Identification of candidate substrates of the leucine rich repeat kinase 2 by mass spectrometry-based phosphoproteomics

William Carl Edelman

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

Judit Villén, Chair

Leo Pallanck

Deborah Nickerson

Program Authorized to Offer Degree:

Department of Genome Sciences

University of Washington

Abstract

Identification of candidate substrates of the leucine rich repeat kinase 2 by mass spectrometry-based phosphoproteomics

William Carl Edelman

Chair of the Supervisory Committee:

Assistant Professor, Judit Villén

Department of Genome Sciences

Mutations in the kinase domain of the leucine rich repeat kinase (LRRK2) have been implicated in heritable forms of Parkinson's disease (PD). Specifically, a glycine to serine mutation (G2019S) has demonstrated hyperactive autophosphorylation, neuronal toxicity, and locomotor deficits in the fruit fly *Drosophila melanogaster*— all of which are related to its pathogenicity in PD. My dissertation focuses on identifying novel substrates of LRRK2 through analysis of proteome-wide changes in protein abundance as well as identifying changes in phosphorylation of proteins *in vitro* and in the *in vivo* fruit fly model. Using mass spectrometry, I provide quantitative information on thousands of proteins and phosphorylation sites. *In vitro* kinase assays on peptides derived from fly heads or a neuroblastoma cell line provide evidence for direct substrates of LRRK2, while the *in vivo* experiment in flies expressing LRRK2 identifies both direct and indirect phosphorylation substrates of the kinase. Herein, I present evidence for novel, LRRK2-mediated phosphorylation sites in the *Drosophila melanogaster* and the neuroblastoma models of PD. I also show changes in protein expression upon expression of human LRRK2 in the fly model.

Table of contents

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

1.1 Protein Kinases: what are they and why are they important?

Kinases are a biologically important group of enzymes that transfer a negatively charged phosphoryl (PO_3^{2-}) group to another protein— an enzymatic process known as phosphorylation. This process was first observed by Burnett *et al.* in 1954[1], and subsequently shown that phosphorylation can change the activity of a protein[2]. Kinases have continued to evolve, duplicate and expand in numbers throughout bifurcations in the tree of life and exist in all phylogenetic domains from archaea and bacteria to eukarya. Kinases and their substrates are considered to be the prime mediators of cell signaling—the mechanism through which cells sense their environment, send signals throughout a protein network and change behavior through alterations in gene expression and regulatory networks of kinase-substrate relationships. Understanding these relationships and their biological function is essential for understanding fundamental biological processes and understanding the dysfunction of these mechanisms that can lead to diseases, such as cancer, neurodegeneration, hypertension, autoimmune diseases and specific diseases like Parkinson's disease (PD) [3-5].

In humans, protein kinases are one of the largest and most well-studied family of genes. A typical eukaryotic kinase harbors 12 evolutionarily conserved subdomains including a eukaryotic-specific protein kinase (ePK) catalytic domain, which is critical for the phosphorylation reaction. Protein kinases all have a similar domain structure to carry out phosphorylation, which includes a bi-lobal structure, with $MgATP^{1-}$ binding in between. A region known as a P-loop (glycine rich) binds ATP while the A-loop (activation) binds the peptide, aligning the phosphate acceptor for transfer from ATP **(Figure 1.1)**. With more than 500 members in the human genome, kinases represent about 2% of all human genes. They consist of 9 canonical evolutionary gene families and one group of atypical kinases **(Table 1.1)** [7].

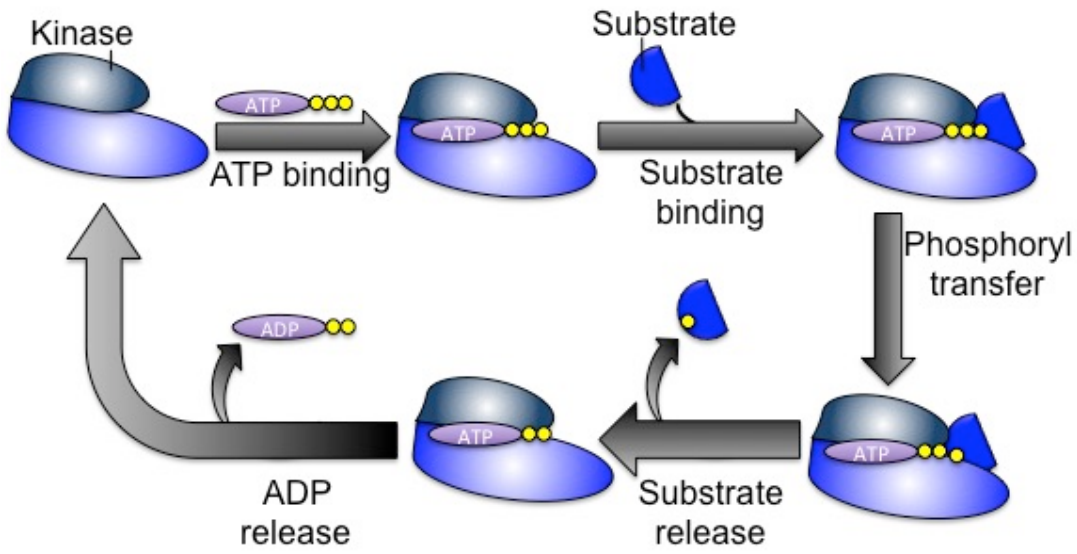


Figure 1.1 Kinase mechanism
The basic mechanism of kinase activity is shown. Concept from Ubersax *et al.* (2007)[6]. From top left to right, clockwise: kinases bind ATP within active site, then bind the substrate followed by the transfer of the γ-phosphate (yellow) to the substrate. The substrate is released from the complex followed by release of ADP. The kinase is left unchanged to carry out the kinase reaction again.

Table 1.1 Groups of kinases[7,8]

Group	Membership	Example families	General function or salient feature
AGC	63	PKA,PKG,AKT1/2/3	Core intracellular signaling
CAMK	74	CaMK1/2/4, MAPKAPK2/3/5, MLCK	Calmodulin/calcium signaling
CK1	12	Casein kinase1, tau tubulin kinase	High eukaryotic conservation
CMGC	61	CDK,MAPK,GSK3	Cell cycle, splicing
STE	47	MAPK, Ste7	MAPK cascade, homologous with yeast sterile kinases
TK (tyrosyl kinase)	90	EGFR, FGFR, InsulinR	Almost all tyrosine kinase specific activity

TKL (tyrosyl kinase-like)	43	MLK1-4, LISK, LIM	Most similar to TK, activity on serine and threonine
RCG (receptor guanylyl cyclase)	5	Retinal guanylyl cyclase 1, heat-stable enterotoxin receptor	Generates cGMP
Atypical	40	ATM, BCR, TRRAP	Lack sequence homology to ePK domain
Other	83	KSR, titin, CASK	Diverse group, catalytically inactive

Protein kinases are the primary mediators of signal transduction for a host of different cellular processes including the cell cycle, cell metabolism, transcription, cytoskeletal dynamics, cell motility, cell death, and cell-to-cell communication during development. On a larger scale, kinases are also critical to the maintenance of cellular homeostasis and proper function of the nervous system, immune system and many specific physiological processes in the human body. Further, even complex behaviors like memory and learning are being attributed to the activity of some protein kinases[9]. Thanks to the technological revolution surrounding DNA and genome sequencing, researchers have catalogued nearly all the protein kinases across all the major groups of life[10]. This level of interest speaks to the importance of kinases, their interactions and substrates— all of which can lead to the better understanding of each protein, and how they are connected to each other. Despite decades of efforts, specific substrates and the complete repertoire of signaling in each cell type across species is far from being understood. To address these in more detail will require novel high throughput technologies, model organism tissue, and/or cellular analysis to understand the full function of even pairs of proteins.

Because of their involvement in so many critical biological processes, kinases are considered to be the drivers of a number of human diseases; from metabolic disorders, like diabetes[11] to neurodegenerative diseases[12] and cancers[13], many kinases have taken center stage in the focus of biomedical research[14]. Some diseases arise from mutations, or genetic changes in the genome that impart a gain of function or overexpression of a particular kinase. For example, when a kinase exhibits a gain of function mutation, it is likely that the conformational change results in an active state, leading to deregulated over activity of the kinase. In particular, an alteration to one of the kinase subdomains may

lead to a change in function and, in turn, change in the behavior or survival of the cell. In the case of the leucine rich repeat kinase (LRRK2), a non-synonymous point mutation has been connected to the development of Parkinson's disease.

1.2 Approaches to identifying kinase-substrate relationships

Determining kinase-substrate relationships is important to understand the cellular functions and promiscuity of a kinase. Although the technology to establish kinase-substrate relationships has dramatically changed in the past 15 years, the basic biological questions remain when assaying kinases. Typically, these questions include: what are the kinase's substrates? which amino acids are specifically phosphorylated? what genetic/chemical or environmental change can modify the kinase activity? To answer these questions, the field of identifying kinase substrate relationships researchers leverage biochemical, genetic, and analytical chemistry technologies. Early approaches to identify kinase substrate relationships consist of an *in vitro* kinase assay. It began with the isolation of protein extract containing a kinase, and detecting its activity on fractions of protein isolate as a substrate. When Burnett and Kennedy (1954) [1] characterized kinase activity in an *in vitro* kinase experiment, protein phosphorylation was detected with [γ^{32}P]-ATP and radiation counting. With these bulk detection methods, it was possible to demonstrate phosphorylation from one isolate to another, but was limited to the highly abundant proteins and did not resolve site-specific mapping to the primary amino acid sequence. The early use of *in vitro* kinase assays with [γ^{32}P]-ATP employed the use of thin layer chromatography (TLC), one-dimensional or two-dimensional gel electrophoresis and a film-based (X-ray or photo stimulated film) imaging system. To roughly localize phosphorylation sites, many of these early studies required a second dimension TLC and imaging and/or Edman degradation to identify the site and the protein sequence. Although radiolabeling was a very sensitive detection method, it required the handling of potentially hazardous radioisotopes, and the experiments were time consuming, requiring further mutational analysis for confirmation.

Other methods utilize a chemical genetics approach. Analog sensitive kinase assays can use phosphoproteomics to compare cells an engineered kinase that only accepts a bulky ATP molecule to a wild-type protein. This is advantageous as it can help discern the target kinase activity from endogenous

or extraneous kinases. These non-target kinases are typically present *in vitro* when using whole protein extracts [15] or *in vivo* as was demonstrated by Chen *et al*. (2005) [16].

Newer molecular biotechnology methods and high throughput assays have propelled the discovery of possible kinase substrates forward. For example, genome-wide genetic screening through the use of siRNA libraries and phenotype assays in yeast and other model organisms are capable of establishing a connection between proteins in a high-throughput fashion[17] but it does not resolve whether the interaction is as a result of direct phosphorylation event or mediated by other proteins in a signaling pathway. Further, it is limited to model organisms such as yeast, worms and flies, for which genetic tools are available.

Peptide microarrays and phage display are also high-throughput methods that offer a readout of kinase activity on a large set of putative substrates but they do not establish the sites of phosphorylation unless they are adapted to analysis by mass-spectrometry based methods. Peptide arrays are useful in determining kinase specificity on the basis of activity on primary sequences and scanned by autoradiography, fluorescence[18] and immunoblotting[19]. Similarly, peptide and protein microarrays have been used for high-throughput screening of substrates for purified kinases on a global scale. Ptacek *et al*. (2005)[20] used proteome chip technology in the systematic analysis of 87 yeast kinases against the yeast proteome to develop a first generation phosphorylation map for yeast. Although it is possible to a create assays for even larger proteomes like the human proteome the surface based approaches have a few drawbacks including an undersampling of interactions that require an adapter protein to aid the interaction, non-native protein conformation due to binding the chip and non-localization of the phosphorylation sites

1.3 Mass spectrometry quantitative phosphoproteomics

Today, phosphoproteomic MS methods are in widespread use, and are sensitive analytical methods for the detection of peptide phosphorylation from nearly any experiment amenable to protein digestion and peptide purification. The combination of electrospray ionization, liquid chromatography tandem MS (ESI-LC-MS/MS) methods and phosphopeptide enrichment (described in detail in Chapter 2) has enabled the quantitative identification of tens of thousands of peptides and specific phosphorylation

sites in only 2-3 hours. A variety of experimental designs to test kinase-substrate phosphorylation use MS as the analytical tool of choice for analyzing the products of kinase-substrate experiments. For example, the co-purification of kinase-substrate pairs in affinity purification MS analysis can be used to detect phosphorylation on candidate substrates and test the hypothesis that these pairs represent kinase-substrate relationships. However, the limitation here is due to the transient nature of kinase-substrate interactions that may preclude their detection[21,22]. Alternatively, *in vitro* kinase assays on proteins or peptides, as in the classical methods, offer a direct interaction of kinase-substrate relationships. With advances in MS-based methods in terms of duty-cycle, quantification and precise localization of phosphorylated residues, *in vitro* methods are fast, direct and convenient. The *in vitro* phosphoproteomic approach can be combined with *in vivo* phosphoproteomic methods, as in the kinase assay linked phosphoproteomics method (KALIP), where endogenously phosphorylated peptides are extracted and treated with phosphatase prior to re-phosphorylating them with kinases of interest and comparing them to endogenous, or *in vivo,* phosphoproteomics of phosphorylation events modulated by kinase of interest[23]. KALIP leverages the overlap in both *in vitro* and *in vivo* assays to find kinase substrates. This offers high confidence in putative substrates, but requires adequate tools in both systems. In *vitro,* it requires a kinase amenable to activity outside of the context of a cell, where it may rely on secondary messenger or an unknown co-activator. *In vivo* methods require adequate expression of the kinase and appropriate cellular context. The biological system may need to produce a specific phenotype adequately assay kinase activity or may require a specific stimulus for analysis.

1.4 Dissertation goals

In this dissertation, it is my goal to convey the accessibility of modern phosphoproteomics methods and demonstrate its application towards the identification of candidate leucine-rich repeat kinase-2 substrates, a key protein in the understanding of Parkinson's disease (PD). In Chapter 2, I aim to introduce the accessibility of methods for studying protein phosphorylation. It was published as a chapter in a phosphoproteomics methods book as "A practical recipe to survey phosphoproteomes" [24]. In Chapter 3, I report the methods and results of *in vitro* and *in vivo* kinase assays in models used in PD research. The *in vitro* studies are the first to apply a chemical labeling strategy to identify LRRK2

substrates with MS phosphoproteomics methods. The *in vivo* work represents the first MS phosphoproteomic analysis of brain tissue from the LRRK2 Parkinsonian fly model. In Chapter 4, I close with a discussion of the major findings from investigating LRRK2 substrates, and evaluate them against the current body of knowledge for the PD kinase and provide perspective to guide future studies towards a more complete understanding of the role of LRRK2 in pathogenic phenotypes, particularly neurodegeneration including the accumulation of pathogenic protein aggregates so common to other age-related neurodegenerative diseases like Alzheimer's disease.

Chapter 2: A Practical Recipe to Survey Phosphoproteomes

Shotgun Proteomics, A practical recipe to survey phosphoproteomes, 1156, 2014, 389-405, Edelman, W.C., Haas, K.M., Hsu, J.I., Lawrence, R.T., & Villén, J.

2.1 Introduction

The reversible phosphorylation of serine, threonine, and tyrosine is a chief regulatory mechanism within the cell. Phosphorylation can rapidly alter a protein's structure, enzymatic activity, and interactions with other proteins, and thus plays the lead role in cellular signaling events. Phosphorylation-dependent signaling pathways are ubiquitous, extending from cell surface receptors and cytosolic kinases to thousands of substrates throughout every subcellular domain. Signaling pathways drive a myriad of cellular functions including cell motility, division, growth, metabolism, gene expression, and apoptosis. At present, more than 200,000 distinct phosphorylation sites have been reported, offering a vast regulatory landscape primed for further exploration[25]. In this chapter, we review mass spectrometry-based phosphoproteomics technology and its applications, and provide a protocol to comprehensively survey the phosphorylated species of a biological sample.

Phosphoproteomic applications are well suited to answer many fundamental questions in signal transduction research. At one extreme, phosphoproteomics can be used to power signaling network analyses. For example, Bodenmiller *et al .* systematically profiled the yeast kinome, determining the relationship between 97 kinases and 27 phosphatases and their substrates[26]. In another study, the phosphorylation patterns amongst a panel of nine mouse tissues were examined, revealing a diverse

signaling landscape across nearly 36,000 phosphorylation sites[27]. At the other extreme, phosphoproteomics is used to identify bona fide substrates of a particular protein kinase. This can be accomplished by genetically or pharmacologically manipulating kinases prior to analysis, as was the case with work on Cdk1[28], Polo and Aurora mitotic kinases[29], and mTOR[30,31]. Additional methods have also been developed to profile kinase activities by in vitro assays performed on complex peptide mixtures[23,32-34] or to globally ascertain phosphorylation site stoichiometry[35]. Finally, one of the most fascinating applications is the large-scale characterization of dynamic biological processes. Mass spectrometry-based phosphoproteomics was used to quantify the phosphorylation of more than 20,000 sites throughout mitosis[36] and, more recently, to dynamically quantify more than 37,000 distinct sites during the adipocyte insulin response[37]. These studies offer only a glimpse of the potential applications of phosphoproteomics. We anticipate that the diversity of uses will grow as the technique becomes a more routine feature of the biologist's toolbox.

As a discipline, phosphoproteomics began in earnest around a decade ago when phosphorylated peptides were first enriched using immobilized metal affinity chromatography (IMAC) followed by liquid chromatography-tandem mass spectrometry (LCMS/MS) analysis to identify 383 phosphorylation sites in yeast[38]. Since then, the technology has improved exponentially, combining optimized enrichment and fractionation strategies, superior mass spectrometers with higher sensitivity, speed and resolution, and statistically robust computational algorithms. It is now achievable for a non-expert with modest resources to profile more than 5,000 phosphorylation sites in a single experiment.

To transform phosphoproteomics into a mainstream technology, researchers had to overcome three main challenges. First, phosphorylated species spanned six to seven orders of magnitude in cellular protein abundance and varied in phosphorylation stoichiometry. Second, coverage of protein sequences was incomplete. Enzymatic digestion of proteins generates peptides of diverse lengths. Peptides that are too short cannot be uniquely mapped to a protein, and those that are too long or highly charged pose a challenge for mass spectrometry analysis. Third, MS/MS fragmentation of phosphopeptides was poor because it generates a phosphate neutral loss fragment dominant over sequence-informative b- and y-type ions.

The dynamic range and abundance problem has mostly been addressed by incorporating into the proteomic workflow an enrichment step that selects for phosphorylated species. This enrichment is typically performed at the peptide level and can target the entire phosphoproteome or just a specific subset. Enrichment methods using immunoprecipitation, chromatographic separation, or metal ion affinity reagents prior to mass spectrometry have been successfully used in phosphoproteomics **(Table 2.1)**.

Table 2.1 Phosphopeptide enrichment methods
Commonly used phosphopeptide enrichment methods. We indicate the recommended amount of peptide starting material, their potential bias as to which phosphopeptides are enriched, and other (non-phospho) peptides that might co-purify. To guide the selection of one method over another, we also indicate the throughput level, the relative cost per sample and the recommended applications

Enrichment method	Scale	Enrichment bias	Co-purified peptides	Throughput	Cost	Applications
Metal ion affinity (IMAC)	0.1-0.5 mg	Multiply phosphorylated	Acidic	High	$	Global phosphorylation
Immuno-affinity purification	5-50mg	Phosphorylaion motif	-	Medium	$$$	Phosphorylation motif
Acid-base interaction (TiO2, ZrO2)	0.1-0.5mg	Multiply phosphorylated	Acidic	High	$	Global phosphorylation
SCX chromatography	1-10mg	Peptides of low solution charge	Acetylated, protein C-terminal	Low	$$	Global phosphorylation
Precipitation with divalent cations (Ca2+, Ba2+)	0.2-1mg	Acidic phosphopeptides	Acidic	High	$	Global phosphorylation, compatible with detergents
SCX + IMAC/TiO2	1-10mg	-	-	Low	$$	Global phosphorylation at deep coverage
ERLIC	1-10mg	Multiply phosphorylated	Sialylated glycopeptides	Low	$$	Global phosphorylation

Peptide immunoprecipitation uses antibodies against a phosphorylation motif (e.g., phosphotyrosine) to isolate peptides containing the motif[39]. IMAC enrichment exploits phosphate's coordination to metal cations (e.g., Fe^{3+}, Ga^{3+}) to purify phosphorylated peptides from nonphosphorylated[38,40,41]. Similarly, enrichment with metal oxides (e.g., TiO_2, ZrO_2) relies on a Lewis acid–base interaction between the oxide and the phosphate group[42,43]. All of these methods start with a peptide complex mixture to obtain a phosphopeptide complex mixture. Alternatively, chromatographic techniques can simultaneously fractionate complex peptide mixtures and enrich for phosphopeptides. Strong cation-exchange (SCX) chromatography separates peptides on the basis of solution charge. At acidic pH, a phosphate group contributes to the peptide solution charge with a negative charge, facilitating the separation between phosphopeptides and their non-phosphorylated counterparts[44]. SCX chromatography has received considerable attention and is commonly combined with the enrichment methods exposed above for deep phosphoproteome coverage[45-48]. Electrostatic repulsion hydrophilic interaction chromatography (ERLIC, also called eHILIC and ion-pair normal phase) uses high organic concentration on an anion exchange

column to promote the hydrophilic interaction of phosphopeptides with the column while decreasing the usual electrostatic repulsion to the stationary phase[49].

Proteome coverage can be increased by performing separate digestions with different enzymes[50]. This strategy has been mainly applied to single-protein phosphorylation analysis, given the added workload and sample amount requirements.

To alleviate poor MS/MS fragmentation of phosphopeptides by collision-induced dissociation (CID), several mass spectrometry data acquisition methods have been developed (Table 2). Neutral loss-dependent MS3[44] and multistage activation (also known as pseudo-MS 3)[51] rely on extended CID fragmentation schemes on the phosphate loss product ion to obtain sequence- informative fragment ions. However, with increased sensitivity of current ion trap mass analyzers, the neutral loss peak dominance is not such a burden and may no longer necessitate MS3 approaches, which, in fact, are slower and less sensitive than classic CID MS/MS fragmentation[52]. Alternatively, electron transfer dissociation (ETD)[53] induces chemical fragmentation of the peptide sequence, generating c- and z-type ions without affecting amino acid side chains, and has thus been successfully applied to the analysis of peptides with labile modifications, including phosphorylation[54]. Higher energy collisional dissociation (HCD) is yet another fragmentation method that has been applied to phosphopeptide analysis[55,56]. HCD produces b- and y-type ions, and fragment analysis is performed at high mass precision on a mass analyzer that permits the detection of low m/ z fragments. Consequently, HCD is one of the best fragmentation methods for unambiguously localizing phosphorylation sites, in particular near peptide termini. Overall, any of these fragmentation methods performed on today's mass spectrometers is suitable for phosphorylation analysis. However, given its high speed and sensitivity, CID seems the best choice for analyzing highly complex phosphopeptide mixtures, despite the neutral loss problem.

Table 2.2 MS/MS fragmentation methods used in phosphoproteomics
Mass spectrometry acquisition strategies and fragmentation methods used in phosphoproteomics. A list of advantages and disadvantages for each method is provided, as well as their recommended applications. Some of these methods may be combined on some instrument configurations

MS fragmentation method	Advantages	Disadvantages	Best used
CID MS/MS	• High sensitivity • Fast scan times	• Intense neutral loss product	With large-scale proteomics experiments
ETD MS/MS	• High peptide sequence coverage • No neutral phosphate loss • Precise site localization	• Limited availability on instrumentation • Poor fragmentation of peptides with z=2	With highly charges peptides
HCD MS/MS	• Low noise spectra • High mass precision • Covers low m/z fragment ions • Precise site localization	• Lower sensitivity than CID MS/MS • Slower scan times	With large-scale proteomics experiments
MS^3 and pseudo MS^3	• High sensitivity • Resolves problematic neutral phosphate loss	• Lower sensitivity than CID MS/MS • Slower scan times	With phosphopeptide samples of lower complexity (e.g., single-protein analysis)
Precursor ion scanning	• High sensitivity • Sequencing only when a phosphopeptide is detected	• Requires switching between negative and positive modes • Slow	For determining whether a peptide is phosphorylated

Apart from the automation of sample preparation and LC-MS/ MS analysis, phosphoproteomics has achieved a high level of routine automation in the realm of spectral data interpretation. Adequate interpretation of phosphopeptide spectra requires the correct localization of a peptide phosphorylation site within the peptide sequence. To localize a phosphorylation site, the presence of site-determining ions in the fragmentation spectrum is critical. Several algorithms (e.g., Ascore[57], PTM score[46], phosphoRS[58], and LuciPHOr[59]) have been developed to provide a confidence metric of correct site localization, all revolving around the idea of scoring matches to site-determining ions.

Paramount to understanding cellular signaling is the ability to compare the phosphoproteomes between two or more biological conditions. Phosphoproteomics via LC-MS/MS enables quantitative comparisons between controls and genetic mutants, drug treatments, and disease states. Quantitative proteomic methods have been reviewed elsewhere within this book.

In the last 10 years, proteomics researchers devoted much effort to developing technology to analyze phosphoproteomes, from methods for phosphopeptide enrichment to LC-MS/MS data acquisition, to automated data interpretation and statistics. Currently, robust and automated phosphoproteomics workflows are available, setting a prime time for global cellular signaling research. It is certain that phosphoproteomics research within the next years will make important contributions in

deciphering kinase–substrate relationships, assembling signaling networks, and overall expanding the protein phosphorylation knowledge base.

The protocol presented is a starting point for phosphoproteomics experiments using enrichment of phosphopeptides via IMAC and analysis by LC-MS/MS. On a practical note, this protocol is robust and portable and can be combined with a wide variety of experimental pipelines and sample types.

2.2 MATERIALS

Solid-Phase Extraction for Peptide Desalting on a Cartridge

1. C18 solid-phase extraction cartridges.
2. Vacuum manifold.
3. pH paper strips.
4. Sample acidification solution: 10 % trifluoroacetic acid (TFA).
5. Column wash solution: 100 % acetonitrile (ACN).
6. Column equilibration and sample wash solution: 0.1 % TFA.
7. Sample wash solution: 0.5 % acetic acid (AA).
8. Column wash and sample elution solution: 50 % ACN and 0.5 % AA.
9. Elution tubes: 2 mL microcentrifuge tubes or 15 mL conical tubes, depending on elution volume.
10. Vacuum concentrator equipped with cold trap and vacuum pump.

Solid-Phase Extraction for Peptide Microscale Desalting on a Stage Tip

1. Stage tip assembly components: Empore C18 polymer disks (3M), plastic pipette tips (200 µL size), blunt-end cannula with any rod that snugly fits into the cannula shaft.
2. Liquid passing mechanism: Syringe or centrifuge.
3. Stage tip wash solution: 100 % methanol (MeOH).
4. Stage tip wash solution: 100 % ACN.
5. Stage tip wash and sample elution solution: 70 % ACN and 1 % formic acid (FA).
6. Stage tip equilibration and sample wash solution: 1 % FA.
7. Sample resuspension solution: 10 % FA, or 1 % FA (see Subheading 3.2, step 4).
8. Vacuum concentrator equipped with cold trap and vacuum pump.

IMAC

1. Ni-NTA magnetic agarose beads (5 % suspension, Qiagen).

2. Metal stripping solution: 40 mM EDTA pH 8.

3. Metal loading solution: 100 mM FeCl 3.

4. Wash solution: Deionized water.

5. Wash and sample resuspension solution: 80 % ACN, 0.1 % TFA.

6. Phosphopeptide elution solution: 70 % ACN, 1 % NH 4OH, 29 % H 2O.

7. Acidifying solution: 10 % FA.

8. Shaker.

9. Magnet bar.

10. 0.2 mL microcentrifuge tubes.

11. Stage tip assembly components: Empore C18 polymer disks, plastic pipette tips (200 µL size), blunt-end cannula with any rod that snugly fits into the cannula shaft.

12. Liquid passing mechanism: Syringe or centrifuge.

13. Stage tip wash solution: 100 % MeOH.

14. Stage tip wash solution: 100 % ACN.

15. Vacuum concentrator equipped with cold trap and vacuum pump.

LC-MS/MS

1. Nano-liquid chromatography system consisting of autosampler and pumps.

2. Analytical column: Capillary column fabricated from fused silica capillary tubing (50–100 µm inner diameter) 15–40 cm in length and packed with 3 µm C18 reversed-phase material (see **Note 1**).

3. Pre-column: Capillary pre-column of the same inner diameter and packing material as the analytical column, 1–2 cm in length. Phosphoproteomics Using IMAC 396

4. Mass spectrometer: LTQ Orbitrap Velos (Thermo Fisher Scientific).

5. LC solvents: 0.1 % FA in water and 0.1 % FA in ACN.

6. LC-MS/MS sample resuspension solution: 3 % ACN, 4 % FA.

Methods

The IMAC protocol has been adapted from Ficarro *et al*. [60], and it is designed for the enrichment of phosphopeptides from digested whole-cell protein extract. In our hands, IMAC enrichment alone can yield up to 8,000 unique phosphopeptide identifications. Optionally, to obtain deeper phosphoproteome coverage, we suggest including SCX fractionation in the experimental workflow.

To prepare the peptide sample prior to IMAC, we recommend supplementing the lysis buffer with protease inhibitors (e.g., leupeptin and aprotinin, or a protease inhibitor cocktail) to prevent protein degradation, and phosphatase inhibitors (e.g., sodium β-glycerophosphate, sodium pyrophosphate, sodium orthovanadate, calyculin A, and/or okadaic acid) to prevent protein dephosphorylation. We also recommend avoiding the use of detergents or eliminating them prior to digestion. Enzymatic digestion is typically conducted with trypsin or Lys-C. While sufficient peptide data can be collected from samples digested with one enzyme, we have found that dividing the protein sample for use with two different enzymes increases the number of identified phosphorylation sites. We are providing two equivalent methods for sample cleanup via solid-phase extraction, depending on the scale of the experiment. We recommend using solid-phase extraction cartridges (Subheading 3.1) to desalt peptide sample amounts between 40 µg and 1 g and stage tip desalting (Subheading 3.2) for sample amounts less than 40 µg. An overview of the method described in this chapter is provided in **Fig. 1**.

Solid-Phase Extraction for Peptide Desalting on a Cartridge

1. Acidify the peptide sample to pH 2 by adding 10 % TFA to achieve a final concentration of 0.1–0.5 %. Verify pH using a pH strip (see **Note 2**).

2. To desalt the peptide sample via C18 solid-phase extraction cartridge, choose a cartridge size with packing material about 20 times the total mass of the peptides (see **Note 3**). Each peptide sample should receive its own cartridge, unless combining chemically labeled samples. Insert the cartridges into the knob slots of the vacuum manifold for use.

3. Wash the cartridge C18 polymer: Fill the cartridges to the top (approximately five column volumes) with the designated solution, open the knob valves, and run the liquid through the cartridge, repeating as indicated (see **Note 4).** Use this technique for the following sequence of washes: three washes with 100 % ACN and one wash with 50 % ACN and 0.5 % AA. Washing can be done at a fast flow-through speed.

4. Equilibrate the C18 cartridge to peptide sample conditions by passing through three cartridge volumes of 0.1 % TFA. Equilibration can also be done at fast flow-through speed.

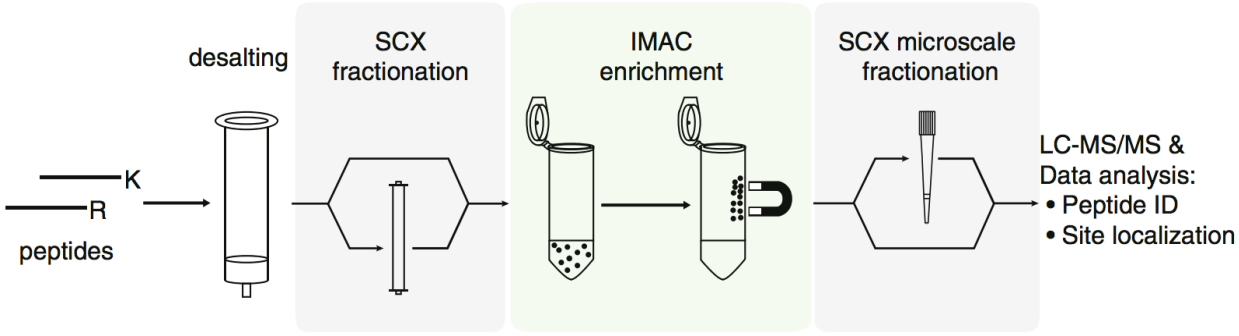


Figure 2.1 Experimental procedure for phosphopeptide enrichment and analysis by mass spectrometry.
Prior to enrichment, peptides derived from enzymatic digestion must be desalted on a C18 solid-phase extraction cartridge or a C18 stage tip (see Subheadings 3.1 and 3.2). Phosphopeptide enrichment is performed by immobilized metal affinity chromatography (IMAC), using a magnetic Fe 3+ -coated solid support (shown on the green rectangle) (Subheadings 3.3 and 3.4). Phosphopeptides are analyzed by LC-MS/MS to identify their sequence and their site/s of phosphorylation (Subheadings 3.5 and 3.6). IMAC enrichment alone can yield up to 8,000 unique phosphopeptide identifications. To achieve deeper phosphoproteome coverage, it is advisable to incorporate SCX fractionation into the workflow. This optional SCX chromatography step (shown on grey rectangles) provides peptide separation and partial phosphopeptide enrichment. SCX fractionation can be performed before IMAC on a column format or after IMAC on a stage tip format, depending on starting sample amounts and proteome complexity (see **Note 13)**.

5. Load samples into the designated cartridges using the above technique. Adjust the vacuum manifold pressure to a slow flow-through speed to ensure efficient peptide binding to the C18 polymer (see **Note 5)**.

6. Wash the peptides bound to the polymer three times with 0.1 % TFA and once with 0.5 % AA. Washing can be done at a fast flow-through speed.

7. Use the above technique to elute the peptides from the polymer at slow flow-through speed with five column volumes of 50 % ACN and 0.5 % AA, placing a fresh microcentrifuge tube or conical tube to collect the peptide eluate from this step.

8. Reserve a 20 µg aliquot of each desalted peptide sample in a vial insert for analysis by mass spectrometry to verify sample integrity. Divide the remaining peptide samples into 1 mg aliquots in fresh microcentrifuge tubes for later use in IMAC.

9. Dry all aliquots to completion in the vacuum concentrator. Resuspend the 20 µg aliquot in 20 µL of 3 % ACN and 4 % FA and analyze by LC-MS/MS (see Subheading 3.5). Store dried 1 mg aliquots at −20 °C.

Solid-Phase Extraction for Peptide Microscale Desalting on a Stage Tip

1. Prepare stage tips for each peptide sample with Empore C18 disks and 200 µL pipette tips. Cut out four small circular wedges of C18 polymer using the blunt-end cannula. Push the polymer into the pipette tip using the rod (see **Note 6**).

2. Wash the C18 polymer according to the following sequence, using a syringe or a centrifuge to pass the liquid through (see **Note 7**): two times with 20 all of 100 % MeOH, letting the MeOH sit for 5 min after the first addition; once with 20 µL of 100 % ACN; and once with 20 µL of 70 % ACN and 1 % FA.

3. Equilibrate the polymer by passing 20 µL of 1 % FA two times through the column.

4. Prepare the sample for stage tip desalting. If the sample is in solution containing ACN, add 10 µL of 10 % FA and dry the sample with the vacuum concentrator for about 10 min to evaporate the ACN. When the sample is dried down, resuspend in 20 µL of 1 % FA.

5. Load the peptide sample on the stage tip micro-column by passing the peptide solution through the polymer (see **Note 8**).

6. Wash the sample by passing 20 µL of 1 % FA twice through the column.

7. Elute the sample into an MS insert by passing 20 µL of 70 % ACN and 1 % FA.

8. Dry the sample to completion in the vacuum concentrator. Store at −20 °C.

Preparation of Fe 3+ Magnetic Beads for IMAC Phosphopeptide Enrichment

1. Use 1 mL of Qiagen Ni-NTA magnetic agarose beads (5 % suspension).

2. Wash beads three times with 1 mL of deionized water turning the test tube several times over between washes. To remove supernatant pull beads to the side of the tube with the help of a magnet bar and aspirate liquid with a pipette.

3. Incubate the beads for 30 min in 1 mL 40 mM EDTA pH 8.

4. Repeat the three washes with 1 mL of deionized water.

5. Load the beads with Fe 3+ by incubating them with 1 mL of 100 mM FeCl 3 for 30 min.

6. Wash the beads twice with 1 mL of deionized water.

7. Wash the beads three times with 1 mL 80 % ACN and 0.1 % TFA (see **Note 9**).

8. Resuspend the beads in 1 mL 80 % ACN and 0.1 % TFA to make a 5 % suspension.

IMAC Phosphopeptide Enrichment with Fe 3+ Magnetic Beads

1. Transfer the Fe 3+-loaded bead slurry to microcentrifuge tubes. Use 1 μL of the 5 % bead suspension per 10 μg of peptides (see **Note 10**). Pipette out the liquid from the slurry to isolate the beads.

2. Dissolve the lyophilized peptide sample in 80 % ACN and 0.1 % TFA at 1 μg/μL, and transfer the solution to the microcentrifuge tubes containing the beads. Incubate the peptides with the beads on a shaker for 30 min.

3. During the incubation time, construct C18 stage tips, which will later be used to filter out the magnetic beads during phosphopeptide elution (see **Note 11**). To construct a stage tip, press the blunt-end cannula into a two-layer sheet of Empore C18 and push the excised disk into the end of a 200 μL pipette tip until a snug fi t is achieved. Wash the C18 disk by passing through 20 μL of MeOH, followed by 20 μL of 100 % ACN. Add 60 μL of 10 % FA to a 2 mL microcentrifuge tube, and suspend the stage tip in the tube using a centrifuge adapter.

4. After the 30-min bead incubation, wash the phosphopeptide bound beads three times with three volumes of 80 % ACN and 0.1 % TFA to remove any non-phosphorylated peptides that are not bound to the beads.

5. Elute the phosphopeptides from the beads by adding three volumes of 70 % ACN, 1 % NH 4OH, and 29 % water. Incubate for 2 min, and transfer the eluent containing phosphopeptides to stage tips for filtering (see **Note 12**). To filter off any remaining beads, push the liquid through the stage tip into the microcentrifuge tube containing 10 % FA using a syringe or a centrifuge.

6. Lyophilize the phosphopeptide elution with the vacuum concentrator.

7. To desalt by stage tip, see Subheading 3.2.

As an optional procedure, SCX chromatography can be conducted before or after IMAC to obtain deeper coverage of the phosphoproteome. Fractionation is recommended for highly complex peptide samples, such as those derived from mammalian cell proteomes. SCX chromatography separates peptides based on their charge state[44], and, because of phosphate's negative charge, it provides partial phosphopeptide enrichment (see **Note 13**).

LC-MS/MS

The LC-MS/MS method will depend on the instrument model and configuration. Here, we provide an example of a 120-min data-dependent acquisition method using the Easy nanoLC (Proxeon) nano-liquid chromatography system coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo).

1. High-performance liquid chromatography (HPLC) method: LC gradient should be set to yield maximum phosphopeptide separation. For tryptic digest phosphopeptide samples, use a gradient of 10–30 % ACN in 0.1 % FA delivered at ~250 nl/min over 100 min. End the gradient with a high percentage of ACN (≥90 % for 10 min) to wash the column of remaining peptides and then a low percentage of ACN (5 % for 10 min) to re-equilibrate the column for subsequent runs (see **Note 14**).
2. Mass spectrometry acquisition method: For each cycle, schedule one full MS scan in the Orbitrap followed by 20 MS/MS scans after CID fragmentation (see **Note 15**). These scans target the 20 most intense ions present in the MS scan. Use dynamic exclusion to prevent the selection of precursors that have been fragmented on the preceding 30 s (see **Note 16**). Exclude precursors of charge 1+ or unknown charge.
3. Load sample onto column and run LC-MS/MS method: Resuspend the dried phosphopeptide sample in 3 % ACN and 4 % FA at 1 μg/μL, and load 1 μg of sample onto the analytical column for LC-MS/MS analysis following the acquisition method described above.

Data Analysis: Peptide Identification and Site Localization

Peptide spectra can be analyzed using a variety of database search algorithms to identify matching peptide sequences (e.g., Sequest[61], Mascot[62], X!Tandem[63], OMSSA[64]).

1. Regardless of the search algorithms used, there are key parameter specifications to include as follows: the digestive enzyme used, constant modifications (cysteine carboxymethylation), variable peptide modifications (methionine oxidation and serine/threonine/tyrosine phosphorylation), and, in quantitative experiments, isotope label modifications. The complete, sequenced protein database for the organism of study should also be supplied when performing the search. These can be accessed on a variety of servers; one commonly used is the UniProt protein knowledge base.

2. Searching against concatenated target/decoy databases (for example, using as a decoy all peptide sequences written in reverse) allows estimation of the false discovery rate (FDR) for the peptide data set[65]. The final list of peptide identifications is filtered to eliminate false identifications based on an FDR cutoff, typically set at 1 %.

3. Most search algorithms do a poor job in confidently identifying the correct phosphorylation positional isomer for a phosphopeptide. Thus, it is important to use a phosphorylation site localization algorithm to address this limitation. There are a number of phosphorylation site localization algorithms, such as Ascore[57], SLoMo[66], phosphoRS[58], and LuciPHOr[59]. Most of these algorithms focus on site-determining ions (i.e., those that are unique to a particular localization) and provide a site localization confidence score. In general, they are more accurate in determining the precise position of phosphorylation on a given peptide than database search algorithms.

4. Finally, when the phosphoproteomic experiment involves different experimental conditions, quantification algorithms provide a way to extract peak intensities and calculate relative peptide abundances.

5. Proteomic software packages such as MaxQuant[67], Scaffold[68], and Proteome Discoverer (Thermo Scientific) integrate tools for peptide identification, site localization, and peptide relative quantification.

Notes

1. Analytical column length and bead particle size will depend upon the maximum pressure allowed by the HPLC. Longer columns and smaller beads enhance peptide separation but result in higher HPLC pressures.

2. The purposes of acidifying the peptide samples are to inactivate the enzyme used in digestion and to increase binding to the C18 solid-phase extraction material.

3. Total mass of the peptide sample is estimated after cell lysis/protein extraction and before protein reduction, alkylation, and digestion, commonly by bicinchoninic acid method.

4. For all steps using the solid-phase extraction cartridges, be sure to close the vacuum manifold knob valves before the liquid in the cartridge runs past the polymer. Do not let the polymer dry out during any step.

5. An additional, optional step to maximize the quantity of column- bound peptides is to reload the eluent through the column. This requires separately collecting the eluent during the first sample loading step.

6. Do not force the polymer wedge into the pipette tip or it will be difficult to pass liquid through the polymer. Insert the polymer so that it is resting in place but not tightly compacted in the pipette tip. Incubation with methanol in the next step will cause the polymer to expand; thus, the wedges will become more snugly fit.

7. If using a syringe to push the liquid through, pass the liquid relatively quickly through the C18 polymer and take care to avoid dislodging the polymer from the pipette tip. If using a centrifuge, spin the pipette tip in a microcentrifuge tube at 700–1,000 rpm for 30–90 s, depending on how quickly the Phosphoproteomics Using IMAC 402 liquid passes through the polymer. In either case, do not let the polymer dry out; retain a small amount of liquid above the polymer. These stipulations apply to all wash, equilibration, sample load, and elution steps, unless further specified.

8. For sample load and elution steps, pass the liquid slowly by syringe or centrifuge about 500 rpm for approximately 2 min.

9. If the beads are to be stored for future use, wash and resuspend them in 30 % ethanol. At the time of the experiment, wash the beads three times with an 80 % ACN and 0.1 % TFA solution.

10. Optimal ratio of peptide to beads varies across biological samples. The values provided are a good starting point, but we recommend optimizing this for each sample.

11. It is important to filter out the magnetic beads that accidentally ended up in the phosphopeptide elution, since they could cause damage to the HPLC during sample analysis.

12. The basic environment assists the detachment of the phosphopeptides from the beads. However, be sure to limit the time of phosphopeptide exposure in the basic environment to 3 min to prevent side reactions involving the phosphate group. The purpose of the formic acid is to neutralize the pH of the elution.

13. Peptide fractionation by SCX chromatography: SCX chromatography can be performed prior to IMAC enrichment using an SCX column and an HPLC elution gradient to obtain high-resolution peptide separation[44]. SCX fractionation can also be performed in a solid-phase extraction cartridge format for a rapid but lower resolution fractionation[69]. Alternatively, SCX can be conducted after IMAC in microscale solid-phase extraction format by utilizing a stage tip packed with an SCX disk[70]. This method is recommended for less complex samples, such as yeast proteomes, or for smaller sample amounts. Stagetip fractionation can be done alone or in conjunction with peptide desalting, in which case the stage tip is constructed with a C18 polymer disk stacked on top of an SCX disk. The advantages of either solid-phase extraction SCX format are that they are simpler, less expensive, and less technical to operate than the SCX chromatography HPLC system. However, SCX chromatography produces higher resolution fractions and allows a greater number of obtainable peptide fractions and total phosphopeptide identifications compared to the SCX disk stage tips. For inclusion of these optional procedures in the phosphopeptide enrichment protocol refer to previous literature[52,70].

14. For mammalian samples or other highly complex samples, a longer acquisition time will improve peptide separation and increase the depth of phosphoproteome coverage. These parameters are flexible and may depend upon sample complexity, column material, and mass spectrometer scanning speed.

15. Peptides scheduled for isolation and sequencing can be fragmented by several methods, depending on the mass spectrometer used. These include CID, ETD, and HCD. CID frequently suffers from the neutral loss of phosphate as dominant fragment in the spectrum[44,52]. ETD reduces the risk of phosphate loss during fragmentation and is optimally suited for high charge states. HCD enables higher mass accuracy and the observation of lower m/ z ions, yet is a slower fragmentation method than CID.

16. Dynamic exclusion time should be adjusted to match chromatographic peak widths.

Chapter 3: Identification of novel LRRK2 substrates

3.1 Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder, affecting about 1.5% of the elderly population over 65[71]. PD is characterized by stereotypical movements including, tremors, shuffled gait, "pill rolling" hand movements and bradykinesia (stiffness, rigidity and slow movement). It is a disorder that can affect the facial expressions, sleep disturbances[72,73] and in some cases, hyposmia (loss of sense of smell) [74], which usually precedes the onset motor phenotypes. Non-motor phenotypes include cognitive impairment/dementia[75], anxiety[76] and orthostatic hypotension[77]. On a microscopic level, common/typical manifestations of PD involve the degeneration of dopaminergic neurons in the substantia nigra pars compacta where α-synuclein, ubiquitin and TAU proteins aggregate to form inclusions known as Lewy bodies[78-80].

There are several risk factors that have been associated with PD, including environmental exposure to toxins, age, and several heritable genes known as "PARK" loci. The PARK8 locus on chromosome 12 was initially discovered through a linkage analysis of PD in a Japanese family[81] and codes for the leucine rich repeat kinase 2 protein (LRRK2). Disease-segregating mutations on LRRK2 were later identified in populations through positional cloning[82,83].

Mutations in LRRK2's primary structure are thought to be the most prevalent monogenic cause of late-onset Parkinson's disease which occur in sporadic PD cases and in some cancers[84,85]. Other disease associations at the LRRK2 genetic locus have included inflammatory bowel disorder and leprosy[86-88]. Of all the disease-segregating mutations discovered in LRRK2, 100 *LRRK2* human polymorphisms have been reported[89,90], while fewer than 10 mutations are considered pathogenic.

LRRK2 protein

LRRK2 is a multi-domain 286 kDa serine/threonine protein kinase, associated with an array of cellular functions, including cytoskeletal remodeling, vesicular trafficking and lipid membrane dynamics, Wnt signaling[91,92], autophagy[93], and the regulation of protein translation[94]. In addition to an active kinase domain, LRRK2 contains a Ras of complex (Roc) GTPase domain and the C-terminal of Roc (COR), which place LRRK2 in the ROCO protein superfamily. LRRK2 harbors scaffolding domains for

interactions with other proteins to potentially mediate substrate specificity: armadillo repeats (ARM), ankyrin-repeats (ANK), leucine-rich repeats (LRR), and a WD40 domain (**Figure3.1**). Data concerning LRRK2's biochemical activity supports the protein as an active, but weak, GTPase which promotes the GTP-bound active kinase state.

Several PD-associated mutations have been discovered in LRRK2's ROC/GTPase; COR; and kinase domains, respectively: N1437H, R1441C/G/H/S; Y1699C, S1761R; G2019S, I2020T, which , may all increase LRRK2 kinase activity, however this is still an active area of research and depends on variables only present in a living system such as LRRK2 dimerization status and subcellular localization[95].

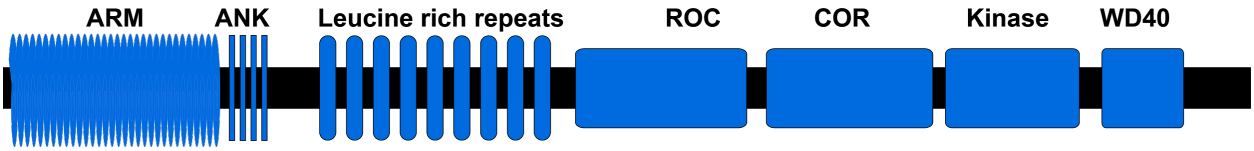


Figure 3.1. Schematic of LRRK2 domains.
ARM: armadillo repeats, ANK: ankyrin repeats, ROC: Ras-like G domain, COR: C-terminal of Roc domain, WD40: 40-mer tryptophan-aspartic acid repeat.

The most prevalent LRRK2 mutation is the glycine to serine change at position 2019 (LRRK2-G2019S) in the activation-loop segment of the kinase, leading to a ~3-fold increase in kinase activity[96-98]. The mutation accounts for about 6.6% of familial Parkinson's disease cases of people from European origin[99], 29.7% in populations with Ashkenazi Jewish ancestry[100] and up to 37% in North African Arab-populations[101]. In all three populations, pathogenic LRRK2-G2019S presents itself in an autosomal dominant, pleomorphic way, with age-dependent penetrance. Patients heterozygous for this mutation exhibit characteristics that are indistinguishable from patients with sporadic PD in the general population[99]. LRRK2-G2019S is typically, but not always, associated with α-synuclein positive Lewy bodies[102] while kinase-dead forms of the protein delay cell death and stops the formation of protein inclusions[103].

Cellular and animal models to study LRRK2 in the context of PD

Alterations in LRRK2 kinase activity are likely to confer a disruptive outcome, therefore investigators over the past fifteen years have explored the function of LRRK2 and have experimentally described the pathobiology for LRRK2-G2019S mutation in various model systems.

In SH-SY5Y neuroblastoma cells, LRRK2-G2019S pathogenicity is attributed to its hyperactive kinase function. LRRK2-G2019S reduces neurite outgrowths, typically present in neuronal cells grown in culture[104]. In rat primary cortical neurons, LRRK2 aggregates to form pathological inclusion bodies only when the kinase activity of LRRK2 is preserved, but not in a kinase-dead form of the protein[103]. LRRK2 can also form halo-shaped aggregations around protein inclusions[105].

Animal models have been essential for understanding the extent of LRRK2's hyperactive form in generating a movement phenotypes. Studies in mice, rats, and the fruit fly *Drosophila melanogaster* attempted to recapitulate PD phenotypes with varying results.

Some mouse models of LRRK2-G2019S show hyperkinetic activity with abnormal mitochondrial morphology and dynamics[106,107]—more similar to pathologies described by models of PINK1 and PARKIN mutations (PARK2 and PARK6 loci), but lack the movement-deficit phenotype. Other mouse models using a higher expression HSV amplicon as a gene transfer tool have demonstrated severe nigral neurodegeneration by LRRK2-G2019S[108]. In rats, LRRK2-G2019S causes substantial nigral dopaminergic neurodegeneration[109] and significant reduction in motor performance on the rotarod test[110], while both phenotypes have not been reported in the same model.

Drosophila however, represents the single model that displays both age-dependent phenotypes for locomotion and neuronal loss[111]. The described fly model expresses wild-type (WT) or pathogenic human LRRK2-G2019S under an inducible promoter system in retina, brain or a pan-neuronal pattern. Evidence for retinal degeneration in both WT-LRRK2 and LRRK2-G2019S transgenic fly lines, supports the idea that LRRK2 plays a cytotoxic role *in vivo*. Interestingly, Parkinson's disease also produces many abnormalities in the human visual system[112] including reduced dopaminergic neuron innervation in the eye[113], misfolded α-synuclein and phosphorylated α-synuclein in the inner retinal layer [114,115]. It is not clear, however, to what extent LRRK2 specifically plays a role in the visual system in humans with PD. Further, when LRRK2-G2019S is expressed in dopaminergic neurons and pan-neuronal expression, flies

exhibit neurodegeneration and a movement deficit akin to that in humans. This model revealed that non-mammalian systems like flies recapitulate phenotypes similar to human PD even when other mammalian systems fail to do so, making the fly a suitable model for studying LRRK2-induced parkinsonism.

LRRK2 substrates

With LRRK2 as an established disease-risk protein kinase, a deeper understanding of its function through its substrates holds promise to better understand its role in the cell, and may aide in the development of therapeutics and drugs. Although some work has been done to identify biologically relevant substrates of LRRK2, many substrates have been identified from PD models that cannot fully recapitulate PD relevant phenotypes, and only a few of them have been fully validated *in vivo*.

The first LRRK2 substrate identified was LRRK2 itself, through its intramolecular autophosphorylation activity and the phosphorylation of MPB (myelin basic protein) [96]. These phosphorylation events confirmed LRRK2 as an authentic kinase and provided a straightforward way to assess the activity of LRRK2. Subsequently, others have taken a fractionation approach and classical Edman degradation to search for and localize LRRK2 activity on purified substrates— LRRK2-G2019S is known to phosphorylate the cytoskeletal proteins moesin, ezrin and radixin at homologous sites (moesin T558) [98]. This finding led to the frequently used LRRK2-optimized peptide, “LRRKtide” (RLGRDKYKTLRQIRQ), to assay LRRK2 activity. This peptide substrate was improved upon by development of the more reactive, optimized generic LRRK2 peptide substrate “Nictide” (RLGWWRFYTLRRARQGNTKQR) [116]. LRRK2 also phosphorylates members of Lewy Body aggregates: tubulin-associated Tau (associated with Alzheimer’s disease) [117] is phosphorylated on several residues in vitro [118] and α-synuclein on S129[119].

Evidence indicates that LRRK2 can act as a mitogen-activated protein kinase through sequence homology, which may include MKK3, 4, 6 and 7 as substrates[120]. LRRK2-mediated interactions suggest that LRRK2 may play a role in cytoskeletal remodeling via a role in Wnt signaling through an interaction with GSK3 to mediate the phosphorylation of TAU protein[92]. In autophagic membrane remodeling of cells, LRRK2 is associated with autophagic structures and pathways to regulate the neurite processes[93]. Studies of membrane dynamics are providing further evidence that LRRK2 can modulate synaptic vesicle

(SV) dynamics by phosphorylating proteins in the pre-synapse like EndoA at S75[121] and Snapin at T117 in a direct kinase assays[122]. It is possible that EndoA, Snapin and LRRK2 together are part of a membrane remodeling network in synapses, where LRRK2 acts as a binding partner to phosphorylate the presynaptic ATPase, N-ethylmaleimide sensitive fusion (NSF) protein at T645 in the ATP binding pocket. Further, LRRK2 is known to increase NSF's ATPase activity through phosphorylation of T645.

Apart from the focus of LRRK2 and PD, the kinase can also phosphorylate p53 on T304 and T377, following a T-X-R motif. As a tumor suppressor, p53 is an intriguing candidate as it links the finding that LRRK2 is also implicated in cancer and apoptotic processes. Evidence from *in vivo* and *in vitro* studies suggest that LRRK2 can also phosphorylate Akt1 at its activation site on S473, supporting a connection between altered LRRK2 function and cell survival mechanisms [123].

In this work, I apply a mass-spectrometry based phosphoproteomics strategy to uncover novel potential substrates *in vivo* and *in vitro*. Using the transgenic *Drosophila melanogaster* fly PD model described above, *in vivo* I quantitatively compare LRRK2 expressers to non-expressers through proteomics and phosphoproteomics data. I couple this approach to an orthogonal one that utilizes purified human LRRK2 in *in vitro* reactions on tryptic peptides from fly head extracts for neuronal tissue. I also apply this same *in vitro* approach to human neuroblastoma cell lines to uncover potential substrates there. A similar approach has shown to be a successful strategy for identifying potential substrates of the spleen tyrosine kinase (Syk) [23]. My strategy has the benefit of capturing the direct interaction between the peptide substrates and the kinase, while adding the context of a living system to kinase activity. Here, *in vivo* proteomics adds potential members of a pathway, including direct and downstream effectors of LRRK2. Applying this strategy to the PD fly model, will contribute to an understanding of how LRRK2 confers the neurodegenerative phenotypes and the motor phenotypes through LRRK2 kinase activity.

Other proteomics studies have revealed a number of LRRK2 substrates including a subset of Rab GTPases ribosomal proteins. Steger *et al.* (2016) [124] combined a phosphoproteomics, genetics and pharmacological approach to find Rab GTPases as key LRRK2 substrates from mouse embryonic fibroblasts. Specifically, the strongest evidence from the literature shows that LRRK2 can phosphorylate Rab10 GTPase at T73. Martin *et al.* (2014) [125] took an *in vitro* approach to confirm the phosphorylation of co-purified proteins from a LRRK2 tandem affinity purification (TAP) followed by liquid chromatography-

tandem mass spectrometry (LC-MS/MS) to confirm phosphorylation sites. Pathogenic LRRK2-G2019S showed increased phosphorylation of substrates belonging to the 40S ribosomal subunit including s11 (T28,T46 and T54), s15 (T136), and s27 proteins[125].

The application of mass spectrometry to detect, localize and quantify phosphorylation events provides an excellent way to reduce bias in detection of potential substrates. The advantage I present here is that *in vitro* kinase assays on peptides are not subject to missing substrates due to limitations of the biological conditions needed to expose a residue to phosphorylation. Here, all linear sequences are likely to be exposed for phosphorylation, revealing additional potential substrates of interest in a direct kinase-substrate assay. This study is a complement to the growing knowledge-base of what the substrates of LRRK2 are, and provides insight into LRRK2 function in terms of PD and other associated neurodegenerative diseases.

3.2 Materials and methods

Drosophila lines

Fly stocks were maintained on standard cornmeal and molasses food at 25°C in multiple different stock bottles, and turned over regularly to prevent overcrowding. Transgenic UAS-LRRK2 flies were a gift from Dr. Wanli Smith at John's Hopkins University School of Medicine. Specifically, UAS-LRRK2-wild-type (III) and UAS-LRRK2-G2019S-2 (no balancer) were maintained. All crosses were carried out by mating virgin Elav-Gal4 female flies to UAS-LRRK2 males. Flies were described previously in Liu *et al.* (2008) [111].

Reagents

GST-tagged LRRK2 spanning residues 907 to 2527 was purchased (Life Technologies). Nictide and GST-LRRK2-G2019S was purchased from the MRC (Dundee, UK). Nictide sequence is RLGWWRFYTLRRARQGNTKQR.

Autophosphorylation reaction

Human wild type GST-tagged LRRK2 constructs were incubated for 60 minutes at 30ºC in a kinase reaction buffer containing 25mM Tris pH 7.5, 10mM $MgCl_2$, 1mM EGTA, 1mM Na_3VO_4, 5mM B-glycerophosphate and 2mM dithiothreitol with or without 1mM ATP. The reaction was placed on ice and diluted with H_2O and Tris pH 8.9 to 50 mM Tris. LRRK2 was digested with trypsin at an enzyme-to-

substrate ratio of 1:100 at 37ºC for 3 hours. Digestion was quenched by acidification with 10% TFA to pH 2.

Identification of peptides and modifications from autophosphorylation

Raw mass spectra were searched using Comet version 2015.01[126] against a database containing full length human LRRK2 (NCBI version gi:83722282) and silkworm database from SilkDB (14,623 protein entries). This combination of databases was to check for any contamination that may have arisen from the expression of LRRK2 in the insect cells. Searches allowed variable modifications for two oxidized methionine residues and two phosphorylated serine, threonine or tyrosine residues.

Fly head peptide extraction

Approximately 200 flies are harvested and snap-frozen in liquid nitrogen, vortexed to dismember the flies and isolated on a pre-chilled wire sieve apparatus with appropriate pore size. Fly heads were homogenized in 50 mM Tris, 75 mM NaCl, 9 M urea, and pH 8.0 buffer containing protease inhibitors (Roche, complete EDTA-free) using a 1 mL dounce homogenizer on ice for 10 minutes of continuous reciprocation. Homogenate was spun at 12,000 g for 10 minutes to remove debris. The supernatant was treated with 5 mM dithiothreitol for 30 minutes at 55°C to reduce protein disulfide bonds, and then with ten 15 mM of iodoacetamide for 30 minutes at room temperature in the dark to alkylate cysteine residues. The alkylation reaction was quenched with a second addition of 5 mM DTT with 15 min incubation at room temperature. The concentration of urea was diluted five-fold with 50 mM Tris pH 8.2 before a 12 hour trypsin digestion at a ratio of 1:100 (enzyme: substrate). The peptides were desalted on a 50 mg tC18 SepPak cartridge prior to labeling of peptides.

Reductive dimethyl labeling of tryptic peptides

For each of the heavy and light samples, 1.25 mg of desalted and lyophilized peptides were resuspended in 1.25 mL 1M MES buffer pH 5.5. Peptides were labeled in solution by adding 50 µl of 4% formaldehyde (Sigma, “light” label) and 50 µl of 600 mM sodium cyanoborohydride ($NaBH_3CN$, light) (Sigma) or “heavy” formaldehyde (COD_2, heavy) (Cambridge Isotopes) and sodium cyanoborodeuteride ($NaBD_3CN$, heavy) (Cambridge Isotopes). The reaction was incubated for 2 minutes after the addition of

formaldehyde and incubated for 10 minutes with subsequent addition of sodium-cyanoborohydride. Each reagent was added a second time with occasional agitation by manually flicking the tube and incubated for 1 hour. The labeling reactions were quenched with an equal volume of 10% TFA at a pH of ~2, followed by desalting on 50 mg tC18 Sep Pak cartridge (Waters) and lyophilizing the eluant.

GST-LRRK2-WT kinase reaction

Samples of heavy and light labeled tryptic peptides were incubated with, or without GST-LRRK2-WT for 4 hours at room temperature with shaking in triplicate. Buffer was composed of 25mM Tris pH 7.5, 75mM NaCl, 10mM $MgCl_2$, 1mM EGTA, 1mM sodium orthovanadate, 5mM beta-glycerophosphate, 2mM DTT, and 1mM ATP. pH was adjusted to ~7.4 with 3µl sodium hydroxide. After the kinase reaction each sample's pH was about 6.5-6.8 and adjusted to ~7.1 for subsequent 3-hour digestion with 15 ng of trypsin at 37ºC.

Identification and quantification of peptides, proteins and modifications

All LC-MS/MS data were acquired in ".raw" file format and converted to the mzXML test format using readW. The data were searched against the FlyBase *Drosophila melanogaster* translated ORF database (version 6.05, 30,444 entries) using the Comet algorithm. Variable modifications included two oxidized methionines per peptide and three phosphorylated serine, threonine or tyrosine residues. Peptides were quantified with an in-house quantification algorithm. Log_2(Heavy/Light) peptide ratios were calculated by taking the area under the curve for each identified labeled peptide. Ratios were normalized to mean of Log_2(heavy/light) peptide rations from LC-MS/MS analysis prior to IMAC enrichment to account for any mixing error.

***In vitro* human GST-LRRK2-G2019S kinase assays**

Kinase reactions were performed in a buffer of 100mM Tris-HCl pH 7.5, 10mM $MgCl_2$, 0.1mM EGTA and 1mM ATP. Human GST-LRRK2-G2019S was reacted with tryptic fly peptides extracted from heads or SH-SY5Y tryptic peptides in the presence of the nictide peptide at a molar ratio of 1:800 (LRRK2: nictide), and 1: 15.9 kinase to peptides (w/w). Replicate reactions were incubated with (n=10) or

without (n=10) LRRK2-G2019S. Nictide positive controls (n=1) were phosphorylated with the same reaction buffer, kinase and in parallel to peptide reactions without the tryptic peptides. Each replicate of the kinase-positive condition was sampled for LC-MS/MS analysis to collect positive control data of LRRK2 activity through the detection of the phospho-nictide species.

Immobilized metal ion affinity chromatography (IMAC):

Peptides were dissolved in 80% acetonitrile, 0.1% TFA and incubated with 50µl of a 5% slurry of Fe-NTA magnetic agarose beads. To prepare Fe-NTA beads, Ni-NTA beads (Qiagen) were washed three times with H_2O then incubated with 40mM EDTA pH 8.0 for 30 min at room temperature. Beads were then washed with H_2O three times and incubated with 10mM iron chloride for 30 min. Samples were enriched for phosphopeptides in part of a 96-well plate format, using the KingFisher Flex magnetic particle-handling robot (Thermo Scientific) programmed to incubate peptides with 150 µl 5% bead slurry for 30 min, wash 3 times with 150 µl of 80% acetonitrile, 0.1% TFA and elute with 50µl of 1:1 acetonitrile:2.5% NH_4OH. The eluents were acidified with a 10% FA acid solution before filtration through a stage-tip constructed with two layers of C18 Empore material (3M).

Identification and quantification of peptides, proteins and modifications

Data were searched using the Comet search algorithm against the Swissprot human database from April 8th, 2015 (20,204 entries) or the *Drosophila melanogaster* database (version 6.09, 30,452 entries). Variable modifications included two oxidized methionines per peptide and three phosphorylated serine threonine or tyrosine residues. Label-free quantification was achieved by normalizing to the total summed intensities.

***In vivo* identification and quantification of peptides, proteins and modifications**

Data were searched using MaxQuant[67] version 1.5.2.8 using a Drosophila melanogaster database from April 7th, 2015 18,799 entries, with methionine oxidation and serine, threonine and tyrosine phosphorylation as variable modifications. Reductive dimethylation labels were specified as

"DimethLys0"; DimethNter0" for the light label and "DimethLys6"; "DimethNter6" for the heavy label. Data from stagetip fractionation, strong-cation exchange and IMAC were analyzed in one implementation of MaxQuant where fractions were grouped by experiment for a combined search and quantification. Protein identifications, sites and peptide spectral matches were filtered to a 1% false-discovery rate.

Skyline

PRM MS/MS data were visualized with the Skyline[127] software package.

Mammalian cell culture and peptide extraction

SH-SY5Y cells (ATCC) were grown in DMEM (Gibco) with high glucose, L-glutamine with added penicillin/streptomycin (Gibco) and fetal bovine serum to 10%, cells were passaged and split 1:5 when necessary and harvested at a confluence of about 60-70%. SH-SY5Y cells were quickly washed three times with ice-cold phosphate buffered saline and snap-frozen in liquid nitrogen. Cells were mechanically scraped into ice-cold buffer containing 50mM Tris pH 8.0, 75mM NaCl, 9M urea containing protease inhibitors (Roche, complete EDTA-free). Samples were sonicated on ice 3 times for 30s each with a 1 minute pauses in between. Cells were spun down at 12,000 g for 10 minutes to precipitate insoluble material and cell debris. To measure the protein content, the supernatant was assayed through the bicinchoninic acid method. Samples were immediately reduced and alkylated with 5mM DTT at 55ºC and alkylated with 15mM iodoacetamide at room temperature followed by a second addition of 5mM DTT to quench alkylation. The concentration of urea was diluted five-fold with 50mM Tris pH 8.2 before an overnight trypsin digestion at a weight/weight ratio of 1:100 (enzyme: substrate).

LC-MS/MS data filtering and phosphorylation site localization

Database search results were filtered by Percolator[128] and phosphorylation sites were localized with an in-house implementation of the Ascore algorithm[57].

Statistical significance

Student's t-test was conducted across replicate samples to determine replicate p-values used for cutoffs featured in volcano plots.

Gene ontology enrichment

Enrichment analysis was performed using the GO-rilla online tool[129,130]. Functional ontology was determined by uploading a subset of protein names (p-val<=0.01, Log_2(+kinase/-kinase) >=0.05) and comparing to a background of all protein names obtained in the complete Log_2(+kinase/-kinase) calculated dataset.

Phosphopeptide LC-MS/MS analysis

In general, peptides were resuspended in 4% formic acid, 3% acetonitrile and loaded onto a 100µm x 3 cm pre-column packed with Reprosil C18 1.9 µm, 120 Å particles (Dr. Maisch GmbH). Peptides were separated by reverse phase chromatography on an EASY-nLC-II (Thermo Scientific), with a 100µm ID x 30 cm column packed with same C18 particles. The gradient was 3-30% acetonitrile in 0.15% formic acid. Full MS spectra were collected in the Orbitrap, while MS/MS spectra were collected in the ion trap by selecting the 20 most abundant ions (30ms injection time) with collision-activated dissociation (CAD). The autophosphorylation reaction was run in a 90minute gradient of 200 femtomoles of digested LRRK2; *in vitro kinase* reactions were analyzed in a 120 minute gradient.

Peptides analyzed for protein level LC-MS/MS analysis

Peptides were injected onto a 40cm x 100µm column as above and analyzed on an EASY-nLC-1000 coupled to a hybrid quadrupole-orbitrap QExactive mass spectrometer (Thermo Scientific) with data dependent settings over a 90 minute gradient.

Nictide analyses

Pre-imac samples and parallel positive controls were analyzed using the LTQ Velos Orbitrap as above with the exception of the acquisition method, which used parallel reaction monitoring (PRM). An MS1 full-scan was programmed to precede a top-18 most abundant ion MS/MS and parent masses of +5 nictide (m/z=550.70) and phosphorylated +5 nictide (mz=566.70) were programmed into the acquisition method in addition to collecting MS/MS scans on the top 18 most abundant ions.

3.3 Results

In vitro LRRK2-WT autophosphorylation

To identify novel autophosphorylation sites of LRRK2 and ensure our LRRK2 is active, I devised an autophosphorylation kinase assay in which human recombinant LRRK2 is incubated in kinase reaction buffer in the presence or absence of ATP, the protein is digested and peptides are analyzed by LC-MS/MS **(Figure 3.2)**.

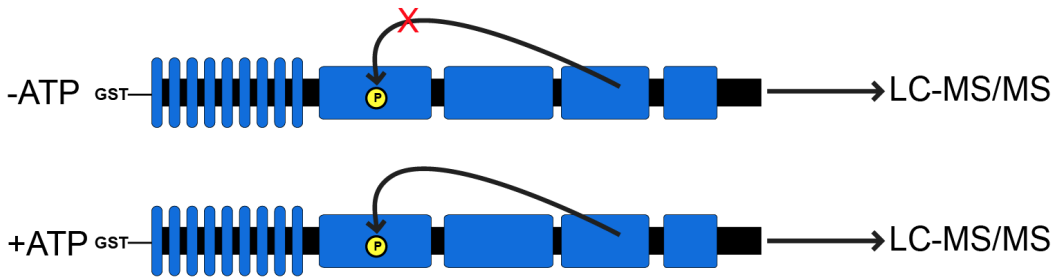


Figure 3.2 Autophosphorylation experimental approach.
LRRK2 autophosphorylates through an intramolecular mechanism. Two conditions were compared: 1) An autophosphorylation reaction without ATP (above, red "X" depicts no phosphorylation at an arbitrary site). 2. An autophosphorylation reaction with ATP to allow for the kinase domain (arrow origin) to phosphorylate another region of the protein (depicted here is the ROC domain). The two conditions were subject to LC-MS/MS for analysis.

Using this assay, I identified 93 peptides in the +ATP condition and 79 peptides in the -ATP control, which corresponds to about 31% coverage of the protein (+/-10 ppm filter the comet search results for LRRK2 phosphopeptides). Of these peptides, I found 20 and 1 non-redundant phosphopeptide sequences mapping to LRRK2 for the +ATP and -ATP conditions, respectively. These 20 phosphopeptides correspond to 11 unique phosphorylation sites **(Table 3.1)** Notably, phosphorylation at T1503 is used as LRRK2 kinase activity reporter[131].

Of the autophosphorylation sites identified, T1452 showed the greatest signal, followed by T1491, T1849, and T1503 **(Table 3.1)**. Sites T1452, T1491 and T1503 exist in the ROC domain, and site T1849 in the COR domain of the protein **(Figure 3.3)**, confirming the previous finding that the major target of LRRK2 autophosphorylation is the ROC domain[132].

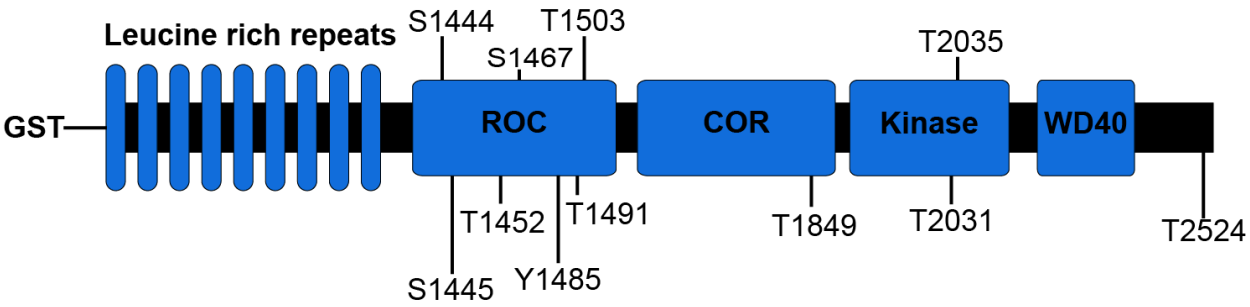


Figure 3.3 Positions of identified LRRK2 autophosphorylation sites.
Schematic of the GST-tagged LRRK2 used in the assay. The N terminal GST sequence is attached to the LRRK2 native sequence which spans resides 907-2527. Sites S1445, T2031 and T2035 had Ascores <13. For all other sites Ascore >=13.

Table 3.1 Autophosphorylation sites of LRRK2
Quantified unique phosphopeptides by charge state identified through a comet search. Filtered for +/- 10 ppm.

Ascore sequence	Charge	LRRK2 Sites	Log(Control Area)	Log(ATP area)
ARASS#SPVILVGT#HLDVSDEKQR	4	S1444,T1452	-	4.87
ARASSS#PVILVGT#HLDVSDEKQR	4	S1445,T1452	-	4.87
ASS#SPVILVGT#HLDVSDEKQR	3	S1444,T1452	-	5.49
ASS#SPVILVGT#HLDVSDEKQRK	3	S1444,T1452	-	4.77
ASSS#PVILVGTHLDVSDEKQR	3	S1445	4.00	4.62
ASSSPVILVGT#HLDVSDEK	2	T1452	-	4.87
ASSSPVILVGT#HLDVSDEKQR	2	T1452	-	4.65
ASSSPVILVGT#HLDVSDEKQR	3	T1452	-	6.64
ASSSPVILVGT#HLDVSDEKQRK	3	T1452	-	5.40
DYHFVNAT#EESDALAK	2	T1491	-	4.93
GFPAIRDY#HFVNAT#EESDALAK	3	Y1485,T1491	-	4.77
GFPAIRDYHFVNAT#EESDALAK	3	T1491	-	6.26
GFPAIRDYHFVNAT#EESDALAKLR	4	T1491	-	5.18
KACMS#KITK	2	S1467	-	5.55
LRKT#IINESLNFK	2	T1503	-	4.71
LRKT#IINESLNFK	3	T1503	-	6.19
LT#IPISQIAPDLILADLPR	2	T1849	-	5.42
M*GIKT#SEGTPGFRAPEVAR	3	T2031	-	4.25
MGIKT#SEGTPGFR	2	T2031	-	4.53
MGIKT#SEGTPGFRAPEVAR	3	T2031	-	5.90
MGIKTSEGT#PGFRAPEVAR	4	T2035	-	6.05
MRRT#SVE	2	T2524	-	7.30

Identification of *in vitro* LRRK2 substrates:

Kinase assay reactions:

Next, I applied the active kinase used in the autophosphorylation assay to find novel substrates of LRRK2 in the fly. Since previous work[111] has established *Drosophila melanogaster* as a model of PD through expressing human LRRK2(WT or G2019S mutation), finding substrates for LRRK2 among fly peptides would provide further insight into the conserved mechanisms through which human LRRK2 impart the model's disease phenotypes. In order to achieve a comparative, quantitative output of newly phosphorylated peptides, I used a chemical labeling strategy at the peptide level to distinguish between two treatments in triplicate experiments. I chose to utilize a reductive dimethylation strategy[133] to differentially label peptides that would be subject to two conditions in the *in vitro* kinase assay **(Figure 3.4)**. In this way, peptides from digested fly head protein extract are labeled with stable isotopes of carbon and hydrogen in dimethyl groups on free amines (N-termini and lysine residues).

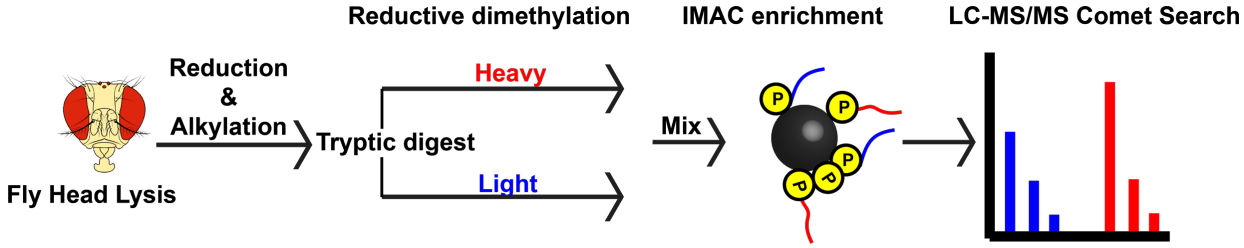


Figure 3.4 Experimental overview of *in vitro* kinase assay using reductively di-methylated tryptic peptides
Fly heads were harvested, extract was reduced and alkylated before digested with trypsin. After the sample was divided into two, peptides were differentially labeled either heavy or light using reductive dimethylation labeling chemistry. Heavy peptides were reacted with GST-LRRK2-WT (red). Light peptides were reacted without LRRK2 under otherwise identical conditions. The two conditions were mixed for IMAC phosphopeptide enrichment and LC-MS/MS analysis. This was repeated in triplicate. A single mock condition was analyzed where neither heavy nor light condition contained LRRK2 for a "mock" experiment.

The chemical reaction resulted in labeling efficiencies of ~ 99% for both heavy and the light peptides pools. The heavy labeled peptides were reacted with LRRK2 in the presence of ATP, while the light peptides were not reacted with LRRK2. Prior to phosphopeptide enrichment of the kinase reaction, a sample was taken for quantitative analysis which enabled the normalization of phosphopeptides after phosphopeptide enrichment and LC-MS/MS. The triplicate searches and quantification resulted in an average of 13,900 labeled peptides (about 6,950 unique peptides) across three experiments covering

about 1,750 proteins. After phosphopeptide enrichment, replicate samples resulted in about 5940 quantified phosphopeptide heavy: light ratios counting each phosphopeptide charge-state as a separate quantification, present in one or more replicates. Applying student's t-test ($p<0.01$, change of $Log_2(H/L)>=0.5$, normalized by pre-IMAC quantitative distribution median) reveals significantly changing phosphopeptides **(Figure 3.5)** mapping to 8 unique proteins, including 8 phosphorylation sites **(Table 3.2).**

Table 3.2 Candidate substrates by GST-LRRK2-WT
Peptides found to be significantly phosphorylated by LRRK2 ($p<0.01$, $Log_2(H/L)>=0.5$). *Flybase ID FBpp0305847 from comet search mapped to alternative equivalent entry. Phosphorylation site denoted by "p" before residue.

Ascore Sequence_charge	Flybase polypeptide	Site	Symbol	Example Human Ortholog
DDASSSSTHDADGASLSGKpSSPVER_3	FBpp0305848*	S171	mtd	TLDC2
GALNRDDNNDDDDATTLAPNSNEDpYDTRPQYSFAYDVR_4	FBpp0309533	Y71	Crys	ZNF160
RApSSGVGPTNAAATVAAATGAVGAVNPSNNYNAAGSAADR_3	FBpp0303278	S622	par-1	MARK4
RTNpSVPMGDGDIGAGSTTSGDSGEAQEVVLPK_3	FBpp0307490	S1121	scrib	SCRIB
SHTAGSTDpSAEKNNAANK_2	FBpp0289693	S35	CG42492	Otopetrin
TGpTLTQNR_2	FBpp0312297	T369	JYalpha	ATP4A
TKEEGGTSETASEAASEAATPAPAATPAPAASApTGSK_3	FBpp0088688	T46	Mlc2	MYL5
VQPDM*EELTENAEEGDEEDKQpSEQKLGR_4	FBpp0310442	S189	Syt1	SYT3

Although LRRK2 is known to phosphorylate mostly threonine residues, the significant changes include serine and threonine residues and a single tyrosine residue. A search for a motif in peptides that were significantly changing did not result in a significant motif nor did they fit the F/Y-x-T-x-R/K motif described in the literature[134].

Upon converting the FlyBase polypeptide ID to the Uniprot protein entries, the polypeptides mapped changing mapped to several potential isoforms which were included in a GO enrichment analysis for both the background (all IDs in dataset) versus the foreground (p-value <0.01, $Log_2(H/L) \geq 0.5$). Changes observed corresponded to enriched terms including "calcium ion binding" (Syt1, Mlc2 and Crys), "locomotion" (Mlc2, scrib, JYalpha, and Par-1) and "neuromuscular junction" (Syt1, scrib, and par-1) (p-value threshold < 10^{-3}). JYalpha also maps to a atpalpha, a protein with a conserved, and highly overlapping amino acid sequence. The peptide detected here maps to both proteins.

The most extreme, changes **(Figure 3.5)** includes members of the ATPase and synaptotagmin protein families, JYAlpha and Syt1. The most dramatic change (~40-fold) was observed on the phosphopeptide "TGTpLTQNR" through a well-localized T369 (Site localization Ascore >13) site. This peptide maps to JYalpha/ Atpalpha protein isoforms in the fly, with clear homology to the human p-type potassium ion ATPases (ATP1A3, ATP1A1, ATP1A4, ATP12A, ATP4A) and homology to ATP13A2, a

known recessive early-onset PD locus. Interestingly, ATP13A2 is known to regulate the expression and the ubiquitylation of a member of synaptotagmin Syt11 in humans, whose human ortholog is found in this group of changing peptides, as Syt3[135] Phosphorylated S1121 was identified mapped to SCRIB, a leucine-rich repeat protein involved in protein binding (Ascore = 10.3) was also identified as the second-highest changing site. The third change, S171 (non-localized) corresponds to the "mtd" protein, whose orthologs include TLDC2 protein in humans, which harbors a TLDc domain known for regulating neuroprotection against oxidative stress[136].

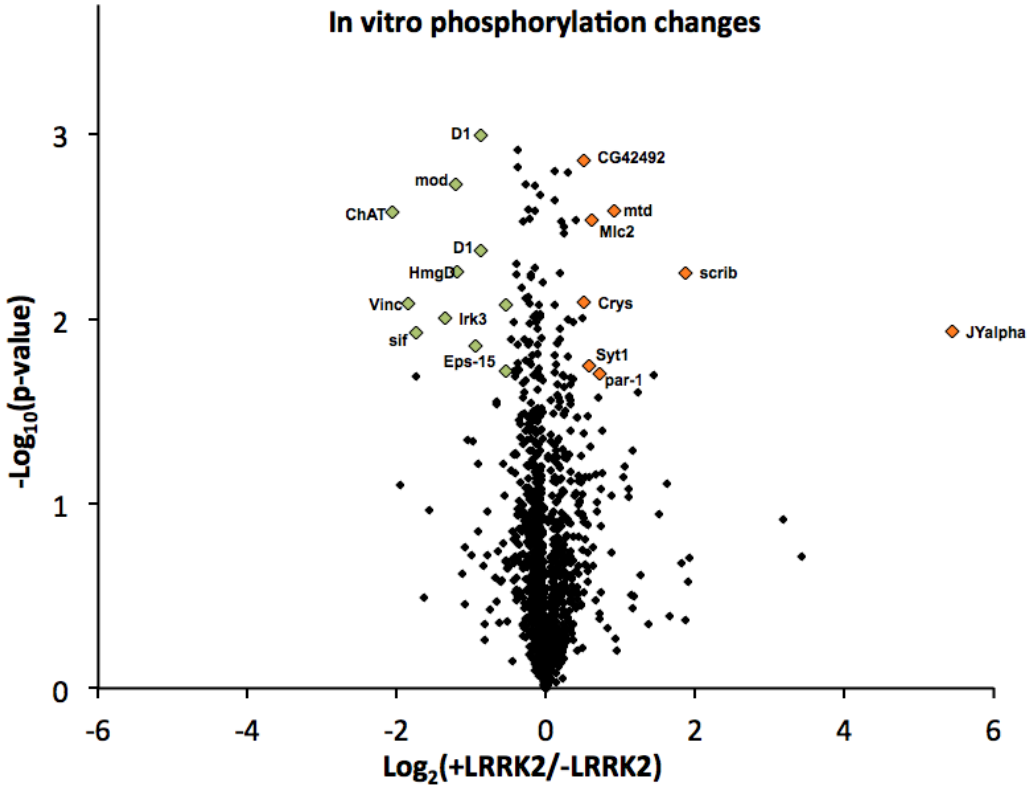


Figure 3.5 Volcano plot of average changes of phosphorylation sites.
Points on plot represent phosphorylation sites with an Ascore value >= 13 and normalized average $Log2_2$(+GST-LRRK2/-LRRK2). Green and orange points represent decreasing and increasing sites, respectively ($p<0.05$).

To identify other interesting candidates not detected in all of the replicates, I applied a less stringent filter across the three replicates for over 40-fold changes not present in a mock reaction, revealing several interesting candidate substrates over a mean 40-fold change **(Table 3.3).** The greatest change from these criteria resulted in identifying a phosphorylation site on porin (human VDAC3 ortholog)

on T173. Porin is a voltage-gated anion channel, though to regulate mitochondrial organization[137,138]. Upon closer inspection, this site nearly resembles the F/Y-x-T-x-R/K motif as a Y-Tp-x-R/K site. Similarly, ACTA1 follows a F/Y-Tp-x-x and CCDC94 follows F/Y-Tp-x-R/K motif.

Table 3.3 Other LRRK2 substrates found in one or more replicates
Phosphopeptides changing ≥ 40-fold in one or more replicates. These phosphopeptides were not found in the mock control samples. Asterisk (*) indicates oxidized methionine.

Ascore sequence_charge	Flybase polypeptide	Sites	Symbol	Rep1	Rep2	Rep3	Example Human Orthologs
ADSSpTEDSSSEDDAPKK_3	FBpp0305035	T320	Nopp140	-	-5.56	-	NOLC1
LTTNNFALGYpTTK_2	FBpp0304343	T173	porin	-	-	-8.60	VDAC1, VDAC2
QYpTVRLMAPFNM*RCK_2	FBpp0086387	T32	CG8435	-	-8.05	-	CCDC94, CCDC130
ETDADEVDGRYSPPPpTSSAR_3	FBpp0307423	T237	CG2246	-	-	-7.63	PRPSAP2, PRPSAP1
LQAGCQpSQSLPpTEV_2	FBpp0308679	S791, T796	CG12290	-	-	-7.58	-
GYSFpTTTAER_2	FBpp0312019	T202	Act87E	-	-5.20	-7.57	ACTA1, ACTA2,ACTB
SNLSSSAApSpSM*LGLNSAGSSPAHQLHASMTGGIGYGMK_3	FBpp0309144	S545, S546	Nup153	-	-5.78	-	-

To explore where the identified sites reside in the protein structure, I mapped JYalpha T369 and porin T173 onto available structural data from respective orthologs **(Figures 3.6 & 3.7)**. Site T369 maps to a disordered region of the protein neighboring the binding of a Mg^{2+} ion and the site of ADP binding. This suggest that phosphorylation at T369 may be a functionally relevant regulatory mechanism for ATPase activity in the protein. The orthologous phosphorylation site was not identified in in the PhosphoSitePlus® database for human, pig or rat, but has been identified in mouse. Porin's T173 site is in an apical region of the beta-barrel wall adjacent to the pore's aperture-controlling alpha helix. This site may be of potential regulatory importance for controlling the aperture of the pore and its ion permeability.

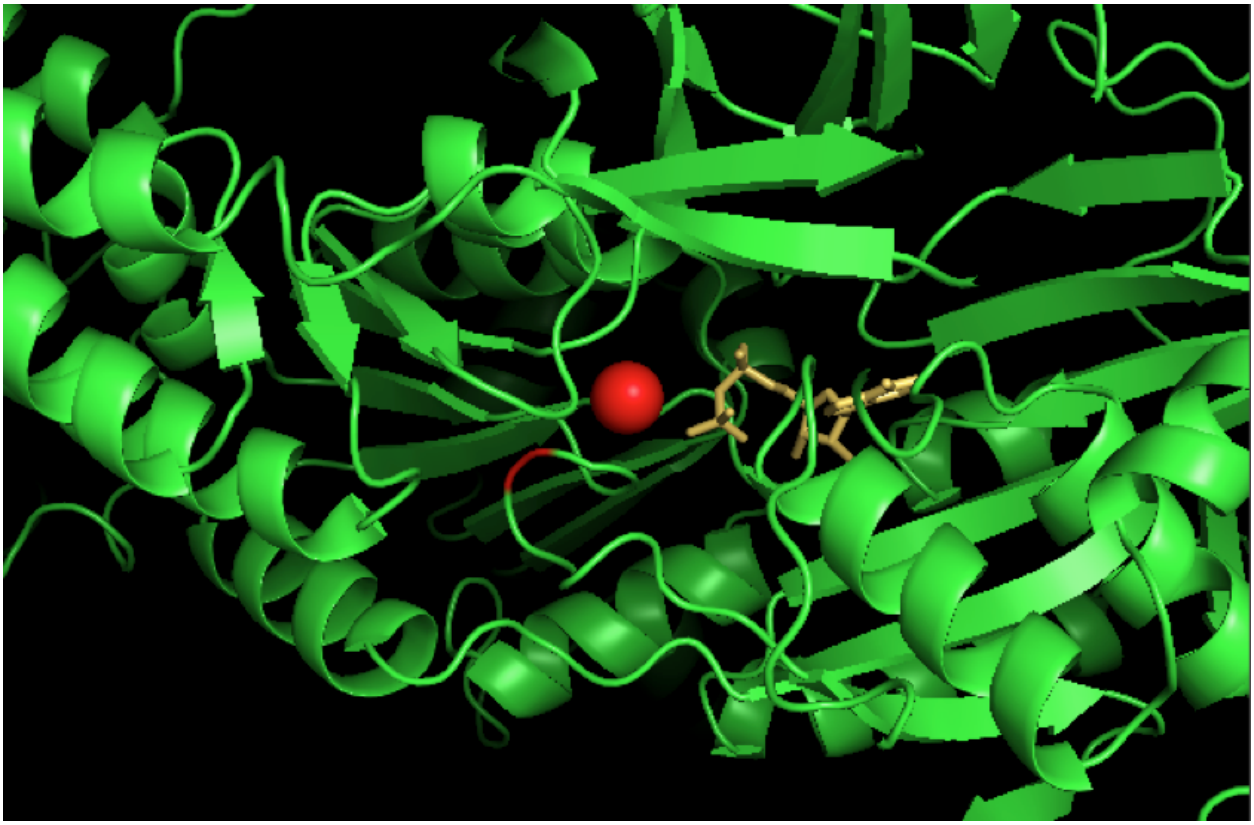


Figure 3.6 JYalpha is an *in vitro* LRRK2 substrate.
A model of the porcine JYalpha ortholog (sodium/potassium atpase) is shown from PDB (ID: 3WGU). The phosphorylated T398 residue is shown as a red segment in a close-up of regulatory region and catalytic domain of the protein showing the relative positions of the phosphorylation site of interest, an Mg^{+2} ion binding site and the ADP in the binding pocket.

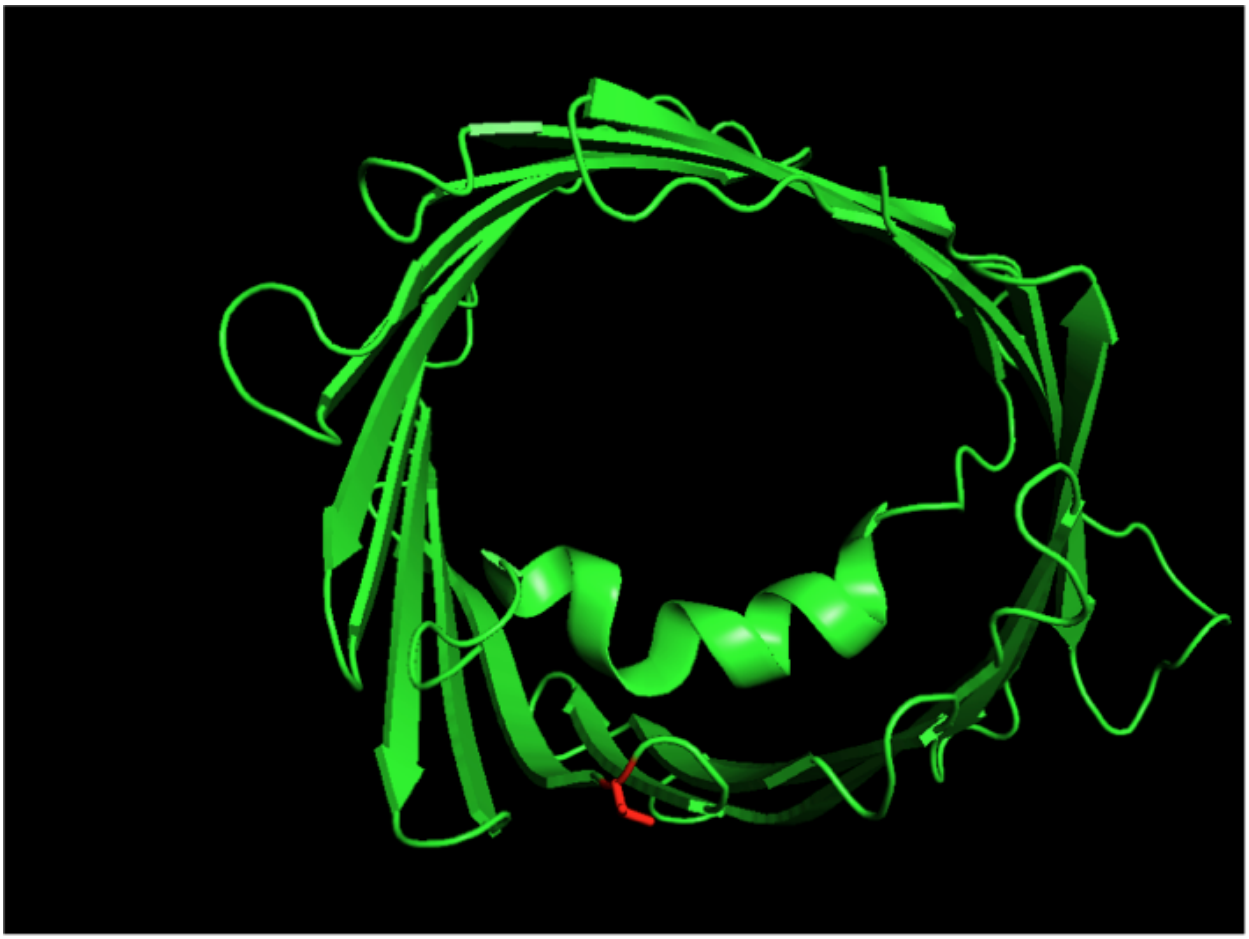


Figure 3.7 Porin is an *in vitro* LRRK2 substrate.
The structure of the human ortholog in mouse (PDB ID: 3EMN) protein is shown. The site of possible in vitro phosphorylation is shown in red.

In vitro detection of LRRK2-G2019S substrates: Application of an internal peptide substrate

To follow-up with additional evidence for a possible ATPase as LRRK2 substrate and to identify additional substrates, I devised a series of *in vitro* kinase assays on peptide extracts from two PD models: *Drosophila* peptides from head protein extracts and human tryptic peptides from SH-SY5Y neuroblastoma cells. These are systems others have used in addressing substrates of LRRK2[139-141]. This approach takes advantage of the synthetic "nictide" peptide substrate as an internal positive control, since nictide is a sensitive readout of LRRK2 kinase activity. I employed the hyperactive LRRK2-G2019S kinase in a label free approach with an automated IMAC phosphopeptide enrichment protocol to minimize the variability between replicates and expanded the number of replicates (n=10). **Figure 3.8** provides an experimental overview where two conditions were compared in a label-free *in vitro* kinase assay with nictide against

tryptic peptides. The main comparison for candidate substrates consisted of quantitatively comparing enriched phosphopeptides from a reaction with LRRK2-G2019S (left side of **Figure 3.8**) or without LRRK2-G2019S (right side of **Figure 3.8**). Alongside these enrichments, the sample was analyzed prior to phosphopeptide enrichment, to asses the kinase activity on nictide by measuring the relative signal intensity of nictide to phosphonictide.

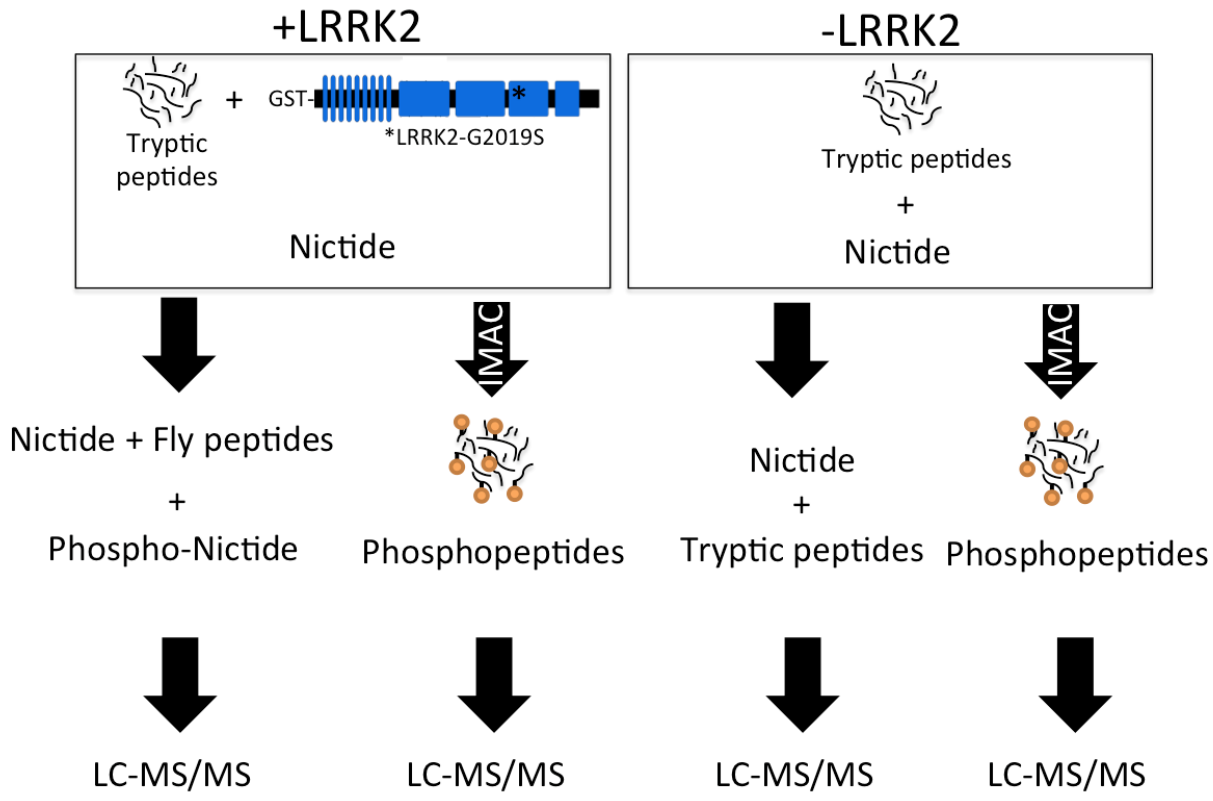


Figure 3.8 Experimental approach for *in vitro* kinase assays with GST-LRRK2-G2019S.
Tryptic peptides from either *Drosophila melanogaster* head protein extract or SH-SY5Y neuroblastoma cell protein extract were reacted with or without GST-LRRK2-G2019S. Pre-IMAC samples were analyzed for each to detect phosphonictide as a positive control of GST-LRRK2-G2019S activity.

Activity of LRRK2-G2019S on a model peptide substrate measured by mass spectrometry

First, I addressed whether nictide could detected by LC-MS/MS analysis. The relative intensities of each charge state reveal that +5 nictide has the greatest intensity in the mass spectrometer, and is therefore the most salient ion **(Figure 3.9).** I also discovered that the peptide sequence isn't amenable for a database search as it rarely produces a discernible fragmentation pattern for MS2 interpretation. Going further, it was also determined through trial and error that attempting to inject 1fmol of nictide into the mass spectrometer results in an undetectable signal due to the loss of nictide in sample handling and

stagetip desalting steps. Therefore, I began to optimize the nictide by designing experiments that would result in the injection of about 50pmoles of nictide to readily detect the peptide. This initial test also confirms that the nictide used in the following experiments did not contain contamination of phosphorylated nictide. The commercial nictide aliquot was deemed to only have the non-phosphorylated nictide.

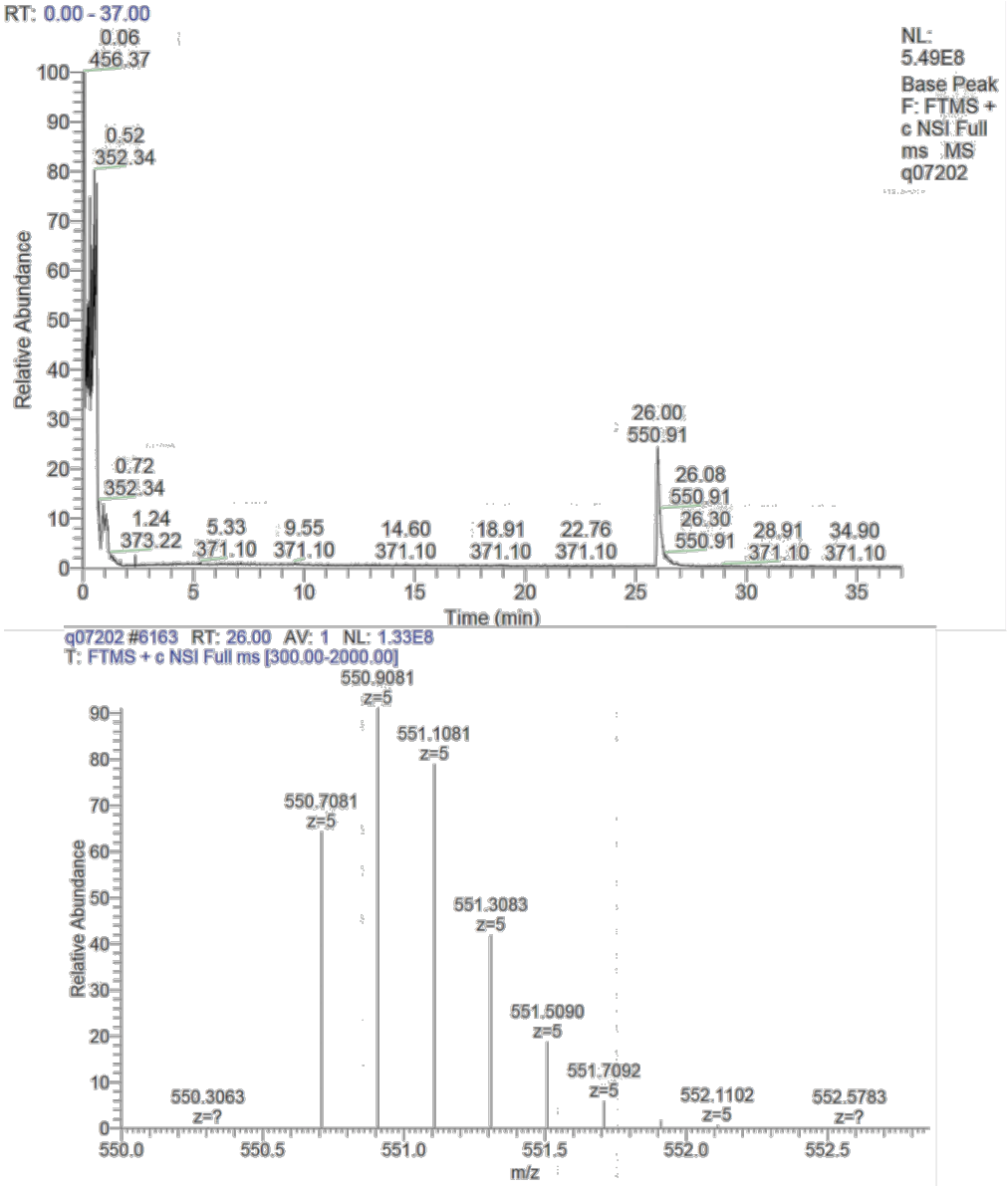
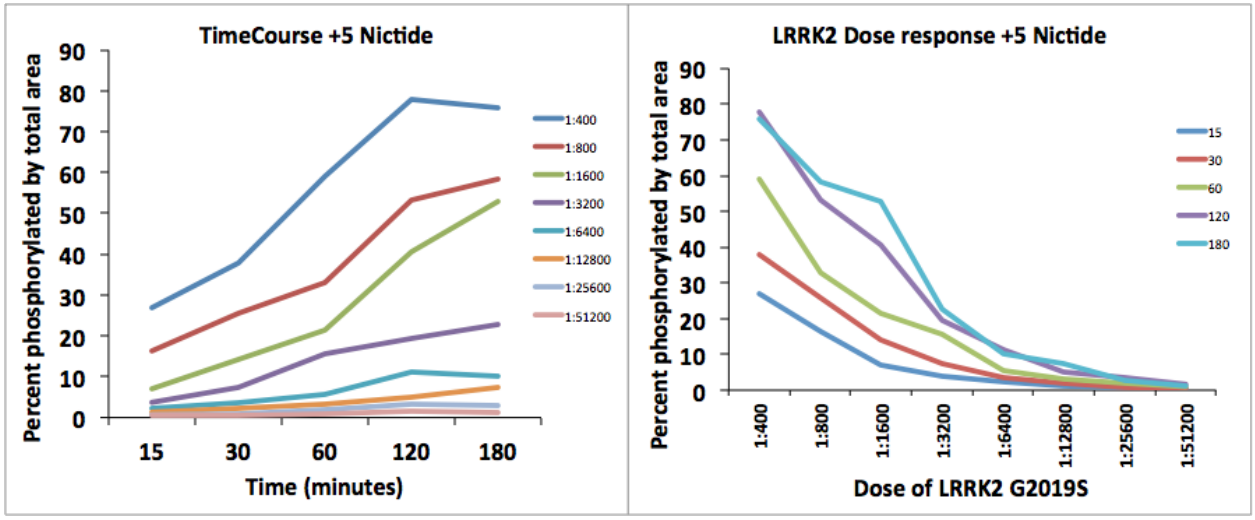


Figure 3.9 Nictide is detected by mass spectrometry
Top: Intact nictide shown eluting at 26 minutes from a 37 minute gradient. Bottom: isotopic envelope of the +5 version of the nictide.

I next tested the phosphorylation of nictide by GST-LRRK2-G2019S to find the kinase's optimal concentration and optimal reaction time **(Figure 3.10)**. I aimed to find the conditions where ~20% phosphorylation occurred of the total nictide peak area. Time points 15,30,60,120 and minutes show an increase from 15 to 120 minutes and a leveling off at 180 minutes. The apparent decrease by some of the curves at 180 min are likely due to small changes from the preceding time point, slight variability in the quantification of the peak area, or both. These results suggested that the use of a 1:800 GST-LRRK2-G2019S:Nictide ratio at 30 minute time point would be effective in the detection of phospho-nictide in a background of tryptic peptides from fly or neuroblastoma cells (PD model peptides).

I also analyzed a kinase reaction on nictide via a parallel reaction monitoring (PRM) assay to observe the intensity and quality of nictide and phosphonictide peptides as they would be used in the



kinase assay with the added phosphopeptide enrichment on PD model peptides.

Figure 3.10 Nictide is phosphorylated in a time and dose-dependent manner
Left: a time course of different LRRK2 concentrations (LRRK2:nictide). Right: same experiment viewed as a function of the LRRK2:nictide ratio.

The phosphonictide ions were clearly detected eluting adjacent to the nictide peaks in the gradient resulting in about 30% phosphorylation of nictide **(Figure 3.11).** Upon closer inspection with Skyline to analyze the spectra, the +5 phosphonictide was readily detected at minute 17.1 in the 30 minute gradient with many fragment ions also detected **(Figure 3.12)**. This result indicates that fragmentation of the

nictide is possible with a PRM assay and could be amenable for database searches using Comet, provided the appropriate spectra are detected.

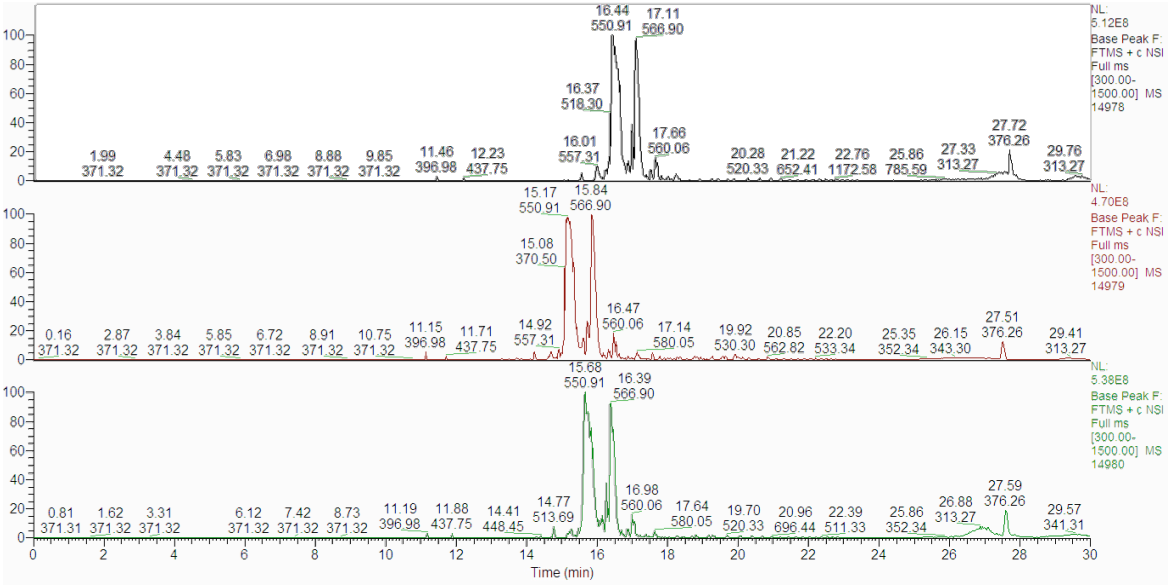


Figure 3.11 Chromatograms of nictide phosphorylation in 1:800 (E:S) ratio, 30 minute reaction. Nictide peptide and phosphonictide are represented in peak labeled 550.91 and 566.90, respectively. Each panel represents a replicate phosphorylation experiment.

Next, I tried to determine if a nictide: LRRK2-G2019S kinase assay could produce detectable phosphopeptides after a phosphopeptide enrichment with immobilized metal affinity chromatography (IMAC). In order to confirm the activity of LRRK2 in an *in vitro* assay with a background of tryptic peptides, the phosphorylated nictide peptide would have to be detected among other peptides in a tryptic peptide pool. I performed a series of *in vitro* kinase assays and compared the recovery of phosphonictide before and after IMAC. Phosphonictide is more reliably detected in the IMAC flow-through **(Figure 3.13 C)** compared to after IMAC **(Figure 3.13 A),** or with a targeted assay **(Figure 3.13 B).** The signal in the pre-IMAC sample was about two orders of magnitude greater than the IMAC enriched sample. This result suggests that the phosphonictide is not amenable to IMAC enrichment and must be detected prior to IMAC **(Figure 3.13)**. These data confirm that the detection of nictide and phosphonictide is possible in a background of tryptic peptides, and enabled the application of nictide to a larger number of replicates in

both fly and neuroblastoma *in vitro* kinase assays. The following label free experiments confirmed the activity of GST-LRRK2-G2019S in pre-IMAC samples.

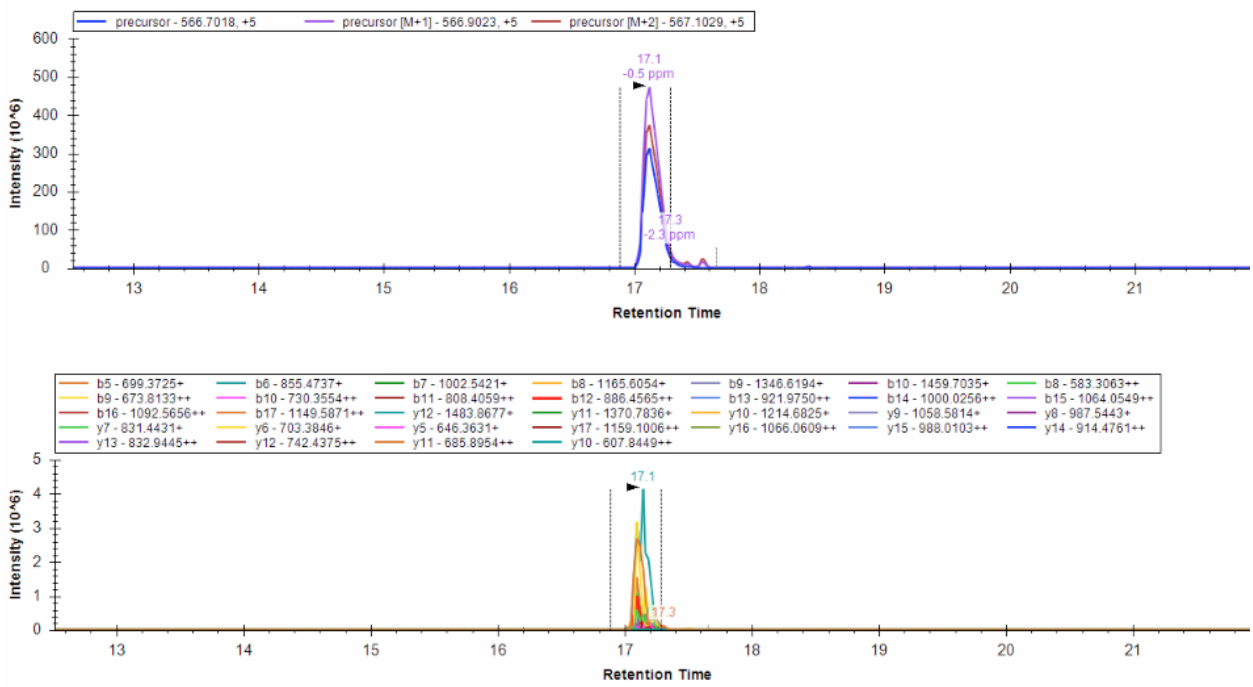


Figure 3.12 Extracted phosphonictide ion chromatograms viewed in Skyline.
Top: +5 peptide extracted ion chromatogram. Bottom: fragment ion from nictide precursor PRM fragmentation.

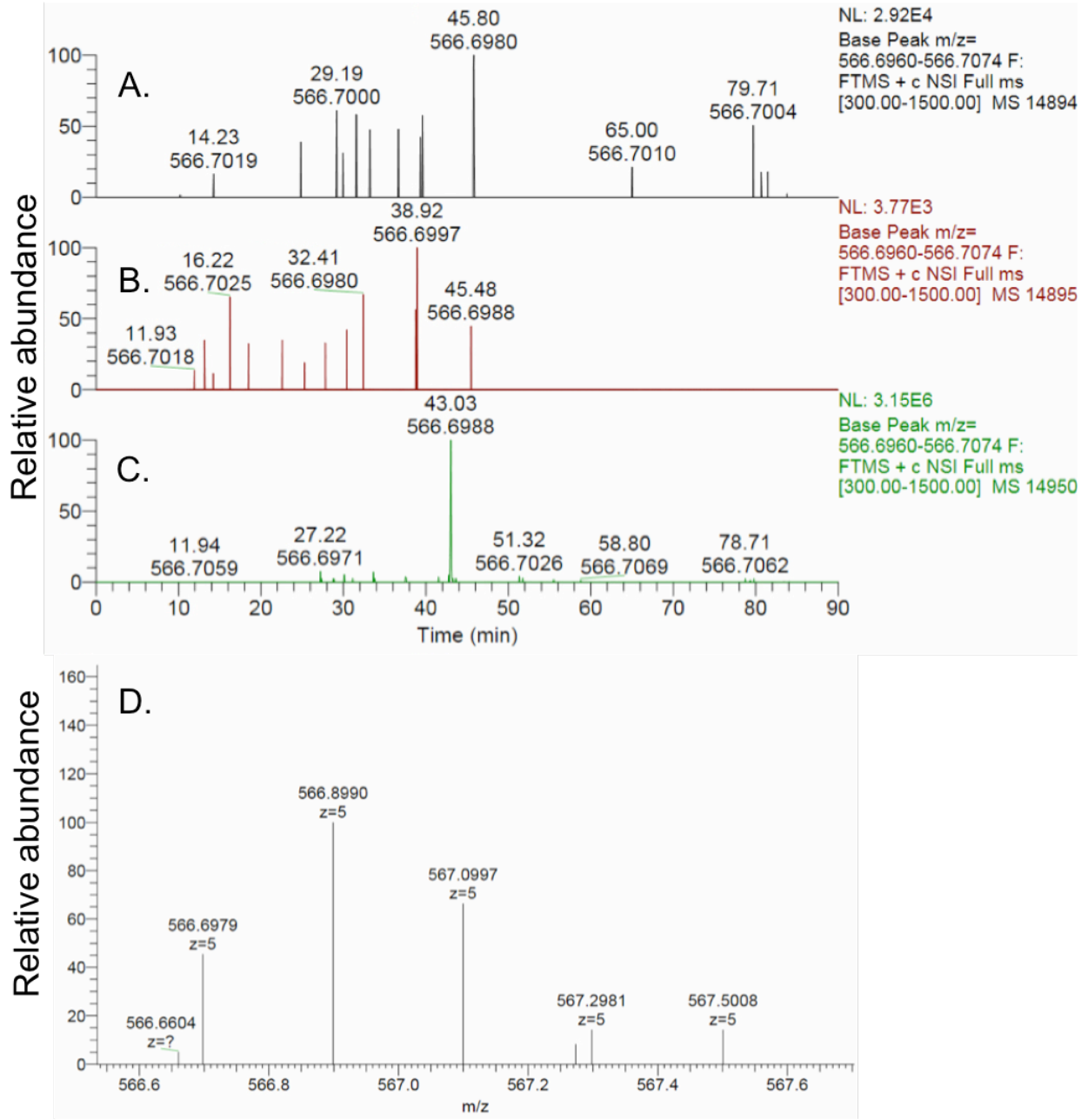


Figure 3.13 Phosphonictide recovery after IMAC compared to pre-IMAC.
(A) Phosphonictide in the presence of tryptic fly peptides with a phosphonictide mass filter; m/z=566.7017. (B) Same as top panel in a PRM assay to target the phosphopeptide mass. (C) A pre-IMAC analysis of the same sample. (D) Isotopic envelope of the phosphonictide corresponding to the appropriate charge state of +5.

Detection of LRRK2-G2019S substrates from whole fly head tissue and neuroblastoma cells

Label-free in vitro assay on fly peptides:

In order to increase the signal from potential substrates, I employed the GST-tagged LRRK2-G2019S hyperactive kinase mutant in a series of *in vitro* experiments. I compared reactions on fly head tryptic peptides with or without GST-LRRK2-G2019S in the presence of nictide. In parallel, a third reaction provided additional evidence of an active LRRK2 reaction through a positive control without PD model peptides **(Figure 3.9-1).** I calculated label-free quantification (LFQ, based on peak areas) values for peptide phosphorylation after automated IMAC on about 190µg of fly head peptides in 10 replicates per condition.

I identified 23,395 (S.D. = 632.57, n =20) total unique peptides and successfully quantified 22,536 peptides (859 missing quantifications) which resulted in 14,998 label-free phosphopeptide quantifications of peptides with quantifications in both +kinase condition and the -kinase condition (represented as $\log_2$ transformed [+kinase/-kinase] ratios). The average phosphopeptide enrichment was about 92% (S.D. = 2.7). The data represent 9,869 total unique phosphorylation sites and 43 candidate LRRK2 substrates (p-value=<0.01, $\log_2$[+kinase/-kinase] >=0.5). The median LFQ coefficient of variation was (CV) = 0.26 and 0.24 for the control and kinase treated phosphopeptide quantifications, respectively.

The nictide spike-in was detectable throughout the experiment in the presence of peptides (data not shown). The pre-IMAC samples reveal evidence of nictide (not shown) and phosphonictide, although the signal was very low in many cases. For example, **Figure 3.14** shows +kinase (replicate 1) with a peak at retention time (RT) 35.83 mz= 566.7091 with corresponding +5 charge state and isotopic envelope likely corresponding to the phosphonictide species. Similarly, the parallel positive control **(Figure 3.15)** reveals a high signal, further confirming the conditions were favorable for GST-LRRK2-G2019S activity. The degree of phosphorylation is estimated to be about 27% as calculated by the area of the phosphonictide peak divided by the summed area of nictide and phosphonictide species.

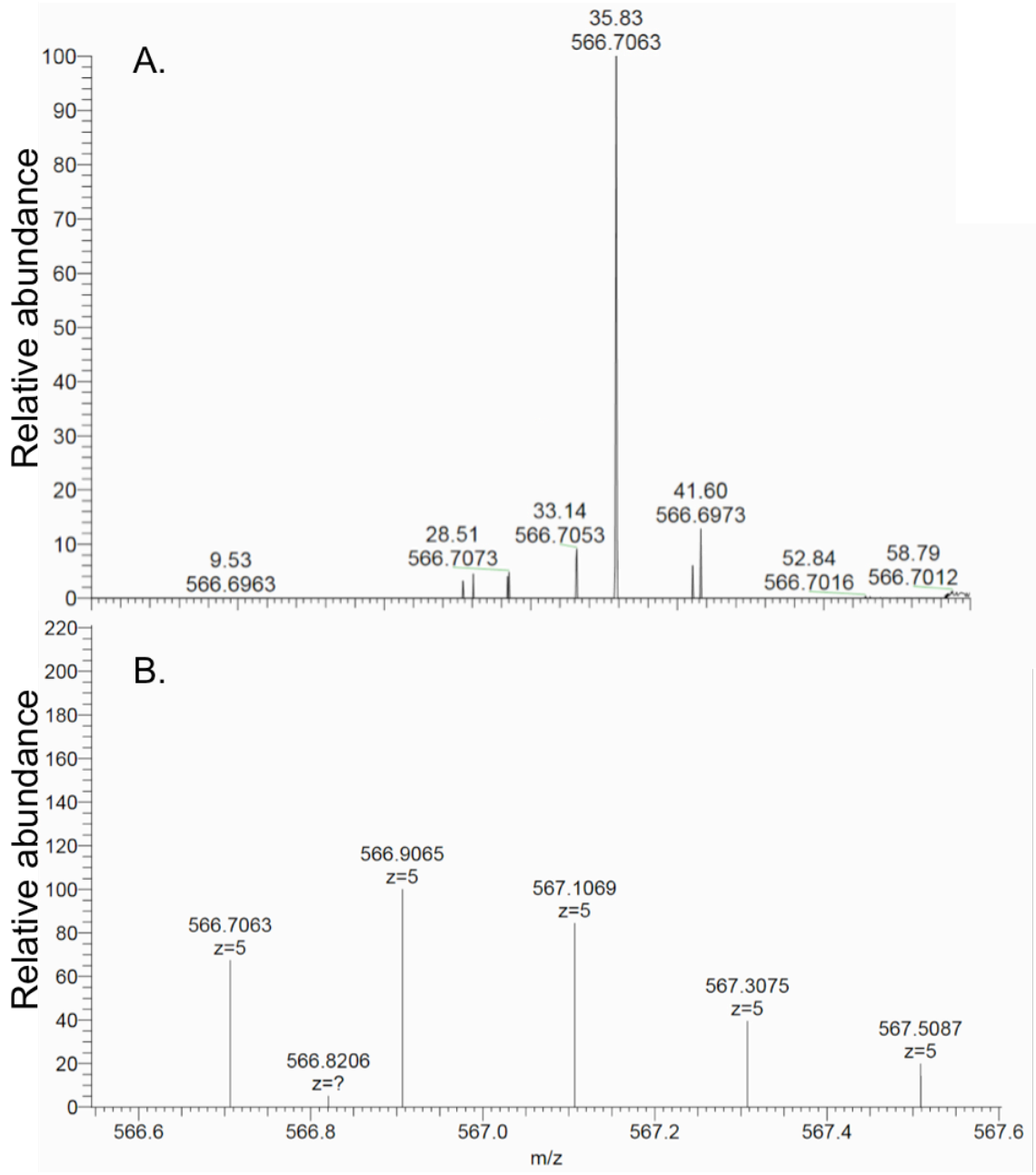


Figure 3.14 *In vitro* kinase reaction on fly peptides: spike-in nictide is phosphorylated.
(A) Chromatographic peak of phosphonictide eluting at 35.83 minutes. (B) MS1 of +5 phosphonictide isotopic envelope.

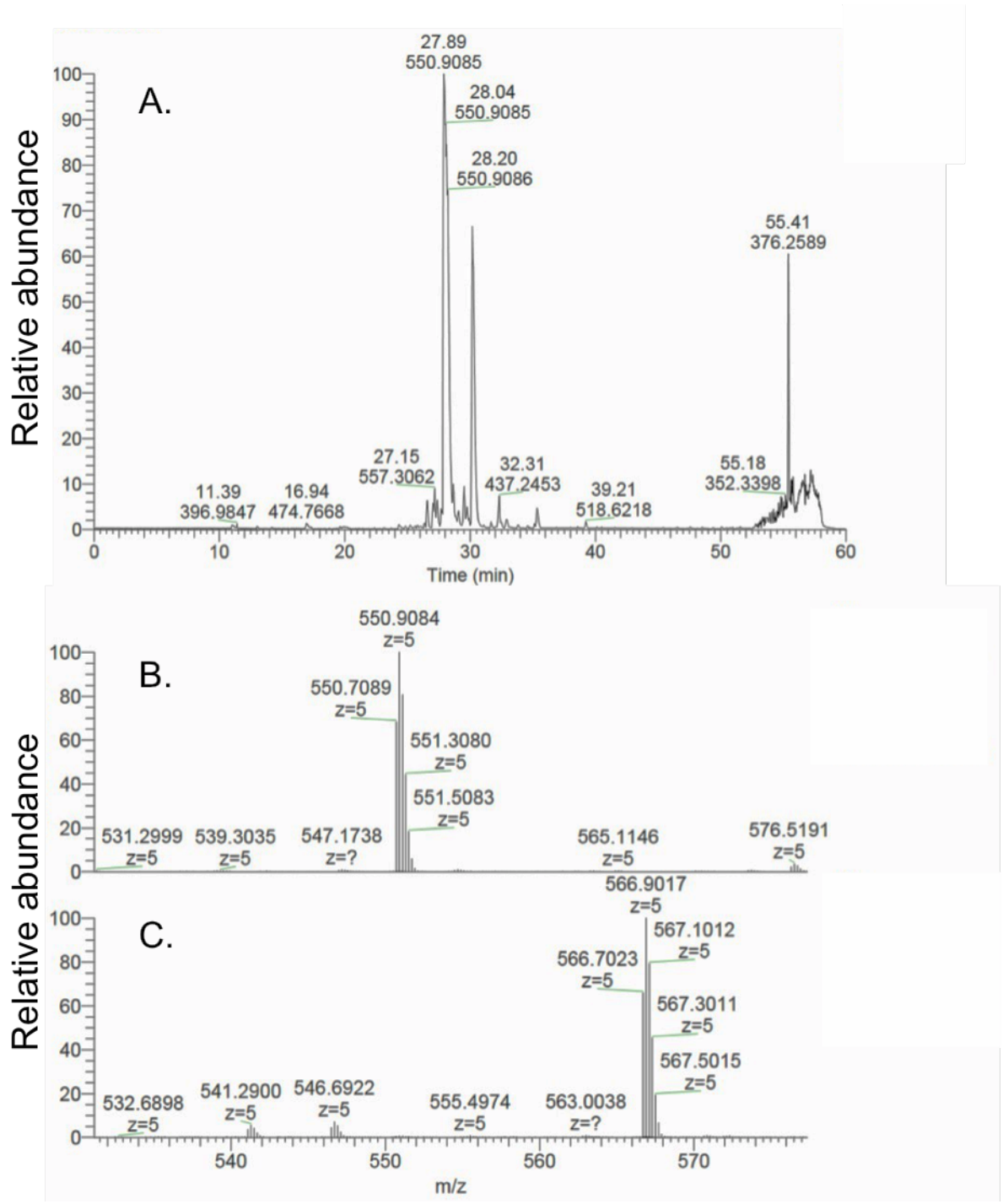


Figure 3.15 MS evidence for phosphonictide in parallel control reaction without fly head peptides. (A) LC-MS/MS chromatographic peaks of nictide and phosphonictide eluting at about 27.89 minutes and 30.2 minutes, respectively. (B) MS1 of +5 nictide isotopic envelope. (C) MS1 of +5 phosphonictide isotopic envelope.

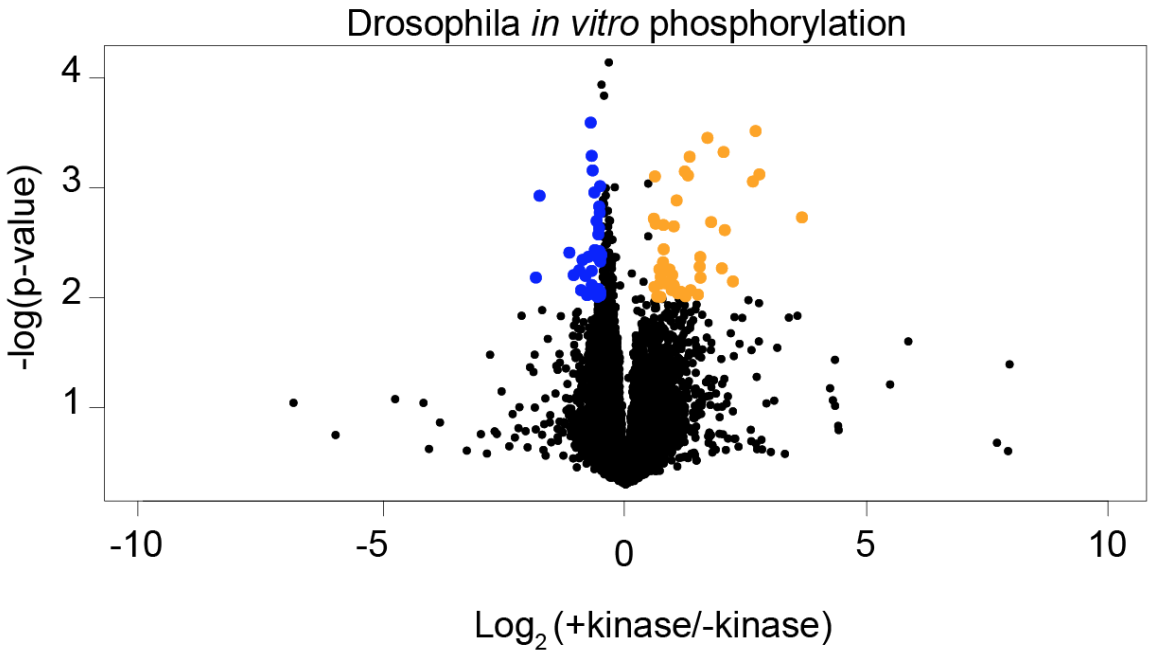


Figure 3.16 Volcano plot of LFQ phosphopeptides from fly peptide kinase assay
IMAC replicates (n=10). Orange: Log_2(+kinase/-kinase) >=0.5, p<0.01 (t-test). Blue: Log_2(+kinase/-kinase) <=-0.5, p<0.01.

Relatively few changes were observed given the filtering criteria **(Figure 3.16)**, with more positive changes in the kinase-treated condition, as expected. **Table 3.4** shows a subset of the peptides that were changing the most by at least a 2.5-fold change. Both serine and threonine phosphorylated residues were found in this list.

I next analyzed the data to find representative gene ontology in the data. Inputting the changes in orange, against the rest of the proteins represented in the volcano plot in Figure 3.17 resulted in a gene ontology enrichment of “calcium transport”, “divalent inorganic cation transporter activity” and “calcium-transporting ATPase activity”. Not represented in these data is the atpase, JYalpha T369. Although the site is represented in both replicate datasets, it was barely changing (Log2 average ratio = 0.09).

Other proteins in this list reveal candidate substrates involved in synaptic vesicle turnover. The S367 site on stnB was shown to be changing over 12-fold on a protein part of the endocytic complex. It is known to regulate synaptic vesicle dynamics, and is important for developmental stages of synaptogenesis[142,143]. EndoA, whose site S278 was changing over four-fold, is a known LRRK2 substrate for a different phosphorylation site at S75[121]. Previous human studies did not assay Drosophila S278 because it is not conserved with human EndoA, although it may serve as another phosphorylation

mechanism when observing impediments to endocytosis in the fly as observed in Matta *et al*. (2012)[121]. Further, Ank2 was phosphorylated over 6.5-fold, a mediator of axonogenesis, synaptic stability synaptic vesicle turnover[144].

Some phosphorylation sites were changing to a very high degree that did not pass the 0.01 filter, yet are noteworthy. A list of 20 quantified phosphopeptides with a log2(+kinase/-kinase)>=5 contains some candidate substrates not found in enough replicates for a p-value and three associated with them <=0.065. They include CG43340 T2415 (an uncharacterized gene), Mlc2 S67 and stnA S744. Again, a phosphorylation site of the stoned locus emerges as a possible candidate substrate, a protein which functions as a member of the endomembrane/ synaptic vesicle dynamics though the protein syt1, which was found earlier in this study (kinase assay on labeled peptides) [145].

Table 3.4 Candidate substrates from Drosophila head *in vitro kinase* assay
Phosphopeptides exhibiting >=2.5-fold change, p<0.01.

Ascore Sequence_Charge	Flybase polypeptide entry	Sites	Symbol	Average Log2 Ratio	Example Human Ortholog
AAAEDKAAAADAAGDAADNGT#SKDGEDAADAAAAAPAK_3	FBpp0307124	T54	Bacc	1.70	-
ARS#CT#PFPVTGYTYEPASQLALDAYVAR_3	FBpp0310815	S228,T230	CG45076	1.35	-
AS#SPVSEGTDSTTIPGPVPKPEIPIDIS#FDANAPVPKPR_4	FBpp0310629	S2627,S2653	Ank2	1.33	ANK2
AS#TPGS#EDDEDLLLGS#PR_2	FBpp0300236	S5081,S5085	futsch	1.54	MAP1S
DEPVFTSLLIRPDES#THDITSQPQAATGLER_3	FBpp0310530	S367	stnB	3.66	STON2
ELSPS#ATESELEEDAAIAEQSR_2	FBpp0291927	S305	Lmpt	1.50	FHL2
GLPS#PDGDKDRDPLLIDLFER_4	FBpp0300236	S427	futsch	2.23	MAP1S
LS#SEDKFKPSNLLETSESLDLLSGEIK_3	FBpp0309607	S176	Vap-33A	2.00	VAPA
NASPIPFGT#SLNEQTSWPPR_3	FBpp0305827	T781	CG1674	2.64	-
NLFDS#SLNLLPQQQQQHQPDLAK_3	FBpp0293846	S1117	pyd	2.06	TJP2
NNS#PTDVSVVNCGSSENMSNGTIVSHNVLNSSDAILEK_4	FBpp0310599	S190	CG40498	2.77	-
RLSETS#S#IEYADNIPDELTIPEIDVER_3	FBpp0293012	S1178,S1179	PMCA	2.70	ATP2B2
SITS#GGT#GGGASMLAAAALR_3	FBpp0084879	S964,T967	trp	1.77	TRPC1
STS#PLPAIGVPAIAPASPIYATSTK_3	FBpp0291577	S348	by	2.03	-
TLLDLNLDGGGGGLNEDGTPS#HISSSASPLPSPMR_3	FBpp0306589	S278	EndoA	1.55	SH3GL3
TRLWPVIGRDS#GEEGASPTGDSDDS#SDGSGAGFPFGVRPLIPLDSK_4	FBpp0306803	S95,S109	vir-1	1.55	-

Label-free *in vitro* assay on neuroblastoma peptides:

As an orthogonal approach to find conserved LRRK2-dependent phosphorylation sites, I also included a search in the mammalian neuroblastoma model frequently used for studying LRRK2 biology. Here, I identified 22672 (S.D. = 763.08, n =20) total unique peptides and quantified 21,709 peptides which resulted in 13,354 label-free phosphopeptide quantifications (represented as $\log_2$ transformed [+kinase/-kinase] ratios). The average phosphopeptide enrichment was about 87.17% (S.D. = 13.39). This represent 9,045 total unique phosphorylation sites and 11 candidate LRRK2 substrates (p-value=<0.01 $\log_2$(+kinase/-kinase) >=0.5). The median LFQ CV was 0.23 and 0.22 for the control and kinase treated conditions, respectively. The nictide spike-in was ubiquitously detected in all samples.

Phosphonictide was also readily detectable in all +kinase replicates **(Figure 3.17)** and parallel positive control samples **(Figure 3.18).** Here, I achieved about a 21% nictide phosphorylation calculated as in the fly counterpart to this experiment.

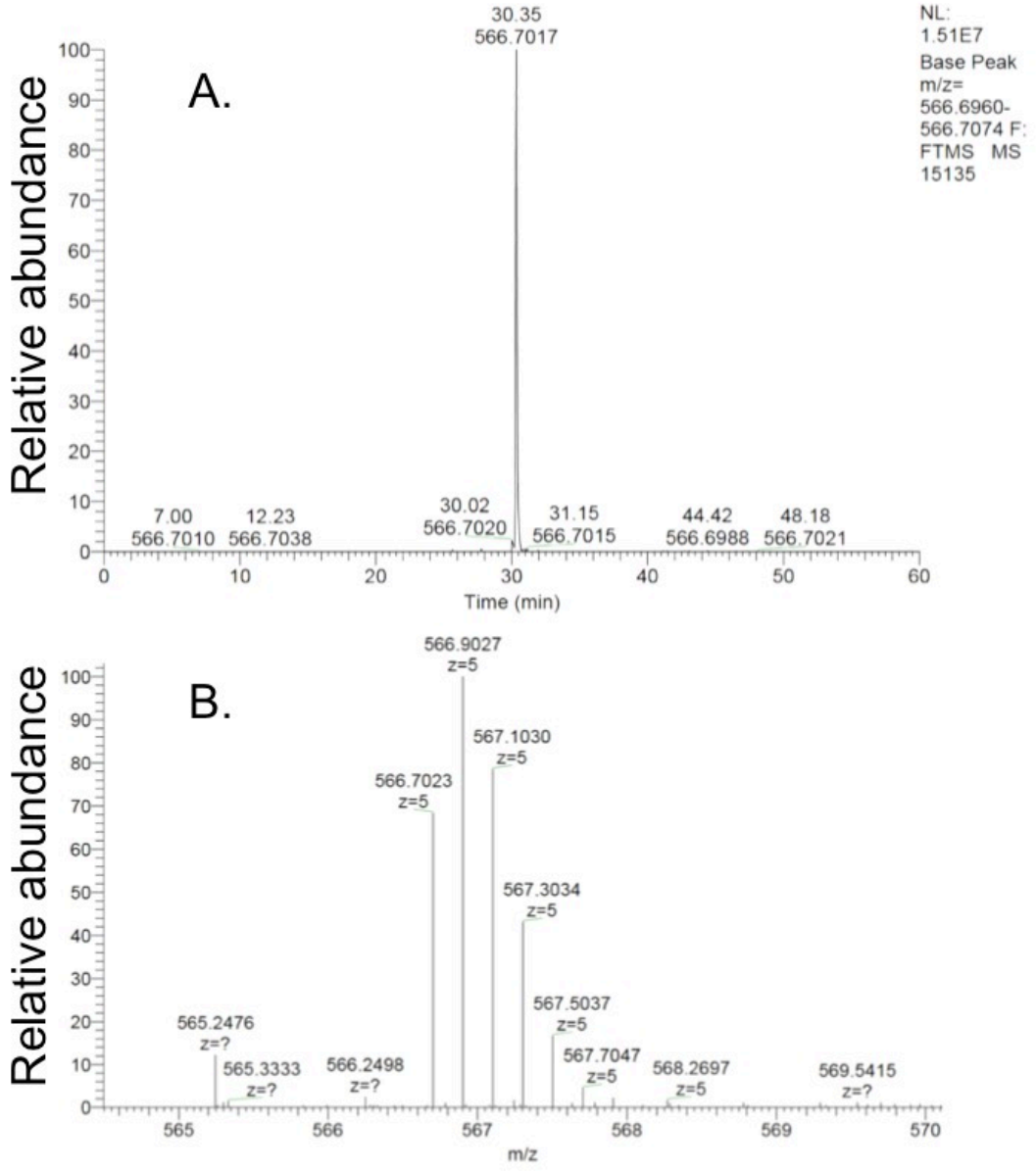


Figure 3.17 *In vitro* kinase reaction on SH-SY5Y neuroblastoma peptides: spike-in nictide is phosphorylated.
(A): chromatographic peak of +kinase replicate 1 phosphonictide eluting at 30.35 minutes. (B): MS1 of +5 phosphonictide isotopic envelope.

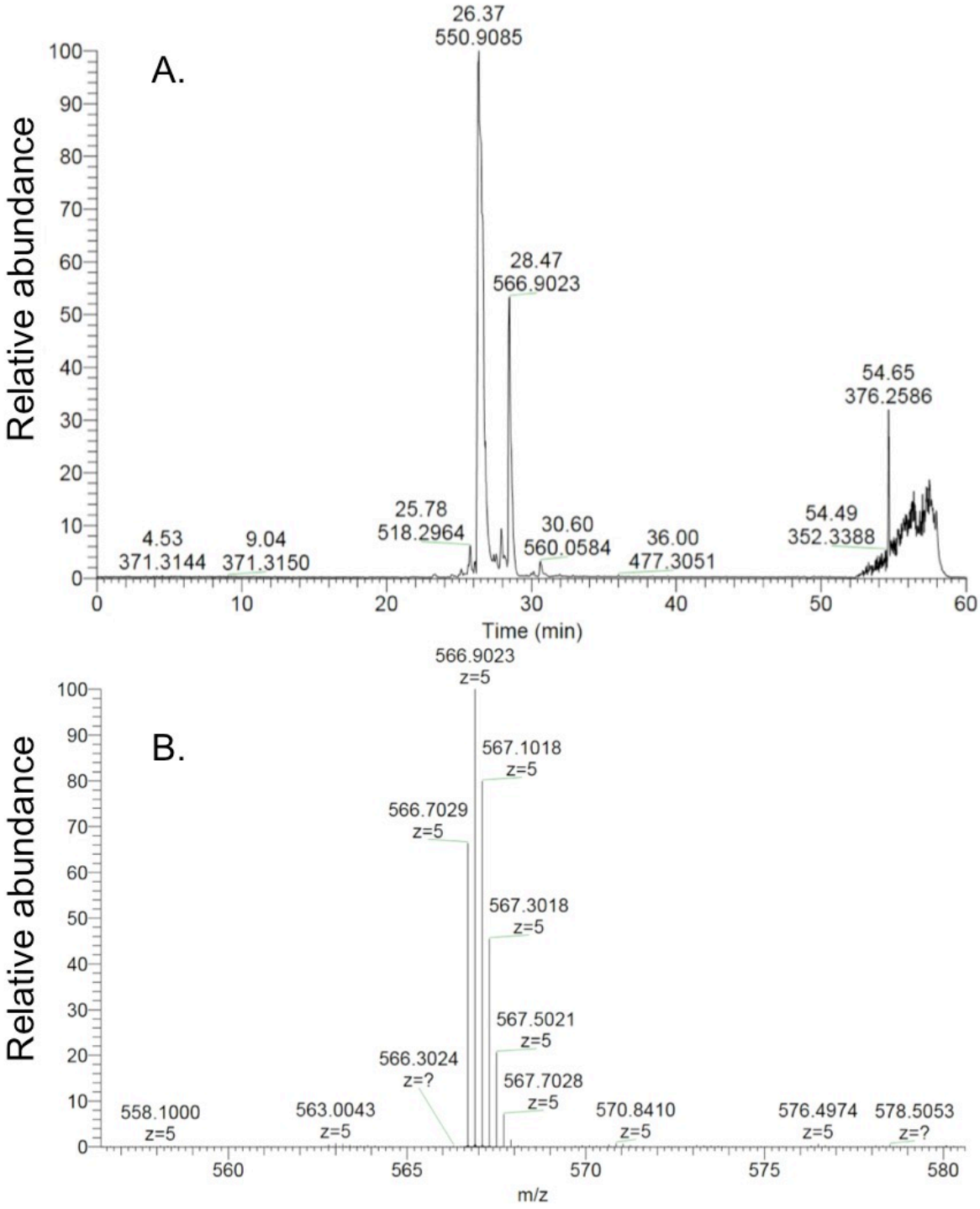


Figure 3.18 MS evidence for phosphonictide in parallel control reaction without SH-SY5Y peptides. Top: LC-MS/MS chromatographic peaks of nictide and phosphonictide eluting at about 26.37 minutes and 28.47 minutes, respectively. Bottom: MS1 of +5 phosphonictide isotopic envelope.

Gene ontology enrichment for peptides in **Figure 3.19** (orange points) as performed above resulted in the enrichment of several functional groups implicated in enhancer sequence-specific DNA

binding like HNRNPC, HDAC2, NIPBL, SMARCC2, LDB1 and. HNRNPC S299 was changing almost 3-fold, and is a protein involved in modulating mRNA stability. Other features in the 11 changing phosphopeptides include phosphorylation mapping to microtubule-associated protein-2 (MTAP2) which exhibited over a 3.5-fold change. This protein has been shown to associate with Lewy bodies previously, and other microtubule associated proteins play a strong role in LRRK2 biology. LRRK2 has been shown to negatively regulate MAP1B and its fly homog, futsch. The microtubule associated protein futsch is a well-known LRRK2 interactor in the fly model, and has been shown to be a LRRK2 substrate. Both of these proteins in their respective models demonstrate the ability to rescue LRRK2-G2019S hyperphosphorylation phenotypes such as cytotoxicity and apoptosis[146,147].

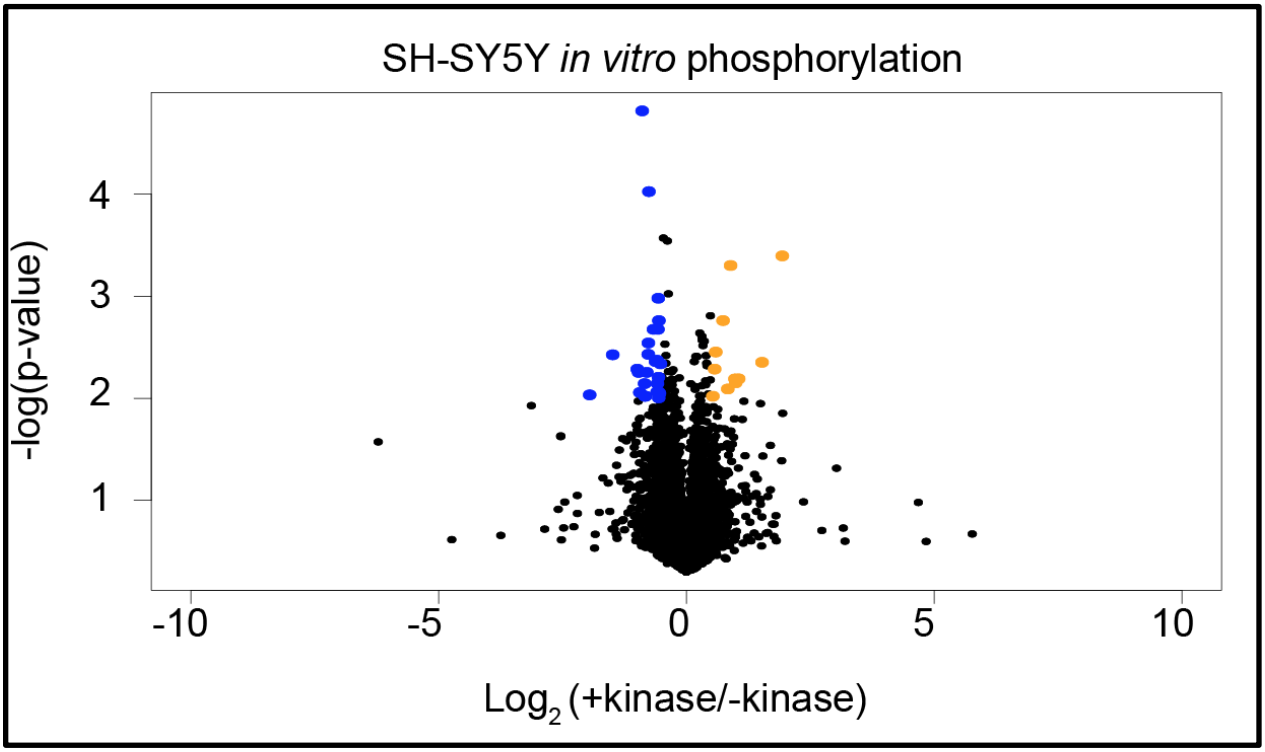


Figure 3.19 Volcano plot of LFQ phosphopeptides from kinase assay IMAC replicates. Orange: Log_2(+kinase/-kinase) >=0.5, p<0.01 (Ttest). Blue: Log_2(+kinase/-kinase) <=-0.5, p<0.01.

Table 3.5 Phosphopeptide hits from SH-SY5Y protein extract *in vitro kinase* assay Phosphopeptides exhibiting >=2.5-fold change, p<0.01. Asterisk(*) indicates oxidized methionine, "#" indicates phosphorylation site after residue.

Ascore Sequence_Charge	Uniprot Entry	Sites	Entry Name	Average Log2 Ratio
APQQQPPPQQPPPPQPPPQQPPPPPS#YSPARNPPGASTYNK_4	Q9BUJ2	S716	HNRL1	0.98
DDEKEAEEGEDDRDS#ANGEDDS_2	P07910	S299	HNRPC	1.53
GESRPET#PK_2	Q6KC79	T713	NIPBL	0.53
LYS#PK_2	Q86U70	S19	LDB1	0.84
M*LPHAPGVQM*QAIPEDAVHEDS#GDEDGEDPDKR_4	Q92769	S394	HDAC2	0.74
QNALLFAEEEDGEAGAEDKRSQEET#PGHR_4	O15240	T424	VGF	1.06
RKS#VPSETVVEDSR_2	P11137	S833	MTAP2	1.94
RS#PS#PSPTPEAK_2	Q8TAQ2	S302,S304	SMRC2	0.99
SLS#DSESDDSK_2	Q13185	S95	CBX3	0.57
TQPHHST#PTK_3	Q12789	T514	TF3C1	0.89
VNS#TTKPEDGGLKK_3	Q9Y4B6	S255	VPRBP	0.59

Motif analysis for fly and neuroblastoma cells

No motifs could be discerned from either dataset (orange points in volcano plots) using motif-X, even under the least stringent criteria. A few peptides do exhibit the motif described by[134]. In the fly peptide experiment CG1674 T781 shows over a six-fold increase adhering to the F/Y-X-T-X partial sequence, which belongs to a gene ontology group of Z disc function; a muscle sarcomere functioning in the attachment of actin filaments. Another interesting feature of these data come from a the Bacc T54 site where another partial match to X-T-X-K. The only known feature of this protein from mutant phenotypes is that of inter-male aggressive behavior and behavioral response to ethanol. Similar motifs from the neuroblastoma experiment were not observed in this filtered subset.

Neuroblast model reveals phosphorylated proteins associated with lewy bodies

Of LRRK2 related PD cases, LRRK2-G2019S is more frequently associated with Lewy body (LB) pathology than other LRRK2 mutations, suggesting a link between hyperphosphorylation and LBs[148]. Although the neuroblast-model candidate list did not have clear overlap in phosphorylation sites with the in-vivo data or the *in vitro* fly model data. I sought to explore the connection between LB-associated proteins and the increasing phosphopeptides detected.

Table 3.6 Phosphopeptides mapping to Lewy body-associated proteins.
Phosphopeptides changing Log_2(+kinase/-kinase) ≥ 0.5 and found in one or more replicates corresponding to proteins described in Lewy bodies[149]. Asterisk(*) denotes non-localized phosphorylation site Ascore<13.

Name	Uniprot ID	Ascore Sequence_charge	Sites	Average Log2 ratio	p-value
TAU	P10636	SGYSSPGS#PGT#PGSR_2	S519,T522	1.01	-
TAU	P10636	AKTDHGAEIVY#KSPVVSGDTS#PR_4	Y711*,S721*	0.74	-
SQSTM	Q13501	EVDPSTGELQSLQMPES#EGPSSLDPSQEGPTGLK_3	S361	3.48	-
MTAP2	P11137	S#GTSTPTTPGSTAITPGTPPS#YSSR_3	S1588*,S1608*	2.20	-
MTAP2	P11137	RKS#VPSETVVEDSR_2	S833	1.94	0.00
MTAP2	P11137	FAALEQPEVERRPS#PHDEEEFEVEEAAEAQAEPK_4	S1324	1.38	-
MTAP2	P11137	KTEVQAHS#PSR_2	S1480	1.32	-
MTAP2	P11137	ES#QPSPPAQEAGYSTLAQSYPSDLPEEPS#SPQER_4	S626,S653	1.17	-
MTAP2	P11137	VDHGAEIIT#QSPGRSS#VAS#PR_3	T1780,S1787	0.97	-
MTAP2	P11137	SGTSTPT#TPGSTAITPGT#PPSYSSR_2	T1594*,T1605*	0.93	0.21
MTAP2	P11137	T#PGT#PGTPSYPR_2	T1613,T1616	0.91	-
MTAP2	P11137	AGKSGTS#T#PTTPGSTAITPGT#PPSYSSR_3	S1591*,T1592*	0.90	0.08
MTAP2	P11137	ADEGKKETS#PESSLIQDEIAVK_4	S1155*	0.78	-
MAP1B	P46821	AAEAGGAEEQYGFLTTPT#K_2	T1069	7.77	-
MAP1B	P46821	TATCHSS#SSPPIDAASAEPYGFR_3	S1817*	2.36	-
MAP1B	P46821	EKTATCHSSSS#PPIDAASAEPYGFR_3	S1819*	1.64	-
MAP1B	P46821	ASVS#PM*DEPVPDS#ESPIEK_3	S1378*,S1387*	1.28	-
MAP1B	P46821	AS#VSPM*DEPVPDSES#PIEK_2	S1376,S1389	1.11	-
MAP1B	P46821	VSAEAEVAPVSPEVT#QEVVEEHCASPEDK_4	T1302*	1.03	-
MAP1B	P46821	QGVDDIEKFEDEGAGFEESS#ETGDYEEK_3	S937*	0.93	0.03
MAP1B	P46821	SLMSS#PEDLTKDFEELKAEEVDVTK_3	S832*	0.88	0.03
MAP1B	P46821	SEQSSMSIEFGQESPEQS#LAMDFSR_3	S1657*	0.88	-
MAP1B	P46821	TTKS#PSDSGYSYETIGK_3	S1915*	0.87	0.28
MAP1B	P46821	S#LMSS#PEDLTKDFEELKAEEVDVTK_3	S828,S832	0.75	0.19
MAP1B	P46821	TATCHS#SSSPPIDAASAEPYGFR_4	S1816*	0.60	-
MAP1B	P46821	TLEVVS#PSQSVTGSAGHTPYY#QSPTDEK_3	S1322*,Y1337*	0.58	-
MAP1B	P46821	VSPSKSPS#LSPSPPS#PLEK_3	S1258*,S1265*	0.57	-
MAP1B	P46821	ESS#PLYSPT#FSDSTSAVK_3	S1793*,T1799*	0.57	0.33
MAP1B	P46821	VQSLEGEKLS#PKS#DISPLT#PR_2	S1779,S1782	0.56	0.08
MAP1B	P46821	DSISAVSSEKVS#PSKS#PSLSPS#PPSPLEK_3	S1252*,S1256*	0.55	0.04
MAP1B	P46821	QS#PDHPTVGAGVLHITENGPTEVDY#SPSDMQDSSLSHK_4	S1666,Y1689	0.54	-
1433T	P27348	DNLTLWTS#DSAGEECDAAEGAEN_2	S230*	1.94	0.01

I searched for phosphopeptides mapping to proteins associated with Lewy bodies to add insight into *in vitro* LRRK2-G2019S activity on peptides. 397 phosphopeptides (multiple charge states) were found mapping to 15 proteins involved in LBs. 32 phosphopeptides were increasing in the +kinase condition (Log_2(+kinase/control) = ≥0.5) corresponding to three microtubule-associated proteins: MAP1B, MAP-2 and TAU. Additionally, a non-localized site (Ascore ≤13) on 14-3-3 protein theta, and a site on sequestosome-1 **(Table 3.6)**. The most represented protein on the list of Lewy body associated protein are phosphopeptides corresponding to the microtubule-associated protein MAP1B— a protein well known to be involved in the development of the nervous system, but recent studies demonstrate its role in axonal elongation and neuronal growth, axonal guidance and many functions related to both Parkinson's and Alzheimer's disease protein aggregation[150]. Microtubule-associated protein 2, MAP-2, was the second most represented protein on this list with 9 localized phosphorylation sites on six unique phosphopeptides. MAP-2 is a cytoskeletal component known for its role in the development of neurites and its role in regulating neuronal cytoskeletal function[151]. TAU protein was represented by potentially

novel LRRK2-mediated phosphorylation sites including S519, T522 (Ascore ≥ 13), Y711 and S721 (Ascore≤13). Known LRRK2 sites were not identified here (T181, T149, or T153) [117,118]. Sequestosome 1 protein (p62) S361 was found to be phosphorylated list with a fold-change over 11-fold over the -kinase control. This site has not been previously described by other studies, although its is well-studied for its role in oxidative stress, protein turnover, neurodegenerative diseases and contribution to protein aggregates[152]. Lastly, 14-3-3 protein at was represented by one site at S230, changing over 3.8-fold, but was not localized. Interestingly, its involvement in Lewy body pathology and apoptotic mechanisms has been well established[153].

***In vivo* detection of LRRK2 substrates and regulated protein**

To detect *in vivo* substrate candidates by querying the proteome and phosphoproteome of a living system, I used the GAL4/UAS system to generate transgenic *Drosophila* expressing human wild-type LRRK2 (hLRRK2) in the whole fly brain. As previously described, the expression of LRRK2 in *Drosophila* exhibits features of LRRK2-linked human Parkinsonism such as neuronal loss, locomotor dysfunction and early mortality[111]. The advantage here is that the fly model can provide LRRK2-activity in a biological context to reveal direct and indirect LRRK2-induced phosphorylation events, and regulation of protein expression through quantifying protein abundances. To my knowledge, this is the first time MS-based proteomic and phosphoproteomic data, has been generated in this manner for this model.

In vivo data were obtained by digesting fly head proteins from LRRK2 expressers and non-expressers, and labeling extracted peptides heavy or light with reductive dimethylation, respectively **(Figure 3.20)**. Three replicates of stage-tip reverse-phase basic fractions were produced, for 5 fractions per replicate alongside immobilized metal-affinity chromatography (IMAC) enrichment performed twice on the same sample per replicate, for a twice-enriched sample. This two-step enrichment was done to increase the amount of phosphopeptides enriched in the final sample. A second pair of replicates were subject to offline strong-cation exchange (SCX) chromatography where I collected 7 (SCX-replicate 4) and 8 fractions (replicate 5) for protein level analysis and IMAC enrichment. In total, five replicates were analyzed for protein level quantification and quantification of phosphopeptides.

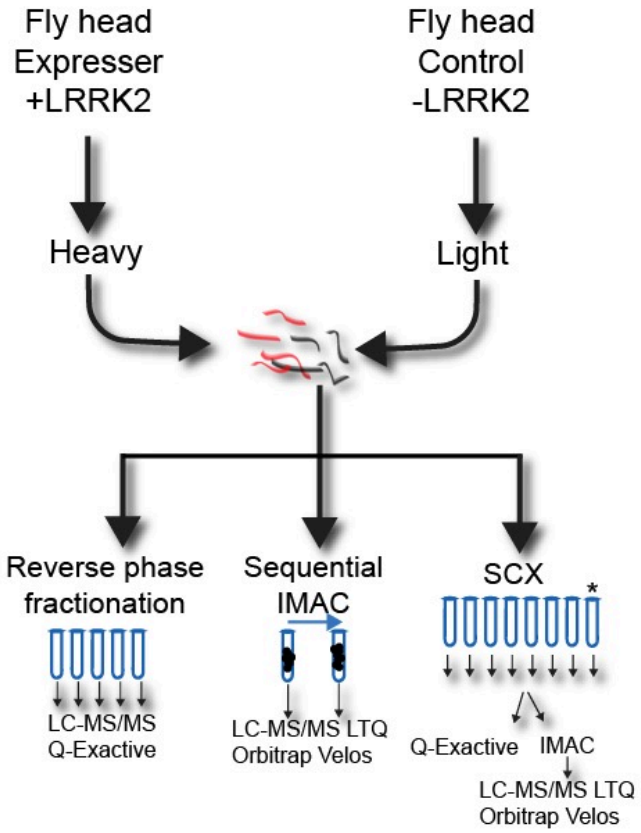


Figure 3.20 Experimental outline for *in vivo* fly head protein analysis and phosphopeptide analysis performed in triplicate.
Fly heads expressing human LRRK2 vs. non-expressers provided tryptic peptides that were labeled with reductive dimethylation chemistry. hLRRK2 peptides from human expressing fly heads were labeled heavy. Control peptides were labeled light and mixed for downstream analysis. Peptides were split into three groups: 1) Peptides were fractionated using reverse phase basic fractionation into 5 fractions. 2) Peptides were enriched for phosphopeptides using IMAC. 3) Peptides were fractionated using strong cation exchange chromatography for subsequent analysis on the Q-exactive instrument or enriched for phosphopeptides with IMAC and analyzed on an LTQ Orbitrap Velos. 1 and 2 were performed in triplicate, 3 was performed in duplicate. Asteriskn represents 1 replicate with 1 omitted fraction.

Protein- level analysis of in vivo LRRK2-WT expression

The proteomic analysis resulted 6,111 identified proteins (FDR<1%) and 1,725 quantified proteins in all replicates. On average, 28% of proteins were found in all replicates for heavy to light comparisons representing about 9% of the annotated protein entries in the UniProt database. A candidate group of proteins changing in abundance are presented in the volcano plot (**Figure 3.21)** where changes versus significance (-Log(p-value)) are plotted. After applying the criteria of p-value < 0.01, and Log_2(hLRRK2/control) $\geq$ 0.5, 10 proteins were found to be increasing in abundance, while 26 proteins

were found to be decreasing in abundance described in **Table 3.7.** Overall the changes in the positive, increasing direction were greater relative to the protein decreases. From the outset of this experiment, I expected to see changes in both directions, due to the known stress response LRRK2 has on cells[154].

A gene-ontology enrichment analysis (GO-analysis) was conducted to group the types of proteins changing in abundance from expressing human LRRK2 in the fly brain. Proteins decreasing in abundance (p-val < 0.05, Log_2(hLRRK2/control) $\leq$ 0.5) included members of the “the extracellular region” annotation (p-value threshold < 10^{-3}). For example, this group includes the odorant binding proteins Obp83g (p-value = 0.015, Log_2(hLRRK2/control) = -0.78) and Obp99c (p-value = 0.018, Log_2(hLRRK2/control) = -0.98). Proteins increasing in abundance (p-val < 0.05, Log_2(hLRRK2/control) $\geq$ 0.5) included those corresponding to the “cellular stress response to stress” GO category (p-value threshold < 10^{-3}, Log_2(hLRRK2/control). The heat shock and oxidative stress response protein TurandotC (TotC) was increased in abundance over 3.8-fold. Two oxidoreductases, CG9150 and CG8757 were also changing over 2.5-fold while PHGPx, a glutathione reductase mitochondrial protein increased by 1.73-fold.

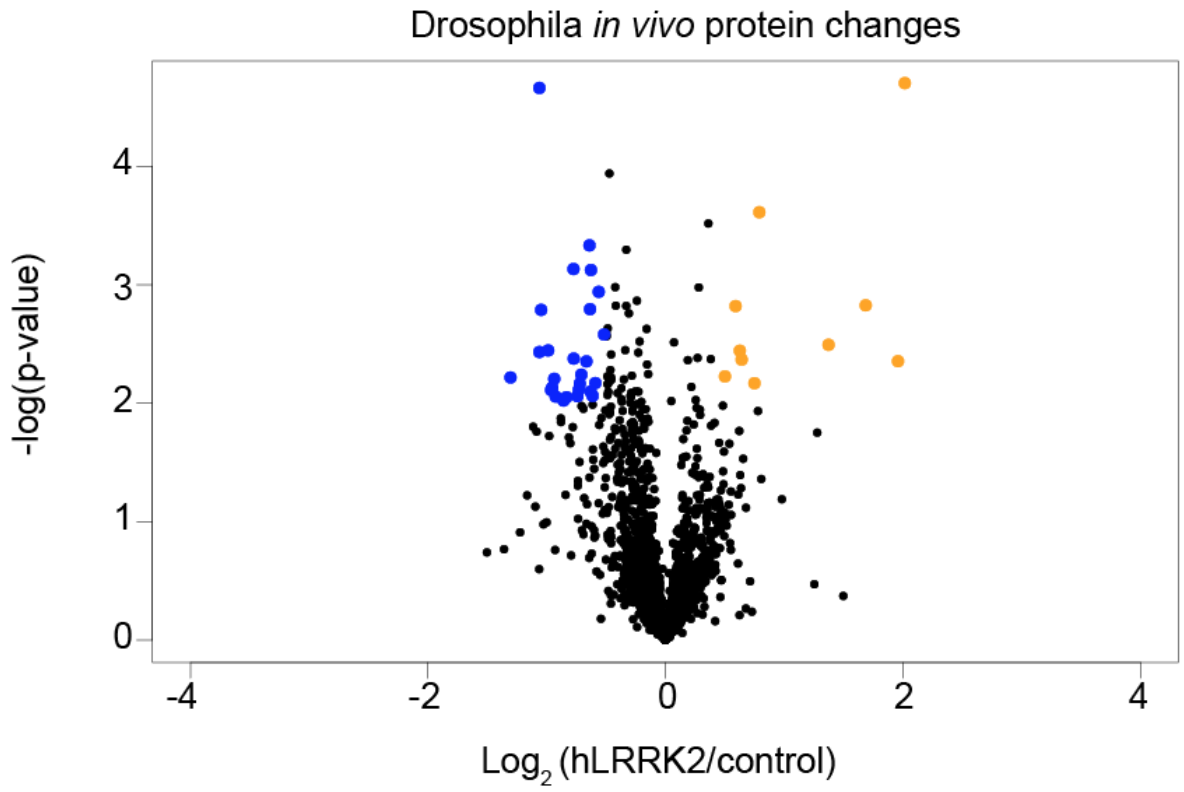


Figure 3.21 Volcano plot of quantified labeled *in vivo* proteins.
Analysis of reductive-dimethylated *Drosophila* head peptides for protein quantification between hLRRK2 expressing flies versus control non-expressing flies. Blue points reflect a greater abundance of given protein in the control versus hLRRK2 flies. Orange points reflect a greater abundance of a given protein in hLRRK2 expressing fly.

Table 3.7 Changes *in vivo* comparing transgenic flies expressing LRRK2 to control flies
Proteins represent significantly changing ratios (p-value <0.01, Log_2(hLRRK2/control) cutoff of +/-0.5.

UniProt Id	Name	Biological process	Log2(H/L) ratio	Fold change
Q8IPD8	**Cuticular protein 30F (RT03505p)**	**neurogenesis [GO:0022008]**	**2.02**	**4.05**
Q8IN43	**Protein Turandot C**	**response to oxidative stress [GO:0006979]; response to UV [GO:0009411]**	**1.96**	**3.89**
Q9VMH9	**CG9150**	**oxidoreductase activity [GO:0016491]**	**1.69**	**3.22**
Q9VU92	**CG8757**	**oxidoreductase activity**	**1.37**	**2.59**

		[GO:0016491]		
Q9VZQ8	**Glutathione peroxidase**	**mitochondrion [GO:0005739]; response to oxidative stress [GO:0006979]**	**0.79**	**1.73**
Q7JZW0	**Cuticular protein 51A (RE08808p)**	**chitin-based cuticle development [GO:0040003]**	**0.75**	**1.68**
B7Z100	**FI16622p1 (Uncharacterized protein, isoform C)**	**NA**	**0.64**	**1.56**
Q9VS84	**FI03225p isoform A-B**	**imaginal disc-derived wing morphogenesis [GO:0007476]**	**0.63**	**1.54**
Q9VF15	**Globin 1, isoform A-F (RH41321p)**	**oxygen transporter activity [GO:0005344]; oxygen transport [GO:0015671]**	**0.59**	**1.51**
O97064	**Ccp84Ag (Cuticle protein)**	**chitin-based cuticle development [GO:0040003]**	**0.50**	**1.42**
Q9VW05	**Cuticular protein 76Bd, isoform B (Cuticular protein 76Bd, isoform D)**	**extracellular matrix**	**-0.52**	**1.43**
Q9VSN3	**Cuticular protein 66D (RE57183p)**	**extracellular matrix**	**-0.56**	**1.48**
Q9V9X4	**Translation initiation factor eIF-2B subunit alpha/beta/delta-like protein**	**cytoplasm**	**-0.59**	**1.51**
A1Z9I0	**CG6357**	**extracellular space**	**-0.61**	**1.53**
Q9VA41	**Niemann-Pick type C-2h,**	**extracellular region**	**-0.63**	**1.54**

	isoform A			
E1JJ50	**Deleted/ uncharacterized protein**	**NA**	**-0.63**	**1.55**
Q9VLP1	**RE73208p**	**NA**	**-0.63**	**1.55**
Q9VKR7	**CG7300, isoform B (CG7300, isoform D)**	**NA**	**-0.64**	**1.56**
Q9VPF3	**4-hydroxyphenylpyruvate dioxygenase**	**4-hydroxyphenylpyruvate dioxygenase activity**	**-0.66**	**1.59**
Q95029	**Cathepsin L**	**extracellular space**	**-0.71**	**1.63**
Q9VMM6	**FI07243p (Obstructor-E, isoform B)**	**extracellular region**	**-0.72**	**1.65**
Q9Y125	**BcDNA.GH08312 (Saposin-related protein)**	**fusome**	**-0.73**	**1.66**
Q9VGY6	**Protein Skeletor, isoforms B/C**	**chromosome**	**-0.74**	**1.67**
Q9VQT7	**RH15675p (RH15676p) (Uncharacterized protein)**	**serine-type endopeptidase inhibitor activity**	**-0.77**	**1.71**
Q8SYQ4	**RE42475p (Uncharacterized protein)**	**cytosol**	**-0.77**	**1.71**
Q8SZN1	**RH02566p (Uncharacterized protein)**	**cysteine-type endopeptidase inhibitor activity**	**-0.83**	**1.78**
Q9VQU0	**Uncharacterized protein**	**serine-type endopeptidase inhibitor activity**	**-0.86**	**1.81**
Q7KTA1	**HL01444p (Nimrod B2)**	**extracellular region**	**-0.92**	**1.90**
Q9VR79	**LD43683p (Obstructor A-B)**	**extracellular space**	**-0.94**	**1.91**
Q9VY87	**Cathepsin B1, isoform A**	**extracellular space**	**-0.95**	**1.94**
A1ZA47	**PDZ and LIM domain protein Zasp**	**actin cytoskeleton**	**-0.96**	**1.95**
Q9VEY0	**CG8927, isoform C**	**extracellular space**	**-0.99**	**1.98**

Q0E8V7	**IP17640p (Uncharacterized protein)**	**mitochondrion**	**-1.05**	**2.07**
Q9U1L2	**CG3699**	**2,4-dienoyl-CoA reductase (NADPH) activity**	**-1.06**	**2.09**
Q9VHK7	**CG8369, isoform A (GH11984p)**	**wing disc dorsal/ventral pattern formation**	**-1.06**	**2.09**
Q9VNL0	**Gasp, isoform A (LD05259p)**	**extracellular region**	**-1.30**	**2.47**

Phosphorylation-site analysis of *in vivo* LRRK2-WT expression in Drosophila heads

The analysis resulted in the identification of 4,939 unique sequence windows containing phosphorylation sites corresponding to one or more phosphorylation sites with a site localization probability $\geq$0.75 **(Figure 3.22)**. Of these windows, 27 sites were increasing (Log_2(hLRRK2 transgenic/control) $\geq$ 0.5, p-value >0.01)) and 54 were decreasing (Log_2(hLRRK2 transgenic/control) $\leq$ 0.5, p-value < 0.01) resulting in a list of potentially regulated phosphorylation sites as an effect of LRRK2 expression in *Drosophila melanogaster* heads. The decreasing phosphorylation sites reflect phosphorylated serine and threonine residues and one phosphotyrosine residue, while the increasing list of sites did not contain any phospho-tyrosine residues. The singleton phosphotyrosine was on an well-localized (probability = 0.99) site on a peptide corresponding to isoforms of the sgg (shaggy) protein kinase, a member of the canonical Wnt signaling cascade.

To capture the ontological classes of regulated phosphopeptides, a gene ontology enrichment analysis was also performed for proteins with regulation at the phosphorylation site. Increasing phosphorylation sites had an enrichment for “structural molecule activity” while decreases in phosphopeptide activity included terms for “muscle-tendon junction” and “microtubule associated complex” (p-value threshold < 10^{-3}). The regulated phosphorylation sites were also analyzed for a gene ontology enrichment together, versus the rest of the detected phosphoproteins and resulted in the enrichment of terms including the “cytoskeletal protein binding” and “microtubule associated complex” (p-value threshold < 10^{-3}). The change in phosphopeptide regulation can be detected as a change without a

concomitant change in protein abundance. In order to visualize how the protein levels are changing for the phosphorylation sites with a change in either direction (p-value < 0.01, Log_2(hLRRK/control) $\geq$ 1) phosphopeptide quantitative ratios were plotted against protein quantitative ratios **(Figure 2.24)**. For those significantly changing phosphopeptides, in both directions the combined gene ontology resulted term enrichments for “cytoskeleton”, “contractile fiber part” and “cell projection”.

These data also demonstrate EndoA to be phosphorylated at S278 *in vivo* matching a site also phosphorylated in the label-free *in vitro* dataset. *In vivo,* S278 was increasing by Log_2(hLRRK2/control) = 0.48 (1.4-fold change), p-value=0.19. *In vitro,* this site was detected on a peptide increasing by Log_2(+kinase/control) = 1.5 (2.8-fold change), p-value = 0.006. This is an interesting finding because of LRRK2’s known phosphorylation at S75. To the best of my knowledge, this is the first description of this phosphorylation site in a living system also confirmed *in vitro* in the fly PD model.

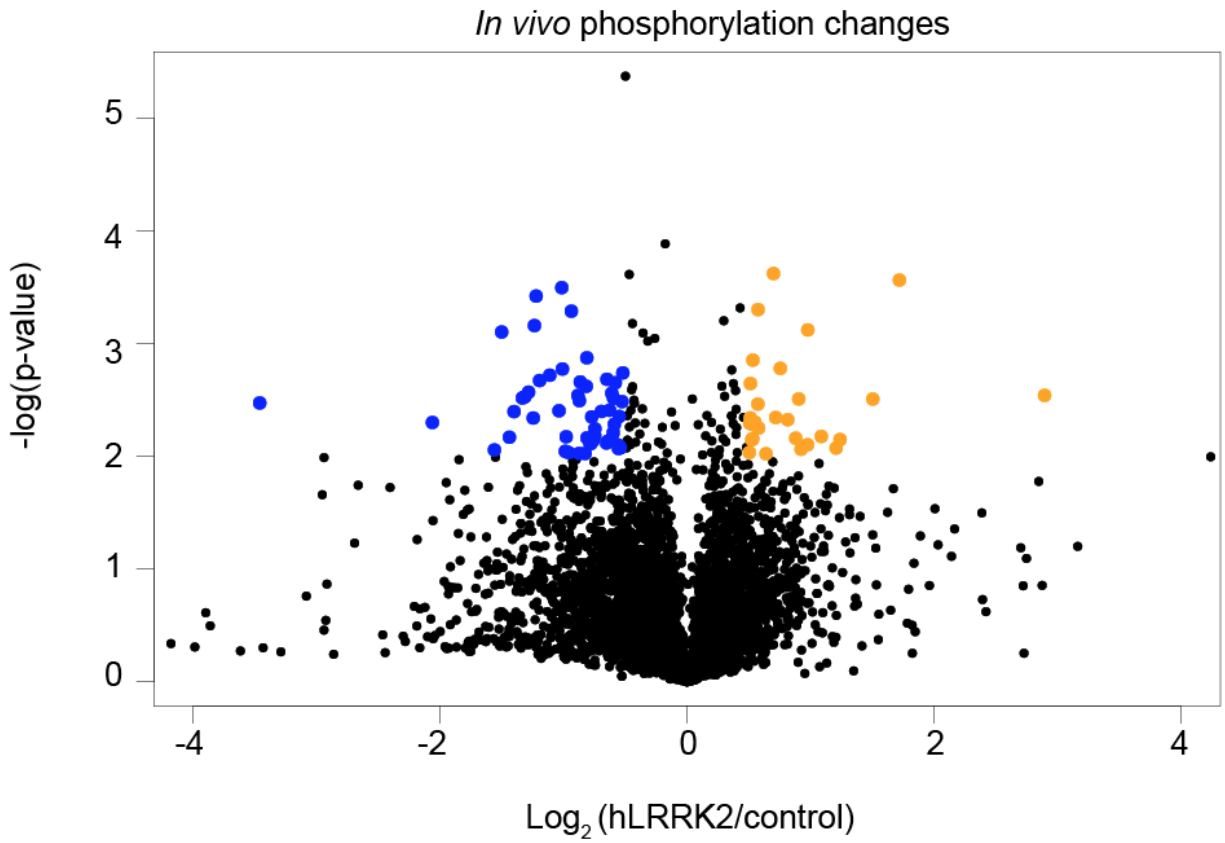


Figure 3.22 Volcano plot of quantified labeled *in vivo* phosphopeptides
Analysis of reductive-dimethylated *Drosophila* head peptides for phosphopeptide quantification between hLRRK2 expressing flies versus control non-expressing flies. Blue points reflect a greater abundance of phosphopeptides in the control versus hLRRK2 flies. Orange points reflect a greater abundance of phosphopeptides in the hLRRK2-expressing fly.

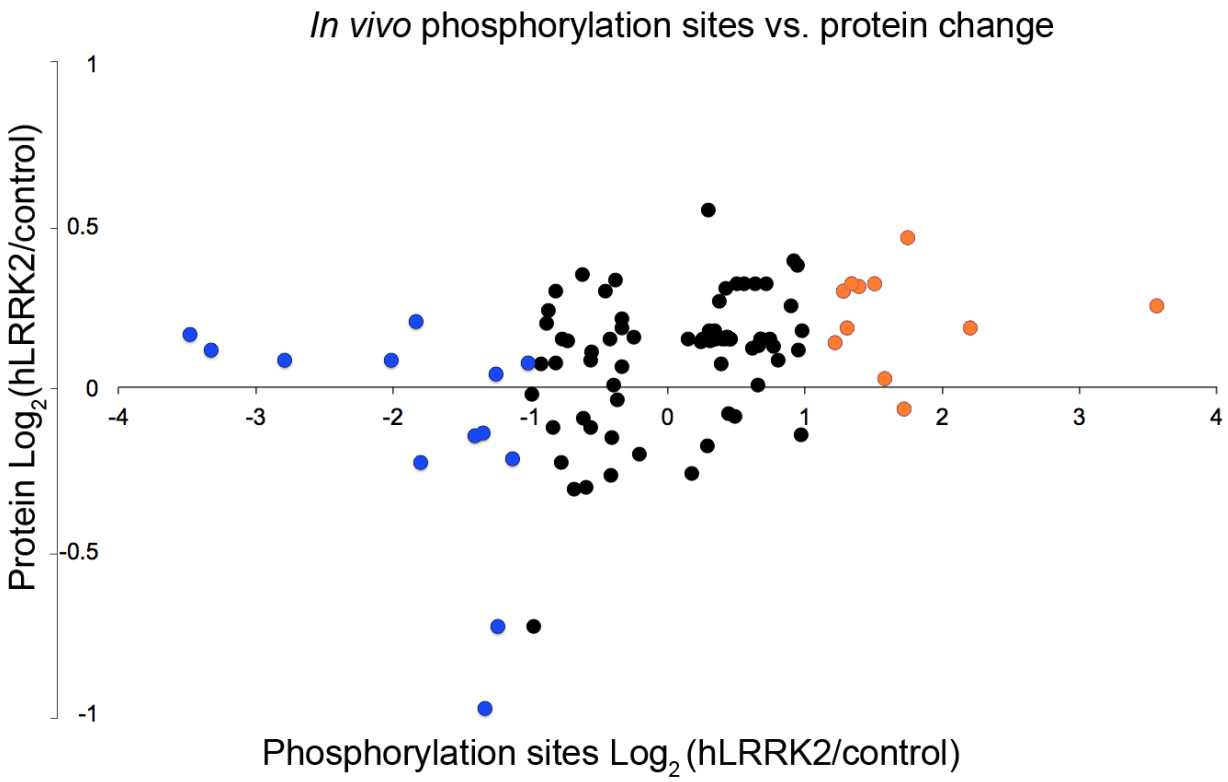


Figure 2.23 Scatter plot comparing phosphopeptide changes vs. protein changes
Phosphopeptides not-corrected for protein abundance were plotted against the changes of the corresponding protein level. Orange points are phosphopeptides changing by Log_2 (hLRRK2/control) $\geq$ 1 (p-value <0.01), blue points are phosphopeptides changing by Log_2(hLRRK2/control) $\leq$ -1 (p-value <0.01).

3.3 Conclusion

In this chapter, I described three approaches for the discovery of LRRK2 candidate substrates with mass spectrometry-based phosphoproteomics: 1) Reductive dimethylation labeling to quantify changes in phosphorylation on fly head peptides by human LRRK2 *in vitro* kinase assays. 2) A label-free approach to quantify changes in phosphorylation on fly head and mammalian SH-SY5Y neuroblastoma peptides through human LRRK2-G2019S hyperactive kinase. 3) An *in vivo* approach with reductive dimethylation comparing proteomic and phosphoproteomic changes, identifying direct and indirect effectors of human LRRK2 expression in the fly brain. The first approach to find *in vitro* substrates resulted in interesting candidate LRRK2 phosphorylation sites such as Atpalpha T398 and porin T173, which have clear orthologs in the human proteome. The second approach expanded the search for

LRRK2 substrates in two separate *in vitro* experiments through more replicates (n=10) on fly peptides and mammalian SH-SY5Y cells. I also applied automated phosphopeptide enrichment, and the nictide as a positive control for enhanced confidence of kinase activity *in vitro.* Although this did not yield Atpalpha or porin as substrates in orthologous fly or human peptide substrates, the experiments on fly peptides uncovered additional candidate substrates involved in LRRK2 biology. Although EndoA S278 was found to be phosphorylated *in vivo* and *in vitro*, there was no overlap between any orthologous phosphorylation sites between fly and human neuroblastoma candidate substrates that met the cutoffs as hits as presented in the volcano plots.

My third approach reveals regulation of proteins and phosphorylation sites as a result of human LRRK2 expression *in vivo* in the *Drosophila* human LRRK2-induced PD model. These data reproduce the LRRK2-mediated induction of an oxidative stress response observed in other studies and exhibited decreases in proteins associated with the extracellular region, like those involved in odorant binding in the drosophila olfactory system. *In vivo,* human LRRK2 expression also induced phosphorylation changes among members relevant to LRRK2 models of PD including components of cytoskeleton, proteins regulating cell projection (in the case of neuronal outgrowths/ neurites) and proteins related to muscle function.

In the next chapter I discuss specific findings from Chapter 3 and their implications for future directions for studying LRRK2 biology and PD. I also discuss the future directions for mass spectrometry-based phosphoproteomics in biological applications based on the methodology described in Chapter 2.

4.1 Discussion

In this chapter I discuss candidate substrates and how they relate to understanding LRRK2-related biology and PD research. I propose future directions of research in light of the findings presented in Chapter 3, then I conclude with a perspective on the future outlook and applications of phosphoproteomic technology for investigating PD biology and related neurodegenerative diseases.

Overall, the identification of LRRK2 substrates as outlined here does not produce the amount of overlap (passing the p-value criteria <0.01) I had expected between the *in vivo* and *in vitro* approaches or between homologous sites in the mammalian model and the fly. For example, the label-free *in vitro*

approaches did not recapitulate the large changes to phosphorylation of Atpalpha T398 (or the orthologous mammalian site at AT1A3 T370). Despite the lack of overlap, each *in vitro* approach provides direct evidence of candidate LRRK2 substrates while *in vivo* proteomics provides a glimpse into how the proteome and phosphoproteome are changing through the expression of human LRRK2 in the fly brain of the PD model. Altogether, this work provides a collection of candidate substrates worth considering in future studies seeking to assay the signaling mediated by LRRK2 and proteomic changes.

In this work I have described the identification of Atpalpha T398 as a possible LRRK2 phosphorylation substrate *in vitro*. This site is intriguing because it is conserved site on multiple disease-relevant human atpases corresponding to ATP1A2 (T378) and ATP1A3 (T370)— proteins involved in alternating hemiplegia of childhood 1 (AHC1) and rapid onset dystonia-Parkinsonism, respectively. Since the phosphopeptide identified could map to multiple conserved atpases in human, and because it was identified *in vitro* it is unknown from this study which human orthologs(s) might correspond to the candidate substrate. Therefore, a connection to a single human protein cannot be made. However, clinical studies have suggested that a mutation in ATP1A2 from a threonine to asparagine at position 378 AHC1 [155-157]. This disease is a severe movement disorder with an early onset in neonates and young infants leading to partial or full limb paralysis and various movement abnormalities including dystonia—a closely related disorder to Parkinson's disease. It is conceivable that the two diseases could share this phosphorylation site as a mechanism of aberrant signaling: where eliminations of phosphorylation leads to phenotypes associated with AHC and early onset disease, and over-phosphorylation by LRRK2-G2019S leads to an age dependent onset and parkinsonian-type movement disorders found in LRRK2-G2019S-induced PD.

The *in vitro* labeled data also reveal porin as a potential substrate. Porin has a human orthologs corresponding to human VDAC1, VDAC2 and VDAC3. Although this site was only found in one replicate, it is an intriguing candidate substrate because the VDAC is known to interact with Parkin, an important protein involved in familial forms of autosomal recessive early onset Parkinson's disease[158]. With sufficient oxidative stress and increased reactive oxygen species, mitochondria can become damaged and need to be removed by autophagy. Parkin interacts with several mitochondrial proteins to trigger autophagy, including VDAC proteins when damaged mitochondria are present in to promote

mitochondrial turnover[159]. Here, phosphorylation of porin hints at the possibility that LRRK2 may also play a role in the signaling of VDAC proteins and mitochondrial turnover. Future studies looking at PD models may benefit from investigating the phosphorylation of VDAC proteins by LRRK2.

In the label-free *in vitro* experiments candidate LRRK2 substrates included a group of proteins corresponding to the gene ontology enrichment criteria of “calcium transport”, “divalent inorganic cation transporter activity” and “calcium-transporting ATPase activity”. This grouping hints at the role LRRK2 is known to play in synaptic vesicle dynamics[160], because calcium activity is an important component in synaptic membranes undergoing exocytosis[161]. Interestingly, the calcium annotation and members of synaptic vesicle dynamics were also found among candidate substrates within the *in vitro* labeled experiments.

Specifically, the candidate substrate proteins like StnB, Ank2 and EndoA are compelling because they further establish a connection of synaptic vesicle trafficking function to LRRK2. In particular, the *in vivo* and *in vitro* overlapping site on EndoA S278 is a potentially novel LRRK2 substrate in the fly PD model. This site represents an additional point of regulation where LRRK2 might modulate the activity of EndoA in the synaptic membrane, potentially adding a level of complexity to the model proposed by Matta *et al.* (2012) [121]. In their model, over-phosphorylation of EndoA at S75 promotes an under-association of EndoA with synaptic membranes, while under-phosphorylation drives more EndoA to associate with synaptic membranes. Here, the finding that LRRK2 phosphorylates EndoA at a different site raises the possibility that phosphorylated S278 might also affect EndoA localization. In the future, this site could be validated as a LRRK2 substrate *in vivo* by comparing LRRK2 kinase dead mutants to LRRK2 or G2019S mutants. Future studies could look at the localization of EndoA S278 phosphomimetic (serine to aspartic acid) or phosphodead (serine to alanine) mutants. Additionally, it would be interesting to test how localization of EndoA would be affected by the same mutations at both residues. For example, is phosphorylation at S278 capable of modulating EndoA localization alone, in the absence of EndoA S75 phosphorylation? These future directions would provide insight into additional mechanisms behind the pathogenicity present in fly models of human LRRK2-induced parkinsonism, because EndoA S278 is not a conserved residue in human EndoA orthologs.

In the human SH-SY5Y neuroblastoma model presented here, the majority of the phosphopeptides indicated as candidate substrates did not neatly fit into known LRRK2 functional categories. The sites changing include protein involving DNA-binding elements not previously implicated in LRRK2 interactions. However, the phosphopeptide corresponding to the heterogeneous nuclear ribonucleoprotein (HNRNPC) S299 is noteworthy because it is a potentially new in LRRK2's role in regulating translation— a functional category attributed to LRRK2 biology through its phosphorylation of 4E-BP1[162].This dataset also suggests MTAP2 as a candidate LRRK2 substrate. Displaying the largest quantitative value in the neuroblastoma dataset, MTAP2 is increasing over 3.8-fold in the +kinase condition. Its membership among proteins known to compose pathogenic Lewy bodies[149], its role in partnering with neuronal potassium channels to maintain proper neuronal function, and regulating neurite growth [163,164] makes it an attractive candidate for consideration in LRRK2's kinase substrate repertoire. Further, MTAP2 phosphorylation is known to alter its localization the actin cytoskeleton, although the specific function of MTAP2 S833 found in this study is not known. In a LRRK2-G2019S hyperactive kinase neuronal model, it is possible that abnormal hyperphosphorylation of this substrate could promote mislocalization and yield pathogenic phenotypes like shortening of neurites and accumulation of Lewy bodies.

The other candidate substrates corresponding to known Lewy body proteins reflects sites of interest to LRRK2-induced PD research. The list offers candidate substrates of special consideration for addressing the PTMs of Lewy body aggregations or models of the pathology in cell culture. These sites provide a starting point for future studies aiming to further understand the role LRRK2 plays in the formation of LBs. In particular, future work could address the impact each site has on the formation of protein aggregates or their association with protein aggregates.

The usage of peptides in these *in vitro* assays offers a glimpse into potential LRRK2 substrates, but it is not the only method available to assay LRRK2 kinase activity. Alternatively, full-length, undigested protein could have been used with *in vitro* kinase assays in a native buffer that would simulate the conditions present in a biological context. Although this method would preserve the tertiary structural elements of proteins in a kinase assay, the drawback is that kinases present may also retain their native function and obscure phosphorylation activity carried out specifically by LRRK2 *in vitro.* Another limitation

of this work is the use of tryptic peptides. It is conceivable that phosphorylation sites adjacent to tryptic cleavage sites, may not be detected as candidate LRRK2 substrates if the kinase depended on the preservation of nearby lysine or arginine residues. Because of the use of tryptic peptides, a subset of potential *in vitro* peptide substrates may have been lost where lysine or arginine are preferred adjacent residues.

In vivo, this work identifies candidate substrates of LRRK2 and downstream effectors of LRRK2 expressed in the fly brain. In addition to uncovering an additional candidate phosphorylation site on EndoA S278, the *in vivo* model presented five proteins increasing in abundance in the human LRRK2-expressing fly brain: Cpr30F, a regulator of transcription and neurogenesis, TotC, a heat and oxidative stress response protein[165], the oxidoreductases CG9150 and CG8757 and the glutathione reductase PHGPx([166].

Although not much is known about Cpr30F, it has been implicated in the regulation of the growth of neural cells[167]. This is this first time it has been linked to LRRK2-induced PD in a living system. The increasing oxidative stress response proteins reaffirm LRRK2's interaction with oxidative stress and cell degeneration— processes frequently cited as a factors leading to Parkinson's disease overall [168,169] and LRRK2-induced PD specifically[170]. The loss of neurons in the substantia nigra is a process thought to be influenced by the generation of unabated reactive oxygen species and free radicals like H_2O_2 and semiquinone from the normal metabolism of dopamine[171]. Normally, the neuroprotective pigment, neuromelanin, inherent to dopaminergic neurons, reduces the ROS burden. However, since LRRK2 kinase activity and ROS levels in cells are linked, a hyperactive or overexpression of LRRK2 can worsen the oxidative stress burden because inhibition of LRRK2 decreases ROS, while increase in LRRK2 expression (as with the *in vivo* experiments) increases ROS. Moreover, H_2O_2 and LRRK2 kinase activity show a positive relationship, where increases in H_2O_2 also promotes LRRK2 kinase activity[172]. These stress response proteins may provide further avenues to perturb the biological system *in vivo* to mimic the stress induced by LRRK2. Because TotC is a oxidative stress response and heat shock protein, future studies could assay the effect of LRRK2 overexpression and TotC inhibition on stress response in terms of both oxidation and protein stability. This may be an interesting line of inquiry given the importance of protein aggregation and ROS in PD.

At the protein level, there were fewer proteins decreasing in abundance. Many were difficult to interpret in the context of this PD model. However, a couple of interesting proteins were included in the gene ontology enrichment analysis that may relate to an interesting facet of neurobiology in PD. It was unexpected to find odorant binding proteins (OBP) Obp83g and Obp99c changing in abundance, however this may be the first example of down-regulation of OBPs in an animal model of PD. In humans, the olfactory system plays an important role in early signs of PD where α-synuclein, the main component of LBs, is first deposited leading to a hyposmia (reduction in olfactory perception) [173,174]. These proteins and their odorant effectors may provide an additional phenotype to assay in the fly model of PD. Together, these *in vivo* putative substrates and effectors offer a glimpse into the mechanisms that might be contributing to the pathogenic state of the LRRK2 fly model.

This work provides a set of candidate substrates that merit validation due to their biological function and connection to PD related phenotypes or their prior association with LRRK2. The substrates and effectors described here merit further validation in future studies in biological systems to better understand their relationship to LRRK2. They may offer points of intervention to reverse pathogenic phenotypes in animal or cell models, as many of the effectors identified have orthologs in mammalian systems, or bear homologous function to merit further inquiry.

Future outlook and applications

The work described in Chapter 3 uses approached that are becoming increasingly applied in assessing kinase during cell signaling which included the use of label free quantitation and automated phosphopeptide enrichment. If I were to speculate on the future in methodological trends for assessing phosphoproteomes, it is clear that the recent developments in automation, high-throughput sample preparation and improved computational tools could greatly impact these analyses by increasing reproducibility and improving peptide identification between independent runs.

Here, I have presented phosphoproteomic data generated through the use of an automated robot for phosphopeptide enrichment, which will likely increase in popularity since it reduces variability when enriching phosphopeptides in high-throughput formats like 96-well plates. A recent study by Humphrey *et al*. (2015) [175] introduced a novel workflow to simultaneously increase throughput and minimize sample handling. They use a 96-well digestion format for trifluoroethanol digestion buffer prior to phosphopeptide

enrichment. This approach does not require peptide-desalting which simplifies sample processing and minimizes sample loss. The use of single-shot MS analysis by the group prior to label-free quantification and peptide matching between runs enhanced peptide identification from separate MS runs. Because this approach is compatible with automated phosphopeptide enrichment, these two methods could be used in clinical applications or when precious and limited samples are the only option. The future in researching PD and other neurological diseases could make use of these tools in situations where post-mortem samples are available or when a clinical sample is appropriate, as is the case when assaying biomarkers of disease.

Overall, the work presented in this dissertation offers methodological insight into assaying phosphoproteomes and provides clues into the biological mechanisms of LRRK2. Chapter 2 describes a method that is approachable to a broad range of biologists with access to mass-spectrometry instrumentation to study phosphorylation in their biological system. Chapter 3 provides insights into new *in vitro* candidate substrates of LRRK2 and *in vivo* effectors in the fly Parkinson's disease model. These candidate substrates represent new avenues of inquiry towards a more complete understanding of LRRK2 signaling and biological models of Parkinson's disease.

References

1. BURNETT, G. & KENNEDY, E. P. The enzymatic phosphorylation of proteins. *J. Biol. Chem.* **211,** 969–980 (1954).
2. Krebs, E. G. & Fischer, E. H. The phosphorylase b to a converting enzyme of rabbit skeletal muscle. *Biochim. Biophys. Acta* (1956).
3. Wirth, A. Rho kinase and hypertension. *Biochim. Biophys. Acta* **1802,** 1276–1284 (2010).
4. Akhurst, R. J. & Hata, A. Targeting the TGFβ signalling pathway in disease. *Nat Rev Drug Discov* **11,** 790–811 (2012).
5. Petersen, A. J., Rimkus, S. A. & Wassarman, D. A. ATM kinase inhibition in glial cells activates the innate immune response and causes neurodegeneration in Drosophila. *Proceedings of the National Academy of Sciences* **109,** E656–64 (2012).
6. Ubersax, J. A. & Ferrell, J. E. Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Biol* **8,** 530–541 (2007).
7. Manning, G., Plowman, G. D., Hunter, T. & Sudarsanam, S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.* **27,** 514–520 (2002).
8. Roskoski, R. A historical overview of protein kinases and their targeted small molecule inhibitors. *Pharmacol. Res.* **100,** 1–23 (2015).
9. Giese, K. P. & Mizuno, K. The roles of protein kinases in learning and memory. *Learn. Mem.* **20,** 540–552 (2013).
10. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298,** 1912–+ (2002).
11. George, S. *et al.* A family with severe insulin resistance and diabetes due to a mutation in AKT2. *Science* **304,** 1325–1328 (2004).
12. Chen, D.-H. *et al.* Missense mutations in the regulatory domain of PKC gamma: a new mechanism for dominant nonepisodic cerebellar ataxia. *Am. J. Hum. Genet.* **72,** 839–849 (2003).
13. Lahiry, P., Torkamani, A., Schork, N. J. & Hegele, R. A. Kinase mutations in human disease: interpreting genotype-phenotype relationships. *Nat. Rev. Genet.* **11,** 60–74 (2010).
14. Cohen, P. Protein kinases--the major drug targets of the twenty-first century? *Nat Rev Drug Discov* **1,** 309–315 (2002).
15. Shah, K., Liu, Y., Deirmengian, C. & Shokat, K. M. Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proc. Natl. Acad. Sci. U.S.A.* **94,** 3565–3570 (1997).
16. Chen, X. *et al.* A chemical-genetic approach to studying neurotrophin signaling. *Neuron* **46,** 13–21 (2005).
17. Sha, D., Chin, L.-S. & Li, L. Phosphorylation of parkin by Parkinson disease-linked kinase PINK1 activates parkin E3 ligase function and NF-kappaB signaling. *Hum. Mol. Genet.* **19,** 352–363 (2010).
18. Houseman, B. T., Huh, J. H., Kron, S. J. & Mrksich, M. Peptide chips for the quantitative evaluation of protein kinase activity. *Nat. Biotechnol.* **20,** 270–274 (2002).
19. Lesaicherre, M. L., Uttamchandani, M., Chen, G. Y. J. & Yao, S. Q. Antibody-based fluorescence detection of kinase activity on a peptide array. *Bioorg. Med. Chem. Lett.* **12,** 2085–2088 (2002).
20. Ptacek, J. *et al.* Global analysis of protein phosphorylation in yeast. *Nature* **438,** 679–684 (2005).
21. Daub, H. *et al.* Kinase-selective enrichment enables quantitative phosphoproteomics of the kinome across the cell cycle. *Mol. Cell* **31,** 438–448 (2008).

22. Belozerov, V. E., Lin, Z.-Y., Gingras, A.-C., McDermott, J. C. & Michael Siu, K. W. High-resolution protein interaction map of the Drosophila melanogaster p38 mitogen-activated protein kinases reveals limited functional redundancy. *Mol. Cell. Biol.* **32,** 3695–3706 (2012).
23. Xue, L. *et al.* Sensitive kinase assay linked with phosphoproteomics for identifying direct kinase substrates. *Proceedings of the National Academy of Sciences* **109,** 5615–5620 (2012).
24. Edelman, W. C., Haas, K. M., Hsu, J. I., Lawrence, R. T. & Villén, J. A practical recipe to survey phosphoproteomes. *Methods Mol. Biol.* **1156,** 389–405 (2014).
25. Hornbeck, P. V. *et al.* PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **40,** D261–70 (2012).
26. Bodenmiller, B. *et al.* Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Science Signaling* **3,** rs4 (2010).
27. Huttlin, E. L. *et al.* A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143,** 1174–1189 (2010).
28. Holt, L. J. *et al.* Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* **325,** 1682–1686 (2009).
29. Kettenbach, A. N. *et al.* Quantitative phosphoproteomics identifies substrates and functional modules of Aurora and Polo-like kinase activities in mitotic cells. *Science Signaling* **4,** rs5 (2011).
30. Yu, Y. *et al.* Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. *Science* **332,** 1322–1326 (2011).
31. Hsu, P. P. *et al.* The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science* **332,** 1317–1322 (2011).
32. Cutillas, P. R. *et al.* Ultrasensitive and absolute quantification of the phosphoinositide 3-kinase/Akt signal transduction pathway by mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 8959–8964 (2006).
33. Yu, Y. *et al.* A site-specific, multiplexed kinase activity assay using stable-isotope dilution and high-resolution mass spectrometry. *Proceedings of the National Academy of Sciences* **106,** 11606–11611 (2009).
34. Kubota, K. *et al.* Sensitive multiplexed analysis of kinase activities and activity-based kinase identification. *Nat. Biotechnol.* **27,** 933–940 (2009).
35. Wu, R. *et al.* A large-scale method to measure absolute protein phosphorylation stoichiometries. *Nat. Methods* **8,** 677–683 (2011).
36. Olsen, J. V. *et al.* Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. *Science Signaling* **3,** ra3–ra3 (2010).
37. Humphrey, S. J. *et al.* Dynamic adipocyte phosphoproteome reveals that Akt directly regulates mTORC2. *Cell Metab.* **17,** 1009–1020 (2013).
38. Ficarro, S. B. *et al.* Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat. Biotechnol.* **20,** 301–305 (2002).
39. Rush, J. *et al.* Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* **23,** 94–101 (2005).
40. Andersson, L. & Porath, J. Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. *Anal. Biochem.* **154,** 250–254 (1986).
41. Posewitz, M. C. & Tempst, P. Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal. Chem.* **71,** 2883–2892 (1999).
42. Sano, A. & Nakamura, H. Titania as a chemo-affinity support for the column-switching HPLC analysis of phosphopeptides: application to the characterization of

phosphorylation sites in proteins by combination with protease digestion and electrospray ionization mass spectrometry. *Anal Sci* **20,** 861–864 (2004).
43. Pinkse, M. W. H., Uitto, P. M., Hilhorst, M. J., Ooms, B. & Heck, A. J. R. Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal. Chem.* **76,** 3935–3943 (2004).
44. Beausoleil, S. A. *et al.* Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 12130–12135 (2004).
45. Gruhler, A. *et al.* Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell Proteomics* **4,** 310–327 (2005).
46. Olsen, J. V. *et al.* Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127,** 635–648 (2006).
47. Villen, J., Beausoleil, S. A., Gerber, S. A. & Gygi, S. P. Large-scale phosphorylation analysis of mouse liver. *Proceedings of the National Academy of Sciences* **104,** 1488–1493 (2007).
48. Villén, J. & Gygi, S. P. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat Protoc* **3,** 1630–1638 (2008).
49. Alpert, A. J. Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. *Anal. Chem.* **80,** 62–76 (2008).
50. Swaney, D. L., Wenger, C. D. & Coon, J. J. Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *J. Proteome Res.* **9,** 1323–1329 (2010).
51. Schroeder, M. J., Shabanowitz, J., Schwartz, J. C., Hunt, D. F. & Coon, J. J. A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. *Anal. Chem.* **76,** 3590–3598 (2004).
52. Villén, J., Beausoleil, S. A. & Gygi, S. P. Evaluation of the utility of neutral-loss-dependent MS3 strategies in large-scale phosphorylation analysis. *Proteomics* **8,** 4444–4452 (2008).
53. Syka, J. E. P., Coon, J. J., Schroeder, M. J., Shabanowitz, J. & Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 9528–9533 (2004).
54. Chi, A. *et al.* Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 2193–2198 (2007).
55. Nagaraj, N., D'Souza, R. C. J., Cox, J., Olsen, J. V. & Mann, M. Feasibility of large-scale phosphoproteomics with higher energy collisional dissociation fragmentation. *J. Proteome Res.* **9,** 6786–6794 (2010).
56. Jedrychowski, M. P. *et al.* Evaluation of HCD- and CID-type fragmentation within their respective detection platforms for murine phosphoproteomics. *Mol. Cell Proteomics* **10,** M111.009910 (2011).
57. Beausoleil, S. A., Villén, J., Gerber, S. A., Rush, J. & Gygi, S. P. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* **24,** 1285–1292 (2006).
58. Taus, T. *et al.* Universal and confident phosphorylation site localization using phosphoRS. *J. Proteome Res.* **10,** 5354–5362 (2011).
59. Fermin, D., Walmsley, S. J., Gingras, A.-C., Choi, H. & Nesvizhskii, A. I. LuciPHOr: algorithm for phosphorylation site localization with false localization rate estimation using modified target-decoy approach. *Mol. Cell Proteomics* **12,** 3409–3419 (2013).
60. Ficarro, S. B. *et al.* Magnetic bead processor for rapid evaluation and optimization of parameters for phosphopeptide enrichment. *Anal. Chem.* **81,** 4566–4575 (2009).
61. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass

spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5,** 976–989 (1994).
62. Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20,** 3551–3567 (1999).
63. Craig, R., Cortens, J. P. & Beavis, R. C. Open source system for analyzing, validating, and storing protein identification data. *J. Proteome Res.* **3,** 1234–1242 (2004).
64. Geer, L. Y. *et al.* Open mass spectrometry search algorithm. *J. Proteome Res.* **3,** 958–964 (2004).
65. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4,** 207–214 (2007).
66. Bailey, C. M. *et al.* SLoMo: automated site localization of modifications from ETD/ECD mass spectra. *J. Proteome Res.* **8,** 1965–1971 (2009).
67. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26,** 1367–1372 (2008).
68. Searle, B. C. Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics* **10,** 1265–1269 (2010).
69. Dephoure, N. & Gygi, S. P. A solid phase extraction-based platform for rapid phosphoproteomic analysis. *Methods* **54,** 379–386 (2011).
70. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2,** 1896–1906 (2007).
71. Lees, A. J., Hardy, J. & Revesz, T. Parkinson's disease. *Lancet* **373,** 2055–2066 (2009).
72. Schenck, C. H., Bundlie, S. R. & Mahowald, M. W. Delayed emergence of a parkinsonian disorder in 38% of 29 older men initially diagnosed with idiopathic rapid eye movement sleep behaviour disorder. *Neurology* **46,** 388–393 (1996).
73. Boeve, B. F. *et al.* Pathophysiology of REM sleep behaviour disorder and relevance to neurodegenerative disease. *Brain* **130,** 2770–2788 (2007).
74. Ponsen, M. M. *et al.* Idiopathic hyposmia as a preclinical sign of Parkinson's disease. *Ann. Neurol.* **56,** 173–181 (2004).
75. Hely, M. A., Morris, J. G. L., Reid, W. G. J. & Trafficante, R. Sydney Multicenter Study of Parkinson's disease: non-L-dopa-responsive problems dominate at 15 years. *Mov. Disord.* **20,** 190–199 (2005).
76. Aarsland, D. *et al.* Neuropsychiatric symptoms in patients with Parkinson's disease and dementia: frequency, profile and associated care giver stress. *J. Neurol. Neurosurg. Psychiatr.* **78,** 36–42 (2007).
77. Allcock, L. M., Ullyart, K., Kenny, R. A. & Burn, D. J. Frequency of orthostatic hypotension in a community based cohort of patients with Parkinson's disease. *J. Neurol. Neurosurg. Psychiatr.* **75,** 1470–1471 (2004).
78. Engelender, S. Ubiquitination of alpha-synuclein and autophagy in Parkinson's disease. *Autophagy* **4,** 372–374 (2008).
79. Ishizawa, T., Mattila, P., Davies, P., Wang, D. & Dickson, D. W. Colocalization of tau and alpha-synuclein epitopes in Lewy bodies. *J. Neuropathol. Exp. Neurol.* **62,** 389–397 (2003).
80. Arima, K. *et al.* Cellular co-localization of phosphorylated tau- and NACP/alpha-synuclein-epitopes in lewy bodies in sporadic Parkinson's disease and in dementia with Lewy bodies. *Brain Res.* **843,** 53–61 (1999).
81. Funayama, M. *et al.* A new locus for Parkinson's disease (PARK8) maps to

chromosome 12p11.2-q13.1. *Ann. Neurol.* **51,** 296–301 (2002).
82. Zimprich, A. *et al.* Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* **44,** 601–607 (2004).
83. Paisán-Ruíz, C. *et al.* Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. *Neuron* **44,** 595–600 (2004).
84. Saunders-Pullman, R. *et al.* LRRK2 G2019S mutations are associated with an increased cancer risk in Parkinson disease. *Mov. Disord.* **25,** 2536–2541 (2010).
85. Agalliu, I. *et al.* Higher frequency of certain cancers in LRRK2 G2019S mutation carriers with Parkinson disease: a pooled analysis. *JAMA Neurol* **72,** 58–65 (2015).
86. Singleton, A. B., Farrer, M. J. & Bonifati, V. The genetics of Parkinson's disease: progress and therapeutic implications. *Mov. Disord.* **28,** 14–23 (2013).
87. Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491,** 119–124 (2012).
88. Marcinek, P. *et al.* LRRK2 and RIPK2 variants in the NOD 2-mediated signaling pathway are associated with susceptibility to Mycobacterium leprae in Indian populations. *PLoS ONE* **8,** e73103 (2013).
89. Paisán-Ruíz, C., Lewis, P. A. & Singleton, A. B. LRRK2: cause, risk, and mechanism. *J Parkinsons Dis* **3,** 85–103 (2013).
90. Rubio, J. P. *et al.* Deep sequencing of the LRRK2 gene in 14,002 individuals reveals evidence of purifying selection and independent origin of the p.Arg1628Pro mutation in Europe. *Hum. Mutat.* **33,** 1087–1098 (2012).
91. Sancho, R. M., Law, B. M. H. & Harvey, K. Mutations in the LRRK2 Roc-COR tandem domain link Parkinson's disease to Wnt signalling pathways. *Hum. Mol. Genet.* **18,** 3955–3968 (2009).
92. Berwick, D. C. & Harvey, K. LRRK2: an éminence grise of Wnt-mediated neurogenesis? *Front Cell Neurosci* **7,** 82 (2013).
93. Alegre-Abarrategui, J. *et al.* LRRK2 regulates autophagic activity and localizes to specific membrane microdomains in a novel human genomic reporter cellular model. *Hum. Mol. Genet.* **18,** 4022–4034 (2009).
94. Martin, I. Decoding Parkinson's Disease Pathogenesis: The Role of Deregulated mRNA Translation. *J Parkinsons Dis* **6,** 17–27 (2016).
95. Cookson, M. R. LRRK2 Pathways Leading to Neurodegeneration. *Curr Neurol Neurosci Rep* **15,** 42 (2015).
96. West, A. B. *et al.* Parkinson's disease-associated mutations in leucine-rich repeat kinase 2 augment kinase activity. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 16842–16847 (2005).
97. Khan, N. L. *et al.* Mutations in the gene LRRK2 encoding dardarin (PARK8) cause familial Parkinson's disease: clinical, pathological, olfactory and functional imaging and genetic data. *Brain* **128,** 2786–2796 (2005).
98. Jaleel, M. *et al.* LRRK2 phosphorylates moesin at threonine-558: characterization of how Parkinson's disease mutants affect kinase activity. *Biochem. J.* **405,** 307–317 (2007).
99. Zabetian, C. P. *et al.* LRRK2 G2019S in families with Parkinson disease who originated from Europe and the Middle East: evidence of two distinct founding events beginning two millennia ago. *Am. J. Hum. Genet.* **79,** 752–758 (2006).
100. Ozelius, L. J. *et al.* LRRK2 G2019S as a cause of Parkinson's disease in Ashkenazi Jews. *N. Engl. J. Med.* **354,** 424–425 (2006).
101. Lesage, S. *et al.* LRRK2 G2019S as a cause of Parkinson's disease in North African Arabs. *N. Engl. J. Med.* **354,** 422–423 (2006).
102. Mamais, A. *et al.* Divergent α-synuclein solubility and aggregation properties in G2019S LRRK2 Parkinson's disease brains with Lewy Body pathology compared to idiopathic

cases. *Neurobiology of Disease* **58,** 183–190 (2013).
103. Greggio, E. *et al.* Kinase activity is required for the toxic effects of mutant LRRK2/dardarin. *Neurobiology of Disease* **23,** 329–341 (2006).
104. Smith, W. W. *et al.* Kinase activity of mutant LRRK2 mediates neuronal toxicity. *Nat. Neurosci.* **9,** 1231–1233 (2006).
105. Alegre-Abarrategui, J., Ansorge, O., Esiri, M. & Wade-Martins, R. LRRK2 is a component of granular alpha-synuclein pathology in the brainstem of Parkinson's disease. *Neuropathol. Appl. Neurobiol.* **34,** 272–283 (2008).
106. Longo, F., Russo, I., Shimshek, D. R., Greggio, E. & Morari, M. Genetic and pharmacological evidence that G2019S LRRK2 confers a hyperkinetic phenotype, resistant to motor decline associated with aging. *Neurobiology of Disease* **71,** 62–73 (2014).
107. Yue, M. *et al.* Progressive dopaminergic alterations and mitochondrial abnormalities in LRRK2 G2019S knock-in mice. *Neurobiology of Disease* **78,** 172–195 (2015).
108. Lee, B. D. *et al.* Inhibitors of leucine-rich repeat kinase-2 protect against models of Parkinson's disease. *Nat. Med.* **16,** 998–1000 (2010).
109. Dusonchet, J. *et al.* A rat model of progressive nigral neurodegeneration induced by the Parkinson's disease-associated G2019S mutation in LRRK2. *J. Neurosci.* **31,** 907–912 (2011).
110. Walker, M. D. *et al.* Behavioral deficits and striatal DA signaling in LRRK2 p.G2019S transgenic rats: a multimodal investigation including PET neuroimaging. *J Parkinsons Dis* **4,** 483–498 (2014).
111. Liu, Z. *et al.* A Drosophila model for LRRK2-linked parkinsonism. *Proceedings of the National Academy of Sciences* **105,** 2693–2698 (2008).
112. Weil, R. S. *et al.* Visual dysfunction in Parkinson's disease. *Brain* (2016). doi:10.1093/brain/aww175
113. Nguyen-Legros, J. Functional neuroarchitecture of the retina: hypothesis on the dysfunction of retinal dopaminergic circuitry in Parkinson's disease. *Surg Radiol Anat* **10,** 137–144 (1988).
114. Bodis-Wollner, I., Kozlowski, P. B., Glazman, S. & Miri, S. α-synuclein in the inner retina in parkinson disease. *Ann. Neurol.* **75,** 964–966 (2014).
115. Beach, T. G. *et al.* Phosphorylated α-synuclein-immunoreactive retinal neuronal elements in Parkinson's disease subjects. *Neurosci. Lett.* **571,** 34–38 (2014).
116. Nichols, R. J. *et al.* Substrate specificity and inhibitors of LRRK2, a protein kinase mutated in Parkinson's disease. *Biochem. J.* **424,** 47–60 (2009).
117. Kawakami, F. *et al.* LRRK2 phosphorylates tubulin-associated tau but not the free molecule: LRRK2-mediated regulation of the tau-tubulin association and neurite outgrowth. *PLoS ONE* **7,** e30834 (2012).
118. Bailey, R. M. *et al.* LRRK2 phosphorylates novel tau epitopes and promotes tauopathy. *Acta Neuropathol.* **126,** 809–827 (2013).
119. Qing, H., Wong, W., McGeer, E. G. & McGeer, P. L. Lrrk2 phosphorylates alpha synuclein at serine 129: Parkinson disease implications. *Biochem. Biophys. Res. Commun.* **387,** 149–152 (2009).
120. Gloeckner, C. J., Schumacher, A., Boldt, K. & Ueffing, M. The Parkinson disease-associated protein kinase LRRK2 exhibits MAPKKK activity and phosphorylates MKK3/6 and MKK4/7, in vitro. *J. Neurochem.* **109,** 959–968 (2009).
121. Matta, S. *et al.* LRRK2 controls an EndoA phosphorylation cycle in synaptic endocytosis. *Neuron* **75,** 1008–1021 (2012).
122. Yun, H. J. *et al.* LRRK2 phosphorylates Snapin and inhibits interaction of Snapin with SNAP-25. *Exp. Mol. Med.* **45,** e36 (2013).
123. Ohta, E., Kawakami, F., Kubo, M. & Obata, F. LRRK2 directly phosphorylates Akt1 as a

possible physiological substrate: impairment of the kinase activity by Parkinson's disease-associated mutations. *FEBS Lett.* **585,** 2165–2170 (2011).
124. Steger, M. *et al.* Phosphoproteomics reveals that Parkinson's disease kinase LRRK2 regulates a subset of Rab GTPases. *Elife* **5,** (2016).
125. Martin, I. *et al.* Ribosomal protein s15 phosphorylation mediates LRRK2 neurodegeneration in Parkinson's disease. *Cell* **157,** 472–485 (2014).
126. Eng, J. K., Jahan, T. A. & Hoopmann, M. R. Comet: an open-source MS/MS sequence database search tool. *Proteomics* **13,** 22–24 (2013).
127. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26,** 966–968 (2010).
128. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S. & MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **4,** 923–925 (2007).
129. Eden, E., Lipson, D., Yogev, S. & Yakhini, Z. Discovering motifs in ranked lists of DNA sequences. *PLoS Comput Biol* **3,** e39 (2007).
130. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10,** 48 (2009).
131. Webber, P. J. *et al.* Autophosphorylation in the leucine-rich repeat kinase 2 (LRRK2) GTPase domain modifies kinase and GTP-binding activities. *J. Mol. Biol.* **412,** 94–110 (2011).
132. Greggio, E. *et al.* The Parkinson's disease kinase LRRK2 autophosphorylates its GTPase domain at multiple sites. *Biochem. Biophys. Res. Commun.* **389,** 449–454 (2009).
133. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. R. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc* **4,** 484–494 (2009).
134. Pungaliya, P. P. *et al.* Identification and characterization of a leucine-rich repeat kinase 2 (LRRK2) consensus phosphorylation motif. *PLoS ONE* **5,** e13672 (2010).
135. Bento, C. F., Ashkenazi, A., Jimenez-Sanchez, M. & Rubinsztein, D. C. The Parkinson's disease-associated genes ATP13A2 and SYT11 regulate autophagy via a common pathway. *Nature Communications* **7,** (2016).
136. Finelli, M. J., Sanchez-Pulido, L., Liu, K. X., Davies, K. E. & Oliver, P. L. The Evolutionarily Conserved Tre2/Bub2/Cdc16 (TBC), Lysin Motif (LysM), Domain Catalytic (TLDc) Domain Is Neuroprotective against Oxidative Stress. *J. Biol. Chem.* **291,** 2751–2763 (2016).
137. Komarov, A. G., Graham, B. H., Craigen, W. J. & Colombini, M. The physiological properties of a novel family of VDAC-like proteins from Drosophila melanogaster. *Biophysj* **86,** 152–162 (2004).
138. Park, J. *et al.* Drosophila Porin/VDAC affects mitochondrial morphology. *PLoS ONE* **5,** e13151 (2010).
139. Plowey, E. D., Cherra, S. J., Liu, Y.-J. & Chu, C. T. Role of autophagy in G2019S-LRRK2-associated neurite shortening in differentiated SH-SY5Y cells. *J. Neurochem.* **105,** 1048–1056 (2008).
140. Liou, A. K. F., Leak, R. K., Li, L. & Zigmond, M. J. Wild-type LRRK2 but not its mutant attenuates stress-induced cell death via ERK pathway. *Neurobiology of Disease* **32,** 116–124 (2008).
141. Smith, W. W. *et al.* Leucine-rich repeat kinase 2 (LRRK2) interacts with parkin, and mutant LRRK2 induces neuronal degeneration. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 18676–18681 (2005).
142. Fergestad, T., Davis, W. S. & Broadie, K. The stoned proteins regulate synaptic vesicle

recycling in the presynaptic terminal. *J. Neurosci.* **19,** 5847–5860 (1999).
143. Mohrmann, R., Matthies, H. J., Woodruff, E. & Broadie, K. Stoned B mediates sorting of integral synaptic vesicle proteins. *Neuroscience* **153,** 1048–1063 (2008).
144. Koch, I. *et al.* Drosophila ankyrin 2 is required for synaptic stability. *Neuron* **58,** 210–222 (2008).
145. Phillips, A. M., Smith, M., Ramaswami, M. & Kelly, L. E. The products of the Drosophila stoned locus interact with synaptic vesicles via synaptotagmin. *J. Neurosci.* **20,** 8254–8261 (2000).
146. Lee, S., Liu, H.-P., Lin, W.-Y., Guo, H. & Lu, B. LRRK2 kinase regulates synaptic morphology through distinct substrates at the presynaptic and postsynaptic compartments of the Drosophila neuromuscular junction. *J. Neurosci.* **30,** 16959–16969 (2010).
147. Chan, S. L., Chua, L.-L., Angeles, D. C. & Tan, E.-K. MAP1B rescues LRRK2 mutant-mediated cytotoxicity. *Mol Brain* **7,** 29 (2014).
148. Kalia, L. V. *et al.* Clinical correlations with Lewy body pathology in LRRK2-related Parkinson disease. *JAMA Neurol* **72,** 100–105 (2015).
149. Shults, C. W. Lewy bodies. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 1661–1668 (2006).
150. Campos, D. V. The MAP1B case: an old MAP that is new again. *Developmental …* (2014).
151. Sánchez, C., Díaz-Nido, J. & Avila, J. Phosphorylation of microtubule-associated protein 2 (MAP2) and its relevance for the regulation of the neuronal cytoskeleton function. *Prog. Neurobiol.* **61,** 133–168 (2000).
152. Tanji, K. *et al.* Phosphorylation of serine 349 of p62 in Alzheimer's disease brain. *Acta Neuropathologica Communications 2014 2:1* **2,** 50 (2014).
153. Kawamoto, Y., Akiguchi, I. & Nakamura, S. 14-3-3 proteins in Lewy bodies in Parkinson disease and diffuse Lewy body disease brains. *Journal of …* (2002).
154. Nguyen, H. N., Byers, B., Cord, B., Shcheglovitov, A. & Byrne, J. LRRK2 mutant iPSC-derived DA neurons demonstrate increased susceptibility to oxidative stress. *Cell Stem Cell* (2011).
155. Kanavakis, E. & Xaidara, A. Alternating hemiplegia of childhood: a syndrome inherited with an autosomal dominant trait. *… Medicine & Child …* (2003).
156. Swoboda, K. J., Kanavakis, E. & Xaidara, A. Alternating hemiplegia of childhood or familial hemiplegic migraine?: a novel ATP1A2 mutation. *Annals of …* (2004).
157. Bassi, M. T., Bresolin, N., Tonelli, A. & Nazos, K. A novel mutation in the ATP1A2 gene causes alternating hemiplegia of childhood. *Journal of medical …* (2004).
158. Pickrell, A. M. & Youle, R. J. The roles of PINK1, parkin, and mitochondrial fidelity in Parkinson's disease. *Neuron* **85,** 257–273 (2015).
159. Sun, Y., Vashisht, A. A., Tchieu, J. & Wohlschlegel, J. A. Voltage-dependent anion channels (VDACs) recruit Parkin to defective mitochondria to promote mitochondrial autophagy. *Journal of Biological …* (2012).
160. Dihanich, S. & Manzoni, C. LRRK2: a problem lurking in vesicle trafficking? *The Journal of Neuroscience* (2011).
161. Jahn, R. & Fasshauer, D. Molecular machines governing exocytosis of synaptic vesicles. *Nature* (2012).
162. Hébert, S. S. & Dorval, V. LRRK2 in transcription and translation regulation: relevance for Parkinson's disease. *Frontiers in neurology* (2012).
163. Harada, A., Teng, J., Takei, Y. & Oguchi, K. MAP2 is required for dendrite elongation, PKA anchoring in dendrites, and proper PKA signal transduction. *The Journal of cell …* (2002).
164. Sandoz, G., Tardy, M. P. & Thümmler, S. Mtap2 is a constituent of the protein network that regulates twik-related K+ channel expression and trafficking. *The Journal of …*

(2008).
165. Ekengren, S. & Hultmark, D. A family of Turandot-related genes in the humoral stress response of Drosophila. *Biochemical and biophysical research …* (2001).
166. Li, D., Blasevich, F., Theopold, U. & Schmidt, O. Possible function of two insect phospholipid-hydroperoxide glutathione peroxidases. *Journal of Insect Physiology* (2003).
167. Neumüller, R. A. *et al.* Genome-wide analysis of self-renewal in Drosophila neural stem cells by transgenic RNAi. *Cell Stem Cell* **8,** 580–593 (2011).
168. Jenner, P. & Olanow, C. W. Understanding cell death in Parkinson's disease. *Ann. Neurol.* (1998).
169. Sherer, T. B. & Greenamyre, J. T. Oxidative damage in Parkinson's disease. *Antioxidants & redox signaling* (2005).
170. Sanders, L. H. *et al.* LRRK2 mutations cause mitochondrial DNA damage in iPSC-derived neural cells from Parkinson's disease patients: reversal by gene correction. *Neurobiology of Disease* **62,** 381–386 (2014).
171. Jenner, P. Oxidative stress in Parkinson's disease. *Ann. Neurol.* (2003).
172. Yang, D. *et al.* LRRK2 kinase activity mediates toxic interactions between genetic mutation and oxidative stress in a Drosophila model: suppression by curcumin. *Neurobiology of Disease* **47,** 385–392 (2012).
173. Braak, H. *et al.* Staging of the intracerebral inclusion body pathology associated with idiopathic Parkinson's disease (preclinical and clinical stages). *J. Neurol.* **249 Suppl 3,** III–1–5 (2002).
174. Sierra, M. *et al.* Olfaction and imaging biomarkers in premotor LRRK2 G2019S-associated Parkinson disease. *Neurology* **80,** 621–626 (2013).
175. Humphrey, S. J., Azimifar, S. B. & Mann, M. High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. *Nat. Biotechnol.* **33,** 990–995 (2015).

VITA

William Edelman was born in a small village in Ecuador. He has roots in New York and Albuquerque, NM. He developed a passion for both biology and theatre at The University of New Mexico, but ultimately chose to study insect systematics in Dr. Kelly Miller's lab in the Museum of Southwestern Biology. He became interested in the field of genomics, DNA sequencing and proteomics through his coursework and curiosity about the biology we cannot see or feel with our senses. He graduated with honors in Biology, BS. In addition to his undergraduate mentor, he lives with gratitude for the mentorship provided by the IMSD program directed by Dr. Margaret Werner-Washburne, Regents' Professor at UNM

In 2016, William earned his PhD in Genome Sciences from the University of Washington after completing his thesis research in the laboratory of Dr. Judit Villén.

William's career goals are to continue to pursue the study of neurodegeneration and cell signaling. He would one day like to teach, mentor and inspire undergraduates in the biological sciences just as others have kindly done for him.