson's Disease, parkin, ARJP, solution structure, nuclear magnetic resonance, zinc-binding, ubiquitination

## **Acknowledgments**

I would like to thank my supervisor, Dr. Gary Shaw, for the opportunity to work in his lab and his support, guidance, encouragement, and patience throughout my project. Thank you for being the best supervisor one could ever ask for.

I would also like to thank my advisory committee, Dr. David Litchfield and Dr. James Choy for your valuable advice and feedback.

To all the past and present the members of the Shaw lab, I really appreciate your support, assistance, and advice during my graduate studies. I could not have done it without all of you, and it was an honor learning so much from many of you and discussing and working together. Especially to Dr. Pascal Mercier, I would like to show my appreciation for all of his assistance regarding structure calculation and making the structure complete at the end.

Last but not least, I would like to thank my family and friends, especially my parents. Their support and love have always allowed me to push myself to the fullest, and I would like to express my sincere appreciation to them for believing in me.

## Table of Contents

Abstract.....	ii
Acknowledgments.....	iii
Table of Contents.....	iv
List of Tables .....	vii
List of Figures .....	viii
List of Abbreviations and Acronyms.....	x
List of Appendices .....	xii
Chapter 1. Introduction .....	1
1.1 Parkinson’s Disease .....	1
1.2 Ubiquitin Protease System.....	2
1.3 E3 Ubiquitin Ligases.....	4
1.3.1 HECT E3 ligases.....	6
1.3.2 RING E3 ligases .....	6
1.3.3 RBR E3 ligases .....	10
1.4 Parkin.....	11
1.4.1 Interactions of parkin .....	11
1.4.2 Structures of parkin.....	13
1.4.3 ARJP mutations in parkin.....	14
1.5 Thesis Overview .....	16
1.6 References.....	18
Chapter 2. Materials and Methods .....	23
2.1 <i>Drosophila melanogaster</i> Parkin C-terminus, IBR-RING2 .....	23
2.1.1 Design of C-terminal domain constructs of Dm Parkin.....	23
2.1.2 C-terminal Parkin into Split-GFP vector .....	24

2.1.3	Site-directed Mutagenesis of Parkin IBR-RING2 domain .....	24
2.2	Expression of IBR-RING2.....	25
2.3	Purification of IBR-RING2.....	27
2.4	NMR Spectroscopy .....	28
2.4.1	Chemical Shift Assignment of IBR-RING2 .....	28
2.4.2	Secondary Structure Prediction.....	29
2.4.3	T <sub>1</sub> /T <sub>2</sub> experiments .....	30
2.4.4	Heteronuclear NOE.....	30
2.5	NMR Structure Determination.....	31
2.6	Solubility Test.....	31
2.7	References.....	33
Chapter 3.	Results and Discussion .....	35
3.1	GST-IBR-RING2 Expression .....	35
3.2	Determination of Optimal Conditions for NMR Data Collection .....	38
3.3	Structural Determination of the Parkin IBR-RING2 Domain .....	43
3.3.1	ESI-MS shows the IBR-RING2 coordinates 4 Zinc ions .....	43
3.3.2	Backbone Chemical Shift Assignment of IBR-RING2 .....	44
3.3.3	Utilization of IBR-RING2 backbone assignment.....	47
3.3.4	Prediction of Zinc Coordinating Cysteines.....	47
3.3.5	Determination of Secondary Structure for parkin IBR-RING2.....	49
3.3.6	Structure Calculation of IBR-RING2 .....	51
3.4	Interaction Studies of IBR and RING2 domains .....	58
3.4.1	The flexible linker of IBR-RING2 is confirmed with protein dynamics studies .....	58
3.4.2	Confirmation of non-interacting IBR and RING2 domains with <sup>1</sup> H- <sup>15</sup> N HSQC spectra.....	63
3.4.3	Split GFP.....	68

3.5 Mutation analysis .....	72
3.6 Summary .....	83
3.7 Closing Thoughts & Future direction .....	84
3.8 References.....	86
Appendices.....	89

## List of Tables

Table 1. Forward and reverse primer sequences used in the amplification of different domains of parkin's C-terminus and mutants of IBR-RING2 and constructs of split GFP ...	26
Table 2. Cysteines involved in Zn <sup>2+</sup> coordination in IBR-RING2 domain, as determined by chemical shifts of C $\alpha$ and C $\beta$ .....	48
Table 3: Structural Statistics for 20 lowest NOE energy structures of IBR-RING2 .....	54



## List of Figures

Figure 1.1: Ubiquitination pathway showing ubiquitin conjugation. ....	5
Figure 1.2:HECT, RING and RBR type E3 ligases, illustrating the ubiquitin transfer step from E2 to E3.....	7
Figure 1.3:Main features of Canonical RING E3 ligases. ....	9
Figure 1.4: Schematic diagram illustrating the domains of parkin, with three dimensional structures of domains in parkin that are present in the protein databank.....	12
Figure 1.5: Zinc coordination of IBR, showing sequential zinc coordination.....	15
Figure 3.1: SDS-PAGE gel of induced vs uninduced GST-IBR-RING2. ....	37
Figure 3.2: Chromatograms of GST purification.....	39
Figure 3.3: SDS-PAGE gels showing two-step purification of IBR-RING2 .....	40
Figure 3.4: Comparison of $^1\text{H}$ - $^{15}\text{N}$ HSQC Spectra and NMR tubes of 300-400 $\mu\text{M}$ IBR-RING2 after the data collection. ....	42
Figure 3.5: Deconvoluted ESI-MS of proteins under denaturing and non-denaturing conditions, which show the mass of both the native and denatured state of IBR-RING2.....	45
Figure 3.6: Two-dimensional $^1\text{H}$ - $^{15}\text{N}$ HSQC spectrum of $^{15}\text{N}$ -isotopically labeled IBR-RING2.....	46
Figure 3.7: Chemical shift index of IBR-RING2 for secondary structure prediction. ....	50
Figure 3.8: Superposition of 20 structures overlay of IBR-RING2, on each of the IBR (blue) and RING2 (red) domains.....	52
Figure 3.9: Representative ribbon structure of IBR-RING2 with the flexible linker between the two domains. ....	53

Figure 3.10: Ribbon structure of parkin IBR domain with template of zinc binding sites.....	56
Figure 3.11: Ribbon structure of parkin RING2 domain.....	56
Figure 3.12: Steady-state heteronuclear NOE values for backbone amides of <sup>15</sup> N-labeled IBR-RING2, obtained at 600 MHz.....	59
Figure 3.13: Backbone <sup>15</sup> N spin relaxation measurement of IBR-RING2 .....	61
Figure 3.14: Superposition of the IBR-RING2 <sup>1</sup> H- <sup>15</sup> N HSQC spectrum with each individual domain.....	64
Figure 3.15: Superposition of the IBR-RING2 <sup>1</sup> H- <sup>15</sup> N HSQC spectrum with each individual domain, both IBR and RING2 displayed.....	66
Figure 3.16: Titration of RING2 into <sup>15</sup> N labeled IBR.....	67
Figure 3.17: Illustration of split GFP system.....	69
Figure 3.18: Solubility test of split GFP constructs shown as labelled at the top.....	71
Figure 3.19: Soluble fraction of cell lysate split GFP fusions with RBR, IBR-RING2, and S100A8/A9. ....	71
Figure 3.20: Disease state substitutions within the parkin IBR-RING2. Ribbon drawings of different domains of parkin.....	73
Figure 3.21: Solubility test of mutants of IBR-RING2, (missense mutations on RING2 only). ....	74
Figure 3.22: Superposition of <sup>1</sup> H- <sup>15</sup> N HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in IBR region (G349E).....	77
Figure 3.23: Superposition of <sup>1</sup> H- <sup>15</sup> N HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the linker region (D412N).....	79
Figure 3.24: Superposition of <sup>1</sup> H- <sup>15</sup> N HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the RING2 region (C449F). ....	80

## List of Abbreviations and Acronyms

PD	Parkinson's disease
ARJP	Autosomal recessive juvenile parkinsonism
PINK	PTEN-induced kinase I
UPS	Ubiquitin protease system
HECT	Homologous to E6-AP C-terminal domain
RING	Really interesting new gene
RBR	RING-between-RING
UbL	Ubiquitin-like
UPD	Unique parkin domain
IBR	In-between-RING
TEV	Tobacco Etch Virus
GFP	Green fluorescent protein
IPTG	Isopropyl $\beta$ -D thiogalactopyranoside
GST	Glutathione S-Transferase
EDTA	Ethylenediaminetetraacetic acid
NMR	Nuclear magnetic resonance
TOCSY	Total Correlation Spectroscopy
NOESY	Nuclear Overhauser effect spectroscopy
CYANA	Combined Assignment and Dynamics Algorithm for NMR Applications
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
MW	Molecular weight
HSQC	Heteronuclear Single Quantum Coherence
RMSD	Root-mean-square deviation
NOE	Nuclear Overhauser effect

HGMD	Human gene mutation database
DCDB	Domain classification database
<i>E.Coli</i>	<i>Escherichia coli</i>
<i>Dm</i>	<i>Drosophila melanogaster</i>
Ala (A)	Alanine
Arg (R)	Arginine
Asn (N)	Asparagine
Asp (D)	Aspartic acid
Cys (C)	Cysteine
Gln (Q)	Glutamine
Glu (E)	Glutamic acid
Gly (G)	Glycine
His (H)	Histidine
Ile (I)	Isoleucine
Leu (L)	Leucine
Lys (K)	Lysine
Met (M)	Methionine
Phe (F)	Phenylalanine
Pro (P)	Proline
Ser (S)	Serine
Thr (T)	Threonine
Trp (W)	Tryptophan
Tyr (Y)	Tyrosine
Val (V)	Valine

## **List of Appendices**

Appendix A: Data from Structure Calculation of IBR-RING2 .....	89
Appendix B: Data from Dynamics Studies of IBR-RING2.....	90
Appendix C: Data from Mutational Analysis of IBR-RING2 .....	92
Appendix D. Multiple Sequence Alignment of IBR-RING2.....	103

# Chapter 1

## Introduction

### 1.1 Parkinson's Disease

Parkinson's disease (PD) is the most prevalent movement disorder and second most common neurodegenerative disease, affecting more than 100,000 Canadians. Over the span of next years, the number of PD patients is expected to double aged 65 and over. PD is caused by the death of neurons responsible for dopamine production in the substantia nigra (Fearnley and Lees, 1990) and common clinical diagnosis of PD patients includes bradykinesia, rigidity, resting tremors and postural instability. Interestingly, the loss of smell is also a commonly noted symptom by PD patients in retrospect (Doty et al., 1995). PD patients start experiencing symptoms at approximately 55 years old, with these become increasingly prevalent. The molecular mechanisms involved in the pathogenesis of the disorder are not clearly understood, except for their proposed relation to aging, environment, and genetic predisposition (Mattson et al., 2002).

There are several categories of Parkinson's disease: sporadic, familial, and symptomatic PDs. The majority of PD patients fall into the sporadic form of Parkinsonism (80%). Familial or genetic forms of Parkinson's disease affect 10% of PD patients and can be autosomal dominant (mutations on *PARK1/4*, *PARK8* genes) or autosomal recessive (mutations on *PARK2*, *PARK6*, *PARK7*, and *PARK9* genes) (Bonifati et al., 2003; Funayama et al., 2002; Kitada et al., 1998; Matsumine et al., 1998; Valente et al., 2004; van Duijn et al., 2001; Zimprich et al., 2004). The current body of knowledge on PD cannot identify the direct molecular causes of the disease. Many of the

symptoms of sporadic PD are similar to those of the familial early-onset form of PD, including autosomal recessive juvenile Parkinsonism (ARJP), where early-onset occurs before the age of 40. A hallmark of idiopathic PD is the observation of cytoplasmic insoluble protein aggregates (Lewy bodies) containing  $\alpha$ -synuclein encoded by *PARK1/4*, which also occurs in autosomal dominant forms of PD (Marin et al., 2004). Interestingly, for the ARJP, Lewy bodies are rarely present (Mori et al., 1998; Takahashi et al., 1994).

The most common form of ARJP (50%) is linked to mutations on the *PARK2* gene on chromosome 6 (Kitada et al., 1998; Matsumine et al., 1998). This gene codes for a 465-residue-protein known as parkin. Other less frequent ARJP related proteins include; PTEN-induced kinase I (PINK), encoded by the *PARK6* gene, and DJ-1 encoded by the *PARK7* gene (Bonifati et al., 2003; Valente et al., 2004). Also, neural loss in ARJP patients is identified to be associated with dysfunction of parkin (Takahashi et al., 1994). The loss of function in parkin due to mutation has been suggested to cause problems for PD patients, leading to the potential neurotoxic accumulation (Sang et al., 2007).

Parkin belongs to the family of E3 ubiquitin-protein ligases in the ubiquitin protease system (UPS), and is responsible for the degradation of misfolded proteins by the proteasome, which prevents the formation of abnormal protein aggregates such as those formed in PD (Shimura et al., 2000).

## **1.2 Ubiquitin Protease System**

For proper survival of cells, it is important to maintain high quality control over proteins by removing short-lived, misfolded, or damaged proteins from the system (Kubota, 2009). Proteasomal degradation is essential for this purpose, and one of the

most well-known mechanisms for this is the ubiquitination proteasome system (UPS) (Lehman, 2009). The UPS is characterized by the transfer of ubiquitin, a 76-residue protein, through a chain of several enzymes until it is covalently attached to a substrate.

Ubiquitin, one of the most highly conserved proteins present in eukaryotic cells, is involved in various cellular processes including: transcription (Pickart, 1997), cell-cycle (Imai et al., 2000), endocytosis (Hicke and Dunn, 2003), DNA repair (Barbour et al., 2006; Hershko and Ciechanover, 1998), and proteolysis (Glickman and Ciechanover, 2002). Therefore, it is not a surprise that when the ubiquitin system encounters abnormalities, many diseases can arise, including but not limited to: Angelman Syndrome, Cystic Fibrosis, and neurodegenerative diseases such as Alzheimer's or Parkinson's Disease (Glickman and Ciechanover, 2002).

Ubiquitination is responsible for joining the C-terminal glycine of ubiquitin to the side chain amine of a lysine within the protein substrate, thereby forming an isopeptide bond. There are seven lysine residues in ubiquitin (K6, K11, K27, K29, K33, K48, and K63) that can form ubiquitin chains (Haglund and Stenmark, 2006). Lys48 and Lys63 are the most common ubiquitination sites, with Lys48 in particular being responsible for proteolytic recognition by the 26S proteasome. And since many enzymes and substrates involved in the ubiquitination pathway have been implicated in cancer, neurological, metabolic and inflammatory disorders, they are attractive targets for therapeutic intervention. Labeling substrates with a single ubiquitin molecule can signal for cellular regulation (Di Fiore et al., 2003; Hicke, 2001)). In contrast, building of ubiquitin chains, with at least four ubiquitin molecules linked through Lys48, signals for protein degradation.

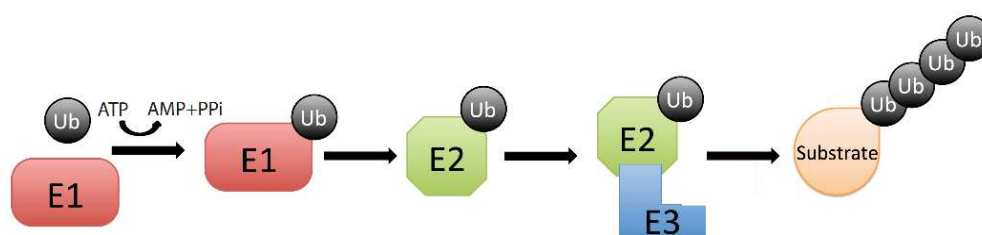


The UPS conjugates ubiquitin through a highly conserved mechanism involving a cascade of enzymes, E1: ubiquitin activating, E2: ubiquitin conjugating, and E3: ubiquitin ligating enzymes, as shown in Figure 1.1. The transfer of ubiquitin is initiated by the ubiquitin activating enzyme (in an ATP-dependent manner) forming a thiolester bond between E1 and ubiquitin. Then, the ubiquitin is transferred to an E2 enzyme through a transthioesterification reaction, which releases E1. The process is complete when the E2-ubiquitin complex associates with the E3 enzyme in order to transfer the ubiquitin to the target protein (Huang et al., 2007).

In the ubiquitination process, E3 enzymes are far more specific and abundant than the other proteins involved (Schwartz and Ciechanover, 2009). Ubiquitination of a particular substrate is thought to recognize a particular E2:E3 combination. In the human genome of substrates, there are two E1s, over 38 E2s, and about 600 to 1000 E3s, known to date that allow ubiquitination.

### **1.3 E3 Ubiquitin Ligases**

Proper ubiquitination of a specific substrate is achieved by having a diverse group of E3 enzymes. There are two major categories of E3 ubiquitin ligases in eukaryotes: the Homologous to E6AP C-Terminus (HECT) type and Really Interesting New Gene (RING) type ligases. The primary difference between the two E3 ubiquitin ligases lies in the process of transferring of ubiquitin onto the substrate. HECT E3 enzymes form a thiolester bond between the ubiquitin and itself first, while the RING E3 ligases allow for



**Figure 1.1: Ubiquitination pathway showing ubiquitin conjugation.**

Ubiquitin is activated by the E1 ubiquitin-activating enzyme in an ATP-dependent manner. Ubiquitin attaches to E1 forming a thiolester bond. Subsequently, it is transferred to the E2 ubiquitin-conjugating enzyme through a transthiolester reaction. Ubiquitin conjugated E2 interacts with an E3 ligase, bringing the substrate closer to the E2, in order for the ubiquitin to attach to the lysine of the substrate. Several repetitions of the ubiquitin conjugation result in ubiquitin chain buildup on the substrate, which signals protein degradation in the UPS.

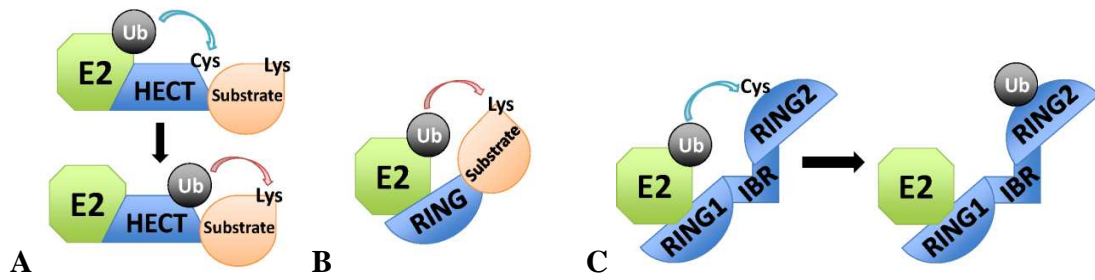
the direct transfer of ubiquitin from the E2 enzyme to the substrate (Passmore and Barford, 2004). Figure 1.2 illustrates the ubiquitin transfer models for the HECT and RING E3 ligases. Catalytic residues have not been identified for RING type ligases, and are suspected not to exist.

### ***1.3.1 HECT E3 ligases***

The HECT domain (about 350 residues long) is usually located at the C-terminus of these E3 ligases, identified based on its similarity to the E6 associated protein (E6AP). A conserved catalytic cysteine (reactive C-terminus) is contained in the protein database in the HECT E3 family, which can form a thiolester bond with ubiquitin. The N-terminus of the HECT domain serves as the E2 binding domain. There are about 30 HECT type E3 ligases, including E6AP, and Nedd4, all of which are known to play roles in protein trafficking, immune response, and regulation of cell growth through involvement in signaling (Rotin and Kumar, 2009).

### ***1.3.2 RING E3 ligases***

The RING type E3 ligases represent the majority of the E3 enzyme family. The mechanism of substrate ubiquitination for RING E3 ligases differs from that observed in the HECT type, as it directly transfers ubiquitin from an E2 to the lysine of a substrate, forming an isopeptide bond without thiolester bond formation (Figure 1.2). RING domains were first discovered as an integral part of several multi-domain protein

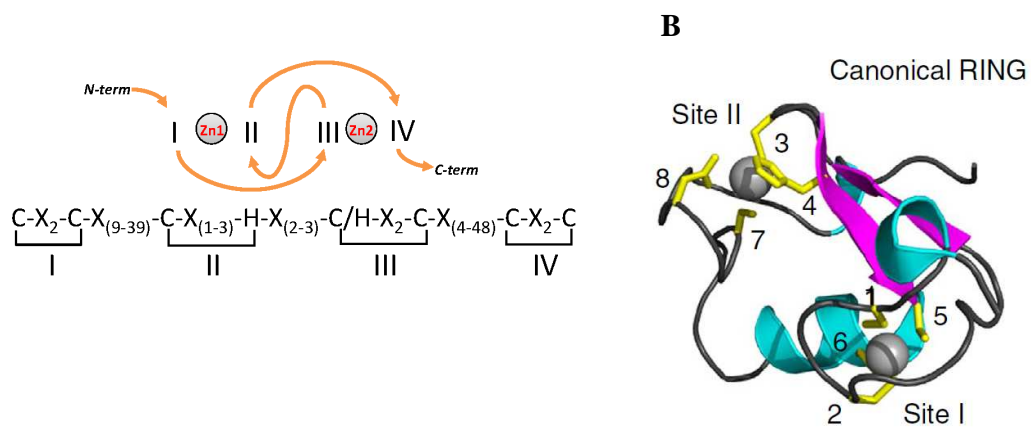


**Figure 1.2:HECT, RING and RBR type E3 ligases, illustrating the ubiquitin transfer step from E2 to E3.**

**A.** HECT type ubiquitin transfer to substrate, involving the catalytic cysteine of E3 ligase (blue & red arrows), **B.** RING type ubiquitin transfer to substrate; E3 does not get directly involved in ubiquitination (red arrow only), **C.** RBR E3 ligase showing hybrid of HECT and RING mechanism. (Adapted from Wenzel and Klevit., 2012)

complexes, including Skp1-cullin-F-box (SCF), and Cullin-elongin B and C (CBC) complexes. Rbx1, the RING E3 ligase was later found to be part of the SCF complexes. RING domains can be located anywhere throughout the E3 ligase sequence and are characterized by the consensus sequence C-X<sub>2</sub>-C-X<sub>(9-39)</sub>-C-X<sub>(1-3)</sub>-H-X<sub>(2-3)</sub>-C/H-X<sub>2</sub>-C-X<sub>(4-48)</sub>-C-X<sub>2</sub>-C, where X represents any given amino acid. The regular pattern of cysteines and histidines allows for the binding of two zinc ions, which stabilizes the protein structure. This coordination is referred to as a cross brace-like motif and the first and third pairs of Cys and second and fourth pairs of Cys/His allow for the proper coordination of zinc within the structure (Figure 1.3).

Some examples of simple RING type ligases linked to diseases are breast cancer 1 (BRCA1), and parkin. Additionally, RING E3 ligases have been found to form dimers, such as BRCA1/BARD1 (Brzovic et al., 2001) and Ring1b/Bmil (Buchwald et al., 2006). The formation of the dimer was suspected to bring the E2 close to the substrate lysine and aid in isopeptide bond formation between ubiquitin and the substrate. Another interesting type of RING E3 ligases are RING-between-RING (RBR) type E3 ligases. The name was assigned to these enzymes due to the presence of an additional RING domain (RING2) at the C-terminus, as well as an in-between-RING domain (IBR) on top of the canonical N-terminal RING domain (RING1). The RING1, IBR and RING2 domains in the RBR E3 ligases contain a repeating pattern of cysteine and histidine residues similar to that modeled in the consensus RING sequence and so were thought to function in typical RING E3 ligase manner.



**Figure 1.3: Main features of Canonical RING E3 ligases.**

**A.** Cross-brace zinc coordination of RING E3 ligase, **B.** Canonical structure of RING, represented by TRAF6 (PDB 3HCS) (Adapted from Spratt et al., 2013).

### **1.3.3 RBR E3 ligases**

Some members of the RBR ligases include: human homolog of *Drosophila* Ariadne (HHARI), heme-oxidized-IRP2 ubiquitin ligase 1(HOIL-1), HOIL-1-interacting protein (HOIP), and parkin (Lucas et al., 2006; Marin et al., 2004). These RBR domains have been demonstrated to be involved in translation and immune signaling. Recently, it has been proposed that RBR E3 ligases might act as a hybrid of RING and HECT-type E3s (as shown in Figure 1.2).

The RBR E3 ligase, HHARI has recently been shown to function as a hybrid of RING/HECT E3 ligases with the E2 conjugating enzyme UbcH7 using GST-pulldown and autoubiquitination assays (Wenzel et al., 2011). In this mechanism the RING2 recruits a ubiquitinated E2 but transfers the ubiquitin to the catalytic cysteine of the RING2 domain. Within the RBR family, there is a conserved cysteine in the RING2 domain, which is not present in RING1/IBR domains or RING E3 ligases. In HHARI, substituting the catalytic cysteine (C357) to alanine or serine eliminated the transfer of ubiquitin to the RING2 domain of E3 and subsequent ubiquitination, showing the importance of this residue for RBR E3 ligase activity. Because of these unique features of the RBR E3 ligase, structural studies of RBR E3 ligases would provide insight to understanding the overall mechanism.

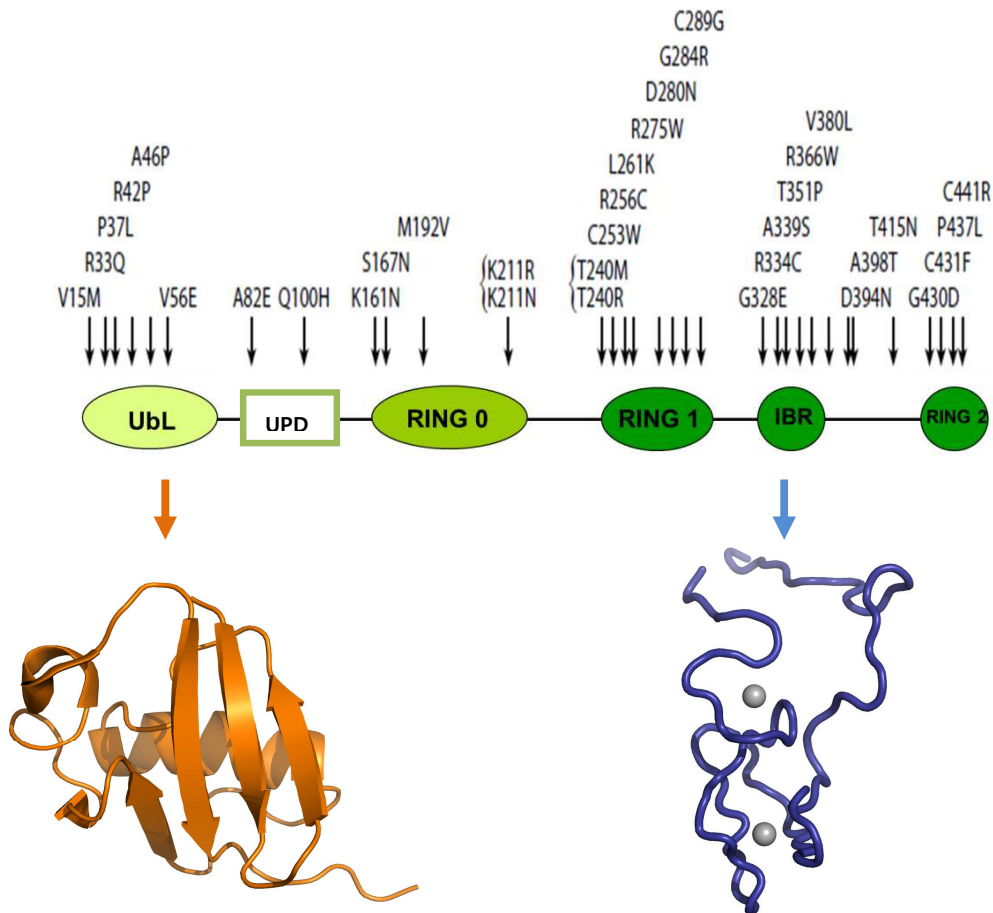
## 1.4 Parkin

Parkin is the most studied RBR E3 ligase, due to its relation to Parkinson's disease (PD). Parkin is an RBR-E3 ligase that has multiple domains containing numerous point mutations that are responsible for 50% or more of ARJP cases (Figure 1.4). The *N*-terminus of parkin contains an ubiquitin-like (UbL) domain and it spans from 1-76 residues. This domain has been proposed to be important in the recruitment of proteasome (Sakata et al., 2003; Tashiro et al., 2003). The UbL domain is followed by the unique parkin domain (UPD) that is only present in parkin, and is approximately 70 residues long. The *C*-terminus of parkin comprised the RBR domains with a RING0 domain preceding it (Beasley et al., 2007; Hristova et al., 2009; Morett and Bork, 1999). Of these domains, only the UbL and IBR domains have been successfully purified and had their three dimensional structures solved (Beasley et al., 2007; Hristova et al., 2009; Morett and Bork, 1999).

### 1.4.1 *Interactions of parkin*

As an E3 ligase, parkin is expected to interact with E2 enzymes. Another member of the RBR E3 ligase, Human Homologue of Ariadne (HHARI), was first recognized to have association with E2 enzymes, Ubch7 and Ubch8, and RING1 of HHARI was identified to be their interacting partner (Ardley et al., 2001; Moynihan et al., 1999). However, when tested with parkin, conflicting data were presented. Zhang et al. (2000) reported that parkin interacts with Ubch8, but not with Ubch7 from immunoprecipitation experiments, while Imai et al. (2000) observed the exact opposite. With co-immunoprecipitation experiments, Shimura and colleagues (2000) identified Ubch7 as a





**Figure 1.4: Schematic diagram illustrating the domains of parkin, with three dimensional structures of domains in parkin that are present in the protein databank.**

Parkin has UbL domain at the *N*-terminus, unique parkin domain (UPD) and RING0 domains in the middle, and RING1, IBR, and RING2 (RBR) at the *C*-terminus. Some of the amino acid substitutions related to ARJP are indicated with black arrows. Orange ribbon diagram shows UbL domain (PDB: 1IYF) and blue ribbon diagram shows IBR (PDB: 2JMO) (Adapted from von Coelln et al., 2004)

binding partner of RING1. More recently, a heterodimeric form of the E2 enzyme Ubc13/Uev2a was suggested as an interaction partner of parkin (Doss-Pepe et al., 2005; Matsuda et al., 2006). Based on the inconsistency of previous work, it still remains to be seen which RING domains in parkin are important for recruitment of E2 enzymes and control of ubiquitination.

#### **1.4.2 Structures of parkin**

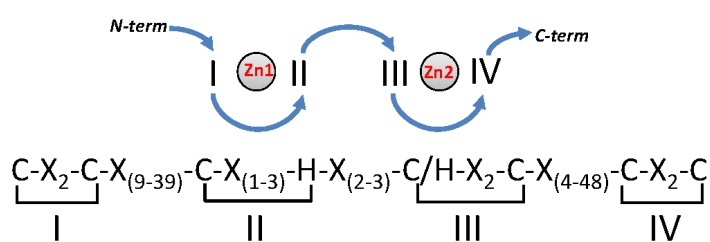
The three-dimensional structure of the UbL domain present at the *N*-terminus of parkin has been solved by nuclear magnetic resonance (NMR) spectroscopy. The structure shows similarity to ubiquitin, possessing the  $\beta$ -grasp fold comprised of five  $\beta$ -sheets and two  $\alpha$ -helices, as shown in Figure 1.4 (Sakata et al., 2003). A distinct feature of this domain of parkin is that it lacks a C-terminal glycine, so that unlike ubiquitin, it cannot conjugate to the lysine of a substrate. Instead, the UbL domain is followed by the UPD, a region postulated to be a disordered linker between the UbL domain and RING0-RBR domains with no obvious other function.

There is not much known about the structures of the RING domains of parkin. However, the proposed linker, the IBR domain, which is situated between the RING1 and RING2 domains has been solved by NMR spectroscopy, as shown in Figure 1.4 (Beasley et al., 2007). There is no distinct secondary structure in the IBR. It contains two zinc ions with a bilobal fold around the zinc coordinating sites. Compared to the canonical cross-brace zinc coordination, the IBR domain is known to coordinate zinc ions “sequentially”, as described by Beasley and colleagues (2007). The sequential zinc coordination is

shown in Figure 1.5, and illustrates that the first and second pairs of cysteines and the third and fourth pairs of cysteines/histidines coordinate zinc. The IBR was proposed to act like a bridge that brings RING2 and RING2 domains together in the full-length parkin assembly, with an expectation that those two are interacting domains. Though the three-dimensional structure of the IBR domain of human parkin has been solved, structures of the RING domains and more importantly, the assembly of the entire C-terminal RBR motif, still remain unknown. It is expected that structural analysis of these domains can answer some questions regarding the catalytic mechanism of parkin in ubiquitination.

#### ***1.4.3 ARJP mutations in parkin***

Today, more than 150 mutations on parkin have been identified in PD patients and some of the missense mutations are illustrated in Figure 1.4. It is unfortunate, however, that a full understanding of how these impact parkin function as an E3 ligase is not known, even with the common observation of its impact on the ubiquitination process. Amino acid substitutions, involving for instance R42P, K161N, Q328E, C341F, G430D, C441R, and W453X reduce protein solubility suggesting the folding of parkin is impacted (Sriram et al., 2005; Wang et al., 2005). Further, R42P, C212Y, Q311X, C341F, and W453X are found in the cytoplasmic inclusions, suggesting that they are unfolded in cells (Sriram et al., 2005; Wang et al., 2005). Auto-ubiquitination assays showed that several substitutions such as C418R and C441R impact parkin's function suggesting zinc binding in the RING2 domain is disrupted (Sriram et al., 2005).



**Figure 1.5: Zinc coordination of IBR, showing sequential zinc coordination.** First zinc site involves first and second pairs of Cys, and the second zinc site is surrounded by third and fourth pairs of Cys/His.

However, other point mutations do not show whether ARJP is caused by a loss of its E3 activity. It is not known how parkin becomes dysfunctional with substitution, or how these alterations affect the protein's structure or stability. To have a better understanding of the impact of these mutations, structure and interaction studies with parkin would be necessary.

## 1.5 Thesis Overview

Mutations in the *PARK2* gene that codes for parkin are strongly linked to ARJP and have many different effects on RBR E3 ligase activity. However, given the structure determination of the parkin IBR domain (Beasley et al., 2007), the roles of the RING1 and RING2 domains in either a RING or RING/HECT E3 ligase mechanism have neither been supported nor opposed by structural analysis. The goal of this project was to determine the three-dimensional structure of IBR-RING2 in order to have a better understanding of parkin's assembly and function. The proposed hypothesis is that the IBR and RING2 of parkin are likely to be interacting, even with the missing RING1 domain, to facilitate ubiquitin transfer to a substrate.

Specific experiments designed to address this hypothesis are:

- 1) Determine the three-dimensional solution structure of parkin IBR-RING2 using NMR spectroscopy.
- 2) Identify whether the IBR and RING2 domains are interacting using binding experiments between IBR-RING2 and individual IBR or RING2 domains

utilizing NMR spectroscopy, NMR relaxation experiments, and a split-GFP system.

- 3) Analyze the impact on the solubility and structure of IBR-RING2 domains caused by mutations (ARJP substitutions) using NMR spectroscopy.

The research outlined in this thesis will provide a better understanding of the function of parkin, by providing the three dimensional structure of the IBR-RING2 domains. The properties relating to their structure and interactions will be useful to propose mechanistic steps for ubiquitination and the pathogenesis in ARJP. Hopefully, this will contribute to therapeutic drug development for patients.

## 1.6 References

- Ardley, H.C., Tan, N.G., Rose, S.A., Markham, A.F., and Robinson, P.A. (2001). Features of the parkin/ariadne-like ubiquitin ligase, HHARI, that regulate its interaction with the ubiquitin-conjugating enzyme, Ubch7. *J Biol Chem* 276, 19640-19647.
- Barbour, L., Ball, L.G., Zhang, K., and Xiao, W. (2006). DNA damage checkpoints are involved in postreplication repair. *Genetics* 174, 1789-1800.
- Beasley, S.A., Hristova, V.A., and Shaw, G.S. (2007). Structure of the Parkin in-between-ring domain provides insights for E3-ligase dysfunction in autosomal recessive Parkinson's disease. *Proc Natl Acad Sci U S A* 104, 3095-3100.
- Bedford, L., Lowe, J., Dick, L.R., Mayer, R.J., and Brownell, J.E. (2011). Ubiquitin-like protein conjugation and the ubiquitin-proteasome system as drug targets. *Nat Rev Drug Discov* 10, 29-46.
- Bonifati, V., Rizzu, P., van Baren, M.J., Schaap, O., Breedveld, G.J., Krieger, E., Dekker, M.C., Squitieri, F., Ibanez, P., Joosse, M., *et al.* (2003). Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *Science* 299, 256-259.
- Brzovic, P.S., Meza, J.E., King, M.C., and Klevit, R.E. (2001). BRCA1 RING domain cancer-predisposing mutations. Structural consequences and effects on protein-protein interactions. *J Biol Chem* 276, 41399-41406.
- Buchwald, G., van der Stoop, P., Weichenrieder, O., Perrakis, A., van Lohuizen, M., and Sixma, T.K. (2006). Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. *EMBO J* 25, 2465-2474.
- Di Fiore, P.P., Polo, S., and Hofmann, K. (2003). When ubiquitin meets ubiquitin receptors: a signalling connection. *Nat Rev Mol Cell Biol* 4, 491-497.
- Doss-Pepe, E.W., Chen, L., and Madura, K. (2005). Alpha-synuclein and parkin contribute to the assembly of ubiquitin lysine 63-linked multiubiquitin chains. *J Biol Chem* 280, 16619-16624.
- Doty, R.L., Bromley, S.M., and Stern, M.B. (1995). Olfactory testing as an aid in the diagnosis of Parkinson's disease: development of optimal discrimination criteria. *Neurodegeneration* 4, 93-97.
- Fearnley, J.M., and Lees, A.J. (1990). Striatonigral degeneration. A clinicopathological study. *Brain* 113 ( Pt 6), 1823-1842.
- Funayama, M., Hasegawa, K., Kowa, H., Saito, M., Tsuji, S., and Obata, F. (2002). A new locus for Parkinson's disease (PARK8) maps to chromosome 12p11.2-q13.1. *Ann Neurol* 51, 296-301.

- Glickman, M.H., and Ciechanover, A. (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol Rev* 82, 373-428.
- Haglund, K., and Stenmark, H. (2006). Working out coupled monoubiquitination. *Nat Cell Biol* 8, 1218-1219.
- Hershko, A., and Ciechanover, A. (1998). The ubiquitin system. *Annu Rev Biochem* 67, 425-479.
- Hicke, L. (2001). Protein regulation by monoubiquitin. *Nat Rev Mol Cell Biol* 2, 195-201.
- Hicke, L., and Dunn, R. (2003). Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annu Rev Cell Dev Biol* 19, 141-172.
- Hristova, V.A., Beasley, S.A., Rylett, R.J., and Shaw, G.S. (2009). Identification of a novel Zn<sup>2+</sup>-binding domain in the autosomal recessive juvenile Parkinson-related E3 ligase parkin. *J Biol Chem* 284, 14978-14986.
- Huang, D.T., Hunt, H.W., Zhuang, M., Ohi, M.D., Holton, J.M., and Schulman, B.A. (2007). Basis for a ubiquitin-like protein thioester switch toggling E1-E2 affinity. *Nature* 445, 394-398.
- Imai, Y., Soda, M., and Takahashi, R. (2000). Parkin suppresses unfolded protein stress-induced cell death through its E3 ubiquitin-protein ligase activity. *J Biol Chem* 275, 35661-35664.
- Kitada, T., Asakawa, S., Hattori, N., Matsumine, H., Yamamura, Y., Minoshima, S., Yokochi, M., Mizuno, Y., and Shimizu, N. (1998). Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* 392, 605-608.
- Kubota, H. (2009). Quality control against misfolded proteins in the cytosol: a network for cell survival. *J Biochem* 146, 609-616.
- Lehman, N.L. (2009). The ubiquitin proteasome system in neuropathology. *Acta Neuropathol* 118, 329-347.
- Lucas, J.I., Arnau, V., and Marin, I. (2006). Comparative genomics and protein domain graph analyses link ubiquitination and RNA metabolism. *J Mol Biol* 357, 9-17.
- Marin, I., Lucas, J.I., Gradilla, A.C., and Ferrus, A. (2004). Parkin and relatives: the RBR family of ubiquitin ligases. *Physiol Genomics* 17, 253-263.
- Matsuda, N., Kitami, T., Suzuki, T., Mizuno, Y., Hattori, N., and Tanaka, K. (2006). Diverse effects of pathogenic mutations of Parkin that catalyze multiple monoubiquitylation in vitro. *J Biol Chem* 281, 3204-3209.



- Matsumine, H., Yamamura, Y., Hattori, N., Kobayashi, T., Kitada, T., Yoritaka, A., and Mizuno, Y. (1998). A microdeletion of D6S305 in a family of autosomal recessive juvenile parkinsonism (PARK2). *Genomics* *49*, 143-146.
- Mattson, M.P., Chan, S.L., and Duan, W. (2002). Modification of brain aging and neurodegenerative disorders by genes, diet, and behavior. *Physiol Rev* *82*, 637-672.
- Metzger, M.B., Hristova, V.A., and Weissman, A.M. (2012). HECT and RING finger families of E3 ubiquitin ligases at a glance. *J Cell Sci* *125*, 531-537.
- Morett, E., and Bork, P. (1999). A novel transactivation domain in parkin. *Trends Biochem Sci* *24*, 229-231.
- Mori, H., Kondo, T., Yokochi, M., Matsumine, H., Nakagawa-Hattori, Y., Miyake, T., Suda, K., and Mizuno, Y. (1998). Pathologic and biochemical studies of juvenile parkinsonism linked to chromosome 6q. *Neurology* *51*, 890-892.
- Moynihan, T.P., Ardley, H.C., Nuber, U., Rose, S.A., Jones, P.F., Markham, A.F., Scheffner, M., and Robinson, P.A. (1999). The ubiquitin-conjugating enzymes UbcH7 and UbcH8 interact with RING finger/IBR motif-containing domains of HHARI and H7-AP1. *J Biol Chem* *274*, 30963-30968.
- Passmore, L.A., and Barford, D. (2004). Getting into position: the catalytic mechanisms of protein ubiquitylation. *Biochem J* *379*, 513-525.
- Pickart, C.M. (1997). Targeting of substrates to the 26S proteasome. *FASEB J* *11*, 1055-1066.
- Ross, C.A., and Pickart, C.M. (2004). The ubiquitin-proteasome pathway in Parkinson's disease and other neurodegenerative diseases. *Trends Cell Biol* *14*, 703-711.
- Rotin, D., and Kumar, S. (2009). Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol* *10*, 398-409.
- Sakata, E., Yamaguchi, Y., Kurimoto, E., Kikuchi, J., Yokoyama, S., Yamada, S., Kawahara, H., Yokosawa, H., Hattori, N., Mizuno, Y., *et al.* (2003). Parkin binds the Rpn10 subunit of 26S proteasomes through its ubiquitin-like domain. *EMBO Rep* *4*, 301-306.
- Sang, T.K., Chang, H.Y., Lawless, G.M., Ratnaparkhi, A., Mee, L., Ackerson, L.C., Maidment, N.T., Krantz, D.E., and Jackson, G.R. (2007). A Drosophila model of mutant human parkin-induced toxicity demonstrates selective loss of dopaminergic neurons and dependence on cellular dopamine. *J Neurosci* *27*, 981-992.
- Spratt, D.E., Julio Martinez-Torres, R., Noh, Y.J., Mercier, P., Manczyk, N., Barber, K.R., Aguirre, J.D., Burchell, L., Purkiss, A., Walden, H., *et al.* (2013). A molecular explanation for the recessive nature of parkin-linked Parkinson's disease. *Nat Commun* *4*, 1983.

Schwartz, A.L., and Ciechanover, A. (2009). Targeting proteins for destruction by the ubiquitin system: implications for human pathobiology. *Annu Rev Pharmacol Toxicol* 49, 73-96.

Shimura, H., Hattori, N., Kubo, S., Mizuno, Y., Asakawa, S., Minoshima, S., Shimizu, N., Iwai, K., Chiba, T., Tanaka, K., *et al.* (2000). Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat Genet* 25, 302-305.

Sriram, S.R., Li, X., Ko, H.S., Chung, K.K., Wong, E., Lim, K.L., Dawson, V.L., and Dawson, T.M. (2005). Familial-associated mutations differentially disrupt the solubility, localization, binding and ubiquitination properties of parkin. *Hum Mol Genet* 14, 2571-2586.

Takahashi, H., Ohama, E., Suzuki, S., Horikawa, Y., Ishikawa, A., Morita, T., Tsuji, S., and Ikuta, F. (1994). Familial juvenile parkinsonism: clinical and pathologic study in a family. *Neurology* 44, 437-441.

Tashiro, K., Tamada, S., Kuwabara, N., Komiya, T., Takekida, K., Asai, T., Iwao, H., Sugimura, K., Matsumura, Y., Takaoka, M., *et al.* (2003). Attenuation of renal fibrosis by proteasome inhibition in rat obstructive nephropathy: possible role of nuclear factor kappaB. *Int J Mol Med* 12, 587-592.

Valente, E.M., Abou-Sleiman, P.M., Caputo, V., Muqit, M.M., Harvey, K., Gispert, S., Ali, Z., Del Turco, D., Bentivoglio, A.R., Healy, D.G., *et al.* (2004). Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science* 304, 1158-1160.

van Duijn, C.M., Dekker, M.C., Bonifati, V., Galjaard, R.J., Houwing-Duistermaat, J.J., Snijders, P.J., Testers, L., Breedveld, G.J., Horstink, M., Sandkuijl, L.A., *et al.* (2001). Park7, a novel locus for autosomal recessive early-onset parkinsonism, on chromosome 1p36. *Am J Hum Genet* 69, 629-634.

von Coelln, R., Dawson, V.L., and Dawson, T.M. (2004). Parkin-associated Parkinson's disease. *Cell Tissue Res* 318, 175-184.

Wang, C., Tan, J.M., Ho, M.W., Zaiden, N., Wong, S.H., Chew, C.L., Eng, P.W., Lim, T.M., Dawson, T.M., and Lim, K.L. (2005). Alterations in the solubility and intracellular localization of parkin by several familial Parkinson's disease-linked point mutations. *J Neurochem* 93, 422-431.

Wenzel, D.M., Lissounov, A., Brzovic, P.S., and Klevit, R.E. (2011). UBC7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature* 474, 105-108.

Zhang, Y., Gao, J., Chung, K.K., Huang, H., Dawson, V.L., and Dawson, T.M. (2000). Parkin functions as an E2-dependent ubiquitin-protein ligase and promotes the degradation of the synaptic vesicle-associated protein, CDCrel-1. *Proc Natl Acad Sci U S A* 97, 13354-13359.

Zimprich, A., Biskup, S., Leitner, P., Lichtner, P., Farrer, M., Lincoln, S., Kachergus, J., Hulihan, M., Uitti, R.J., Calne, D.B., *et al.* (2004). Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* 44, 601-607.

## Chapter 2

### Materials and Methods

#### 2.1 *Drosophila melanogaster* Parkin C-terminus, IBR-RING2

The DNA encoding *Drosophila melanogaster* (*Dm*) parkin (*PARK2*) in pGEX vector was a gift from Dr. Brian Staveley (Memorial University of Newfoundland, Newfoundland and Labrador). This gene was previously cloned into a modified pGEX-6P-2 vector with a Tobacco Etch Virus (TEV)-cleavage site (ENLYFQ(G/S)) by Noah Manczyk (summer student, 2011). Site-directed mutagenesis was performed using primers that are complementary to each other which would bind to the end of TEV-cleavage site as well as the beginning of the IBR (starting from E342) sequence. This generated an *N*-terminal deletion construct, leaving only the IBR-RING2 domain from *Dm* parkin (residues 342-482). The parkin IBR-RING2 construct was verified by DNA sequencing (Robarts Research Institute).

##### 2.1.1 *Design of C-terminal domain constructs of Dm Parkin*

Based on the domain structure of rat parkin identified by limited proteolysis (Hristova et al., 2009), the coding regions for RING1, RING1-IBR and IBR of *Dm* parkin were inserted into a modified pGEX-6P-2 vector with a 5' TEV cleavage site. PCR reactions were performed under standard conditions using Hot Start DNA polymerase (MBI Fermentas) and a touch-down PCR protocol. A total of 50  $\mu$ L reaction mixture was used for all PCR reactions, consisting of PCR buffer with ammonium sulphate (MBI Fermentas), 2.5 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 100 pmol primers, and 1 U of DNA polymerase. All primers used in the PCR reactions are shown in Table 1. The PCR

products were digested with *Bam*HI and *Eco*RI and ligated into the equivalent sites of the modified pGEX-6P-2. DNA sequencing verified the correct sequences of the constructs (Robarts Research Institute). The plasmids were then transformed into *E.coli* JM109 and BL21(DE3) Codon PlusRIL cell lines.

### **2.1.2 C-terminal Parkin into Split-GFP vector**

RING1-IBR-RING2 and IBR-RING2 constructs of *Dm* parkin were inserted into a pETDuet dual expression plasmid containing the *N*- and *C*-terminal halves of GFP under identical T7-promoters (Ghosh et al., 2000). The cloning resulted in the creation of two constructs; (1) *N*-terminal GFP-parkin RING1-IBR-RING2-*C*-terminal GFP, and (2) *N*-terminal GFP-parkin IBR-RING2-*C*-terminal GFP. PCR reactions were performed as described previously (2.1.1) and the primers used for the insertion are listed in Table 1. The PCR products were then digested with *Bam*HI and *Kpn*I and ligated into the equivalent sites of the split-GFP vector. After sequencing, the vectors were transformed into *E. coli* BL21 BL21 (DE3) Codon PlusRIL expression cell line.

### **2.1.3 Site-directed Mutagenesis of Parkin IBR-RING2 domain**

In order to introduce ARJP related mutations, IBR-RING2 was subjected to site-directed mutagenesis using the MBI Fermentas Hot Start kit. After amplification, the parental DNA templates (which are methylated) were digested by Fast Digest *Dpn*I enzyme (Thermo Scientific). The mutated plasmids were then transformed into *E. coli* JM109 competent cells. After verification by the sequencing, the mutated DNA was

transformed into *E. coli* BL21(DE3) Codon PlusRIL. In total, 18 disease-related mutants were designed, and out of those, 16 mutants were successfully made. The mutants were designed for parkin IBR (G349E, Q355C, T372P, G376D, R383Q), linker (D412N, R415Q, A417T, R420C, R420P) and parkin RING2 (T433N, C436R, G447E, G448D, C449F, C459R, and M476L). Table 1 displays all the primers used in generating the mutants. Also, a RING1-IBR-RING2 construct (residues 252-482) was looped out from the *PARK2* gene in the pGEX vector using restriction-free (RF) cloning (van den Ent and Lowe, 2006), using a similar protocol as outline in section 2.1.1.

## **2.2 Expression of IBR-RING2**

The *Dm* parkin IBR-RING2 was overexpressed in *E. coli* BL21(DE3) Codon Plus RIL *Escherichia coli* strain, with an *N*-terminal GST fusion tag. The cells were streaked on Luria-Bertani (LB)-agar plates containing 100 µg/mL ampicillin (Amp) and incubated for approximately 20 hours at 37°C. Single colonies were picked for overnight cultures.

The starter cultures were prepared using sterile LB broth containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol and inoculated with a single colony from the LB-agar plate. These were agitated at 215 rpm for approximately 20 hours at 37°C. Expression was performed by diluting the starter culture 1:100 into 1L of LB media, and grown at 37°C until an  $A_{600}$  of 0.7-0.8 was reached. The cells were then induced with 1 mM IPTG and continued shaking at 215 rpm at 16°C for 16-17 hours.

**Table 1.** Forward and reverse primer sequences used in the amplification of different domains of parkin's C-terminus and mutants of IBR-RING2 and constructs of split GFP

<b>Cloning_pGEX</b>		
<b>Construct</b>	<b>Direction</b>	<b>Primer Sequences</b>
RING1	Forward	5' ATTCATATGGGATCCAATAACATCAAGAATGTTCC 3'
	Reverse	5' ATACTCGAGAAGCTTGAATTCCTATTACTCCTCGGTGGCGAATCTC3'
RING1-IBR	Forward	5' ATTCATATGGGATCCAATAACATCAAGAATGTTCC 3'
	Reverse	5' ATTCTCGAGGAATCAAGCTTCTATTAAGTAGCGCCCGTCCCCTCGG3'
IBR	Forward	5' ATTCATATGGGATCCACACGCGAGGAGTACGATC3'
	Reverse	5' ATTCTCGAGGAATCAAGCTTCTATTAAGTAGCGCCCGTCCCCTCGG3'
<b>Cloning_splitGFP</b>		
RBR	Forward	5' ATTGGATCCTCCGGCGGGGAGAAGG 3'
	Reverse	5' ATTGGATCCGAGTATGTCCTACAGGCAGG 3'
IBR	Forward	5' ATTGGATCCGAGTATGTCCTACAGGCAGG 3'
	Reverse	5' ATTGGATCCGAGTATGTCCTACAGGCAGG 3'
<b>Site-directed Mutagenesis &amp; RF Cloning_pGEX</b>		
<b>Construct</b>	<b>Direction</b>	<b>Primer Sequences</b>
G349E	Forward	5' GAGTATGTCCTACAGGCAGAAGGCGTACTCTGCCCC 3'
	Reverse	5' GGGGCAGAGTACGCCTTCTGCCTGTAGGACATACTC 3'
Q355C	Forward	5' GCAGGTGGAGTATTGTGCCCTGTCCAGGATGCGGCATG 3'
	Reverse	5' CATGCCGCATCCTGGACAGGGGCACAATACTCCACCTGC 3'
G376D	Forward	5' GTGACATGCCAGAACGATTGTGGATACGTGTTCTGC 3'
	Reverse	5' GTGACATGCCAGAACGATTGTGGATACGTGTTCTGC 3'
R383Q	Forward	5' GGATACGTGTTCTGCCAGAATTGTTTGCAGGGCTACC 3'
	Reverse	5' GGTAGCCCTGCAAACAATTCTGGCAGAACACGTATCC 3'
D412N	Forward	5' GCGAGTACACCGTGAACCCAAATCGAGCTGCC 3'
	Reverse	5' GGCAGCTCGATTTGGGTTACGGTGTACTCGC 3'
R415Q	Forward	5' CCGTGGACCCAAATGGCGCTGCCGAGGCGCG 3'
	Reverse	5' CGCGCCTCGGCAGCGCCATTTGGGTCCACGG 3'
A417T	Forward	5' GACCCAAATCGAGCTACCGAGGCGCGATGGGATGAG 3'
	Reverse	5' CTCATCCCATCGCGCCTCGGTAGCTCGATTGGGTC 3'
R420C	Forward	5' CGAGCTGCCGAGGCGTGCTGGGATGAGGCCAGC 3'
	Reverse	5' GCTGGCCTCATCCCAGCACGCCTCGGCAGCTCG 3'
R420P	Forward	5' CGAGCTGCCGAGGACCGTGGGATGAGGCCAGC 3'
	Reverse	5' GCTGGCCTCATCCCAGGTGCCTCGGCAGCTCG 3'
RBR	Forward	5' CTGTATTTCCAGGGGGCGGCAATAACATCAAGAATGTTCC 3'
	Reverse	5' GGAACATCTTGTATGTTATTGCCGCCCCCTGGAAATACAG 3'

\*ARJP mutations in RING2 region were designed by Noah Manczyk (Summer Student, 2011), and are not listed in the table.

To prepare  $^{15}\text{N}$  or  $^{15}\text{N}$ - $^{13}\text{C}$  labeled protein, M9 minimal media was used supplemented with 1 g/L  $^{15}\text{N}$ -ammonium chloride and 2% glucose (for  $^{13}\text{C}$  sample, 2 g/L of  $^{13}\text{C}_6\text{-D}$ -glucose). The media also contained 2 mM  $\text{MgSO}_4$ , 2  $\mu\text{M}$   $\text{FeSO}_4$ , micronutrients, 100  $\mu\text{g/mL}$  ampicillin and 34  $\mu\text{g/mL}$  chloramphenicol. After a 20-hour expression, cells were harvested by centrifugation (6000 g, 4°C, 15min). Cell pellets were then transferred into 50-mL tubes, flash frozen with liquid nitrogen and stored at -80°C until needed.

### **2.3 Purification of IBR-RING2**

Cells were resuspended in lysis buffer (50 mM Tris, 300 mM NaCl, 1 mM DTT, pH 7.5), lysed using an EmulsiFlex-05 homogenizer (Avestin), and centrifuged at 132380 x g for 1 hour. The clarified supernatant was filtered through a MILLEX HV 0.45  $\mu\text{m}$  filter unit (Millipore) and loaded onto a GStrapFF 5 mL column using an AKTA FPLC system (GE Healthcare) at a flow rate of 0.5 mL/min. After washing the column with 20 column volumes of binding/loading buffer (25 mM Tris, 150 mM NaCl, 1 mM DTT, pH 7.5), the protein was eluted with elution buffer (20 mM Tris, 10 mM reduced glutathione, pH 8.0). Fractions containing the eluted GST-parkin IBR-RING2 protein were pooled and subsequently cleaved by TEV protease to remove the GST tag. The protein was then dialysed against 2L of binding/loading buffer. The protein sample was then loaded onto a 5 mL GStrapHP column (in the same manner as described above) with the pure parkin IBR-RING2 protein being isolated in the flow-through. The concentration of the protein was determined by using a Bradford assay, and the protein was stored at 4 °C until use. All other GST-tagged constructs were expressed and purified in the same manner.



## 2.4 NMR Spectroscopy

All NMR spectra were obtained using Varian Unity INOVA 600 (UWO Biomolecular NMR facility) or 800 MHz NMR (NANUC) spectrometers equipped with cold probe technology. The NMR tubes used were standard 5 mm NMR tubes for 600  $\mu\text{L}$  samples and Shigemi microcell NMR tubes for 300  $\mu\text{L}$  samples. All of the NMR samples were prepared with 9:1  $\text{H}_2\text{O}/\text{D}_2\text{O}$  in 25 mM Tris, 150 mM NaCl 5mM DTT at pH 7.5. Also, 1  $\mu\text{L}$  of complete MINI EDTA free protease inhibitor (ROCHE) and 1  $\mu\text{M}$  of EDTA was added to NMR sample to minimize proteolysis during data acquisition, and to prevent possible oxidation, argon gas was gently blown into the sample before closing the cap. NMR samples contained 30  $\mu\text{M}$  DSS as an internal standard. The concentration of IBR-RING2 samples ranged from 300-400  $\mu\text{M}$ , and all NMR spectra were collected at 25°C.

### 2.4.1 *Chemical Shift Assignment of IBR-RING2*

For the sequential backbone assignment of IBR-RING2, five different 3D experiments were used: CBCA(CO)NH (Grzesiek and Bax, 1992), HNCACB (Wittenkind and Mueller, 1993), HNCA (Kay et al., 1990), HNCO (Kay et al., 1994), and HN(CA)CO (Clubb et al., 1992). All of these experiments were collected on a 600 MHz spectrometer. The number of data points and spectral widths for the  $^1\text{H}$ (F3) and  $^{15}\text{N}$ (F2) dimensions were set to 1024 and 7500 Hz and 32 and 1880 Hz, respectively. For CBCA(CO)NH and HNCACB, 46 increments and spectral width of 8000 Hz in were used in the  $^{13}\text{C}$ (F1) dimension. And for HNCO and HN(CA)CO experiments, 28 increments and 3000 Hz were used. For HNCA, 32 increments and 4521.5 Hz were used.

Non-aromatic side chains were assigned using C(CO)NH (Grzesiek et al. 1993),

HC(CO)NH (Grzesiek and Bax, 1993), HCCH-TOCSY(Bax et al., 1990),  $^{13}\text{C}$ -NOESY (Marion et al., 1989), and  $^1\text{H}$ - $^{13}\text{C}$  HSQC (Kay et al., 1992) experiments. The  $^1\text{H}$ - $^{13}\text{C}$  HSQC experiment was collected with 1280 data points and spectral width of 7500 Hz in the  $^1\text{H}$ (F2) and 128 increments and 12000 Hz in the  $^{13}\text{C}$ (F1) dimension. For C(CO)NH experiment, 1280 complex data points and a spectral width of 7500Hz, 32 increments and 1880 Hz, and 52 increments and 11000 Hz in  $^1\text{H}$ (F3),  $^{15}\text{N}$ (F2) and  $^{13}\text{C}$ (F1) dimensions respectively. For the HC(CO)NH, conditions were similar to CCONH, except for F1 dimension, which was adjusted to 56 data points with 7500 Hz. For the HCCH-TOCSY experiment, the data points and spectral widths used were 1280 and 7500 Hz, 32 and 12000Hz, and 128 and 7500 for the  $^1\text{H}$ (F3),  $^{13}\text{C}$ (F2),  $^1\text{H}$ (F1) dimensions respectively. Aromatic side chains were assigned using (HB)CB(CGCD)HD , (HB)CB(CGCDCE)HE (Yamazaki et al., 1993), aromatic TOCSY, and aromatic  $^{13}\text{C}$ -HSQC(Kay et al., 1992).

All spectra were processed using the program NMRPipe (Delaglio et al., 1995) with cosine-squared function in order to minimize the artifacts in the spectra. Linear prediction was used for data processing by doubling the number of points in the F1 and F2 dimensions for the three-dimensional experiments. Manual chemical shift assignment was accomplished using the program NMRView (Johnson and Belvins, 1994) in order to analyze the three dimensional experiments.

#### **2.4.2 Secondary Structure Prediction**

The chemical shift values for  $^1\text{H}\alpha$ ,  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}'$  resonances were used to predict the secondary structure of the parkin IBR-RING2 protein. The method of chemical shift indexing, developed by Wishart and Sykes (1994), uses the values of these assigned

chemical shifts to accurately predict secondary structures by comparing the values to those typically found in a random coil. The analysis was performed using the CSI option in the NMRViewJ program.

#### **2.4.3 *T<sub>1</sub>/T<sub>2</sub> experiments***

The backbone <sup>15</sup>N T<sub>1</sub> and T<sub>2</sub> relaxation rates of IBR-RING2 were measured with the conventional series of <sup>1</sup>H-<sup>15</sup>N-HSQC experiments with varied relaxation delays (Farrow et al., 1994). Longitudinal relaxation time (T<sub>1</sub>) and transverse relaxation time (T<sub>2</sub>) were determined from these experiments, based on the theory that relaxation times are sensitive to the intensity changes of backbone amides. The rotation correlation time (τ<sub>c</sub>) was then calculated from experimental T<sub>1</sub> and T<sub>2</sub> experiments. Using the relationship between T<sub>1</sub>/T<sub>2</sub> and τ<sub>c</sub>, tumbling time of the IBR-RING2 was plotted on a graph by fitting them using the decay curves and determining T<sub>1</sub> and T<sub>2</sub>.

#### **2.4.4 *Heteronuclear NOE***

Residue-specific heteronuclear <sup>15</sup>N{<sup>1</sup>H} NOE values for IBR-RING2 were measured from two dimensional (<sup>1</sup>H, <sup>15</sup>N) correlated spectroscopy with, and without, proton saturation. The time delays used Heteronuclear NOE values were obtained by taking ratios of the peak intensities in the two aforementioned experiments (Kay et al., 1989; Wang et al., 1999). To ensure accuracy, experiments were conducted in triplicate.

## 2.5 NMR Structure Determination

After manual assignment of chemical shifts of IBR-RING2, distance restraints for the structure calculation were extracted from signal intensities of  $^{13}\text{C}$  and  $^{15}\text{N}$ -edited 3D NOESY-HSQC spectra. Dihedral angle constraints were obtained using the TALOS+ program, based on the chemical shifts of  $^1\text{H}\alpha$ ,  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}'$  (Shen et al., 2009). Structures were calculated using a combination of manual and automatic NOE assignment with CYANA software (Herrmann et al., 2002). Seven cycles of combined automated NOESY assignment and structure calculation were performed followed by a final structure calculation. In this step, the knowledge of short atom-atom distances and the amino acid sequence of the protein as well as the arrangement of all atoms in space were calculated by CYANA. In the final structure calculation 100 conformers were calculated and the 20 conformers with the lowest final target function values (based upon minimal NOE, dihedral and van der Waal contact violations) were selected. The final structures were refined in water using explicit restraints for zinc coordination by Dr. Pascal Mercier.

The atomic coordinates for the structures of IBR-RING2 have been deposited in the Protein Data Bank under the accession code 2M48, and the BMRB accession code is 18990.

## 2.6 Solubility Test

To test the solubility of the split GFP parkin proteins and the mutants of IBR-RING2 that were constructed, small-scale expression and solubility tests were conducted. A 10 mL culture of BL21 (DE3) cells expressing split GFP with RING1-IBR-RING2 and split GFP with IBR-RING2 was harvested, resuspended in 1 mL lysis buffer (50 mM Tris, 300 mM NaCl, 1 mM DTT at pH 7.5), and sonicated for 30 seconds. The cell lysate was

then centrifuged at 16000 x g for 10 minutes. The supernatant contained soluble protein, while the cell pellet contained insoluble protein. The pellet was resuspended in 2% SDS and both the supernatant and pellet were analyzed by SDS-PAGE. A similar approach was taken for IBR-RING2 mutants.

## 2.7 References

- Bax, A., Clore, G.M., and Gronenborn, A.M. (1990).  $^1\text{H}$ - $^1\text{H}$  correlation via isotropic mixing of  $^{13}\text{C}$  magnetization, a new three-dimensional approach for assigning  $^1\text{H}$  and  $^{13}\text{C}$  spectra of  $^{13}\text{C}$ -enriched proteins. *Journal of magnetic resonance* 88, 425-431.
- Clubb, R.T., Thanabal, V., and Wagner, G. (1992). A new 3D HN(CA)HA experiment for obtaining fingerprint HN-Halpa peaks in  $^{15}\text{N}$ - and  $^{13}\text{C}$ -labeled proteins. *J Biomol NMR* 2, 203-210.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6, 277-293.
- Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D., and Kay, L.E. (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by  $^{15}\text{N}$  NMR relaxation. *Biochemistry* 33, 5984-6003.
- Ghosh, I., Hamilton, A., and Regan, L. (2000). Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. *J Am Chem Soc* 122, 5658-5659.
- Grzesiek, S., and Bax, A. (1992). Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc* 114, 6291-6293.
- Grzesiek, S., and Bax, A. (1993). Measurement of amide proton exchange rates and NOEs with water in  $^{13}\text{C}/^{15}\text{N}$ -enriched calcineurin B. *J Biomol NMR* 3, 627-638.
- Herrmann, T., Guntert, P., and Wuthrich, K. (2002). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* 319, 209-227.
- Hristova, V.A., Beasley, S.A., Rylett, R.J., and Shaw, G.S. (2009). Identification of a novel  $\text{Zn}^{2+}$ -binding domain in the autosomal recessive juvenile Parkinson-related E3 ligase parkin. *J Biol Chem* 284, 14978-14986.
- Johnson, B.A., and Belvins, R.A. (1994). NMRView: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* 4, 603-614.
- Kay, L.E., Ikura, M., Tschudin, R., and Bax, A. (1990). Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance* (1969) 89, 496-514.

- Kay, L.E., Keifer, P., and Saarinen, T. (1992). Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc* *114*, 10663–10665.
- Kay, L.E., Torchia, D.A., and Bax, A. (1989). Backbone dynamics of proteins as studied by <sup>15</sup>N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* *28*, 8972-8979.
- Kay, L.E., Xu, G.Y., and Yamazaki, T. (1994). Enhanced-sensitivity triple-resonance spectroscopy with minimal H<sub>2</sub>O saturation. *J Mag Res* *109*, 129–133.
- Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M., and Clore, G.M. (1989). Overcoming the overlap problem in the assignment of proton NMR spectra of larger proteins by use of three-dimensional heteronuclear proton-nitrogen-15 Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1. beta. . *Biochemistry* *28*, 6150-6156.
- Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS+: A hybrid method for predicting backbone torsion angles from NMR chemical shifts. *J Biomol NMR* *44*, 213-223.
- van den Ent, F., and Lowe, J. (2006). RF cloning: a restriction-free method for inserting target genes into plasmids. *J Biochem Biophys Methods* *67*, 67-74.
- Wang, G., Wylie, G.P., Twigg, P.D., Caspar, D.L.D., Murphy, J.R., and Logan, T.M. (1999). Solution structure and peptide binding studies of the C-terminal Src homology 3-like domain of the diphtheria toxin repressor protein. *Proc Natl Acad Sci* *96*, 6119–6124.
- Wishart, D.S., and Sykes, B.D. (1994). The <sup>13</sup>C chemical-shift index: a simple method for the identification of protein secondary structure using <sup>13</sup>C chemical-shift data. *J Biomol NMR* *4*, 171-180.
- Wittenkind, M., and Mueller, L. (1993). HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J Magn Reson B* *101*, 201-205.
- Yamazaki, T., Forman-Kay, J.D., and Kay, L.E. (1993). Two-dimensional NMR experiments for correlating <sup>13</sup>C.beta. and <sup>1</sup>H.delta/epsilon chemical shifts of aromatic residues in <sup>13</sup>C-labeled proteins via scalar couplings. *J Am Chem Soc* *115*, 11054-11055.

## Chapter 3

### Results and Discussion

Parkin is an RBR E3 ligase involved in the degradation of proteins that are misfolded through the UPS. The most important portion of parkin is in the C-terminus that comprises the RING0, RING1, IBR, and RING2 domains. The RING2 domain is thought to be the catalytic domain/segment responsible for acquiring a ubiquitin from an E2 enzyme and transferring it to a substrate (Spratt et al., 2013). The RING1 domain is thought to be involved in the interaction with an E2 enzyme. Understanding how this ubiquitin transfer occurs is difficult because three-dimensional structures of the functional RING domains, and more importantly the assembly of the entire C-terminal RBR motif, are not known. In this chapter the structure of the IBR-RING2 region of parkin was determined.

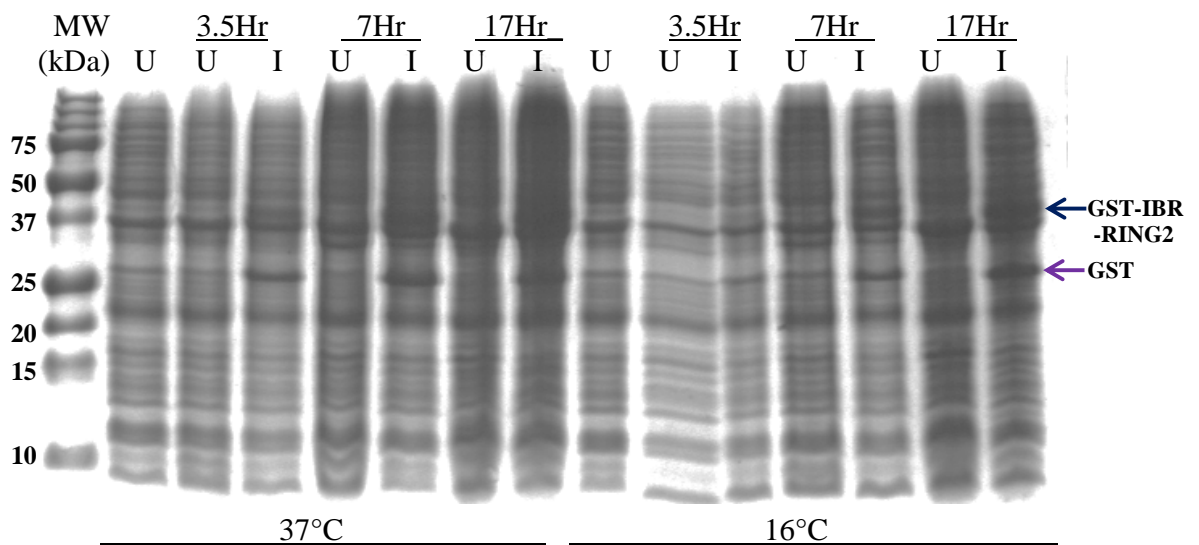
#### 3.1 GST-IBR-RING2 Expression

Prior to the purification of GST-IBR-RING2, expression tests were performed to determine the optimal expression time and temperature for the protein. Figure 3.1 shows an SDS-PAGE gel of an expression test of GST-IBR-RING2 in *E. coli* strain BL21 DE3 CodonPlus-RIL, at 37°C and 16°C. Protein expression was induced for 17 hours, after addition of 1 mM IPTG. As a negative control, comparison was made against an uninduced expression culture.

GST-IBR-RING2 has a molecular weight of approximately 42 kDa and therefore, should display as a band on the SDS-PAGE gel between the 37 and 50 kDa molecular weight markers. Figure 3.1 illustrates the GST-IBR-RING2 band appearing after 3.5



hours of expression at both temperatures. This band shows an increase in intensity after 7 and 17 hours at both temperatures, indicating proper over-expression of the GST-fusion protein. However, another band also shows an increase in intensity near 25kDa and is likely GST protein on its own. The appearance of free GST could result from fusion protein cleavage due to protease activities. This has been observed previously for bacterial expression of full-length parkin (Hristova et al., 2009). After 17 hours of expression, the sample expressed at 37°C shows increased levels of GST protein compared to the GST-fusion protein, while at 16°C, the GST-IBR-RING2 expression is more efficient (compared to all other molecular weight proteins, as well as GST on its own). Thus, the optimal expression conditions for IBR-RING2 were chosen to be 16°C for 17 hours.



**Figure 3.1: SDS-PAGE gel of induced vs uninduced GST-IBR-RING2.**

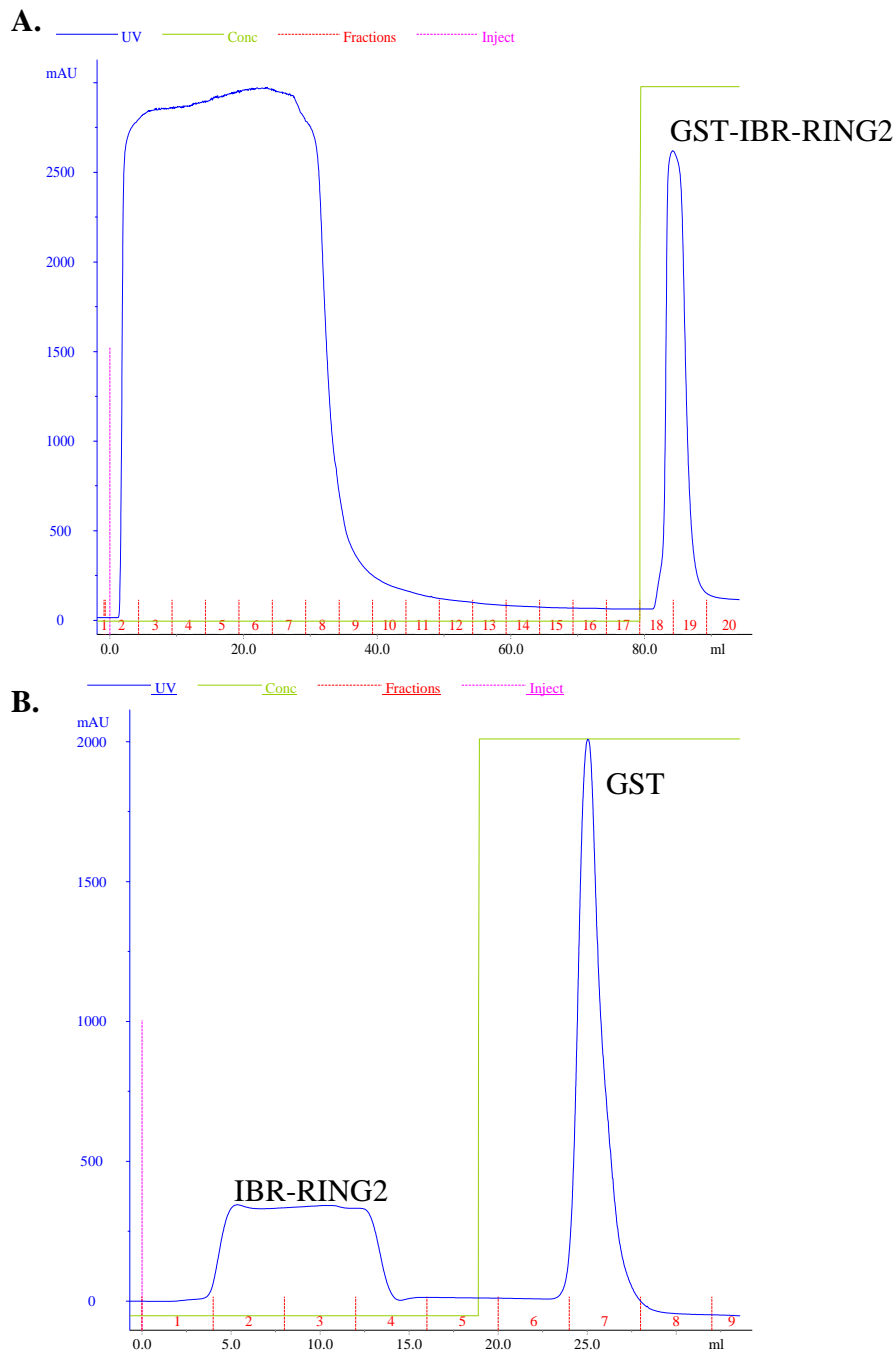
MW refers to the molecular weight ladder. U represents an uninduced sample, and I represents an induced sample; samples are observed at two different temperatures (37°C and 16°C) for expression, and at different time-points (3.5, 7, and 17 hours).

### 3.2 Purification of IBR-RING2

After the cell lysate was centrifuged, GST affinity chromatography was used to purify GST-IBR-RING2. The GST fusion protein was then cleaved with TEV protease and dialyzed extensively to remove excess glutathione. Thorough dialysis was necessary to ensure all glutathione was removed so that the cleaved GST tag could be separated from the IBR-RING2. A second GST purification was subsequently performed and the IBR-RING2 protein eluted in the flow-through fractions while the GST tag remained bound to the column. Figure 3.2 shows typical chromatograms of IBR-RING2 protein purification prior to and following cleavage of the GST tag from the protein. The purity of the protein throughout the process was checked by SDS-PAGE. Figure 3.3 shows segments of SDS-PAGE gels of the complete purification scheme of IBR-RING2. As stated in Chapter 2, all proteins that are GST-tagged were expressed and purified in the same manner. At the end of the purification, the amount of protein in solution was approximately 4-5 mg (from 2L of growth).

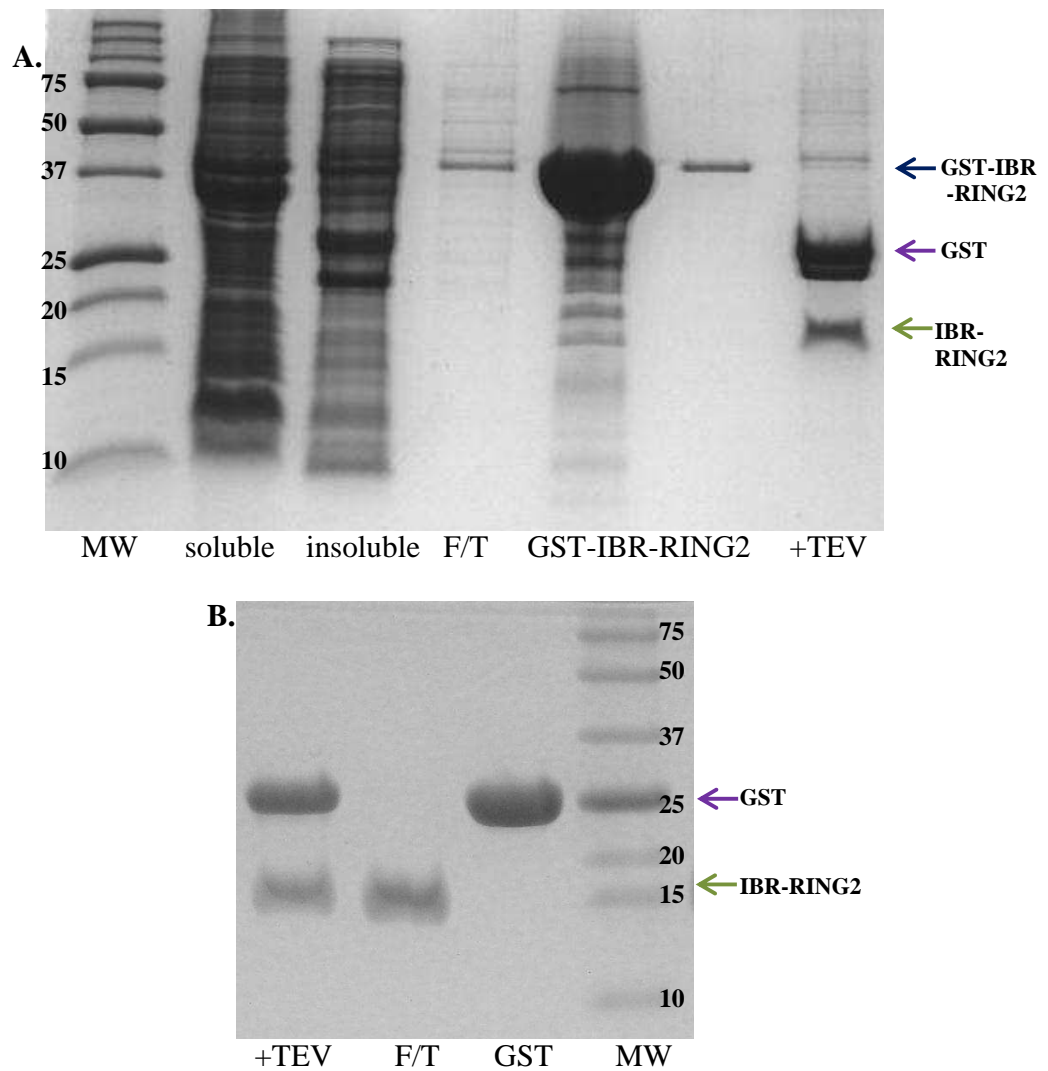
### 3.2 Determination of Optimal Conditions for NMR Data Collection

Prior to collecting numerous 3D NMR experiments for structure determination of IBR-RING2, it was important to determine the optimal conditions for these experiments such as temperature and pH. A series of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of IBR-RING2 were collected under various conditions and it was concluded that 25°C and pH 7.5 were optimal conditions for IBR-RING2 data acquisition. It was also important to ensure that protein samples maintain their integrity during the data collection for NMR experiments. A major concern regarding IBR-RING2 was the oxidation of cysteine residues and degradation of the IBR-RING2 while acquiring NMR spectra. It would be impossible to



**Figure 3.2: Chromatograms of GST purification**

**A.** Chromatogram of a typical GST affinity purification. The absorbance at 280 nm is shown in blue. GST-IBR-RING2 is eluted from the column by the addition of 100% elution buffer (shown in green, 20mM Tris, 10mM GSH, at pH 8.0). **B.** Chromatogram of a second GST purification after TEV cleavage. Sharp and intense elution peak is an indication that thorough dialysis of glutathione was achieved, allowing separation of GST and IBR-RING2.



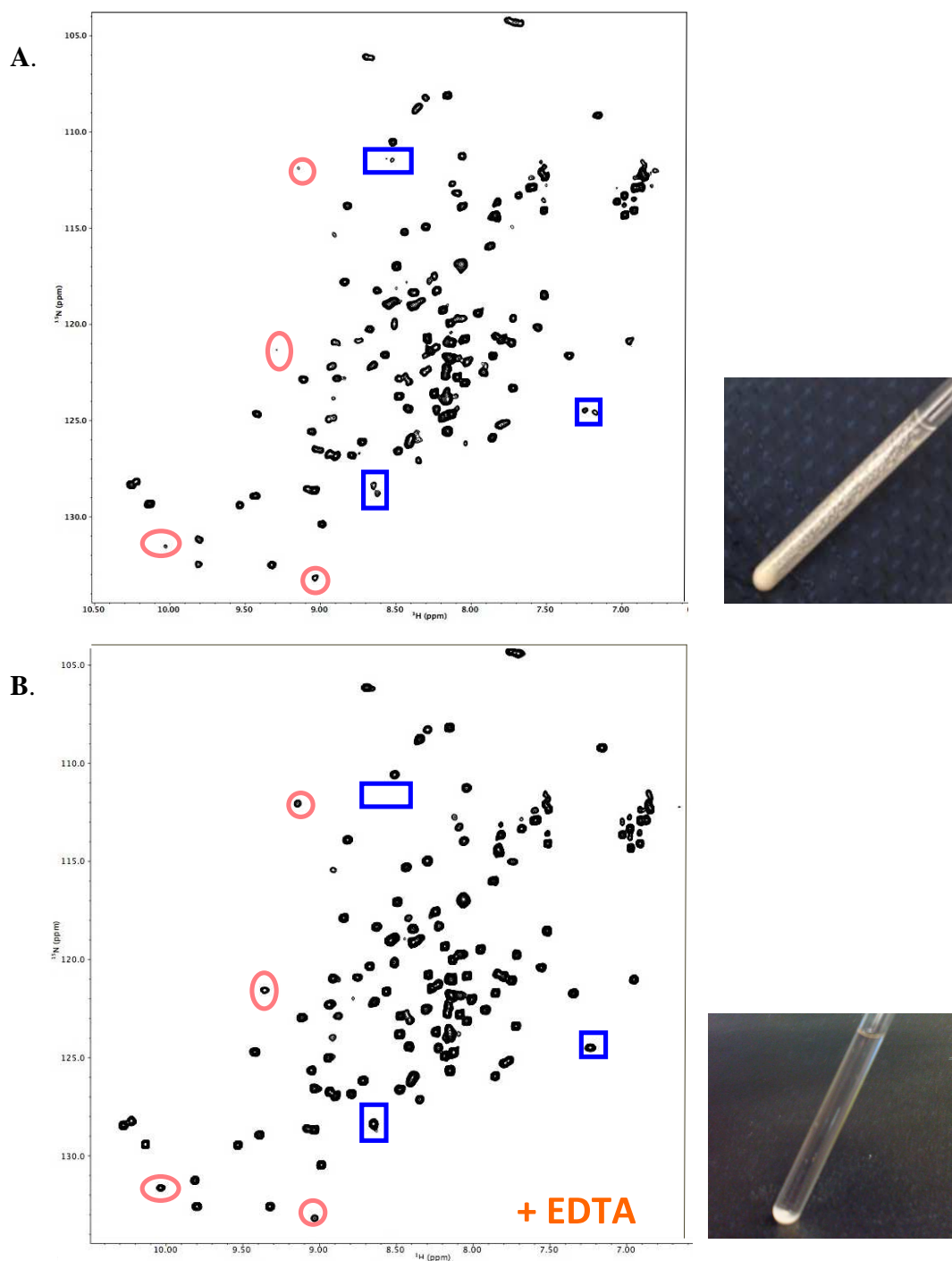
**Figure 3.3: SDS-PAGE gels showing two-step purification of IBR-RING2**

**A.** First, GST purification and TEV cleavage reaction. **B.** TEV-cleaved protein through the GSTrap HP column. Abbreviations used for each lanes are, MW: molecular weight marker, soluble: soluble fraction, insoluble: insoluble fraction, F/T: flow-through, and +TEV: TEV cleaved GST-IBR-RING2.

collect accurate data with degraded or precipitated protein in NMR studies. Therefore, it was important to find out what conditions would ensure that the IBR-RING2 was homogeneous, intact, well folded, and soluble.

As mentioned above, a major concern was the possible oxidation of 17 cysteine residues found in IBR-RING2. In order to avoid oxidation and/or disulphide bond formation, the sample was purged with Argon gas to displace oxygen in the NMR tube. In addition, Shigemi NMR tubes were used which limit the amount of oxygen exposure experienced by the sample due to the insertion of a plunger. Finally, addition of 5 mM DTT to the sample was used to prevent cysteine oxidation.

Despite optimization of sample conditions it became apparent that the IBR-RING2 protein readily precipitated during long NMR data collection times (more than a week). Many of the individual 3D experiments require several days and ideally the sample should be stable for a minimum of two weeks so that most of the required 2D/3D spectra can be collected on the same sample. Figure 3.4 shows a solution of IBR-RING2 in an NMR tube after one week of data collection. The NMR tube became cloudy with precipitated IBR-RING2 within one week. This was a major concern and was thought to be caused by an unknown protease contaminant which resulted in protein unfolding. The contaminant was not inhibited by the protease inhibitor cocktail that was initially added to the NMR sample. These initial experiments used a protease inhibitor cocktail that did not contain EDTA (Ethylenediaminetetraacetic acid) because there was a concern that EDTA might remove zinc ions from the IBR-RING2. It has been previously shown that



**Figure 3.4: Comparison of  $^1\text{H}$ - $^{15}\text{N}$  HSQC Spectra and NMR tubes of 300-400 $\mu\text{M}$  IBR-RING2 after the data collection.**

**A.** For non-optimized NMR sample after ten days (Disappearing peaks: circled, Newly appearing peaks: boxed) **B.** For optimized NMR sample (EDTA addition) after a month (Superimposable to the initial spectrum)

the IBR domain from human parkin unfolds upon the addition of EDTA (Beasley et al., 2007). However, there was still a possibility of the presence of metalloproteases in the sample (which cannot be blocked by any of the inhibitors in the protease inhibitor cocktail) that affect the integrity of IBR-RING2 and were causing it to precipitate. To test this, a small amount of EDTA (1  $\mu\text{M}$ ) was added to the protein sample in an attempt to inhibit the potential metalloproteases, and the stability of the IBR-RING2 was re-assessed. As seen in Figure 3.4, a minor amount of protein precipitated at the bottom, but generally the IBR-RING2 remained soluble in the NMR tube even after a month of experiments. Further,  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments compared before and after this period were almost identical, suggesting that the sample's integrity was maintained. Therefore, it was concluded that the addition of trace amounts of EDTA helped the IBR-RING2 to stay in solution for a longer period of time by inhibiting unknown metalloproteases that might have been present in the NMR sample. Based on this finding, all subsequent NMR samples contained 1  $\mu\text{M}$  EDTA.

### **3.3 Structural Determination of the Parkin IBR-RING2 Domain**

#### **3.3.1 *ESI-MS shows the IBR-RING2 coordinates 4 Zinc ions***

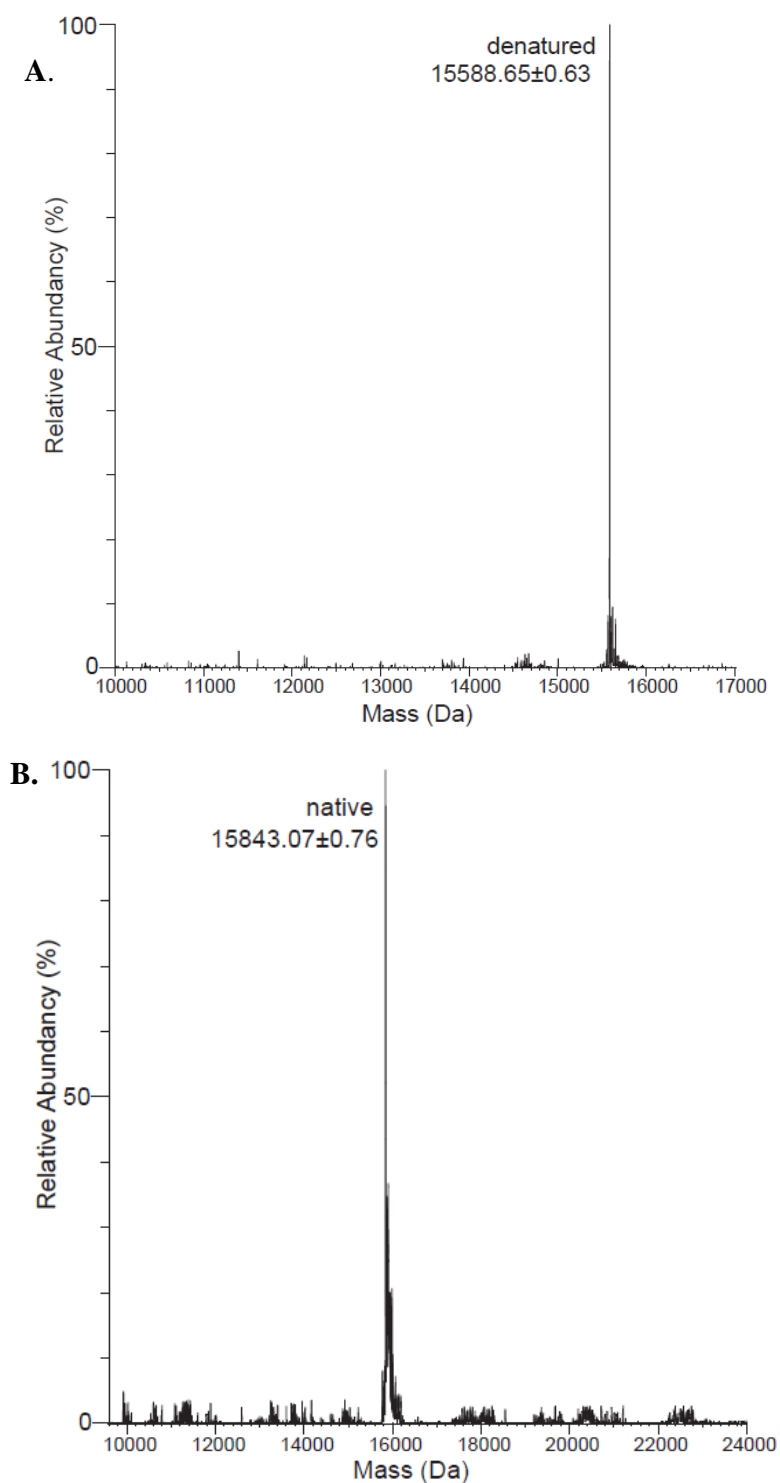
Mass spectrometry is a very powerful technique used in determining the molecular weight of proteins. When comparing two conditions of ESI-MS (non-denaturing and denaturing), it has been demonstrated that full-length parkin binds eight zinc ions (Hristova et al., 2009). The same technique was employed for the IBR-RING2 and showed there are in a total of four zinc ions bound to the protein. The observed



difference between the denaturing and non-denaturing mass of the IBR-RING2, as determined by Paula Pittock (Biological Mass Spectrometry Lab, UWO), was calculated. The native IBR-RING2 had a mass of 15843 Da and the denatured IBR-RING2 was found to have a mass of 15588 Da (Figure 3.5). The difference in mass was 255 Da and division of this value by four resulted in 63.75 Da which is close to the mass of zinc (65 Da), which led to the conclusion that there are four zinc ions coordinated in IBR-RING2.

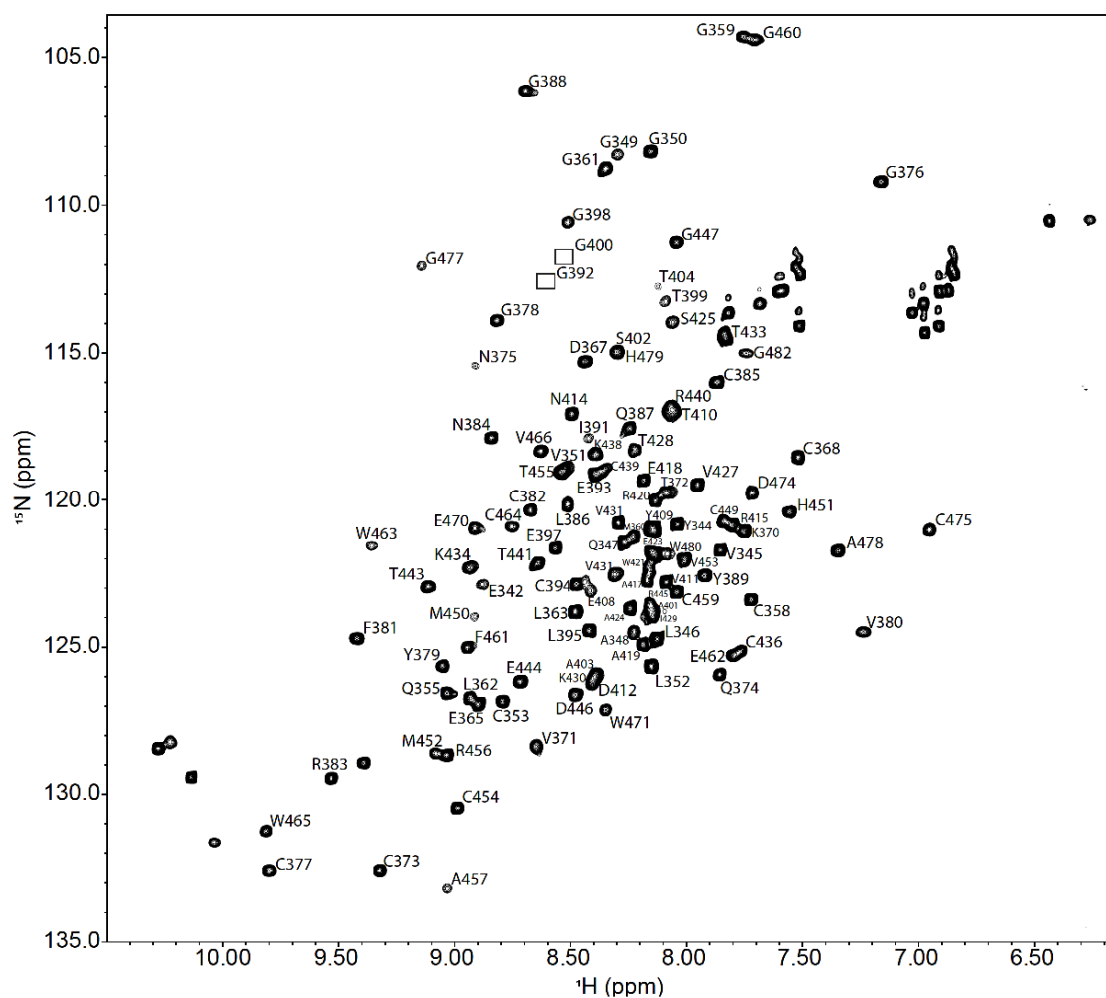
### **3.3.2 Backbone Chemical Shift Assignment of IBR-RING2**

For structural studies with NMR spectroscopy, each resonance needs to be associated with a specific nucleus in the protein of interest. The backbone chemical shift assignment of IBR-RING2 was accomplished manually using the sequential backbone resonance assignment method. A series of multidimensional NMR experiments (HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO) were used in order to assign the HN, N, C $\alpha$  and C $\beta$  of IBR-RING2, as discussed in 2.4.1. A sample of the sequential backbone assignment is shown in Appendix A, Figure A-1. The assigned  $^1\text{H}$ - $^{15}\text{N}$  HSQC (Figure 3.6) is well dispersed confirming the protein was well folded. Approximately 92% of the backbone assignment was completed (103 out of 112 residues). Amide resonances such as N405, S406, and M476 could not be identified from the data because the signal was too weak or overlapping with other peaks.



**Figure 3.5: Deconvoluted ESI-MS of proteins under denaturing and non-denaturing conditions, which show the mass of both the native and denatured state of IBR-RING2.**

**A.** Denaturing IBR-RING2 **B.** Non-denaturing IBR-RING2 result.



**Figure 3.6: Two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of  $^{15}\text{N}$ -isotopically labeled IBR-RING2.**

This data was collected at 25°C at 600 MHz. Backbone amide resonances are labeled, with their one letter amino acid code and sequence number.

### 3.3.3 *Utilization of IBR-RING2 backbone assignment*

Assigning chemical shifts of the amide backbone is the first step in solving the structure of proteins. Additionally, the backbone assignment can serve several other purposes, including: conducting interaction studies by titration, the determination of residues involved in zinc coordination, and domain secondary structure. In this section, chemical shifts were used in the prediction of IBR-RING2 secondary structure, as well as zinc-coordinating cysteines.

### 3.3.4 *Prediction of Zinc Coordinating Cysteines*

With the backbone assignment completed, identification of cysteines exhibiting zinc coordination became possible. Using the sensitivity of chemical shifts to changes in electronic environment, prediction of zinc ligating cysteines and non-ligating cysteine residues can be made. Kornhaber and colleagues (2006) compiled the differences in values of  $C\alpha$  and  $C\beta$  of cysteines when oxidized ( $C\alpha: 55.57 \pm 2.46$  ppm,  $C\beta: 41.17 \pm 3.93$  ppm), reduced ( $C\alpha: 59.25 \pm 3.06$  ppm,  $C\beta: 28.92 \pm 2.11$  ppm), and or zinc-coordinating ( $C\alpha: 59.27 \pm 2.12$  ppm,  $C\beta: 30.89 \pm 1.01$  ppm) from the protein data bank (PDB) and BioMagResBank (BMRB). With this statistical data, the probability of the given cysteine's zinc coordination could be calculated. Through the assignment of  $C\alpha$  and  $C\beta$  chemical shifts of all cysteines in IBR-RING2 (total of 17), the prediction of cysteines involved in zinc coordination was completed (Kornhaber et al., 2006). In Table 2, probabilities for the oxidized, reduced and zinc coordinating states of all cysteines in IBR-RING2 are shown. This method showed there was over 90% probability that residues C353, C377, C382, C385, C394, C439, C454, C459, C464, and C475 were

**Table 2.** Cysteines involved in Zn<sup>2+</sup> coordination in IBR-RING2 domain, as determined by chemical shifts of C $\alpha$  and C $\beta$ .

<b>Cys Chemical Shifts (ppm)*</b>									
<b>IBR</b>									
	<b>C353</b>	<b>C358</b>	<b>C368</b>	<b>C373</b>	<b>C377</b>	<b>C382</b>	<b>C385</b>	<b>C394</b>	<b>C407</b>
Actual <sup>#</sup>	I	I	-	I	I	II	II	II	-
C $\alpha$	57.731	60.268	58.581	58.504	59.988	57.691	58.112	59.835	58.713
C $\beta$	31.865	29.375	28.45	28.311	32.365	32.475	32.359	30.761	28.102
<b>Probabilities</b>									
Oxidized	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Reduced	7.3%	39.8%	<b>98.3%</b>	<b>99.1%</b>	1.2%	7.0%	5.0%	4.1%	<b>99.5%</b>
Zn <sup>2+</sup>	<b>92.7%</b>	<b>60.2%</b>	1.7%	0.9%	<b>98.8%</b>	<b>93.0%</b>	<b>95.0%</b>	<b>95.9%</b>	0.5%

<b>Cys Chemical Shifts (ppm)*</b>								
<b>RING2</b>								
	<b>C436</b>	<b>C439</b>	<b>C449</b>	<b>C454</b>	<b>C459</b>	<b>C464</b>	<b>C467</b>	<b>C475</b>
Actual <sup>#</sup>	I	I	cat	I	I	II	II	II
C $\alpha$	57.016	59.348	59.291	61.571	60.499	57.776	56.999	64.533
C $\beta$	31.698	31.196	28	32.045	30.174	32.93	31.87	29.092
<b>Probabilities</b>								
Oxidized	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Reduced	13.2%	3.5%	<b>99.5%</b>	0.3%	6.8%	7.8%	12.3%	4.9%
Zn <sup>2+</sup>	<b>86.8%</b>	<b>96.5%</b>	0.5%	<b>99.7%</b>	<b>93.2%</b>	<b>92.2%</b>	<b>87.7%</b>	<b>95.1%</b>

\* determined using 3D CBCA(CO)NH and HNCACB experiments

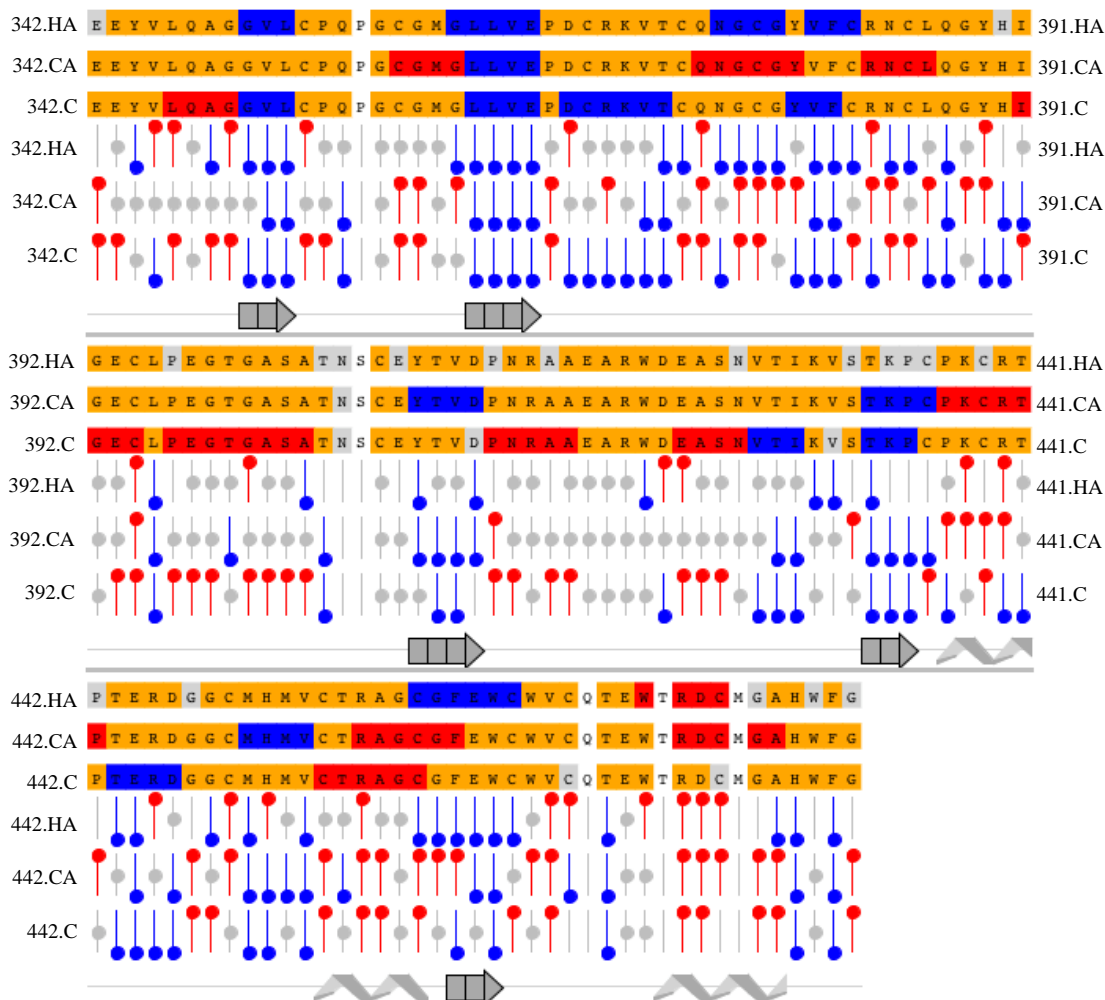
# based on the structure calculation: (I) - involved in 1<sup>st</sup> zinc coordination site, (II) - in 2<sup>nd</sup> zinc coordination site, (cat) – suspected catalytic cysteine, (-) - non zinc coordinating cysteine

coordinated to zinc ions. In addition, there were those other residues (C358, C436, C467) that had over 60% chance of zinc coordination. The data also clearly showed that several cysteine residues (C368, C373, C407, C468) exist in their reduced states based on C $\alpha$  and C $\beta$  chemical shifts (Table 2). In total, this data showed that 13 cysteines in IBR-RING2 are involved in zinc coordination. Zinc coordination in a protein structure typically requires four ligands, and it was predicted that there are more than three zinc ions that can bind to the IBR-RING2 structures. Since histidines can also act as zinc binding ligands, the data is in good agreement with the coordination of four zinc ions by IBR-RING2 shown from the ESI-MS results.

### 3.3.5 *Determination of Secondary Structure for parkin IBR-RING2*

After completing the backbone and side chain chemical shift assignment of IBR-RING2, the secondary structure determination was performed using the chemical shift index (Wishart and Sykes, 1994). This was performed using a macro in the program NMRViewJ that uses chemical shifts of H $\alpha$ , C $\alpha$ , and C for each residue of IBR-RING2 to predict the secondary structure of the protein, as described in 2.4.2. In general, large regions of  $\alpha$ -helices or  $\beta$ -strands were not present in the structure especially in the linker region between IBR and RING2 (Figure 3.7). The predicted secondary structure shows two  $\beta$ -strands in the IBR domain, between G350-L352 and L362-E365, a  $\beta$ -strands in the linker region Y409-D412, and three  $\alpha$ -helices, P437-T441, C454-W471, and T473-A478, and two  $\beta$ -sheets, T433-P435 and F462-W464 in the RING2 region are present.

According to the prediction, RING2 appeared to be more structured than the IBR domain. The data shows that there is no extensive secondary structure in either IBR or



**Figure 3.7: Chemical shift index of IBR-RING2 for secondary structure prediction.**

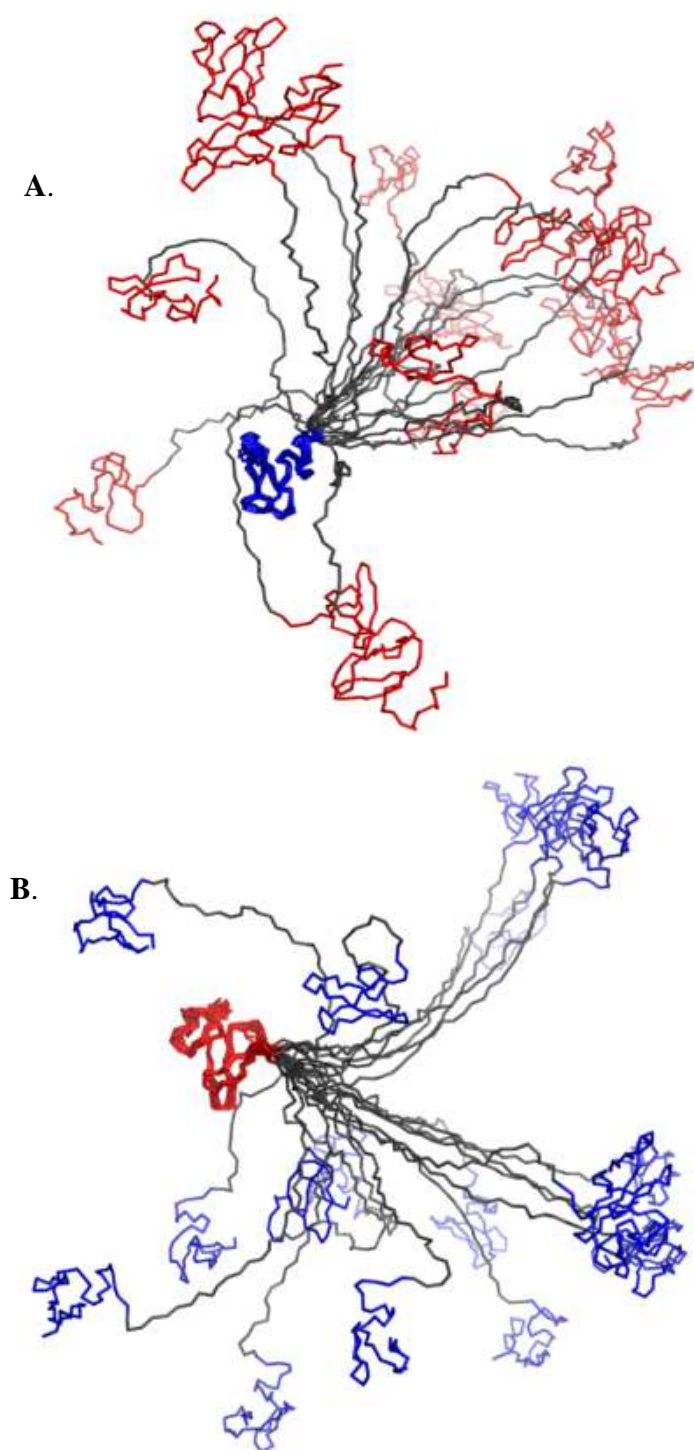
Chemical shift values for  $^1\text{H}\alpha$ ,  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}'$  for each of the residues were compared to those of random coil and given the values -1, 0, and +1. When the value is -1, it is indicative of an  $\alpha$ -helix, and shown in color red, +1 for  $\beta$ -sheet (blue) and 0 for coils (grey). Below the color representation is a secondary structure schematic diagram for IBR-RING2. IBR region contains residues 342-402, and RING2 extends between residues 417-482.

RING2 domains. This is in agreement with the findings for the human parkin IBR structure, that showed little secondary structure (Beasley et al., 2007).

### 3.3.6 *Structure Calculation of IBR-RING2*

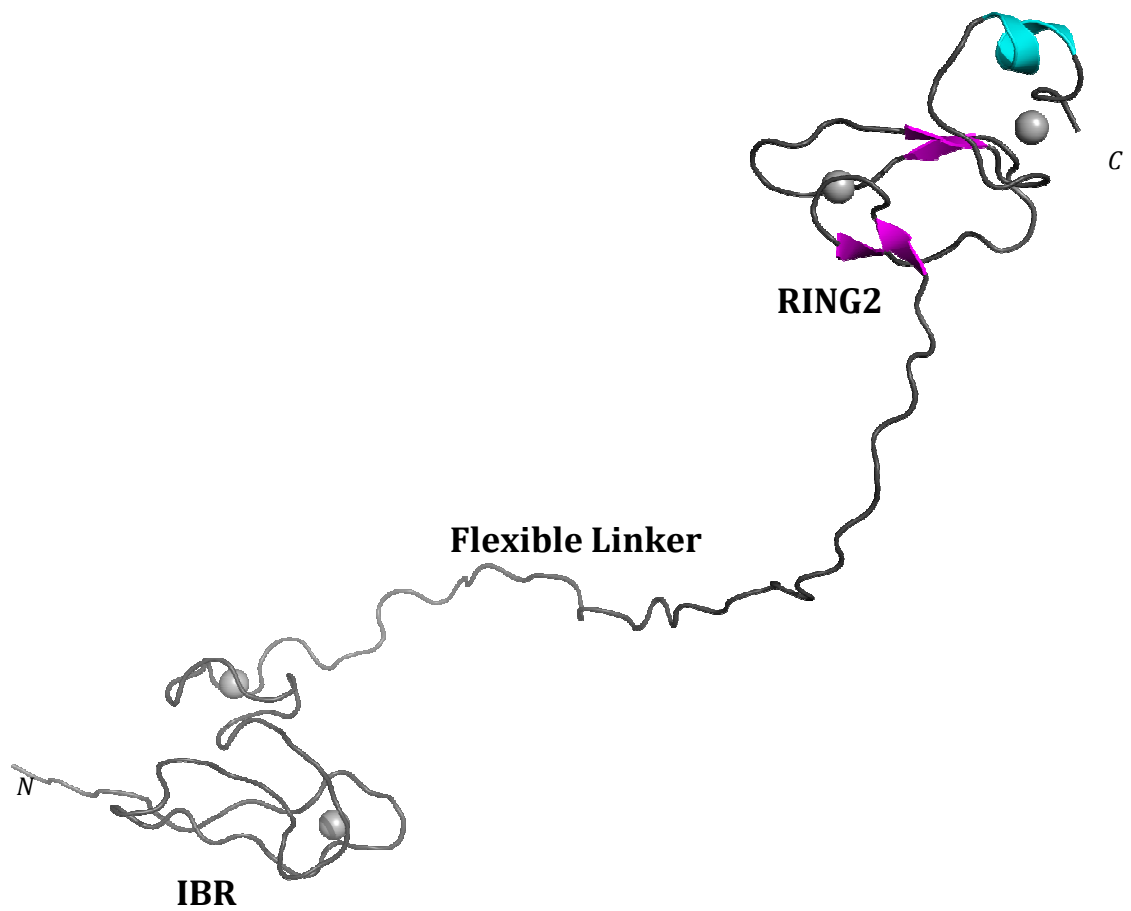
A total of 1156 distance restraints were derived from  $^{15}\text{N}$ -edited,  $^{13}\text{C}$ -edited aliphatic, and  $^{13}\text{C}$ -edited aromatic NOESY spectra. In addition, 49 dihedral restraints were determined from TALOS+ (Shen et al., 2009). These restraints were used for the structure calculation of the IBR-RING2. Figure 3.8 shows a superposition of an ensemble of 20 structures of IBR-RING2, when superimposed using only the IBR (red) or RING2 (blue) domains. These structures represent the lowest energy functions with the least distance, angle, and van der Waal contact violations. This family of structures shows that a flexible linker connects the IBR and RING2 domains. The RMSD of the structured regions for the backbone are  $0.80 \pm 0.22 \text{ \AA}$  (IBR) and  $0.91 \pm 0.20 \text{ \AA}$  (RING2), and those for all heavy atoms are  $1.29 \pm 0.31 \text{ \AA}$  (IBR) and  $1.52 \pm 0.40 \text{ \AA}$  (RING2). A 26-residue flexible linker exists between IBR and RING2, and a ribbon diagram of IBR-RING2 in Figure 3.9 makes this evident. During structure calculations, it became apparent this connecting region contained little regular structure based on the lack of long range inter-residue NOE contacts between either of the domains. Short and sequential short range NOEs also indicated that the linker is flexible. In addition, no long range NOEs between IBR and RING2 were observed, suggesting that they must be isolated from each other in the global structure. The structure of IBR-RING2 shows that the RING2 domain can have multiple orientations with respect to IBR in the 20 structure overlay. Four well-defined  $\text{Zn}^{2+}$ -binding sites are present in IBR-RING2, two in each site.





**Figure 3.8: Superposition of 20 structures overlay of IBR-RING2, on each of the IBR (blue) and RING2 (red) domains.**

Due to the flexible nature of the linker, both domains do not superposition well at the same time, as shown above. **A.** Superposition of IBR-RING2 to IBR using C $\alpha$  atoms. **B.** Superposition of IBR-RING2 to RING2 using C $\alpha$  atoms.



**Figure 3.9: Representative ribbon structure of IBR-RING2 with the flexible linker between the two domains.**

Spheres represent zinc ions bound to each of the domains. Magenta represents the  $\beta$ -sheets and cyan represents the  $\alpha$ -helix in the RING2 structure.

**Table 3:** Structural Statistics for 20 lowest NOE energy structures of IBR-RING2

	<b>Protein</b>
<b>Completeness of Resonance Assignments</b>	
Backbone (N, CA)	(103/112) – 92.0%
Sidechain (C, H)	(1073/1219) – 82.8%
HN	(122/133) – 91.7%
HA	(151/158) – 95.4%
HB	(202/202) – 100%
<b>NMR distance and dihedral constraints</b>	
Distance constraints <sup>1</sup>	
Total NOE	611 / 447 / 1156
Intra-residue	156 / 143 / 360
Inter-residue	
Sequential ( $ i - j  = 1$ )	201 / 117 / 353
Medium-range ( $ i - j  < 4$ )	65 / 48 / 115
Long-range ( $ i - j  > 5$ )	189 / 139 / 328
Intermolecular	- / - / 0
Hydrogen bonds	0 / 0 / 0
Zinc restraints	24 / 24 / 48
Total dihedral angle restraints <sup>2</sup>	
$\phi$	22 / 25 / 49
$\psi$	22 / 25 / 49
<b>Structure statistics</b>	
Violations (mean and s.d.)	
Distance constraints (Å)	0.040 ± 0.002
Dihedral angle constraints (°)	0.467 ± 0.080
Max. dihedral angle violation (°)	7.5
Max. distance constraint violation (Å) <sup>3</sup>	0.7
Deviations from idealized geometry <sup>4</sup>	
Bond lengths (Å)	0.006 ± 0.000
Bond angles (°)	0.550 ± 0.022
Impropers (°)	0.758 ± 0.056
Average pairwise r.m.s. deviation (Å) <sup>5</sup>	
Heavy	1.288 ± 0.313 / 1.519 ± 0.399
Backbone	0.800 ± 0.220 / 0.914 ± 0.216
<b>Ramachandran statistics<sup>6</sup></b>	
Residues in most favored regions	70.5% / 74.0% / 73.5%
Residues in additional allowed regions	25.3% / 24.9% / 24.3%
Residues in generously allowed regions	2.7% / 0.9% / 1.5%
Residues in disallowed regions	1.5% / 0.2% / 0.6%

<sup>1</sup> For IBR residues 342-396 / RING2 residues 430-482 / Entire sequence E342-G482

<sup>2</sup> Psi/Psi dihedral restraints determined using TALOS+

<sup>3</sup> Total of 4 NOE violations > 0.5 Å over all 20 models

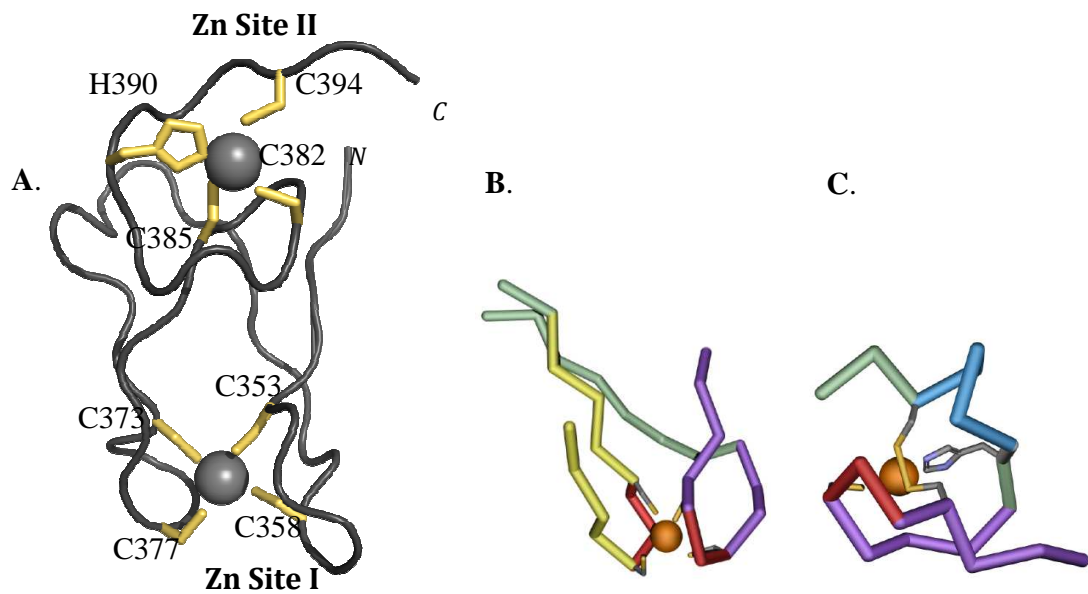
<sup>4</sup> As reported by Xplor-NIH

<sup>5</sup> Using residues V351-I392 for IBR / residues K430-W480 for RING2 (all 20 models)

<sup>6</sup> As reported by Procheck using residues V351-I392 for IBR / residues K430-W480 for RING2 / the entire sequence E342-G482.

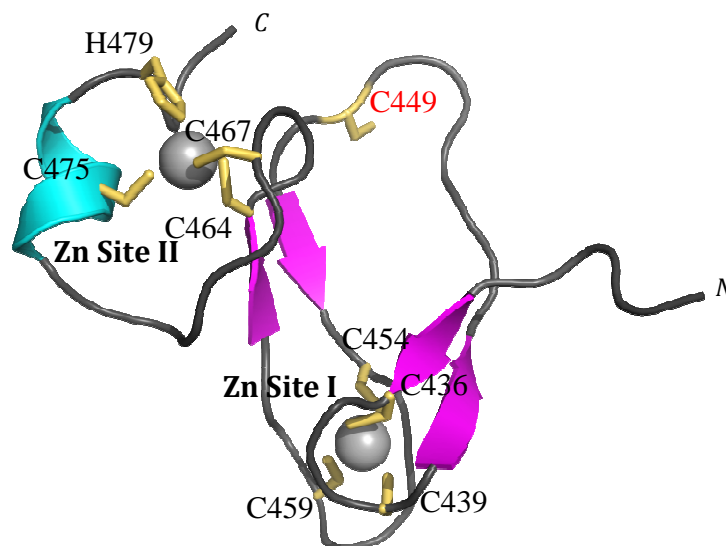
Individual IBR and RING2 domains within the IBR-RING2 structure are shown in Figures 3.10 and 3.11. As shown previously in the human parkin IBR structure (Beasley et al., 2007), the IBR did not display any notable secondary structure. The *N*-terminus at the beginning of the IBR is not as structured, and can be attributed to it being the end part of the linker region between RING1 and IBR. The IBR domain shows a bilobal fold with two zinc ions bound to the protein. It contains two zinc sites in “scissor-like” and “Gag-knuckle-like” folds, as described by Beasley et al. (2007). The long ends of the loop for the first zinc site (utilizing C353, C358, C373, and C377) are the origin of the ‘scissor-like’ fold also referred to as zinc ribbon (Krishna et al., 2003). The IBR second zinc binding site (using C382, C385, H390, and C394) is referred to as the zinc knuckle-like fold, based on the compact zinc binding site which occurs due to the sudden turn (zinc knuckle) required to wrap around the zinc ion.

The structure of fly RING2 in IBR-RING2 is reminiscent of the human IBR structure. The RING2 of IBR-RING2 coordinates zinc just like IBR, which coordinates zinc ions sequentially; 1st, 2nd, 3rd, and 4th zinc coordinating residues comprise the first zinc site (C436, C439, C454, C459), while the second zinc coordinating site is composed of 5th, 6th, 7th and 8th zinc coordinating residues (C464, C467, C475, H479). In comparison, the canonical RING structure (like TRAF6 or BRCA1) coordinates zinc through the cross-brace coordination motif that facilitates 1st, 2nd, 5th and 6th zinc coordinating residues for the first zinc site, while the second zinc coordinating site is composed of 3rd, 4th, 7th and 8th zinc coordinating residues. This results in a more compact structure, since it allows more contacts to be established within the structure.



**Figure 3.10: Ribbon structure of parkin IBR domain with template of zinc binding sites.**

A. IBR domain with zinc coordinating cysteines are highlighted in yellow, showing limited secondary structure. B. Zinc ribbon structure representative figure. C. Representation of gag-knuckle-like fold.



**Figure 3.11: Ribbon structure of parkin RING2 domain.**

RING2 domain is illustrated with  $\alpha$  helix and  $\beta$  sheets, with catalytic cysteine and zinc coordinating cysteines highlighted in yellow.

The RING2 in IBR-RING2 structure adopts the sequential zinc coordination, resulting in more elongated and less compact structures compared to the canonical RING. This structural difference between the canonical RING and RING2 of parkin leads to the conclusion that RING2 is not a canonical RING, and would have a special function in RBR E3 ligase.

Based on the prediction of zinc coordination (Table 2), C449 was suspected to be in a reduced state, and not coordinating a zinc ion. As discussed in section 1.3.3, RBR E3 ligases are suspected to have RING2 with catalytic cysteines, allowing the RING/HECT hybrid function of the enzyme. The HHARI C357 residue that was identified to have catalytic activity (Wenzel and Klevit, 2012), and the sequence alignment of HHARI with parkin showed that C449 of parkin is the conserved catalytic cysteine among all the human RBR ligases. With this in mind, the structure of RING2 was analyzed. As predicted, C449 was not one of the zinc coordinating residues, and it was not part of a region with secondary structure (Figure 3.11). It was rather located on the surface loop of RING2, well exposed for any potential catalytic activities. The existence of a solvent exposed catalytic cysteine on RING2 supports that it is not a canonical RING, and it would play the role of a catalytic domain suggested by the RING/HECT hybrid mechanism of RBR type E3 ligases.

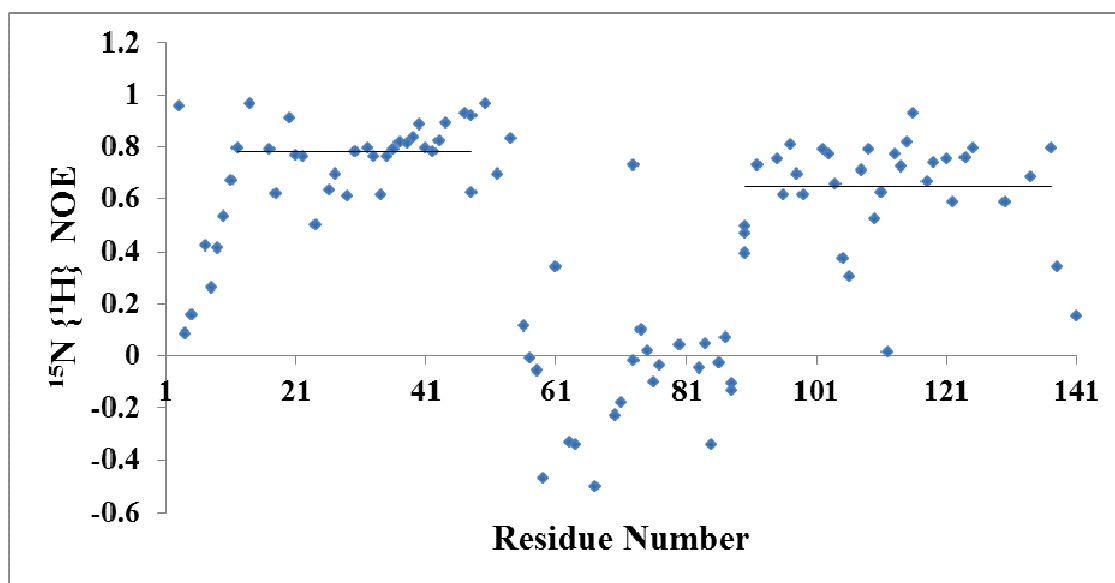
There was no notable direct interaction between the IBR and RING2 domains from the structure analysis. To further support the finding that the linker between the IBR and RING2 domains is unstructured and dynamic, additional experiments were conducted.

### 3.4 Interaction Studies of IBR and RING2 domains

#### 3.4.1 *The flexible linker of IBR-RING2 is confirmed with protein dynamics studies*

To further characterize the two domains in IBR-RING2, a  $^{15}\text{N} \{^1\text{H}\}$  heteronuclear NOE experiment was performed. It was expected that this experiment would allow investigation into whether the two domains are interacting (displaying a globular structure), or acting as two separate domains (showing elongated structure) (Merkley and Shaw, 2004). Since each of the domains are approximately 40 amino-acid-long, it was expected that the value of the average NOE would be smaller if each two domain acted separately compared to the NOE if the domains were interacting and behave as a more globular species.

Figure 3.12 displays the graph of heteronuclear NOE values for each residue in IBR-RING2. By determining the ratio of the intensities with saturation and no saturation, heteronuclear NOE values were calculated. Flexibility of the linker region was confirmed from the experiment, since it showed near-zero/negative NOE values, a common observation when there is a flexible hinge between the domains. Additionally, this observation was comparable to the values obtained for calmodulin (Barbato et al., 1992), cytokinesis protein Cdc4p (Slupsky et al., 2001), and Ubc1 (Merkley and Shaw, 2004) which all possesses a flexible linker region. This indicates that the linker between the IBR and RING2 domains has high mobility on the nanosecond time scale. The average NOE for the IBR and RING2 regions were 0.80 and 0.65, respectively. These NOE values were similar to that of the 15 kDa domain, such as ubiquitin-like conjugating enzyme Ubc9 (which had the NOE of 0.72) (Liu et al., 1999). It was expected that a



**Figure 3.12: Steady-state heteronuclear NOE values for backbone amides of  $^{15}\text{N}$ -labeled IBR-RING2, obtained at 600 MHz.**

Negative NOE values in the IBR-RING2 domain reflect increased flexibility with respect to the two other domains (Positive NOE: structured region, Negative NOE: unstructured and flexible). NOE values were determined as the ratios of the peak intensities, measured from spectra recorded with and without proton spin saturation.



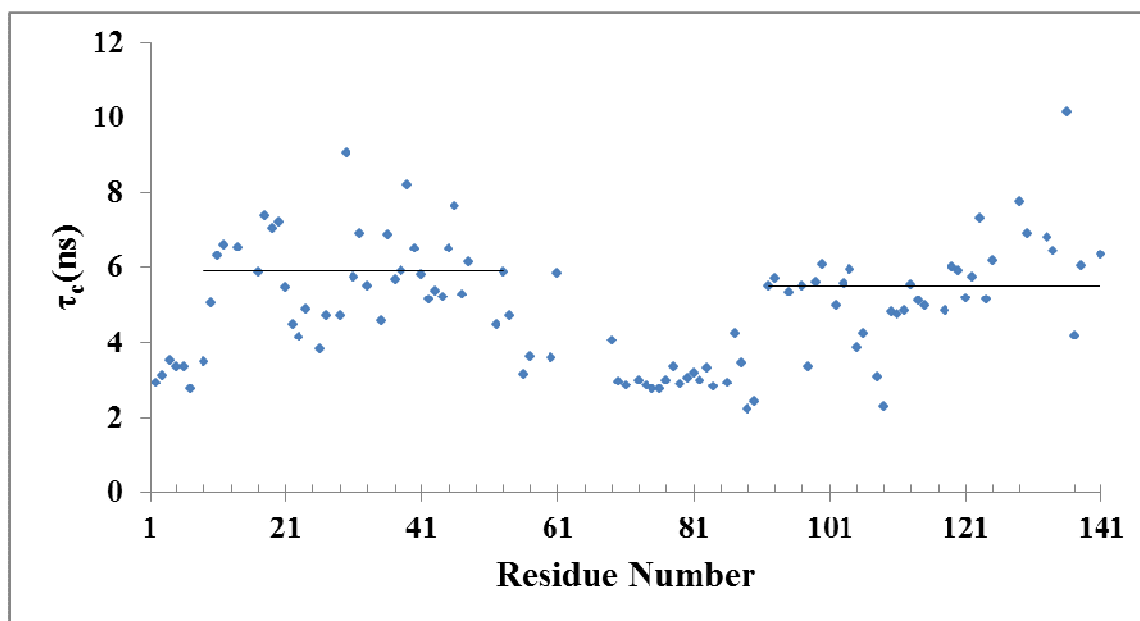
heteronuclear NOE closer to 0.7 (close to 5 kDa proteins) would be obtained based on calculated molecular weight of each IBR or RING2 domains. Therefore, it was not clear from the heteronuclear NOE experiment that differences between a compact or extended forms of the IBR-RING2 could be determined.

Another approach taken to examine the flexibility of the linker between IBR-RING2 further was involving  $T_1/T_2$  relaxation experiments. The objective of this set of experiments was to gain more information about the particle size of IBR-RING2. A particle in a solution is rotated by one radian at different rates, depending on the size of the particle. The time it takes for the particle to rotate is referred to as rotational correlation time ( $\tau_c$ ), a value which can be useful in biochemistry when approximating the size of a molecular weight. It has been shown that for a rigid protein molecule that is smaller than 25 kDa,  $\tau_c$  can be described as a function of the ratio between the longitudinal and transverse  $^{15}\text{N}$  relaxation times (Kay et al., 1989).

$$\tau_c \approx (4\pi\nu_{\text{N}})^{-1} * [(6T_1/T_2) - 7]^{1/2}, \text{ where } \nu_{\text{N}} \text{ is the } ^{15}\text{N} \text{ resonance frequency.}$$

Backbone  $^{15}\text{N}$  spin relaxation experiments are used to obtain information about the residue-specific dynamics of proteins. In this project, longitudinal and transverse measurements were conducted by fitting the  $T_1$  and  $T_2$  graph to the exponential decay curve.

Figure 3.13 shows a graph for  $\tau_c$  that is calculated by dividing  $T_1$  by  $T_2$ , for each residue of IBR-RING2. Average  $\tau_c$  values for the IBR and RING2 regions (shown by dark lines) were 5.9 ns, and 5.5 ns, respectively. The linker region's  $\tau_c$  value was much lower (below four), and also less scattered.



**Figure 3.13: Backbone  $^{15}\text{N}$  spin relaxation measurement of IBR-RING2** (25 mM Tris, 150 mM NaCl, 5 mM DTT at pH 7.5, 25°C for 300  $\mu\text{M}$  IBR-RING2).  $T_1$  and  $T_2$  relaxation time values were obtained from the peak intensities measured from spectra recorded with the shortest delay time to that of spectra with longer delay times. Then  $T_1/T_2$  ( $\tau_c$ ) was calculated and plotted on a graph as shown.

The  $\tau_c$  value of small molecules moving randomly in solution would be shorter than that of large molecules that form a protein complex. This indicates that the linker region is moving more freely in solution than the two domains in IBR and RING2. Moreover, the prediction of  $\tau_c$  can be approximated to be half the value of its molecular weight (Anglister et al., 1993).

$$\tau_c [\text{ns}] \approx 1/5 T_2 [\text{s}] \approx 1/2 \text{ MW} [\text{kDa}], \text{ where MW is molecular weight and } [\ ] \text{ shows units}$$

Based on this equation, a globular protein that has molecular weight close to IBR-RING2 (~16 kDa) should have a  $\tau_c$  value of about 8 ns. Also, the rotational correlation time values compiled by Aramini and colleagues (2010) showed that the 7.2 kDa protein PsR76A has a  $\tau_c$  of 5.1 ns, and that of 15.8 kDa protein ER541-37-162 is 10.0 ns. These predicted values of  $\tau_c$  compared to the acquired value of IBR-RING indicates that the protein is not globular. Therefore, these findings show that IBR and RING2 are two separate domains with an approximate size of 7-8 kDa.

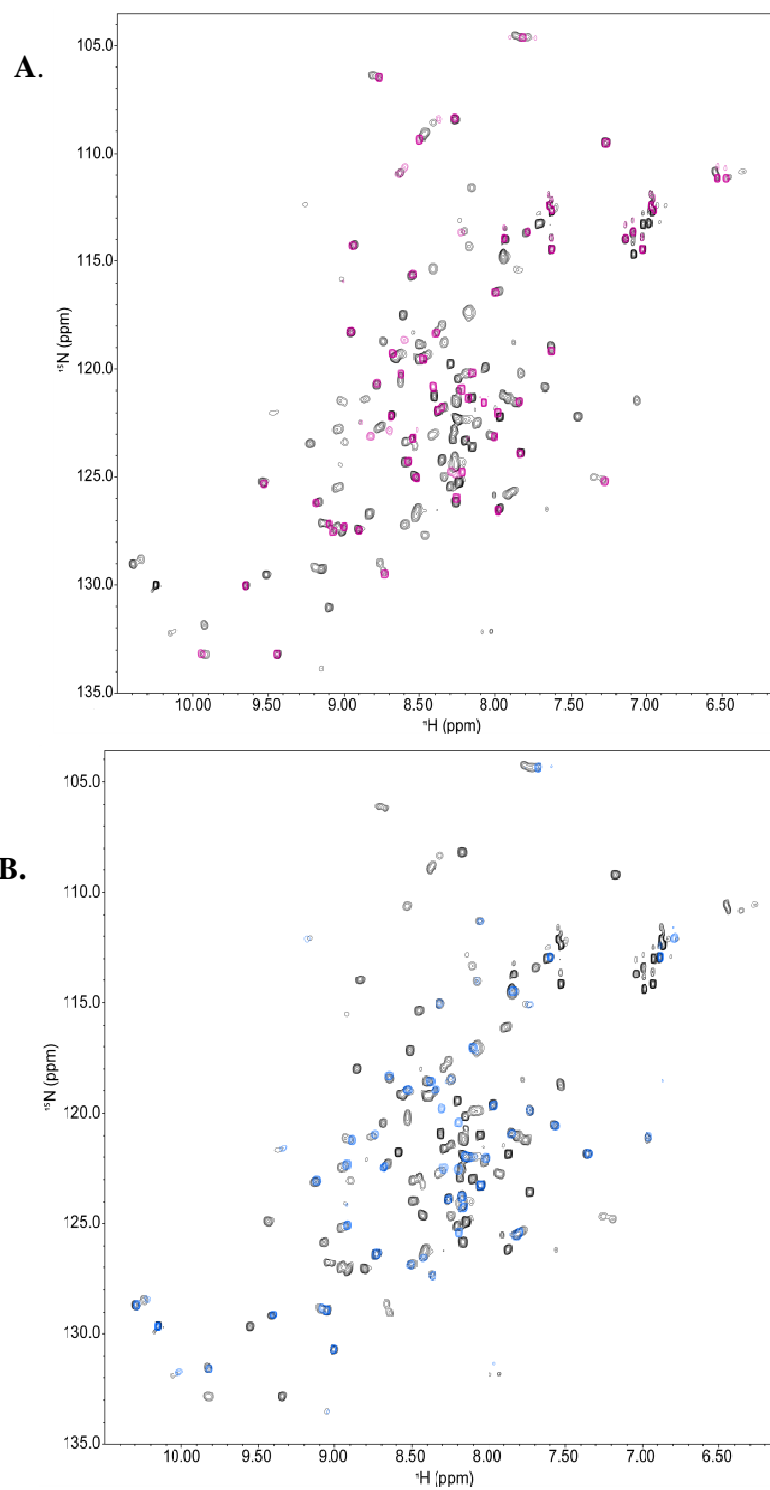
In these experiments, it was important to monitor the change in the intensities of the peaks during data collection. Thus, stability of the sample was crucial. However, the unstable nature of the IBR-RING2 made it difficult to maintain the sample for the duration of data collection and could have produced inconclusive results. This explains some of the unexpected results, such as values greater than 1 for heteronuclear NOE or residues that show non-exponential decay in  $T_1/T_2$  experiments that were not included in the graphs. However, the original data of heteronuclear NOE graph with outliers is presented in Appendix B, Figure B-1. The  $T_1$  and  $T_2$  decay curves of well fitted residues

and poorly fitted residues are also presented to show the quality of collected data in Appendix B, Figure B-2.

### ***3.4.2 Confirmation of non-interacting IBR and RING2 domains with $^1\text{H}$ - $^{15}\text{N}$ HSQC spectra***

To further probe the interaction between IBR and RING2, two additional experiments using NMR spectroscopy were performed with the individual domains of IBR and RING2. These were conducted in an attempt to demonstrate that in the absence of the linker, the spectrum of IBR-RING2 would display residues at the same/similar position to the individual domains of IBR or RING2. This observation would indicate that the environment of the residues of each domain are not impacted by other domains, confirming the that the IBR and RING2 domains do not interact with each other in solution.

The first supplementary experiment conducted involved superimposing the HSQC of individual domains, IBR(342-402) and RING2(417-482), to the full length IBR-RING2 protein for comparison. The second involved conducting a titration experiment, using labeled IBR with non-labeled RING2. A change in any peak position in either the IBR or RING2 domains was expected, as a sign of possible interaction. The absence of this observation would confirm that the IBR and RING2 domains do not interact with each other in solution, whether or not they are connected by a linker.



**Figure 3.14: Superposition of the IBR-RING2  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum with each individual domain.**

Black contours: IBR-RING2, Pink contours: IBR, and Blue contours: RING2. **A.** Superposition of IBR-RING2 with IBR, **B.** Superposition of IBR-RING2 with RING2

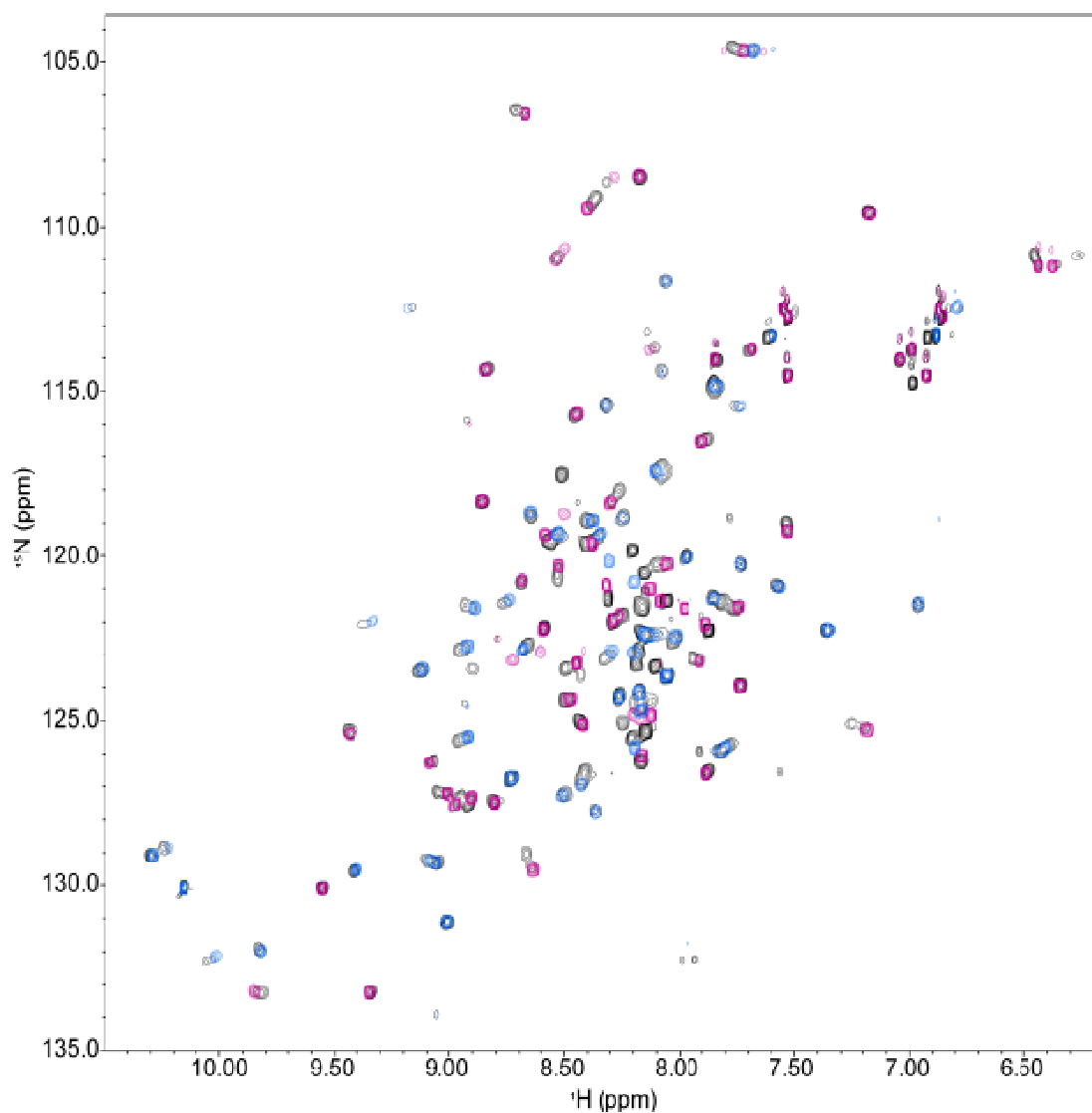
### 3.4.2.1 Individual domains of IBR and RING2

$^1\text{H}$ - $^{15}\text{N}$  HSQC of IBR (residues 342-402) and RING2 (residues 417-482) were collected using the same condition as for IBR-RING2. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for IBR and RING2 against the spectrum of IBR-RING2 showed that most of the peaks overlapped (Figure 3.14) suggesting that there is no interaction between the two domains. Also, when both spectra are superimposed on IBR-RING2 (Figure 3.15), some of the peaks that are not superimposable by both domains are found in the linker (403-416), therefore they would not be expected to superimpose as the linker is not intact in the separate domains. The postulate that IBR and RING2 are two non-interacting domain connected by a flexible linker is strongly supported by this experiment.

### 3.4.2.2 Titration experiment with labeled IBR and non-labeled RING2

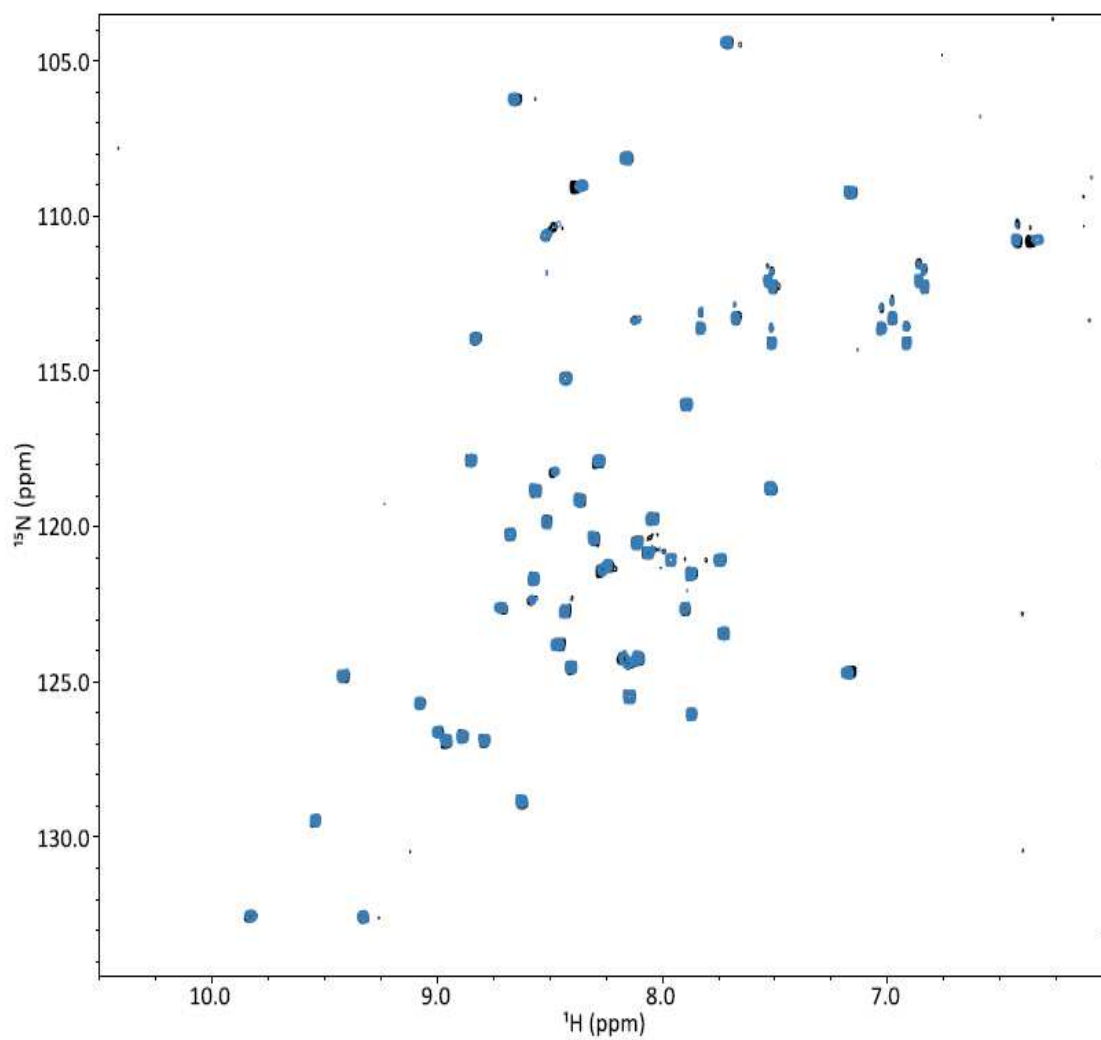
To confirm that there is no interaction between the IBR and RING2 domains, unlabeled RING2 was titrated into  $^{15}\text{N}$ -labeled IBR. The peaks corresponding to the IBR domain did not show any significant change in chemical shift even in the presence of two equivalents of RING2 (Figure 3.16). When there is an interaction, peaks that are involved in the site of interaction would be expected to change in chemical shift or peak intensity. Since the spectra are overlapping almost perfectly, this further supports that no interaction occurs between the isolated IBR and RING2 domains of parkin.

In summary, the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for the individual IBR and RING2 domains were superimposable with IBR-RING2 spectrum. Also, the chemical shift changes were absent when parkin RING2 was titrated into  $^{15}\text{N}$ -labeled IBR. These



**Figure 3.15: Superposition of the IBR-RING2  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum with each individual domain, both IBR and RING2 displayed.**

Black contours: IBR-RING2, Pink contours: IBR, and Blue contours: RING2. Most peaks of IBR-RING2 are overlapping with each other or at a close proximity.



**Figure 3.16: Titration of RING2 into  $^{15}\text{N}$  labeled IBR.**

Black spectrum is the original spectrum of IBR, and blue is IBR with the addition of RING2. Since most of the peaks are perfectly overlapping, it appears as if there are not many black peaks in the spectrum.

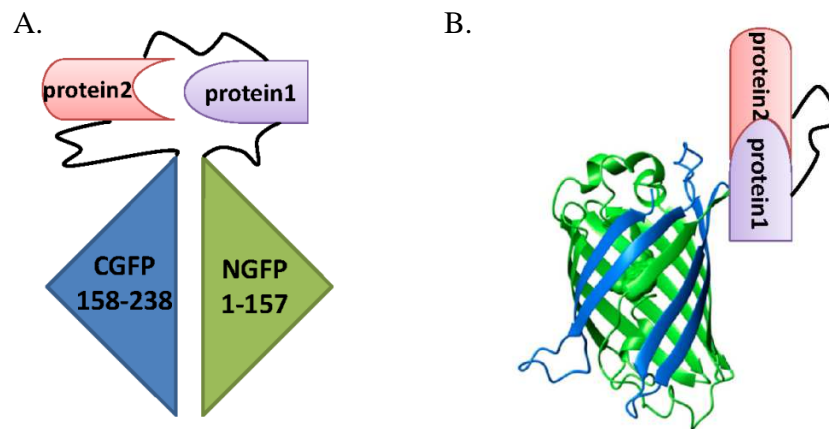


cumulative results indicate that, in the context of native parkin, the IBR and RING2 domains do not associate with each other.

### 3.4.3 *Split GFP*

Numerous attempts were made in order to obtain soluble parkin RING1 domain or RING1-IBR-RING2 domain to possibly examine chemical shift changes in  $^1\text{H}$ - $^{15}\text{N}$ -labeled IBR-RING2 spectrum in the presence of unlabeled RING1. This was to identify the region of interaction in IBR-RING2 with RING1 when all the components of the RBR domain of parkin are present. This would help identify the relationship between the RING1 and RING2 domains. Since the RING1 is expected to recruit the complex of E2 enzyme bound to ubiquitin and pass the ubiquitin onto the RING2, identification of the spatial arrangement of RING1 and RING2 might allow some details of this mechanism to be revealed. In the absence of the RING0 domain, purification of the RING1 or RBR of parkin was proven to be very difficult, with the notable exception of RING2 or IBR-RING2. Therefore, an alternative method was explored to monitor the interactions between the RING1 and IBR-RING2 proteins. This method involved cloning the parkin RING1-IBR-RING2 into a split GFP vector and observing the reformation of GFP should an interaction between the RING1 and RING2 domains occur. The advantage of this approach is that only a small fraction of the protein needs to remain soluble to observe the green fluorescence for detecting interaction.

Split green fluorescent protein was shown not to reassemble when split in a loop between residues 157 and 158 and produced in *trans* in an expression system (Ghosh et al., 2000). Upon insertion of two interacting domains that facilitate the assembly of the split GFP, the fluorescence of the split GFP can be obtained. With this property, the



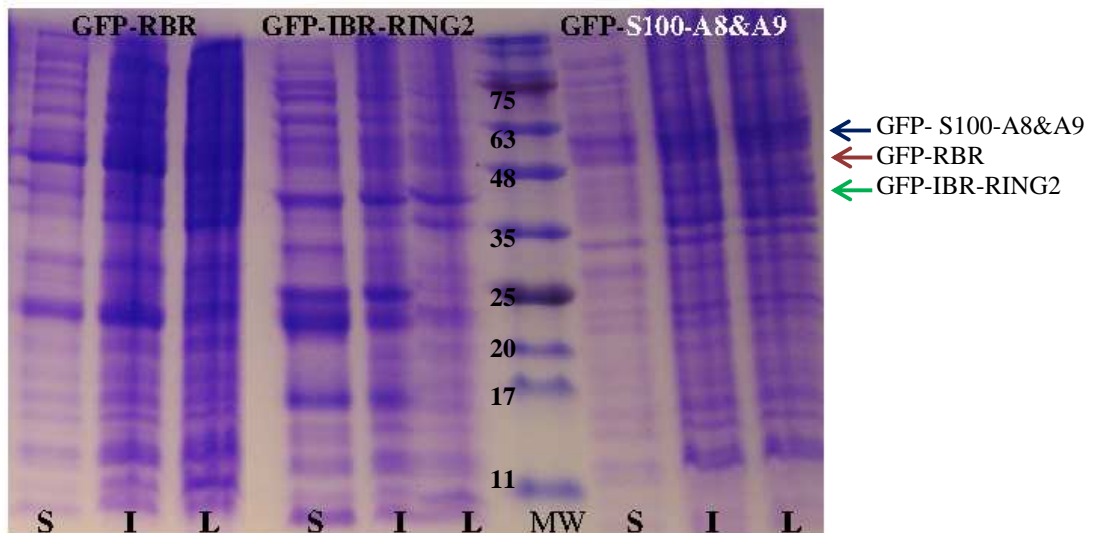
**Figure 3.17: Illustration of split GFP system.**

**A.** Insertion of two interaction partners to be expressed with split GFP **B.** Upon interaction of the two proteins, split GFP reassembles itself and will fluoresce.

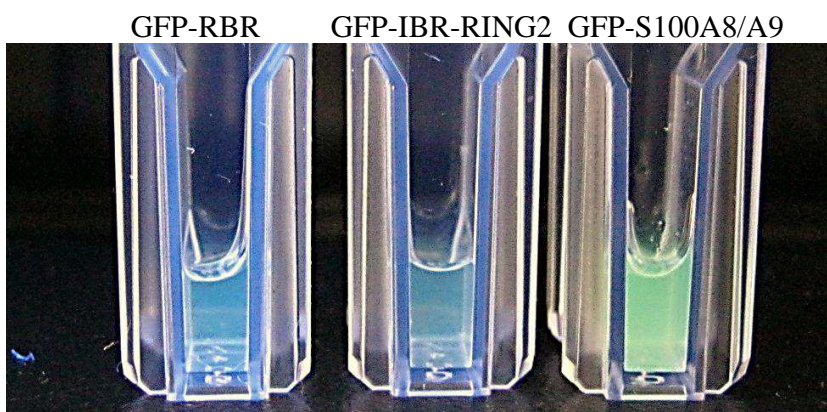
detection of direct protein-protein interactions can be monitored. Figure 3.17 illustrates how the *N* and *C* termini of GFP assemble based on the interaction of two proteins being tested. Also, it has been reported that the interaction strength ( $K_D$ ) required to be tested with split GFP is about 1 mM (Magliery et al., 2005), which makes it possible to detect weak interactions by split GFP analysis. This is the primary reason why split GFP was used to screen for any IBR-RING2 interaction. For the purpose of my research, the insertion of RBR and IBR-RING2 into this split GFP system was conducted in the hope of detecting any possible interaction between RING1 and RING2, or any other combination possible.

The split GFP-parkin RING1-IBR-RING2 construct was designed with the hope of monitoring a possible interaction between the RING1 and IBR-RING2 domains. It was suggested that the RING1 and RING2 domains might be in close proximity in full-length parkin (Beasley et al., 2007). To test this, IBR-RING2 and RING1-IBR-RING2 were incorporated into the split GFP Duet vector and tested for expression. The calcium binding proteins (S100A8 and S100A9) known to form a tight ( $K_D < 1 \mu\text{M}$ ) functional heterodimer, were used as a positive control.

All soluble fractions of the cell lysates were confirmed to have the expression of the desired protein by SDS-PAGE (Figure 3.18). Observation under ultraviolet (UV) light was used to determine whether any of them fluoresced as an indication of an interaction between RING1 and IBR-RING2 or IBR with RING2 interaction. The positive control S100A8:S100A9 showed bright green fluorescence under the UV light, indicating an interaction between the two proteins (Figure 3.19). The split GFP-parkin IBR-RING2 and split GFP-parkin RING1-IBR-RING2 proteins did not fluoresce, indicating that despite



**Figure 3.18: Solubility test of split GFP constructs shown as labelled at the top.** First three lanes are for RBR, next three are for IBR-RING2 and last three lanes are for S100A8/A9 complex in split GFP. At the bottom, lanes are identified by samples from the solubility test, labeled with soluble fraction (S), insoluble fraction (I), and crude cell lysate (L), respectively.



**Figure 3.19: Soluble fraction of cell lysate split GFP fusions with RBR, IBR-RING2, and S100A8/A9.**

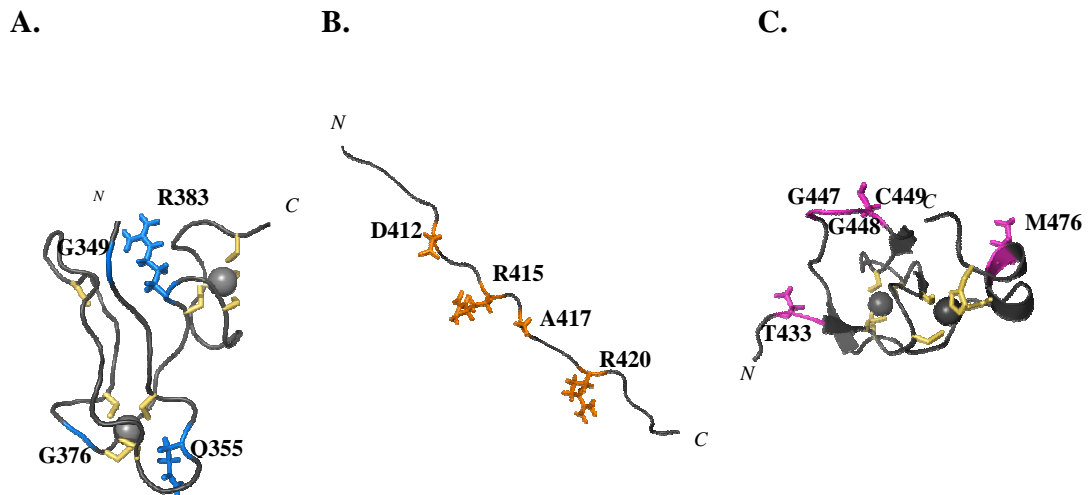
The UV lamp was shone to maximize the fluorescence from split GFP to observe any fluorescence indicating an interaction between the inserted domains. It is clear only the positive control, S100A8/A9 complex, fluoresced in green, whereas the other samples did not.

the long linker between the IBR and RING2 domains, they do not form a complex for interaction. An interaction between RING1 and IBR-RING2 was not detected using the split GFP vector; this can be interpreted that no direct interaction occurs between RING1 and RING2.

### 3.5 Mutation analysis

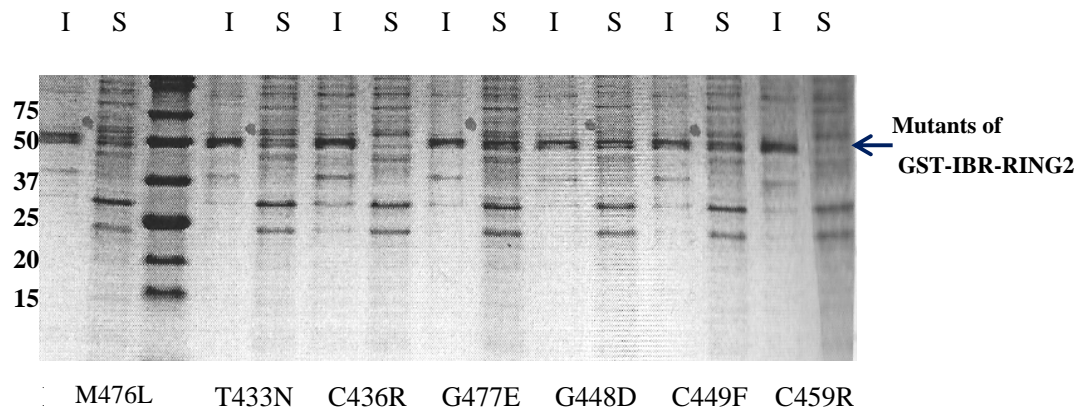
Determining the clinical significance of missense mutations of parkin was difficult, primarily due to the large number of different mutations possible on the *PARK2* gene (Pankratz and Foroud, 2007). The significant number of varying mutations can be attributed to the likeliness of less detrimental effects of missense mutations on enzyme function as compared to exon deletion or duplication (Pankratz and Foroud., 2007), which were shown to have direct impact on parkin function. It still remains unclear whether homozygous or compound heterozygous parkin mutations are more likely to be the cause of diseased state parkin (Kay et al, 2007).

Within the parkin IBR-RING2, there are over 20 missense mutations that have been identified in ARJP patients included in HGMD. These mutations include: G349E, Q355C, T372P, G376D, R383Q in the IBR domain, D412N, R415Q, A417T, R420C, R420P in the linker, and T433N, C436R, G447E, G448D, C449F, C459R, and M476L in RING2 domain, as shown in Figure 3.20. In this study, all of the soluble mutations were successfully designed, expressed, and purified for NMR studies. Determining the effects of each of these mutations on the structure was the main goal of the mutational analysis. All constructs were expressed and purified under conditions identical to the wild-type IBR-RING2. Figure 3.21 shows the solubility test of IBR-RING2 for the mutations in RING2 region. It was clear that the solubility of IBR-RING2 is only impacted when the



**Figure 3.20: Disease state substitutions within the parkin IBR-RING2. Ribbon drawings of different domains of parkin.**

A. IBR, B. linker C. RING2 showing positions of the ARJP causative mutation sites, with sticks to indicate the side chain positions in the protein (blue in IBR, orange in linker, and pink in RING2). Also, zinc coordinating cysteines and histidines are shown in yellow.



**Figure 3.21: Solubility test of mutants of IBR-RING2, (missense mutations on RING2 only).**

I: insoluble, S: soluble fraction, and the mutations tested are listed at the bottom of the gel. Clearly C436R and C459R are not soluble due to the mutations on the zinc coordinating cysteine. C449F, on the other hand, is still soluble.

zinc coordination is disrupted, as C436R and C459R were shown to be insoluble, while the rest were identified to be soluble.

In total, 16 IBR-RING2 mutants were successfully created by mutagenesis, and expression tests were used to identify the insoluble mutants. As expected, the residues identified to be involved in zinc coordination resulted in poor solubility of the IBR-RING2. This was likely a result of improper zinc coordination causing the IBR-RING2 to be unstable and insoluble (Figure 3.21). Multiple  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the soluble IBR-RING ARJP substituted proteins were collected to determine the effect of these substitutions on the structure of parkin IBR-RING2. It was speculated that changes in the structure would be localized in the region where the substitution is present, since there is not much interaction between IBR and RING2 (Mutations in the IBR region would only affect the peaks of IBR, leaving the peaks of RING2 unchanged). There were some unique mutants that did not change the HSQC greatly, only affecting a couple of peaks, and those mutants were located in the linker region between IBR and RING2 (D412N, R415Q, A417T, R420C, and R420P).

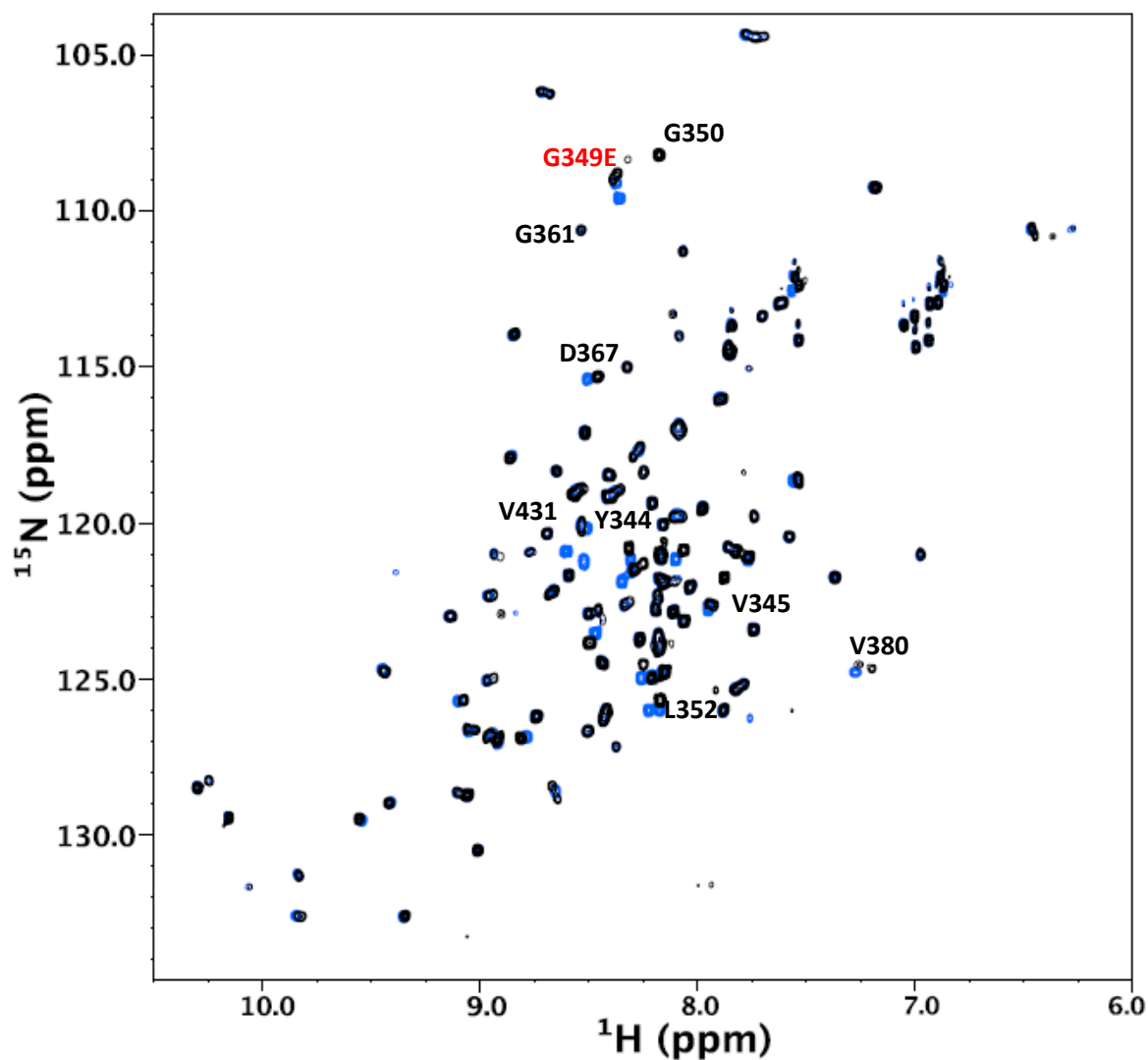
Analysis of the disease-causing mutants' effect on the IBR-RING2 structure was done by comparing the  $^1\text{H}$ - $^{15}\text{N}$  HSQC of  $^{15}\text{N}$  labeled samples with wild type (WT). For example, observing changes in amino acids distanced from regions of the substitution that are considered to be important residues would indicate that the substitution caused a change in the structure by globally affecting residues throughout the IBR or RING2 domains. The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra (Figures 3.22-24) for all IBR-RING2 mutants had well-dispersed peaks and patterns similar to the peak positioned for wild-type IBR-



RING2. This indicated that these mutations did not affect the three-dimensional structure of IBR-RING2.

One of the point mutations in the IBR region linked to ARJP, G349E (human G328E), is located on the surface of the IBR domain. It was inferred previously with the human IBR domain (Beasley et al., 2007) that this substitution would disrupt the interaction surface of the loop where G349 is located in, resulting in the ARJP disease state. In fruit fly neurons, the parkin mutants was found to be insoluble in the cell lysate, further suggesting that the high expression level of this disease state of parkin cannot be tolerated (Wang et al., 2007). This interesting observation correlates with the lower yield obtained, when expressed in *E. coli*, of this ARJP substitution of parkin IBR-RING2. It has also been reported that ubiquitination of PLC $\gamma$ 1, one of parkin's substrates, was lowered when compared to the level of G349E with the wild type (Dehvari et al., 2008). The G349E ARJP substitution was also identified to be a pseudo-dominant parkin mutation, since it is often found on one allele in patients with ARJP (Periquet et al., 2001). The residues affected by this point mutation are in close proximity, such as V345, G350, L352, and V380 (Figure 3.22). This indicates that this point mutation does not impact the domain's fold but does cause localized chemical environment changes. Insolubility of this mutant in high expression levels could indicate its potential to act as a neurotoxin, but that conclusion could neither be supported nor opposed by this mutational analysis as the protein was soluble when expressed in different organisms.

The missense mutation in the linker region D412N (human: D394N) was reported to be a benign polymorphism (Kay et al., 2007), despite being listed as an ARJP-related mutation in the HGMD. This is due to the presence of this mutation in patients



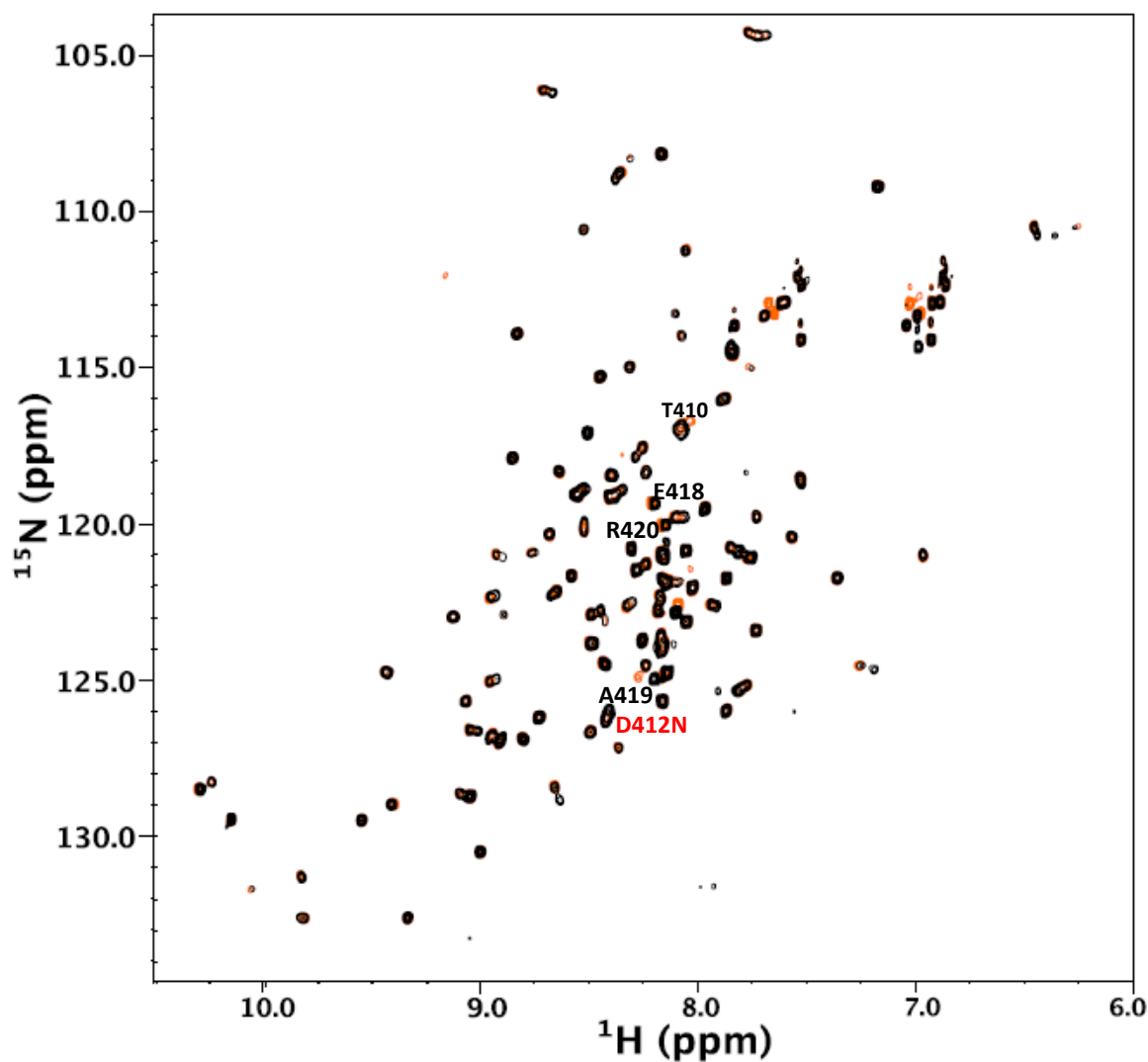
**Figure 3.22: Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in IBR region (G349E).**

The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and blue peaks belong to the spectrum of G349E mutant of IBR-RING2.

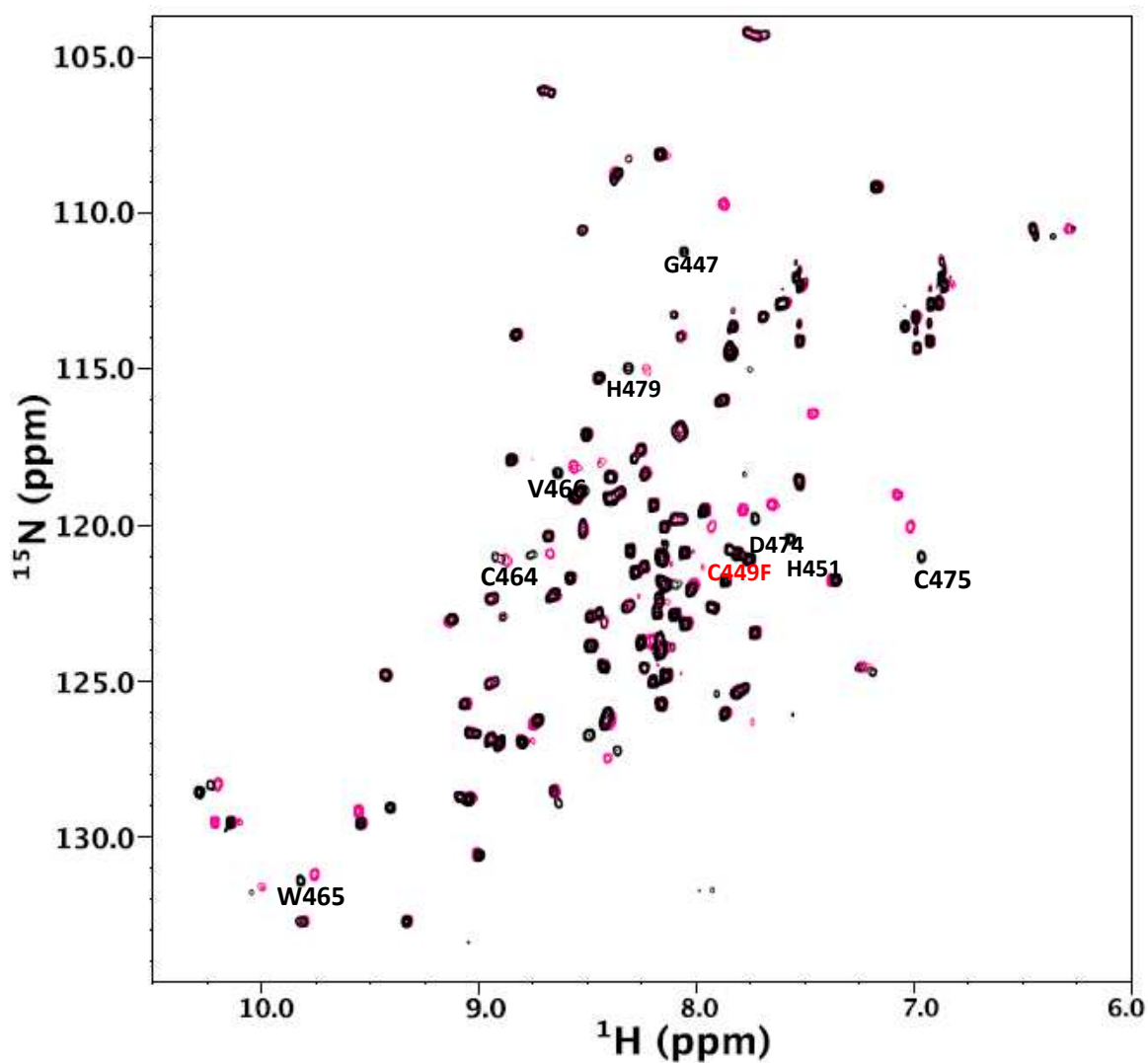
distinguished with a sporadic form of PD, thus making it difficult to conclude that there is a direct link between this substitution and the disease (Figure 3.23). No significant changes were observed in peak positions with the D412N substitution for either IBR or RING2. While this may support the previous finding that D412N is not directly linked to ARJP, it does not exclude the possibility that other domains of full-length parkin may be impacted by this substitution. For example, the D412N substitution could disrupt interactions between other domains and the linker which in turn would inhibit function of parkin. Alternatively, this substitution might also be completely benign and not affect parkin function.

In the RING2 region, the missense substitution C449F (human C431F) was shown to impair parkin E3 ubiquitin-protein ligase activity toward the interacting substrates, such as ZNF746 (Shin et al., 2011) and Bcl-2 (Chen et al., 2010). This cysteine residue of RING2 is suspected to be a catalytic site for RBR type E3 ligase function. According to the structure, C449F is exposed to the solvent, since it does not coordinate a zinc ion (Figure 3.11). It is well conserved throughout orthologous and paralogous organisms, providing compelling evidence that this cysteine is catalytic. For this reason, substitution of this residue would be expected to have a detrimental impact on the function of parkin. As expected, there are many chemical shift changes that occur in HSQC when this residue (C449) is mutated (Figure 3.24), compared to other mutations in the RING2 region as shown in Appendix C.

Of all the HSQCs of IBR-RING2 mutants, Q355C (Appendix C, Figure C-1) resulted in major shifts in all peak positions, as compared to the wild-type protein. This can be rationalized by the location of the residue in structure. As shown in Figure 3.20,



**Figure 3.23: Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the linker region (D412N).**  
The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and orange peaks belong to the spectrum of D412N mutant of IBR-RING2



**Figure 3.24: Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the RING2 region (C449F).**  
The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and pink peaks belong to the spectrum of C449F mutant of IBR-RING2.

Q355 is very close to the first zinc coordination site of the IBR domain (near the *N*-terminus), and the introduction of an additional cysteine in this position may have disrupted proper zinc coordination. As shown by removal of cysteines with C436R and C459R (Figure 3.21), disruption of proper zinc coordination in IBR-RING2 had a detrimental impact on the protein folding and solubility. The introduction of an additional cysteine residue near the zinc coordination site I of IBR may have modified the proper folding of the protein, although it did not result in the protein becoming insoluble. Interestingly, introduction of an additional cysteine residue (R420C) much further away from the zinc coordination site, however, did not result in major shifts in peak position. The R420 mutation is in the linker region (Figure 3.20), where neither IBR nor RING2 domains are in close proximity, so additional cysteines only impacted residues in its vicinity.

It can be shown through the mutational analysis of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra that ARJP substitutions caused minor changes in the IBR-RING2 structures. Most of the peaks that display changes are local as evident from the spectra (Figures 3.22-24). Substitutions in the IBR region only impact the resonance peaks in the IBR domain, while substitutions in RING2 domain changed the resonance peaks in RING2 only. This is consistent with the result that IBR-RING2 are two separate domains, since if there was an interaction between these two regions, they would display some changes in other regions.

An interesting trend observed in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra collected for ARJP substitutions in the linker region between the IBR and RING2 domains is the presence of significantly fewer peaks affected by any of the substitutions. This would indicate that

substitutions in the linker do not affect the chemical environment of residues found in either the IBR or RING2 domains of parkin, further supporting the flexibility of the linker.

Both of these findings support what was expected based upon the structure calculation along with NMR titration data analysis. The  $T_1/T_2$  experiment results are also in agreement with discovery of the flexibility of the linker in IBR-RING2.

According to the solubility test of IBR-RING2 mutants in the RING2 domain, when the zinc-coordinating cysteine is affected, the protein becomes insoluble due to its inability to bind to the required structural zinc ion. The insolubility of these substitutions (C436R and C459R) did not allow for  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra to be collected. However, these findings further demonstrate the importance of zinc ions in the proper folding of IBR-RING2. Without the zinc, parkin IBR-RING2 cannot fold properly and can fall out of solution. Interference of zinc coordination, such as addition of EDTA into the protein (causing the precipitation of human IBR), also agrees with this finding.

Clearly, it is evident that a structural rearrangement does not occur due to the soluble ARJP substitutions. Also, it is not presently clear from the data the cause of PD by these ARJP substitutions, as they did not adversely affect the protein's fold. Despite this, the mutational analysis does provide new information that IBR-RING2 has a flexible linker and that there is no interaction between the two domains.

### 3.6 Summary

GST fusion tagged IBR-RING2 has been successfully expressed and purified. The protein fold of IBR-RING2 was confirmed with an  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiment. MS analysis showed that there are a total of four zinc ions bound to the IBR-RING2. A zinc coordinating cysteine prediction of IBR-RING2 also confirmed the number of zinc ions in the structure to be four.

Multiple NMR experiments were performed to determine the structure of the IBR-RING2, by making the appropriate chemical shift assignments. After assigning the backbone and side-chains, the chemical shift index was used to predict the secondary structure of IBR-RING2. Also, the structure calculation with CYANA confirmed that there was not much secondary structure in either IBR-RING2. The structures of IBR and RING2 were very similar and the linker between them did not have much structure. The structure of IBR-RING2 supported that RING2 is most likely not involved in the recruitment of an E2 enzyme, and it also confirmed the existence of solvent exposed catalytic cysteine for RING2.

Flexibility of the linker region between IBR and RING2 were confirmed with multiple experiments: NMR rate analysis (heteronuclear NOE and T1/T2 experiments), titration experiment, overlay of single domain  $^1\text{H}$ - $^{15}\text{N}$  HSQC, and split-GFP system. This lead to the conclusion that IBR and RING2 are two non-interacting domain connected by a linker.

The ARJP substitution study of IBR-RING2 by  $^1\text{H}$ - $^{15}\text{N}$  HSQC monitored the possibility in protein cold changes. The IBR and RING2 showed mostly local changes



around the point mutations, while the mutations in the linker region did not show much change. Moreover, the importance of zinc-coordination for these proteins was again confirmed by drastic changes in zinc coordinating cysteine impacting the structure greatly.

### **3.7 Closing Thoughts & Future direction**

The finding that last two C-terminal domains of parkin (IBR-RING2) are not interacting implies that there are other domains that must come into contact in full length parkin. Furthermore, this result suggests that the original view on the IBR's function of bringing RING1 and RING2 together for proper RBR domain function may not hold true. The recently published structures using X-ray crystallography of the majority of parkin (RING0-RING1-IBR-RING2) showed the assembly of RING0 and RING1 in the linker region between IBR and RING2 (Trempe et al., 2013; Wauer and Komander, 2013; Riley et al., 2013). As the structures of IBR and RING2 from the crystal structure were very similar to the structure in this thesis determined using NMR spectroscopy, it supports the completeness of the NMR spectroscopy method of structure calculation.

The flexible nature of the linker suggests the importance of this linker in the role of parkin as an E3 ligase, since it makes contact with both RING0 and RING1. The linker may serve to stabilize or solubilize these two domains, as they were found to be insoluble without IBR and RING2 domains attached.

Originally, the future direction of this project was to determine structures for other parts of the domain of parkin, including RING1, then RING0, and even the N-terminus, UbL. However, with the recent finding of the structure of parkin (RING0-

RING1-IBR-RING2) with X-ray crystallography, it became redundant to repeat the structure calculation with NMR spectroscopy. Still, with the completion of the chemical shifts assignment of residues in IBR-RING2, direct interacting partners of IBR-RING2 can be analyzed, or even titration experiments with other proteins are possible. It would be interesting to develop a method to see how N-terminus of parkin, UbL, regulates the protein to function as an E3 ligase through contacting the rest of the protein.

It has already been shown that the E2 enzymes, UbcH7 and UbcH8, do not interact with the RING2 domain (Spratt et al., 2013), suggesting that it would be another RING domain that is the binding partner of E2. In this work, it was found that expression and purification of RING0 or RING1 domain without IBR-RING2 is difficult. In the future, it would also be interesting to monitor the behavior of RING0 linked directly to RING2. If these constructs behave well, monitoring the residue-specific interaction between RING0-RING2 would contribute to the better understanding in the inter-molecular behavior of the C-terminus of parkin.

### 3.8 References

- Aramini, J.M., Tubbs, J.L., Kanugula, S., Rossi, P., Ertekin, A., Maglaqui, M., Hamilton, K., Ciccocanti, C.T., Jiang, M., Xiao, R., *et al.* (2010). Structural basis of O6-alkylguanine recognition by a bacterial alkyltransferase-like DNA repair protein. *J Biol Chem* 285, 13736-13741.
- Barbato, G., Ikura, M., Kay, L.E., Pastor, R.W., and Bax, A. (1992). Backbone dynamics of calmodulin studied by <sup>15</sup>N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* 31, 5269-5278.
- Beasley, S.A., Hristova, V.A., and Shaw, G.S. (2007). Structure of the Parkin in-between-ring domain provides insights for E3-ligase dysfunction in autosomal recessive Parkinson's disease. *Proc Natl Acad Sci U S A* 104, 3095-3100.
- Chen, D., Gao, F., Li, B., Wang, H., Xu, Y., Zhu, C., and Wang, G. (2010). Parkin mono-ubiquitinates Bcl-2 and regulates autophagy. *J Biol Chem* 285, 38214-38223.
- Dehvari, N., Isacson, O., Winblad, B., Cedazo-Minguez, A., and Cowburn, R.F. (2008). Presenilin regulates extracellular regulated kinase (Erk) activity by a protein kinase C alpha dependent mechanism. *Neurosci Lett* 436, 77-80.
- Ghosh, I., Hamilton, A., and Regan, L. (2000). Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. *J Am Chem Soc* 122, 5658-5659.
- Hristova, V.A., Beasley, S.A., Rylett, R.J., and Shaw, G.S. (2009). Identification of a novel Zn<sup>2+</sup>-binding domain in the autosomal recessive juvenile Parkinson-related E3 ligase parkin. *J Biol Chem* 284, 14978-14986.
- Kay, D.M., Moran, D., Moses, L., Poorkaj, P., Zabetian, C.P., Nutt, J., Factor, S.A., Yu, C.E., Montimurro, J.S., Keefe, R.G., *et al.* (2007). Heterozygous parkin point mutations are as common in control subjects as in Parkinson's patients. *Ann Neurol* 61, 47-54.
- Kay, L.E., Torchia, D.A., and Bax, A. (1989). Backbone dynamics of proteins as studied by <sup>15</sup>N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 28, 8972-8979.
- Kornhaber, G.J., Snyder, D., Moseley, H.N., and Montelione, G.T. (2006). Identification of zinc-ligated cysteine residues based on <sup>13</sup>C<sub>alpha</sub> and <sup>13</sup>C<sub>beta</sub> chemical shift data. *J Biomol NMR* 34, 259-269.
- Krishna, S.S., Majumdar, I., and Grishin, N.V. (2003). Structural classification of zinc fingers: survey and summary. *Nucleic Acids Res* 31, 532-550.
- Liu, Q., Yuan, Y.C., Shen, B., Chen, D.J., and Chen, Y. (1999). Conformational flexibility of a ubiquitin conjugation enzyme (E2). *Biochemistry* 38, 1415-1425.

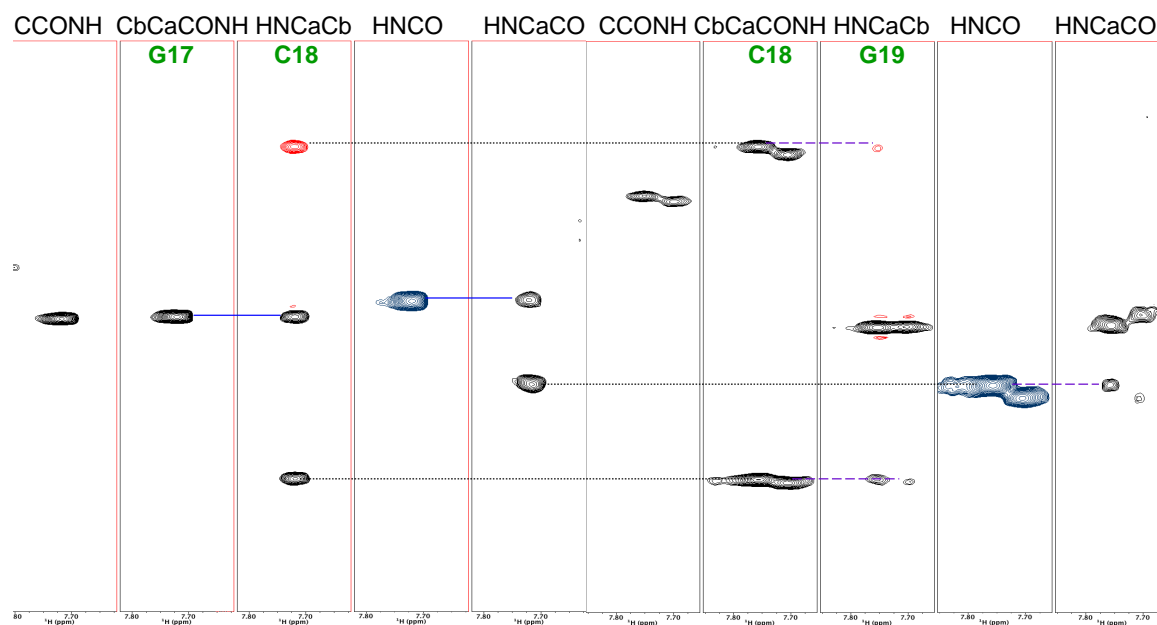
- Magliery, T.J., Wilson, C.G., Pan, W., Mishler, D., Ghosh, I., Hamilton, A.D., and Regan, L. (2005). Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. *J Am Chem Soc* *127*, 146-157.
- Merkley, N., and Shaw, G.S. (2004). Solution structure of the flexible class II ubiquitin-conjugating enzyme Ubc1 provides insights for polyubiquitin chain assembly. *J Biol Chem* *279*, 47139-47147.
- Pankratz, N., and Foroud, T. (2007). Genetics of Parkinson disease. *Genet Med* *9*, 801-811.
- Periquet, M., Lucking, C., Vaughan, J., Bonifati, V., Durr, A., De Michele, G., Horstink, M., Farrer, M., Illarioshkin, S.N., Pollak, P., *et al.* (2001). Origin of the mutations in the parkin gene in Europe: exon rearrangements are independent recurrent events, whereas point mutations may result from Founder effects. *Am J Hum Genet* *68*, 617-626.
- Riley, B.E., Lougheed, J.C., Callaway, K., Velasquez, M., Brecht, E., Nguyen, L., Shaler, T., Walker, D., Yang, Y., Regnstrom, K., *et al.* (2013). Structure and Function of Parkin E3 Ubiquitin ligase reveals aspects of RING and HECT ligases. "submitted for publication".
- Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS+: A hybrid method for predicting backbone torsion angles from NMR chemical shifts. *J Biomol NMR* *44*, 213-223.
- Shin, J.H., Ko, H.S., Kang, H., Lee, Y., Lee, Y.I., Pletinkova, O., Troconso, J.C., Dawson, V.L., and Dawson, T.M. (2011). PARIS (ZNF746) repression of PGC-1alpha contributes to neurodegeneration in Parkinson's disease. *Cell* *144*, 689-702.
- Slupsky, C.M., Desautels, M., Huebert, T., Zhao, R., Hemmingsen, S.M., and McIntosh, L.P. (2001). Structure of Cdc4p, a contractile ring protein essential for cytokinesis in *Schizosaccharomyces pombe*. *J Biol Chem* *276*, 5943-5951.
- Spratt, D.E., Julio Martinez-Torres, R., Noh, Y.J., Mercier, P., Manczyk, N., Barber, K.R., Aguirre, J.D., Burchell, L., Purkiss, A., Walden, H., *et al.* (2013). A molecular explanation for the recessive nature of parkin-linked Parkinson's disease. *Nat Commun* *4*, 1983.
- Trempe, J.F., Sauve, V., Grenier, K., Seirafi, M., Tang, M.Y., Menade, M., Al-Abdul-Wahid, S., Krett, J., Wong, K., Kozlov, G., *et al.* (2013). Structure of parkin reveals mechanisms for ubiquitin ligase activation. *Science* *340*, 1451-1455.
- Wang, H.Q., Imai, Y., Kataoka, A., and Takahashi, R. (2007). Cell type-specific upregulation of Parkin in response to ER stress. *Antioxid Redox Signal* *9*, 533-542.
- Wauer, T., and Komander, D. (2013). Structure of the human Parkin ligase domain in an autoinhibited state. *EMBO J*.

Wenzel, D.M., and Klevit, R.E. (2012). Following Ariadne's thread: a new perspective on RBR ubiquitin ligases. *BMC Biol* 10, 24.

Wishart, D.S., and Sykes, B.D. (1994). The  $^{13}\text{C}$  chemical-shift index: a simple method for the identification of protein secondary structure using  $^{13}\text{C}$  chemical-shift data. *J Biomol NMR* 4, 171-180.

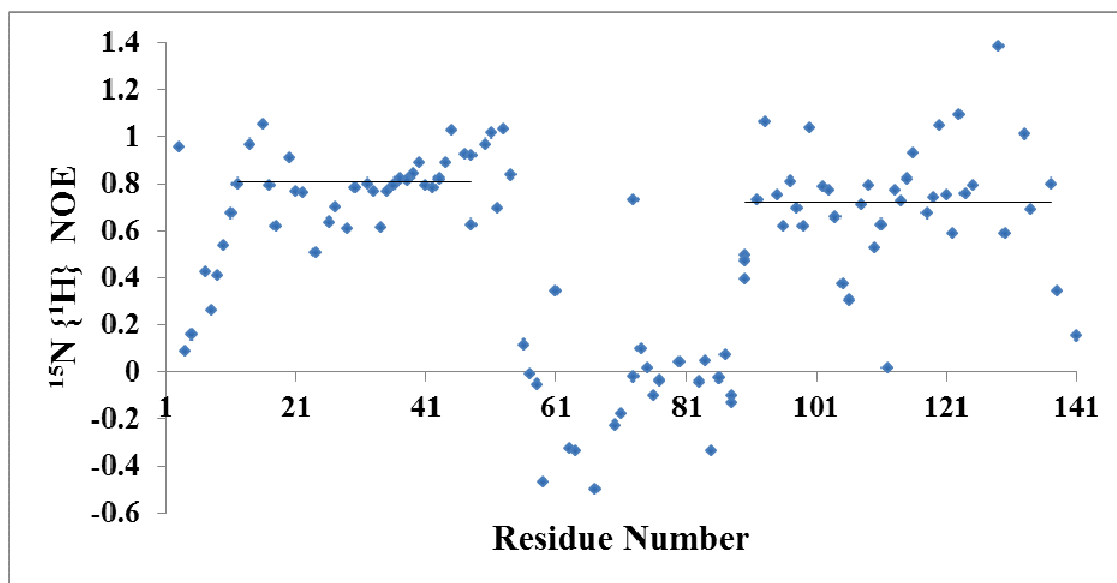
## Appendices

### Appendix A: Data from Structure Calculation of IBR-RING2

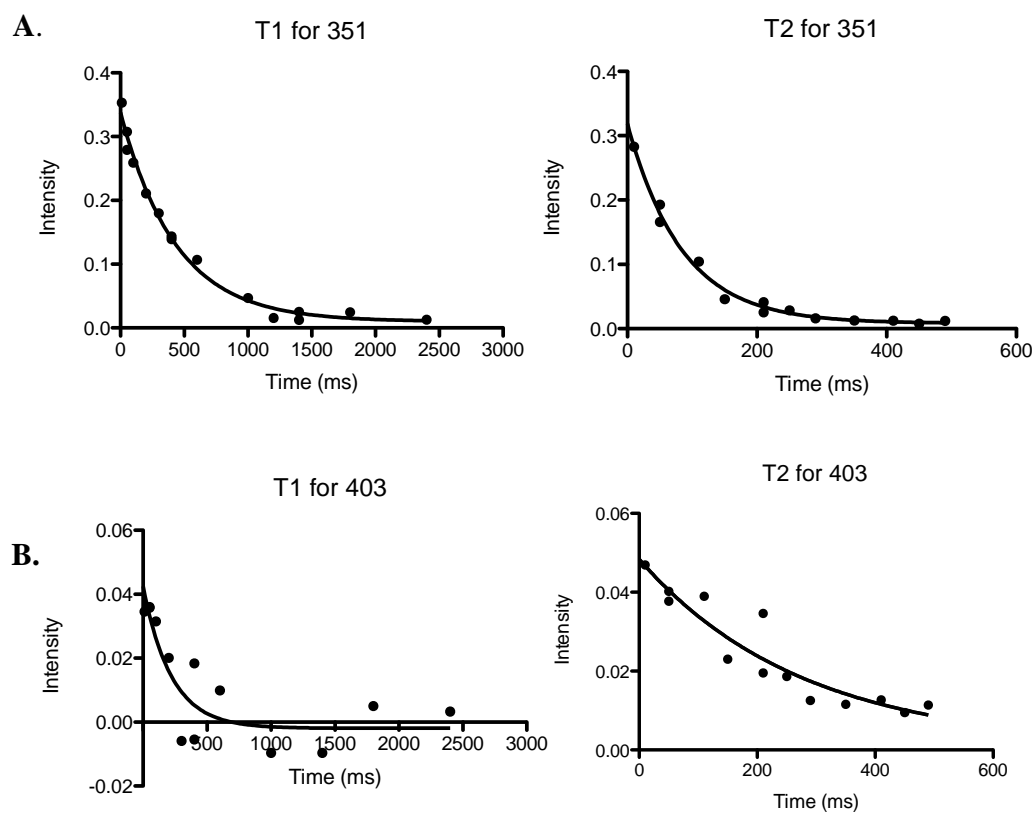


**Figure A-1. Backbone assignment of IBR-RING2.**

The strip plot of spectra shows <sup>15</sup>N planes for the residues near the start of IBR domain of IBR-RING2. For each of planes, the five panels are shown, first CCONH, then CBCA(CO)NH, HNCACB, HNCO, and HNCACO. The connection of *i*-1 residue chemical shifts with *i* residue chemical shifts are indicated with blue, purple, and dotted lines. CCONH was used as a guide to distinguish amino acids that have very close chemical shifts, with the additional information on C<sub>γ</sub>, C<sub>δ</sub>, and C<sub>ε</sub>.

**Appendix B: Data from Dynamics Studies of IBR-RING2**

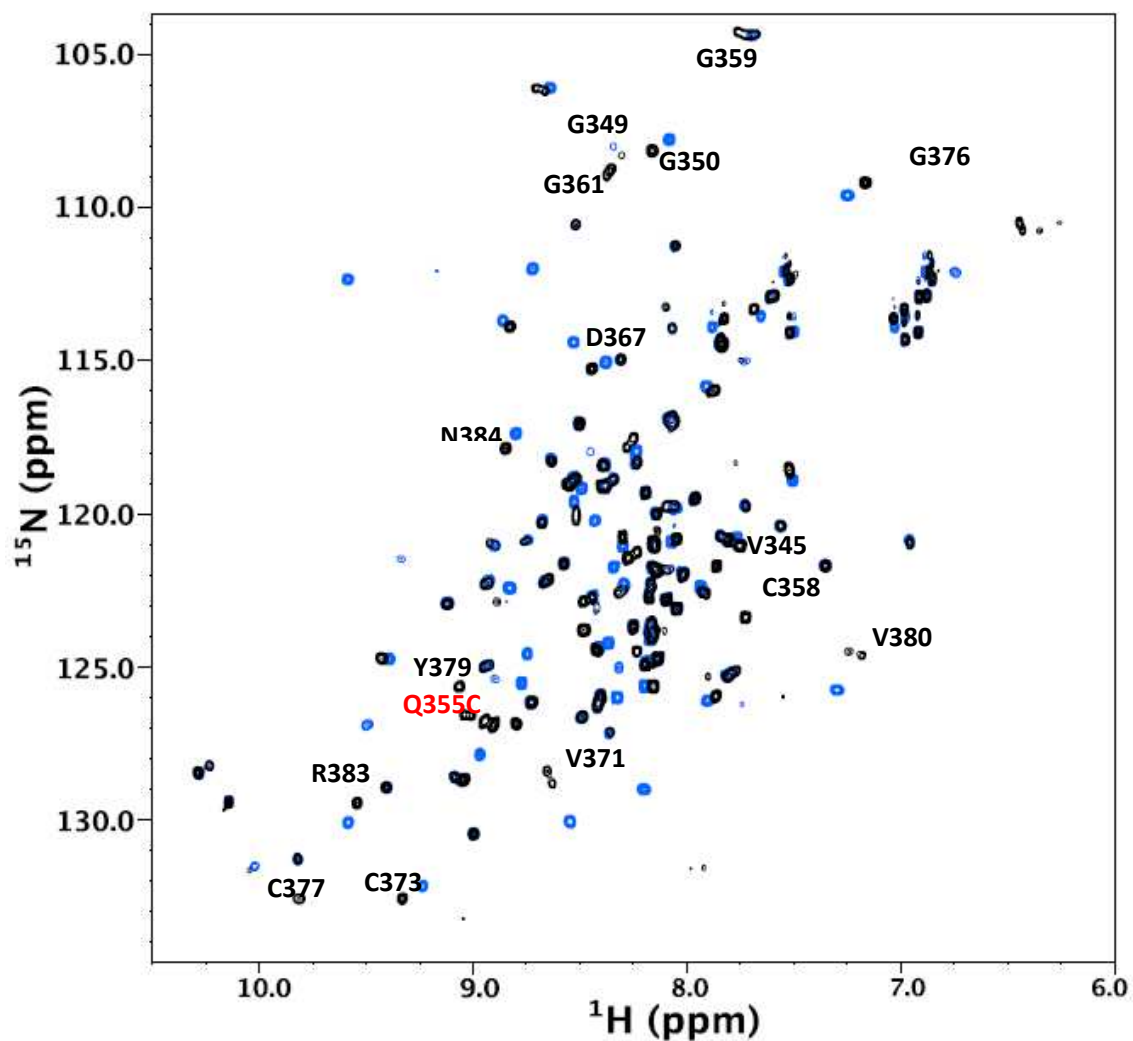
**Figure B-1. Original steady-state heteronuclear NOE values for backbone amides of  $^{15}\text{N}$ -labeled IBR-RING2.** Negative NOE values in the IBR-RING2 domain reflect increased flexibility with respect to the two other domains. Values above 1 were not deleted in this graph, as shown in Figure 3.12.



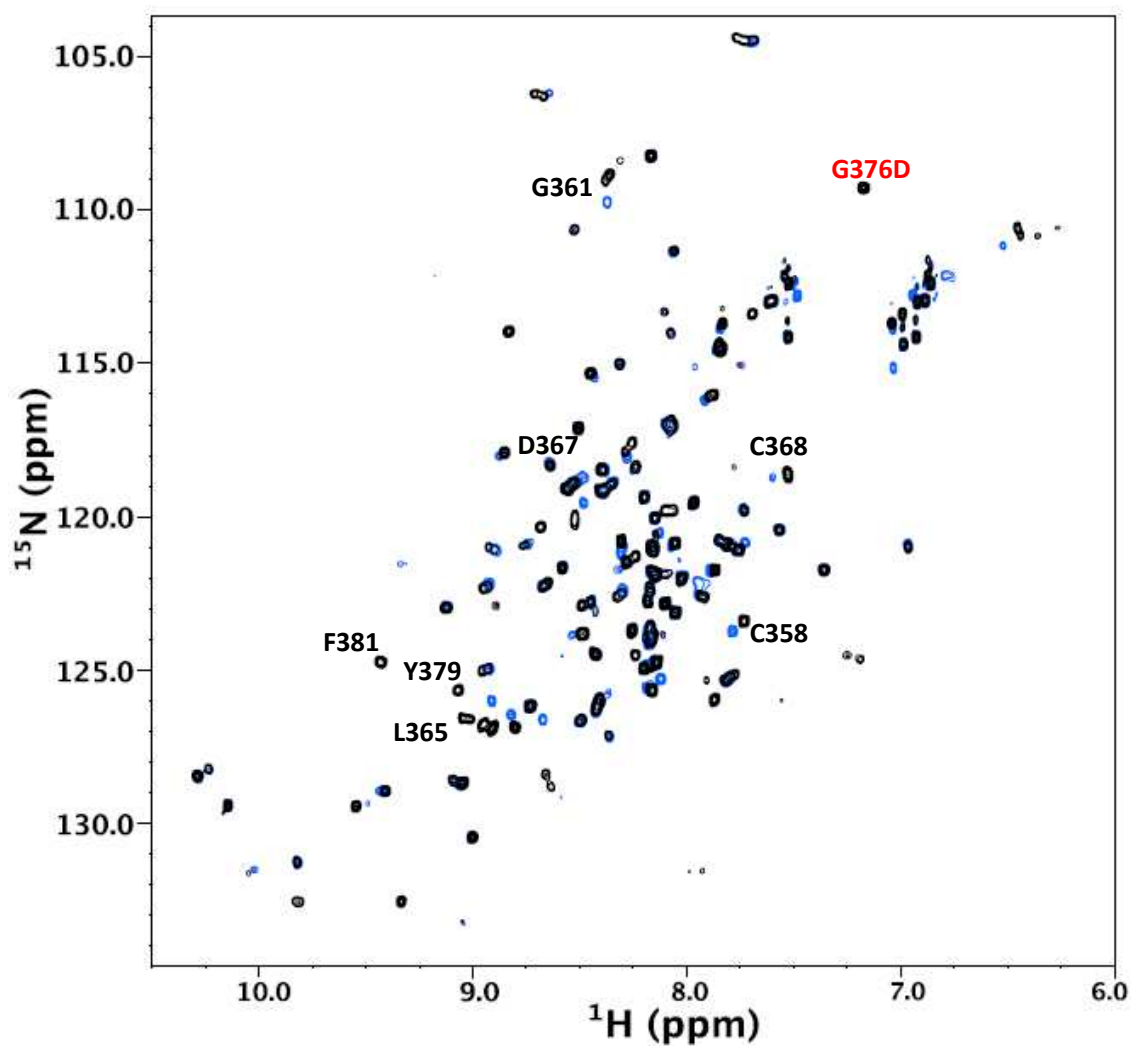
**Figure B-2. T1 and T2 graphs for two different residues on IBR-RING2. A.** T1 and T2 decay curves for V351, and data fits well to the decay curves. **B.** T1 and T2 decay curves for A403, and data does not fit well to the decay curves.



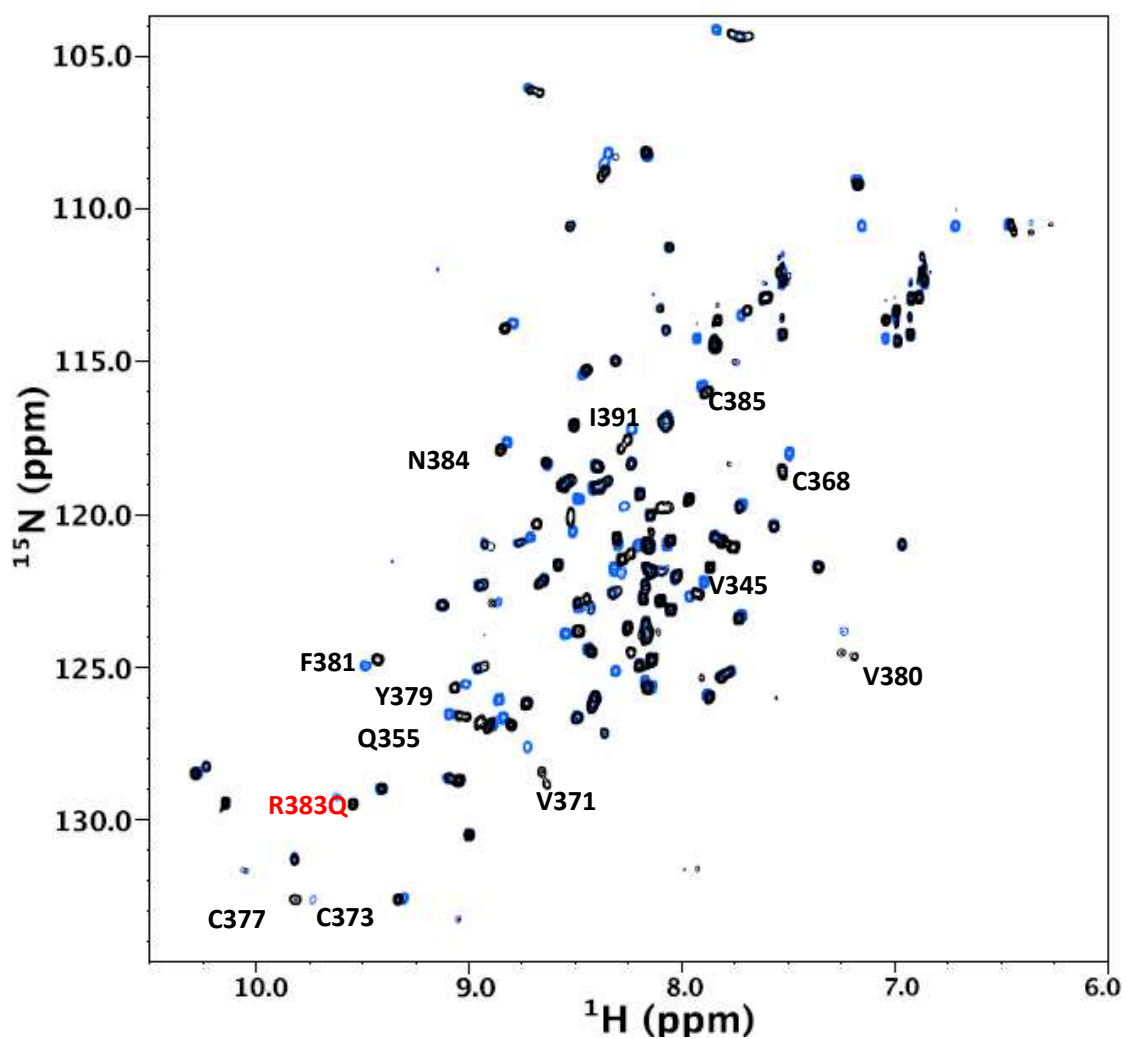
## Appendix C: Data from Mutational Analysis of IBR-RING2



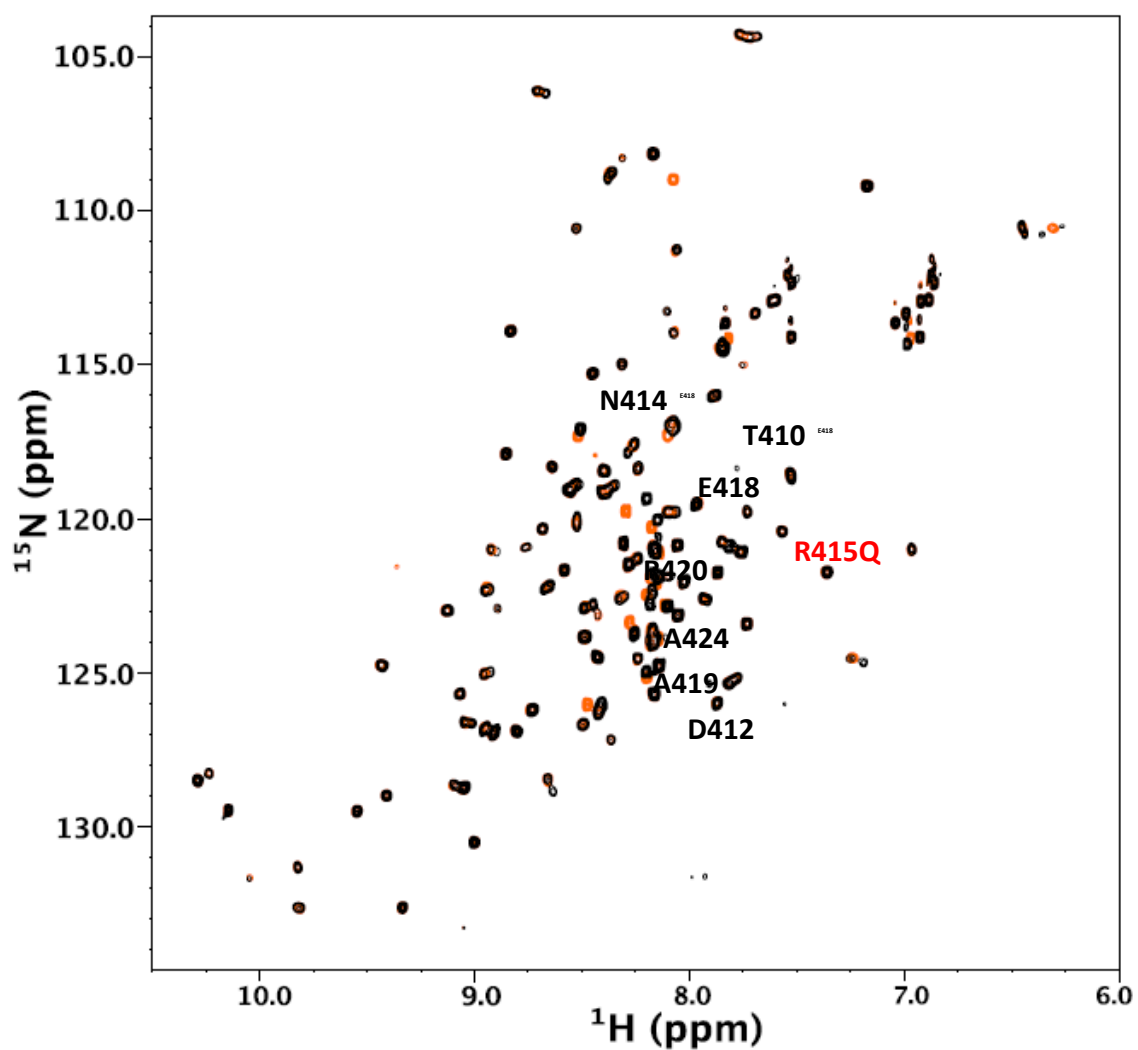
**Figure C-1. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in IBR region (Q355C).**  
The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and blue peaks belong to the spectrum of Q355C mutant of IBR-RING2.



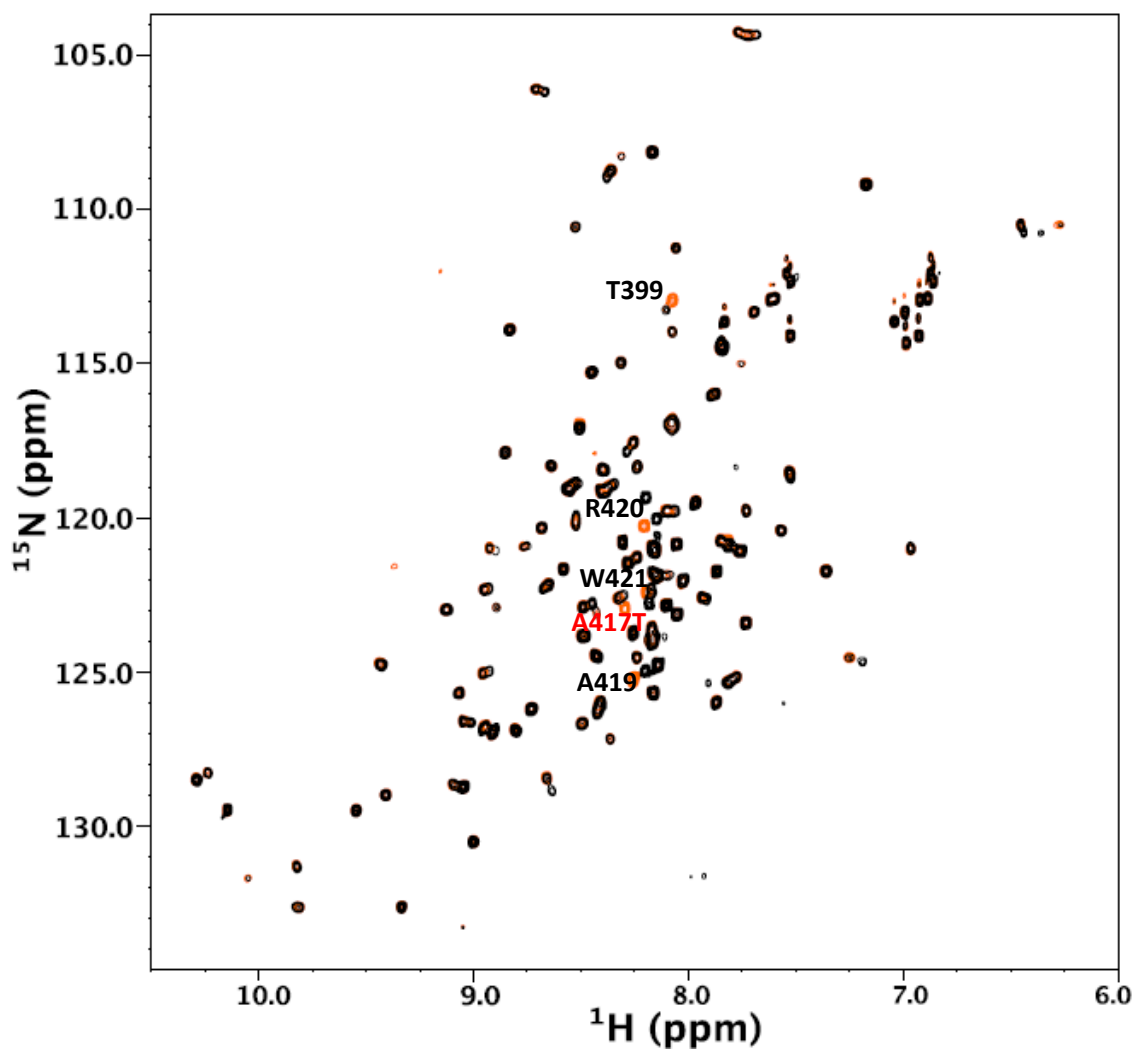
**Figure C-2. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in IBR region (G376D).**  
The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and blue peaks belong to the spectrum of G376D mutant of IBR-RING2.



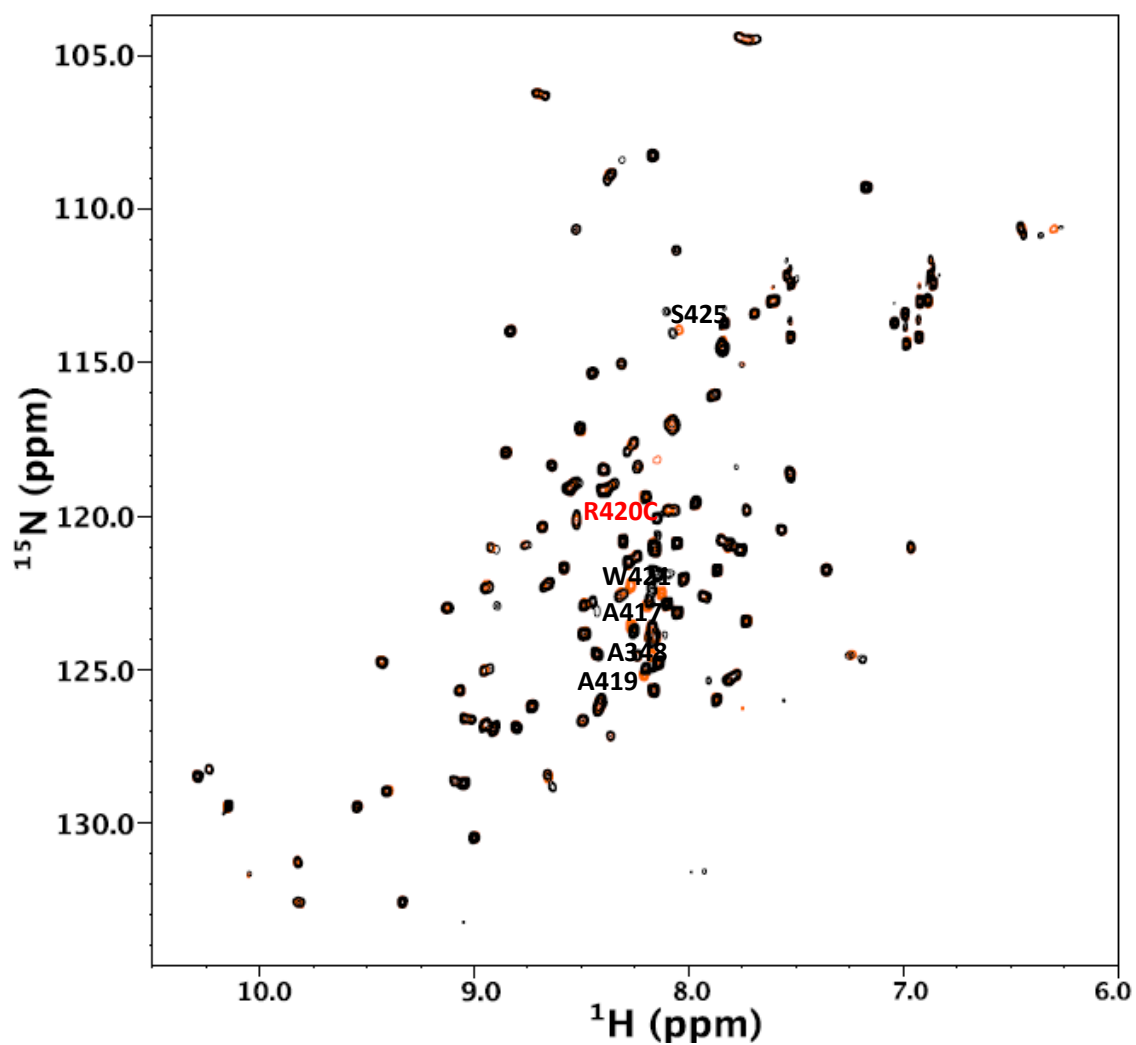
**Figure C-3. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in IBR region (R383Q).**  
The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and blue peaks belong to the spectrum of R383Q mutant of IBR-RING2.



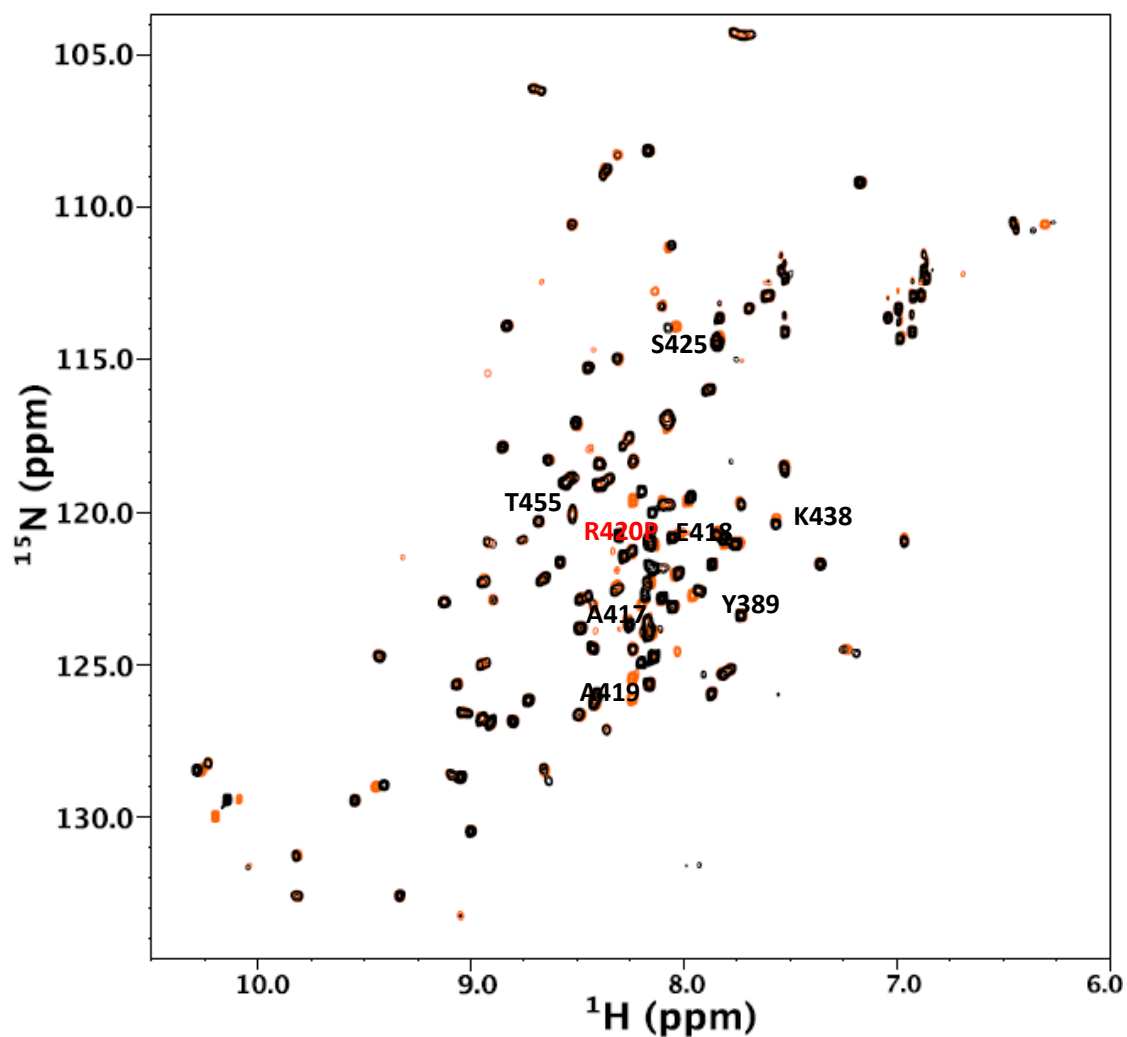
**Figure C-4. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the linker region (R415Q).**  
The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and orange peaks belong to the spectrum of R415Q mutant of IBR-RING2



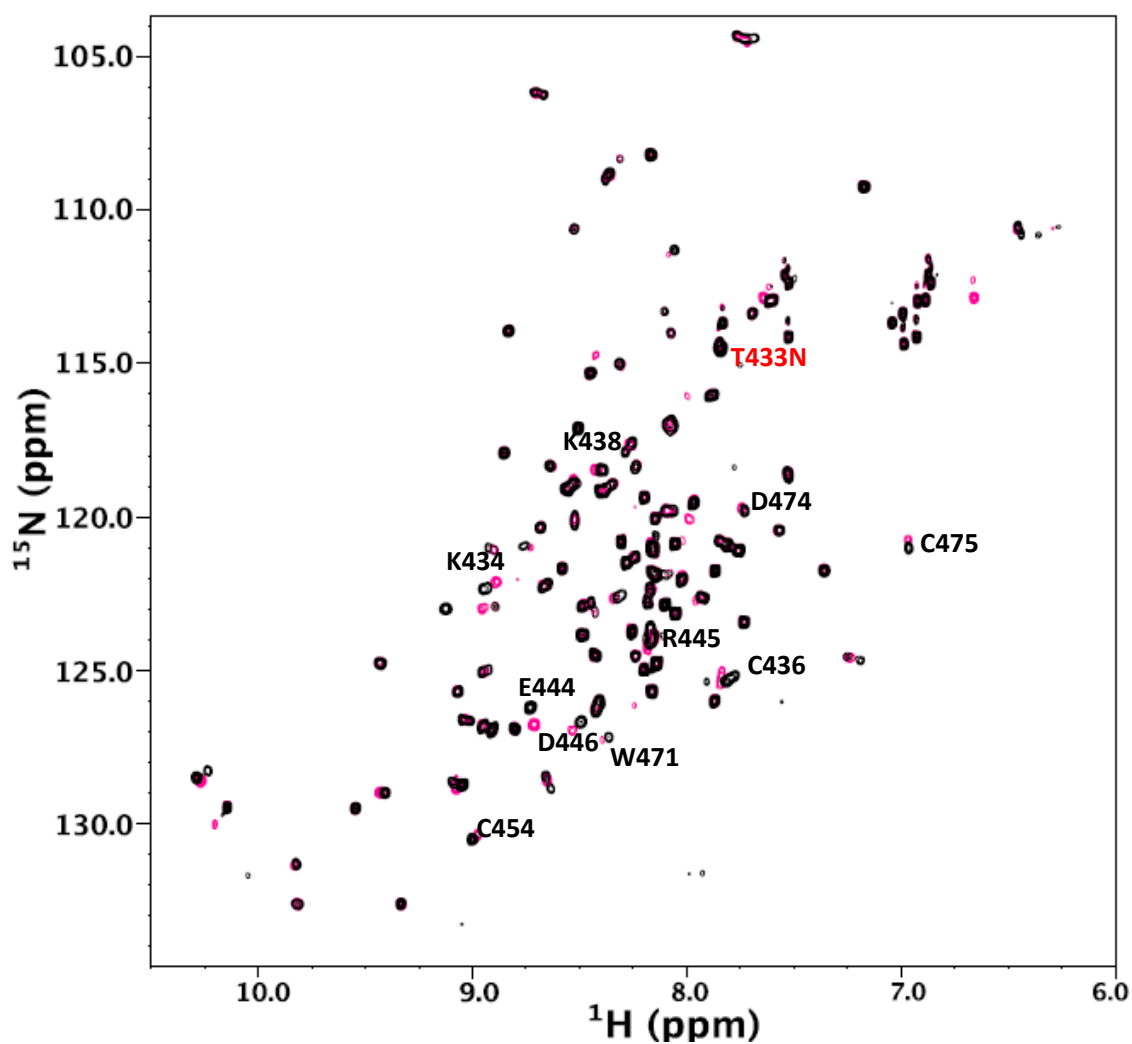
**Figure C-5. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the linker region (A417T).**  
The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and orange peaks belong to the spectrum of A417T mutant of IBR-RING2



**Figure C-6. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the linker region (R420C).**  
The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and orange peaks belong to the spectrum of R420C mutant of IBR-RING2

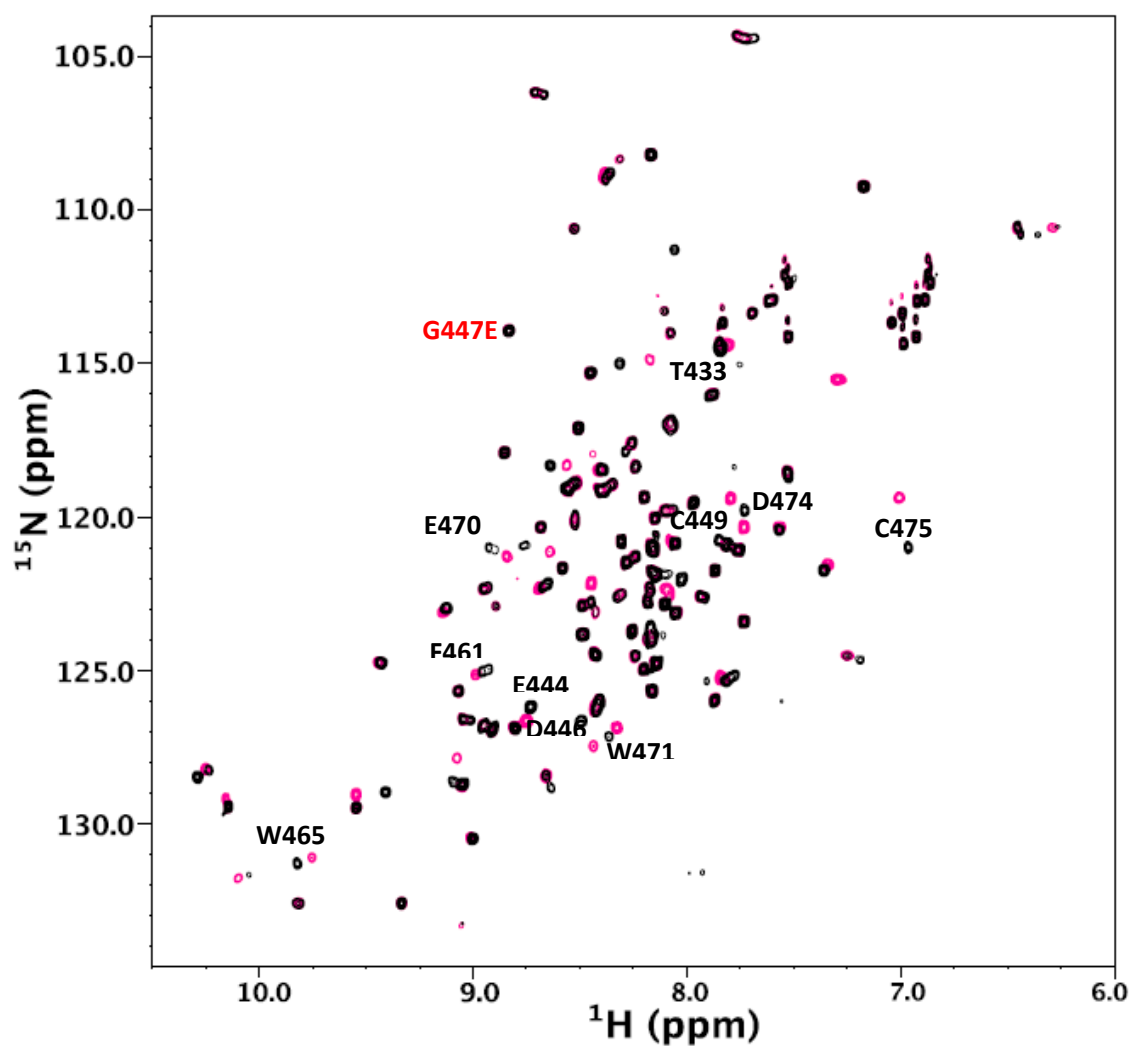


**Figure C-7. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the linker region (R420P).**  
The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and orange peaks belong to the spectrum of R420P mutant of IBR-RING2



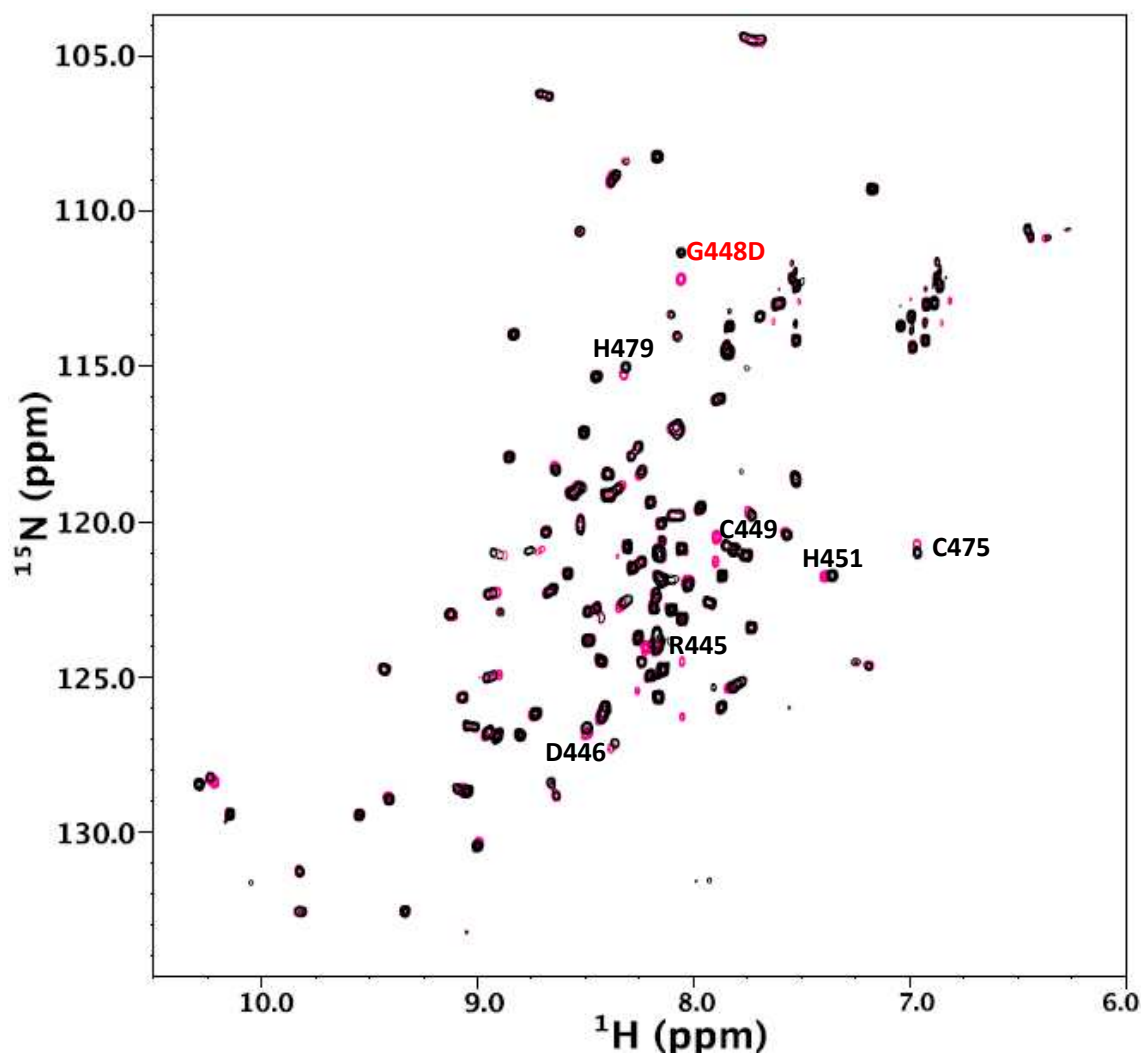
**Figure C-8. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the RING2 region (T433N).** The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and pink peaks belong to the spectrum of T433N mutant of IBR-RING2.





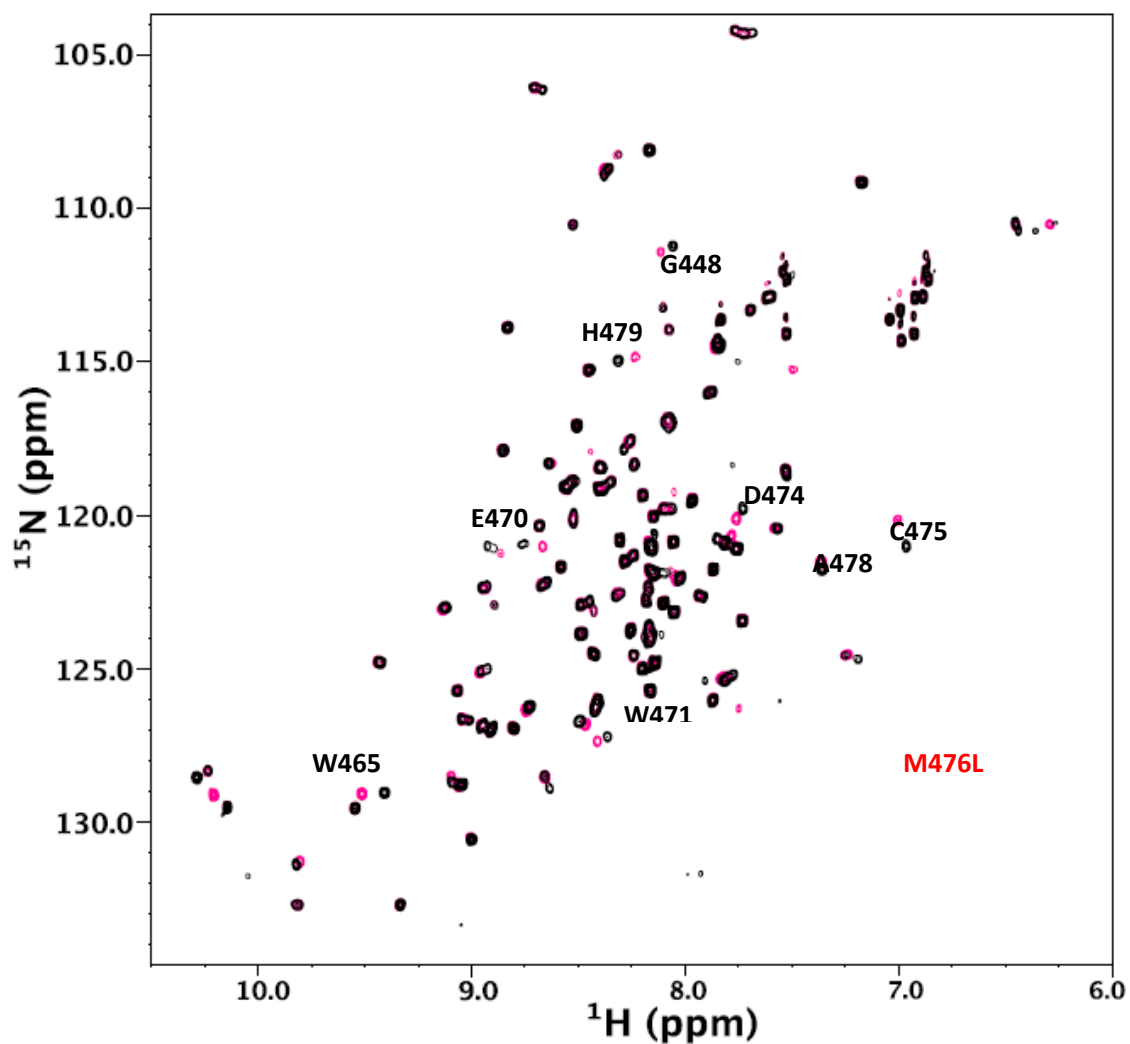
**Figure C-9. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the RING2 region (G477E).**

The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and pink peaks belong to the spectrum of G477E mutant of IBR-RING2.



**Figure C-10. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the RING2 region (G448D).**

The point mutation is labeled in red, and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and pink peaks belong to the spectrum of G448D mutant of IBR-RING2.



**Figure C-11. Superposition of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type IBR-RING2 with mutated IBR-RING2 in the RING2 region (M476L).**

The point mutation is indicated in red on the side (peak not found), and residues that are impacted (shifted) are labeled. Black peaks belong to the wild type IBR-RING2 and pink peaks belong to the spectrum of M476L mutant of IBR-RING2.

## Appendix D. Multiple Sequence Alignment of IBR-RING2

### A. IBR

```

Human   GGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYHEGEC SAVFE
Chimp   GGVLCPRPGCGAGLLPEPDQRKVTCEGGNGLGCGFAFCRECKEAYHEGEC SALFE
Rat     GGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGCGFVFCRDCKEAYHEGECDSMFE
Mouse   GGVLCPRPGCGAGLLPEQGQRKVTCEGGNGLGCGFVFCRDCKEAYHEGDCDSLLE
Horse   GGMLCPSPGCGAGLLPEPSRRKVTCEGGNSLGC GFVFCRDCKEAYHEGECNALLE
Pig     GGVLCPRPGCGAGLLPEPGQRKVTCEGGNSLGC GLVFCRDCKEAYHEGEC SALFE
Opossum GGVLCPSPGCGAGLLPGPEVRKITEP SNGLGCGFEFCRECKEAYHEGECNTLFE
Chicken GLLCPTPSCGAGLLPEPEVRKIVCEP GNGIGCGFVFCRECKEAYHEGECSSFLS
Zebrafish GGVLCPTPGCGAGLLPEPDLRRIVCEP GNGIGCGSVFCRECKEAYHEGECNSLLS
Fly(D.mela) GGVLCPPPGCGMGLLVPEPDCRKVTCQN- - - GCGYVFCRNC LQGYHIGEC LPEGT

```

### B. Linker

```

Human   ASGTTTQA - YRVDERAAEQARWEAASKETIKK
Chimp   ASGTTTQA - YRVDERAAEQARWEAASKETIKK
Rat     ASGAT SQA - YRVDRRAAEQARWEAASKETIKK
Mouse   P SGAT SQA - YRVDKRAAEQARWEAASKETIKK
Horse   ASGAVTQA - YRVDERAAEQARWEAASKETIKK
Pig     ASAAVAQA - YRVDQKAAEQARWEAASKETIRK
Opossum ASGAAAQA - FMVDEQAAERARWEAASKETIKK
Chicken TQGAVAQKGYVVDENAAMQARWEAASKETIKK
Zebrafish PPGAMAQKGYVVDEHAAMQARWEAASRETIKK
Fly(D.mela) GASATNSCEYTVDPNRAAEARWDEASNVTIKV

```

### C. RING2

```

Human   TTKPCPRCHVPVEKNGGCMHMKCPQPQCRLIEWCWC GCEWNRVCMGDHWF DV
Chimp   TTKPCPRCHVPVEKNGGCMHMKCPQPQCRLIEWCWC GCEWNRVCMGDHWF DV
Rat     TTKPCPRCNVPIEKNGGCMHMKCPQPQCKLEWCWC GCEWNRACMGDHWFDV
Mouse   TTKPCPRCNVPIEKNGGCMHMKCPQPQCKLEWCWC GCEWNRACMGDHWFDV
Horse   TTKPCPRCHVPVEKNGGCMHMKCPQSQCQLEWCWC GCEWNRACMGDHWFDV
Pig     TTKPCPRCHVPVEKNGGCMHMKCPQPQCQLEWCWC GWEWNRDCMGDHWFDV
Opossum TTKPCPRCHIPVEKNGGCMHMKCPQPQCKFEWCWC SLEWNRTCMGDHWFDV
Chicken TTKPCPNCHIPVEKNGGCMHMKCPRPQCRFEWCWC GLEWNRTCMGNHWFD-
Zebrafish TTKPCPNCHIPVEKNGGCMHMKCPHPQCRFEWCWC GLEWNRTCMGDHWFE-
Fly(D.mela) STKPCPKCRTPTERDGGCMHMVCTRAGCGFEWCWC QTEWTRDCMGAHWFG-

```

## Figure D-1. Multiple Sequence Alignment of IBR-RING2

Multiple sequence alignment (MSA) of A. IBR, B. Linker, and C. RING2 region using Jalview program. Original MSA of parkin performed by Steve Beasley (Dr. Shaw lab) was modified. All the conserved zinc coordinating amino acids are highlighted in yellow.

# Curriculum Vitae

## Yeong Ju Noh

---

### EDUCATION

<i>York University</i> <b>BSc in Chemistry</b>	<b>2007 - 2011</b>
Honours: Passed as a member of the Dean's Honour Role, "First Class with Distinction"	
<i>University of Western Ontario</i> <b>MSc in Biochemistry</b>	<b>2011-present</b>

---

### AWARDS

• York University Entrance Scholarship	\$ 12,000	<b>2007-2010</b>
• York University Science and Engineering Scholarship	\$ 2,000	<b>2007-2008</b>
• Queen Elizabeth II Aiming for the Top Scholarship	\$ 14,000	<b>2007-2011</b>
• NSERC Undergraduate Summer Research Award	\$ 6,000	<b>2009</b>
• NSERC Undergraduate Summer Research Award	\$ 6,000	<b>2010</b>
• York University James Wu Internship	\$ 12,000	<b>2010-2011</b>
• Western Graduate Research Scholarship - Biochemistry	\$ 15,000	<b>2011-2013</b>
• Ontario Graduate Scholarship Award	\$ 15,000	<b>2012-2013</b>

---

### VOLUNTEER AND TEACHING EXPERIENCE

<i>York University</i> <b>Volunteer</b> – “York University’s 2008 Science Olympics” Helped to supervise high school students and set-up/organize activities	<b>2008</b>
<b>Volunteer</b> – “Trillium Health Center at Mississauga, West Toronto” Helped nurses, patients and physiotherapists in rehabilitation, pediatrics, and surgical units	<b>2006-2009</b>
<b>Volunteer</b> – “York University Chem1000 tutor” Helped first year students with understanding concepts, problem solving, and writing lab reports	<b>2008-2009</b>
<i>University of Western Ontario</i> <b>Teaching Assistant</b> – for “UWO Biochemistry 3387G” Prepared and supervised weekly laboratory sessions and graded all laboratory reports	<b>2012</b>

---

**WORK EXPERIENCE**

---

- York University – NSERC USRA student*  
**Research Assistant** **2009-2010**  
Responsible for expression and purification of pillin proteins
- York University – James Wu Internship student*  
**Research Assistant** **2010-2011**  
Conducted the expression and purification of TrbB for the formation of protein crystal, and set up crystal screens
- York University–BCHM 4000 student* **2011**  
**Undergraduate Thesis**  
Performed research that was a continuation of my internship  
Obtained several crystals of protein, TrbB, which diffracted with 2.3Å

---

**POSTER PRESENTATION**

---

- “Structure of C-terminal Region of Parkin, IBR-RING2” at the University of Western Ontario 2012  
as part of the requirements for Master’s Program