ABSTRACT

Title of Document: MICROWAVE-SUPPORTED ACID HYDROLYSIS FOR PROTEOMICS

Joseph Romaine Cannon
PhD
2012

Directed By: Catherine Fenselau, Professor, Department of Chemistry and Biochemistry

Our goal is to develop, optimize and demonstrate workflows that incorporate rapid Asp-selective chemical proteolysis into proteomic studies of complex mixtures. This can be further divided into several specific aims. The first aim is to develop and optimize the sample preparation, mass spectrometric, and bioinformatic methods required for complex mixture analysis of peptides resulting from acid digestion both in solution and in polyacrylamide gels. Second, the optimized methods will be applied to three model systems. In the first application, the large peptides derived from microwave-supported acid hydrolysis of human ribosomes isolated from MCF-7 breast cancer cells are analyzed. Secondly, acid hydrolysis will be applied to characterize Lys63 linkages in polyubiquitins. Finally, all the above methods will be

combined for the analysis of extracellular vesicles shed by myeloid derived suppressor cells from a murine mammary carcinoma model.

After optimizing the mass spectrometric and bioinformatic methods required for analysis of peptides resulting from acid hydrolysis, the most comprehensive analysis using this digestion technique to date was achieved both for in gel and in solution analysis. In gel digestion resulted in identification of over twelve hundred peptides representing 642 proteins, and in solution digestion via mass biased partitioning allowed identification of over 300 proteins. Mass biased partitioning also resulted in two distinct peptide populations from the high and low mass analyses implemented.

Nearly 90% of the predicted human ribosomal proteins were identified after acid hydrolysis. High resolution analysis of both precursor and product ions resulted in an average sequence coverage of 46% among identified proteins. It was also demonstrated that microwave-supported acid hydrolysis facilitates a more informative method for analysis of Lys63 linked polyubiquitin. After acid hydrolysis, ~629 Da mass shifts were found to be indicative of isopeptides. These isopeptides were easily identified from complex mixtures using tandem mass spectrometry and diagnostic b ions. Extracellular vesicles from a murine carcinoma model were then analyzed using in gel microwave-supported acid hydrolysis, mass biased partitioning after in solution digestion, and the sample was interrogated for the presence of ubiquitinated peptides.

MICROWAVE-SUPPORTED ACID HYDROLYSIS FOR PROTEOMICS

By

Joseph Romaine Cannon

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
PhD
2012

Advisory Committee:
Professor Catherine Fenselau, Chair
Professor Marco Colombini, Dean's Representative
Associate Professor Douglas Julin
Assistant Professor Nicole LaRonde-LeBlanc
Assistant Professor Shuwei Li

Dedication

This work is dedicated to my wife, Carol, without whom this would not have been possible.

Acknowledgements

The work presented in this dissertation utilized the knowledge, experience, and guidance of many scientists and non-scientists, alike. First, I would like to acknowledge the guidance and training provided by my advisor, Catherine Fenselau. Her scientific guidance allowed me to complete the work presented herein, and she has been the driving force in making me the scientist I am today.

Additionally, this work would not have been possible without the intellectual guidance and support of several other scientists, specifically, two of our collaborators and friends, Dr. Peter Gutierrez and Dr. Nathan Edwards. Further collaborative efforts were conducted in conjunction with Dr. David Fushman and Mark Nakasone, and Dr. Suzanne Ostrand-Rosenberg. I was also fortunate enough to benefit from the expertise of the director of the Proteomics Core Facility, Dr. Yan Wang. Her guidance with experimental development and instrumentation was crucial for completion of this dissertation.

My fellow graduate students in the Fenselau lab, Dr. Colin Wynne, Dr. Karen Lohnes, Avantika Dhabaria, Meghan Burke, and Rebecca Rose, and Maria Oei provided intellectual stimulation, moral support, and valuable friendships. I would also like to acknowledge helpful discussion and support from Dr. Joshua Cherry. I would especially like to thank my lab partner and biggest supporter, Waeowalee Choksawangkarn, for her unwavering confidence and friendship.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACN	Acetonitrile
AGC	Automated Gain Control
CID	Collision induced dissociation
CNBr	Cyanogen bromide
Da	Dalton
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECV	Extracellular vesicle
ESI	Electrospray ionization
Evalue	Expectation value
FBS	Fetal bovine serum
FDR	False discovery rate
FT	Fourier transform
FT-ICR	Fourier transform ion cyclotron resonance
GO	Gene ontology
HPLC	High pressure liquid chromatography
HSQC	Heteronuclear singel quantum coherence
IP	Immunoprecipitation
kDa	kilodalton
LC	Liquid chromatography
LC-MS	Liquid chromatography mass spectrometry
LC-MS/MS	Liquid chromatography tandem mass spectrometry
LTQ	Linear trap quadrupole
MALDI	Matrix assisted laser desorption ionization
MDSC	Myeloid derived suppressor cell
MS	Mass spectrometry
MWCO	Molecular weight cutoff
NMR	Nuclear magnetic resonance
PBS	Phosphate buffered saline
PPM	Part per million
PSI	Pounds per square inch
PTM	Post translational modification
RNA	Ribonucleic acid
RP-HPLC	Reversed phase high pressure liquid chromatography
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
TOF	Time of flight
UniprotKB	Uniprot Knowledge Base

List of Appendix Tables

Chapter 1: Introduction

Introduction

Proteomic experiments can be divided into three distinct but equally important parts. In the order in which they occur, those parts are sample preparation, mass spectrometry, and bioinformatic analysis. This work will focus on sample preparation using microwave-supported acid hydrolysis, and the changes it requires in the mass spectrometric and bioinformatic parts of the experiment. First, it will examine two methods to enhance analysis of microwave-supported acid hydrolysis produced peptides from complex mixtures. In the first method, the acid hydrolysis reaction is coupled to polyacrylamide gel electrophoretically fractionated protein mixtures. The second method enhances analysis of proteins digested in solution by partitioning the resulting peptides based on size, facilitating mass spectrometric and bioinformatic methods catered to the two distinct peptide populations. The work will also demonstrate several important applications of microwave-supported acid hydrolysis. It will demonstrate successful analysis of the human ribosome, a novel method for extracting and analyzing oligomeric state specific peptides from Lys63

linked polyubiquitin, and finally, all of these methods will be combined for analysis of extracellular vesicles isolated from myeloid derived suppressor cells.

Mass Spectrometry Based Proteomics

The word "proteome" was coined by Marc Wilkins in 1996[1] and can be defined as the entire protein complement of a given organism under specified conditions. The study of the proteome, or "proteomics" is the natural next of kin of all the previously generated genomic sequencing data. Unlike an organism's genome, which is generally static, the proteome is ever changing in response to stimuli. In addition to the changing expression levels that proteins demonstrate, there are myriad post translational modifications (PTM) that lead to specific consequences. Despite its official definition proteomics has come to signify a specific type of experiment in which a protein or set of proteins is analyzed (and sometimes quantitated) using mass spectrometry. Mass spectrometry (MS) is well suited for the task of identifying proteins and peptides since it characterizes molecular populations based on mass, a property that with sufficiently high accuracy can be directly indicative of empirical formula, and rather fortuitously, eighteen of the twenty commonly occurring amino acids have unique elemental compositions (and masses). Proteomic analysis by MS offers a unique analytical perspective on biological problems. It allows simultaneous analysis of a large number of proteins in a relatively unbiased manner. Proteomic methods display an impressive degree of analytical versatility, allowing characterization of entire cell lines[2] at one extreme to in depth analysis of single proteins at the other. This work will first introduce the tools required for the mass

spectrometry and bioinformatic parts of a proteomic workflow in terms of two commonly used experimental approaches for proteomic analysis, termed "bottom up" and "top down".

Electrospray Ionization

Advances in 'soft' ionization methods have allowed introduction of large biomolecules into the gas phase without artefactual fragmentation.[3, 4] The advent of these new technologies and their effect on allowing biomolecular analysis culminated in reception of the Nobel Prize in Chemistry in 2002 by Tanaka and Fenn.[5] Electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) have facilitated rapid advancement in proteomic analytical technology, and are still the most widely used ionization techniques in biological mass spectrometry. Although both ionization methods were used in this work, this dissertation will focus on ESI because of its proclivity for complex mixture analysis when coupled with liquid chromatography.

Electrospray ionization is the method of choice for sample introduction into the MS following fractionation of a complex mixture by liquid chromatography (LC), which has long been known to offer simultaneous separation and concentration of analytes of interest with a wide variety of possible stationary phases. In positive polarity ESI, a potential is applied between the inlet of the mass spectrometer and an acidified liquid sample (often directly eluting from an LC column), which is forced through a small diameter capillary drawn to a point. In the presence of this potential,

the acidic sample is desolvated en route to the MS inlet. The electrosprayed peptides are multiply protonated during the desolvation process. For small peptides this will often lead to doubly and triply charged ions, and the number of acquired charges increases proportionally with size (with some discrepancies observed due to gas phase protein structure).[6] Neutral protein and peptide molecular weights can be determined from the observed mass to charge ratios based on the following formula:

$$m = [(m + zH/z) * z] - [z^1 H_{mass}] \qquad \text{(Equation 1)}$$

In the relationship above, the mass, m, is equal to the observed charge density multiplied by the number of associated charges, and subsequently the product of the charge and the elemental proton mass is subtracted. Additionally, large species will produce a distribution of charge states that are representative of different gas phase conformations, which results in dilution of the total ion signal for a single polypeptide species. The extent of this phenomenon is dependent on polypeptide length.

Coupling with High Pressure Liquid Chromatography

As mentioned previously, high pressure liquid chromatography (HPLC) is well suited for coupling with ESI. In HPLC, analytes are partitioned between mobile and stationary phases depending on their affinity for each. For proteomics applications, the most common stationary phases capitalize on hydrophobic interactions with proteins and peptides, often called reversed phase chromatography. These stationary phases are usually composed of a porous silica based bead of known

particle (between 1.7 and 5µm) and pore (100-1000Å) size attached to a hydrocarbon of defined length (common lengths are butyl (C4), octyl (C8), and octadecyl (C18)). Liquid chromatography tandem mass spectrometry (LC-MS/MS) experiments are also made possible by the convenient solvent system used in reversed phase HPLC (RP-HPLC). In positive ion mode LC-MS/MS experiments, peptides are loaded onto a RP column in a solvent that is nearly all aqueous. Solvents are substituted with 0.1% formic acid to aid in the ESI process. Once on the column, peptides or proteins are eluted by steadily increasing the proportion of less polar organic solvent (usually acetonitrile (ACN)) going through the column. Introduction into the MS via ESI occurs upon elution from the HPLC column.

LTQ Orbitrap

Electrospray has been used in conjunction with many types of mass analyzers, but this dissertation will focus on two that are often paired together. Ion traps and Fourier transform based analyzers, specifically Orbitraps, have become a popular instrument configuration in recent years. The two mass analyzers have duty cycle and resolution limitations that offset one another, but that function well when combined.

Ion Traps

Wolfgang Paul was awarded the 1989 Nobel Prize in physics for inventing the ion trap (sometimes referred to as the Paul trap).[7, 8] Ion traps function based on oscillating electric and radio frequency potentials applied in three dimensions to maintain ion populations in regions of high stability. This method proves useful for MS/MS in which ions of interest (or more accurately m/z regions of interest) are isolated, fragmented, and their resulting product ions are detected upon collision with an electron multiplier. Ions outside the m/z region of interest are ejected based on the principle of mass selective instability.[9] This method allows coupling of a specific precursor ion with its associated product ions that form following collision induced dissociation (CID) with an inert gas, such as He. The most common method used to produce the aforementioned fragment ions is CID, although there are alternative methods.[10] The ion trap is filled with a bath gas, and by rapidly oscillating the end cap electrode potentials, the isolated peptide population undergoes multiple collisions with the bath gas molecules present in the trap, and acquires energy which is redistributed throughout the peptide. Energetic redistribution following this slow heating method results in peptide bond cleavage, with the weakest bonds being cleaved in highest abundance. **Figure 1** demonstrates commonly observed backbone cleavage patterns and the associated nomenclature for the resulting ions.

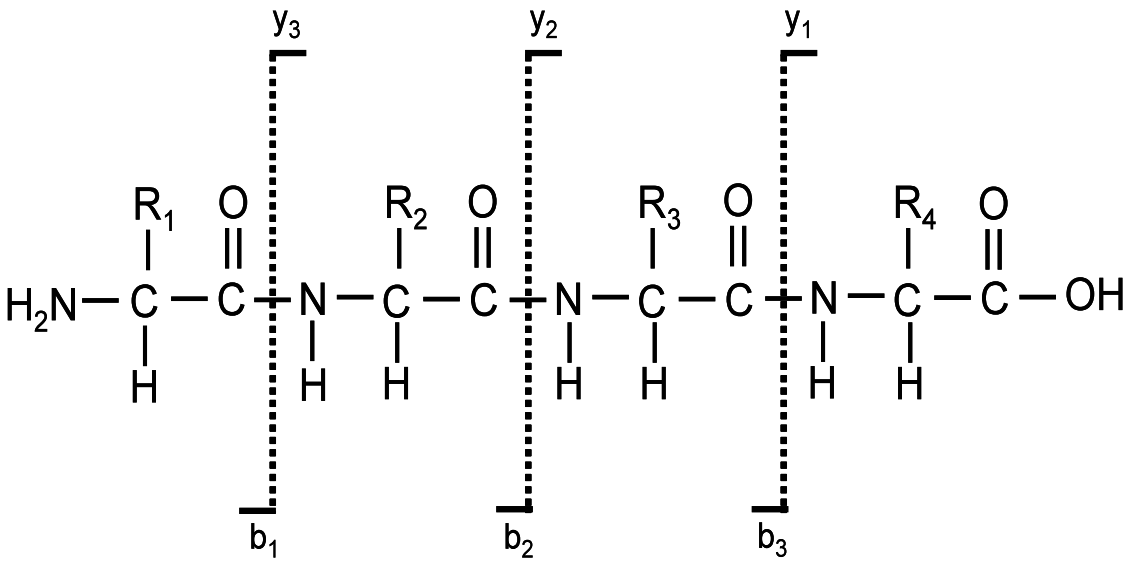

Figure 1. Location of bond cleavage resulting in b and y ions, and its associated nomenclature.

Using CID, backbone cleavage occurs predominantly at the peptide bond, forming b and y ions that contain the amino and carboxy termini, respectively. Using these specific cleavage patterns in conjunction with predicted ion populations from a given sequence it is possible to search precursor masses along with their resulting fragment ions against databases produced *in silico* in a high throughput fashion.[11] Ion traps are specifically well suited for this type of analysis due to their rapid duty cycle. This allows analysis of multiple peptides in a short time frame (on the order of tens of milliseconds), a prerequisite for complex mixture analysis.

The Orbitrap

The Orbitrap was invented by Makarov and coworkers[12] and has since become a staple of high resolution proteomic analysis. Unlike Fourier transform ion

cyclotron resonance mass spectrometers (FT-ICR-MS), which also provide high resolution data, Orbitraps do not require a superconducting magnet to function. Similar to ion traps, charged molecules in the gas phase are confined between electrodynamic potentials. The unique shape of the Orbitraps inner and outer electrodes between which the ions are confined (**Figure 2**) forces them to oscillate about a central axis at a given frequency, ω, dependent on the mass (m), charge (q), and an instrument specific electric field constant, k, with the following relationship:[13]

$$\omega = \left(\frac{kq}{m}\right)^{1/2} \qquad \text{(Equation 2)}$$

Frequencies are acquired in the form of a transient image current, and subsequently undergo a fast Fourier transformation (FT) to arrive at the observed mass spectrum.

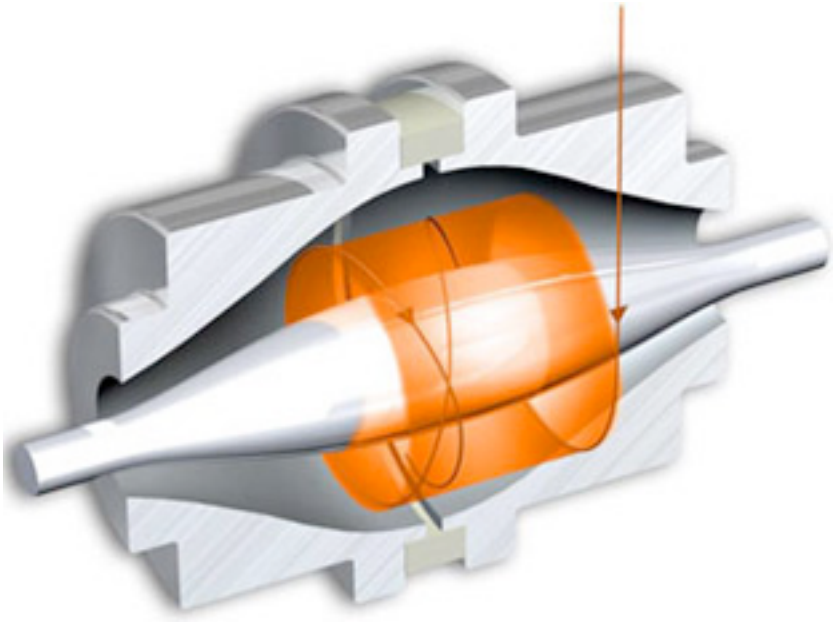

Figure 2. Cutaway view of the spindle shaped inner electrode and the surrounding outer electrode of the Orbitrap mass analyzer.

Like all methods that utilize FT, the resolution of the transformation product exhibits a linear dependence on the frequency transient acquisition time. Additionally, resolution decreases with changes in charge density according to the following relationship:[13]

$$\frac{m}{\Delta m} = \frac{1}{2\Delta\omega}\left(\frac{kq}{m}\right)^{1/2} \qquad \text{(Equation 3)}$$

In contrast to the high scan speed and unit resolution of the ion trap analyzer, Orbitraps achieve the high resolution required for large polyeptide analysis at the expense of acquisition time (**Figure 3**).

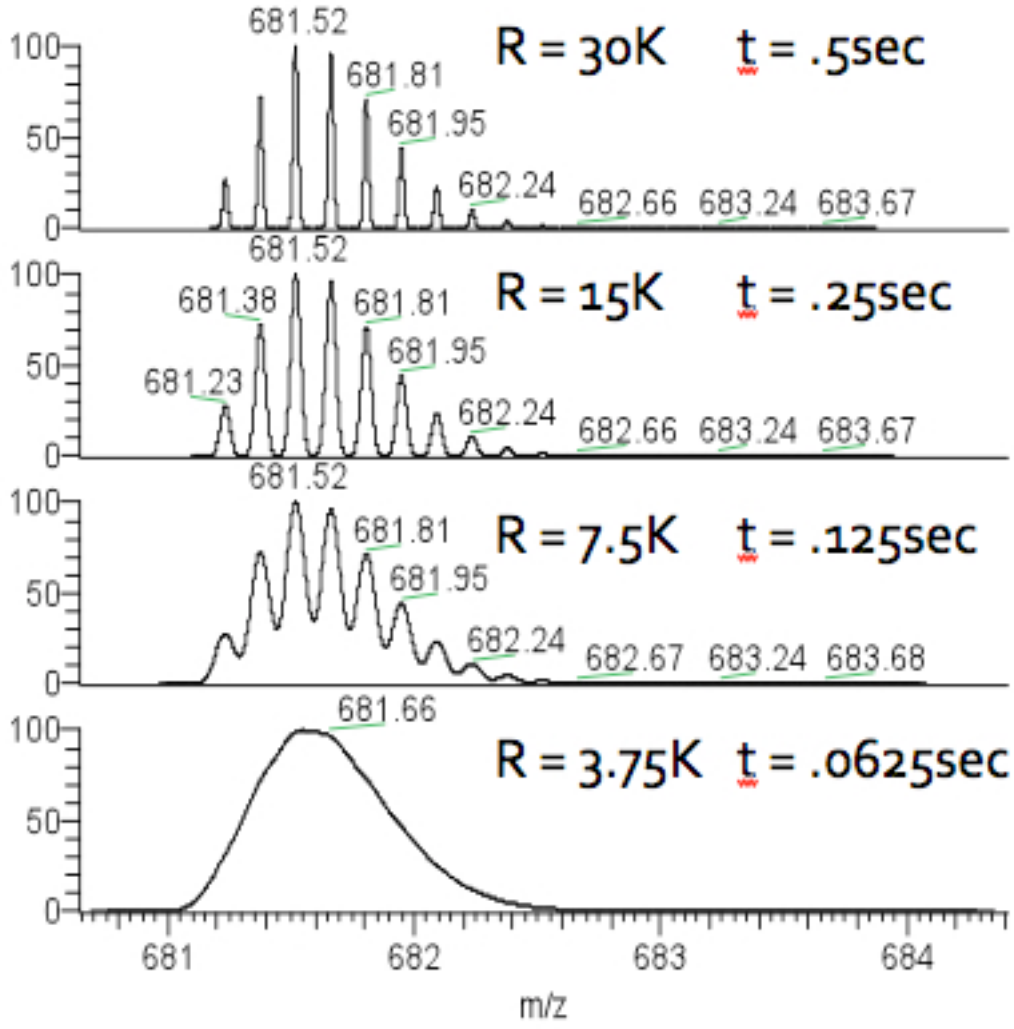

Figure 3. Simulated isotopic spacing at differing resolutions, *R*, of a randomly selected ribosomal peptide VNVPKTRRTFCKKCGKHQPHKVTQYKKGKDSLYAQGKRRY at charge state z = 7. The time required for transient resolution, *t*, is shown at right.

Hybrid instruments have joined the two analyzers to optimize analysis by allowing high resolution and high mass accuracy precursor analysis (in the Orbitrap) with concomitant isolation, fragmentation, and detection of multiple ions of interest at low resolution (in the ion trap).

Methodologies

Bottom Up

Proteomic methods utilize the tools outlined above in several ways dependent on the system being interrogated. Generally, there are two methodologies that differ in the presence or absence of a proteolytic step prior to introduction into the mass spectrometer. Bottom up methods use trypsin, a protease that hydrolyzes peptide bonds C terminal to unmodified Arg and Lys residues unless they are followed by Pro, to produce a high number of small peptides that can be fragmented and detected in the mass spectrometer with high efficiency. After cleavage C terminal to basic residues and introduction into the gas phase, peptides are likely to sequester protons at each terminus (at the primary amine on the amino terminus and at the side chain primary amine on Lys or the guanidyl group on Arg). The chemical result of this circumstance is that both b and y fragment ions produced during the CID process are likely to have an associated proton and be detected. Tryptic digestion has become such an integral part of proteomics that some bioinformatic search methods will utilize predicted relative ion abundances based on well characterized fragmentation schemes of tryptic peptides rather than just observing the presence or absence of an expected peak.[14, 15] Although the small peptide size is convenient for fragmentation using CID, the large number of peptides produced from most protein mixture samples makes LC-MS/MS analyses very complex. Tryptic peptides generally provide

identification of a relatively small percentage of a protein's sequence. The combination of low sequence coverage and high number of similarly sized peptides requires that confident protein identification be restricted to those in which two or more tryptic peptides are observed with high confidence. Despite this stringent bioinformatic requirement, identification of thousands of peptides and hundreds of proteins in a single experiment has become routine.

Top Down

Top down methods lack a proteolytic step, and introduce intact proteins into the mass spectrometer directly. Analysis without prior proteolysis provides 100% sequence coverage of the identified protein automatically, and also provides information on relative abundances of post translationally modified isoforms.[16] The intact mass is analyzed at high resolution to obtain an accurate neutral mass, and subsequently the protein is fragmented with product ions also being detected at high resolution. High resolution analysis is required due to the fact that large molecular weight species acquire a high number of charges in ESI-MS. With isotopic resolution, charge state can be rapidly determined according to the following formula:

$$z = \frac{1}{(m/z_{^{13}C\,n} - m/z_{^{13}C\,n-1})} \qquad \text{(Equation 4)}$$

The assumed difference between adjacent ^{13}C isotope peaks is approximately 1Da; therefore, the inverse of the difference in m/z of adjacent isotope peaks is equal to the charge state (**Figure 4**).

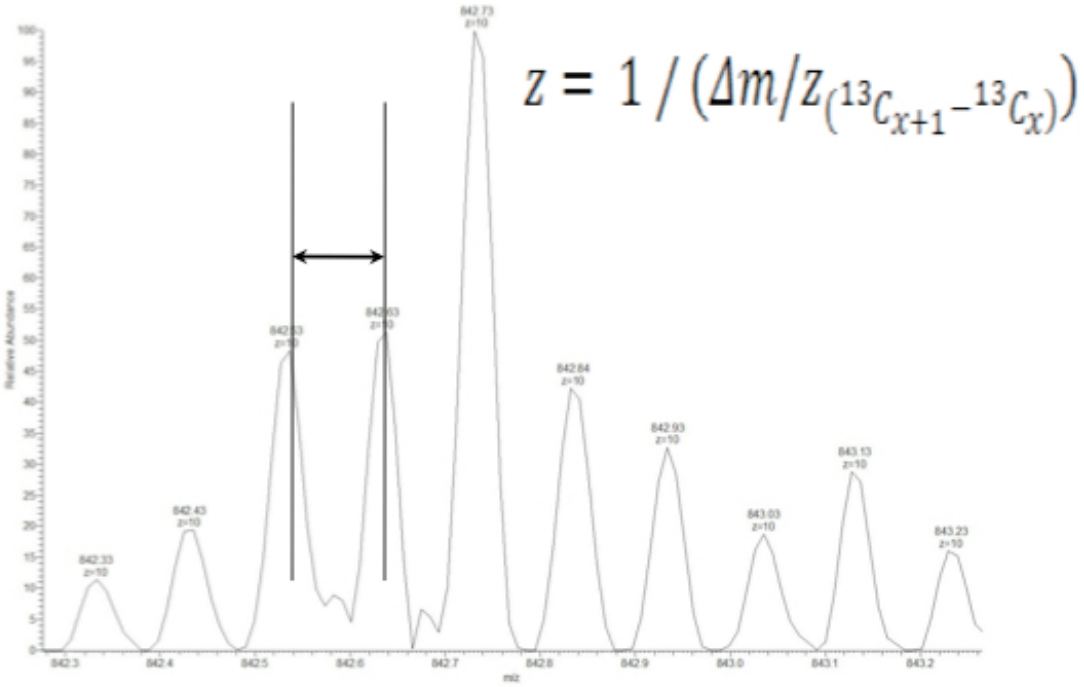

Figure 4. Isotopic spacing of a peptide and the accompanying formula used to determine its charg state.

Accurate charge state determination is required for attaining an ion's neutral mass using Equation 1. Bottom up experiments do not require high resolution since product ions are assumed to have charge states of z = 1 or z = 2, based on the relatively small size of the peptide precursors. Top down analyses were initially limited to single protein infusion-type experiments, in which the protein is put in solution at higher concentration and electrosprayed into the MS directly without any LC fractionation.[17] The time dependence associated with the FT-based mass analyzers required for large protein analysis is poorly suited for liquid

chromatography (LC) for several reasons. Primarily, chromatographic peak widths (in seconds or minutes) may be incompatible with the time required for high resolution accurate mass spectral acquisition. In addition, the ESI process, as stated previously, produces multiple charge states of the same species with polypeptides of higher mass. This phenomenon results in a dilution of total analyte signal across multiple charge states.[18] Complex intact protein mixtures fractionated by LC beforehand are likely to have peaks with overlapping charge state distributions. Tandem mass spectrometry is additionally complicated for complex protein mixtures due to the fact that usually the most abundant peaks, where abundance is the criterion used for precursor selection in most data dependent MS/MS methods, are different charge states of the same precursor. For these reasons, top down methods have not yet been adopted by the proteomics community for complex mixture analysis, barring a few exceptions.[19, 20]

Middle Out

This work will focus, in part, on a compromise between the two methods outlined above, termed middle out (also known as middle down). Middle out peptides are between 3 and 20kDa in mass,[21, 22] and are usually produced using single residue specific proteolysis, although other methods have been employed.[21] This hybrid method utilizes high resolution analysis, although not so high that it can only be performed using an FT-ICR instrument. The method also uses proteolysis, but the resulting peptides are larger on average than those produced using trypsin, and

therefore, provide higher sequence coverage. The benefits and drawbacks of this method will be explored further in the following chapters.

Bioinformatics

Search Algorithms for Bottom Up

Bottom up proteomics experiments produce a large amount of data. Hybrid instruments, like the LTQ-Orbitrap described above, routinely acquire tens of thousands of spectra over the course of a single LC-MS/MS experiment. Datasets like this are too large for manual interpretation of each pair of precursor and associated product ion spectra, so high throughput bioinformatic search methods have become an integral part of proteomics experiments. Proteomic search algorithms utilize *a priori* knowledge of the analyte and sample preparation methods for matching against predicted peptides from a set of protein candidates.[23-25] For MS/MS experiments, inputting a product ion spectrum and the m/z of the ion from which it is derived will return a list of peptides likely to be in the sample based on observed fragment ions (**Figure 5**). First, the specificity of the enzyme used in the proteolysis step is used to compile a list of peptides *in silico* from a set of candidate proteins, which are usually defined by translation of genomic sequences of a given phylogenetic taxon. Subsequently, a precursor ion is searched against the *in silico* produced peptide masses within a user defined mass tolerance. If there are candidate peptides within that mass tolerance, experimentally observed product ions are compared with predicted b and y fragment ions (for CID produced product ions). Candidates are then scored based on the number of observed fragment ions.

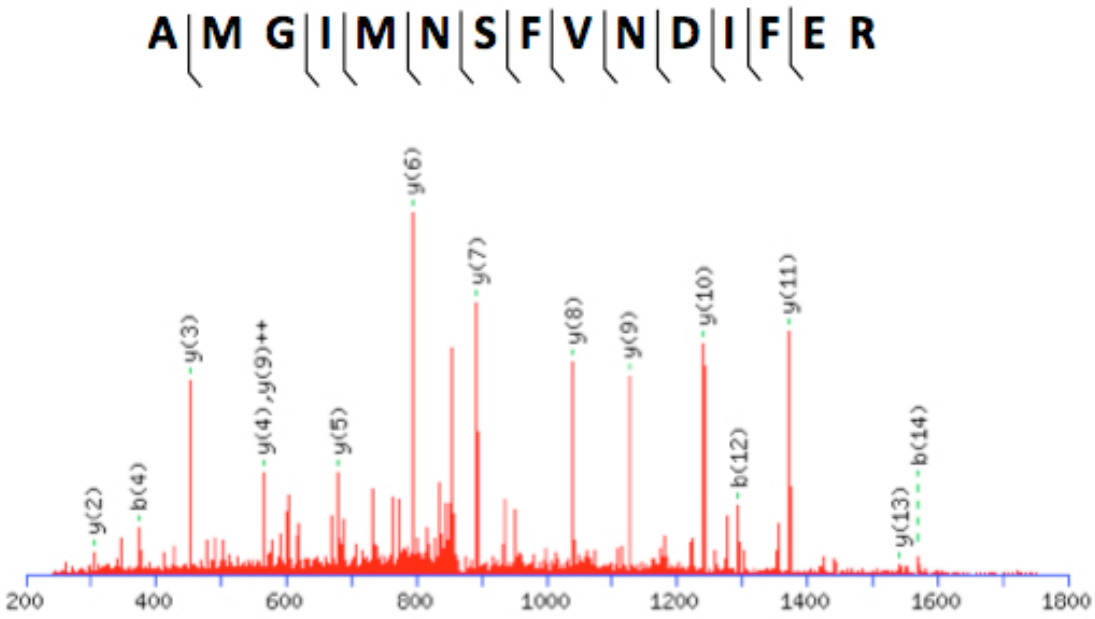

Figure 5. Product ion spectrum matched to a peptide from histone H2B type-1 from *M. musculus*. The nearly complete y ion series facilitated a high confidence assignment.

A common metric for gauging the confidence of peptide identifications is the expectation value (Evalue). Formulae that mathematically describe an algorithm's Evalue differ from one search engine to another, but practically Evalues are the number of additional peptides that match a given precursor and product ion pair with equal or better scores by random chance alone. This is used in conjunction with an additional parameter, the false discovery rate (FDR). False discovery rates are calculated by searching mass spectral data against a database of intentionally false sequences, which are formulated by shuffling or reversing the individual protein sequences in the true forward database. Once these two parameters have been

defined, all 'true' matches in the forward database must have Evalues better than the Evalue of the highest confidence match in the reverse or shuffled database. Applying these stringent parameters achieves a 0% FDR.

Search and Decharging Algorithms for Large Peptides and Proteins

Large peptides and proteins, as stated earlier, require neutral mass determination of both precursor and product ions prior to database searching. Search engines that can match large peptides and proteins to candidate sequences must first decharge all mass spectra. For the large quantity of spectra in a given LC-MS/MS experiment this is performed rapidly by comparing experimentally observed isotope clusters for each MS peak with a cluster predicted from a polypeptide composed of Averagine, a theoretical amino acid weighted for the relative abundance of all twenty commonly occurring amino acids.[26] After attaining the neutral mass of all peaks, the algorithms take advantage of the high mass accuracy measurements for highly specific matching to predicted product ions resulting in a low FDR. The search program utilized for this dissertation, ProSightPC (ThermoFisher, San Jose, CA), was designed to accommodate post translational modifications, and therefore does not associate any additional confidence for high accuracy precursor mass measurement. This peculiarity was best demonstrated by Wynne and Fenselau in the assignment of a single amino acid substitution for DNA binding protein HU Alpha from *Erwinia herbicola* (Figure 3).[27]

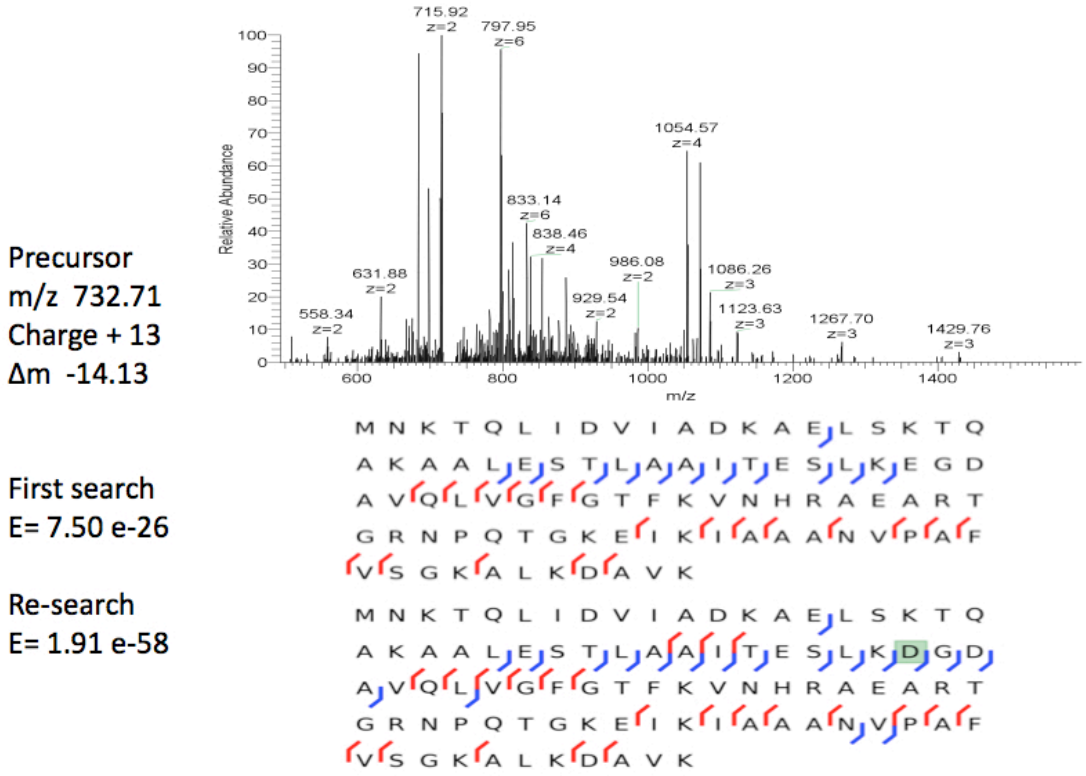

Figure 6. Product ion spectrum of a z = 13 precursor at m/z 732.71. The ion was matched to the DNA binding protein HU alpha with a ~14 Dalton mass difference (top sequence). Applying the observed mass shift to a region lacking overlapping b and y ions resulted in an increase in confidence due to a higher number of overlapping b and y ions. Adapted from reference 27.

In this example, the spectrum was matched to a sequence based on the high accuracy fragment ions, but with a precursor mass difference of ~14 Daltons (Da). Examination of the matched fragment ions shows that there are no ions overlapping the region containing "EGDAV". The authors then localized the ~14Da mass shift to this region. Implementing a mass shift to the only amino acid that can accommodate this mass change (a Glu to Asp substitution) resulted in an increase in the total number of matched ions with many of them overlapping one another. Following

identification using these methods, the results must be corrected to attain a FDR as describe above.

Proteomics Databases

The current permeation of proteomics in biological research is due largely to genomic sequencing. Generally, there are a few categories that describe the evidential basis of a protein database. They are either composed of translated genomic data, manually curated, or some combination of the two. Genomic databases are characterized by a higher number of total sequences, but there are frequently redundant entries. Manually curated databases typically have fewer sequence redundancies and a greater amount of functional and biological information per entry. In this work, all final identifications were taken from sequences and entry specific information (cellular component, biological function, etc.) present in the UniProt Knowledgebase (UniProtKB).

Objectives

Scientists undertaking proteomic experiments are generally exposed to a single method for analysis. The typical proteomic experiment employs tryptic digestion either in gel or in solution and low resolution mass spectrometric fragment ion analysis of small peptides providing little more than identification of the proteins present in the sample. Microwave-supported Asp-selective hydrolysis is a rapid and

efficient proteolytic reaction that has been shown to reliably produce peptide sets in which the expected cleavage sites predominate. To date, this method has not been successfully applied to mixtures of greater complexity than the yeast ribosome. The primary aim of this dissertation is to demonstrate, through two methods development experiments applied to complex mixtures and several novel applications, that alternative methods to the tryptic standard, in this case, microwave-supported acid hydrolysis, can provide a comparable amount of information and are sometimes better suited for the sample under interrogation.

Chapter 2: Microwave-supported Acid Hydrolysis

Single Residue Specific Proteolysis

Bottom up methods have provided an extensive and solid foundation for proteomic analysis. These methods rely on the enzyme, trypsin, which cleaves C terminal to Arg and Lys residues unless they are modified or followed by Pro. Despite the ease of use and reliability of bottom up methods, there are numerous benefits to using single residue specific methods that cannot be achieved using trypsin. The most obvious is an overall increase in average peptide length. The extent of the change in average peptide length is dependent on the specificity of the enzyme or chemical used. There are several readily available options that will cleave at a single residue; cyanogen bromide (CNBr) (C terminal of Met), and the enzymes GluC, AspN, ArgC, LysC, and LysN. The relative abundance of each one of these amino acids, as determined by an *in silico* digest of the human proteins present in the UniprotKB shows that of the above options, Met is the least abundant followed by Asp, Lys, Arg, and Glu (**Figure 7**).

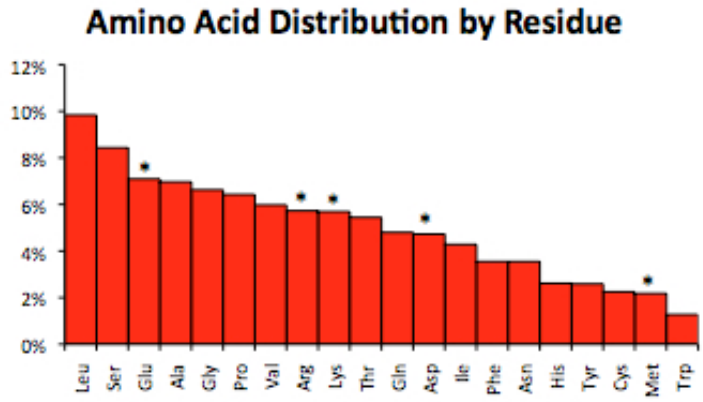

Figure 7. Shown above is a histogram demonstrating the relative abundance of the twenty essential amino acids present in human proteins in the UniProtKB. The abundances of residues commonly exploited for single residue specific proteolysis are denoted by an asterisk.

Despite its theoretical promise, widespread use of CNBr hydrolysis at Met residues for proteomics is marred by several factors. Primarily, the reagent is toxic and requires specialized accommodation for safe use and disposal. Additionally, there are possible undesirable side reactions in which peptide bond hydrolysis does not occur. Despite these drawbacks, several investigators have found CNBr to be a worthwhile method for proteomic analysis. Few methods use CNBr alone, but it is frequently paired with another enzymatic or chemical cleavage method. The high concentrations of acid used in CNBr digestion seem to invite dilution and subsequent acid hydrolysis. Indeed, Baek and co-workers observed a nearly 3-fold increase in the number of proteins identified using this method as well as an increase in average sequence coverage when compared to analysis of the same membrane enriched sample digested with trypsin.[28] Although the relative abundance of Met residues in

the human proteome makes CNBr an attractive option in middle out proteomics, the safety hazards and possible undesirable side reactions associated with its use often outweigh the benefits.

The enzymes AspN, LysC , LysN, and ArgC all produce similar average length peptides to one another based on their respective amino acid distributions. Some enzymes, including the AspN protease isolated from *Pseudomonas fragi*, are metalloproteases, meaning that they require a metal cofactor to cleave proteins.[29] This circumstance makes them incompatible with workflows that utilize metal chelators like EDTA, a commonly used blood anti-coagulant, in their buffers.[30] Despite the drawbacks mentioned above, all of these enzymes have their respective places in the proteomics toolkit. The common theme among all of them, however, is that they are enzymes. All of these enzymes have conditions under which they perform optimally, usually near neutral pH and little to no denaturant. Once optimal conditions are achieved, catalysis is allowed to proceed typically overnight.

The increase in average length displayed by single amino acid specific enzymes and chemical proteolytic methods is accompanied by a decrease in the total number of peptides produced from a complex protein mixture. A reduction in sample complexity facilitates more efficient chromatographic fractionation prior to introduction into the mass spectrometer. Karger and colleagues first alluded to this fact in their extended range proteomic analysis publications using single residue specific proteolysis of standard proteins with the enzyme LysC.[31-33] The importance

of the idea was later underscored by Mechtler and co-workers, when they demonstrated a linear relationship between chromatographic peak capacity and the number of proteins identified in a given LC-MS/MS experiment.[34] The significance of a decreased number of peptides for such a complex sample is obvious, but previous work by Swatkoski and co-workers demonstrated that fewer peptides from an equivalent sample will benefit even samples of low to moderate complexity using the *S. cerevisiae* ribosome as a model system.[35] The same publication utilized *in silico* digestions to show a high number of peptides with similar size and chemical properties that are likely to co-elute using reversed phase liquid chromatography (**Figure 8**).

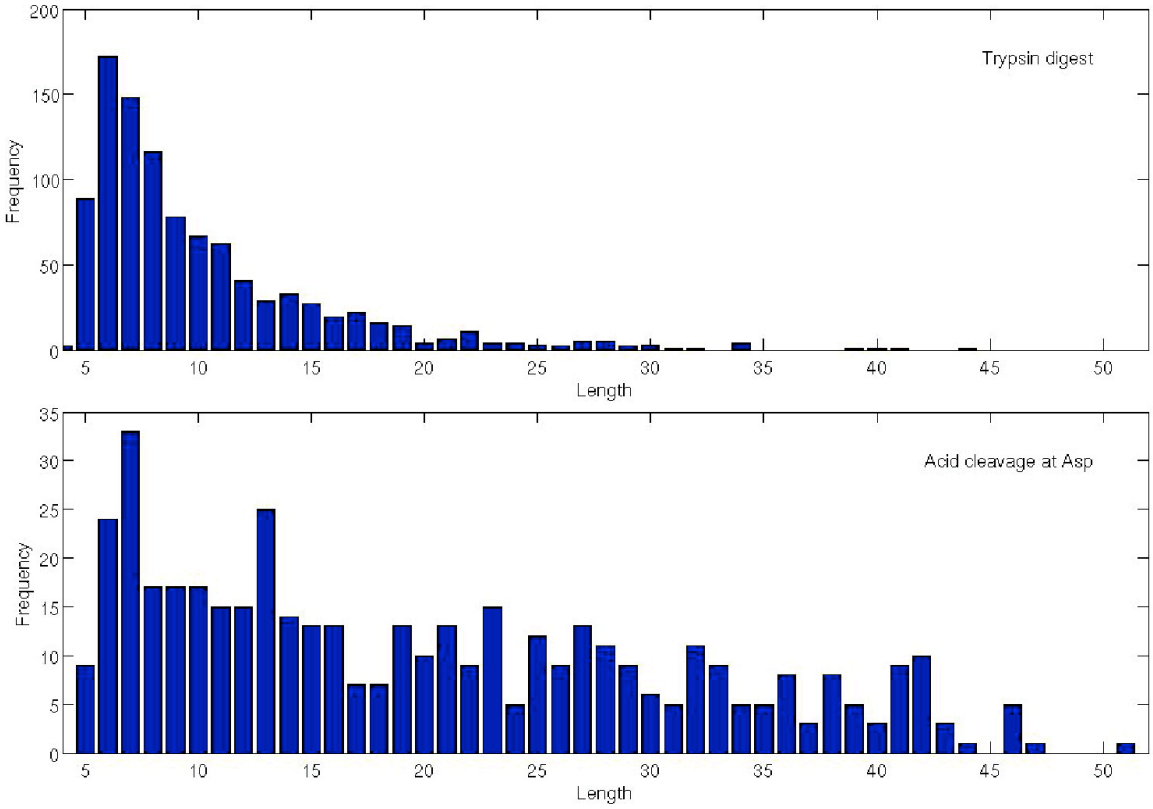

Figure 8. Comparison of peptide lengths of *S. cerevsiae* ribosomal proteins following digestion with trypsin (top) and Asp specific digestion (bottom). Note the difference in Y axis. Reprinted with permission from reference 35.

Longer peptides also benefit proteomic analysis by their size alone. An increase in sequence coverage provides more information on a given sequence of a protein. This is particularly well suited for proteins that display multiple modifications that work in concert. Histones, which are extensively post translationally modified proteins responsible for packaging chromatin within a cell's nucleus, are a perfect example of this phenomenon. According to the UniProtKB, histone H2A.X (accession number P16104) from *H. sapiens*, can contain up to six modified residues with an additional four modifications inferred by sequence

similarity. This ~15kDa protein contains only two internal Asp residues, resulting in three peptides following Asp specific acid hydrolysis. The high number of post translational modifications working in concert with one another require an analytical method that can observe as many PTMs as possible on a single peptide. Additionally, Histones are DNA binding proteins, and in order to perform their required function they must contain a high occurrence of basic residues. The relative abundance of Lys and Arg residues in DNA binding proteins is increased significantly compared to the entire human proteome. These two reasons, the ability to observe long peptides containing multiple PTMs that work cooperatively, and a high incidence of basic residues that would result in peptides too small for analysis following tryptic digestion, make Asp-selective hydrolysis an attractive option. An additional model system of nucleotide binding proteins, the human ribosome, will be examined later in this work.

Microwave-supported Acid Hydrolysis

Acid Hydrolysis

High temperature incubation in acidic solutions has been employed for indiscriminate protein hydrolysis into their individual amino acids. Schultz and co-workers observed in 1962 that under acidic conditions used for total protein hydrolysis, lone Asp residues were produced at a rate at least 100 times greater than any other residues.[36] Later characterization of the reaction by Inglis and co-workers

led to the hypothesis that hydrolysis proceeds via two different mechanisms where the bond is cleaved between Asp and the adjacent N- or C-terminal amino acid.[37] At temperatures greater than 108°C, the reaction will proceed indiscriminate of the acid used as long as the pH is below the pK_a of the beta carboxyl group of the Asp side chain. Li and co-workers demonstrated in 2001 the utility of Asp specific acid cleavage for proteomics experiments.[38] Using a heating block to keep the temperature at 108°C, the reaction was allowed to proceed for up to 4 hours in 2% formic acid. Laboratories later demonstrated the temperature could be easily maintained with a microwave oven, a technique already adopted in the synthetic organic chemistry community. Microwave irradiation was found to allow selective protein hydrolysis to be achieved in minutes, instead of the hours previously required.[39, 40]

Two mechanisms are illustrated in **Figure 9**, which allow proteolysis to occur C-terminal and N-terminal to Asp residues. The Asp side chain cyclizes, forming 6- or 5-membered rings with carbonyl groups on either side. Model systems favor hydrolysis of the 5-membered ring,[41, 42] supporting previously observed results from Inglis and others that peptides resulting from C-terminal Asp cleavage are more abundant.

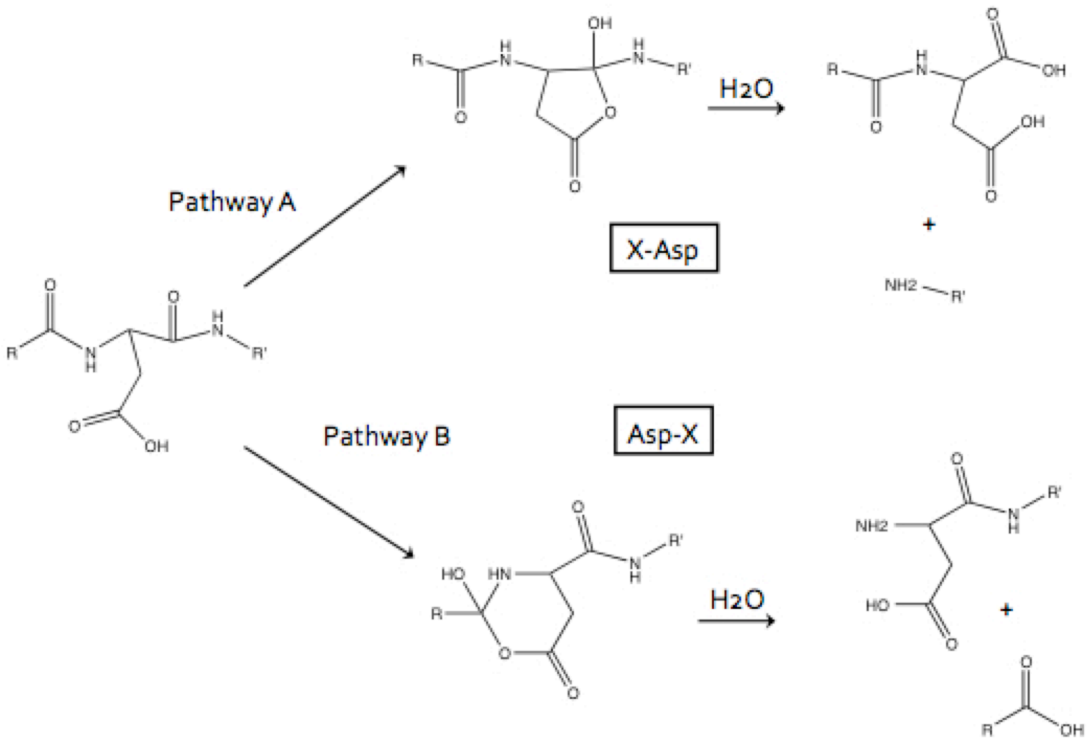

Figure 9. Mechanisms of Asp-selective acid hydrolysis. Adapted from reference [38].

Acid hydrolysis has been demonstrated using several acids,[43-45] consistent with the idea that Asp side chain protonation is a requirement. Inglis' initial characterization defined pH 2.0 (achieved with both HCl and formic acid), at a temperature of 108°C over the course of a 2 hour incubation period as optimal reaction conditions.[37] Hydrochloric and formic acid are also known to denature proteins, facilitating protonation and proteolysis. Formic acid and acetic acid are employed by most laboratories at concentrations ranging from 0.05% to 12.5%.[22, 30, 35, 39, 40, 43, 44, 46-51] Low abundance side reactions include N-terminal pyroglutamate formation, formylation, and presence of the dehydrated cyclic intermediates. Proteomics search engines will readily allow for all of which. Fenselau and co-workers observed that incubation with formic acid resulted in formylation even under

short incubation times with simple mixtures, a phenomenon not observed using acetic acid.[47] Koomen and co-workers demonstrated that selective acid hydrolysis can even be implemented with the acidic matrices used for MALDI-MS.[52] Hydrolysis with stronger acids like trifluoroacetic acid at higher concentrations leads to non-specific cleavage, likely due to a higher proportion of hydronium catalyzed hydrolysis.[53, 54]

In general, incubation time increases with sample complexity. The best results have been obtained when total protein concentration does not exceed 0.1mg/mL, irrespective of sample complexity. Protein disulfide bond reduction has been demonstrated by the addition of dithiothreitol (DTT) to the reaction vessel.[44, 46] The high temperature acidic conditions hydrolyze both O-linked glycans and phosphorylations.[38, 43]

Microwave-supported Acid Hydrolysis

Microwaves lack sufficient energy for molecular rearrangement, and are therefore non-ionizing. In an electromagnetic field, dipoles will attempt to constantly align themselves with the electric field. This frequent dipolar rotation provides dielectric heating. In addition, microwave absorption and heating is solvent dependent.[55] The physics governing microwave energy absorption are beyond the scope of this dissertation; however, Lidstrom, et al can provide a comprehensive review of microwave-supported chemistry.[55]

Microwaves are a type of electromagnetic radiation with a wavelength between 1 and $1e^{-3}$ meters. The specific frequency of 2.45 GHz is commonly used to support chemical reactions. Initial experiments in this laboratory utilized a domestic microwave oven operated at 745W;[39] however, subsequent attempts from this laboratory as well as others progressed to research grade microwaves with tunable power settings.[30, 35, 43, 44, 46-48, 50] Asp-selective acid digestion has been performed in domestic microwaves, and with non-microwave induced heating,[38, 40, 45, 52] but microwaves specifically made for chemical reactions have convincing benefits. Kitchen microwaves irradiate non-uniformly. While this is a non-issue for food preparation, which takes up significant space in a domestic microwave, the microscale sample volumes typically used in proteomics experiments necessitate more focused exposure. Commercial microwaves offer focused irradiation in a confined area with uniform energy distribution. As an added benefit, research grade microwaves facilitate temperature regulation to within $\pm 5\ ^{o}$C.

After demonstrating that microwave-supported acid hydrolysis works on standard proteins and bacteria, Swatkoski et al appraised its effectiveness for complex mixtures using both MALDI and ESI approaches,[35] Ribosomes were purified from *S. cerevisiae* and digested for 20 minutes in 12.5% acetic acid at 140 $\pm 5^{o}$ C. Using an automated bottom up approach, 247 peptides were confidently observed, representing 58 of the 78 expected ribosomal proteins. Ribosomal and other nucleotide binding proteins contain high numbers of Arg and Lys residues in their primary sequences. This makes them poorly suited for proteomic analysis with tryptic digestion, because many of the peptides produced are too small for MS analysis. **Figure 10** illustrates

smaller (red) and large (green) peptides predicted for each protein from the ribosome, and the sequence coverage provided. The bottom graph shows peptides produced by trypsin and the top graph shows those produced with Asp-selective acid digestion.

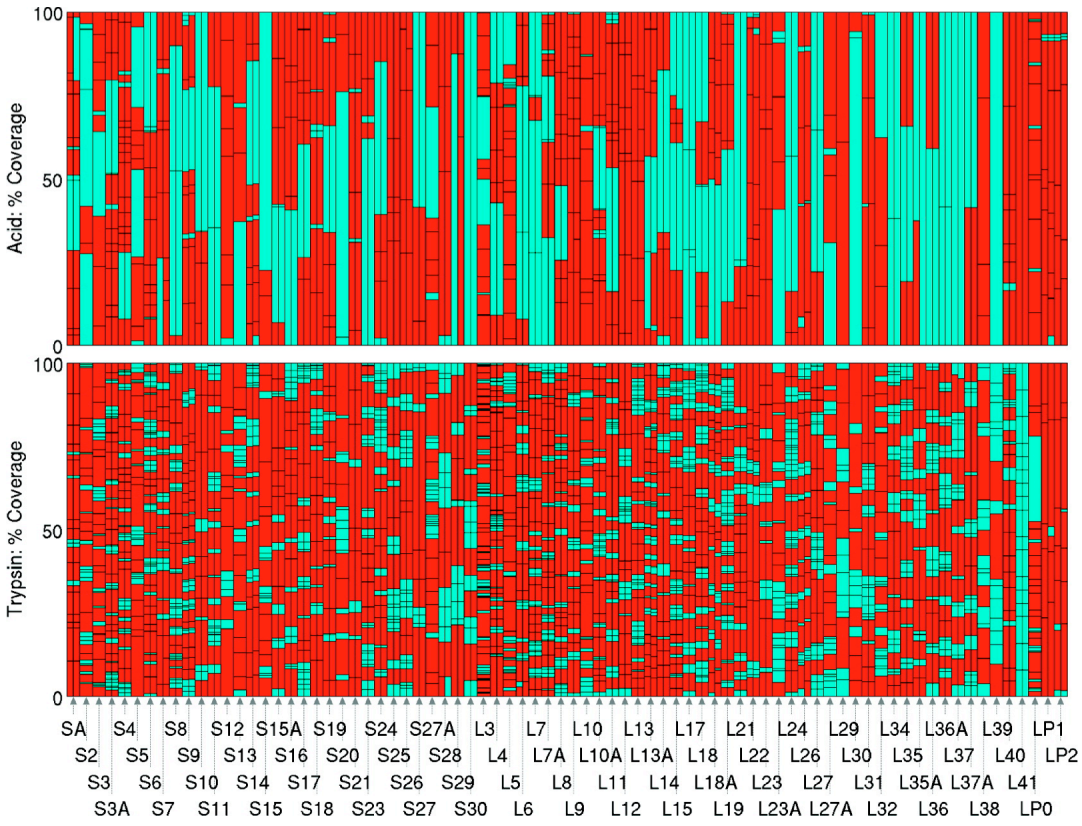

Figure 10. Predicted peptides resulting from Asp specific (top) and tryptic digestion of the *S. cerevisiae* ribosome. Peptides in green are outside the 500-5,000Da mass range. Individual proteins are shown on the X axis and the relative sequence coverage provided by each peptide is shown on the Y axis. Reprinted with permission from reference 35.

The chemical properties of acid hydrolysis produced peptides are distinct from those produced with trypsin. Trypsin's specificity for the C terminal side of Lys and Arg provides a basic site on the C terminus of every product. Protonation of the C-terminal amino acid side chain and amino terminus, typically gives product ion spectra ideally suited for high throughput database searching.[15] Swatkoski and Fenselau demonstrated product ion spectra of peptides without a C terminal basic amino acid could be interpreted and searched against databases.[35] Interestingly, the majority of peptides from the *S. cerevisiae* ribosome had two or greater internal basic residues, which aided in sensitivity for nanospray MS analysis. Swatkoski and co-workers evaluated the hydrolysis reaction by searching the observed data without designating any enzyme specificity. This step allowed evaluation of cleavage specificity based on fragmentation data, a stronger basis for identification compared to peptide mass fingerprinting, applied previously. In agreement with Li et al,[38] 70% of the peptides adhered to the expected cleavage either side of Asp. Mascot, a popular proteomics search algorithm, now offers acid cleavage ("formic acid" in Mascot) as a parameter for enzymatic specificity.

Basile and co-workers have implemented microwave-supported acid hydrolysis in line with chromatographic separation.[44] Proteins were subjected to microwave irradiation for 5 minutes in 12.5% formic acid at 130 ^{o}C via a 5 μL fused silica reaction loop. The method was used to analyze an *E. coli* lysate in which ten proteins were successfully identified by LC-MS/MS. The same laboratory later demonstrated combining microwave-supported acid hydrolysis with electrochemical oxidation.[46]

The previously outlined experiments demonstrate that acid hydrolysis, whether supported by microwave irradiation or not, can reliably and quickly produce peptides that are the result of hydrolysis on either or both sides of Asp residues. Additionally, despite lacking a C terminal basic site to aid in ion trap and quadrupole fragmentation, acid hydrolysis peptides are amenable to high throughput bioinformatic search algorithms. The most compelling results from these experiments have been from standard proteins or with simple mixtures. This dissertation will go over several methods development and optimization type experiments aimed at extending the benefits of microwave-supported acid hydrolysis to more complex mixtures.

Method Development – In Gel Digestion of a Complex Mixture

Introduction

Denaturing polyacrylamide gel electrophoresis has become a mainstay for fractionation of complex protein mixtures based on molecular weight. Arguably the most applied technique in proteomics couples tryptic digestion with electrophoretic separation. Microwave-supported acid hydrolysis has been shown to be an effective method for rapid proteolysis in proteomic workflows in solution, and has even been applied on pure proteins in gel and molecular weight standards. Despite its rapid nature and ease of applicability, microwave-supported acid hydrolysis has yet to be

used on truly complex mixtures, whether in gel or not. Presented here is a demonstration of in gel microwave-supported acid hydrolysis on a whole cell lysate from RPMI-8226 multiple myeloma cells.

Currently, comprehensive proteomics employs almost exclusively workflows that utilize a two-residue specific enzymatic proteolysis step coupled with some method of fractionation at either the protein or peptide level. Proteolysis is usually performed with trypsin, a protease that cleaves C terminal to Arg and Lys, unless followed by Pro. While this method has proven to be quite reliable and effective, there is still a need for alternative options in the proteomics toolkit.
In gel digestion using trypsin for protein identification has become a standard procedure for many labs. The ease of applicability of this method has made it readily available even for scientists that normally work outside the proteomics arena. Despite its simplicity, few methods can rival the protein level separation efficiency of electrophoresis A recent publication from this laboratory demonstrated a higher number of identifications using in gel tryptic digestion when compared to other easily implementable methods that have garnered recent acclaim.[56, 57] As of yet, there are few examples in the literature of in gel digestion being performed using microwave-supported acid hydrolysis.[38, 40, 45] While these works are important to show the feasibility of in gel microwave-supported acid digestion, they do not demonstrate the potential of the reaction for applications analyzing complex mixtures. Additionally, there are only a few examples of this proteolytic method being used successfully for mixtures containing more than a few proteins, in gel or not.[22, 35, 44, 46] This work will

demonstrate the feasibility of using microwave-supported acid digestion on complex mixtures in gel using a whole cell lysate from RPMI-8226 multiple myeloma cells.

Materials and Methods

Cell Growth, Harvesting, and SDS-PAGE Fractionation. RPMI-8226 cells and RPMI-1640 media were purchased from ATCC (Manassas, VA). Cells were grown in suspension in RMPI-1640 media supplemented with 10% fetal bovine serum (FBS) and antibiotics. The cells were maintained in a 5% CO_2 atmosphere. RPMI-8226 cells were incubated at 4°C for 30 minutes in hypotonic buffer (2.5mM imidazole, pH 7.0) prior to lysis using N_2 cavitation. The lysate was spun at top speed in a benchtop centrifuge for 10 minutes to pellet insoluble cellular debris, and the supernatant recovered. Soluble protein was precipitated using the method of Wessel and Flugge[58] and concentration was estimated using the RC/DC Protein Assay (BioRad, Hercules, CA) post resolubilization in gel loading buffer.

Four 170µg aliquots were diluted in Laemmli loading buffer (BioRad, Hercules, CA) and fractionated electrophoretically on a 8-16% gradient Tris-HCl polyacrylamide gel (BioRad, Hercules, CA). The gel was then fixed and stained. Eight bands were excised from the gel with each band composed of equivalent molecular weight regions from each of the four lanes (32 bands total). The bands were then pooled and destained according to the method of Shevchenko and co-workers.[59] The pooled gel bands were macerated to aid in solvent uptake and eventual peptide extraction. Following destaining, the bands were dehydrated in neat ACN and

rehydrated in 12.5% (v/v) acetic acid in water for 30 minutes. Just prior to digestion additional solvent was added to ensure the gel pieces were completely submerged.

Microwave-supported Acid Hydrolysis. For in gel digestion the destained gel pieces were digested for 30 minutes at 140°C while irradiating with 300W of microwave energy. The post digestion supernatant was recovered and the peptides were recovered following the previously referenced protocol.[59] Briefly, peptides were extracted by adding a 1:2 solution of 5% formic acid:ACN and incubating at 37°C on a shaker for 30 minutes. The post digestion supernatant and extraction solution containing the acid proteolysis peptides were combined and lyophilized to near dryness. The solution was diluted to ~100µL prior to LC-MS/MS analysis.

LC-MS/MS. For each of the eight pooled gel samples, 95µL of the digested mixture were injected and fractionated via RP-nLC prior to MS analysis. Peptide solution was injected via a Shimadzu Prominence NanoLC (Columbia, MD) and autosampler. The sample was concentrated online at 5µL/min in 100% solvent A (97.5% water/2.5% ACN/0.1% formic acid) for 20min. Subsequently, the peptides were fractionated on a 0.15mm x 150mm C18 column (Grace Vydac, Deerfield, IL) using a linear gradient that declined from 100% solvent A to 65% solvent A over the course of ninety minutes with a flow rate of 500nL/min. Mass spectrometric analysis was performed using a Thermo LTQ Orbitrap XL (San Jose, CA). Precursors were acquired at high resolution (30,000) and the top nine most abundant multiply charged MS2 ion trap scans were acquired simultaneously. Peptides were fragmented with the

following CID settings: normalized collision energy was set to 35, activation time was set to 30ms, activation Q was set to .250, and isolation width was set to 3Da.

Bioinformatics. All spectra were searched and combined using the PepArML combiner provided by Dr. Nathan Edwards (https://edwardslab.bmcb.georgetown.edu/PepArML/).[60] Acid digested peptides were searched against the IPI Human database. Dehydration at Asp, pyro-glutamic acid formation from N terminal Glu and Gln, and oxidation of Met were added as variable modifications. These parameters were searched against six different search engines and the results were combined and an estimated peptide specific false discovery rate was generated. Peptides with FDR less than 10% were included in the list of identified peptides and therefore proteins. The raw data was also searched against the IPI Human database without designating enzymatic specificity using the Mascot search engine (MatrixScience, Manchester, UK). Mascot searches used a 10 part per million (ppm) precursor mass tolerance and 0.6Da fragment ion mass tolerance. Charge states +2, +3, and +4 were interrogated using the same variable modifications listed above.

Results and Discussion

The proteomics community has accommodated the fact that polyacrylamide gel electrophoresis permeates nearly all aspects of *in vitro* biological research by making bottom up methods specifically tailored to gel fractionated samples.

Similarly to in solution digestion methods, microwave-supported acid hydrolysis is catching up to the tryptic standard. Although acid hydrolysis has been used on gel fractionated samples previously,[38, 40, 45] there have yet to be any successful analyses of complex mixtures. In this report, a gel fractionated whole cell lysate from multiple myeloma cells resulted in 1241 distinct peptides representing 642 proteins (**Appendix Table 1** and **Appendix Table 2**). The average sequence coverage per protein was 8.25%. Despite the higher protein concentration required for in gel acid digestion, the efficiency of electrophoretic fractionation is maintained (**Figure 11**). The resulting data validates this conclusion by observing greater than 75% of all proteins identified were confined to a single band.

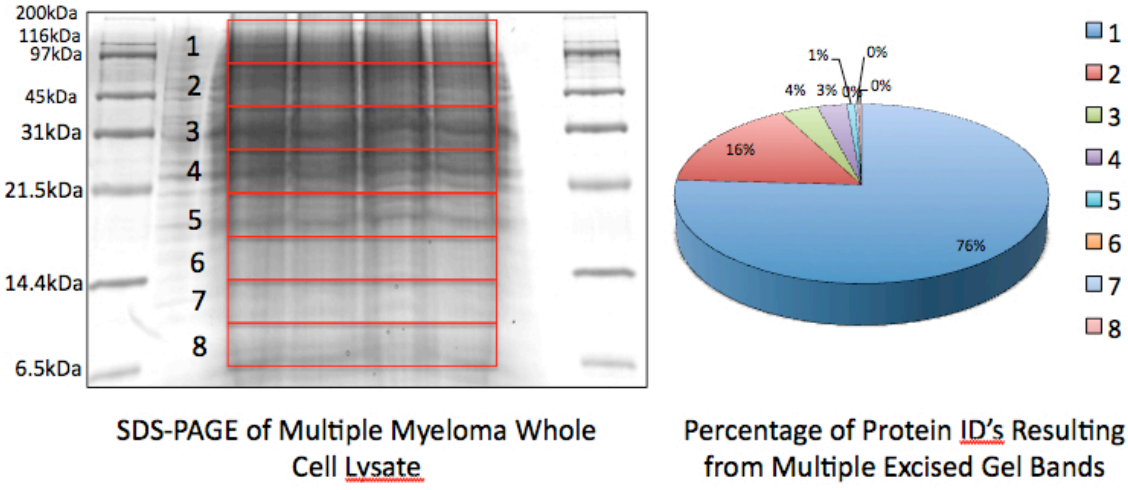

Figure 11. Coomassie stained gel demonstrates lowered stain resolution due to high protein concentration. Excised bands are outlined in red. Despite lower resolution, a high proportion of protein identifications were confined to a single excised gel band (right).

Microwave-supported Acid Digestion for Bottom Up Proteomics. In both of the experiments outlined above it can be concluded that microwave-supported acid hydrolysis in gel is a viable method for proteomic analysis of complex mixtures. When compared to the standard tryptic methods used by most laboratories, acid hydrolysis provides a faster alternative, albeit at lower sensitivity. Since the result of kinetic selectivity is a distribution of peptide products regardless of digestion time, there is an inherent decrease in the absolute quantity of any single peptide *ion* of interest. Practically, this means that a given portion of a protein's sequence will be observed in at least one specific ion and at most four (DXXXXXXD, DXXXXXX, XXXXXXD, and XXXXXX) due to the dual terminal specificity.[30] It is also likely that the decrease in sensitivity can be attributed to the peptide mass distribution resulting from microwave-supported acid digestion.[22] Contrary to the ease of implementation of in gel bottom up proteomics, top down methods have not enjoyed the same widespread success despite significant effort.[61, 62] This is due to a decrease in extraction efficiency from the polyacrylamide gel matrix as the length of the polypeptide increases.

In this work we demonstrate that microwave-supported acid digestion is a viable workflow for the analysis of complex mixtures in gel. Specifically, we show that this method is applicable to a complex mixture such as a whole cell lysate and will provide results in a small fraction of the time used for the current tryptic standards. In gel microwave-supported acid digestion has been shown to provide a more comprehensive analysis than previously shown, albeit with lower sensitivity than one would achieve using enzymatic methods. Despite the lowered sensitivity,

the benefits of electrophoretic fractionation are still observed. Proteins can be effectively denatured and fractionated, cleaned of interfering non-proteinaceous materials, and visualized prior to rapid chemical digestion well suited for LC-MS/MS analysis.

Method Development – Mass Biased Partitioning (adapted from reference 63)

Introduction

All of the previously mentioned methods take advantage of acid hydrolysis produced peptides that are similar in size to those produced using trypsin. Although this strategy has a clear advantage over enzymatic methods due to the increase in speed of analysis, it ignores additional benefits provided by single residue specific proteolysis methods. A strategy is presented for enhancing the middle-out analysis of higher mass peptides recovered from a complex protein mixture. Following a 30 min digestion of multiple myeloma cell lysate by an acid cleavage reaction that is selective for aspartic acid, a 3,000Da membrane filter is used to bifurcate the peptide product mixture, and the heavier fraction is subjected to collisional activation with precursor selection that excludes charge states below +4. Filtration and charge state selection are shown to provide significant increases in the number of peptides identified in the mass range above 3,000Da and in information about protein sequences.

Middle-out proteomic strategies exploit the advantages of analyzing heavier peptides (3,000-20,000Da) in proteomic analyses.[64, 65] These include improved chromatographic fractionation, higher sequence coverage, and characterization of cohabiting and potentially interactive modifications.[32, 66] Experimentally, analysis of peptides in the mass range 3,000 to 20,000Da simplifies the complex mixtures offered by bottom up strategies, while avoiding the diminished performance of top down experiments.[21, 67] Most often, single-residue specific enzymatic reactions are used to produce mid-range peptides. Previously we have demonstrated the use of a chemical method that cleaves proteins selectively at aspartic acid in less than 30 min.[22] Like other proteolytic agents, microwave-supported acid cleavage of complex protein mixtures produces complex peptide mixtures that contain lower mass peptides as well as heavier peptides. During automated analysis by LC-MS/MS, an abundance of low mass peptides in the mixture can suppress or obscure precursor selection and activation of the mid-range peptides. Here we report the use of molecular weight cutoff filters to separate a complex peptide mixture into high mass and low mass fractions. In addition, we limit precursor selection in high resolution tandem mass spectrometry experiments to charge states of 4+ or greater in order to optimize high throughput analysis of mid-range peptides. The combination of these two experimental modifications is evaluated here as mass biased partitioning.

Materials and Methods

RPMI-8226 Cell Culture, Isolation, and Lysis. RPMI-8226 multiple myeloma cells were grown and harvested as published.[68] The cells were lysed in 2.5mM imidazole, pH 7, using N_2 cavitation at 1250psi (Parr Instrument Co, Moline IL). Following centrifugation at 10,000rpm for 10 minutes, the supernatant was collected and stored at -80°C until digestion.

Microwave-supported Acid Hydrolysis and Mass Biased Partitioning. Protein was precipitated from solution using the method of Wessel and Flugge.[58] The pellet was redissolved in buffer and the concentration was determined by RC/DC assay (BioRad). A 50µg aliquot was diluted to 0.1µg/µL with acetic acid and water to arrive at a final solvent concentration of 12.5% (v/v) acetic acid. The acidified sample was digested in a CEM Discover Microwave (Matthews, NC) at a maximum temperature of 140°C for 30 minutes while irradiating with 300W. Following digestion the samples were allowed to cool to room temperature. Amicon 3kDa and 10kDa molecular weight cutoff (MWCO) filters (Millipore, Billerica, MA) were equilibrated with 12.5% acetic acid prior to use. The acid digestion products were then fractionated through the filter according to the manufacturer's instructions. The high mass fraction retained above the filter was diluted with 200µL of Milli-Q water, aspirated 30x to maximize recovery, and lyophilized to reduce the volume to

approximately 100µL. The low mass filtrate was also collected and reduced to 100 µL.

LC-MS/MS. For the low mass peptide sample, 1/20th of the total volume was injected for LC-MS/MS analysis. Reversed phase chromatography was carried out using a Shimadzu Prominence LC and Autosampler (Columbia, MD). Following injection, the peptides were desalted and concentrated online for 20 minutes at a flow rate of 10µL/min with 100% Solvent A (97.5/2.5/0.1 H_2O/ACN/formic acid). The peptides were then fractionated using a 0.150mm x 150mm Grace Vydac Everest C18 column packed with 5µm particles with 300Å pores (Deerfield, IL). The flow rate was set to 300 nL/min and the concentration of Solvent B (97.5/2.5/0.1 ACN/H_2O /formic acid) was increased in a linear fashion from zero to 35% over the course of 180 minutes. Low mass peptides were introduced into an LTQ-Orbitrap XL and MS1 scans were acquired at 30,000 resolving power. Precursor peaks were limited to the 8 most abundant multiply charged peptides. Product ion spectra were recorded in the LTQ at unit resolution.

For the high mass portion, 100µL of sample was injected for each analysis. Chromatographic conditions were identical to those outlined above. Survey scans were acquired at 30,000 resolving power and the three most abundant multiply charged precursors were isolated and fragmented in the ion trap, and subsequently detected at 7,500 resolving power in the Orbitrap.

Charge State Inclusion and Rejection. For mass biased partitioned samples chromatographic conditions were identical to those outlined above. For the low mass fraction, MS parameters were identical aside from charge state selection. Data dependent analysis was set to only isolate and fragment precursors with charge states of +2 or +3. For higher mass peptide analysis, only precursors with charge states greater than +4 were selected for isolation and fragmentation. Automated gain control targets were set to 5×10^5 and 5×10^6 for the survey and product ion scans, respectively. Isolation width was set to 10Da.

Bioinformatics. Mid-mass peptides were identified using ProSightPC2.0, which incorporates the Xtract algorithm for precursor and product ion neutral mass calculations (ThermoFisher). Precursor mass tolerance using ProSightPC was set to 250Da. The large mass window allows for the identification of the N- or C-terminal Asp-cleavage products observed in microwave-supported acid hydrolysis with ProSightPC's AspN *in silico* digest option. In addition, the large window allows for the identification of dehydrated and oxidized species. Fragment mass tolerance was set to 15ppm. Mid-mass spectra were searched against a database of reviewed human proteins from the UniProtKB digested *in silico* using ProSight's Database Manager utility, with a maximum of 9 missed cleavages and maximum peptide mass of 20kDa. Spectra were also searched against a shuffled version of the same database. Identifications were assigned automatically with E-values less than or equal to 10^{-8}, the threshold at which no peptides were matched in the shuffled database search. Other peptides from the high mass fraction that were matched with E values $< 10^{-4}$

were validated manually. RAWmeat (Vast Scientific, Cambridge, MA) was used to obtain scan and charge statistics from data files.

The low mass peptide fraction was searched using the PepArML search engine.[10] PepArML combined results from 6 search engines against the IPI Human database incorporating cleavage specificity at either side of Asp residues. Dehydration at Asp, pyro-glutamic acid formation from N terminal Glu and Gln, and oxidation of Met were added as variable modifications. *In silico* digestions were performed using *ad hoc* software developed in house.

Results and Discussion

In order to prepare a sample enriched in mid-range peptides, fractionation of the peptide mixture was evaluated using both 3kDa and 10kDa filters. The heavier fraction of peptides recovered from each filter was analyzed by LC-MS/MS, excluding selection of +2 and +3 precursor ions and implementing 7500 resolution for product ion analysis as described in the experimental section. Increasing the resolution for product ion spectra improves product ion charge-state determination and boosts the specificity of product ion matching, however, at the expense of increased duty cycle. A larger proportion of the 10kDa filter's peptides were heavy ($\geq$ 3kDa), compared to the 3kDa filter's peptides, although the absolute number of heavy peptides was greater for the 3kDa filter. Consequently the 3kDa filter was selected for the rest of the study.

A combined total of 349 proteins were identified from 624 peptides when both the heavy and light peptide fractions were analyzed using optimal conditions as described in Materials and Methods. Fifty-six peptides were identified in the high mass fraction (3kDa filter) from 38 proteins, whereas 568 peptides were observed in the low mass portion, from 333 proteins. **Figure 12** shows the molecular masses of peptides identified from injections of the heavier retentate and the lighter filtrate. In this figure, the proportion of each fraction's identified peptides in each of 500Da bins is plotted side-by-side. The masses of the peptides identified from each injection display distinct distributions, reflecting successful fractionation of light and heavy masses.

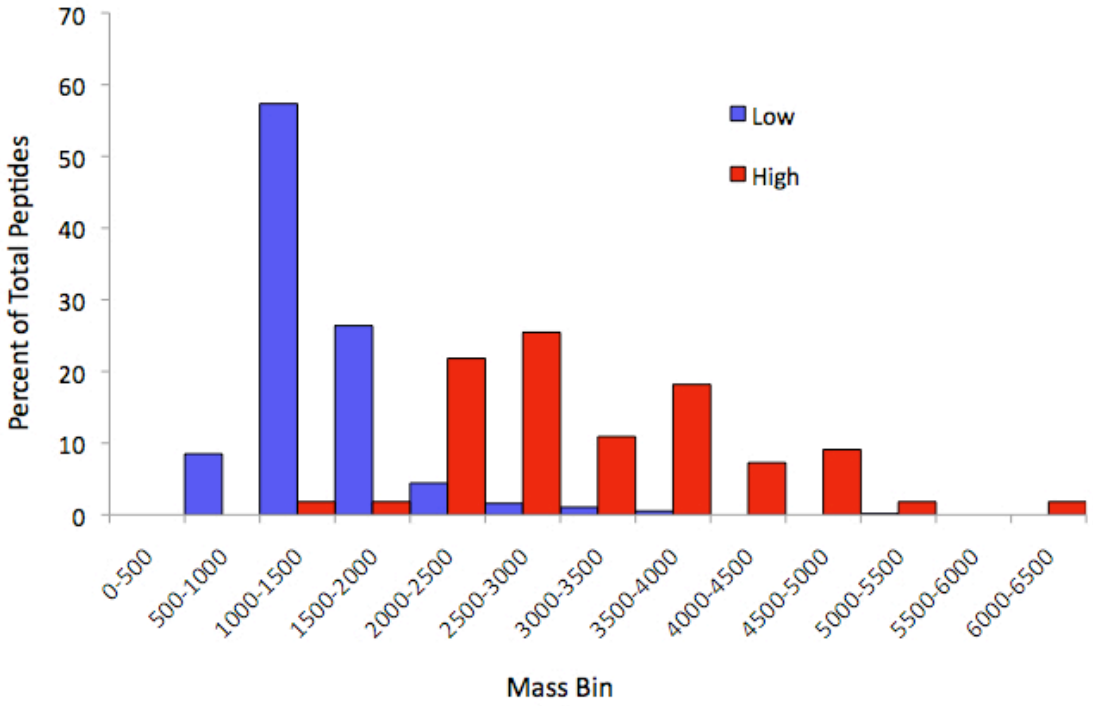

Figure 12. Peptide mass distributions of high (red) and low (blue) mass peptide fractions resulting from mass biased partitioning after microwave-supported acid hydrolysis.

Of the total, only 9% of the peptides identified have masses > 3,000Da. To some extent, this reflects the smaller number of heavier peptide products that can be formed. A theoretical analysis of AspC cleavage of the proteins identified indicates that 27% of the peptide products would be expected to weigh more than 3,000Da (with 0 missed cleavages and a maximum peptide mass of 20kDa). By contrast, 91% of the peptides identified had masses below 3,000Da, while 73% were predicted. Experimental under-sampling clearly occurs in our experimental analysis of the mid-range peptides.

The distribution of experimental peptide identifications was also compared to those from two control experiments. Whole cell lysate was subjected to acid cleavage

and analyzed twice, first using the LC-MS/MS parameters optimized for the lighter filtrate and also those optimized for the heavier peptide mixture. One hundred and ninety-eight peptides were identified when product ion characterization was carried out with unit mass resolution in the ion trap (our low mass mode). Fifteen of these peptides had masses exceeding 3,000Da, however the use of low resolution compromised their identification. Twenty-eight peptides were identified using the higher resolution Orbitrap analysis, of which thirteen had masses exceeding 3,000Da. The more complex mass biased sample preparation appears to provide more peptide identifications overall, and, more to the point, a nearly four-fold increase in identifications of mid-range peptides with high reliability.

We have reported previously, in a study of human ribosomal proteins, that analysis of heavier peptides obtained using Asp-selective acid cleavage provides higher sequence coverage than the shorter peptide products of tryptic proteolysis.[22] The advantageous coverage offered by longer peptides is confirmed in the present study of whole cell lysate. Thirty-three per cent of singlet peptide identifications in the high mass fraction provided greater than 10% sequence coverage of their parent proteins, while only 10% of the individual peptides identified in the low mass fraction provided greater than 10% coverage. Overall, the average protein coverage for the high mass peptide fraction is 9.4%, and for the low mass peptides is 4.6%. This trend for Asp-selective acid cleavage is in agreement with the general view of single-residue middle-out analysis of protein.[21, 22, 67]

The proportion of peptides produced by a proteolytic reaction that fall into the middle-out range of 3 to 20 kDa depends on the specificity of the reagent, the

frequency of the targeted residue(s) and experimental control of missed cleavages. Of particular interest is the distinction between the products of trypsin, which cleaves at two amino acid residues--lysine and arginine, and methods that cleave at a single amino acid. Karger and co-workers have suggested that the false discovery rate decreases with larger peptides.[32] Goodlett and co-workers have examined the notion that precursor ion densities are not distributed uniformly across the m/z landscape in their gas phase fractionation experiments.[69, 70] To explore the density of overlapping peptide masses in our experiment, human proteins in the UniProtKB were digested *in silico* based on cleavage with trypsin, and also on cleavage at the C-terminal side of aspartic acid, and duplicate peptide sequences were removed. For each *in silico* peptide, the number of additional peptides within 10ppm of its mass was counted, and the peptides tabulated in various non-specificity bins. The results are summarized in **Figure 13**, where it can be seen that 70% of all Asp-C peptides in the mass range 500 to 1000 share their 10ppm window with 5 or fewer additional peptides, while 19% of the tryptic peptides 500 to 1,000Da have this advantage. The 70% precursor specificity is achieved with tryptic products only at the molecular mass range of 3,500-4,000Da.

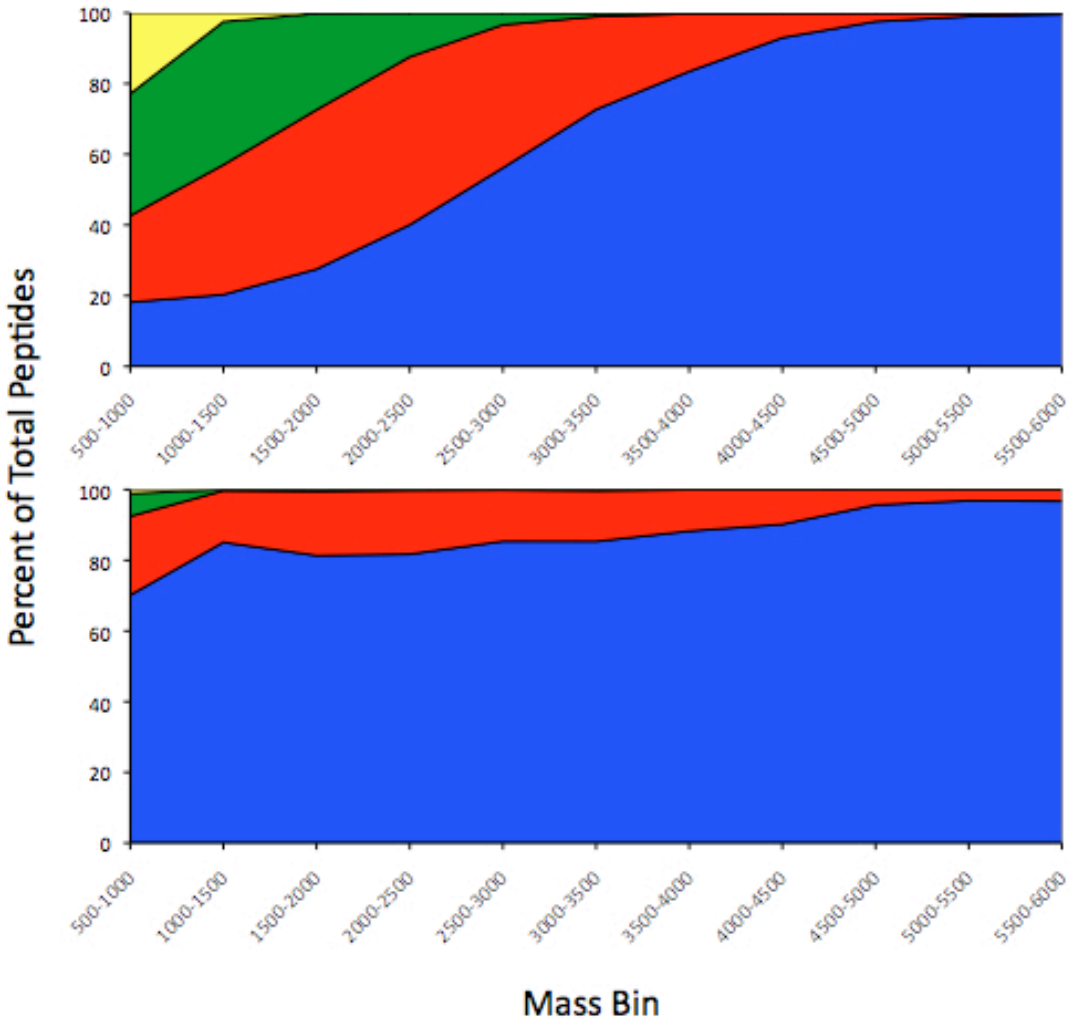

Figure 13. The proportion of peptides with masses that overlap within 10ppm, produced by *in silico* digestion of a subset of human proteins in UniProtKB using trypsin (top) and Asp C (bottom). Portions in blue indicate regions with 0-5 additional peptides, regions in red have 6-10, in green have 11-20, and in yellow have greater than 20.

Consideration of **Figure 13** indicates that the molecular masses of peptides in the middle-out mass range are more unique than those of their bottom up counterparts.

Theoretically, Asp-specific proteolysis (and likely any single residue proteolysis) produces a higher proportion of peptides with nearly unique masses than trypsin. This allows more reliable identifications with lower false discovery rates due to the lower number of searchable peptide candidates within a given precursor mass tolerance.

In summary, we have demonstrated that microwave-supported acid hydrolysis is a viable method for rapid proteolysis of complex protein mixtures, and that mass biased partitioning, comprising both membrane cut-off filtering and precursor charge state selection, provides enhanced access to, and identification of, mid-sized peptides. Care must be taken with the analysis of the larger peptides. In our experience there is often just one identified per protein. The use of high resolution in the determination of both precursor and fragment ion masses, and the reduced probability for overlapping masses provides some of the needed specificity.

Chapter 3: Middle-Out Analysis of the Human Ribosome (adapted from reference 22)

Introduction

Despite tremendous advances based on Fourier transform ion cyclotron resonance mass analyzers, there are still technical barriers to the implementation of top-down tandem mass spectrometry as the method of choice for automated analysis of complex protein mixtures. In addition to the challenge to continue to extend the mass range to accommodate protein requirements,[71, 72] the resolution and mass accuracy required for deconvolution of very high charge state species are achieved on many Fourier transform ion cyclotron resonance analyzers at ion accumulation times incompatible with interfaced chromatographic flow rates. In the meantime, real advantages have been recognized for the analysis of long polypeptides, so-called middle-out analysis, in the mass range 3,000 to 20,000Da. Such proteolytic products are reported to provide valuable information about the cooperative occurrence of PTMs.[71, 73-75] Midrange peptides are reported to be fractionated with improved resolution by HPLC.[32] Longer peptides carry higher numbers of charges when electrosprayed, which enhance both CID and ETD.[35, 48, 76]

These polypeptides are usually produced from protein mixtures by enzymatic or chemical methods that cleave with selectivity for a single residue.[32, 35, 75, 77, 78] One of these, microwave-supported acid hydrolysis, has been consistently shown to cleave polypeptides on one or both sides of Asp residues. This method has the advantage that it does not discriminate between derivatized and underivatized Lys and Arg

residues; however, its major advantage is the speed with which it provides proteolysis in a variety of solvents.

Without requiring customized modifications, the Orbitrap mass analyzer has been shown to provide resolution sufficient to decharge midsized polypeptide precursors in a time frame compatible with capillary HPLC peak widths. Additionally, the Orbitrap is able to analyze the multiply charged fragments resulting from collisional or electron transfer activation of highly charged precursors with high resolution (15,000 and up) at a duty cycle compatible with chromatography. The Orbitrap used in this work was coupled to a linear ion trap, whose trapping capability maintains sensitivity while providing robust multicollision activation. The longer duty cycle that results from measuring both precursor and product ions with high resolution means that fewer peptides can be analyzed in a given elution time and recommends the use of mixtures of fewer peptides such as those provided by the middle-out strategy.

The computational requirements include the capability to extract and deconvolute charge states from isotope patterns of precursor and product ions and to search the resulting fragmentation patterns against predictions from databases of protein sequences. In the present work, ProSightPC 2.0 was used, which also provides the option to specify acid cleavage at Asp.

The ribosome is an important multiprotein complex, currently under intense scientific scrutiny.[79] Pulse chase experiments have shown that the half-life of the eukaryotic ribosome exceeds that of the cell,[80] and modest protein modifications have been hypothesized to occur in response to changes in cellular health and drug

treatment. In the human ribosomal database, Asp residues account for 3.81% of total residues. By comparison, Arg, Lys, and Arg-plus-Lys account for 9.30, 12.36, and 21.65% of total residues, respectively. **Figure 14**a-c shows the distribution of peptides of various residue lengths predicted for molecular weights greater than 500Da with at most 1 missed cleavage, with respect to (a) tryptic digestion (3,397 total), (b) Lys-C digestion (2,406 total), and (c) acid digestion (991 total). We use an *in silico* Asp-C cleavage for this figure, instead of the total acid digest, to eliminate double counting of sequences that differ only in the presence of an Asp residue. Observed acid cleaved peptides have also been filtered throughout to eliminate peptides differing only in terminal Asp residues. In the small ribosomal proteome, comprising mostly basic proteins, Asp-selective cleavage is expected to provide a peptide set of limited size with molecular masses (lengths) across the middle-mass range. The present evaluation of a novel middle-out workflow has been carried out on the ribosomal proteome with the expectation that it will be applied in future studies of differential modification in that system.

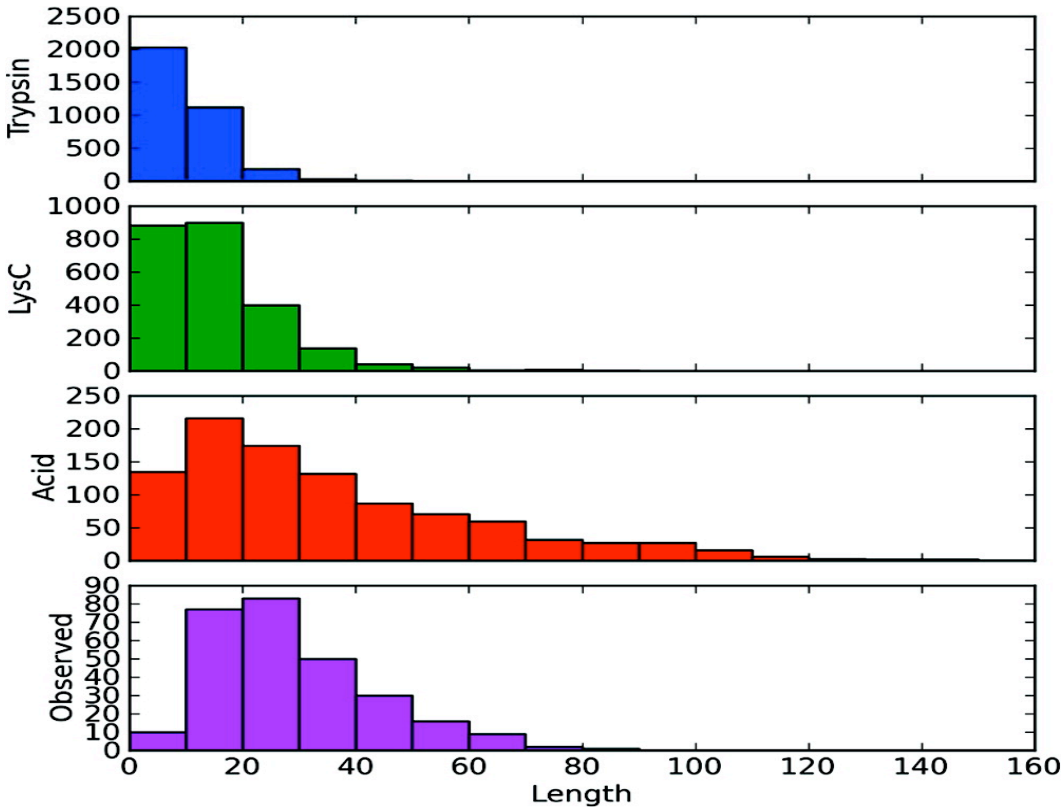

Figure 14. Distribution of peptide products by length, predicted from the 84 proteins in the human ribosome cleaved by (a) trypsin, (b) Lys-C, and (c) Asp-C acid cleavage. (d) Distribution of Asp-C peptides identified experimentally in an acid cleavage digestion. Reprinted with permission from reference 22.

Materials and Methods

Cell Culture and Ribosome Isolation. MCF7 breast cancer cells were grown to confluence in Improved Minimal Essential Media (IMEM) with L-glutamine supplemented with 1% penicillin-streptomycin antibiotic solution and 10% heat

inactivated FBS. Cells were maintained at a temperature of 37 °C in a 5% carbon dioxide atmosphere until confluence.

Ribosomes were isolated from confluent cells. All procedures were carried out on ice unless otherwise indicated. Homogenization of the cell pellet with a Kinematica mechanical homogenizer (Littau, Lucerne; Switzerland) in two volumes homogenization buffer (50mM Tris-HCl, pH 7.5; 5mM $MgCl_2$, 25mM KCl, 200mM sucrose) was followed by centrifugation at 10,000g for 10 minutes at 4°C in a benchtop centrifuge. The supernatant was collected and the remaining pellet rehomogenized on ice and centrifuged. The supernatant was layered 1:1 over a sucrose cushion buffer (50mM Tris-HCl, pH 7.5; 5mM $MgCl_2$, 25mM KCl, 2 M sucrose) and the ribosomal pellet isolated by centrifugation at 260,000g at 4°C for 2 hours in a swinging bucket rotor.

Protein Extraction. Ribosomal proteins were extracted from a 500µL suspension following a variation of the method described by Hardy et al.[81] In brief, one volume of the ribosomal suspension was mixed with 0.25 volumes of 1M $Mg(OAc)_2$ followed by the addition of 1 volume glacial acetic acid. Each solution was incubated for 1 hour and the precipitated rRNA was pelleted by centrifugation at 10,000rpm at 4°C for 10 minutes in a benchtop centrifuge. The supernatant was then collected. Molecular weight cutoff filters (Microcon Ultracel YM-3, Millipore, Billerica, MA) were used to concentrate the samples and reduce the acid content. After desalting and concentrating the samples, the protein concentrations were determined using the RC/DC protein assay (Bio-Rad). Dilutions of 0.1 mg/mL

concentrations were prepared for each sample. To each of these samples acetic acid and dithiothreitol were added to achieve 12.5% acetic acid and 5mM dithiothreitol.

Microwave-supported Acid Hydrolysis. The method optimized and reported by Swatkoski et al[35] for digestion of ribosomal proteins was used. Fifty microliters of each solution was irradiated at 300W for 20 minutes in SPS mode in a Discover Benchmate microwave (CEM) at 140°C. Samples were collected and injected directly into the nanoLC through an autosampler.

LC-MS/MS. Seven injections were made into a Shimadzu Prominence NanoLC (Shimadzu, Columbia, MD) interfaced to the LTQ Orbitrap mass spectrometer (ThermoFisher, San Jose CA) via an Advance CaptiveSpray Plug-and-Play source (Michrom Bioresources, Auburn CA). The samples were loaded onto a 0.3mm x 5mm Peptrap 300 C18 pre-column (Dionex, Sunnyvale, CA) in 5% Solvent A (97.5% water/2.5% ACN/0.1% formic acid) and desalted for 15 minutes. Peptides were eluted into a 0.1mm x 150mm C18 analytical column (Grace Vydac) and separated with a linear gradient of 5−15% solvent B (97.5% ACN/2.5% water/0.1% formic acid) in 5 min, then to 69% B in 115 min. Flow rate was 500nL/min. Automated gain control settings were optimized using replicate injections of the human ribosome digest. Survey scans were acquired in the Orbitrap with resolving power of 30,000 at m/z 400 and an automated gain control (AGC) target level of 5 x 10^5. The three most abundant ions were selected for fragmentation using CID in the linear ion trap. Precursor ions were isolated using a 3Da window, and fragmented

with He gas for 30ms with a normalized collision energy of 35. The product ion AGC target level was set to 5 x 10^4 and fragment ion scans were acquired in the Orbitrap with resolving power of 15,000 at m/z 400. Dynamic exclusion parameters were set to exclude ions previously selected for fragmentation for 3 minutes. All data were acquired in reduced profile mode to accommodate further downstream processing.

Bioinformatics. Following spectral acquisition .RAW files were processed using ProSightPC 2.0 provided by Professor Neil Kelleher, University of Illinois.[65] Each .RAW file was processed in High Throughput mode. Spectra were decharged with cRAWler using the THRASH algorithm. A FASTA format protein sequence database of 79 human ribosomal proteins was extracted from the Ribosomal Protein Gene Database[82] and configured for acid-cleavage analysis with ProSightPC 2.0. Spectra were searched in Absolute Mass mode using a 2.5Da precursor window based on the peptide monoisotopic mass. An additional search, using a loose precursor window of 250 Da, was carried out to search for evidence of PTMs and peptide isoforms. Significant identifications, when sufficient b and/or y ions were matched despite discrepancies between the predicted precursor mass and the observed mass were manually checked using ProSightPC's Sequence Gazer tool. The sequence positions of the b and/or y ions matched help to localize the mass-shift from putative PTMs and single amino acid substitutions.[19, 27] Mass tolerance for fragment ions was set at 15ppm. False discovery rates (FDR) were calculated using a randomly shuffled version of the ribosomal protein sequence database previously described. Mascot (Matrix Science) searches were also used to analyze the data. Searches were carried

out specifying “no enzyme.” Up to 9 missed cleavages were allowed, with precursor tolerance of 10ppm and product ion tolerance of 0.05 Da. Variable modifications were selected to include N-terminal acetylation, N-terminal pyro-glutamate formation from Gln and Glu, Met oxidation and Ser/Thr/Tyr phosphorylation.

Results and Discussion

Since resolution attained with Fourier transform based mass analyzers is linearly dependent on transient acquisition time, it is important to optimize the duty cycle to acquire the maximum number of high resolution product ions scans per unit time, particularly with complex mixtures containing coeluting peptides throughout the chromatographic separation. Automated gain control (AGC) is a user tunable parameter designed to prevent space charge effects, a phenomenon in which high concentrations of ions in the confined space of the ion trap coulombically repel one another and decrease resolution during analysis. Optimization of the AGC settings was based on several important assumptions associated with the chemical nature of ribosomal proteins and peptides. Importantly, all proteins present in the mixture are assumed to have stoichiometric abundance. This allowed optimization to be carried out without any concern for acquiring spectra of low abundance peptides. Additionally, the ribosomal peptides are highly basic and thought to sequester protons[15]. The high gas phase basicity of Arg and Lys residues (which exhibit increased abundance in ribosomal proteins when compared to the rest of the human proteome) was presumed to increase the half-life of ribosomal peptides in the electric

fields within the mass spectrometer. These assumptions allowed a decrease in the maximum number of ions used to fill the trap in each product ion scan (based on the increased stability in the gas phase), and an additional decrease in the total ion trap fill time in the event there is an insufficient number of ions to fill the trap (based on the assumed stoichiometric abundance of all proteins and peptides). **Figure 15** demonstrates the increase in the average number of scans per unit time after applying the optimized settings.

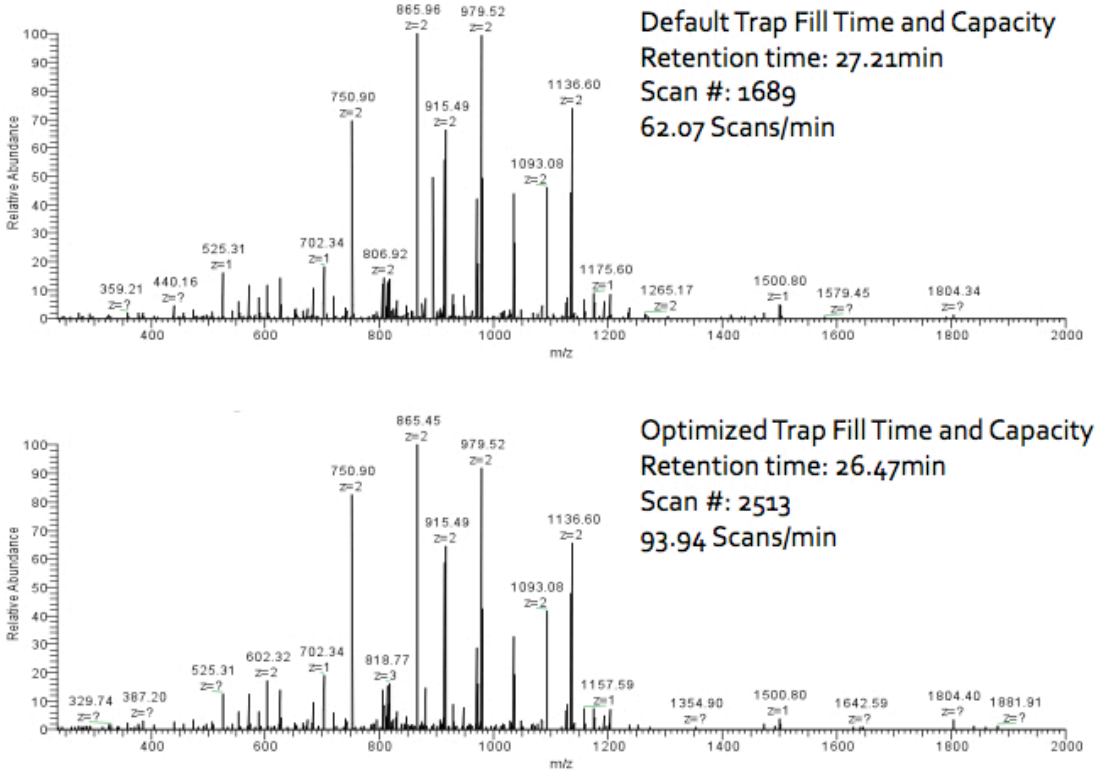

Figure 15. Product ion spectra of the same precursor ion acquired using default (top) and optimized (bottom) AGC parameters of the same peptide.

The changes implemented in the AGC parameters allowed an approximately 1.3 fold increase in the total number of acquired scans, which in turn, resulted in an increase in the total number of identified peptides using an identical chromatographic separation.

Peptide identifications were accepted at E-values of 1×10^{-3} or less, corresponding to a FDR of <1%. Three hundred sixty-six distinct peptide sequences were identified using ProSightPC 2.0, including a number of peptide sequences differing only in the addition or removal of N- or C-terminal Asp. Two hundred seventy-six unique peptides remain after accounting for this redundancy. Forty-four percent of the peptides had masses greater than 3kDa. Twenty-eight percent of the predicted peptides were identified experimentally. **Figure 14**d indicate that microwave-supported acid cleavage produces peptides from ribosomal proteins that span a range of lengths, including significant representation in the range 3000 to 10 000 Da. It should be noted that the full range of charge states could not be explored at the resolution used in this study. We are working to extend the range of middle-out analysis on the LTQ-Orbitrap at higher resolution.

Microwave-supported acid hydrolysis exhibits kinetic selectivity (not binding specificity) for cleavage at aspartic acid.[43] In a previous study of yeast ribosomal proteins we reported that acid cleavage occurred with fidelity at Asp at 83% of the termini in peptides identified. Mascot was used with "no enzyme" specified to make this analysis. The present data set was also searched with Mascot, specifying "no enzyme." In this search, 188 peptides were identified, including 168 already

identified in the ProSight search. Among the 376 peptide termini analyzed, 34 comprised protein termini and 19 (5%) carried termini produced by cleavage at residues other than Asp. The majority of these occurred next to Pro,(4) Asn,(5) and Glu(6).

The 366 peptides identified represented 70 of the 79 human ribosomal proteins present in the ribosomal database with at least one significant peptide identification, and 65 of the 79 ribosomal proteins were identified with at least two distinct significant peptide identifications. The average sequence coverage observed for these 70 proteins was 46.2%. **Figure 16** illustrates the coverage accumulated from 7 injections, which ranges between 98.6% and 6.4% for individual proteins.

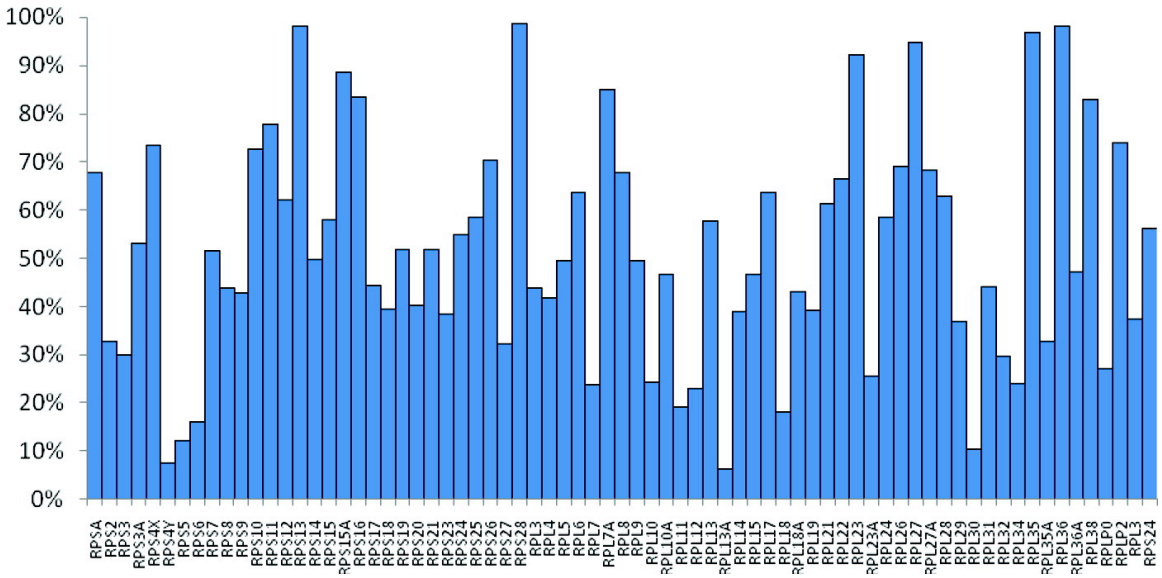

Figure 16. Sequence coverage of seventy proteins by the peptide products of acid digestion combined from seven LC−MS/MS analyses.

The present analysis allows us to test the hypothesis that larger acid digestion products provide an increase in sequence coverage. Ribosomal protein L21 is represented by just two unique peptides. However, as **Figure 16** indicates, those peptides account for 61% of the total sequence. **Figure 17** provides a comparison of the sequence coverage for 70 proteins provided by the peptides with molecular masses above 3,000Da identified using ProSight, and coverage provided by peptides identified with molecular masses below 3,000Da. The set of peptides with masses below 3,000Da comprises 205 peptides, whereas the set with masses above that threshold contains 161. The more limited set of longer peptides provides higher coverage, 36% on average, while the average coverage provided by the more abundant shorter peptides is 21%.

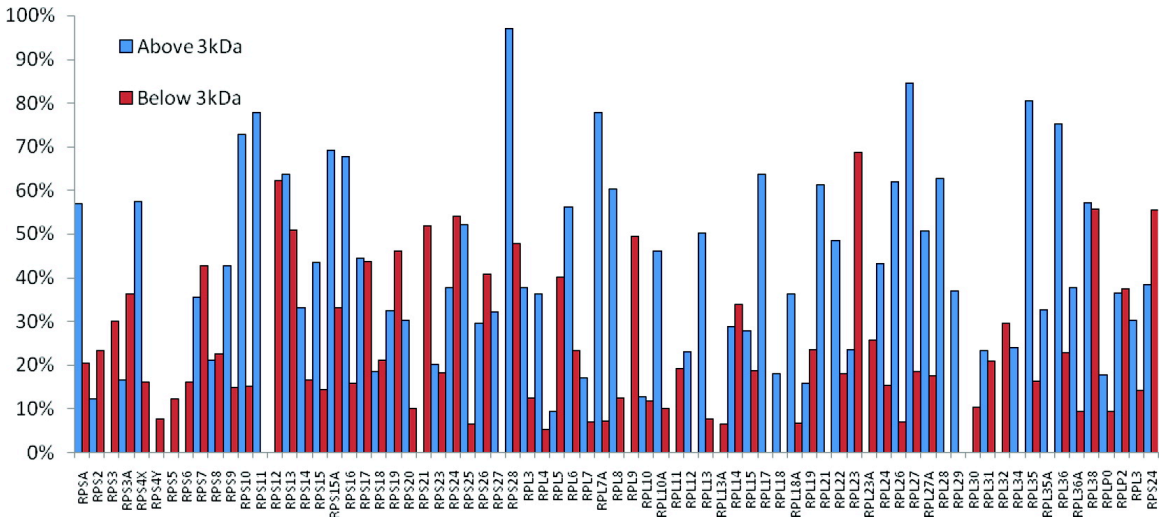

Figure 17. Sequence coverage of seventy proteins by subsets of the peptides in Appendix Table 5.

Another positive byproduct of acid digestion is the presence of multiple basic residues in the interior of many of the polypeptide products. Enzymatic methods that

cleave at basic sites create peptides with sequences that localize positive charges at the termini and must rely on proton transfer for backbone fragmentation.[15] Peptides produced from acid digestion, on the other hand, usually contain multiple internal basic sites and readily produce multiply charged precursor and fragment ions. **Figure 18** illustrates the formation of highly charged fragment ions by collisionally induced dissociation of a polypeptide containing 14 Arg and Lys residues. Although no long series of sequence ions is produced, the fragment ions observed, used in conjunction with accurate mass measurements of precursor and product ion masses, are sufficient to allow the search program to identify the polypeptide with high reliability.

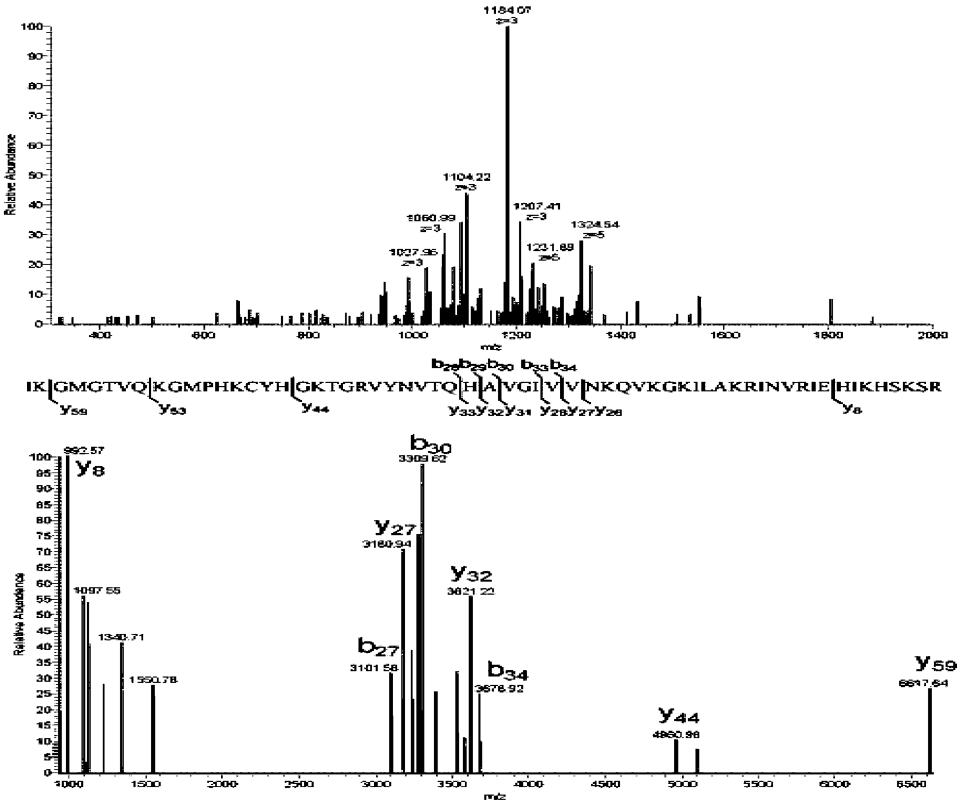

Figure 18. Product ion spectra of a peptide with a calculated monoisotopic mass of 6854.84Da: (top) high resolution product ion scan; (middle) sequence and

fragmentation assigned by ProSightPC 2.0; (bottom) decharged product ion scan. Reprinted with permission from reference 22.

The largest polypeptide observed has a neutral mass of 9174.51Da and carries 12 charges on ions at m/z 765.01. This charge state was determined readily at the resolution used.

Evidence for a number of modified peptides was collected in the course of this study using the loosely constrained precursor search.[27] One example (**Figure 19**), curated and localized manually using ProSight's Sequence Gazer tool, involves a mass-shift in an acid digest product of ribosomal protein L10. Initially identified with E-value 1.672×10^{-7} based on 4 b and y ions, its E-value improved to 1.31×10^{-25}, from 12 b and y ions, after application of an ~27 Da mass shift indicative of a single amino acid substitution from Ser to Asn.

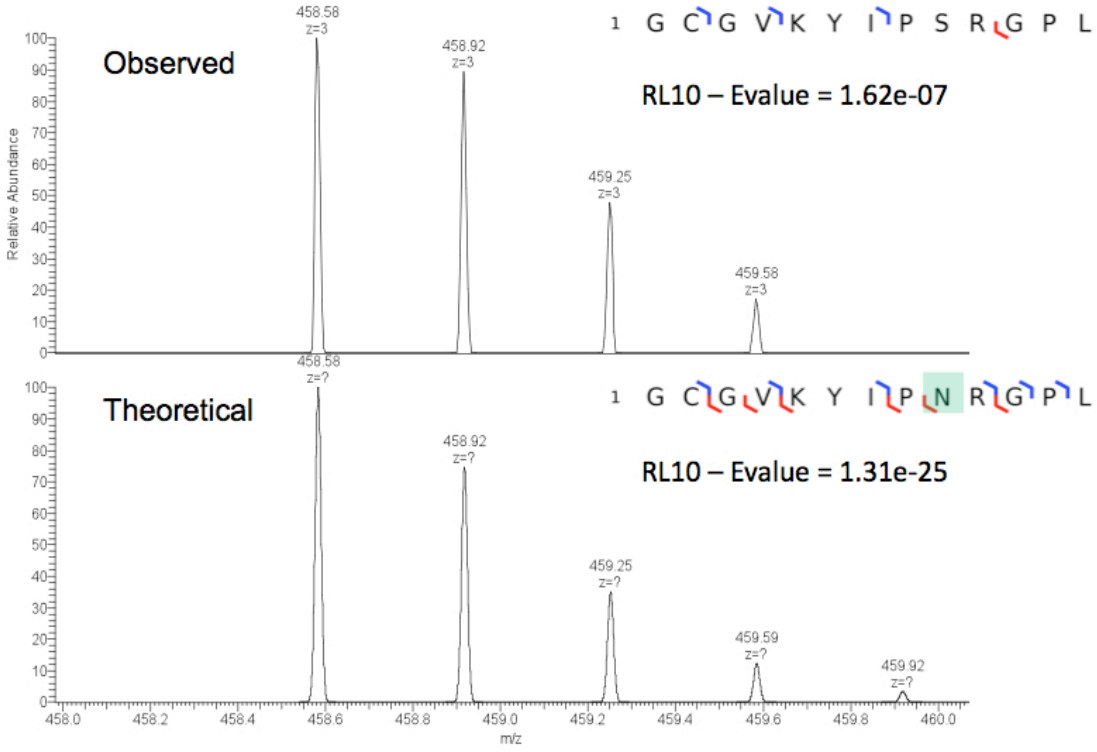

Figure 19. Comparison of the observed (top) and theoretical (bottom) isotope clusters from peptide [194-206] from ribosomal protein L10. Also shown are the sequences, matched fragment ions, and corresponding Evalues. The single amino acid substitution is highlighted in green in the theoretical sequence and satisfies the observed precursor mass difference.

The substitution was later confirmed as a naturally occurring variation.[83]

Conclusions. The novel middle-out workflow presented here demonstrates a rapid automated method for analysis of a small protein mixture, allowing high confidence identifications of nearly 30% of the expected peptidome, recognition of protein modifications and an average of 46% sequence coverage of 70 of 79 possible parent proteins. We plan to introduce modifications to improve the analysis further,

notably the use of still higher resolution on the LTQ-Orbitrap instrument to permit deconvolution and analysis of peptides with higher charge states.

Chapter 4: Analysis of Oligomeric State Specific Peptides from Lys63 Linked Polyubiquitin

Introduction

Ubiquitin is a 76 residue protein conserved among all eukaryotes that most famously targets substrate proteins for proteasomal degradation. Many ubiquitin monomers are joined to one another and to the substrate protein of interest via isopeptide bonds that join the C terminal glycine residue to one of seven possible lysines on ubiquitin, and a solvent accessible lysine on the substrate. Although proteasomal degradation is the most well known outcome of ubiquitination, there are many more that are directly linked to the length and linkage pattern of the modifying ubiquitin moiety. In particular, Lys63-linked polyubiquitin is thought to affect at least four different non-degradative pathways within the cell; DNA damage repair, cellular signaling, intracellular trafficking, and ribosomal biogenesis.[84]

Proteomics has become an integral tool for studying ubiquitination and its effects in a high throughput manner. Typically, ubiquitinated proteins are immunoprecipitated and subsequently digested using trypsin. Cleavage with trypsin displays a convenient GlyGly tag on isopeptide bonded lysines, which is easily accommodated as a variable post translational modification through nearly all proteomics search engines. Alternatively, the substrate peptides can be

immunoprecipitated using an anti-GlyGly antibody for identification of the site containing peptides only.[85] Using these methods, the data provided will reveal solely the site of ubiquitination. There is no information provided about the nature of the modification, itself. Currently, methods used for determining the length of a modifying ubiquitin moiety require observation of the intact mass of the substrate protein in addition to the polyubiquitin chain. Due to the large size of the monomer (~8.5kDa), addition of only a few ubiquitins to all but the smallest proteins will likely render them unsuitable for top down mass spectrometric analysis. Denaturing polyacrylamide gel electrophoresis has become a standard method used to formulate ubiquitin ladders, where the protein of interest is observed in several bands separated by electrophoretic shifts indicative of the addition of a single ubiquitin monomer. These ladders allow one to accurately tell how many ubiquitins are on a given substrate, but fail to establish the linkage pattern between each unit, or even distinguish whether there is, in fact, a polyubiquitin chain or multiple monoubiquitinated sites. Even immunoblotting with linkage specific antibodies only confirms the presence or absence of a given linkage type at one isopeptide bond in a potentially extensively branched molecule which could contain several.

Serendipitously, the primary sequence of ubiquitin contains no Asp residues between the C terminus and Lys63. This means that if there is a polyubiquitin species containing only Lys63 linkages that is cleaved using Asp specific hydrolysis, the sum of the C terminal peptide masses (minus the loss of one water molecule for every isopeptide bond formed) is a direct indicator of the length of the polyubiquitin chain. For example, as depicted in **Figure 20**, a pentamer would have a monoisotopic mass

of the C terminal peptide (DYNIQKESTLHLVLRLRGG) mass multiplied by the number of contributing monomer units minus 1 water molecule per isopeptide bond.

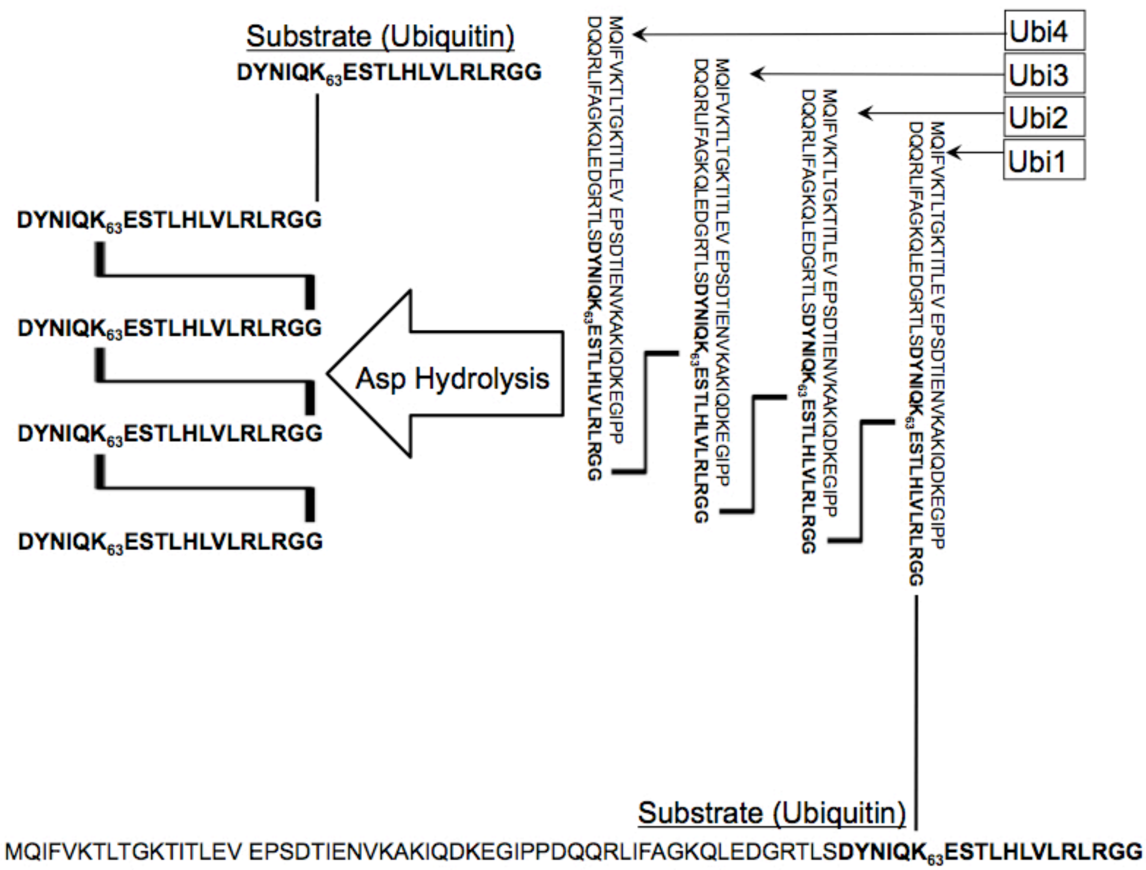

Figure 20. A schematic representation demonstrating the unbroken chain of C terminal ubiquitin following Asp specific hydrolysis. Shown is a Lys63 linked tetramer conjugated to a ubiquitin substrate (the mass is equal to that of the pentamer). Neutral mass of the resulting peptide is equal to the mass of the C terminal peptide (2211.21Da) multiplied by the number of monomeric units (5) minus one water molecule per isopeptide bond (4*18.01).

$$10984.01\text{Da} = ((2211.21\text{Da} \times 5)-(4 \times 18.01\text{Da}))$$

This method is clearly well suited to ubiquituin polymers that are not conjugated to a substrate, but it can also be useful for ubiquitin oligomers that are acting as protein modifiers. A prerequisite for using this method in a discovery type way would be to have *a priori* knowledge of the actual sites of ubiquitination or of which proteins are modified. This is a task for which trypsin has proven quite successful. Once the ubiquitination sites are identified, one can confirm the number of Lys63 linked ubiquituins on it based on the change in observed mass from the unmodified peptide alone. In a typical complex mixture proteomics experiment, the sample would contain far too many peptides for this type of analysis, but the added constraint of having a known ubiquitination site on the peptide of interest significantly decreases the number of molecular weight possibilities. Of particular importance is that the peptide that also contains Lys63 following Asp specific proteolysis is the most C terminal on ubiquitin, meaning that it will be conjugated to the substrate regardless of any type of further linkage. In the least informative scenario, high resolution accurate mass analysis with this longer "tag" will provide a higher level of confidence for previously arrived at tryptic digestion results. In situations in which there are multiple consecutive Lys63 linkages on a substrate it will provide an unprecedented look at the length of the ubiquitin moiety associated with a specific protein substrate under given cellular or *in vitro* conditions.

Polyubiquitin chains have been examined mass spectrometrically with the goal of distinguishing specific linkages before. Aebersold and co-workers were able to use bottom up and selected reaction monitoring to extract quantitative data

concerning the overall character of ubiquitin linkages from a complex mixture.[86] This method provides statistical information on the abundance of different linkages in the complex mixture, but fails to offer any information on the ubiquitin moiety associated with a specific substrate protein, a relationship that would be greatly beneficial to characterize. In a work that utilized non-tryptic enzymatic digestion methods for linkage specific analysis, Goldberg, Gygi, and co-workers were able to observe evidence for the formation of at least di-ubiquitin for all seven possible Lys residues.[87] Clemmer and co-workers were able to distinguish Lys63 and Lys48 linked dimers of intact ubiquitin using differences in collisional cross section with ion mobility mass spectrometry.[88] Oldham and co-workers demonstrated the successful linkage specific analysis of diubiquitin using a top down approach.[89] These analyses demonstrated the untapped versatility offered by non-tryptic methods, with the two more recent approaches using higher molecular weight species that can define linkage patterns. Unfortunately with most commercially available mass spectrometers, the resolution required for confident observation of intact species composed of more than a few ubiquitins is not yet attainable. The same problem occurs when a ubiquitin chain is conjugated to a substrate protein.

Materials and Methods

Ubiquitin Expression. All expression, purification, and chain ligation was performed in the lab of Dr. David Fushman by Mark Nakasone. Competent *E. coli* BL-21(DE3) Rosetta™ cells were transformed to express wild type, K63R mutant,

K63 (where all Lys other than Lys63 are mutated to Arg), K0 (where all seven Lys are mutated to Arg), and an additional mutant form of ubiquitin in which an extra Asp residue is present on the C-terminus to prevent chain elongation (D77). Following expression cells were lysed by sonication, and the lysate was cleared via centrifugation at 22,000rpm. The supernatant was precipitated by the addition of 1.5%(v/v) perchloric acid, and again centrifuged. The ubiquitin containing supernatant was then dialyzed using a 3kDa molecular weight cutoff membrane against 2L of 50mM ammonium acetate (pH 4.5). The ubiquitin containing solution was then loaded onto a 5mL cation exchange column (GE Life Sciences, Pittsburgh, PA) and eluted over a 22 column volume gradient from 0-40% solvent B (50mM ammonium acetate, 1M NaCl, pH 4.5). Ubiquitin containing fractions were pooled and the purity was confirmed using SDS-PAGE.

Formation of Dimeric and Trimeric Ubiquitin. All polyubiquitin chains were synthesized enzymatically from the monomers described above using the method of Pickart and Rassi[90] and Dong[91] et al with slight modifications. After expression in *E. coli* cells, UBE1 (E1 activating enzyme), Ubch5b (used as substrate), and MMS2 (necessary for UBC13 E2 activity) were purified using a 5mL His Trap column (GE Life sciences). The Lys63 specific E2, UBC13, was expressed as a fusion protein in BL-21(DE3) cells. Yeast ubiquitin C-terminal hydrolase (YUH1) was expressed without a tag and purified according to the method outlined by Johnston and co-workers. To generate dimeric K63 linked chains D77 and K63R monomers were allowed to react in the presence of UBE1, MMS2, UBC13, TCEP

and ATP. Trimeric K63-linked chains were synthesized by removing D77 from the Lys63 dimer, and then using the same reaction conditions with the D77 monomer (2+1) to obtain the trimer. After a period of twenty-four hours ubiquitin chain reactions were stopped by the addition of 10mL of cation buffer A (50mM ammonium acetate, pH 4.5), the solution was spun at 13,000rpm to remove precipitated E1 and E2 enzymes, then slowly injected on a 5mL cation exchange column at 0.2mL/min using an FPLC. Once no absorbance was detected the attached ubiquitin and polyubiquitin species were eluted with 50mM ammonium acetate containing 1M NaCl, pH 4.5. The eluted ubiquitin was then exchanged into PBS pH7.4, concentrated to volume of 1mL and the dimeric and trimeric species were resolved by size exclusion chromatography (SEC). The purity of the SEC fractions were confirmed using 15% SDS-PAGE and MALDI.

Mono and Polyubiquitinated UbcH5b. Ubiquitinated Ubch5b was created using the same reaction conditions as the chains with the exception of the concentration of monomeric ubiquitin and Ubch5b. To create mono-ubiquitinated Ubch5b, the same reaction outlined above was performed with 75µM Ubch5b and 1mM of K0 ubiquitin in a 2mL volume. The Lys63 polyubiquitinated UbcH5b[92] was created under similar conditions with 75µM Ubch5b and 1.5mM Lys63 ubiquitin. After twenty-four hours each reaction was diluted with 10mL of PBS with 0.5M NaCl, pH 7.4, then loaded on a 1mL His-Trap column. The elution, which contained a mixture of Ubch5b and Ubch5b with a varying number of ubiquitins, was subsequently concentrated and exchanged into PBS, 10mM DTT, pH7.4 to a final

volume of 1mL. This was fractionated by SEC, and different oligomeric states of ubiquitinated Ubch5b were detected using 15% SDS-PAGE.

In Gel Acid Digestion of Dimeric and Trimeric Ubiquitin. Once in our hands, electrophoretically separated ubiquitin dimers and trimers were excised using a transfer pipette from a polyacrylamide gel and destained in 50/50 ACN/water. After destaining the gel pieces were dehydrated in ACN and subsequently rehydrated in 12.5% acetic acid solution for 30 minutes. Just prior to digestion additional solvent was added to ensure the gel pieces remained saturated. Gel pieces were digested for 5 minutes at 140ºC using 300W of microwave energy to maintain temperature in a CEM Discover microwave (Matthews, NC). After digestion the peptides were extracted directly and desalted using Millipore C18 ZipTips (Billerica, MA) according to the manufacturer's instructions. Mass spectra were acquired in linear mode on a Kratos Axima CFR MALDI-TOF MS (Shimadzu, Columbia, MD).

In Solution Digestion of Di- through Hexa-Ubiquitin. Lys63 linked polyubiquitin (n = 2-7) was formulated as indicated above. The protein was purified by SEC and confirmed by SDS-PAGE. The resulting protein solution was precipitated[58] and redissolved in 12.5% acetic acid. Digestion was carried out at 140°C in a CEM Discover microwave for 20 minutes while irradiating with 300W of energy. The resulting peptides were then cleaned up using C18 Ziptips (Millipore) according to the manufacturer's instructions prior to analysis by MALDI-TOF.

Immunoprecipitation and LC-MS/MS. Enzymatic digestion was performed using sequencing grade trypsin acquired from Promega (Madison, WI), and digestion was conducted according to the supplied protocol. Lysine63 specific antibody (Millipore, Billerica, MA) was conjugated to agarose beads using the ThermoPierce Direct IP kit (Rockford, IL) according to the manufacturer's instructions. All IP reactions were performed according the manufacturer's protocol. High resolution mass spectrometry was performed using a Thermo LTQ-Orbirap XL. Intact peptide masses were acquired at maximum resolution and product ion measurement was performed at 15,000 R_p (at m/z 400). Prior to acquisition of mass spectra, peptides were separated using reversed phase chromatography with either an Agilent Zorbax C3 column or a GraceVydac C18 column at a flow rate of 500nL/min. Data acquired from tryptic digestion was searched with the MASCOT (Manchester, UK) search engine against human proteins from the UniProtKB. Precursor mass tolerance was set to 10ppm and fragment ion mass tolerance was set to .05Da. The diglycine tag was used as a variable modification for Lysine residues.

Results and Discussion

MALDI Characterization and In Gel and In Solution Digestion of Oligoubiquitin. Due to the rapid nature of microwave-supported acid digestion in conjunction with its kinetic selectivity, short digestion times (<90 seconds) were tested initially in solution to confirm that the expected cleavage sites were observed using wild type monoubiquitin (**Figure 21**).

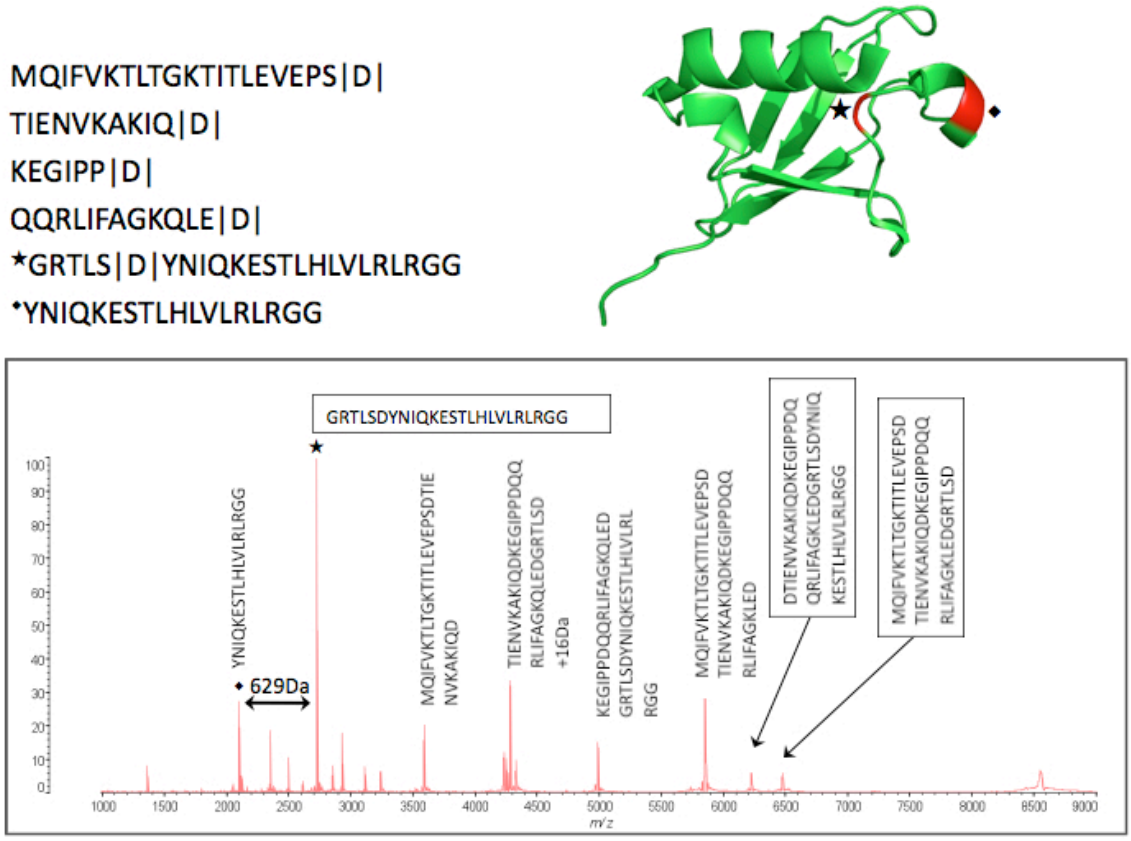

Figure 21. Shown on the top left of the figure is the amino acid sequence of ubiquitin with lines shown at possible microwave-supported acid cleavage sites. An asterisk and a diamond show the sites of cleavage for the most abundant (1 missed cleavage) and second most abundant (0 missed cleavages) peaks of the C terminus containing peptides, respectively. Also shown are the corresponding sites on an NMR solution structure (PDB ID: 1d3z).[93]

Interestingly, the peptide with the highest relative abundance is the segment from Gly53 to the C terminus. This peptide contains 1 missed cleavage site. Typically, the most abundant peptides are from the termini, but without any missed cleavages, a

result that generally agrees with the idea of a globular protein in solution as a wound up ball of spaghetti with the termini exposed. At the molecular level, microwave-supported acid hydrolysis must have some dependence on solvent accessibility to potential cleavage sites. This inference is underscored by the mechanisms proposed by Inglis and coworkers, which require cyclization of the carboxylic acid side chain of Asp to adjacent residues.[37] Hydrogen bond donation to and from the adjacent backbone nitrogen atoms makes it unlikely the cyclization process would occur, inhibiting the formation of the proposed cleavage products. Nuclear magnetic resonance and X-ray crystal structures display considerable evidence that the most C terminal Asp residue on ubiquitin is in involved in hydrogen bonding in a 3-10 helix,[93] which could explain the lack of abundance of the C terminal peptide without any missed cleavages (**Figure 22**).

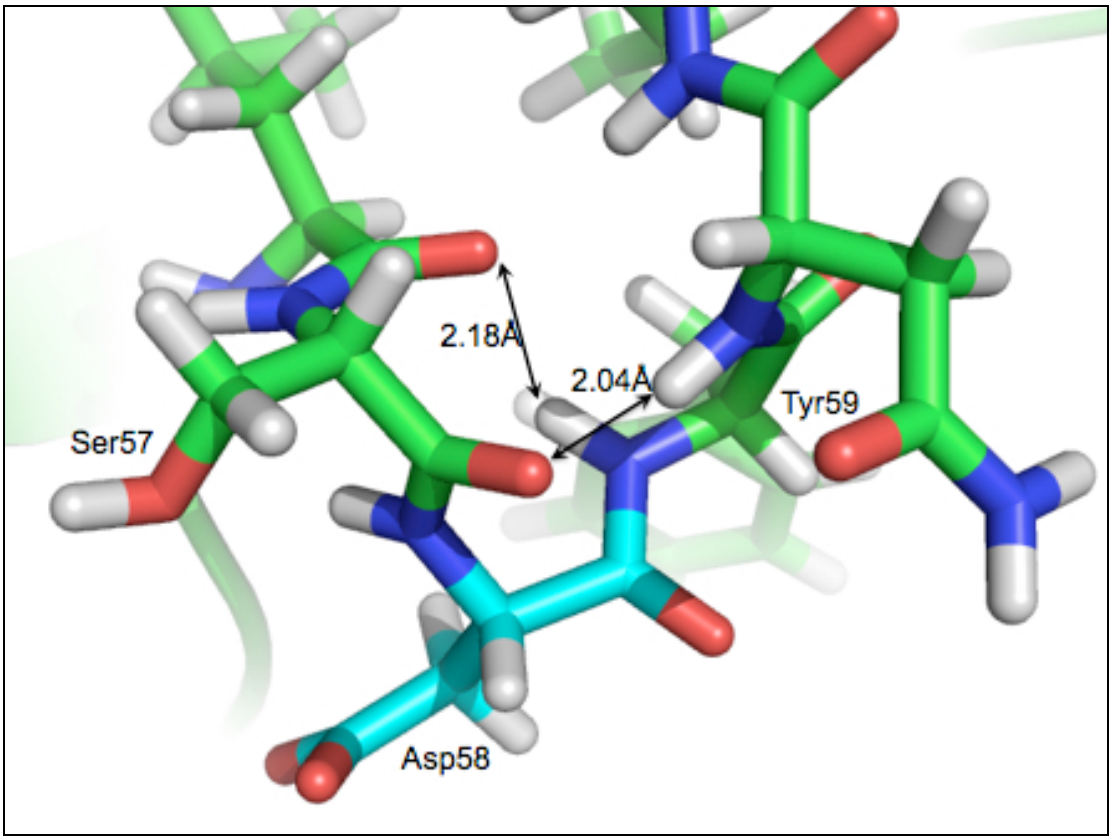

Figure 22. Zoomed in region containing Asp58 from the NMR solution structure of monoubiquitin.[93] Measurements of the backbone nitrogen (on Tyr59) and backbone oxygen (on Ser57) to nearby hydrogen bonding partners are shown.

Strong correlation between the majority of peaks from two dimensional (^{15}N, ^{1}H) heteronuclear single quantum coherence (HSQC) spectra acquired by Mark Nakasone for monoubiquitin in water and in 12.5% (v/v) acetic acid at room temperature supports this conclusion. Even under the harsh conditions of microwave-supported acid digestion, the low abundance of the most C terminal peptide suggests that some semblance of secondary structure is maintained prior to complete thermal denaturation.

Initial analysis was simplified by the fact that the expected intact protein products were isolated by SDS PAGE into distinct gel bands at high concentrations. This allowed for effective in gel acid digestion and subsequent extraction of the large oligomeric state specific peptides. Observation of the C terminal inclusive peptide was paramount since it contains the site of conjugation to the substrate and/or preceding ubiquitin moiety. Two particular doublets of peaks were found close to the expected masses from the calculation described in **Figure 20**. The two most abundant peaks in the region of interest were found to be the result of C terminal Asp cleavage as expected, and they were each accompanied by a peak 115Da lower, indicating the loss of one Asp residue (**Figure 23**). The two doublets were separated by ~629 Da, which correlates to the mass difference between the C terminal peptide with one missed cleavage and that with zero missed cleavages (DGRTLS). This same pattern of increasing by one missed cleavage was observed in the spectrum of the acid digested trimer (where one branch contained a K63R mutation), with an additional incremental mass shift.

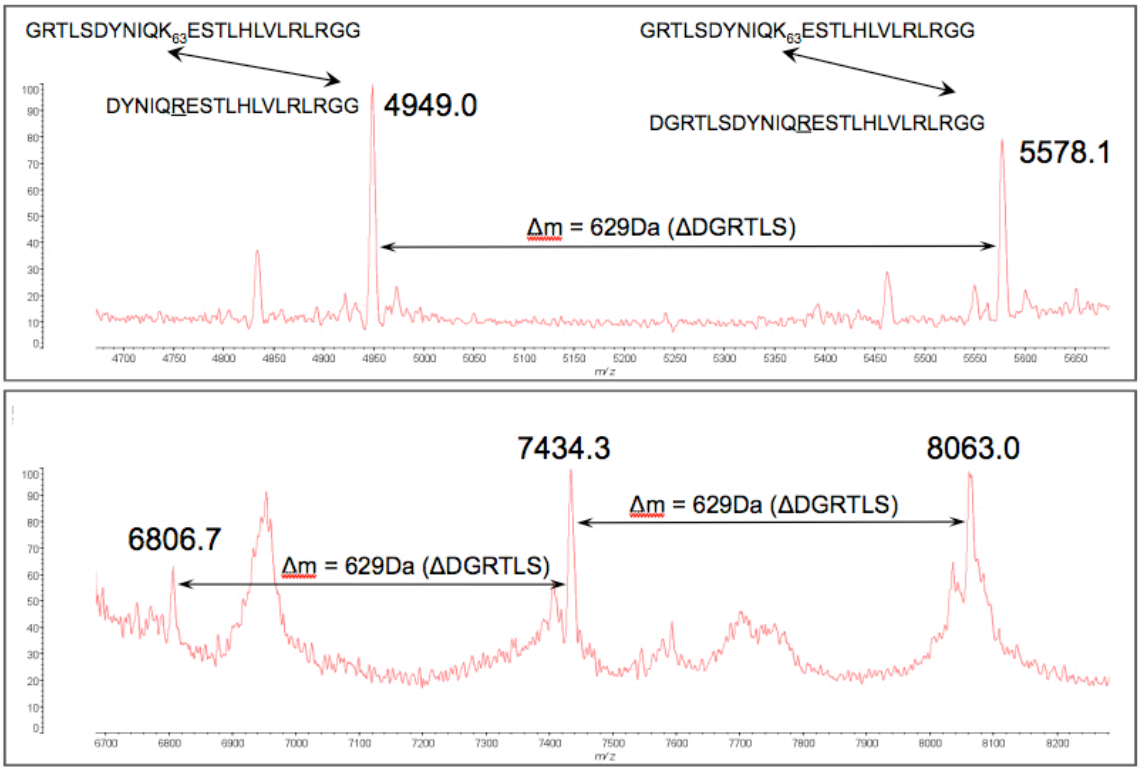

Figure 23. Shown above are regions of interest from MALDI spectra of in gel microwave-supported acid digestion products of di- (top) and tri-ubiquitin. The characteristic 115Da pairs are observed as well as mass differences that correlate to 1 (top) and 2 additional missed cleavages (bottom).

LC-MS/MS and Diagnostic B ions. Following the initial characterization of the reaction products by MALDI, it was essential to demonstrate that this method is compatible with LC-MS/MS methods. Indeed, it could be expected that the increase in sensitivity provided by chromatographic separation would benefit the analysis of larger peptides that have higher ionization energy requirements in MALDI. To confirm the expected results, multiply-ubiquitinated species engineered in the same manner were digested and subsequently separated chromatographically prior to high-resolution accurate mass precursor and product ion analysis. As stated in the

introduction, current methods in ubiquitin proteomics allow for identification of the ubiquitinated peptide, but fail to provide any information on the PTM itself. This is achieved by digesting with trypsin and observing a GlyGly tag from the C terminus of ubiquitin on the Lysine of the conjugated peptide. Due to the length of the tag left behind on the conjugated peptide using acid digestion (at least 17 residues), it is impossible to carry over this methodology. Asp selective digestion, however, does allow for highly specific analysis by capitalizing on high resolution accurate mass measurements of diagnostic b ions from the N terminus of the conjugated ubiquitin peptide. It was previously demonstrated using peptide mass fingerprinting that the most likely cleavage site for observing the C terminal peptide of ubiquitin occurred at Asp52. From this point forward (toward the N terminus), it is possible to formulate reconstructed ion chromatograms for each individual b ion up until the isopeptide bonded Lys63 (**Figure 24**). These ions prove to be diagnostic for isopeptides due to the fact that they are present from each possible C terminal peptide regardless of its position in the ubiquitin oligomeric species, whereas this is not the case with any given y ions.

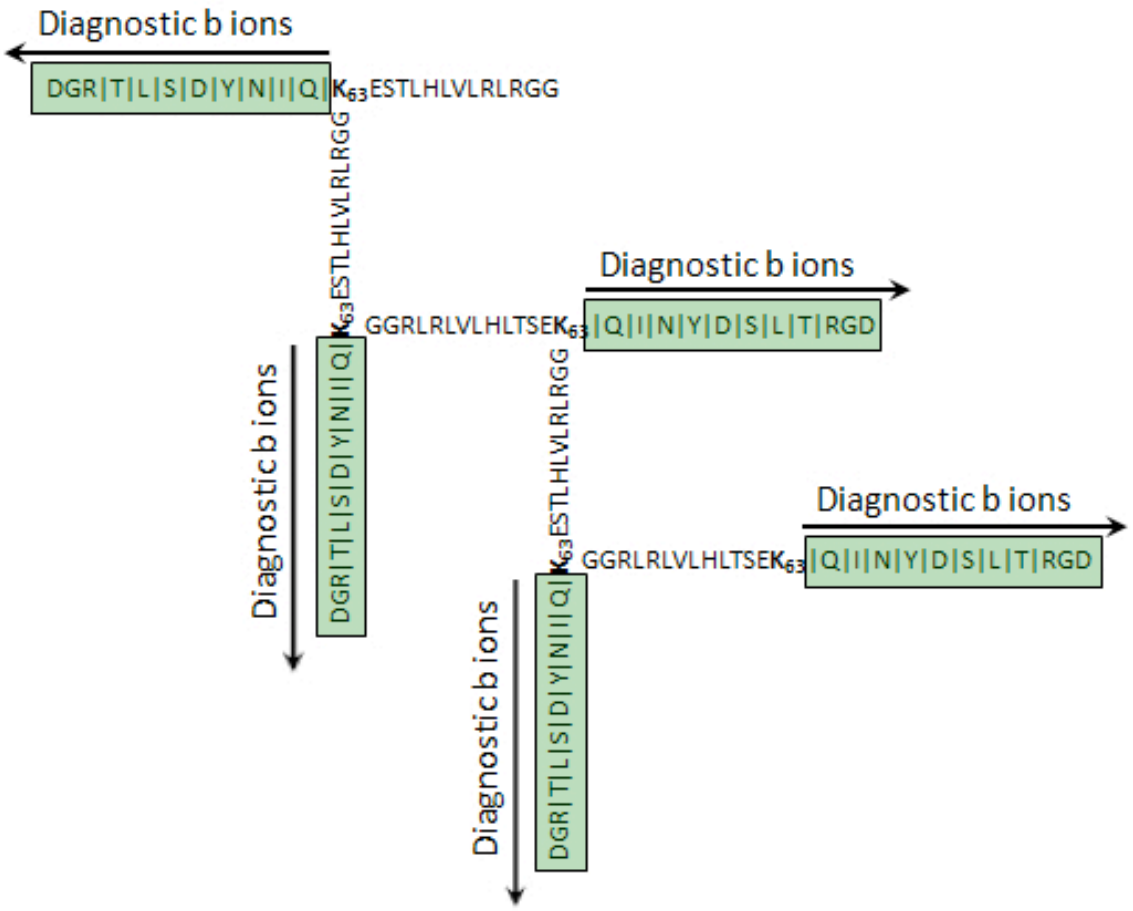

Figure 24. Shown above is a primary sequence representation of a Lys63 linked pentaubiquitin peptide resulting from Asp-specific proteolysis (each constituent peptide has 1 missed cleavage at Asp58). The b ions stemming from the residues highlighted in green plus any additional ions toward the next possible cleavage site from the C terminus are unchanged regardless of the oligomeric unit of Lys63 linked ubiquitins.

Especially with branched peptides, as the peptide gets larger and contains more branches the product ion spectral complexity increases, and so does the number of possible product ions – most of which are likely to be internal and unmatchable to any predicted fragments. Currently, there are no bioinformatic methods to automate

the identification of these branched oligomers. Ubiquitin is a small protein that produces a small number of peptides under the previously outlined conditions with Asp selective hydrolysis. This number is sufficiently small that with high resolution accurate mass measurement of modified precursor masses and known sites of ubiquitination from tryptic digestion experiments run in parallel, manual interpretation of the changes in precursor mass should be attributable to a previously identified peptide in simple mixtures.

Lys63 Linkage Specific Protein and Peptide Immunoprecipitation.

Immunoprecipitation has been used extensively in biochemistry for the purification and analysis of individual proteins and protein complexes. Analysis of ubiquitin, specifically, has benefited greatly from the production of antibodies that are specific to oligomeric states of ubiquitin, and more recently, specific ubiquitin-ubiquitin linkages. Lys63 linkage specific peptides have been formulated in two different ways using intact diubiquitin and a synthetic branched peptide as immunogens. In this case the branched peptide derived antibody is better suited for analyzing oligomeric state specific peptides since it can be used at both the protein and the peptide level. The enhanced specificity can be exploited for immunoprecipitating Lys63 linkage containing (poly)ubiquitinated proteins as well as for pulling down Lys63 linkage containing peptides that result from Asp specific digestion of the complex mixture. This process proves to be crucial for successful analysis of the oligomeric state specific peptides due to an inherent increase in dynamic range associated with polyubiquitin analysis. For example, if one is attempting to observe a pentamer

specific ubiquitin peptide, there is only one informative peptide per polyubiquitin. In addition to the one peptide of interest, there are four peptides that are of no interest per constituent monomeric unit. This phenomenon can be effectively demonstrated by observed the change in relative abundance of a single peptide of interest from the ubiquitinated substrate as an increasing number of ubiquitins are attached (**Figure 25**).

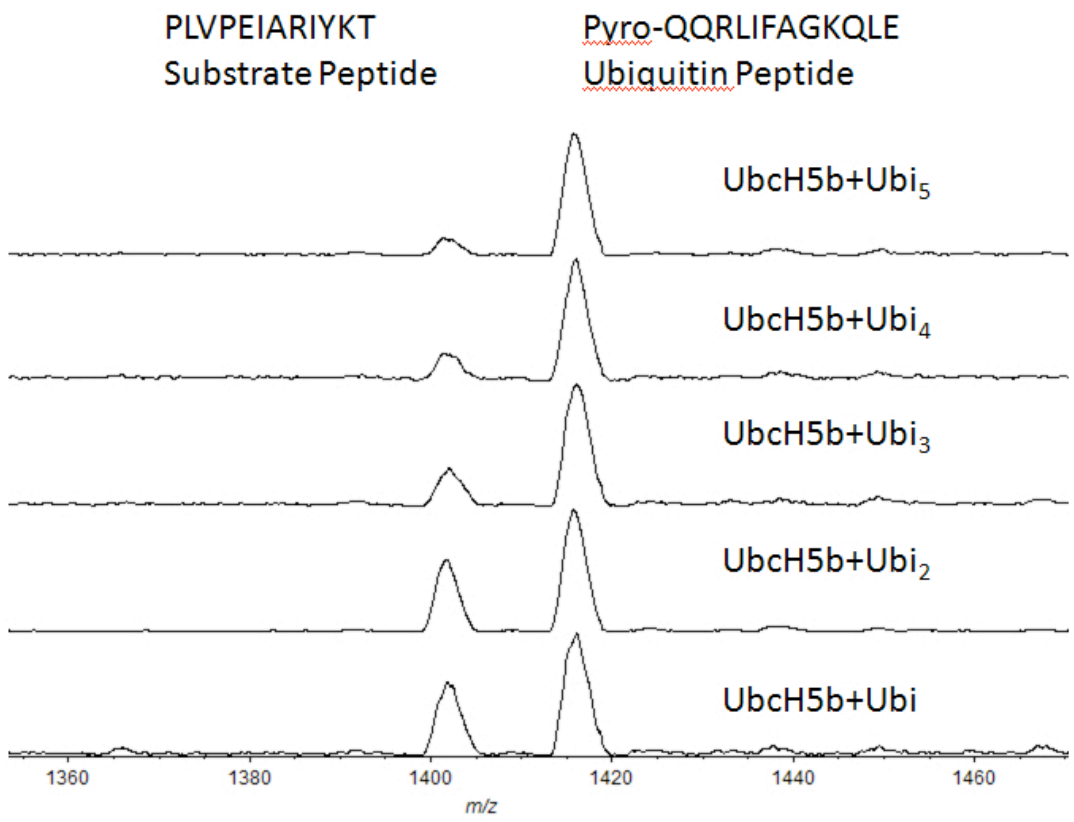

Figure 25. A series of MALDI spectra excised from gel bands correlating to mono- through penta-ubiquitinated UbcH5b. The relative abundance of a single peptide from UbcH5b (left peak) is compared to that of a ubiquitin peptide (right peak).

A small increase is easily overcome, but the problem is also dependent on the number of ubiquitin moieties present in the complex mixture, and the relative number of linkages associated with each one.

A mixture of wild type ubiquitin oligomers was digested and subsequently injected for LC-MS/MS analysis. Higher order oligomeric state mixtures of wild type polyubiquitin species were used to further demonstrate the viability of the above outlined immunoprecipitation method. Size exclusion chromatographic fractions containing mixtures of engineered Lys63-only linkage containing polyubiquitin were digested and subsequently immunoprecipitated with the Lys63 linkage specific antibody. Higher order ubiquitin oligomers were fractionated using RP-LC prior to high resolution precursor and fragment ion detection. All isopeptides eluted with similar retention times, as confirmed by viewing the extracted ion chromatograms of diagnostic b ions. Deconvolution of the averaged spectra over the elution profile of all isopeptide bond containing peptides revealed several high molecular weight species (**Figure 26**).

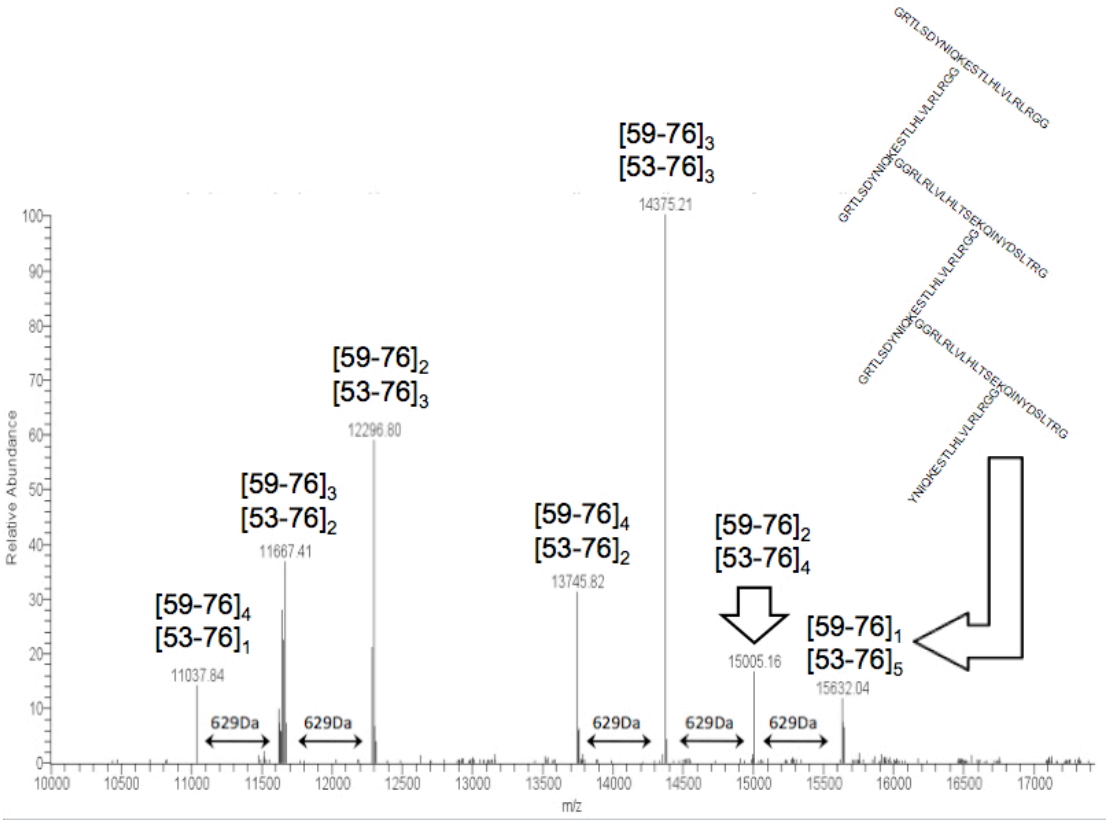

Figure 26. Deconvoluted spectra averaged over the course of the elution profile of all isopeptides. Labeled peaks demonstrate assignment of Lys63 containing isopeptides. Peaks separated by arrows in spectrum are ~629Da apart, showing the single missed cleavage mass shift associated with the Lys63 containing peptide is observed frequently in higher order oligomers. Primary sequence structure with isopeptide bonds of the highest molecular weight species is shown on the right.

The high molecular weight species depicted in **Figure 26** show characteristic mass shifts associated with the Lys63 containing peptide. The 629Da mass shift indicative of a change in the number of missed cleavages observed from in gel digestion is carried over to higher molecular weight oligomers. The mass shifts are the result of

the Lys63 linked part of the polyubiquitin chain. In **Figure 26**, a quartet of peaks are separated by 629Da each, indicative of a missed cleavage on three Lys63 branches. The triplet of peaks displays the same phenomenon, showing there are at least two Lys63 branches. Larger branched peptides are observed with several charge states and require higher sensitivity for unambiguous identification. Microwave-supported acid hydrolysis has kinetic selectivity for either side of Asp residues,[43] often producing peptides that differ only by the presence or absence of an Asp residue at either terminus. While this is not a major concern for typical proteomics experiments examining unbranched peptides,[22, 30, 35] the problem of having multiple ions per cleavage site is compounded by the increased number of cleavage sites per isopeptide. Additionally, the signal for each distinct mass is spread out over multiple charge states, thereby decreasing the sensitivity of the analysis. Matrix assisted laser desorption ionization MS circumvents this problem by producing predominantly singly charged ions. Analysis was performed on a mixture of di- through hexa-ubiquitin using MALDI-TOF and sequential singly charged 629Da doublets are clearly observed, analogous to a 1D gel of multiply ubiquitinated species (**Figure 27**).

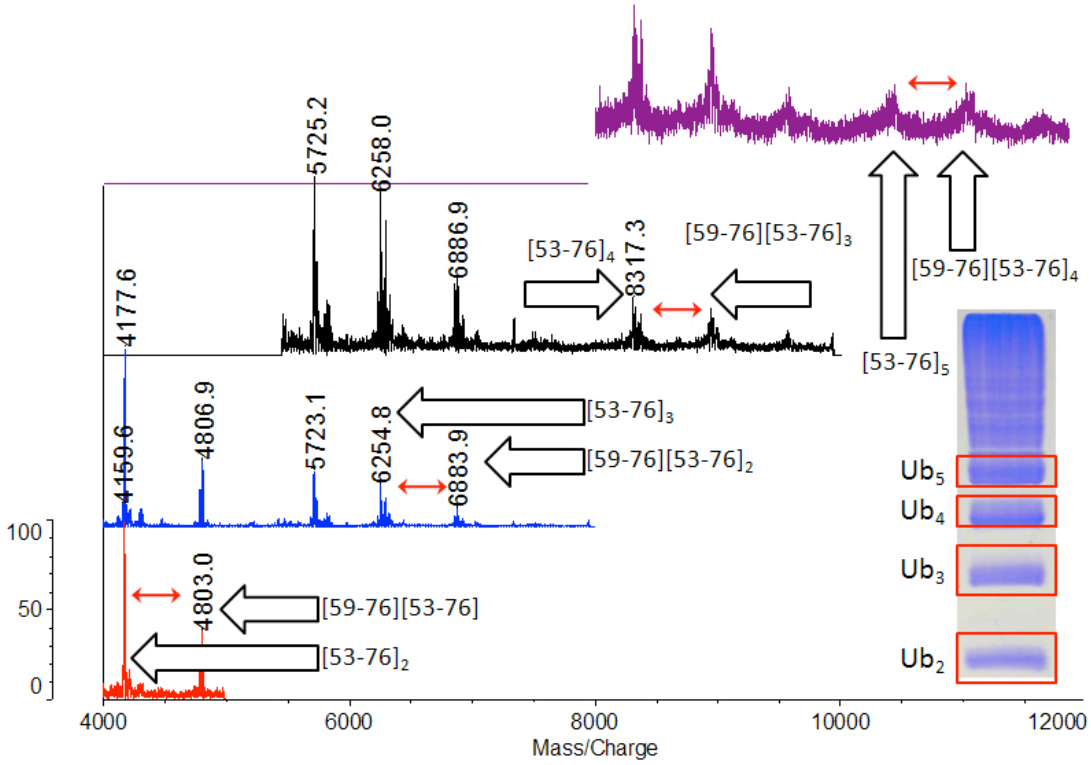

Figure 27. MALDI spectra demonstrating a stepwise increase in C terminal ubiquitin peptides from di- through penta-ubiquitin. Red double-headed arrows highlight characteristic 629Da doublets. In the bottom right inset is a gel providing an equivalently informative ubiquitin ladder.

Identification of Ubiquitination Sites on UbcH5b. The E2 ligase, UbcH5b, is implicated directly in the ubiquitination of the p53 tumor suppressor.[94] To test the hypothesis that one could attribute a specific linkage combination to a known substrate protein based on precursor mass alone, a mixture of different oligomeric states of Lys63 ubiquitin was attached to the E2 enzyme UbcH5b enzymatically. Initially, work was done using monoubiquitinated UbcH5b. High resolution fragment ion detection revealed the presence of several peaks that contained isopeptide specific

b ions. Further analysis of the precursors that produced the diagnostic b ions demonstrated predicted masses of specific acid digestion isopeptides. Two peptides containing three possible lysines on UbcH5b, on the N terminal peptide, ALKIRHKELN (following removal of the initiator methionine) (**Figure 28**) and on Lys128, from the peptide DPLVPEIARIYKT were assigned with precursor masses matching within 5ppm (**Figure 29**) (both sites were confirmed using tryptic digestion (**Figure 30**)).

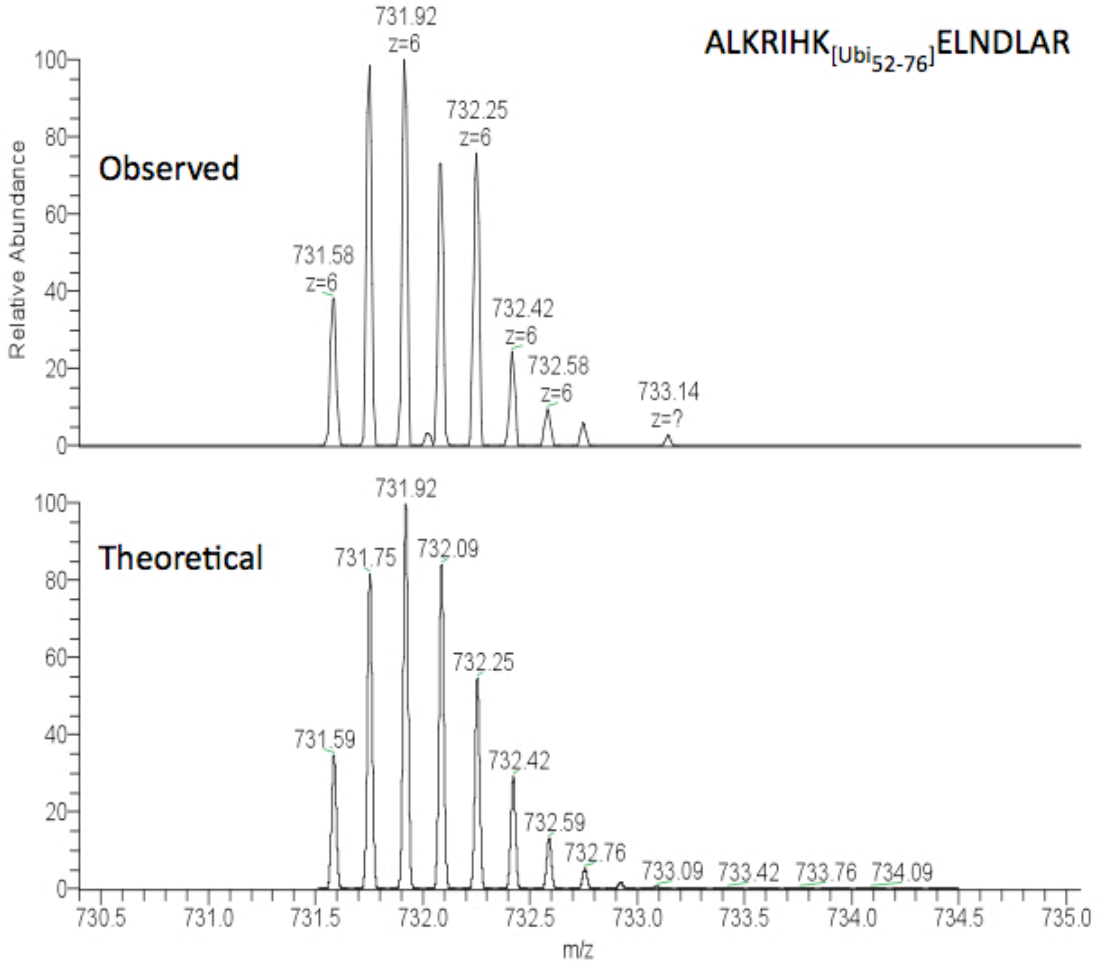

Figure 28. Isotopic cluster comparison of the observed (top) and theoretical (bottom) ubiquitinated peptides [2-15] from UbcH5b.

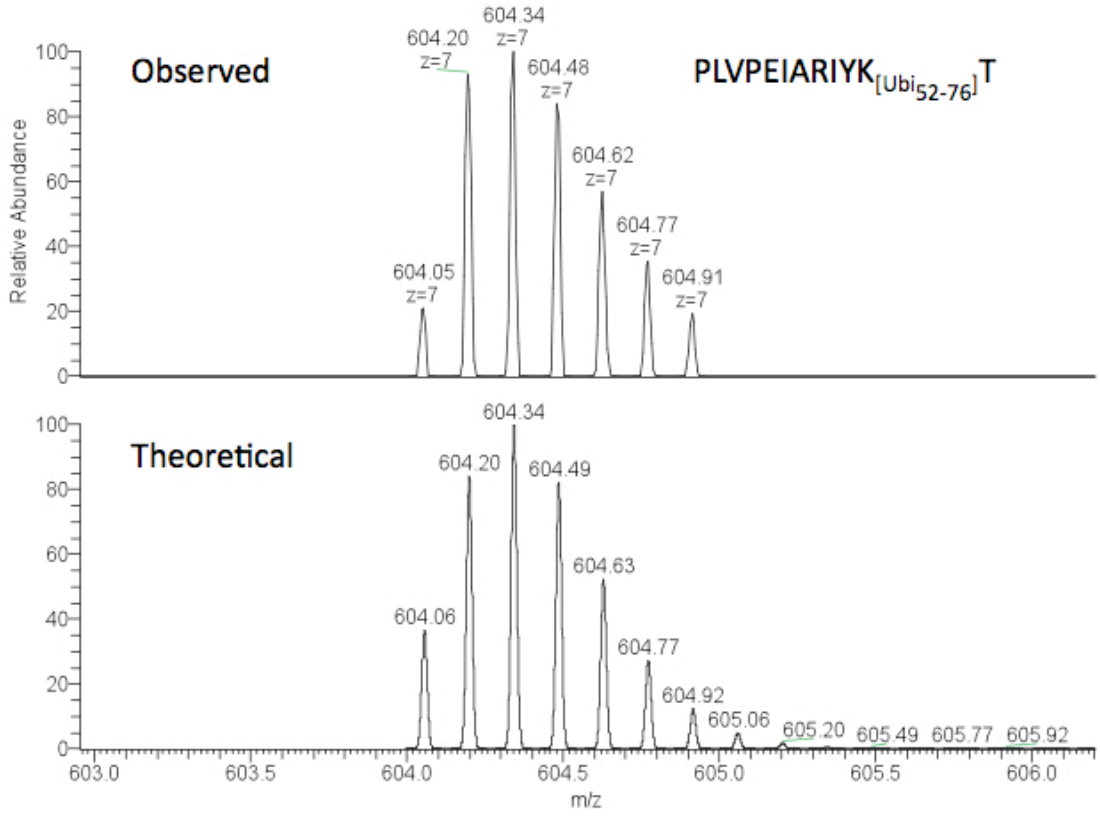

Figure 29. Isotopic cluster comparison of the observed (top) and theoretical (bottom) ubiquitinated peptide [118-129] from UbcH5b.

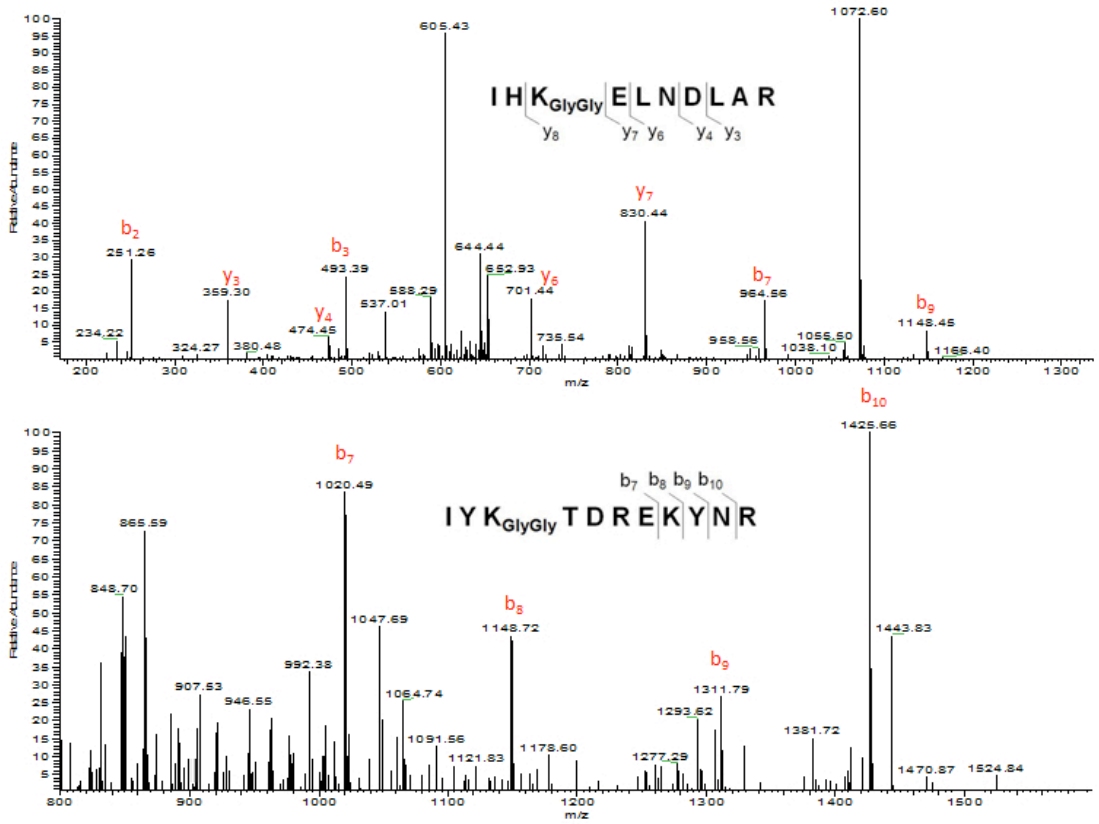

Figure 30. Product ion spectra of tryptic peptides from UbcH5b displaying the characteristic di-glycine tag indicative of a ubiquitination site.

A useful property of this method is the inherent high mass of peptides that contain Lys63 isopeptide bonds. For example, Lys63 linked diubiquitin will produce a C terminal isopeptide approximately twice the length of the smallest C terminal acid cleavage product ((2096Da x 2)-18Da), with the substrate conjugated species being even larger. The minimal mass of >4kDa allows for fast and efficient mass biased partitioning to exclude peptides of little significance to the analysis using commercially available molecular weight cutoff filters.[63] This technique is further validated by the fact that Asp specific hydrolysis produces peptides that are longer on average than those produced using the standard tryptic digestion method. This method of mass biased partitioning via molecular weight cutoff filters was applied to

the sample prior to immunoprecipitation of the Lys63 linkage containing peptides to reduce complexity and minimize nonspecific binding to the antibody. After antibody elution and injection, the observed m/z ratios can be deconvoluted, corrected to attain the monoisotopic mass, and then all possible mass combinations of the 0 and 1 missed cleavage peptides (the two most commonly produced) can be subtracted. Since UbcH5b is the substrate, an *in silico* digestion of the primary sequence provided candidate neutral masses to search the leftover mass against. Performing this type of analysis on the conjugated UbcH5b sample resulted in observation of mono- and di-ubiquitin on the acid digestion peptide DDPLVPEIARIYKT. The only lysine, Lys128 is ubiquitinated (**Figure 31**).

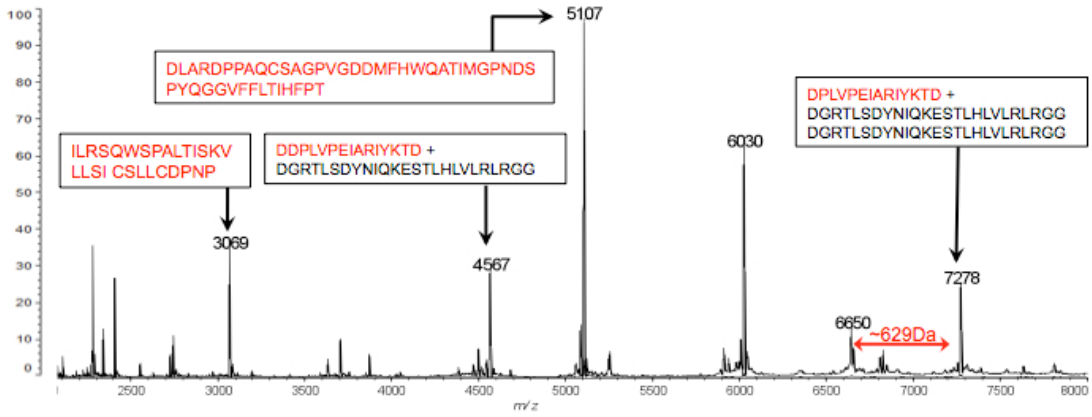

Figure 31. MALDI spectrum of the acid hydrolysis products of mixture of UbcH5b conjugated to differing numbers of ubiquitins. Peptides derived from UbcH5b are shown in red font, peptides from ubiquitin are shown in black font.

In the above Figure, characteristic 629Da mass shifts were used to identify potential peptides that were ubiquitinated. Tryptic and monoubiquitination experiments using

K0 ubiquitin allowed for *a priori* knowledge of the ubiquitinated peptides. Mono and di-ubiqutinated UbcH5b at Lys128 were both observed in the above spectrum.

Conclusions. Ubiquitin is vital protein modifier that participates in many cellular processes that are not yet fully understood. As demonstrated with Lys48 linked ubiquitin and proteasomal degradation, the length as well as the linkage pattern of the modifying ubiquitin moiety play crucial roles in determining a substrate protein's fate.[95] Despite a significant contribution to the ubiquitin field thus far, current tryptic proteomic methods have been unable to discern anything more than the site of modification on a substrate protein. Presented is an Asp specific hydrolysis method that can effectively measure the length of a Lys63 linked ubiquitin moiety, and identify the site of ubiquitination provided the substrate protein is known. Although Asp specific hydrolysis is well suited for ubiquitinated single proteins and simple mixtures on its own, parallel digestion with trypsin for confirmation of the ubiquitination site simplifies manual assignment. Manual assignment is further simplified by the presence of peaks correlating to 0 and 1 missed cleavages on each monomer unit of the ubiquitin chain. Here the two digestion methods were used in parallel to identify two ubiquitination sites following conjugation *in vitro*. Using the above outlined method, there is the potential for analysis of polyubiquitin isolated directly from lysates after enzymatic removal from their substrates, which could provide an unprecedented view of relevant polyubiquitin lengths in vivo.

Chapter 5: Extracellular Vesicle Analysis with Microwave-supported Acid Hydrolysis

Introduction

Extracellular vesicle (ECV) describes several types of extracellular organelle. Apoptotic blebs, exosomes, and shedding microvesicles are all included in this umbrella term. Generally, the three different subtypes listed are defined by their size and morphology.[96] In all cases, the biomolecules (whether protein, nucleic acid, carbohydrate, or lipid)[97] contained in the ECVs are apportioned there by the parent cell, meaning that all molecules found inside an ECV are also found in parental cells. This complexity make extracellular vesicles an excellent model system for further evaluation of the optimized methods outlined in the previous chapters. Previously outlined examples of model systems to evaluate our methods have done so on chemically unique samples (ribosomes or poly-ubiquitins) or without observing the combined effects of multiple methods (in gel acid hydrolysis or mass biased partitioning of a whole cell lysate). The complexity and communal interest in extracellular vesicles make them an excellent model system for evaluating a microwave-supported acid hydrolysis driven analysis. Studies are underway around the world to ascertain how closely the proteins in exosomes and the other ECVs reflect proteins in the parent cells, and if there are any signature classes of proteins

commonly excreted. Presented here is a proteomic analysis of ECVs shed by myeloid derived suppressor cells using microwave-supported acid hydrolysis both in gel and in solution, according to the optimized methods outlined in Chapters 2 and 4.

Materials and Methods

Extracellular Vesicle Protein Preparation. Extracellular vesicles were provided by the laboratory of Dr. Suzanne Ostrand-Rosenberg (University of Maryland, Baltimore County). Myeloid derived suppressor cells (MDSC) were isolated from BALB/c mice 3-4 weeks post inoculation with 4T1 mammary carcinoma cells that were transfected to stably express Interleukin 1ß (IL-1ß). Following isolation, the MDSC were cultured overnight and intact cells were removed. Extracellular vesicles (ECVs) were isolated from the supernatant using sucrose density gradient centrifugation. Fractions correlating to densities between 1.1 and 1.2g/mL were isolated and resuspended in phosphate buffered saline (PBS). After being transferred to our laboratory, the isolated ECVs were solubilized at 60°C in Laemmli buffer. Proteins were precipitated using chloroform and methanol,[58] and then resuspended in either gel loading buffer for in gel digestion or in 12.5% acetic acid for mass biased partitioning experiments.

In Gel Microwave-supported Acid Hydrolysis. For in gel digestion, four aliquots, each containing 45µg of protein from ECVs were fractionated on an 8-16% gradient Tris-glycine gel (Figure 32).

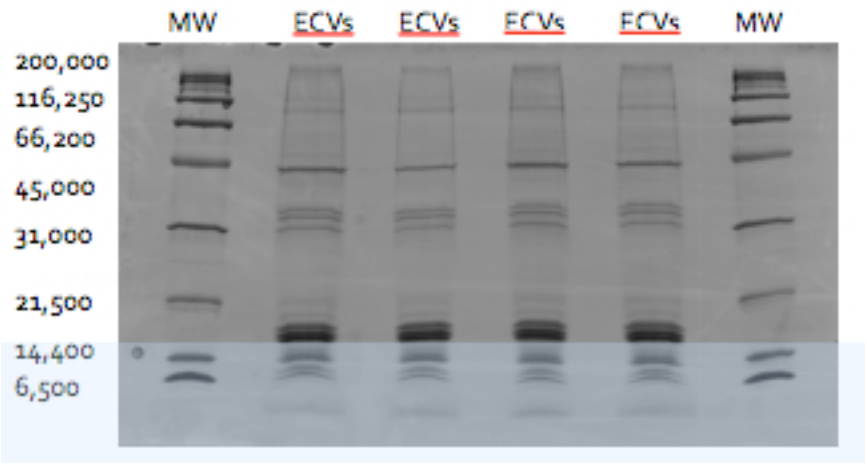

Figure 32. SDS-PAGE gel of fracationated ECVs.

Following electrophoretic fractionation, eight bands per well were excised and destained according to the optimized methods outlined previously.[59] Briefly, gel pieces were pooled and destained in 50/50 ACN/water for 30 minute intervals until the blue color of the coomassie stain was gone or significantly reduced. The pooled gel pieces were then macerated and dehydrated in ACN, and subsequently rehydrated in 12.5% acetic acid. Prior to digestion, additional solvent was added to ensure the gel pieces were covered during the digestion. Gel pieces were digested for 30 minutes at 140°C while using 300W of microwave energy to maintain temperature in a CEM Discover microwave. Peptides were extracted using a 2:1 solution of ACN: 5% formic acid while vortexing for 30 minutes. The peptide containing extract was lyophilized to near dryness and resuspended in 0.1% formic acid prior to LC-MS/MS.

Mass Biased Partitioning and Lys63 Immunoprecipitation. Extra cellular vesicles in PBS were precipitated as above and resuspended in the appropriate amount of 12.5% acetic acid to achieve a final protein concentration of 0.1mg/mL.

This ensured that all proteins were efficiently solubilized prior to digestion. Protein was then digested as above. Following digestion, Millipore 3kDa MWCO filters (Billerica, MA) were equilibrated in 12.5% acetic acid prior to use. The digest was then fractionated by size using the 3kDa MWCO filters according to the manufacturer's instructions. The higher molecular weight retentate was first immunoprecipitated (IP) using a Lys63 ubiquitin specific antibody (Millipore) coupled to an agarose solid support (ThermoPierce, Rockford, IL) and recovered according to the manufacturer's instructions. This specific antibody was selected, in part, because it was formulated against a short branched peptide immunogen rather than intact di-ubiquitin. The smaller branched peptide ensures that epitope binding residues must be contained within the sequence present following Asp specific hydrolysis of a ubiquitinated substrate. The flow through of the IP reaction was analyzed directly by LC-MS/MS, and the lower mass MWCO filtrate was lyophilized to near dryness.

LC-MS/MS. The peptides resulting from in gel digestion were injected into an LTQ-Orbitrap XL (ThermoFisher, San Jose, CA) via a Shimadzu Prominence (Columbia, MD) autosampler and LC. Following injection, peptides were loaded and concentrated onto a C18 trap column and subsequently fractionated using a C18 150mm x .150mm GraceVydac (Deerfield, IL) analytical column. Peptides were fractionated using a linear gradient from 0-35% solvent B (97.5% ACN/2.5% water/0.1% formic acid) over the course of ~3 hours. The low mass peptides resulting from mass biased partitioning after in solution digestion were analyzed in

triplicate in the same manner. Higher mass and IP recovered peptides were analyzed with high resolution precursor and product ion detection on either the above mentioned C18 analytical column or a 100mm x .100mm C8 column. Survey scans for the high mass flow through after the IP reaction were performed at 30,000 R_p and product ion scans at 7,500. Data dependent parameters were adjusted such that only precursors with charge states of +4 or greater were selected for fragmentation. Immunoprecipitated Lys63 containing peptides were analyzed with identical data dependent parameters, but the precursor R_P was set to 100,000.

Bioinformatics. Low resolution fragment ion data resulting from both in gel acid hydrolysis and the low mass fraction after mass biased partitioning were searched using PepArML[60] and the IPI Mouse database. Results from six different search engines were combined and peptide specific estimated FDRs were generated. Dehydration of Asp, pyro-glutamate formation from Gln and Glu, and Met oxidation were chosen as variable modifications. The high mass fraction after mass biased partitioning was searched against a database of reviewed mouse proteins from the UniProtKB using ProSightPC, (ThermoFisher). A precursor mass tolerance of 250Da was set to accommodate dual terminal specificity of acid digestion. A more stringent Evalue threshold of 1×10^{-8} was set based on previously calculated FDRs from the optimized workflow for the high mass peptides. Later searches for specific PTMs were performed using Mascot, and Evalues better than 0.05 warranted manual inspection of the mass spectrum. Peptide matches below this threshold were not considered. Additionally, raw data files were interpreted manually for assignment of

ubiquitin modification. Protein functional and gene ontology classification was performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID).[98]

Results and Discussion

In Gel Microwave-supported Acid Hydrolysis. Based on the gel shown in **Figure 32**, one would assume that the protein complement of ECVs from IL-1 beta expressing MDSCs is a relatively simple mixture. Despite the amount of information that can be deduced from the Coomassie-stained gel image, in gel digestion coupled with mass spectrometry is more sensitive. In gel digestion using microwave-supported acid hydrolysis resulted in identification of 76 unique proteins from 96 peptides that demonstrated the expected speficity. Additional peptides were identified using search parameters that accommodated non-specific hydrolysis at either or both termini. Consistent with the gel image, high sequence coverage was achieved for three histone proteins, H4, H3.2, and H3.3, whose molecular masses around 15kDa correlate to the darkest bands on the gel.

Mass Biased Partitioning and Ubiquitin. Interrogation of the ECV sample using mass biased partitioning resulted in a more informative analysis. The low mass fraction provided 202 total peptides, and an additional 20 peptides were identified in the high mass fraction. Consistent with previous results obtained in the methods

development of mass biased partitioning outlined in Chapter 2, there are distinct differences in the peptide mass distributions of each fraction (**Figure 33**).

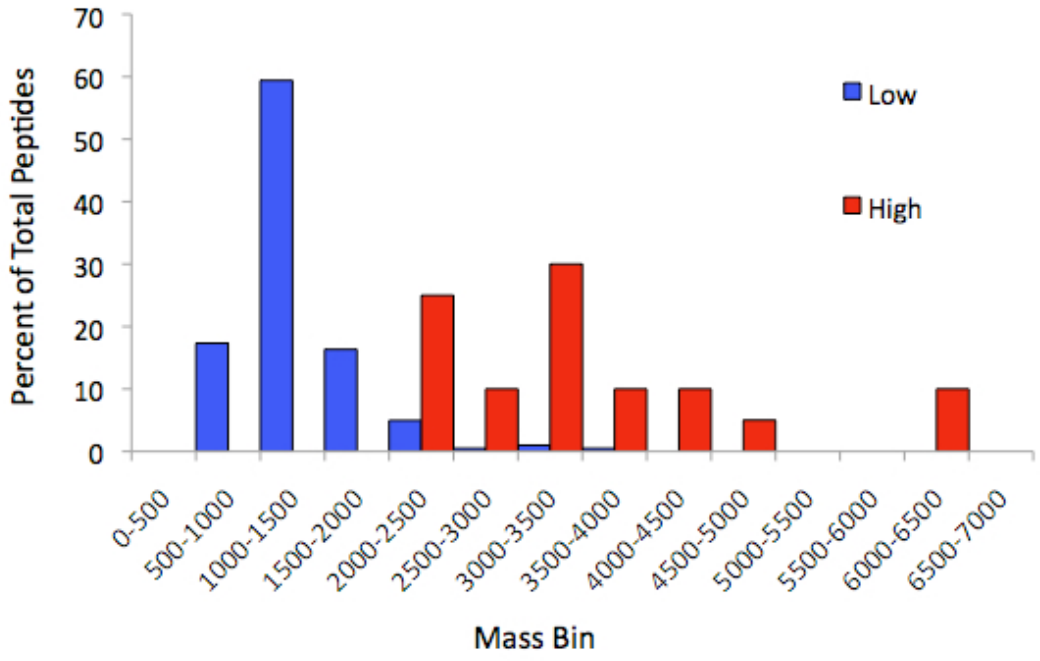

Figure 33. Peptide mass distributions of low (blue) and high (red) mass fractions resulting from mass biased partitioning of ECVs after microwave-supported acid hydrolysis.

The large peptides allowed unambiguous identification of two proteins with high sequence similarity. Peptide [91-130] is specific to both mouse histones H2A type 1 and H2 type 2-A, and differs in the two peptides only by a single Lys to Arg substitution at position 100 (**Figure 34**). Analysis of the large (>4kDa) peptides resulted in identification of both peptides. Additional evidence for the presence of both proteins was found in the in gel analysis, based on peptide [44-72] with a separate Leu to Met substitution at position 51.

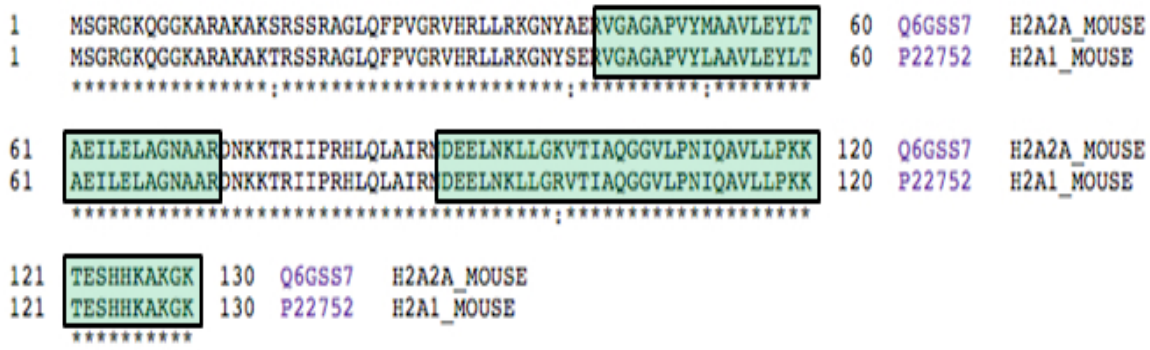

Figure 34. Sequence alignment of mouse histones H2A type 1 and H2A type 2. The four residue differences are indicated at their specific locations by two dots below the aligned sequences. Peptides observed are indicated by shading.

Proteolytic methods alternative to trypsin are important in studies of histone sequence and modification.[99] For one thing, many modifications occur on Lys and Arg, which would interfere with proteolysis. In addition, the frequency of Lys and Arg produces short peptides which contribute little to a given protein identification. *In silico* tryptic digestion of the two proteins assuming no missed cleavages results in a four-residue peptide containing the observed site of amino acid substitution (LLGK or LLGR). The two tryptic peptides have neutral masses of 429.30Da (for the Lys containing sequence) and 457.30Da. In most ESI analyses, tryptic peptides will acquire two charges resulting in charge densities for the two tetrapeptides of 215.65 (for the Lys containing sequence) and 229.65. In most cases, ESI analyses only detect ions in the range of m/z 400-2000 to avoid superfluous detection of uninformative peptide sequences and avoid competition from ambient small molecules during the ionization process.

Additional analysis of the low mass fraction was targeted to look for post-translational modifications and revealed two separate monoacetylations on both Lys27 and Lys29 of ubiquitin (**Figure 35** and **Figure 36**).

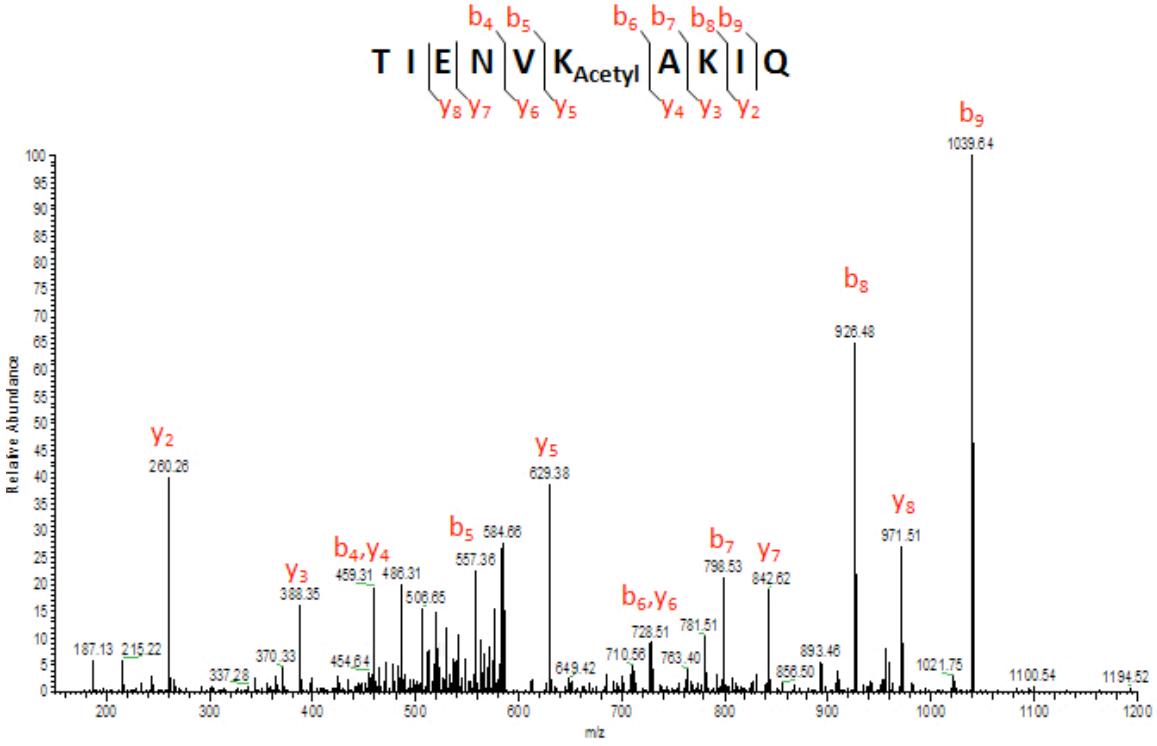

Figure 35. Product ion spectrum demonstrating the site of acetylation on Lys27 of ubiquitin.

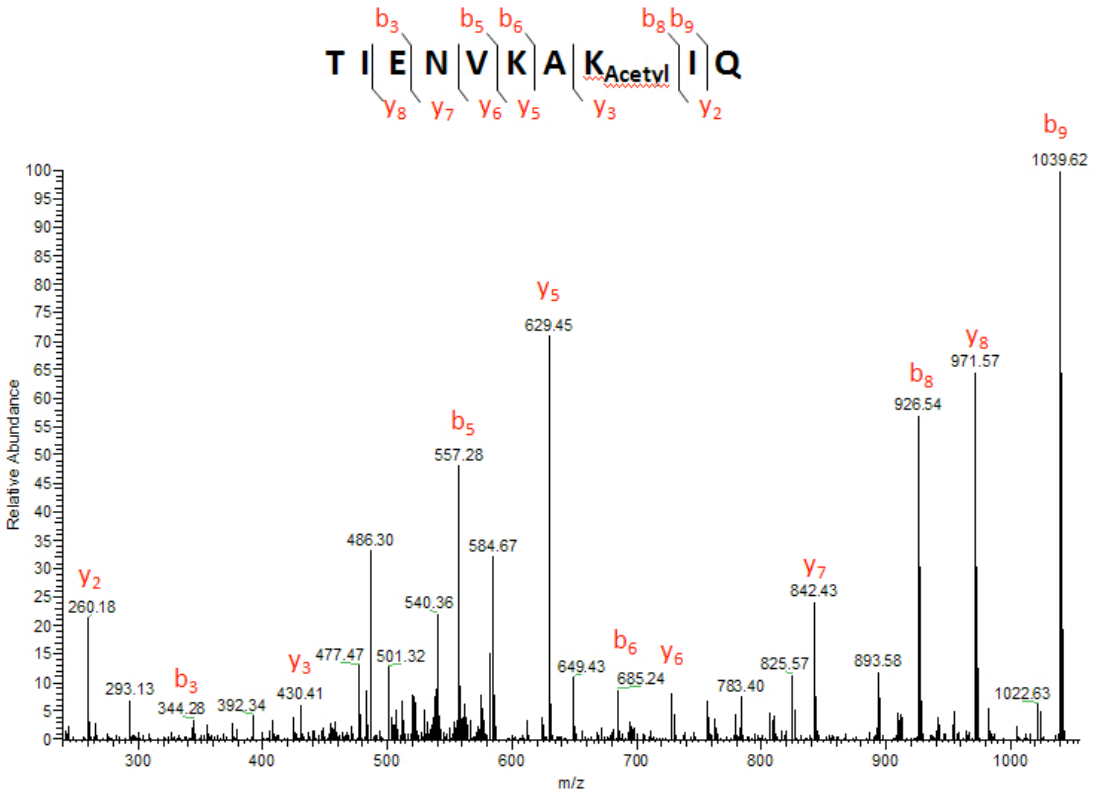

Figure 36. Product ion spectrum demonstrating the site of acetylation on Lys29 of ubiquitin.

Additionally, N terminal peptides from ferritin heavy chain, S100A9, and isocitrate dehydrogenase were identified with post translationally acetylated N termini following removal of the initiator Met residue. An important factor that aids in assignment of these PTMs is the observation of several additional high confidence peptides from the same proteins. Proteins identified with only a single post translationally modified peptide were confirmed by manual inspection. For ubiquitin, all peptides aside from the most C terminal peptide were observed. This is consistent with the biochemistry associated with ubiquitin, in which the C terminal end is conjugated to a substrate protein.

Evidence from both in gel microwave-supported acid hydrolysis and mass biased partitioning indicates protein ubiquitination. Ubiquitin peptides were observed in five of the eight excised gel bands, and in the low mass fraction recovered from digestion in solution the most abundant peptide (as determined by LC peak area integration of extracted ion chromatograms) was [40-51] assigned to ubiquitin (**Figure 37**).

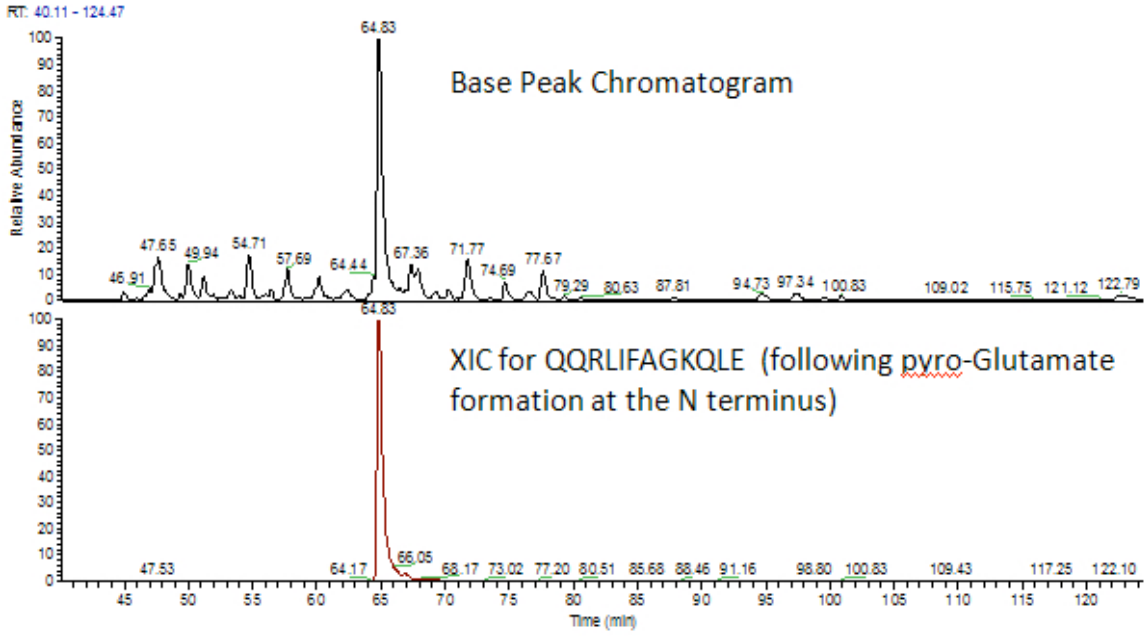

Figure 37. Base peak chromaogram (top) and extracted ion chromatograms (XIC) for peptide [40-51] of ubiquitin.

Taken together, this evidence spurred further investigation. Ubiquitination has been proposed to sort proteins into both exosomes and the proteoasome.[9] Manual inspection of the LC-MS/MS data file acquired from injection of the IP eluate was used to assign ubiquitination sites on substrate proteins. Similar to assignment of the two acetylated peptides on ubiquitin, a prerequisite was that the modified peptide was not the only one observed from a given protein. Analysis of the relatively few peaks

observed from the IP eluate revealed one peptide in common to both mouse histone H3.3 and H3.2. A methylated peptide [107-136] was found to match within 2ppm to the predicted mass of the ubiquitinated sequence at Lys116 (**Figure 38**).

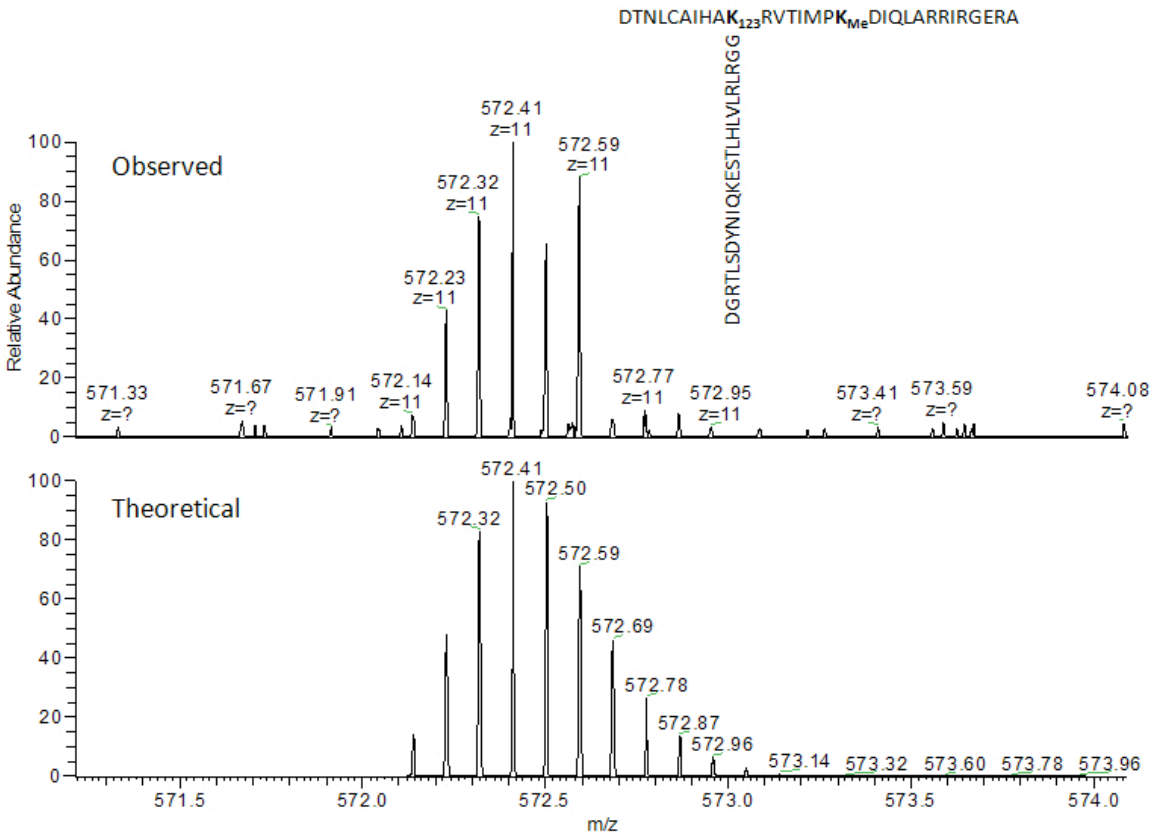

Figure 38. (A) Observed (top) and theoretical (bottom) isotope clusters of the ubiquitinated (at Lys123) and methylated (at Lys116) peptide [107-136] from histone H3.

Histone H3 is known to be monoubiquitinated in order for somatic recombination to occur.[100] Somatic recombination is a process in which variable, diverse, and joining (V(D)J) gene segments are rearranged and recombined to facilitate diverse immunoglobulin production.[101] Grazini and co-workers hypothesize that recombinationally active gene loci are favored for repair of the DNA breaks via

ubiquitination at specific sites.[100] Removal of this activity results in mice that are severe combined immunodeficient due to lack of fully mature B and T lympocytes.[102] Although it is not absolutely certain that Histone H3 carries only mono-ubiquitin at only one site, the absence of observed ubiquitinated peptides from ubiquitin, itself, strengthens the conclusion. It is of interest that several enzymes that function directly in the protein ubiquitination cycle - E3 ubiquitin ligase DZIP3, ubiquitin conjugating enzyme E2N, ubiquitin like modifier activating enzyme 1 (UBA1), protein VBRBP[103] and four different sequences present in the 26S proteasome were identified as well. Included also is a protein implicated directly in hydrolysis of Lys63 linked polyubiquitin chains, Josephin-2 (see **Table 2**). The presence of these proteins in conjunction with the high abundance of peptides from ubiquitin, itself, all support the hypothesis that ubiquitination is an important step in assigning proteins for extracellular vesicle export.[95]

Both the in gel and mass biased partitioning experiments using microwave-supported acid hydrolysis resulted in a total of 398 distinct peptides representing 240 proteins (Appendix Table 6). The higher mass peptides provide significantly increased sequence coverage. The average sequence coverage achieved from the high mass fraction (~22.5%) is more than 6 fold greater than that achieved from the low mass fraction (~3.6%). Combining the results from both the in gel and in solution experiments further illustrates this point. Even higher average sequence coverage was achieved for proteins that were observed in the high mass portion in addition to either the in gel or low mass portion following mass biased partitioning (~34%). All of these results are consistent with the simple reality that longer peptides provide

greater sequence coverage. For this sample the increase in sequence coverage allowed identification of two distinct isoforms of histone H2, and allowed probable identification of a specific ubiquitination site on histone H3, which is known to have significant biological implications.

Combining protein identifications from all peptides displaying the expected specificity provides a mixture with sufficient complexity for functional analysis. Using the DAVID gene ontology software, all genes represented by a single UniProtKB accession number (at one gene per accession number) are compared to a background gene list composed of the mouse genome. When compared to the background, the gene list generated from all of the proteins identified in the extracellular vesicle experiments demonstrates significant enrichment in certain functional gene ontologies (GOs). The GO categories are clustered to reduce redundancy. Specifically, based on annotation in the SwissProt database, 71 of the 240 interrogated genes were implicated in nucleotide binding demonstrating high enrichment with a P value of 9.0×10^{-20} (where a P value <0.01 is indicative of greater enrichment than would be observed by random chance alone). Additional enrichment was observed for gene products involved in ubiquitin-like modification (23 gene products resulting in a P value of 1.8×10^{-6}). As a final test of the methods efficacy, the comprehensive acid hydrolysis workflow was compared to an in gel tryptic analysis (which has been shown to be the most effective method for comprehensive analysis).[56] The results are summarized in **Error! Reference source not found.**. Importantly, it is shown that the although the total number of identifications are lower, the average sequence coverage per protein is greater than

two-fold higher. Taken together, these data suggest that the protein component of extracellular vesicles shed by MDSC from mammary carcinoma bearing mice is highly enriched in proteins involved in ubiquitination, and it is likely that many of which originate from the nucleus and are involved in DNA/RNA maintenance.

Table 1. Average sequence coverage and number of unique and shared proteins observed using the comprehensive microwave-supported acid hydrolysis workflow in comparison to in gel tryptic digestion of the same sample.

	Acid Hydrolysis	Trypsin
Unique Proteins	134	329
Average Sequence Coverage	8.56	3.67
Shared Proteins	125	

Table 2. Sequence coverage achieved using in gel microwave-supported acid hydrolysis and mass biased partitioning from ECVs.

Protein Description	**Sequence Coverage**
1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma-2	0.87%
14-3-3 protein epsilon	9.02%
26S protease regulatory subunit 10B	2.31%

26S proteasome non-ATPase regulatory subunit 2	3.08%
40S ribosomal protein S12	11.36%
40S ribosomal protein S16	15.75%
40S ribosomal protein S23	18.88%
40S ribosomal protein SA	5.42%
6-phosphofructokinase, muscle type	3.21%
6-phosphogluconate dehydrogenase, decarboxylating	2.48%
60 kDa heat shock protein, mitochondrial	6.28%
60S acidic ribosomal protein P2	18.26%
60S ribosomal protein L30	10.43%
60S ribosomal protein L9	10.42%
78 kDa glucose-regulated protein	3.82%
A-kinase anchor protein 10, mitochondrial	1.21%
Actin-related protein 2/3 complex subunit 1B	3.49%
Actin-related protein 3B	2.39%
Actin, aortic smooth muscle	19.36%
Actin, cytoplasmic 1	31.47%
Activator of 90 kDa heat shock protein ATPase homolog 1	2.66%
Activin receptor type-2A	1.75%
Adenine phosphoribosyltransferase	6.67%
Adenylyl cyclase-associated protein 1	5.06%
ADP-ribosylation factor 3	5.52%
ADP-ribosylation factor 4	5.56%

Alanyl-tRNA synthetase, cytoplasmic	1.55%
Aldehyde dehydrogenase family 16 member A1	1.87%
Aldehyde dehydrogenase, mitochondrial	2.31%
Alpha-enolase	16.13%
Annexin A11	2.78%
Aspartate aminotransferase, mitochondrial	2.56%
AT-rich interactive domain-containing protein 1A	1.05%
Ataxin-10	3.16%
ATP synthase subunit alpha, mitochondrial	2.17%
ATP synthase subunit beta, mitochondrial	5.29%
ATP-citrate synthase	2.47%
ATP-dependent RNA helicase DDX39	3.75%
Autophagy-related protein 2 homolog B	0.67%
Barrier-to-autointegration factor	15.73%
Calcium homeostasis endoplasmic reticulum protein	0.85%
Calmodulin	10.07%
Calreticulin	6.25%
Cathelin-related antimicrobial peptide	4.05%
CD9 antigen	11.06%
Chloride intracellular channel protein 1	6.22%
Chromodomain-helicase-DNA-binding protein 9	0.52%
Clathrin heavy chain 1	2.81%
Cofilin-1	11.45%

Coiled-coil domain-containing protein 151	3.04%
Coronin-1A	13.23%
Coronin-7	2.06%
Creatine kinase M-type	3.15%
Cytoplasmic dynein 1 heavy chain 1	0.52%
Cytoplasmic tyrosine-protein kinase BMX	3.07%
Cytosolic phospholipase A2 delta	4.12%
DNA repair protein REV1	2.48%
DNA replication licensing factor MCM6	3.05%
DNA topoisomerase 1	1.04%
Dynamin-2	2.41%
Dynein heavy chain 12, axonemal	0.68%
E3 ubiquitin-protein ligase DZIP3	1.00%
Elongation factor 1-alpha 1	25.76%
Elongation factor 1-gamma	2.97%
Endoplasmin	1.50%
Eukaryotic initiation factor 4A-II	4.42%
Eukaryotic translation initiation factor 3 subunit E	2.92%
Eukaryotic translation initiation factor 3 subunit G	4.06%
Eukaryotic translation initiation factor 3 subunit L	2.30%
Eukaryotic translation initiation factor 4H	6.85%
Exportin-1	0.93%
Fatty acid synthase	0.88%

FH2 domain-containing protein 1	0.96%
Fibrinogen beta chain	2.70%
Fibrinogen gamma chain	2.29%
Fibronectin	2.02%
Filamin-A	3.10%
Filamin-C	0.44%
Forkhead box protein J2	4.60%
Fructose-bisphosphate aldolase A	4.95%
Gamma-enolase	5.07%
Gasdermin-C	4.49%
Gelsolin	1.54%
Glucose-6-phosphate 1-dehydrogenase X	6.02%
Glucose-6-phosphate isomerase	13.44%
Glutamate dehydrogenase 1, mitochondrial	1.61%
Glutamate--cysteine ligase regulatory subunit	2.92%
Glyceraldehyde-3-phosphate dehydrogenase	30.63%
Glycogen phosphorylase, liver form	2.00%
GTP-binding nuclear protein Ran	12.50%
GTP-binding protein 1	2.40%
GTP-binding protein SAR1a	7.58%
GTP-binding protein SAR1b	7.58%
Guanine deaminase	4.63%
Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	4.41%

Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	4.41%
Guanylate cyclase soluble subunit beta-1	1.77%
H-2 class I histocompatibility antigen, alpha chain (Fragment)	3.36%
Heat shock 70 kDa protein 1-like	5.93%
Heat shock cognate 71 kDa protein	10.84%
Heterogeneous nuclear ribonucleoprotein A/B	3.86%
Heterogeneous nuclear ribonucleoprotein A1	5.31%
Heterogeneous nuclear ribonucleoprotein F	3.37%
Heterogeneous nuclear ribonucleoprotein K	3.02%
Heterogeneous nuclear ribonucleoprotein L	10.24%
Heterogeneous nuclear ribonucleoproteins A2/B1	8.22%
Hexokinase-3	1.08%
Histone chaperone ASF1A	8.82%
Histone H2A type 1	53.08%
Histone H2A type 2-A	53.08%
Histone H2A.Z	24.22%
Histone H2B type 1-A	12.60%
Histone H2B type 1-M	58.73%
Histone H2B type 3-B	79.37%
Histone H3.2	39.71%
Histone H3.3	39.71%
Histone H4	45.63%
Huntingtin	0.55%

Importin subunit beta-1	2.40%
Integrator complex subunit 3	1.83%
Interleukin enhancer-binding factor 2	5.64%
Josephin-2	7.98%
Kelch domain-containing protein 2	3.45%
Keratin, type I cytoskeletal 10	2.11%
Keratin, type I cytoskeletal 24	1.56%
Keratin, type II cytoskeletal 1	6.12%
Keratin, type II cytoskeletal 6A	6.33%
Kinesin-like protein KIF9	1.52%
Lactotransferrin	3.96%
Leukotriene-B(4) omega-hydroxylase 2	1.53%
Lipoxygenase homology domain-containing protein 1	0.87%
Lysozyme C-2	6.76%
Macrophage mannose receptor 1	1.03%
Macrophage migration inhibitory factor	19.13%
Major vault protein	1.39%
Metastasis-associated protein MTA3	2.71%
Microtubule-associated protein 2	0.55%
Microtubule-associated serine/threonine-protein kinase 2	1.33%
MMS19 nucleotide excision repair protein homolog	3.30%
Moesin	1.73%
Molybdenum cofactor biosynthesis protein 1	2.52%

Murinoglobulin-1	1.22%
Myosin light polypeptide 6	20.53%
Myosin regulatory light chain 12B	5.81%
Myosin-9	1.28%
Myosin-Ig	1.46%
NAD-dependent malic enzyme, mitochondrial	3.06%
NEDD4-binding protein 1	1.12%
NEDD9-interacting protein with calponin homology and LIM domains	2.00%
Neutral alpha-glucosidase AB	1.38%
Neutrophil elastase	3.40%
Neutrophil gelatinase-associated lipocalin	9.00%
NHL repeat-containing protein 2	1.66%
Nucleolin	3.11%
Nucleoside diphosphate kinase B	26.97%
Olfactomedin-4	2.18%
One cut domain family member 3	4.69%
Peptidyl-prolyl cis-trans isomerase B	6.94%
Peptidyl-prolyl cis-trans isomerase G	1.20%
Peroxiredoxin-2	14.14%
Phosphoglycerate kinase 1	5.52%
Plastin-2	2.39%
Pleckstrin	4.57%
Polyubiquitin-B	17.05%

Proteasome subunit alpha type-5	10.79%
Proteasome subunit beta type-4	3.41%
Proteasome subunit beta type-8	3.62%
Protein CDV3	2.85%
Protein diaphanous homolog 2	1.00%
Protein disulfide-isomerase A3	3.17%
Protein disulfide-isomerase A4	3.76%
Protein Dok-7	2.98%
Protein MEMO1	4.38%
Protein phosphatase 1G	1.66%
Protein S100-A6	7.87%
Protein tyrosine phosphatase domain-containing protein 1	2.14%
Protein VPRBP	1.26%
Prothymosin alpha	24.32%
Proto-oncogene vav	1.30%
Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	1.38%
Pyruvate kinase isozymes M1/M2	19.77%
Rab GDP dissociation inhibitor beta	2.02%
Ras GTPase-activating-like protein IQGAP1	0.72%
Ras-related C3 botulinum toxin substrate 2	5.21%
Ras-related protein Rab-7a	4.83%
Ras-related protein Rap-1A	5.98%
Ras-related protein Rap-2b	6.01%

Receptor-type tyrosine-protein phosphatase U	1.66%
Reticulon-4 receptor	2.96%
Retinoic acid early-inducible protein 1-alpha	4.35%
Rho-related GTP-binding protein RhoB	4.59%
Rho-related GTP-binding protein RhoG	5.24%
SAFB-like transcription modulator	0.87%
Scavenger receptor cysteine-rich domain-containing protein LOC284297 homolog	0.88%
Selection and upkeep of intraepithelial T-cells protein 5	1.44%
Serine/threonine-protein kinase 38-like	1.72%
Serine/threonine-protein kinase D1	1.85%
Seryl-tRNA synthetase, cytoplasmic	3.32%
Sister chromatid cohesion protein PDS5 homolog B	1.87%
Small nuclear ribonucleoprotein Sm D1	9.24%
Small nuclear ribonucleoprotein Sm D2	11.86%
Stefin-3	9.28%
Stress-70 protein, mitochondrial	3.68%
Stress-induced-phosphoprotein 1	2.21%
Structural maintenance of chromosomes protein 3	1.07%
Succinyl-CoA ligase [GDP-forming] subunit beta, mitochondrial	4.85%
Synaptojanin-2	1.67%
T-complex protein 1 subunit alpha	1.44%
T-complex protein 1 subunit eta	3.31%

T-complex protein 1 subunit gamma	1.65%
T-complex protein 1 subunit theta	1.64%
Talin-1	2.24%
TAR DNA-binding protein 43	3.38%
Thrombospondin-1	0.77%
Tight junction protein ZO-2	1.46%
Transitional endoplasmic reticulum ATPase	1.36%
Transketolase	5.78%
Translocon-associated protein subunit beta	10.93%
Transportin-1	1.78%
Triosephosphate isomerase	5.22%
Tubulin alpha-1B chain	7.10%
Tubulin beta-5 chain	11.04%
Tumor suppressor p53-binding protein 1	0.56%
Tyrosine-protein kinase Lyn	1.56%
Tyrosine-protein phosphatase non-receptor type 6	3.70%
Tyrosyl-tRNA synthetase, cytoplasmic	2.46%
U6 snRNA-associated Sm-like protein LSm3	11.76%
Ubiquitin-conjugating enzyme E2 N	13.16%
Ubiquitin-like modifier-activating enzyme 1	3.69%
Uncharacterized protein C15orf39 homolog	1.27%
Uncharacterized protein C1orf190 homolog	3.77%
Uncharacterized protein C2orf54 homolog	1.99%

Uncharacterized protein KIAA1671	2.92%
UPF0704 protein C6orf165 homolog	1.45%
Valyl-tRNA synthetase	0.79%
Vinculin	0.75%
WD repeat-containing protein 1	4.95%
Zinc finger and BTB domain-containing protein 7C	2.26%
Zinc finger protein 37	1.85%
Zinc finger protein 638	1.02%

Chapter 6: Conclusions and Prospectus

Microwave-supported acid hydrolysis has previously been shown to be a viable method for analysis of single proteins and samples of low complexity.[35, 43] Presented in this work, were several experiments that further developed this digestion method and the associated sample preparation, MS, and bioinformatic techniques required for successful analysis. First, two modifications on sample preparation both before and after microwave-supported acid hydrolysis were implemented. Specifically, microwave-supported acid hydrolysis was extended for in gel digestion of a whole cell lysate, resulting in identification of over 600 proteins. Although in gel digestion with acid hydrolysis is not new, it has previously only been used in well resolved gel bands and spots consisting of single or only a few proteins. Next, mass biased partitioning was used to enhance analysis of peptides produced from acid digestion by fractionating them into two distinct populations based on their size. The 3kDa molecular weight cutoff filters conveniently bifurcate the sample into aliquots with specific MS analysis parameters. The high mass parameters used were then shown to be applicable to successful analysis of the human ribosome. Proteins that

exhibit a high incidence of basic amino acids in their primary sequence in order to bind the nucleic acid phosphate backbone are poorly suited for tryptic digestion. Microwave-supported acid hydrolysis resulted in identifcation of 70 of the 79 predicted ribosomal proteins (~89%).[22] Average sequence coverage provided by peptides greater than (n = 161 peptides) and less than 3kDa (n = 221 peptides) was ~36% and ~21%, respectively. Next a novel application for ubiquitinated proteins and Lys63 polyubiquitin chains was demonstrated. Examination of ubiquitin's primary sequence demonstrates that Asp specific hydrolysis will reveal the site of substrate ubiquitination. Additionally, for examination of polyubiquitins joined exclusively via Lys63 linkages, it will produce a single peptide composed of the C terminal peptides of all constituent monomers, the mass of which is directly indicative of polyubiquitin length. Finally, extracellular vesicles isolated from mammary carcinoma bearing mice were examined utilizing all of the above outlined techniques. Combining in gel digestion, mass biased partitioning, and potential ubiquitination analysis identified 389 peptides representing 240 proteins. Far more important than a numerical total of proteins and peptides identified, the microwave-supported acid hydrolysis workflows were able to provide information sufficient for drawing testable biological hypotheses – the main goal for discovery type high throughput proteomic experiments.

The work outlined in this dissertation has demonstrated that microwave-supported acid hydrolysis is a viable workflow for rapid, high throughput analysis of complex proteomic samples both in gel and in solution. It also has specific benefits for the analysis of nucleotide binding proteins and Lys63 linked polyubiquitin and

polyubiquitinated proteins that cannot be achieved using standard tryptic methods. Currently, the proteomics community relies nearly exclusively on tryptic digestion. Thanks to bioinformatic and instrumental advancements in high resolution Orbitrap and TOF mass spectrometry the observable mass range is ever increasing. Microwave-supported acid hydrolysis has been shown to be an effective way of observing, in a high throughput manner, peptides that are longer on average, and potentially more informative than their tryptic counterparts.

Appendices

Appendix Table 1. Peptides identified with the expected cleavage specificity following in gel microwave-supported acid hydrolysis of multiple myeloma whole cell lysate. In rare cases (denoted by an asterisk) the confidently identified peptide was not mapped to a parent protein in the database.

UniProt	Sequence	Theoretical Mass	Observed Mass	Mass Diff	Estimated FDR
Q8WU90	LSLYIPR	860.51	860.51	0.00	0.040106306
Q9BX68	GNEVAKAQQATPGGAAPTIFSRIL	2396.28	2396.27	-0.01	0.035242291
Q9BX68	QQLLGHLLLVAKQTAKAEGLG	2170.25	2170.27	0.03	0.001593625
O95833	TVCAHFRQAPIP	1338.69	1338.65	-0.03	0.022452989
P62258	LVYQAKLAEQAERY	1680.88	1681.90	1.02	0.001593625
P62258	EMVESMKKVAGM	1320.62	1320.62	0.00	0.002276608
P62258	TLSEESYK	955.45	955.45	0.00	0.005398111
P62258	STLIMQLLR	1073.63	1073.62	-0.01	0.001593625
P26640	LAAVTALLLPFRYVL	1659.01	1660.01	1.00	0.002276608
P26640	LSVFYPGTLLETGH	1532.79	1532.79	0.00	0.042135504
P26640	EAKLQQTEAELRKV	1623.89	1623.89	0.00	0.039757576
P26640	EAIALFQKML	1144.63	1144.63	0.00	0
Q9Y4L1	LFERVPGPVQQALQSAEMSL	2199.14	2199.16	0.03	0.001593625
Q9Y4L1	EAAAMGAVYQAAALSKAFKVKPFVVR	2704.49	2706.48	1.99	0.003157895
P42765	LSEFAAKAALSAGKVSPETV	1975.06	1976.06	1.00	0.001593625
P42765	EHARPQTTLEQLQKLPPVFKK	2469.39	2469.36	-0.02	0.012337217
P42765	PSIMGIGPVPAISGALKKAGLSLK	2304.36	2304.35	-0.01	0.00757257

P42765	LVEVNEAFAPQYLAVERSL	2147.13	2148.11	0.98	0.001593625
P42772	DVAGYLRTATG	1104.56	1104.52	-0.03	0.026450742
Q14974	VVMASLLRMFQSTAGSGGVQE	2167.08	2168.09	1.02	0.001593625
Q14974	EVMQLLLENLGNENVHRSVKPQILSVFG	3144.67	3145.66	0.99	0.012337217
Q9UJW8	ACAGSTR	664.30	664.87	0.58	0.078844169
Q9Y6B6	ETIANVPILILGNKI	1588.95	1589.97	1.02	0.002276608
P52292	QIQQVVNHGLVPFLVSVLSKA	2258.28	2259.27	0.99	0.001593625
P34897	PKTGREIPYTFE	1436.73	1437.72	0.99	0.090652699
P34897	ALLERGYSLVSGGT	1421.75	1421.75	-0.01	0.026450742
Q9Y6Q1	DGHKVIMSLQQK	1382.73	1382.78	0.05	0.063943162
Q13541	EPPMEASQSHLRNSPE	1789.80	1790.81	1.01	0.001593625
P34932	PFVEAEKSNLAY	1366.68	1366.67	0.00	0.002276608
P34932	EAVTRGCALQCAILSPAFKVREFSIT	2791.45	2791.43	-0.02	0.063654574
P34932	PAIAQFSVQKVTPQS	1599.86	1599.86	-0.01	0.001593625
Q562R1	LYANTVLSGGSTMYPGIA	1829.89	1827.90	-1.99	0.021633931
P11021	QGNRITPSYVAFTPEGERLIG	2287.16	2288.18	1.02	0.001593625
P11021	AAKNQLTSNPENTVF	1632.81	1632.81	-0.01	0.001593625
P11021	IKFLPFKVVEKKTKPYIQV	2304.40	2304.42	0.02	0.006197432
P11021	AGTIAGLNVMRIINEPTAAAIAYGL	2499.35	2499.34	-0.01	0.001593625
P11021	LFRSTMKPVQKVLE	1674.95	1674.94	-0.01	0.00757257
P11021	EIVLVGGSTRIPKIQQLVKEFFNGKEPSRGINP	3632.02	3632.03	0.01	0.022452989
P11021	LTGIPPAPRGVPQIEVTFEI	2133.18	2134.19	1.01	0.002276608
P11021	VNGILRVTAE	1070.61	1070.60	-0.01	0.001593625
P11021	TRNELESYAYSLKNQIG	1984.99	1984.97	-0.02	0.001593625
P11021	KETMEKAVEEKIEWLESHQ	2343.14	2344.15	1.00	0.012337217
O43324	ILLYYGLHRFIV	1505.88	1505.88	0.00	0.022452989
Q16706	EGSFPQGQLSMLQEK	1693.80	1695.91	2.11	0.092548077
P52565	EHSVNYKPPAQKSIQEIQEL	2319.19	2320.19	1.00	0.001593625
P52565	ESLRKYKEALLGRVAVSA	1971.13	1972.12	1.00	0.012337217
P52566	ESLIKYKKTLLG	1373.83	1373.82	0.00	0.002276608
P11142	AAKNQVAMNPTNTVF	1604.80	1604.80	0.00	0.002276608
P11142	AGTIAGLNVLRIINEPTAAAIAYGL	2481.40	2481.40	0.00	0.001593625
P11142	KKVGAERNVLIF	1372.82	1372.81	-0.01	0.002276608
P11142	GIFEVKSTAG	1007.53	1007.52	0.00	0.001593625
P11142	PVEKALR	811.49	811.49	0.00	0.003157895
P11142	PVEKALRD	926.52	926.52	0.00	0.010505051
P11142	IVLVGGSTRIPKIQKLLQ	1962.24	1962.22	-0.01	0.002276608
P11142	IVLVGGSTRIPKIQKLLQD	2077.26	2078.28	1.01	0.004582951

P11142	FFNGKELNKSINP	1506.78	1506.78	0.00	0.016024653
P11142	VTPLSLGIETAGGVMTV LIKRNTTIPTKQTQTFTT YS	3967.13	3968.11	0.97	0.005398111
P11142	NQPGVLIQVYEGERAM TK	2032.04	2034.03	1.99	0.001593625
P11142	IERMVQEAEKYKAE	1722.86	1722.88	0.02	0.007002188
P11142	KVSSKNSLESYAFNMK ATVE	2232.11	2232.09	-0.02	0.012337217
P52597	PPLKFMSVQRPGPY	1615.85	1615.84	-0.01	0.001593625
P36578	KAAAAAAALQAKS	1170.67	1171.67	0.99	0.001593625
P28562	MVMEVGTL	894.42	894.53	0.11	0.079587629
P61088	ESNARYFHVVIAGPQ	1668.84	1668.81	-0.03	0.024608501
Q93009	PANYILHAVLVHSG	1489.80	1490.79	0.99	0.016024653
Q93009	PELAASGATLPKF	1300.70	1301.71	1.01	0.003157895
P01889	TELVETRPAG	1071.56	1071.55	0.00	0.028343211
P54136	LENPPLLVTPSQQAKFG	1837.99	1839.97	1.98	0.015837821
P38117	LVLLGKQAI	953.63	953.62	0.00	0.00757257
P38117	GGLETLRLKLPAVVTA	1636.99	1636.98	0.00	0.031793343
P38117	LVAKLKEIGRI	1238.81	1238.80	0.00	0.001593625
Q08945	PVEAFAQNVLSKA	1372.74	1372.72	-0.01	0.002276608
Q08945	EISFVNFARGTTTTRSF	1914.96	1914.95	-0.01	0.003775366
P36776	FQVTEEVKALTAEIVKTI R	2174.23	2175.23	1.00	0.024896266
O15160	EILAHRLGLIPIHA	1533.91	1533.90	-0.01	0.020577308
Q12905	ETSFSEALLKRNQ	1503.77	1503.77	0.00	0.004582951
Q12905	LAPNSAEQASILSLVTKI NNVI	2294.28	2295.29	1.01	0.001593625
P52815	LNELLKKTLKIQ	1439.91	1439.90	0.00	0.02661101
Q86U42	ARSIYVGNV	977.53	977.52	-0.01	0.040106306
Q9UHV9	EKPAAKENSEGAGAKA SSAGVLVS	2239.14	2240.14	1.00	0.087783172
Q9C0J9	DSRGGGSGGGPGGGAA AAAAALLGP	1950.95	1951.08	0.13	0.073445128
P82970	MPKRKAAGQGD	1139.59	1137.65	-1.94	0.004582951
O15355	AKLTTEEVIKELAQIAG RPTE	2296.26	2297.26	1.00	0.003157895
O15355	FIQSKISQR	1105.62	1105.62	0.00	0.001593625
O60641	AFAAPSPATTASPAKVD	1600.81	1600.83	0.02	0.006197432
Q9Y3C6	PTGTGRGGASIYGKQFE	1724.85	1724.85	0.00	0.02661101
P54577	VYRLSSVVTQH	1287.69	1287.69	0.00	0.010505051
P54577	EKWGGNKTYTAYV	1497.72	1497.72	0.00	0.059104342
P54577	LKNSVEVALNKLL	1439.87	1439.86	-0.01	0.012337217
P54577	PPAGSAPGEHVFVKGYE KGQP	2151.07	2152.07	0.99	0.012337217
Q9Y221	IPLGFGVAAKSTQ	1287.72	1287.72	0.00	0.002276608
Q9Y221	PMAIVVFHQA	1127.58	1127.57	-0.01	0.012337217
P60953	PSTIEKLAKNKQKPITPE TAEKLAR	2790.60	2791.60	1.00	0.042949177
Q9Y2H5	KVLIPERYIDLEPD	1698.92	1697.94	-0.98	0.078844169
O43447	GTGVASIYRGPFA	1294.67	1294.66	-0.01	0.015837821
Q9UKL4	PPESIGGPGGTGGGGSG	1963.90	1962.08	-1.82	0.034355828

	GGKRED				
P54652	GIFEVKSTAG	1007.53	1007.52	0.00	0.001593625
P54652	PVEKALR	811.49	811.49	0.00	0.003157895
P54652	PVEKALRD	926.52	926.52	0.00	0.010505051
P54652	IVLVGGSTRIPKIQKLLQ	1962.24	1962.22	-0.01	0.002276608
P54652	IVLVGGSTRIPKIQKLLQ D	2077.26	2078.28	1.01	0.004582951
P54652	FFNGKELNKSINP	1506.78	1506.78	0.00	0.016024653
P54652	EAVAYGAAVQAAILIGD	1612.85	1613.77	0.92	0.016024653
P68371	LQLERINVYYNEATGGK YVPRAVLV	2864.55	2865.56	1.00	0.012337217
P68371	SVRSGPFGQIFRP	1446.77	1446.76	-0.01	0.003775366
P68371	RIMNTFSVVPSPKVS	1660.90	1660.89	0.00	0.001593625
P68371	TVVEPYNATLSVHQLVE NT	2113.07	2115.05	1.98	0.001593625
P38646	IKNVPFKIVRASNG	1541.90	1541.89	-0.01	0.001593625
P38646	AGQISGLNVLRVINEPT AAALAYGL	2510.39	2510.38	0.00	0.001593625
P38646	ISILEIQKGVFEVKSTNG	1961.08	1961.08	0.00	0.003775366
P38646	QALLRHIVKEFKRETGV	2006.14	2006.14	0.00	0.003775366
P38646	ANGIVHVSAK	994.56	994.55	0.00	0.001593625
P38646	TETKMEEFK	1141.53	1141.53	0.00	0.054851348
Q01469	EYMKELGVGIALRKMG AMAKP	2274.20	2274.20	-0.01	0.039757576
P13284	ALQPPHEYVPWVTVNG KPLE	2273.18	2274.17	0.99	0.016024653
P54709	IIGLKPEGVPRI	1290.80	1290.83	0.03	0.02661101
Q01518	SLLAGPVAEYLKISKEIG G	1944.09	1945.06	0.96	0.012337217
Q01518	VVGIVEIINSK	1169.70	1169.70	-0.01	0.001593625
P46782	ELINAAKGSSNSYAIKK K	1903.05	1904.05	1.00	0.003157895
P46782	ELERVAKSNR	1182.65	1183.65	1.00	0.029044586
P11940	EAVAVLQAHQAKEAAQ KAVNSATGVPTV	2769.48	2769.47	-0.01	0
P05387	MRYVASYLLAALGGNS SPSAK	2171.10	2171.11	0.00	0
P05387	YVASYLLAALGGNSSPS AK	1867.97	1867.97	0.00	0.074812968
P05387	DRLNKVISELNGKNIE	1841.00	1843.00	2.00	0.020057307
P05387	RLNKVISELNGKNIE	1725.97	1725.98	0.00	0.001593625
P05387	VIAQGIGKLASVPAGGA VAVSAAPGSAAPAAGS APAAAEEKK	3667.99	3668.96	0.97	0.001593625
Q07666	PAARMSRSSGRSGSM	1536.72	1536.72	0.00	0.070288918
Q07666	PSFTHAMQLLTAEIEKIQ KG	2241.18	2241.18	0.00	0.001593625
Q07666	PKYAHLNM	988.48	988.48	0.00	0.062583818
P62736	EAQSKRGILTLKYPIEHG IITNW	2648.44	2648.43	-0.01	0.01536
P62736	EAQSKRGILTLKYPIEHG IITNWDD	2860.49	2861.49	1.00	0.074812968
O76021	SASASLSSAAATGTSTST PAAPTARKQL	2590.32	2591.32	1.00	0.001593625
P05455	SIESAKKFVETPGQKYK	2169.13	2170.12	0.99	0.055479138

	ET				
Q9NP55	ILKPGGGTSGGLLGGLL GKVTSVIPGLNNII	2914.72	2915.73	1.01	0.022133939
Q01851	DHISSPSLALMAGAGGA GAAAGGGGAH	2242.05	2242.11	0.06	0.055479138
P13645	LKNQILNLTT	1156.68	1156.68	-0.01	0.001593625
P13645	LTQLLNNMRSQYEQLA EQNRK	2576.31	2576.31	0.00	0.002276608
P13645	AEAWFNEKSKELTTEI	1894.93	1894.93	0.00	0.002276608
P13647	EINKRTTAENEFVMLKK	2032.08	2032.08	0.00	0.012721794
P13647	AAYMNKVELEAKV	1464.76	1464.76	0.00	0.012337217
P13647	AELSQMQTHVS	1229.57	1230.58	1.00	0.001593625
P13647	SIIAEVKAQYE	1249.66	1250.66	1.00	0.001593625
P13647	YQELMNTKLAL	1322.69	1322.69	0.00	0.004582951
O95467	DFETEPETAPTTEPETEP E	2129.88	2130.16	0.28	0.087783172
P13667	PPIPVAKI	833.54	833.55	0.02	0.003157895
P13667	PPIPVAKID	948.56	948.56	0.00	0.006197432
P13667	ATSASVLASRF	1108.59	1108.58	0.00	0.003775366
P13667	YEGSRTQEEIVAKVREV SQP	2304.17	2304.16	-0.01	0.026450742
P13667	YNGPREKYGIV	1294.67	1294.66	-0.01	0.019836639
P13667	YMIEQSGPPSKEILTLKQ VQEFLK	2805.50	2805.50	0.00	0.001593625
P13667	VIIIGVFKGES	1160.68	1160.68	0.00	0.001593625
P13667	KKNPVKFEGG	1102.61	1102.60	-0.01	0.010505051
P56537	VLKVEVFRQTVA	1387.82	1387.81	0.00	0.082482646
P56537	TTSTELSVVESVFKLNE AQPSTIATSMR	3025.53	3025.53	0.00	0.002276608
P31146	VRVSQTTW	975.51	975.51	0.00	0.060606061
P31146	VGTGAAMLTLGPEVHP	1548.80	1548.79	-0.01	0.001593625
P31146	PALTAEEWLGGR	1298.66	1298.67	0.00	0.001593625
P31146	AGPLLISLK	910.59	910.58	0.00	0.001593625
P31146	AVSRLEEEMRKLQATV QELQKRL	2754.52	2754.51	-0.01	0.012337217
P62829	NTGAKNLYIISVKGIKG RLNRLPAAGVG	2879.68	2880.69	1.00	0.052436195
P62829	LWPRIASNAGSIA	1354.74	1354.74	0.01	0.006197432
P31153	ALKEKVIKAVVPAKYL	1769.12	1770.11	0.99	0.016423358
P31153	TIYHLQPSGRFVIGGPQG	1926.01	1927.01	1.00	0.002276608
P31153	AGLTGRKIIV	1026.65	1026.65	-0.01	0.029101856
P31153	LRPGVIVR	908.59	908.59	0.00	0.001593625
Q9NZD4	MALLKANK	903.52	905.50	1.98	0.084638861
P83916	LIAEFLQSQKTAHET	1714.89	1715.90	1.01	0.003157895
P83916	PERIIGATD	970.51	971.54	1.03	0.093906094
Q9HCS4	MPQLGGGGGGGGGGSG GGGGSSAGAAGGG	2060.86	2060.19	-0.67	0.022452989
Q9HCS4	MPQLGGGGGGGGGGSG GGGGSSAGAAGGGD	2175.88	2177.13	1.25	0.012337217
O00178	AGKSTLLGVLTHGELD	1609.87	1609.89	0.03	0.073445128
O00178	DEGGPSGGPAVGAPPPG	1417.65	1418.84	1.20	0.028593311
P13796	EEMMELREAFAKV	1563.74	1563.74	0.00	0.008347245

P13796	EFIKIFHGLKST	1400.78	1400.77	-0.01	0.002276608
P13796	SKAYYHLLEQVAPKG	1702.90	1702.90	0.00	0.001593625
P13796	VVRGNPKLNLAFIANLF NRYPALHKPENQ	3332.83	3334.82	1.99	0.024896266
P13796	PKISTSLPVL	1053.64	1053.64	0.00	0.003775366
P48643	GYEQAARVAIEHL	1455.75	1455.74	-0.01	0.010505051
P23246	EMEKQQREQVEKNMK	1915.92	1915.95	0.02	0.032956222
P23246	AYHEHQANLLRQ	1478.74	1478.73	-0.01	0.01536
P31276	SAAESGIGGGGGGGGG GTGGAGGGCSGASPGK APSMD	2831.18	2829.69	-1.49	0.07975334
P07237	LAQQYGVRGYPTIKFFR NG	2214.17	2214.17	0.00	0.012337217
P07237	TASPKEYTAGREA	1379.67	1379.67	0.00	0.001593625
P07237	IVNWLKKRTGPAATTLP	1865.09	1866.10	1.01	0.019836639
P07237	GAAAESLVESSEVAVIG FFK	2010.03	2010.06	0.03	0.001593625
P07237	SAKQFLQAAEAI	1275.68	1275.67	-0.01	0.001593625
P07237	GVVLFKKF	936.58	936.57	-0.01	0.001593625
P07237	GVVLFKKFD	1051.61	1051.60	0.00	0.019836639
P07237	EGRNNFEGEVTKENLL	1829.89	1829.90	0.01	0.001593625
P07237	KQPVKVLVGKNFE	1484.87	1484.86	-0.01	0.001593625
P07237	HENIVIAKM	1069.56	1069.55	-0.01	0.002276608
P07237	STANEVEAVKVHSFPTL KFFPASA	2576.33	2576.32	-0.01	0.001593625
P07237	GFKKFLESGGQ	1196.62	1196.61	-0.01	0.001593625
O00299	DEEIELAYEQVAKALK	1847.95	1848.97	1.01	0
O00299	EEIELAYEQVAKALK	1714.91	1715.91	1.00	0
P56730	MTLARFVLALMLGALP EVVGFD	2362.28	2361.27	-1.00	0.017501509
P15374	AIRVTHETSAHEGQTEA PSI	2133.05	2133.90	0.86	0.052436195
O14979	LTEYLSRFGEVV	1411.73	1411.73	0.00	0.009079653
Q6P3S6	GSPILNGGSLSPGTAAV GGSSLD	1994.99	1994.11	-0.88	0.026450742
O43242	GKTAAAAAEHSQREL	1538.78	1538.77	-0.01	0.020057307
Q13158	FEAGAAAGAAPGEE	1246.55	1244.66	-1.89	0.068678915
P23526	IGLAAWGRKAL	1154.69	1154.69	0.00	0.019836639
P23528	KKNIILEEGKEILVG	1682.00	1682.00	0.00	0.001593625
P23528	PYATFVKMLP	1181.62	1181.61	0.00	0.001593625
P23528	PYATFVKMLPD	1280.65	1281.64	1.00	0
P23528	LVFIFWAPESAPLKSKMI YASSK	2628.40	2628.40	0.00	0.01536
P15531	SVESAEKEIGLWFHPEE LV	2198.09	2199.09	1.00	0.001593625
O43390	EAKIKALLERTGYTL	1686.97	1686.99	0.03	0.001593625
O43390	VTTGQRKYGGPPP	1356.71	1357.70	0.99	0.008347245
O43390	SVYSGVQPGIGTEVFVG KIPR	2189.18	2190.19	1.01	0.016423358
O43390	ELVPLFEKAGPIW	1479.81	1479.81	-0.01	0.002276608
P23588	TEQQSPTSGGGKVAPAQ PSEEGPGRK	2579.26	2579.25	-0.01	0.073160357
Q9H773	TAAPGRFSFSPEPTLE	1705.83	1704.87	-0.96	0.003157895

Q9H773	IRRLHAEFAAER	1467.81	1467.80	0.00	0.020057307
Q13200	KAPVQPQQSPAAAPGG T	1603.83	1603.82	-0.01	0.001593625
Q13200	EELRPLPVSVRVGQAV	1729.98	1730.01	0.03	0.009079653
P25100	DSSAGGSSAGGGGGSA GGAAPSEGPAVGGVPG GAGGGGGVVGAGSGE	3492.54	3490.82	-1.72	0.047058824
Q13310	ESLKELFSQFGKTLSVK VMR	2308.26	2310.25	1.99	0.049226442
P62917	PGRGAPLAKVVFR	1366.82	1366.81	-0.01	0.002276608
Q16630	YGSAIETLVTAISLIKQS KVSA	2278.28	2279.27	0.99	0.002276608
P60228	LVKVIQQESYTYK	1597.87	1598.86	0.99	0.001593625
Q16543	PLLEAVPKTG	1023.60	1023.59	-0.01	0.003775366
Q13405	QELKAWLLEKGF	1443.78	1444.78	1.00	0.002276608
P33240	QIAMLPPEQRQSILILKE QIQKSTGAP	2999.65	3000.65	1.01	0.016024653
P62269	LTKRAGELTE	1116.61	1116.61	0.00	0.002276608
P62269	GKYSQVLANGL	1148.62	1148.61	-0.01	0.002276608
P07814	PPLGALLAVEHVK	1342.80	1342.79	0.00	0.002276608
P33316	YTI PPMEKAVVKT	1475.81	1475.80	0.00	0.003775366
Q13526	ASFALRTGEMSGPVFT	1669.81	1670.81	1.00	0.004582951
O43707	PVTNLNNAFEVAEKYL	1820.93	1821.96	1.03	0.002276608
P31943	PPRKLMAMQRPGPY	1640.86	1640.86	-0.01	0.001593625
P31948	PTNMTYITNQAAVYFEK G	2046.97	2046.96	-0.01	0.001593625
P31948	AIHFYNKSLAEHRTP	1782.92	1782.91	-0.01	0.073160357
P31948	PAMRLILEQMQK	1456.79	1456.78	-0.01	0.001593625
P31948	PQALSEHLKNPVIAQKI QKLM	2385.36	2385.35	-0.01	0.002276608
P31949	GYNYTLSKTEFLSFMNT ELAAFTKNQK	3145.54	3147.54	2.00	0.001593625
P31949	SFLKAVPSQKRT	1360.78	1361.78	1.00	0.010505051
P25398	VNTALQEVLKTALIH	1648.95	1648.96	0.01	0.001593625
O96018	GETQPMTEVD	1087.45	1086.53	-0.92	0.04361223
Q15046	PMRQRQLFEEQAKAKA AG	2058.08	2058.08	0.00	0.026450742
P09417	QVTAEVGKLLGEEKV	1581.86	1581.86	0.00	0.001593625
Q99471	VGTGYYVEKTAE	1315.63	1315.63	0.00	0.002276608
Q05823	DGAGGASGLASPGC	1118.47	1117.60	-0.87	0.092548077
O95137	VPIRTVSSAASQGLHMQ ND	2010.00	2011.18	1.19	0.010505051
P09622	TKNILIATGSEVTPFPGIT I	2071.16	2072.17	1.01	0.001593625
P09622	VVAGPMLAHKAE	1221.65	1221.65	0.00	0.001593625
P51149	VTAPNTFKTL	1090.60	1090.61	0.00	0.024896266
P51149	PENFPFVVLGNKI	1472.80	1473.80	1.00	0.001593625
Q15365	AYSIQGQHTISPL	1413.73	1413.72	-0.01	0.005398111
Q9NSE4	KVASVASTLETTFETIST LSGV	2240.18	2241.18	1.00	0.001593625
P43243	GSASAAAKKKLKKV	1385.87	1385.87	0.00	0.028343211
P35232	ERVLPSITTEILKSVVAR F	2139.24	2140.23	0.99	0.001593625

Q15427	EKVSEPLLWELFLQAGP VVNTHMPK	2843.50	2844.52	1.02	0.001593625
P35244	LGLYNEAVKIIH	1368.78	1368.77	0.00	0.007002188
P35269	TGPQSLSGKSTPQPPSG KTTPNSG	2310.15	2310.18	0.04	0.001593625
Q15459	EPTSKKLKTE	1141.63	1141.63	0.00	0.020057307
P17844	VKFVINY	881.50	881.50	0.00	0.028593311
Q99714	PAEYAHLVQAIIENPFLN GEVIRL	2705.45	2706.45	1.00	0.001593625
P62316	RYISKMFLRG	1269.70	1269.70	0.00	0.021633931
P62316	SVIVVLRNPLIAGK	1477.93	1477.93	-0.01	0
O60506	GALAVLQQFK	1073.62	1073.62	-0.01	0.001593625
O60506	EAKIKALLERTGYTL	1686.97	1686.99	0.03	0.001593625
O60506	VTTGQRKYGGPPP	1356.71	1357.70	0.99	0.008347245
O60506	ELVPLFEKAGPIW	1479.81	1479.81	-0.01	0.002276608
P00505	PILGVTEAFKR	1229.71	1229.71	0.00	0.001593625
P00505	AGMQLQGYRYY	1364.62	1364.61	0.00	0.002276608
P00505	GRISVAGVTSSNVGYLA HAIHQVTK	2564.38	2565.39	1.00	0
P33991	LGSAQKGLQV	999.57	999.56	-0.01	0.022133939
P33991	IKKGILLQLFGGTRK	1671.06	1671.09	0.03	0.003775366
P33991	PETRQLVLQTGALVLS	1723.98	1724.03	0.04	0.012337217
P33991	FLTVTGKTVRLL	1346.83	1346.82	-0.01	0
P33992	FMPTILSRF	1110.59	1110.58	-0.01	0.007002188
Q99832	GATILKLL	827.55	827.54	-0.01	0.00757257
Q99832	VVHPAAKTLV	1033.63	1033.62	-0.01	0.003775366
Q99832	LLQLKMIGIKKVQGGAL E	1938.17	1940.17	2.00	0.028343211
Q99832	SVVAGGGAIEMELSKYL R	1878.99	1879.00	0.01	0.001593625
Q99865	EWRGMVLAQAPIMKA WFYITYEKD	2927.45	2926.42	-1.03	0.035242291
P43490	LRHLIVSRSTQAPLIIRP	2069.26	2071.25	1.99	0.070288918
Q96QK1	EQSLVGRFIHLLRSE	1764.96	1764.96	0.00	0.003157895
P78330	AFIGFGGNVIRQQVK	1632.91	1633.92	1.01	0.054851348
P35527	GGILTANEKSTMQELNS RLASYL	2495.27	2495.26	-0.01	0.001593625
P35527	KKGPAAIQKNYSPYYNT I	2055.08	2056.06	0.98	0.001593625
P35527	KKGPAAIQKNYSPYYNT IDD	2267.12	2269.12	2.00	0.078844169
P35527	LTVGNNKTLL	1071.63	1071.63	0.00	0.001593625
P35527	FRIKFEMEQNLRQGV	1893.99	1894.99	1.00	0.012337217
P35527	INGLRQVL	911.56	911.55	0.00	0.012721794
P35527	VNVEINVAPGK	1138.63	1138.63	0.00	0.001593625
P35527	VNVEINVAPGKD	1253.66	1254.64	0.98	0.001593625
P35527	MRQEYEQLIAKNRK	1805.96	1805.95	0.00	0.012337217
P35527	MRQEYEQLIAKNRKD	1920.98	1920.97	-0.01	0.001593625
P51571	ISIIPPLFTVSV	1284.77	1284.77	0.00	0.039757576
Q9NV70	GGTLSRQHNCGTPLPVS SEKD	2182.04	2180.58	-1.46	0.00757257
P35579	LIEKPAGPPGILALL	1500.93	1500.93	0.00	0.001593625

P78417	PTVSALLTSEK	1144.63	1144.63	0.00	0.016423358
Q9NR48	DAVIATASAPPSSSPGRSHSK	2022.01	2023.15	1.13	0.035242291
P27797	FGKFVLSSGKFYG	1435.75	1435.74	-0.01	0.001593625
P27797	DEFTHLYTLIVRP	1602.84	1603.86	1.01	0.001593625
P27797	EFTHLYTLIVRP	1469.80	1469.80	0.00	0.001593625
P27797	EFTHLYTLIVRPD	1584.83	1584.84	0.01	0
P27797	NTYEVKI	865.45	865.45	0.00	0.005398111
P27797	FLPPKKIK	969.64	969.63	0.00	0.04361223
P27797	DYKGTWIHPEI	1357.67	1358.65	0.99	0.016024653
P27797	YKGTWIHPEI	1242.64	1242.63	-0.01	0.003775366
P27797	LWQVKSGTIF	1177.65	1177.64	-0.01	0.001593625
P27797	EAYAEEFGNETWGVTKAAEKQMK	2598.21	2598.19	-0.02	0.002276608
Q9NY33	VQLLEYEASAAGLIRSFSERFPE	2611.33	2612.33	1.00	0.055479138
P51858	EMPEAAVKSTANKYQVFFFGTHETAFLGPK	3326.64	3326.64	0.00	0.017501509
P51858	EPAKEKNEKGALKRRAG	1863.04	1865.04	1.99	0.01536
P53396	LYFTYLEINPLVVTK	1812.01	1813.05	1.05	0.001593625
P53396	EVAPAKKAKPAMPQ	1446.80	1446.81	0.00	0.002276608
P53396	VLINFASLRSAY	1352.75	1352.73	-0.02	0.001593625
P53396	AAAKMFSKAF	1070.56	1070.56	0.00	0.004582951
P35908	PEIQNVKAQEREQIKTLNNKFASFI	2944.58	2946.58	2.00	0.002276608
P35908	GLTAERTSQNSELNNMQ	1891.87	1892.88	1.01	0.001593625
P35908	LLNQEIEFLKVLY	1620.91	1620.91	0.00	0.001593625
P35908	AEISQIHQSVT	1211.61	1211.61	0.00	0.001593625
P35908	SIIAEVKAQYE	1249.66	1250.66	1.00	0.001593625
P35908	LEEALQQAKE	1157.59	1157.59	0.00	0.001593625
P35908	YQELMNVKLAL	1320.71	1320.71	0.00	0.00757257
P60709	EAQSKRGILTLKYPIEHGIVTNW	2634.43	2634.42	-0.01	0.043333333
P60709	DGVTHTVPIYEGYALPHAILRL	2434.30	2434.28	-0.02	0.001593625
P60709	GVTHTVPIYEGYALPHAILRL	2319.27	2319.27	0.00	0.001593625
P60709	GVTHTVPIYEGYALPHAILRLD	2434.30	2435.27	0.97	0.001593625
P60709	YLMKILTERGYSFTTTAEREIVR	2776.46	2776.45	-0.01	0.032956222
P60709	FEQEMATAASSSSLEKSYELP	2304.05	2305.03	0.99	0.017501509
P60709	LYANTVLSGGTTMYPGIA	1827.91	1828.91	1.00	0.001593625
O60869	LATKINEKPQVIA	1423.84	1424.85	1.01	0.040106306
P12004	TLALVFEAPNQEKVS	1644.87	1644.88	0.00	0.001593625
P10606	EEQATGLEREIMLAAKKGL	2068.10	2068.11	0.01	0.001593625
P10606	PYNVLAPKGASGTRE	1558.81	1558.81	0.00	0.002276608
Q92506	DSGYITGTSVEVTGGLFM	1848.85	1846.96	-1.88	0.066064743

Q9Y6E2	LEAVAKFL	889.53	889.52	0.00	0.006197432
Q9Y6E2	ILVAGSMLAPGGTRI	1454.83	1454.83	0.00	0.003775366
P29597	NKCLELSLPSRAAALSF VSLVD	2314.24	2314.19	-0.04	0.026450742
P02786	QTKFPIVNAELSFFGHA HLGTG	2353.19	2353.18	0.00	0.042135504
P55060	ANLQTLTEYLKKTL	1634.92	1634.93	0.01	0.010505051
P55072	LEQVANETHGHVGA	1460.70	1460.70	0.00	0.003775366
P55072	SIAKARGGNIG	1042.59	1042.59	0.00	0.003775366
P55072	RVINQILTEM	1215.66	1216.67	1.00	0.052802223
P55072	PVPEIRR	865.51	865.51	0.00	0.022133939
P53634	LLGTWVFQVGSSGSQR	1720.89	1720.89	0.00	0.001593625
P20264	DAAGAGGGGGGGGGG GGGGAGGGGGGMQPG SAAVTSGAYRG	3030.30	3029.59	-0.70	0.012337217
P20264	AAGAGGGGGGGGGGG GGGAGGGGGGMQPGS AAVTSGAYRGD	3064.30	3065.62	1.32	0.096697175
O14602	EARSLKAYGELPEHAKI NET	2237.14	2237.13	-0.01	0.024896266
P29692	IARARENIQKSLAGSSGP GASSGTSG	2458.25	2459.26	1.00	0.002276608
P29692	LLEEEITKFEEHVQSV	1928.97	1930.97	2.00	0.001593625
O43542	DALLRGGLPL	1023.61	1023.66	0.05	0.069489685
Q13765	TYIVFGEAKIE	1268.67	1268.66	-0.01	0.01536
P55209	PKGIPEFWLTVFKNV	1773.98	1773.98	-0.01	0.026450742
P55209	EPILKHLK	958.60	958.55	-0.04	0.021118721
Q00610	TIRRFQSVPAQPGQTSPL LQYFGILL	2929.62	2931.60	1.98	0.062583818
Q00610	AILGNQMFTHY	1293.62	1294.62	1.01	0.043333333
Q00610	VNTSAVQVLIEHIGNL	1705.94	1705.92	-0.01	0.001593625
O14818	IVVLGVEKKSVAKLQ	1610.01	1610.00	-0.01	0.003775366
Q92903	FANTIPGHGGIMD	1328.62	1327.73	-0.88	0.062583818
P98153	DPAQSGSTPAAEALPGG GRHSRSSLNTVV	2818.40	2820.36	1.96	0.096697175
Q92953	SSQEGCKMENHLFAPEI HSNPGD	2524.07	2522.62	-1.45	0.092548077
Q92973	QFPLPLKERLAAFYGV	1831.00	1832.01	1.01	0
Q9NYU2	EVQGFLFGKLR	1274.71	1275.70	0.99	0.001593625
Q00765	AITKEAKKATVNLLGEE KKST	2258.28	2259.27	0.98	0
Q9Y2C2	ILSEKHGFNLVTSDIHNK	2033.07	2037.75	4.68	0.082482646
P30040	TVTFYKVIPKSKFVLVK F	2143.28	2143.28	0.00	0.00757257
P30040	EFKRLAENSASS	1319.65	1319.64	0.00	0.039757576
P30040	KESYPVFYLFR	1447.75	1447.75	0.00	0.002276608
P30040	ALAGEFIRASGVEARQA LLKQGQ	2412.32	2412.32	0.00	0.003157895
P30042	AAIFPGGFGAAKNLSTF AV	1837.97	1837.96	-0.01	0.002276608
P30044	AFVTGEWGRAHKAEGK VRLLA	2295.26	2295.27	0.01	0.001593625
P30049	VPTLTGAFGILAAHVPT LQVLRPGLVVVHAE	3174.83	3175.82	0.99	0.003157895

P30049	LGAAKANLEKAQAELV GTA	1854.02	1855.02	1.00	0.001593625
P30049	EATRAEIQIRIEANEALV KALE	2448.33	2449.34	1.01	0.002276608
P20618	TRLSEGFSIHTR	1402.73	1402.73	0.00	0.002276608
P22087	QIHIKPGAKVLYLGAAS GTTVSHVS	2516.37	2517.38	1.01	0.001593625
P30101	SFSEAHSEFLKAASNLR	1892.94	1892.92	-0.01	0.001593625
P30101	AGHKLNFAVASRKTFS HELS	2199.15	2200.16	1.00	0.066064743
P30101	GKALERFLQ	1060.60	1060.60	0.00	0.019836639
P30101	PNIVIAKM	900.51	900.50	-0.01	0.001593625
P30101	FISYLQREATNPPVIQEE KPKKKKKAQE	3326.84	3327.85	1.02	0.082482646
O75363	SSLGSVKLD	886.48	886.45	-0.02	0.021633931
P60842	LKATQALVLAPTRELAQ QIQKVVMALG	2889.68	2890.67	0.98	0.012721794
P60842	EMLSRGFK	948.48	948.48	0.00	0.007002188
P60842	VLEVTKKFMR	1249.72	1249.71	-0.01	0.001593625
P60842	FTVSAMHG	848.39	848.38	0.00	0.010505051
P04792	GVVEITGKHEERQ	1480.76	1480.77	0.00	0.048379521
Q13148	AGWGNLVYVVNYPK	1578.82	1578.80	-0.02	0.001593625
Q13148	ASSAVKVKRAVQKTS	1558.92	1558.92	0.00	0.002276608
P39687	YRENVFKLLPQLTYL	1896.05	1896.06	0.01	0.002276608
O75531	FVAEPMGEKPVGSLAGI GEVLGKKLEERGF	3143.67	3143.66	0.00	0.015837821
O75531	KAYVVLGQFLVLKK	1605.00	1605.00	0.00	0.001593625
P14317	AAKGFGGKYGVER	1338.70	1338.71	0.00	0.01536
P22392	SVKSAEKEISLWFKPEE LV	2218.19	2218.18	-0.01	0.002276608
P04908	VGAGAPVYLAAVLEYL TAEILELAGNAAR	2914.58	2916.58	2.00	0.074812968
O95881	GGYIPRILFL	1147.68	1147.67	0.00	0.004582951
Q9H6Z4	VLAPSGATAAGAGD	1138.56	1138.63	0.07	0.018474374
P55809	ESFAMIRGGHV	1184.58	1184.57	0.00	0.001593625
P49257	IKEHLHIVKR	1271.78	1271.78	0.00	0.022133939
P55851	LPCHFTSAFGAGFCTTVI ASPVD	2340.09	2338.68	-1.42	0.082482646
Q15370	AKESSTVFELKRIVEGIL KRPP	2496.44	2496.44	0.00	0.003775366
P49327	PGSAELQKVLQG	1225.67	1225.66	0.00	0.001593625
P49327	LVEAVAHILGIR	1289.78	1290.78	1.00	0.055479138
O75747	QEIRKVAVQQLD	1425.79	1425.81	0.01	0.078844169
Q16378	QGPQRPPPEGLLPRPPG	1774.95	1774.96	0.01	0.001593625
P49411	YVKNMITGTAPL	1306.70	1307.69	0.99	0.005398111
P49411	PELGLKSVQKLL	1323.81	1323.81	-0.01	0.003775366
P49411	TYIPVPAR	915.52	915.51	-0.01	0.004582951
P14625	KIRLISLT	942.62	942.63	0.01	0.02661101
P14625	KEKNLLHVT	1080.63	1080.63	0.00	0.027496757
P14625	EYKAFYKSFSKES	1594.77	1594.76	-0.01	0.012337217
P14625	MMPKYLNFVKGVV	1524.82	1524.82	0.00	0.026450742
P14625	QYVERMKEKQ	1320.65	1320.65	0.00	0.010505051

P14625	GKRFQNVAKEGVKF	1606.89	1606.92	0.02	0.048379521
P61254	MKFNPFVTS	1069.53	1069.52	0.00	0.002276608
P32119	EGIAYRGLFII	1232.69	1233.69	1.00	0.070288918
P32119	EALRLVQAFQYT	1419.75	1419.75	0.00	0.07211848
P49588	PTLLFANAGMNQFKPIF LNTI	2349.26	2350.27	1.01	0.00757257
P49588	TGMGLERLVSVLQNKM SNY	2139.08	2140.08	1.00	0.012337217
P49588	LTGLIAEEKGLVV	1340.79	1340.82	0.03	0.01536
P49588	PVRVVSIGVPVSELL	1562.94	1562.93	0.00	0.001593625
P06744	YSKNLVTE	952.49	952.48	0.00	0.012337217
P06744	LGPLMVTEALKPYSSGG PRVWYVSNI	2833.48	2834.48	1.00	0.002276608
P06744	PSAVAKHFVALSTNTTK VKEFGI	2444.34	2444.33	-0.01	0.001593625
P06744	QYLHRFAAYFQQG	1610.76	1610.76	0.00	0.060906835
P06744	QWGVELGKQLAKKIEP EL	2048.13	2048.12	-0.01	0.003775366
P40227	VLRTNLGPKGTMKMLV SGAG	2029.12	2030.11	0.99	0.020057307
P40227	GTTSNVLIIGELLKQA	1655.95	1655.95	0.00	0.004582951
P40227	SILAIKKQ	899.58	899.58	0.00	0.008347245
P40227	KGFVVINQKGI	1201.72	1201.72	0.00	0.043333333
P40227	ALLIIPKVLAQNSGF	1582.94	1583.95	1.01	0.002276608
P40227	EIMRAGMSSLKG	1260.63	1260.63	0.00	0.004582951
Q9BQ67	QRPFVGHTRSVE	1394.71	1394.70	-0.01	0.02661101
P14866	DQHGGGGGGGGGAGA AGGGGGGENY	1911.75	1913.21	1.46	0.032956222
P14866	QHGGGGGGGGGAGAA GGGGGGENY	1814.73	1814.51	-0.22	0.047058824
P14866	QHGGGGGGGGGAGAA GGGGGGENYD	1894.72	1894.60	-0.12	0.074812968
P14866	PHKTPASPVVHIRGLI	1721.01	1721.00	-0.01	0.001593625
P14866	NQIYIAGHPAFVNYSTS QKISRPG	2647.35	2647.35	0.00	0.001593625
P14866	SRSVNSVLLFTILNPIYSI TT	2337.29	2337.30	0.01	0.001593625
P14866	SVQSAQRAKASLNGA	1486.79	1486.80	0.02	0.001593625
P14866	PGSNPNKRQRQPPLLG	1757.96	1757.96	-0.01	0.063654574
P57737	AAKQQPLTELAAHG	1433.76	1433.76	0.00	0.020057307
P57737	PHRLAVAGE	948.51	948.52	0.00	0.002276608
P49720	AVSGMGVIVHIIEK	1451.82	1451.81	0.00	0.001593625
P61158	ITYFIQQLLR	1293.74	1293.73	-0.01	0.002276608
Q9Y222	DSTPCISV	820.36	818.42	-1.94	0.05671164
P40425	DNMLLAEGVAGPEKGG GSAAAAAAAAASGGGV SP	2863.38	2862.49	-0.89	0.052802223
Q14103	QKEHKLNGKVI	1275.73	1275.72	0.00	0.003775366
P49915	QEKLMQITSLHSLNAFL LPIKTVGVQG	2947.62	2948.63	1.01	0.024896266
P49915	VTPTFLTTGVLSTLRQA	1804.01	1804.01	0.00	0.001593625
Q04837	PVLRQVEGKNPVTIFSL ATNEMWRSG	2944.52	2945.52	1.00	0.001593625
Q04837	VAYQYVKKGSRIYLEG	2114.19	2114.18	0.00	0.020057307

	KI				
Q12874	TSLFAKNPKSKGTKR	1661.96	1661.95	-0.01	0.002276608
O95336	EAAARLLTVPFEKHSTL	1864.02	1864.01	-0.01	0.022133939
P34130	DNAEEGGPGAGGGGCRGV	1540.63	1538.78	-1.86	0.038527188
O60568	EQPSLRPHH	1081.54	1081.54	0.00	0.012721794
Q12931	TGIGMTQEELVSNLGTIARSGSKAFL	2679.39	2679.39	0.00	0.001593625
Q12931	ELTLLHLREF	1251.70	1251.69	-0.01	0.001593625
Q12931	PRAMVGRLNELLVKALERH	2201.26	2201.24	-0.02	0.012337217
P24752	VMVAGGMESMSNVPYVMNRGSTPYGGVKLE	3159.49	3160.44	0.95	0.001593625
P24752	AAKRLNVTPLARIVAFA	1810.09	1810.08	-0.01	0.002276608
P24752	FPIAPVYAASMVLK	1505.83	1505.83	0.00	0.099189883
Q9BT73	EPLIHVFAKNLVAFVSQEAGNRAVLLAVAVK	3284.88	3285.85	0.97	0.012721794
Q14566	LAAAAEPGAGSQHLEVR	1675.86	1675.87	0.00	0.024608501
P26368	VPPPGFEHITPMQYKAMQAAGQIPATALLPTMTP	3603.83	3605.82	1.99	0.012337217
P26368	DQVKELLTSFGPLKAFNLVK	2246.27	2248.26	1.99	0.003775366
P26368	QVKELLTSFGPLKAFNLVK	2114.21	2114.21	0.00	0.001593625
O75223	EVWGVVWKMNKSNLNSL	1985.02	1985.01	-0.01	0.002276608
P32969	IELVSNSAALIQQATTVKNK	2127.19	2127.19	0.00	0.001593625
P32969	GIYVSEKGTVQQA	1378.71	1378.70	-0.01	0.001593625
Q9BVC6	FAPPGQQKREAPV	1423.76	1423.75	0.00	0.003157895
P16930	ALMPFAVPNPKQ	1311.70	1311.69	-0.01	0.012337217
Q13387	PEAAAGPGGVELV	1165.60	1164.72	-0.88	0.05671164
P16989	PAPKSPVGSGAPQAAAPAPAAHVAGNPGG	2501.28	2502.29	1.01	0.001593625
P67809	TKPGTTGSGAGSGGPGGLTSAAPAGG	2070.00	2069.98	-0.02	0.001593625
P67809	VVEGEKGAEAANVTGPGGVPVQGSKYAA	2641.33	2642.34	1.00	0.001593625
Q9Y285	ALFQPQQHPAR	1291.68	1291.67	-0.01	0.002276608
Q9C0K7	IYSVGITACELASGQVPFQD	2079.00	2077.12	-1.88	0.0466541
P26447	VMVSTFHKYSGKEG	1568.77	1568.77	0.00	0.001593625
P26447	EAAFQKLMSNL	1232.62	1232.62	0.00	0.003775366
Q14657	QLSLVVRTMQRFGPPVSR	2053.13	2053.12	0.00	0.001593625
P50502	VAQNPANMSKYQSNPKVMNLISKLSAKFGGQA	3420.76	3421.78	1.01	0.015837821
Q9Y3B8	LNILAEGPNLIIKQP	1631.96	1631.97	0.01	0.00757257
O95725	DITLFHYI	1002.52	1002.55	0.03	0.001593625
P61923	VPLTEQTVSQVLQSAKEQIKWSLLR	2880.61	2880.60	-0.01	0.024896266
Q9UBQ5	QAHQEERPIRQILYLG	1933.02	1933.02	0.00	0.012337217
Q9NVP2	SVLVGPVPAGRHMFVFQA	1911.02	1912.00	0.98	0.001593625
Q96T58	RLNTVASPK	984.57	986.53	1.95	0.064587481

Q969H8	VRPGGVVHSFSHNVGPG	1701.87	1701.86	-0.01	0.001593625
Q96S94	MAAAAAAAGAAGSAAPAAAAGAPGSGGAPSGSQGVLIG	2978.45	2977.51	-0.94	0.034355828
Q9UHX1	TIRQAFAPFGPIKSI	1644.94	1645.93	0.99	0.052436195
P62158	GQVNYEEFVQMMTAK	1773.81	1773.81	0.00	0
*	EMVESMKKVAGM	1320.62	1320.62	0.00	0.002276608
*	TLSEESYK	955.45	955.45	0.00	0.005398111
*	CKLIMQLLR	1116.65	1115.63	-1.02	0.017501509
Q96I99	QITKLYNLFLKI	1475.87	1475.87	0.00	0.003157895
Q96I99	IIFLNGGKPANFL	1402.80	1402.80	0.01	0.002276608
Q86VP6	PFYKITSEALLVTQQLVKVIRPL	2655.57	2656.57	1.00	0.001593625
Q9UNZ2	IVTISQATPSSVSRGTAPS	1857.98	1857.97	-0.01	0.001593625
Q9UNZ2	ESQTLKEANLLNAVIVQRLT	2221.24	2222.24	1.00	0
Q96HN2	GCCAALK	664.30	664.84	0.54	0.090652699
Q96JI7	EALQKLIDDQDISISLLSLR	2233.23	2237.27	4.04	0.096863835
Q8IYB5	KNAIAITNISSSD	1332.69	1332.91	0.22	0.073445128
Q96L73	PLQTSGKAAAPSE	1255.64	1255.62	-0.02	0.005398111
Q8WU39	DRAPLTATAPQL	1252.68	1252.68	0.00	0.001593625
Q8WU39	RAPLTATAPQL	1137.65	1137.65	0.00	0.001593625
Q8WU39	RAPLTATAPQLDD	1349.69	1349.69	-0.01	0.074812968
Q8WU39	QIYEAHQQGRGALEALLCGGPQGACSEKVSATREEL	3781.81	3780.80	-1.02	0
Q8WXS5	PSKGSVAAGLAGAGGGGGGAVGAFGGAAGGAGGGGGGGGGAGAER	3310.58	3310.71	0.13	0.078844169
Q8WXS5	PSKGSVAAGLAGAGGGGGGAVGAFGGAAGGAGGGGGGGGGAGAERD	3425.61	3423.78	-1.83	0.031793343
O00231	ILHSIVKR	964.62	964.61	-0.01	0.052802223
O00231	PIISTHLAKLY	1254.73	1254.73	0.00	0.010505051
Q5VSY0	MASAVLSSVPTTASRFALLQV	2164.16	2164.07	-0.09	0.060606061
Q7KZF4	SHHQKPVNAIIEHVR	1763.95	1763.95	0.00	0.012721794
Q96Q04	DAGAAGGEAGGAGAPGPAEE	1592.67	1591.97	-0.70	0.070288918
Q9NR45	EMAVEFLHELNVPFFKVGSG	2231.11	2231.08	-0.03	0.001593625
Q8TAM2	DNTHVEAIACIGSNHFYS	1976.87	1978.04	1.18	0.017996401
Q8TBB5	DLWVLHLATKTWEQVKSTGGPSGRSGHRMVAWKRQLILFGGFHESTR	5357.79	5359.43	1.64	0.066064743
Q8TDN2	QQAGEVTTAKPEGPS	1481.70	1480.70	-0.99	0.026450742
Q7LBR1	ELSQRLARLR	1222.73	1222.73	0.00	0.06291834
Q8IXH7	PPPVELIRVPAFL	1446.86	1446.84	-0.02	0.029101856
Q96EP5	EIGKLFVGGLD	1146.63	1148.62	1.99	0.027496757
Q96EP5	QAVNMHFH	965.42	965.41	0.00	0.001593625
Q9Y2S7	QVPIQHELFERFLLY	1914.00	1913.99	-0.01	0.054851348

Q8N7A4	DPCHGSGPALM	1081.43	1081.58	0.15	0.004582951
Q8N7I6	MWICPGGGGGGGGG GGGGG	1519.59	1519.88	0.29	0.062583818
P20591	PAAASHPLLLNG	1159.63	1159.63	0.00	0.001593625
P20591	SLRALGVEQ	971.54	971.55	0.01	0.02661101
P20591	IATTEALSMAQEV	1362.67	1362.66	-0.01	0.012337217
P20591	QLSLSEALQREKIFFENH PYFR	2734.39	2734.38	-0.01	0.019836639
P20591	VSIKNFEEFFNLHRTAKS KIE	2536.34	2536.40	0.05	0.012337217
Q0IIN9	MLRQSCSFPVTSLPALG GVCGREGAGAEVPPAA CGCEGRD	3945.80	3945.28	-0.52	0.066725198
Q8NFL0	TFFNLTLKEIHFLKWL	2049.15	2048.19	-0.96	0.028343211
P00558	NGAKSVVLMSHLGRP	1564.85	1565.83	0.98	0.012721794
P00558	KYSLEPVAVELKSLLGK	1873.09	1873.08	-0.01	0.001593625
P00558	KIQLINNML	1101.62	1101.62	0.00	0.003775366
P00558	EEGAKIVK	854.49	854.48	0.00	0.003775366
P00558	LMSKAEKNGVKITLPV	1727.00	1728.00	1.00	0.042949177
P00558	ENAKTGQATVASGIPAG WMGL	2040.01	2040.01	0.00	0.001593625
P00558	KVSHVSTGGGASLELLE GKVLPGV	2333.30	2333.29	0.00	0.010505051
Q53EQ6	RAPAPPPPAEGGYGD	1450.68	1449.77	-0.92	0.066064743
Q8NBS9	PHSKHLYTA	1052.54	1052.54	0.00	0.001593625
Q8NBS9	AKVYVAKV	876.54	876.54	0.00	0.020577308
Q8NBS9	AKVYVAKVD	991.57	991.57	0.00	0.003775366
Q8NBS9	KGTVLALTENNF	1305.69	1305.69	-0.01	0.001593625
Q8NBS9	SLHRFVLSQAK	1284.73	1284.72	-0.01	0.003157895
Q8NBT0	DFSINTKQLASGSM	1497.71	1496.82	-0.90	0.031793343
Q8NEY1	GIKVHGQKAAWED	1437.74	1437.75	0.01	0.047058824
O94776	QFLVVARAVGTFARAL	1700.97	1700.97	0.00	0.001593625
P52272	FFPPERPQQLPHGLGGIG MGLGPGGQPI	2853.47	2854.49	1.01	0.004582951
Q96KP4	GSEIPLPPILLGRLGS	1617.95	1617.93	-0.02	0.001593625
Q96KP4	ELIFARK	857.51	857.51	0.00	0.00757257
Q96KP4	VGAQILLHSHKK	1329.79	1329.78	-0.01	0.02661101
Q96KP4	LTREGGSIPVTLTFQEAT GKNVMLLPVGSA	3085.65	3086.63	0.98	0.003157895
Q8TBM8	NCKELERLTSLYKGG	1709.88	1710.87	1.00	0.060606061
P62979	TIENVKAKIQ	1142.67	1142.67	0.00	0.001593625
P62979	QQRLIFAGKQLE	1412.78	1412.78	0.00	0.001593625
O95876	DRSLVGKLIS	1068.63	1068.66	0.03	0.062583818
P49321	SEAKKLLGLGQKHLVM G	1808.03	1808.03	-0.01	0.04361223
P26599	PVSAQHAKLSL	1149.65	1149.65	0.00	0.001593625
Q9Y2K3	LGKVRSAAARLD	1255.74	1255.62	-0.12	0.063943162
Q71U36	DSFNTFFSETGAGKHVP RAVFV	2412.19	2413.18	1.00	0.040106306
Q71U36	SFNTFFSETGAGKHVPR AVFV	2297.16	2297.17	0.01	0.001593625
Q71U36	EVRTGTYRQLFHPEQLI TGKE	2483.29	2483.28	-0.01	0.073160357

Q71U36	AANNYARGHYTIGKEII	1889.97	1889.97	-0.01	0.015837821
Q71U36	VNAAIATIKTKRTIQFV	1873.11	1873.11	0.00	0.0466541
Q14011	EGKLFVGGLSF	1134.61	1134.60	-0.01	0.002276608
Q14011	TNEQSLEQVFSKYGQIS EVVVVK	2610.35	2610.34	-0.01	0.012337217
P49736	QQIGEKIFASIAPSIYGHE	2070.04	2070.05	0.00	0.001593625
P49736	LTEPIISRF	1074.61	1075.60	0.99	0.010505051
O00232	ESEAFLSNLVVNKTIFA KV	2090.14	2090.13	-0.01	0.001593625
P13639	ITKGVQYLNEIK	1404.80	1404.79	-0.01	0.001593625
P39019	TVKLAKHKELAPY	1496.87	1497.87	1.00	0.002276608
P39019	LDRIAGQVAAANKKH	1572.88	1572.88	0.00	0.078844169
P63173	KEKAEKLKQSLPPGLAV KELK	2333.40	2334.40	0.99	0.003775366
O76094	PEKAKALSKHLPSS	1491.84	1492.84	1.00	0.00757257
P61204	AVLLVFANKQ	1101.65	1101.65	0.00	0.001593625
P18085	AVLLLFANKQ	1115.67	1115.67	0.00	0.001593625
P21926	EVIKEVQEFYK	1392.73	1392.73	0.00	0.002276608
Q9UKM 9	DGGGAGGGGGGGGSG GGGSGGGGGGGSS	1806.67	1807.51	0.83	0.090652699
Q9UKM 9	GGGAGGGGGGGGSGG GGSGGGGGGGSS	1691.65	1692.89	1.24	0.049226442
P61978	SSGPERILSISA	1215.65	1215.64	-0.01	0.082482646
P61978	IETIGEILKKIIPTLEEGL QLPSPTATSQLPLES	3658.03	3660.02	1.99	0.001593625
P61978	LISESPIKGRAQPY	1557.85	1557.84	-0.01	0.001593625
P61978	YSYAGGRGSYG	1136.49	1136.48	-0.01	0.001593625
P61978	LGGPIITTQVTIPK	1436.86	1436.85	-0.01	0.001593625
P61978	LGGPIITTQVTIPKD	1551.89	1552.88	0.99	0.017501509
P61978	LAGSIIGKGGQRIKQIRH ESGASIKI	2716.58	2717.59	1.01	0.028343211
P61978	RIITITGTQ	1001.59	1001.58	-0.01	0.002276608
P61978	QIQNAQYLLQNSVKQY SGKFF	2486.26	2487.25	0.99	0
P09104	LYTAKGLFRAAVPSGAS TGIYEALELR	2853.54	2854.54	1.00	0.00757257
P10599	VKQIESKTAFQEAL	1590.86	1590.86	0.00	0.074812968
P31946	KSELVQKAKLAEQAER Y	1990.08	1990.08	0.00	0.001593625
P31946	TLNEESYK	982.46	982.46	0.00	0.010505051
P31946	STLIMQLLR	1073.63	1073.62	-0.01	0.001593625
P83876	EVLYSIAEKVKNFAVIY LV	2179.23	2180.23	1.00	0.002276608
P07737	LRTKSTGGAPTFNVTVT KT	1978.08	1978.06	-0.02	0.001593625
P07737	LRTKSTGGAPTFNVTVT KTD	2093.11	2094.12	1.01	0
Q0VF49	DPWLSPKY	1004.50	1002.55	-1.95	0.069489685
P29728	ETIRNILLHQLQSARPVI L	2195.29	2196.32	1.03	0.029101856
P22234	VTTKEIVLA	972.59	972.58	0.00	0.052802223
O75347	LEEAEEYKEARLVL	1690.88	1691.89	1.01	0.018962963

P08779	KVRALEEANA	1099.60	1099.60	0.00	0.043333333
P08779	YSPYFKTIE	1146.56	1146.56	0.00	0.005398111
P08779	LSRILNEMR	1130.62	1131.62	1.00	0.002276608
P08779	QYEQMAEKNRR	1434.67	1434.66	-0.01	0.029101856
P00338	LVKVTLTSEEEARLKKSA	2001.15	2001.14	-0.01	0.002276608
P20700	QLLLNYAKKES	1288.70	1289.70	1.00	0.020057307
P20700	QIAQLEASLAAAKKQLA	1735.98	1735.96	-0.02	0.001593625
P20700	QPMGGWEMIRKIG	1484.73	1484.72	-0.01	0.012337217
Q9Y6N9	DLVVAVCPPKEY	1313.67	1312.68	-0.99	0.066725198
P51572	VGNAEVKLEEENRSLKA	1884.99	1885.99	1.00	0.073744437
P11586	PETITWQRVL	1241.68	1241.68	0.01	0.001593625
P62266	EVLVAGFGRKGHAVG	1477.81	1477.82	0.00	0.001593625
P07741	PASFRAAIGLLARHLKATHGGRI	2412.40	2412.40	0.00	0.002276608
P07741	ALEPGQRVVVV	1165.68	1165.68	0.00	0.005398111
P54819	VVFASILAAFSKATS	1510.84	1510.84	0.00	0.06495802
P04406	PSKIKWG	814.47	814.47	0.00	0.028343211
P04406	AGAEYVVESTGVFTTMEKAGAHLQGGAKRVIISAPSA	3702.90	3704.90	2.00	0.003157895
P04406	NFGIVEGLMTTVHAITATQKTV	2330.23	2331.25	1.02	0.001593625
P04406	GPSGKLWR	899.50	899.50	0.00	0.012337217
P04406	NEFGYSNRVV	1183.56	1183.56	-0.01	0.019836639
P04406	LMAHMASKE	1016.48	1016.48	0.00	0
P35754	LVSLQQSGELLTRLKQIGALQ	2294.33	2295.34	1.01	0.008347245
P17174	SAYQGFASGNLER	1398.65	1399.64	0.99	0.002276608
P62888	IIRSMPEQTGE	1259.62	1259.62	0.00	0
P62888	IIRSMPEQTGEK	1387.71	1388.71	1.00	0
P09382	AKSFVLNLGK	1075.64	1076.64	1.01	0.034355828
P62310	GVVLVAPPLRVG	1175.74	1175.74	0.00	0
P61326	VMIRKEAYVHKSVMEELKRII	2571.44	2572.44	1.00	0.016024653
P61326	EHISFTTSKIGSLI	1513.81	1513.82	0.00	0.001593625
P34949	EAATHLKQTMSH	1334.64	1334.63	-0.01	0.002276608
P26038	AELEFAIQPNTTGKQLF	1905.98	1906.97	0.98	0.002276608
P26038	AVLEYLKIAQ	1146.66	1146.66	-0.01	0.004582951
O43678	QVTRALENVLSGKA	1467.80	1467.80	0.00	0.001593625
P30086	LSKWSGPLSLQEV	1442.78	1442.78	0.00	0.001593625
P30086	EQPQHPLHVTYAGAAV	1698.85	1698.85	0.00	0.001593625
P30086	ELGKVLTPTQVKNRPTSISW	2235.24	2235.26	0.02	0.001593625
P30086	SGKLYTLVLT	1093.64	1094.64	1.00	0.012721794
P30086	YVGSGPPKGTGLHRYVWLVYEQ	2505.28	2506.27	0.99	0.020057307
P30086	DYVPKLYEQLSGK	1538.80	1538.80	0.00	0
P30086	YVPKLYEQLSGK	1423.77	1423.77	0.00	0
O75426	MLQRAEGGGGGVGPPAPET	1795.85	1793.87	-1.98	0.0466541

P52209	EGAGHFVKMVHNGIEYG	1825.86	1826.87	1.01	0.002276608
P07205	NGAKAVVLMSHLGRP	1548.86	1547.81	-1.05	0.042135504
P07205	EEGAKIVK	854.49	854.48	0.00	0.003775366
Q13867	FLKKMVAASIKD	1331.76	1331.65	-0.11	0.001593625
P25787	ERSVHKVEPITKHIGLVYSGMGP	2531.33	2531.32	-0.01	0.001593625
P30085	QTMAANAQKNKFLI	1559.81	1559.82	0.01	0.007002188
Q8NFF5	EVATIAAEVTSFSNRFTHVLTAGGIGPTH	2964.51	2964.49	-0.02	0.001593625
P30041	SWGILFSHPR	1198.62	1198.62	0.00	0.001593625
P30041	EKGMPVTARVVFVFGP	1730.92	1731.91	1.00	0.012721794
P30041	KKLKLSILYPATTGRNF	1949.15	1950.14	0.99	0.003775366
P30041	EILRVVISLQLTAEKRVATPV	2316.39	2317.39	1.00	0.001593625
P04264	PEIQKVKSREREQIKSLNNQFASFI	2988.61	2989.63	1.02	0.002276608
P04264	PEIQKVKSREREQIKSLNNQFASFID	3103.64	3105.63	1.99	0.001593625
P04264	KVRFLEQQNQVLQTKWELLQQV	2754.52	2754.52	0.00	0.001593625
P04264	KVRFLEQQNQVLQTKWELLQQVD	2869.55	2871.54	1.99	0.001593625
P04264	FLEQQNQVLQTKWELLQQVDTSTR	2931.51	2933.50	1.99	0.074812968
P04264	TSTRTHNLEPYFESFINNLRRRVD	2964.50	2965.51	1.02	0.022452989
P04264	EINKRTNAENEFVTIKK	2015.08	2015.09	0.01	0.001593625
P04264	EINKRTNAENEFVTIKKD	2130.11	2130.10	-0.01	0.001593625
P04264	SIIAEVKAQYE	1249.66	1250.66	1.00	0.001593625
P04264	IAQKSKAEAESLYQSKYEELQITAGRHG	3134.60	3135.61	1.01	0.001593625
P04264	IAQKSKAEAESLYQSKYEELQITAGRHGD	3249.63	3250.62	1.00	0
P04264	NVKKQISNLQQSIS	1585.88	1585.88	0.00	0.012337217
P04264	NVKKQISNLQQSISD	1700.91	1701.89	0.99	0.001593625
P04264	YQELMNTKLAL	1322.69	1322.69	0.00	0.004582951
P63162	GRIFIGTFKAF	1255.71	1256.71	1.01	0.020057307
P61604	KVLLPEYGGTKVVL	1514.91	1514.91	0.00	0.001593625
P61604	KVLLPEYGGTKVVLDD	1726.95	1726.97	0.02	0.092548077
Q96EZ8	SGPDSQGLLDSSLMASGTASR	1999.93	2001.12	1.20	0.078844169
P61981	REQLVQKARLAEQAERY	2087.12	2087.11	-0.01	0.062583818
P61981	STLIMQLLR	1073.63	1073.62	-0.01	0.001593625
P13010	IESKIQPGSQQA	1284.67	1284.67	0.00	0.001593625
P13010	EAAAVALSSLIHAL	1346.76	1346.75	-0.01	0.001593625
P13010	GITLITKEEASGSSVTAEEAKKFLAPK	2804.52	2804.53	0.02	0.01536
P62277	VKEQIYKLAKKGLTPSQIGVILR	2581.57	2581.55	-0.02	0.001593625
Q9C0A1	MATLNSASTTGTTPSPGHNAPSLPS	2412.12	2413.31	1.19	0.099189883
P08708	TKEMLKLL	974.58	974.58	0.00	0.026450742

P08708	FGSLSNLQVTQPTVGMNFKTPRGPV	2690.39	2691.39	1.01	0
Q9H1R3	DGILFMHKMRVLHL	1690.92	1690.96	0.04	0.026450742
P49773	EIAKAQVARPGG	1177.66	1177.66	0.00	0.001593625
P49773	EIAKAQVARPGGD	1292.68	1292.69	0.01	0
P49773	TIFGKIIRKEIPAKIIFE	2115.28	2115.27	-0.01	0.010505051
P49773	ISPQAPTHFLVIPKKHISQISVAE	2639.48	2639.48	0.00	0.004582951
P49773	LGLNKGYRMVVNEGS	1635.84	1635.84	0.00	0.001593625
P98196	EETGEGPLVNTSD	1328.57	1329.82	1.25	0.059104342
P35637	DQSSMSGSGGGGGGGGGGGSGGGGGYGNQ	2240.84	2241.13	0.30	0.048379521
P35637	QSSMSGSGGGGGGGGGGGSGGGGGYGNQ	2126.79	2128.16	1.37	0.099189883
P35637	QSSMSGSGGGGGGGGGGGSGGGGGYGNQD	2223.81	2225.11	1.30	0.012337217
P31939	KTGLVEFARNLTALGLNLVASGGTAKALR	2940.69	2941.68	0.99	0.029101856
P31939	VPTAKIISREVS	1298.76	1299.77	1.01	0.003775366
P10323	IGSNALRMIQSATPPPPTTR	2107.12	2108.24	1.12	0.096863835
P08621	ERPGPSPLPHR	1223.65	1223.65	0.00	0.003775366
O75821	LQELFRPFGSISRIYLAK	2137.20	2137.20	-0.01	0.02661101
O75821	AARAIAGVSGFGY	1238.64	1239.63	0.99	0.001593625
P17987	GATILKLL	827.55	827.54	-0.01	0.00757257
P17987	AVLAIKYT	877.53	877.52	0.00	0.004582951
P17987	ITKERIQKILATGANVILTTGGI	2409.43	2409.43	0.00	0.028593311
Q96T23	ECRADPK	817.38	818.73	1.36	0.090652699
P40926	IVRANTFVAELKGL	1529.89	1530.89	1.00	0.052436195
Q8WV74	PQKATVVPVLAGVGPL	1544.93	1545.93	1.00	0.003157895
P31483	EMPKTLYVGNLSR	1488.78	1488.77	0.00	0.029101856
P28066	PGAMSRPFGVALLFGGV	1674.89	1674.89	0.00	0.001593625
P28066	EKGPQLFHM	1067.52	1067.52	0.00	0.001593625
Q14978	LYPLVLGFLR	1189.72	1189.72	0.00	0.001593625
Q99470	VRYGSGSGQQSVTGVTSV	1767.88	1768.89	1.02	0.012337217
P14174	PMFIVNTNVPRASVP	1656.87	1656.85	-0.01	0.074812968
P14174	PMFIVNTNVPRASVPD	1755.90	1755.90	0.00	0.031793343
P14174	GFLSELTQQLAQATGKPPQYIAVHVVP	2891.55	2892.57	1.01	0.002276608
P14174	GFLSELTQQLAQATGKPPQYIAVHVVPD	3006.58	3007.57	0.99	0.026450742
P14174	RVYINYY	989.50	989.49	0.00	0.003775366
Q99541	DSVASTITGVM	1095.51	1094.48	-1.03	0.069489685
Q92499	VLMAAETGSGKTGAFSIPVIQIVYETLK	2922.58	2924.58	2.00	0.002276608
Q92804	TGKPKGEATVSF	1220.64	1220.66	0.02	0.00757257
Q9Y3C8	EATRRVVSEIPVLKTNAGPR	2174.23	2174.23	0.00	0.029101856
Q9Y4K0	DLVLNAEMVQQTTYLE	1865.91	1866.10	0.19	0.078844169
Q92806	EKVEEEGAGEGAGGEAGA	1627.70	1628.90	1.20	0.073160357

Q9Y4X5	DGLLCGETGGGGGSALGPGGGGGGGGGGGGGGGPGHEQEE	3121.30	3119.75	-1.55	0.07211848
O60902	DRSSPAVRAAGGGGGGGGGGGGGGGGVGGGGAGGGAGGGRSPVREL	3703.76	3703.90	0.14	0.032956222
O60902	RSSPAVRAAGGGGGGGGGGGGGGGGVGGGGAGGGAGGGRSPVRELD	3703.76	3704.91	1.14	0.031793343
Q02818	EQELEALFTKELEKVY	1950.00	1949.99	-0.01	0.003775366
P07858	IMAEIYKNGPVEGAFSVYS	2074.01	2073.98	-0.03	0.001593625
P07954	PKIANAIMKAA	1126.65	1126.65	-0.01	0.001593625
P07954	AKSKEFAQIIKIGRTHTQ	2055.16	2055.14	-0.02	0.022452989
P29374	SGAVWKQIYMD	1296.62	1295.83	-0.79	0.068935428
P78527	EFVKSVLKIVEKL	1512.93	1512.91	-0.01	0.005398111
Q8N1G1	AGYQPTPLAAPAEPGSKYSLASLD	2385.18	2383.22	-1.96	0.063943162
P78371	LVKSTLGPKGM	1145.65	1145.65	0.00	0.003775366
P78371	KILLSSGR	872.54	872.54	0.00	0.008347245
P78371	LVAQLRAAHSEGNTTAGL	1807.95	1808.95	1.00	0.002276608
Q13224	LQKEEAALAPRSVSLKD	1836.01	1835.95	-0.06	0.042135504
P41091	LQYAAPGGLIGVGTKI	1556.89	1556.88	-0.02	0.002276608
Q99497	VMRRAGIKVTVAGLAGK	1726.04	1726.04	0.00	0.019836639
Q99497	GLILTSRGPGTSFEFALAIVEALNGKEVAAQVKAPLVLK	4007.28	4008.29	1.01	0.012337217
Q12906	ITQSAQHALRLAAFGQLHKVLGM	2489.37	2490.35	0.98	0.003775366
O14980	QGEVVREFMK	1204.59	1204.58	-0.01	0.029101856
Q9Y490	PETQVVLINAVK	1309.76	1309.76	0.00	0.004582951
Q9UQ80	LVVTKYKMGG	1110.61	1110.60	-0.01	0.001593625
Q9UQ80	VAQGTQVTGRKA	1214.67	1214.66	-0.01	0.002276608
Q9UQ80	GEKTIIQNPT	1099.59	1101.58	2.00	0.022452989
Q9UQ80	AMPFTLRAFE	1181.59	1181.58	-0.01	0.007002188
O43497	MHTLLLSALESNMQPHPTELPGPD	2659.26	2658.38	-0.88	0.015837821
Q5TAT6	DPGMTGPTGAAGLPGLHGPPG	1837.88	1836.93	-0.94	0.060606061
Q99460	VMAKFGAILAQGIL	1430.83	1432.80	1.97	0.003157895
P33993	PGVAKSQLLSYI	1274.72	1274.72	0.00	0.002276608
P33993	RLAPRSQYTTGRGSSGVGLTAAVLR	2573.42	2573.40	-0.01	0.074812968
P33993	QPMVPESLADYITAAYVEMR	2299.09	2304.39	5.30	0.096863835
P43004	MASTEGANNMPKQVEVRMH	2160.97	2159.36	-1.61	0.018962963
Q01118	GGSKEKIKQSSSSECSTVD	1937.90	1939.11	1.21	0.064587481
Q6ZN04	DLERNGSGGGGGGSSGGGETL	1801.78	1802.98	1.19	0.068301226
Q15369	AMYVKLISS	1010.55	1010.55	0.00	0.016024653

Q9UHP3	DTQILQQALK	1156.65	1156.58	-0.06	0.020057307
P42126	VQNFVSFISK	1167.63	1167.63	0.00	0.012337217
Q15424	SGAAGAAALSSASSETG TRRLS	2007.00	2006.99	-0.01	0.016024653
P02538	EINKRTAAENEFVTLKK	1972.07	1972.07	0.00	0.028343211
P02538	EINKRTAAENEFVTLKK D	2087.10	2088.11	1.01	0.02661101
P02538	AAYMNKVELQAKA	1435.75	1435.74	-0.01	0.001593625
P02538	EINFLRALY	1119.61	1119.60	0.00	0.031793343
P02538	AELSQMQTHIS	1243.59	1243.58	0.00	0.001593625
P02538	SIIAEVKAQYE	1249.66	1250.66	1.00	0.001593625
P02538	AKNKLEGLE	1000.56	1000.55	0.00	0.070288918
P02538	YQELMNVKLAL	1320.71	1320.71	0.00	0.00757257
Q9NRR4	STVVGTSRLR	1074.61	1074.61	-0.01	0.035242291
Q01105	ETSEKEQQEAIEHI	1651.77	1651.76	-0.01	0.001593625
Q01105	EEALHYLTRVEVTEFE	1945.94	1945.94	0.00	0.001593625
Q01105	ENPYFENKVLSKEFHLN ESG	2362.12	2363.11	0.99	0.001593625
P63208	VEIAKQSVTIKTMLE	1688.94	1689.92	0.99	0.010505051
P62995	LREVFSKYGPIA	1378.76	1378.76	0.00	0.003775366
Q9H6D7	LLQNPYFSKLLLNLSQH VD	2223.20	2223.19	-0.01	0.042135504
P13637	IAARLNIPVSQVNPR	1646.96	1647.94	0.98	0.024896266
P50991	PENVAPRSGATAGAAG GRGKGAYQ	2242.12	2242.11	-0.01	0.074812968
P50991	AIRTSLGPKGM	1145.62	1145.62	-0.01	0.003157895
P20929	MLQVTQAKKSQAIASD	1717.90	1717.98	0.08	0.074812968
P06576	GTEGLVRGQKVL	1255.72	1255.72	-0.01	0.018962963
P06576	NIFRFTQAGSEVSALLG RIPSAVGYQPTLAT	3263.73	3265.72	1.99	0.012337217
P06576	PAPATTFAHL	1024.53	1024.53	0.00	0.00757257
P06576	ATTVLSRAIAELGIYPAV	1844.04	1844.03	-0.01	0.001593625
P06576	PNIVGSEHY	1014.48	1014.47	-0.01	0.003775366
P06576	VARGVQKILQ	1110.69	1110.68	-0.01	0.02661101
P06681	TICGVGNMSANASD	1336.54	1336.78	0.24	0.069489685
Q9BPZ3	PSRSSTSPSIINE	1373.68	1373.67	0.00	0.057481752
Q9BPZ3	LVVKSNLNPNAKEFVPG VKYGNI	2499.38	2500.39	1.01	0.003775366
P40429	KYTEVLKTHGLLV	1499.87	1499.87	0.00	0.003775366
Q96QP1	GAGPTFK	676.35	676.47	0.12	0.092548077
Q2M3C7	DYYAGKNASSILNSAM QQACRKS	2487.16	2487.82	0.66	0.061925062
P08107	AAKNQVALNPQNTVF	1613.85	1613.84	-0.01	0.002276608
P08107	AGVIAGLNVLRIINEPTA AAIAYGL	2479.42	2479.39	-0.02	0.001593625
P08107	PVEKALR	811.49	811.49	0.00	0.003157895
P08107	PVEKALRD	926.52	926.52	0.00	0.010505051
P08107	LVLVGGSTRIPKVQKLL Q	1948.22	1948.21	-0.01	0.004582951
P08107	NQPGVLIQVYEGERAM TK	2032.04	2034.03	1.99	0.001593625
P08107	IERMVQEAEKYKAE	1722.86	1722.88	0.02	0.007002188

P08107	EVQRERVSAKNALESY AFNMKSAVE	2837.41	2838.40	0.99	0.028593311
O43592	QQKANVEAIMLAVMKK LTY	2161.16	2163.15	1.98	0.001593625
Q08J23	PLFPPIEKFYAL	1433.80	1434.82	1.02	0.002276608
O15014	RGGCGVVGGGGSCSSV GGASGGERSV	2194.98	2196.21	1.23	0.062583818
O15014	RGGCGVVGGGGSCSSV GGASGGERSVD	2310.01	2310.94	0.93	0.00757257
A4D1E1	HQLQSDR	864.42	865.01	0.59	0.084638861
Q09028	LVMTHALEWPSLTAQW LP	2092.08	2093.10	1.02	0.0466541
Q09028	FSIHRLVLGTHTS	1466.80	1467.80	1.00	0.059104342
Q09028	AKTIFTGHTAVVE	1372.74	1372.73	-0.01	0.003157895
Q96KC9	DAEISITSEVSGTLK	1548.79	1547.96	-0.83	0.068301226
Q9Y4E6	DKQGSEEGLAMTTSISL QEAF	2223.04	2221.23	-1.81	0.031793343
Q99729	AGKMFVGGLSW	1151.58	1151.57	-0.01	0.001593625
Q99729	AASVEKVL	815.48	815.47	0.00	0.004582951
O14933	LQKKPPPYLRNLSS	1639.94	1639.93	-0.01	0.040106306
Q7LDG7	PARTRLNGAKMKQLFSI LEELAMVTSLRPPVQAN P	3908.11	3909.07	0.96	0.029101856
P62837	ALKRIHKELNDLARD	1791.01	1792.00	0.99	0.074812968
P62837	PLVPEIARIYKT	1398.82	1398.82	0.00	0.001593625
Q7Z6J2	DSEVAPAAPVPTPGPPA AAATPGPPA	2305.16	2305.09	-0.07	0.074812968
Q9BX82	MNVEVVKVMPQ	1304.65	1305.66	1.02	0.099189883
P61758	EAQALLEKNLSTATKNL	1824.99	1824.99	0.00	0.001593625
Q9UJZ1	PEYAVTQLAQTTMRSEL GKLSL	2435.27	2436.28	1.01	0.02661101
P60660	VMRALGQNPTNAEVLK VLGNPKS	2435.33	2435.33	0.00	0.001593625
P60660	FEHFLPMLQTVAKNK	1801.95	1801.95	-0.01	0.001593625
P60660	YVEGLRVF	981.53	981.53	0.00	0.002276608
P17066	PVEKALR	811.49	811.49	0.00	0.003157895
P17066	PVEKALRD	926.52	926.52	0.00	0.010505051
P17066	FFNGKELNKSINP	1506.78	1506.78	0.00	0.016024653
P17066	VAPLSLGLETAGGVMT TLIQRNATIPTKQTQTFT TYS	3909.06	3910.12	1.07	0.012721794
Q16777	VGAGAPVYMAAVLEYL TAEILELAGNAAR	2948.53	2950.53	2.00	0.074812968
Q7Z5K2	DVKLEFFGFE	1211.59	1211.69	0.10	0.020057307
Q96AE4	YGYGGQKRPLE	1266.64	1266.63	0.00	0.002276608
P05161	RVPLASQGLGPGSTVLL VV	1862.10	1863.10	1.00	0.01536
P62913	TGNFGFGIQEHI	1318.63	1318.63	-0.01	0.001593625
P62913	FYVVLGRPGFSIA	1424.78	1424.78	0.00	0.001593625
Q14697	PPIARLSVSGR	1151.68	1151.70	0.02	0.00757257
Q14697	EPGAWEETFKTHS	1499.67	1499.66	-0.01	0.00757257
Q14697	SKPYGPMSVGL	1134.57	1134.58	0.01	0.001593625
Q14697	VFQYELYNPMALYGSV PVLLAHNPHR	3027.54	3029.54	1.99	0.017996401

Q14697	ISSNTAGKTLFGKMM	1584.80	1584.81	0.01	0.005398111
Q14697	VRWMSETGII	1190.61	1190.64	0.02	0.008347245
Q14697	SGYRVHEELRNLGLYV KTR	2289.23	2289.26	0.02	0.020057307
Q14697	MNEPSVFNGPEVTMLK	1791.85	1791.85	-0.01	0.001593625
Q14697	AQHYGGWEHR	1239.55	1239.55	0.00	0.001593625
Q14697	ALLVHPVS	834.50	834.50	0.00	0.003775366
Q14697	SGAHGVQVYLPGQGEV WY	1945.93	1946.93	1.00	0.001593625
Q14697	PETSVLVLRKPGINVAS	1779.03	1779.02	-0.01	0.001593625
Q86UH8	LLLGHPTAVGCAKGTD	1551.81	1551.96	0.16	0.001593625
Q9ULD6	GLNSGDHSDSAKSVSSL NPVK	2080.02	2085.10	5.08	0.074812968
P06889	YYCQAWNSSSVLFGGG TKLTVLG	2450.19	2450.21	0.01	0.026450742
Q16576	TRSNTTSKPSHLV	1426.75	1426.75	0.00	0.090652699
P22626	AAMAARPHSI	1023.53	1023.52	0.00	0.002276608
P22626	YFEEYGKI	1047.49	1047.49	0.00	0.001593625
P22626	SRGGGGNFGPGPGSNFR GGS	1820.83	1820.83	0.00	0.001593625
P68104	FIKNMITGTSQA	1309.67	1309.66	-0.01	0.001593625
P68104	KPLRLPLQ	963.62	963.62	0.00	0.003157895
P68104	NVGFNVKNVSVK	1303.72	1303.71	-0.01	0.001593625
P68104	GPKFLKSG	832.48	832.48	0.00	0.003157895
P68104	GPKFLKSGD	947.51	947.51	0.00	0.003157895
P68104	MRQTVAVGVIKAV	1370.81	1370.83	0.02	0.001593625
P68104	KKAAGAGKVTKSAQKA QKAK	1998.21	1998.22	0.01	0
Q7Z7F0	DERNGSGTLTGSH	1329.59	1328.73	-0.86	0.099189883
Q9UHD9	QLVLIFAGKILK	1324.85	1324.85	0.00	0.001593625
Q9UHD9	INAAIERLLGSQPS	1467.80	1468.81	1.01	0.004582951
Q86WA8	LSFVTASCLD	1036.49	1036.49	0.00	0.006197432
A6NIZ1	TAGTEQFTAMR	1211.56	1211.56	0.00	0.001593625
Q96K17	SLTSLRKLAEQFPRQVL	1985.14	1985.14	-0.01	0.012337217
P55957	LATALEQLLQAYPR	1585.88	1585.89	0.01	0.012721794
Q9NP62	FNSYVQSPAYHSPQED	1867.80	1866.04	-1.77	0.070288918
P08238	SGKELKID	888.49	888.51	0.01	0.006197432
P08238	IIPNPQERTLTLV	1492.86	1492.86	0.00	0.003775366
P08238	TGIGMTKAD	892.43	892.43	0.00	0
P08238	LINNLGTIAKSGTKAFM EALQAGA	2418.29	2419.31	1.01	0.001593625
P08238	KKKKTKKIKEKYI	1662.09	1663.09	1.00	0.039757576
P08238	QEELNKTKPIWTRNP	1835.95	1835.94	-0.01	0.001593625
P08238	QEELNKTKPIWTRNPD	1950.98	1950.97	-0.01	0.001593625
P08238	QEELNKTKPIWTRNPDD	2048.00	2047.99	-0.01	0.074812968
P08238	KENYKKFYEAFSKNLK LGIHE	2585.36	2585.35	-0.02	0.016024653
P08238	ASRMEEV	820.37	820.39	0.01	0.061725614
Q92882	IVQLLLAKGART	1281.81	1283.74	1.93	0.073160357
O43516	DKPKGAGAGGGGGGFG GGGGFGGGGGGGGGS	3060.38	3058.61	-1.77	0.073160357

	FGGGGPPGL				
O43516	KPKGAGAGGGGGGFGGGGGFGGGGGGGGGGSFGGGGPPGL	2945.36	2943.73	-1.63	0.017501509
P25788	IVKEVAKIIYIVH	1523.94	1523.94	0.00	0.002276608
Q6IC83	MGSKLTCCLGPSGGLNC	1655.71	1654.92	-0.79	0.072407883
P51991	GRVVEPKRAVSRE	1481.84	1481.84	0.00	0.02661101
P51991	SVKPGAHLTVKKIFVGGIKE	2107.25	2108.25	1.00	0.003775366
P62937	GEPLGRVSFELFA	1420.74	1420.73	0.00	0.001593625
P62937	FTRHNGTGGKSIYGEKFE	2026.99	2026.98	-0.01	0.001593625
Q6NVY1	FHEGVRAVLIDKD	1497.79	1497.91	0.11	0.092548077
O60814	TGISSKAMGIMNSFVN	1655.80	1657.78	1.98	0.043333333
O60814	TGISSKAMGIMNSFVND	1770.83	1770.85	0.02	0.017501509
P47914	QTKAQAAAPASVPAQAPKRTQAPTKASE	2786.47	2786.46	0.00	0
P62495	QRLQSKVLKLV	1293.81	1293.81	-0.01	0.042135504
Q6UXW0	DLVLGAEEAHGSRL	1465.75	1466.83	1.07	0.005398111
O43895	DVRIWIGTSYTMYGIYEMIPKEKLVT	3103.58	3103.17	-0.41	0.074812968
P25705	VPVGEELLGRVV	1265.73	1265.73	-0.01	0.001593625
P25705	AKLKEIVTNFLAGFEA	1749.97	1750.96	0.99	0.001593625
P23508	MMAAAAAAAAAGSSSSGGGGGGSGSSSSSS	2277.91	2277.17	-0.74	0.055479138
Q6ZS46	DVLLSGRGGGGGGGGGARTGGGEGE	2085.98	2084.17	-1.81	0.047058824
P62805	GVLKVFLENVIR	1385.84	1385.84	0.00	0.096697175
P62805	AVTYTEHAKRKTVTAM	1821.94	1822.93	0.99	0.004582951
P62805	VVYALKRQGRTLYGFGG	1884.04	1884.03	0.00	0.001593625
Q6XQN6	PAFFEHLRAL	1199.65	1199.64	0.00	0.039757576
Q96FL8	AVTLAIAVINVTGVSVGFGLSSAC	2248.21	2248.13	-0.09	0.090652699
P06733	LFTSKGLFRAAVPSGASTGIYEALELR	2853.54	2854.55	1.01	0.001593625
P06733	EGGFAPNILENKEGLELLKTAIGKAGYT	2914.54	2914.53	-0.01	0.001593625
P06733	EGGFAPNILENKEGLELLKTAIGKAGYTD	3029.57	3030.56	0.99	0.001593625
P06733	VAASEFFRSGKY	1360.68	1360.67	-0.01	0.001593625
P06733	VAASEFFRSGKYD	1475.70	1476.69	0.99	0.001593625
P06733	PSRYISPD	933.46	933.45	0.00	0.026450742
P06733	LYKSFIK	897.53	897.53	0.00	0.006197432
P99999	LIAYLKKATNE	1262.72	1262.72	-0.01	0.001593625
P04075	PYQYPALTPEQKKELS	1890.97	1890.96	-0.01	0.084638861
P04075	IAHRIVAPGKGILAA	1485.91	1485.91	-0.01	0.001593625
P04075	IAHRIVAPGKGILAAD	1600.94	1600.93	-0.01	0.001593625
P04075	DGRPFPQVIKSKGGVVGIKV	2080.22	2081.21	0.99	0.031793343
P04075	GRPFPQVIKSKGGVVGIKV	1965.19	1965.18	-0.01	0.001593625

P04075	GRPFPQVIKSKGGVVGIKVD	2080.22	2080.21	-0.01	0
P04075	KGVVPLAGTNGETTTQGL	1741.92	1742.92	1.00	0.001593625
P60981	PFKHFVGMLPEK	1428.76	1428.76	0.00	0.052436195
P07910	PRSMNSRVFIGNLNTLVVKKS	2359.32	2360.29	0.97	0.026450742
P61586	EHTRRELAKMKQEPVKPEEGR	2529.32	2530.32	1.00	0.005398111
Q8N6N2	SLTGTPSGGGGMGHEGRGQSGELGD	2297.99	2298.89	0.90	0.032956222
P14618	PILYRPVAVAL	1210.74	1210.74	0.00	0.001593625
P14618	PILYRPVAVALD	1325.77	1325.76	-0.01	0.010505051
P14618	TKGPEIRTGLIKGSGTAEVELKKGATLKITL	3208.88	3209.89	1.02	0.001593625
P14618	DGLISLQVKQKGA	1355.78	1355.77	0.00	0.002276608
P14618	GLISLQVKQKGA	1240.75	1240.74	-0.01	0.001593625
P14618	FLVTEVENGGSLGSKKGVNLPGAAV	2442.31	2443.34	1.03	0.001593625
P14618	FLVTEVENGGSLGSKKGVNLPGAAVD	2557.34	2557.35	0.01	0
P14618	MVFASFIRKAS	1255.67	1255.67	-0.01	0.001593625
P14618	MVFASFIRKASD	1370.70	1370.69	-0.01	0.003157895
P14618	LRVNFAMNVGKARGFFKKGD	2236.20	2237.20	0.99	0.07211848
P14618	VVIVLTGWRPGSGFTNTMRVVPVP	2581.42	2581.41	-0.01	0
Q92945	RGGGGPGGGGPGGGSAGGPSQPPGGGGPGIRK	2553.25	2553.25	-0.01	0.001593625
Q92945	QPESKKLASQG	1154.59	1154.59	0.00	0.002276608
Q9HCE6	AHQPGAERNLLYED	1593.75	1592.92	-0.83	0.026450742
Q14031	GPSGLPGPPGALGD	1190.59	1190.98	0.39	0.02661101
P63092	LLAEKVLAGKSKIED	1612.94	1613.01	0.07	0.096697175
Q5TF58	MALAFGCPPGGGGGGCPGGGGGGGGAGPGPSPVTAALRD	3193.45	3194.64	1.19	0.084638861
P06748	ENEHQLSLRTVSLGAGAK	1890.99	1890.99	0.00	0.001593625
P06748	ENEHQLSLRTVSLGAGAKD	2006.02	2006.01	-0.01	0.001593625
P06748	VKLLSISGKRSAPGGGSKVPQKKVKLAA	2803.71	2804.73	1.01	0.002276608
P06748	EEAEEKAPVKKSIR	1594.87	1595.86	0.99	0.001593625
P06748	QEAIQDLWQWRKSL	1764.89	1764.89	0.00	0.078844169
Q9BV86	EIYHVYSFALR	1378.70	1378.70	0.00	0
P39023	PSKPVHLTAFLGYKAGMTHIVREV	2650.44	2650.42	-0.03	0.063654574
P37802	VGRPQPGRENFQNWLK	1925.00	1925.00	-0.01	0.049226442
P37802	PNWFPKKSKENPRNFS	1975.01	1976.00	0.99	0.015837821
Q9Y266	PEINTKKINPENSKLS	1810.98	1810.97	-0.01	0.020057307
Q9Y266	EQKKQEILKKFM	1530.86	1530.85	-0.01	0.00757257
P13693	GVTPYMIFFK	1201.62	1201.61	-0.01	0.02661101
Q96AB3	SGLLGLFQGQNSLLH	1582.85	1583.84	0.99	0
P08865	VLKFLAAGTHLGGTNL	1610.91	1610.91	0.00	0.001593625
Q6PCT2	TKPGQTESRGRLQGVAE	2779.53	2779.42	-0.11	0.052436195

	LRLAGLELT				
O60888	SEVLMMIKTQSSLVPAL T	1963.04	1963.03	0.00	0.015837821
Q9BZZ5	SAEFNLVNNALLSIFKM	1910.00	1911.00	1.00	0.001593625
Q9BZZ5	EVLTKEVEELILTESKK VLE	2310.29	2311.29	0.99	0.002276608
P28070	FQYLKQVLGQMVI	1565.86	1565.87	0.01	0.063943162
O60422	GGGGGGGGGAGGAGG AGSAGGGAD	1526.61	1524.87	-1.74	0.001593625
Q14549	MQRAGGGSAPGGNGG GGGGGPGTAFSID	2346.04	2347.22	1.18	0.031793343
P19338	VKLAKAGKNQG	1112.67	1112.66	0.00	0.07975334
P19338	EIEPAAMKAAAAAPASE	1608.78	1610.76	1.98	0.002276608
P19338	SEEEAMETTPAKGKKA AKVVPVKAKNVAE	3039.63	3040.63	1.00	0.003775366
P19338	ARTLLAKNLPYKVTQ	1715.01	1715.02	0.01	0.012337217
P19338	AAEIRLVSK	985.59	985.59	0.00	0.001593625
P19338	GKSKGIAYIEFKTEA	1640.88	1640.86	-0.01	0.016024653
P19338	AEKTFEEKQGTEI	1508.74	1509.75	1.01	0.003157895
P19338	GRSISLYYTGEKGQNQ	1799.88	1799.87	-0.01	0.002276608
P19338	TTEETLKESF	1183.56	1183.56	-0.01	0.001593625
P19338	RETGSSKGFGFV	1270.63	1270.64	0.01	0.001593625
P55327	VTATSAYKKTSETLSQA GQKASAAFSSVGSVITK KLE	3774.01	3775.02	1.02	0.012337217
O75061	LFGGGGAAGPTQAGQS GVED	1756.80	1755.88	-0.92	0.04361223
B4DJU8	KIQLHLT	851.52	851.52	-0.01	0.072407883
Q5SXX8	DTGPRAMGSLSRPSFS	1662.78	1662.91	0.13	0.026450742
Q8IWT0	EFFIPREVKVLSI	1557.89	1559.88	1.99	0.002276608
Q5HYB6	EKKMELQEIQLKEAKHI AEEA	2494.31	2494.21	-0.10	0.019836639
Q53G59	GVWYSVAPMNVRRGL AGATTLGD	2406.21	2405.70	-0.51	0.093906094
P62826	RKVKAKSIVFHRKKNL QYY	2405.42	2406.41	1.00	0.026450742
P62826	PNLEFVAMPALAPPEVV M	1923.98	1923.98	0.00	0.002276608
P62826	PALAAQYEH	998.48	998.48	0.00	0.001593625
P62826	PALAAQYEHD	1113.51	1113.51	0.00	0.001593625
Q5ST30	DAEVVVGTTRPETLPG	1639.84	1640.81	0.97	0.032956222
Q9Y5Y2	FIQEFPGSPAFAALTSIA QKIL	2348.28	2349.29	1.01	0.001593625
P0C7M2	ESLRSHFEQWGTLT	1671.80	1671.80	0.00	0.052436195
P0C7M2	GRVVEPKRAVSRE	1481.84	1481.84	0.00	0.02661101
P0C7M2	SQRPGAHLTVKKIFVGG IKE	2164.25	2164.14	-0.11	0.096863835
Q15084	VIELTPSNFNREVIQS	1844.96	1844.96	0.00	0.003775366
Q15084	AALSALRQLVK	1168.73	1168.72	-0.01	0.002276608
Q15084	RLGGRSGGYSSGKQGR S	1708.87	1708.87	0.00	0.070288918
P22314	SNGEQPLSAMVSMVTK	1677.81	1677.84	0.03	0.001593625
P22314	AAELVALAQAVNARAL PAVQQNNL	2444.35	2444.39	0.04	0.001593625

P22314	TAAAAVRQMNPHIRVTSHQNRVGP	2610.37	2610.38	0.01	0.068678915
P22314	VNNPLHL	805.44	805.44	0.00	0.070288918
P22314	FIVAASNLRAENY	1466.75	1467.75	1.00	0.002276608
P23284	EKKKGPKVTVKVYF	1631.98	1631.97	0.00	0.001593625
P23284	VGRVIFGLFGKTVPKTV	1817.09	1817.09	0.00	0.001593625
P23284	GTGGKSIYGERFP	1367.68	1367.68	-0.01	0.001593625
P23284	ENFKLKHYGPGWVSMANAGK	2215.10	2215.09	-0.01	0.001593625
P23284	TNGSQFFITTVKTAWL	1812.94	1812.94	0.00	0.001593625
Q8IVT5	AAAMGEK	676.32	676.47	0.15	0.078844169
P62854	KAIKKFVIRNIVEAAAVR	2025.26	2025.25	-0.01	0.012337217
P62854	RTPPPRFRPAGAAPRPPPKPM	2293.27	2294.28	1.01	0.00757257
A6NKD9	EERAALAATGAASGGGGGGGGAGSRSSI	2313.11	2314.27	1.16	0.069489685
A6NKD9	EERAALAATGAASGGGGGGGGAGSRSSID	2410.12	2408.16	-1.96	0.066725198
Q9BYG9	ENEHQLSLRTVSLGAGAK	1890.99	1890.99	0.00	0.001593625
Q9BYG9	ENEHQLSLRTVSLGAGAKD	2006.02	2006.01	-0.01	0.001593625
Q9BYG9	VKLLSISGKRSAPGGGSKVPQKKVKLAA	2803.71	2804.73	1.01	0.002276608
Q9BYG9	EEAEEKAPVKKSIR	1594.87	1595.86	0.99	0.001593625
Q9BYG9	IKAKMQASIEKAH	1453.81	1453.80	0.00	0
P52739	HLVIHTGD	890.46	890.50	0.04	0.02661101
Q6PIQ7	RFSGSKSGNTASLTISGLQAE	2110.07	2110.06	0.00	0.001593625
Q6PIQ7	FYPGAVTVAWKA	1308.69	1308.68	0.00	0.001593625
Q8TAF3	SIYSLAMNQLGTIIVSGSTEKVLRVWD	2995.57	2993.98	-1.59	0.042135504
P62857	VLTLLESEREARRLR	1840.06	1840.06	-0.01	0.008347245
A5A3E0	ESGPSIVHRKCL	1306.68	1306.67	-0.01	0.012337217
Q9Y262	QKVYELQASRVSS	1476.76	1477.75	0.99	0.012337217
P02768	ETYVPKEFNAETFTFHA	2011.93	2012.93	1.00	0.001593625
O75083	EGKLLEAKGPVT	1222.69	1223.68	0.99	0.082482646
	DGKGSASGQGSC	1052.42	1050.57	-1.85	0.042135504
Q9H299	EMRALAGNPKATPPQIVNG	1945.02	1946.04	1.02	0.001593625
P49459	ILQNRWSPTY	1276.66	1277.66	1.00	0.017996401
B4E2M5	MTKLHQAVAAG	1141.59	1142.64	1.04	0.01536
P17542	PVVGAGGGGGGGGGGAPP	1276.62	1275.78	-0.83	0.087221095
Q65ZC9	PWGKGTLVTVSSGGGGSGGGGSGGGGSD	2273.03	2273.28	0.25	0.073160357
A8MUW5	DPYGGGGGGGGGGGGGGGYRRY	1886.80	1884.95	-1.85	0.07975334
Q92614	LESLEAANQSLQAD	1469.70	1469.76	0.06	0.02661101
P41250	TVNKTPHTATLR	1337.74	1337.74	0.00	0.028343211
O75179	DNSGGGGGGGGGGGGGGTSSNNSEEEE	2295.83	2294.27	-1.57	0.078844169
O75179	NSGGGGGGGGGGGGG	2277.82	2277.17	-0.65	0.092548077

	GGTSSNNSEEEED				
P50990	TGANVVVTGGKVA	1171.66	1171.65	0.00	0.001593625
P50990	GVNTFKVLTR	1133.66	1133.65	-0.01	0.001593625
P50990	QIIMAKPAGGPKPPSGK K	1787.01	1787.01	0.00	0.004582951
P10809	ARALMLQGV	957.54	957.54	0.00	0.003157895
P10809	AVAVTMGPKGRTVIIEQ SWGSPKVTK	2739.51	2740.52	1.01	0.001593625
P10809	GVTVAKSID	888.49	888.49	0.00	0.003157895
P10809	KYKNIGAKLVQ	1260.76	1260.75	-0.01	0.001593625
P10809	GTTTATVLARSIAKEGF EKISKGANPVEIRRGVM LAV	3870.15	3871.13	0.98	0.012337217
P10809	AVIAELKKQSKPVTTPE EIAQVATISANG	2992.64	2992.64	-0.01	0.001593625
P10809	KEIGNIIS	872.50	872.49	0.00	0.003775366
P10809	AMKKVGRKGVITVK	1513.95	1513.94	-0.01	0.001593625
P10809	AYVLLSEKKISSIQSIVP ALEIANAHRKPLVIIAE	3813.21	3814.23	1.02	0.003775366
P10809	GEALSTLVLNRLKVGLQ VVAVKAPGFG	2735.61	2736.61	1.00	0.001593625
P10809	GEALSTLVLNRLKVGLQ VVAVKAPGFGD	2850.63	2852.62	1.99	0.003157895
P10809	LGKVGEVIVTK	1141.71	1141.70	0.00	0.001593625
P10809	LGKVGEVIVTKD	1256.73	1256.73	0.00	0
P10809	LGKVGEVIVTKDD	1353.75	1352.72	-1.03	0.074812968
P10809	AMLLKGKG	816.49	816.49	0.00	0.00757257
P10809	KAQIEKRIQEIIEQL	1838.06	1838.05	-0.01	0.001593625
P10809	VTTSEYEKEKLNERLAK LS	2237.19	2237.18	-0.01	0.002276608
P10809	GVAVLKVGGTS	986.58	986.57	0.00	0.001593625
P10809	GVAVLKVGGTSD	1101.60	1101.60	-0.01	0.002276608
P10809	FVNMVEKGII	1148.63	1148.62	0.00	0.003157895
P10809	PTKVVRTALL	1096.70	1096.69	-0.01	0.003775366
P10809	PTKVVRTALLD	1211.72	1211.72	0.00	0
P10809	AAGVASLLTTAEVVVT EIPKEEK	2354.29	2354.26	-0.04	0.001593625
P10809	DPGMGAMGGMGGGM GGGMF	1705.60	1707.43	1.82	0.038527188
P10809	PGMGAMGGMGGGMG GGMF	1606.57	1605.93	-0.64	0.096697175
P07900	SGKELHINLIPNKQ	1589.89	1589.89	0.00	0.001593625
P07900	TGIGMTKAD	892.43	892.43	0.00	0
P07900	LINNLGTIAKSGTKAFM EALQAGA	2418.29	2419.31	1.01	0.001593625
P07900	TGEPMGRGTKVILHLKE	1865.02	1865.01	-0.01	0.029101856
P07900	QEELNKTKPIWTRNP	1835.95	1835.94	-0.01	0.001593625
P07900	QEELNKTKPIWTRNPD	1950.98	1950.97	-0.01	0.001593625
P07900	QEELNKTKPIWTRNPDD	2048.00	2047.99	-0.01	0.074812968
Q08828	MAGAPRGGGGGGGGA GEPGGAER	1897.85	1892.99	-4.86	0.092548077
Q2M3C6	GAVIILSLAPMVASTVA NGPRSPWD	2503.33	2504.73	1.41	0.082482646
Q14D33	ITEGKEKEGGLVTAGHD	1739.87	1740.91	1.04	0.032956222

	TLVKNQRGPK	1139.68	1141.64	1.97	0.003775366
Q8NAP3	DNIQTGVENVVL	1299.67	1297.70	-1.97	0.029101856
P19838	DSFGGGSGAGAGGGGM FGSGGGGGGTGSTGP	2371.94	2370.31	-1.63	0.047058824
P60174	EREAGITEKVVFEQTKV IA	2128.15	2129.17	1.02	0.001593625
P60174	EREAGITEKVVFEQTKV IAD	2243.18	2243.18	0.00	0.005398111
P60174	GFLVGGASLKPEFV	1419.78	1421.76	1.99	0.001593625
P60174	GFLVGGASLKPEFVD	1534.80	1535.81	1.00	0
Q8N2G4	MQAPRAAPAAPLSYD	1557.76	1556.90	-0.86	0.017996401
P06454	LKEKKEVVEEAENGR	1756.93	1756.93	0.00	0.001593625
P06454	EEAESATGKRAAE	1329.62	1330.61	0.99	0.001593625
Q63HN8	AAPLEAASVPSAD	1179.58	1179.63	0.05	0.028593311
A2J423	IWGQGTMVTVSSGGGG TGGGGSGGGGS	2239.98	2241.16	1.19	0.090652699
A6NL28	EEKMELQEIQLKEAKHI AEEA	2477.25	2478.25	1.00	0.001593625
Q13813	QLIAAGHYAKG	1110.58	1110.57	-0.01	0.001593625
Q08211	FVNYLVRINEIKSEEVPA FGVASPPPLT	3085.65	3086.60	0.95	0.001593625
Q08211	AVIEAEHTLREL	1379.74	1380.74	1.00	0.003157895
Q8N554	SALAVKWPRD	1123.61	1123.49	-0.12	0.062583818
O15050	WSTQEIEACLQ	1306.59	1306.65	0.07	0.020577308
Q13838	FLLKPELLRAIV	1410.90	1410.89	-0.01	0.007002188
Q13838	ISSYIEQTR	1095.56	1095.55	-0.01	0
P23468	SSIPNNKEIPSHHPTDPV ELR	2366.20	2366.12	-0.07	0.082482646
Q9P1F3	GKLSVKFGVLFRD	1446.83	1446.77	-0.07	0.070288918
Q6IS14	IFTGKKYE	984.53	984.52	-0.01	0.004582951
Q6IS14	LGKEIEQKY	1106.60	1106.59	0.00	0.001593625
P32004	QNNVTIMANLKVKD	1568.83	1568.82	-0.01	0.039757576
P32004	GSFIGQYSGKKEKEAAG GND	2023.96	2024.07	0.11	0.022133939
A8MYX 2	MSQSVLAGGGIPEPHLG CPGTYR	2326.12	2321.21	-4.91	0.096863835
Q6IMJ6	AFIHVHTHTHTCTHRGA D	2039.95	2041.05	1.10	0.093906094
P62942	YAYGATGHPGIIPPHAT LVF	2081.07	2082.08	1.01	0.002276608
P62942	VELLKLE	842.51	842.51	0.00	0.001593625
P59510	ICVQGQCMAAGCD	1297.49	1298.72	1.22	0.004582951
P59510	CELRGMLLAKKSECSSQ CGQGYRTLD	2875.34	2875.43	0.09	0.082482646
Q8NER1	MLLVEPLNRLLQD	1552.86	1552.93	0.06	0.024608501
Q99757	RVVNSETPVVV	1197.67	1197.67	0.00	0.028593311
Q99757	LAIEYEVSAVPTVLAMK NG	2004.06	2004.07	0.01	0.001593625
Q99757	QLEAFLKKLIG	1241.74	1241.73	0.00	0
Q5UCC4	DTGGQGGGGGGGGGG GSGLCCVPPSL	2041.86	2041.03	-0.83	0.04361223
Q00839	SAGRSGAGLEQEAAAG G	1487.70	1487.70	0.00	0.002276608

	FIMKTPKAIAAKAIID	1730.02	1729.88	-0.13	0.093906094
Q66K41	GAVVSPIPTLASGAPGEPQSKAVPAAPPLGPPLQPPPTPD	3757.01	3756.26	-0.75	0.099189883
Q2VWA4	AGAAAEALGGAGAGGAGAAPKAGLSGLFWPAGRKD	3032.56	3031.57	-0.99	0.057481752
	MGAGVGSSVAW	1036.46	1036.52	0.06	0.02661101
Q3SY00	CGSHQEPAAGLGAAAAGGAAAGRAAAGPD	2432.13	2431.33	-0.80	0.099189883
A2AB27	MGERVLWPLRAGQHWRHRPLPAGLQD	3057.61	3057.30	-0.31	0.024896266
Q9BY44	LAPTPAPQSTPRNTVSQSISG	2108.09	2109.07	0.98	0.001593625
B4DWP8	MLLSGAPPAAPPAPPP	1498.79	1500.77	1.99	0.068301226
B4DX20	VLRTNLGPKGTMKMLVSGAG	2029.12	2030.11	0.99	0.020057307
B4DX20	GTTSNVLIIGELLKQA	1655.95	1655.95	0.00	0.004582951
B4DX20	VLTEVVVDSVLAVRRPGYPID	2260.26	2261.14	0.88	0.087221095
B4DX20	KGFVVINQKGI	1201.72	1201.72	0.00	0.043333333
B4DNX1	AAKNQVALNPQNTVF	1613.85	1613.84	-0.01	0.002276608
B4DNX1	LVLVGGSTRIPKVQKLLQ	1948.22	1948.21	-0.01	0.004582951
B4DNX1	NQPGVLIQVYEGERAMTK	2032.04	2034.03	1.99	0.001593625
B4DNX1	IERMVQEAEKYKAE	1722.86	1722.88	0.02	0.007002188
B4DNX1	EVQRERVSAKNALESYAFNMKSAVE	2837.41	2838.40	0.99	0.028593311
B4DNX1	SRRGPGPGGFGAQGPKGGSGSGPTIEEV	2595.28	2594.16	-1.12	0.090652699
Q96GF2	PGSGGGGGGGGGGSSSGSSSSD	1649.61	1647.90	-1.72	0.092548077
Q96GF2	AGKSTIGGQIMYLTGMV	1725.88	1725.91	0.03	0.012337217
Q96GF2	EGKTIAIGKVLKLVPEK	1804.12	1804.13	0.01	0.002276608
P07437	RISVYYNEATGGKYVPRAILV	2368.29	2368.28	-0.01	0.001593625
P07437	SVRSGPFGQIFRP	1446.77	1446.76	-0.01	0.003775366
P07437	RIMNTFSVVPSPKVS	1660.90	1660.89	0.00	0.001593625
P07437	TVVEPYNATLSVHQLVENT	2113.07	2115.05	1.98	0.001593625
B4DUA0	IIVNWVNETLREAKKSSSISSFK	2635.43	2637.37	1.94	0.016024653
B4DUA0	PKISTSLPVL	1053.64	1053.64	0.00	0.003775366
Q14934	LGGPGGGAGGAGGGRVLECPSIR	1994.01	1995.02	1.01	0.074812968
Q6UB99	DSQHSTPVPTAPTSACSPSFF	2144.95	2145.50	0.56	0.078844169
P40925	LTAKELTEEKESAFEFLSSA	2229.11	2230.09	0.99	0
P20839	PVVLSPSHTVG	1091.60	1091.59	-0.01	0.002276608
Q9BTE6	PGSQVLVRV	953.57	954.56	1.00	0.012337217
	DEVPVQLFTGAQMESRGGCKSHRQ	2657.24	2658.82	1.58	0.061925062
B7Z620	GGGHDPR	676.30	676.72	0.42	0.096697175
P47897	TAPRAMAVLESLRVIIT	2668.51	2668.51	0.00	0.006197432

	NFPAAKSL				
P47897	ETKGFHQVPFAPIVFIER T	2197.17	2198.16	0.99	0.012337217
P68363	DSFNTFFSETGAGKHVP RAVFV	2412.19	2413.18	1.00	0.040106306
P68363	SFNTFFSETGAGKHVPR AVFV	2297.16	2297.17	0.01	0.001593625
P68363	EVRTGTYRQLFHPEQLI TGKE	2483.29	2483.28	-0.01	0.073160357
P68363	AANNYARGHYTIGKEII	1889.97	1889.97	-0.01	0.015837821
P68363	VNAAIATIKTKRSIQFV	1859.10	1859.10	0.01	0.001593625
P26641	AYLKTRTFLVGERVTLA	1937.11	1937.10	-0.01	0.003157895
P26641	PFAHLPKSTFVL	1355.76	1355.75	-0.01	0.012337217
P26641	TLSVALPYFWEHF	1608.80	1608.79	0.00	0.001593625
P26641	YESYTWRKL	1244.62	1244.61	-0.01	0.002276608
Q6ZTA4	AAAGPACGGAGGSAAG GLGGGAGGGG	1784.79	1784.81	0.02	0.07975334
P50395	PRTFEGI	818.43	818.43	0.00	0.070288918
P50395	FTGHALALYRT	1248.66	1248.66	0.00	0.016024653
P50395	LGTESQIFISRTY	1513.78	1513.77	-0.01	0.001593625
Q6NUN9	GSGPGTGGGGSGSGGG GGGSGGGSARD	1962.80	1963.21	0.41	0.032956222
O00148	FLLKPELLRAIV	1410.90	1410.89	-0.01	0.007002188
O00148	ISTYIEQSR	1095.56	1095.55	-0.01	0.022133939
P25205	LTTLVAFPSSSVYPTK	1709.92	1709.93	0.00	0.001593625
P25205	QLAKSLAPSIHGH	1340.72	1340.72	0.00	0.00757257
P25205	LENGSHIRG	981.50	982.51	1.01	0.001593625
Q15717	ELRSLFSSIGEVESAKLI R	2115.17	2116.16	0.99	0.001593625
Q15717	QTTGLSRGVAFIRF	1534.83	1535.83	1.00	0.002276608
O00303	LQQVGGASARIQ	1226.67	1226.66	-0.01	0.001593625
P62491	EARAFAEKNGLSFIETS AL	2035.04	2036.03	0.99	0.001593625
P27824	TTAPPSSPKVTYKAPVP TGEVYFA	2507.29	2507.32	0.03	0.003775366
O43852	AFLGAEEAKTF	1182.59	1182.58	-0.01	0.002276608
O43852	FGEALVRH	927.49	927.49	-0.01	0.001593625
P29401	QQKLQALK	938.55	938.56	0.01	0.002276608
P29401	GHPVPKQAFT	1080.57	1080.57	0.00	0.069489685
P29401	PAPLQHQM	920.45	920.45	0.00	0.006197432
P29401	KIATRKAYGQALAKLG HAS	1983.14	1983.11	-0.02	0.002276608
P29401	LAMFRSVPTSTVFYPS	1817.90	1817.90	0.00	0.003157895
P29401	FQVGQAKVVLKSK	1430.86	1430.86	0.00	0.003157895
P29401	QVTVIGAGVTLHEALA AAELLKKEKINIRVL	3279.93	3280.94	1.01	0.012337217
P29401	PFTIKPL	814.50	814.49	0.00	0.029101856
P29401	AIAQAVRGLITKA	1310.80	1310.80	-0.01	0
P14314	RVWAAIR	870.52	870.52	0.00	0.0466541
P14314	AAQEARNKFEEAERSLK	1976.01	1976.00	-0.01	0.01536
P31350	PQQLQLSPLKGLSLV	1619.96	1620.95	0.99	0.028343211
A5XEI0	VQSLIRAYQVRGHHIVK	2213.25	2212.22	-1.04	0.066725198

LD

Appendix Table 2. Table of peptides identified without designating enzymatic specificity after in gel microwave-supported acid hydrolysis.

UniProt	Left AA	Sequence	Right AA	Observed Mass	Theoretical Mass	Mass Diff	E-value
P06733	A	GNSEVILPVPAFNVINGGSHAG	N	2148.09	2148.10	-0.01	0.019
P09104	A	GNSDLILPVPAFNVINGGSHAG	N	2148.09	2148.10	-0.01	0.014
P23528	C	TLAEKLGGSAVISLEGKPL	-	1882.08	1882.08	0.00	0.000034
P02538	D	SIIAEVKAQYE	E	1249.65	1249.66	0.00	0.038
P10809	D	AAGVASLLTTAEVVVTEIPKE	E	2098.15	2097.16	0.99	0.026
P35908	D	SIIAEVKAQYE	E	1249.65	1249.66	0.00	0.038
P23284	D	NFVALATGEKGFGYK	N	1600.82	1600.82	-0.01	0.0013
P32969	D	IELVSNSAALIQQATTVK	N	1886.05	1885.05	1.00	0.00017
P35527	D	LEMQYETLQEELMALKK	N	2097.05	2096.05	0.99	0.0052
P05386	D	KINALIKAAGVNVE	P	1438.85	1438.85	0.00	0.0076
P11021	D	AGTIAGLNVMRIINE	P	1570.85	1570.85	0.00	0.0047
P04264	D	LEIATYRTLLEGEE	S	1635.83	1635.84	0.00	0.047
P30046	E	SWQIGKIGTVMTFL	-	1580.85	1579.84	1.01	0.019
O60361	E	GLNVVKTGRVMLGETNPA	D	1856.00	1855.00	1.00	0.005
P00558	E	PVAVELKSLLGK	D	1252.77	1252.78	-0.01	0.00073
P00558	E	PAKIEAFRASLSKLG	D	1586.91	1586.91	-0.01	0.00045
P06733	E	LLKTAIGKAGYT	D	1234.72	1234.73	-0.01	0.03
P09651	E	PEQLRKLFIGGLSFETT	D	1936.04	1935.05	0.99	0.035
P14625	E	RIMKAQAYQTGK	D	1393.75	1393.75	0.00	0.029
P19338	E	PAAMKAAAAAPASE	D	1255.62	1255.62	0.00	0.0032
P28065	E	PPLVLAAANVVRNISYKYRE	D	2273.27	2272.27	1.00	0.041
P28838	E	PPLVFVGKGITF	D	1273.74	1273.74	0.00	0.01
P55884	E	PPVPAQGEAPGEQAR	D	1503.75	1502.75	1.01	0.0053
P60709	E	GYALPHAILRL	D	1222.71	1222.72	-0.01	0.046
P62249	E	PRTLQYKLLEPVLLLGKERFAGV	D	2639.56	2639.55	0.01	0.0098
P68363	E	PYNSILTTHTTLEHS	D	1712.83	1712.84	-0.01	0.0022
Q14566	E	PGAGSQHLEVR	D	1149.59	1149.59	0.00	0.026
Q14697	E	PSVFNGPEVTMLK	D	1418.73	1417.73	1.01	0.0045
Q15056	E	PPYTAYVGNLPFNTVQG	D	1836.90	1836.90	0.00	0.0031
P62314	E	PVQLETLSIRG	N	1211.69	1211.69	0.00	0.019
Q07021	E	PELTSTPNFVVEVIK	N	1671.90	1671.91	0.00	0.028
P04406	F	GIVEGLMTTVHAITATQKTV	D	2069.12	2069.12	0.00	0.00063
P30049	G	AAKANLEKAQAELVGTA	D	1683.92	1683.92	0.00	0.00088
P35527	G	PAAIQKNYSPYYNTI	D	1741.86	1741.87	0.00	0.016

P60709	G	DGVTHTVPIYEGYALPHAILRL	D	2434.29	2434.30	-0.01	0.021
P60709	H	TVPIYEGYALPHAILRL	D	1925.07	1925.08	-0.01	0.00022
P68104	K	AAGAGKVTKSAQKAQKAK	-	1742.01	1742.02	0.00	0.0095
P27037	L	AFAVFLISC	S	969.51	969.50	0.01	0.047
P14174	M	PMFIVNTNVPRASVP	D	1640.87	1640.87	0.00	0.009
P14174	M	PMFIVNTNVPRA	S	1357.72	1357.72	0.00	0.0097
P00558	N	GAKSVVLMSHLGRP	D	1450.81	1450.81	0.00	0.04
P04264	N	VKKQISNLQQSIS	D	1471.84	1471.84	0.00	0.014
P25398	N	TALQEVLKTALIH	D	1435.84	1435.84	0.00	0.04
P35527	N	HKEEMSQLTGQNSG	D	1544.69	1544.69	0.00	0.031
P60174	N	GRKQSLGELIGTLNAAKVPA	D	2022.16	2022.16	0.00	0.0025
P60174	N	GAFTGEISPGMIK	D	1306.66	1306.66	0.00	0.044
P68104	N	GQTREHALLAYTLGVKQLIVGVNKM	D	2738.52	2738.53	-0.01	0.0000051
P62266	P	GVRFKVVKVANVSLLALYKGKKERPRS	-	3041.83	3041.83	-0.01	0.0029
P06732	P	VSPLLLASGMAR	D	1213.68	1213.69	-0.01	0.031
P07737	P	SVWAAVPGKTFVNITPAEVGVLVGK	D	2538.42	2538.42	-0.01	0.0012
P28838	P	SVVLVGLGKKAAGI	D	1310.82	1310.83	-0.01	0.047
P35527	P	AAIQKNYSPYYNTI	D	1644.81	1644.81	0.00	0.026
P43308	P	VVIGSTSAPGQGGILAQREF	D	1986.04	1986.05	-0.01	0.0044
P61088	P	ALQIRTVLLSIQALLSAPNP	D	2118.26	2117.26	1.01	0.00095
P04908	R	VGAGAPVYLAAVLEYLTAEILELAGNAAR	D	2915.59	2914.58	1.01	0.0000001
P05387	R	YVASYLLAALGGNSSPSAK	D	1867.97	1867.97	0.00	0.00073
P22626	R	GGGGNFGPGPGSNFRGGS	D	1577.69	1577.70	-0.01	0.017
P62805	R	GVLKVFLENVIR	D	1385.84	1385.84	0.00	0.015
Q16777	R	VGAGAPVYMAAVLEYLTAEILELAGNAAR	D	2932.53	2932.54	-0.01	0.0018
P13645	R	ALEESNYELEGK	I	1380.64	1380.64	0.00	0.000018
P04264	R	FLEQQNQVLQTKWELLQQVDTSTR	T	2932.50	2931.51	0.99	0.01
P07737	W	AAVPGKTFVNITPAEVGVLVGK	D	2166.24	2166.24	0.00	0.00021
A0M8Q6	Y	PGAVTVAWKA	D	998.55	998.55	0.00	0.0084
P00558	Y	SLEPVAVELKSLLGK	D	2148.09	2148.10	-0.01	0.0039

Appendix Table 3. All peptides identified in the high mass fraction following mass biased partitioning. Commonly observed mass differences from the theoretically observed mass can be explained by loss of a terminal Asp (~115Da), loss of a terminal Asp and formation of the cyclic intermediate in acid hydrolysis (~133Da), acetylation following loss of the initiator Met (~89Da), acetylation (~42Da), and loss of Asp and oxidation (~97Da).

UniProt	Sequence	Theoretical Mass	Observed Mass	Mass Diff	E Value
Q07666	DPSFTHAMQLLTAEIEKIQKG	2356.21	2241.18	-115.03	5.21E-40
P06733	DPSRYISPDQLADLYKSFIK	2355.21	2240.19	-115.02	1.96E-39
P14618	DYPLEAVRMQHLIAREAEAAIYHLQLFEELRRLAPITS	4462.36	4347.34	-115.02	1.03E-32
P48735	DEMTRIIWQFIKEKLILPHV	2508.39	2375.36	-133.03	1.05E-28
O95861	MASSNTVLMRLVASAYSIAQKAGMIVRRVIAEG	3492.87	3403.84	-89.03	1.39E-28
P42765	MALLRGVFVVAAKRTPFGAYGGLLK	2634.52	2545.49	-89.03	1.95E-28
P10809	DGTTTATVLARSIAKEGFEKISKGANPVEIRRGVMLAV	3985.18	3870.16	-115.02	3.28E-26
P04406	DNEFGYSNRVVDLMAHMASKE	2412.08	2297.06	-115.02	7.91E-24
P09651	DFGNYNNQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGYGGSSSSSSYGSGRRF	6052.68	5937.65	-115.03	2.46E-22
P60709	DLAGRDLTDYLMKILTERGYSFTTTAEREIVR	3732.91	3617.89	-115.02	5.84E-22
P63261	DYLMKILTERGYSFTTTAEREIVR	2891.49	2776.47	-115.02	1.15E-21
P06733	DGTENKSKFGANAILGVSLAVCKAGAVEKGVPLYRHIA	3883.08	3768.05	-115.03	8.33E-20
Q14568	DLINNLGTIAKSGTKAFMEALQAGA	2533.32	2418.3	-115.02	5.79E-19
P60709	DGVTHTVPIYEGYALPHAILRL	2434.3	2319.28	-115.02	1.32E-17
P04406	DSTHGKFHGTVKAENGKLVINGNPITIFQER	3406.77	3291.74	-115.03	1.48E-16
P62906	DELVYNIHLAVNFLVSLLKKNWQNVRALYIKSTMGKPQRLY	4874.68	4741.65	-133.03	4.35E-16
P63261	DLTDYLMKILTERGYSFTTTAEREIVR	3220.64	3105.62	-115.02	7.32E-16
P06576	DEGLPPILNALEVQGRETRLVLEVAQHLGESTVRTIAM	4154.22	4021.19	-133.03	2.79E-15
P31949	MAKISSPTETERCIESLIAVFQKYAGK	2999.55	2910.53	-89.02	3.21E-14
P04406	DAGAEYVVESTGVFTTMEKAGAHLQGGAKRVIISAPSA	3817.93	3702.91	-115.02	3.6E-14
P04075	DESTGSIAKRLQSIGTENTEENRRFYRQLLLTA	3795.95	3662.91	-133.04	4.12E-14
P62805	DAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG	3902.03	3787	-115.03	8.53E-14
P10809	DAVAVTMGPKGRTVIIEQSWGSPKVTK	2854.54	2739.51	-115.03	7.04E-13
Q99497	DGLILTSRGPGTSFEF	4122.31	4007.29	-115.02	2.26E-12

Q5VTE0	ALAIVEALNGKEVAAQVKAPLVLK MGKEKTHINIVVIGHV	1773.99	1685	-88.99	4.6E-11
P29401	DRFVLSKGHAAPILYAVWAEAGFLAEAELLNLRKISS	4055.2	3940.18	-115.02	4.97E-11
Q14697	DPEGHFETPIWIERVVIIGAGKPAAVVLQTKGSPESRLSFQH	4595.43	4480.39	-115.04	8.66E-11
P60174	DEREAGITEKVVFEQTKVIA	2261.19	2128.15	-133.04	1.34E-10
P60709	DGVTHTVPIYEGYALPHAILRLDLAGR	2946.57	2849.48	-97.09	1.72E-10
P10809	DAAGVASLLTTAEVVVTEIPKEEK	2469.32	2354.3	-115.02	2.23E-10
P26640	DPEAEAALELALSITRAVRSLRA	2451.34	2336.31	-115.03	4.21E-10
P06733	MSILKIHAREIF	1456.82	1367.79	-89.03	4.21E-10
P27348	MEKTELIQKAKLAEQAERY	2278.2	2320.21	42.01	1.1E-09
P22626	DKIVLQKYHTINGHNAEVRKALSRQEMQEVQSSRSGRGGNFGFG	4928.5	4813.48	-115.02	1.61E-09
P04406	DPSKIKWGDAGAEYVVESTGVFTTMEKAGAHLQGGAKRVIISAPSA	4729.42	4596.39	-133.03	1.94E-09
Q6PEY2	DSFNTFFSETGAGKHVPRAVFV	2412.19	2297.16	-115.03	2.59E-09
P10809	DAYVLLSEKKISSIQSIVPALEIANAHRKPLVIIAE	3928.24	3813.22	-115.02	4.19E-09
P51991	DFGNYSGQQQSNYGPMKGGSFGGRSSGSPYGGGYGSGGGSGGYGSRRF	4733.04	4618.03	-115.01	7.85E-09
P26641	DPGSEETQTLVREYFSWEGAFQHVGKAFNQGKIFK	4029.96	3914.95	-115.01	1.12E-08
O94964	DSYASEIKELQLVLAEAH	2015.02	1800.03	-214.99	0.00000114
P46060	DEEEEEEEEEEEEEEPQQRGQGEKSATPSRKIL	3901.71	3767.68	-134.03	0.000000129
P60174	DELIGQKVAHALAEGLGVIACIGEKL	2646.44	2513.41	-133.03	0.000061
P29401	DQVTVIGAGVTLHEALAAAELLKKEKINIRVL	3411.98	3279.93	-132.05	6.86E-08
P51665	DQMVVVYLASLIRSVVALHNLINNKIANR	3262.83	3130.79	-132.04	0.0000125
P04406	MGKVKVGVNGFGRIGRLVTRAAFNSGKV	2917.65	2786.62	-131.03	0.00000225
P21284	DVSTQTSTISKKLRLKEMITQTEGMSFD	3173.59	3050.54	-123.05	0.000000258
P14618	DLRVNFAMNVGKARGFFKKG	2254.22	2139.19	-115.03	0.000000356
P23528	DLVFIFWAPESAPLKSKMIYASSK	2727.43	2612.4	-115.03	0.000000756
P14625	DPRGNTLGRGTTITLVLKEEAS	2327.24	2212.21	-115.03	0.0000426

P22626	DFGNYNQQPSNYGPMKSGNFGGSRNMGGPYGGGNYGPGGSGGSGGYGGRSRY	5198.22	5083.2	-115.02	6.66E-08
P07237	DGKLSNFKTAAESFKGKILFIFI	2573.43	2458.41	-115.02	0.00000247
Q07666	DATVGGPAPTPLLPPSATASVKMEPENKYLPELMAEK	3848.96	3733.95	-115.01	0.0000536
Q13098	DVNYVVENPSLDLEQYAASYSGLMRIERLQFIA	3802.89	3703.91	-98.98	0.0000177
P22234	MATAEVLNIGKKLYEGKTKEVYELL	2839.54	2750.52	-89.02	0.0000204

Appendix Table 4. All peptides in the low mass fraction following mass biased partitioning.

UniProt	Sequence	Theoretical Mass	Observed Mass	Mass Diff	Estimated FDR
P46777	IICQIAYARIEG	1348.72	1348.72	0.00	0.00804829
P46777	YMRYLMEE	1133.49	1133.49	0.00	0.015540016
Q16222	ENGVHELVKNGI	1289.67	1289.70	0.02	0.048727666
P62258	LVYQAKLAEQAERY	1680.88	1680.88	0.00	0.021094265
P62258	EMVESMKKVAGM	1320.62	1320.62	0.00	0.00804829
P62258	STLIMQLLR	1089.62	1089.62	0.00	0.013125513
P62258	STLIMQLLRD	1188.65	1188.65	0.00	0.062854199
Q9Y4L1	FNFHINYG	1010.46	1010.46	0.00	0.070858283
Q9Y4L1	RVESVFETLVE	1306.68	1306.67	0.00	0.021094265
Q9Y4L1	AEPISEPEKVETGSEPG	1754.82	1754.82	0.00	0.00804829
P34897	PEMWELLQREK	1457.73	1459.63	1.90	0.081632653
P34897	ALLERGYSLVSGGT	1421.75	1421.75	-0.01	0.00804829
P34932	PAIAQFSVQKVTPQS	1599.86	1599.87	0.01	0.00804829
P11021	QGNRITPSYVAFTPEGERLIG	2287.16	2288.18	1.02	0.00804829
P11021	AAKNQLTSNPENTVF	1632.81	1632.81	0.00	0.00804829
P11021	AAKNQLTSNPENTVFD	1747.84	1747.84	0.00	0
P11021	NGVFEVVATNG	1105.54	1105.54	0.00	0.021094265
P11021	FSETLTRAKFEELNM	1814.89	1814.88	0.00	0.070351759
P11021	LFRSTMKPVQKVLE	1674.95	1675.95	1.00	0.020319303
P11021	VNGILRVTAE	1070.61	1070.61	0.00	0.055818852
P11021	VNGILRVTAED	1185.64	1185.63	0.00	0
P11021	QNRLTPEEIERMVN	1710.84	1711.82	0.99	0.070858283
Q15029	VSISKFF	826.46	826.46	0.00	0.04972973
Q15029	PMLLELAKQ	1041.59	1041.59	0.00	0.00804829
O96019	IGSYTVRAGYAGE	1342.65	1342.64	-0.01	0.00804829
O96019	FPTAIGMVVER	1218.64	1218.64	0.00	0.016768293

P11142	AAKNQVAMNPTNTVF	1604.80	1604.80	0.00	0.00804829
P11142	AAKNQVAMNPTNTVFD	1719.83	1720.85	1.02	0
P11142	KKVGAERNVLIF	1372.82	1372.82	0.00	0.00804829
P11142	FFNGKELNKSINP	1506.78	1506.78	0.00	0.035756853
P11142	FFNGKELNKSINPD	1621.81	1621.81	0.00	0.011100833
P11142	IERMVQEAEKYKAE	1722.86	1722.87	0.01	0.016768293
P11142	KCNEIINWL	1131.57	1131.57	0.00	0.017830609
P52597	PPLKFMSVQRPGPY	1615.85	1615.85	0.00	0.00804829
Q93009	VMLFLKMY	1043.55	1043.55	-0.01	0.057621792
Q93009	PMLLQFFKSQGYR	1613.84	1613.62	-0.22	0.086289549
P38117	LTSKLSVISVE	1174.68	1174.68	0.00	0.058001036
Q08945	PVEAFAQNVLSKA	1372.74	1372.73	0.00	0.00804829
P36776	MGAALTGAESHELQ	1413.66	1413.66	0.00	0.041498559
O15143	AAAGMLSFGGRL	1149.60	1149.59	0.00	0.00804829
Q12905	ETSFSEALLKRNQ	1503.77	1503.77	0.00	0.00804829
Q01130	VEGMTSLKV	962.51	962.50	-0.01	0.072420635
Q9H175	DEEPSPTASCSLTGAQGSETQ	2093.87	2094.15	0.28	0.025157233
O15488	DEVIEVNLI	1042.55	1041.66	-0.90	0.083855422
P54577	VYRLSSVVTQH	1287.69	1287.70	0.00	0.086289549
P54577	EKWGGNKTYTAYV	1497.72	1497.73	0.01	0.053219798
P54577	PPAGSAPGEHVFVKGYEKGQP	2151.07	2151.07	-0.01	0.016768293
Q9Y221	IPLGFGVAAKSTQ	1287.72	1289.70	1.98	0.017830609
Q9Y221	PMAIVVFHQA	1111.58	1111.58	0.00	0.00804829
O95373	PNTIIEALRGTM	1314.70	1314.70	0.01	0.00804829
P38646	IKNVPFKIVRASNG	1541.90	1541.90	0.00	0.00804829
P38646	IENMVKNAEKYAEE	1666.79	1666.79	0.00	0.02318094
Q6P2Q9	PLEVHLL	819.49	819.49	0.00	0.013125513
Q01518	SPSKAGAAPYVQAF	1392.70	1392.70	0.00	0.045733408
Q01518	DVVGIVEIINSK	1284.73	1285.78	1.05	0.044742729
Q01518	VVGIVEIINSK	1169.70	1169.70	0.00	0.011100833
P46782	IKLFGKWST	1078.62	1078.62	0.00	0.016768293
P46782	ELINAAKGSSNSYAIKKK	1903.05	1903.06	0.01	0.043206367
P54840	FHVELTSPPTTEGFK	1688.84	1695.92	7.08	0.098585121
P05455	VKNRSVYIKGFPT	1507.85	1507.84	-0.01	0.052462527
Q9Y230	AFLFNELKGETM	1398.69	1398.70	0.01	0.021094265
P46926	TILANARFF	1051.58	1051.58	0.00	0.042261565
P13645	LKNQILNLTT	1156.68	1156.68	0.00	0.06676783
P13667	PPIPVAKI	833.54	833.54	0.01	0.013125513
P13667	PPIPVAKID	948.56	948.56	0.00	0.00804829
P13667	ATSASVLASRF	1108.59	1108.59	0.00	0.012173913
P13667	VIIIGVFKGES	1160.68	1160.68	0.00	0.00804829
P13667	ESGKKFAMEPEEF	1509.68	1509.69	0.01	0.011100833
P31146	VGTGAAMLTLGPEVHPD	1663.82	1663.83	0.01	0

P31146	PALTAEEWLGGR	1298.66	1298.66	0.00	0.00804829
P31146	AGPLLISLK	910.59	910.58	0.00	0.021094265
P15153	SKPVNLGLW	1012.57	1013.57	1.00	0.016768293
Q9BYT8	FARFSGTNVET	1227.59	1227.59	0.00	0.038596491
P13796	EEMMELREAFAKV	1563.74	1564.75	1.01	0.016768293
P13796	SKAYYHLLEQVAPKG	1702.90	1703.89	0.98	0.026022305
P13796	PKISTSLPVL	1053.64	1053.64	0.00	0.01458671
P13796	PKISTSLPVLD	1168.67	1168.67	0.00	0.065106816
P07237	GAAAESLVESSEVAV IGFFK	2010.03	2010.05	0.02	0.00804829
P07237	SAKQFLQAAEAI	1275.68	1275.68	0.00	0.00804829
P07237	SAKQFLQAAEAID	1390.71	1391.70	1.00	0
P07237	GVVLFKKF	936.58	936.58	0.00	0.00804829
P07237	GVVLFKKFD	1051.61	1051.61	0.00	0.00804829
P07237	EGRNNFEGEVTKENL L	1829.89	1829.90	0.01	0.00804829
P07237	KQPVKVLVGKNFE	1484.87	1484.87	0.00	0.02318094
P07237	KQPVKVLVGKNFED	1599.90	1599.90	0.00	0
P13804	QLHAAVGASRAAV	1232.66	1232.66	0.00	0.00804829
P48735	RAGTFKMVFTPK	1381.75	1381.76	0.00	0.024065864
Q6P3S6	GSPILNGGSLSPGTAA VGGSSL	1897.97	1898.02	0.04	0.048087432
Q13011	AFFQVKEV	966.52	966.52	0.00	0.048087432
P23526	IGLAAWGRKAL	1154.69	1154.69	0.00	0.024065864
P23526	AIVCNIGHF	972.49	972.49	0.00	0.042261565
P23528	KKNIILEEGKEILVG	1682.00	1682.00	0.00	0.011100833
P23528	PYATFVKMLP	1165.62	1165.63	0.01	0.070858283
P23528	PYATFVKMLPD	1280.65	1280.65	0.00	0
P15531	FCIQVGRNIIHGS	1442.75	1442.74	0.00	0.015540016
P23588	GGTGGGSTYVSKPVS WA	1609.77	1609.77	0.00	0.00804829
Q9H773	LPLAVLSKM	970.59	970.59	0.00	0.00804829
Q13200	ELEMLVERLGEK	1426.75	1426.76	0.02	0.043206367
Q13247	LKNGYGFVEFE	1301.63	1301.62	-0.01	0.013125513
Q13247	FMRQAGEVTYA	1271.60	1271.60	0.00	0.057621792
P25054	NGNETESEQPKESNE NQEKEAEKTI	2861.28	2860.25	-1.03	0.048727666
Q9BZ68	KVQEVSSDGGCEAAL GTHYR	2105.98	2105.09	-0.89	0.099316629
P62917	PGRGAPLAKVVFR	1366.82	1366.82	0.00	0.020319303
Q16630	VGEEFNQEAEYGGH	1564.64	1564.66	0.01	0.00804829
P60228	LVKVIQQESYTYK	1597.87	1597.86	-0.01	0.041498559
P62269	GKYSQVLANGL	1148.62	1148.62	0.00	0.00804829
P62081	MFSSSAKIVKPNGEK P	1718.90	1718.90	0.00	0.04972973
P07814	PPLGALLAVEHVK	1342.80	1342.80	0.00	0.012173913
P15880	FFLGASLK	881.50	881.50	0.00	0.013125513
O43719	PLVLNEIRE	1081.61	1081.62	0.00	0.013125513
P31943	PPRKLMAMQRPGPY	1640.86	1640.87	0.00	0.00804829
P31943	PPRKLMAMQRPGPY	1755.89	1756.89	1.00	0.017830609

	D				
P31948	AIHFYNKSLAEHRTP	1782.92	1782.92	0.00	0.094348435
P31948	AIHFYNKSLAEHRTPD	1897.94	1897.95	0.00	0.036556604
P31948	PAMRLILEQMQK	1456.79	1456.80	0.01	0.021094265
Q15911	DPCSPSPGASGSAGKSG	1460.62	1460.80	0.18	0.078392945
O43815	DPSHMVASFSKGYTSIFNMETQQRILTLESNV	3661.74	3661.98	0.24	0.086289549
Q15181	WKVIAINV	941.57	941.57	0.00	0.013125513
Q15185	RFSEMMNNMGG	1272.50	1272.51	0.00	0.048727666
Q9NX94	RAATKAPGMEPSGSVAGLGEL	1998.02	1996.66	-1.36	0.07925636
O94903	QVELSMGMSA	1083.46	1083.48	0.02	0.075098814
Q15365	SSSPEVKGYWASL	1409.68	1409.69	0.01	0.012173913
P25786	LTTKNVSIGIVGK	1328.80	1328.80	0.00	0.036556604
Q99613	FESHITSYKQNPEQSA	1864.86	1864.86	0.00	0.094348435
Q9NWN3	DEGKTKKGVLEAP	1370.74	1371.79	1.05	0.06676783
P17844	VKFVINY	881.50	881.50	0.00	0.00804829
P62316	SVIVVLRNPLIAGK	1477.93	1477.94	0.01	0
O60506	ERAIEALKEFNE	1429.72	1429.72	0.00	0.00804829
O60506	GALAVLQQFK	1073.62	1073.62	-0.01	0.00804829
P00505	PILGVTEAFKR	1229.71	1229.71	0.00	0.00804829
P00505	PILGVTEAFKRD	1344.74	1344.74	0.00	0.032667877
P00505	AGMQLQGYRYY	1348.62	1348.63	0.01	0.021094265
P33991	PLAKEEENVGI	1197.62	1197.63	0.00	0.024065864
P33992	FMPTILSRF	1110.59	1110.59	0.00	0.083855422
Q99829	PFLEFFRQG	1139.58	1139.58	0.01	0.026022305
P35527	LTVGNNKTLL	1071.63	1071.63	0.00	0.058001036
P35527	VNVEINVAPGK	1138.63	1138.64	0.01	0.00804829
P78417	PTVSALLTSEK	1144.63	1144.63	0.00	0.00804829
Q8N0Y2	SGMIPLAGTAPGAEGPAPG	1665.80	1666.93	1.12	0.078392945
P35659	FIKTTVKELIS	1277.76	1277.76	0.00	0.026022305
P35749	LGEELEALKTELE	1472.76	1473.78	1.02	0.039650146
P27797	FGKFVLSSGKFYG	1435.75	1435.75	0.00	0.00804829
P27797	FGKFVLSSGKFYGD	1550.78	1550.78	0.00	0.013125513
P27797	EFTHLYTLIVRP	1469.80	1470.81	1.01	0.00804829
P27797	EFTHLYTLIVRPD	1584.83	1584.84	0.01	0
P27797	DYKGTWIHPEI	1357.67	1357.67	0.00	0.011100833
P27797	YKGTWIHPEI	1242.64	1242.64	0.00	0.011100833
P27797	LWQVKSGTIF	1177.65	1177.65	0.00	0.094348435
P27797	LWQVKSGTIFD	1292.68	1292.68	0.00	0.017830609
P27797	EAYAEEFGNETWGVTKAAEKQMK	2598.21	2599.22	1.02	0.020319303
P02461	DAIKVFCNMETGETCISANPLNVPRKHWWT	3459.65	3460.37	0.72	0.054881266
P63104	YYRYLAEVAAG	1274.63	1274.63	0.00	0.047330765

P53396	ISYVLPEHMSM	1305.61	1305.62	0.01	0
P10515	IEAFKNYTL	1097.58	1097.58	0.01	0.038596491
P60709	IKEKLCYVAL	1178.67	1178.68	0.00	0.01458671
P60709	FEQEMATAASSSSLE KSYELP	2304.05	2304.05	0.01	0.07925636
P60709	LYANTVLSGGTTMYP GIA	1827.91	1827.91	0.00	0.043206367
P60709	ESGPSIVHRKCF	1340.67	1340.67	0.00	0.05075594
P12004	TLALVFEAPNQEKVS	1644.87	1644.88	0.01	0.00804829
P12004	VPLVVEYKIA	1129.67	1129.67	0.00	0.00804829
O14561	MPPLTLEGIQ	1097.58	1097.58	0.01	0.094348435
O14561	PEKLSVNSHFMK	1415.72	1415.72	0.00	0.090995261
P55060	LLTEMVNRFQSG	1393.70	1393.70	0.00	0.08
P55072	NSVVSLSQPKM	1188.62	1188.62	0.00	0.020319303
P55072	RVINQILTEM	1215.66	1216.67	1.00	0.058001036
P55072	LEFLAKMTNGFSGA	1484.73	1486.73	1.99	0.043206367
P53675	AILGNKMFTHY	1293.65	1293.61	-0.04	0.00804829
P04350	SVRSGPFGQIFRP	1446.77	1446.77	0.00	0.013125513
P04350	SVRSGPFGQIFRPD	1561.80	1561.80	0.00	0
P04350	RIMNTFSVVPSPKVS	1660.90	1660.90	0.01	0.015540016
P04350	TVVEPYNATLSVHQL VENT	2113.07	2114.05	0.98	0.00804829
P04350	LVSEYQQYQ	1156.54	1157.52	0.98	0.054082715
Q13765	TYIVFGEAKIE	1268.67	1269.69	1.02	0.072420635
Q00610	FPVAMQISEKH	1285.65	1285.65	0.00	0.047330765
Q00610	AILGNQMFTHY	1293.62	1293.62	0.01	0.011100833
Q00610	NAIITMMNHPT	1241.59	1241.59	0.00	0.013125513
Q9NYU 2	EVQGFLFGKLR	1274.71	1275.72	1.01	0.00804829
Q9NYU 2	SALFINGLHM	1101.56	1101.56	0.00	0.013125513
P45974	PSLAEHLSHFGI	1306.67	1306.66	0.00	0.024065864
P45974	SISESVPVGPKVR	1353.76	1353.76	0.00	0.086289549
O00487	KMTPEQLAIKNVGKQ	1683.93	1683.94	0.00	0.021094265
P30040	KESYPVFYLFR	1447.75	1447.75	0.00	0.012173913
P20618	TRLSEGFSIHTR	1402.73	1402.73	0.00	0.021094265
P30101	SFSEAHSEFLKAASNL R	1892.94	1893.94	1.00	0.021094265
P30101	GPVKVVVAENF	1157.64	1157.65	0.00	0.048087432
P30101	PNIVIAKM	884.52	884.52	0.00	0.00804829
P22102	LAGFAVGAMER	1120.57	1120.56	-0.01	0.064516129
P60842	VLEVTKKFMR	1249.72	1249.72	0.00	0.00804829
P57059	DLMPCSLGTFVLVQ	1537.75	1539.62	1.87	0.065106816
P47756	SYRSPWSNKY	1286.60	1286.61	0.00	0.01458671
Q02543	LTTAGAVTQCYR	1282.63	1282.64	0.01	0.00804829
P55786	VFSPIGERLGW	1259.67	1259.66	-0.01	0.021094265
O95881	GGYIPRILFL	1147.68	1147.67	0.00	0.017830609
P55809	FALVKAWKA	1032.61	1032.60	-0.01	0.015540016

P55809	ESFAMIRGGHV	1184.58	1184.58	0.00	0.021094265
P49327	GFKEQGVTFPSG	1252.61	1252.61	0.00	0.044742729
P49327	PGSAELQKVLQG	1225.67	1225.67	0.00	0.00804829
P49327	SLLGMEFSGR	1095.54	1095.54	0.00	0.00804829
P49327	LVEAVAHILGIR	1289.78	1289.79	0.00	0.048727666
P49411	YVKNMITGTAPL	1306.70	1306.69	0.00	0.094348435
P49411	PELGLKSVQKLL	1323.81	1323.81	0.00	0.032667877
P14625	KIRLISLT	942.62	942.62	0.00	0.071713147
P14625	EYKAFYKSFSKES	1594.77	1594.77	0.00	0.025157233
P14625	EEEETAKESTAEKDEL	1818.80	1818.81	0.00	0.098585121
P61254	MKFNPFVTS	1069.53	1069.53	0.00	0.00804829
P08134	GKQVELALW	1042.58	1043.58	1.00	0.015540016
P08134	GKQVELALWD	1157.61	1156.61	-1.00	0.041498559
P49588	LTGLIAEEKGLVV	1340.79	1340.78	-0.01	0.026022305
P06744	PQNMFEFW	1097.46	1097.47	0.00	0.042261565
P40227	KGFVVINQKGI	1201.72	1201.71	0.00	0.069556452
P40227	EIMRAGMSSLKG	1260.63	1260.63	0.00	0
P57737	LPVEVLQFHPTS	1365.73	1365.73	0.00	0.011100833
P49720	FQKIFPMG	966.50	966.50	0.00	0.090995261
P49915	LTSKPPGTTEWE	1344.66	1343.66	-1.00	0.013125513
P84090	LSCLVYRA	923.49	923.49	0.00	0.00804829
P24752	FPIAPVYAASMVLK	1505.83	1505.84	0.01	0.011100833
Q14562	ASLITVMQIHLTEPPGD	1820.93	1822.70	1.77	0.058001036
Q14566	LAAAAEPGAGSQHLEVR	1675.86	1675.87	0.00	0.00804829
Q14566	VSKLSTPGARAETNSRVSGV	2015.08	2016.08	1.00	0.09762901
P32969	MKTILSNQTV	1133.61	1133.61	0.00	0.00804829
P32969	GIYVSEKGTVQQA	1378.71	1378.71	0.00	0.00804829
P16989	PAPKSPVGSGAPQAAAPAPAAHVAGNPGG	2501.28	2502.28	1.00	0.053219798
P67809	TKPGTTGSGAGSGGPGGLTSAAPAGG	2070.00	2070.01	0.02	0.021094265
Q9NZR2	DYVNRRLYWA	1354.68	1353.63	-1.05	0.020319303
P01031	AYSPGQTVSLNMATGM	1642.73	1641.76	-0.97	0.045733408
P26447	EAAFQKLMSNL	1250.63	1250.63	-0.01	0.057621792
O14950	YLRELLTTMG	1195.63	1195.62	0.00	0.017830609
Q9H1A4	DLKLGPYV	903.51	904.09	0.58	0.064516129
Q969H8	VRPGGVVHSFSHNVGPG	1701.87	1701.87	0.00	0.044742729
Q969H8	VRPGGVVHSFSHNVGPGD	1816.90	1817.90	1.00	0.00804829
Q96GV9	KMIKEPAD	946.48	947.56	1.09	0.090995261
Q7LG56	WALRWIAD	1029.54	1029.48	-0.06	0.047330765
Q9P227	RSVCSGASGRRAGAGD	1505.71	1503.77	-1.94	0.011100833
Q8WU39	DRAPLTATAPQL	1252.68	1252.67	0.00	0.094348435

Q8WU39	RAPLTATAPQL	1137.65	1137.65	0.00	0.015540016
Q8WU39	RAPLTATAPQLD	1252.68	1252.68	0.00	0
Q8WWH4	DEFTKNGITSK	1238.61	1238.63	0.01	0.086289549
Q9ULL4	HVTVPLALMFE	1255.66	1255.61	-0.05	0.021094265
O60281	CTELVLKQLQEMKPTVSLKKLEVHSN	2994.62	2993.58	-1.05	0.083855422
Q6N069	LSLLQIQMR	1100.64	1100.64	0.00	0.013125513
Q8N7Z3	APVMTMTASLLPLVPD	1670.86	1671.90	1.04	0.08
P20591	IATTEALSMAQEV	1362.67	1362.67	0.00	0.012173913
P20591	TYSWLLKERS	1281.67	1281.67	0.00	0.00804829
Q8NFZ5	DPGSGGWEEAPRAAAALCTLYHEAGQRLRRLQ	3478.73	3480.67	1.94	0.09823911
Q8NHS2	MPTLSVFMD	1071.46	1071.63	0.17	0.070858283
P00558	NGAKSVVLMSHLGRP	1564.85	1564.85	0.00	0.00804829
P00558	NGAKSVVLMSHLGRPD	1679.88	1679.88	0.00	0.00804829
P00558	KIQLINNML	1085.63	1085.63	0.00	0.015540016
Q8NBS9	KGTVLALTENNF	1305.69	1305.69	0.00	0.00804829
Q8NBS9	KGTVLALTENNFD	1420.72	1420.72	0.00	0
Q8NBS9	SLHRFVLSQAKDEL	1641.88	1641.88	-0.01	0.098585121
Q9BU89	PEVLEILKQYSS	1404.75	1404.75	0.00	0.021094265
Q96KP4	VKQLGGSVELV	1127.66	1127.66	0.00	0.013125513
P24534	KSYIEGYVPSQA	1340.66	1340.65	-0.01	0.00804829
P24534	VAVFEAVSSPPPA	1269.66	1270.66	1.00	0.090995261
P24534	MLEEQITAFE	1209.56	1209.56	0.00	0.072420635
P24534	YVQSMDVAAFNKI	1484.73	1485.74	1.00	0.098585121
P62979	QQRLIFAGKQLE	1412.78	1412.78	0.00	0.012173913
P62979	QQRLIFAGKQLED	1527.80	1527.81	0.00	0.015540016
Q8N163	PELLLLR	852.54	852.55	0.00	0.015540016
P49736	LTEPIISRF	1074.61	1074.61	0.00	0.00804829
P13639	PIFKVFD	864.47	864.47	0.00	0.015540016
P13639	ITKGVQYLNEIK	1404.80	1404.80	0.00	0.00804829
Q8N5D0	DGLIRQY	863.45	863.48	0.03	0.072420635
P61204	AVLLVFANKQ	1101.65	1101.65	0.00	0.00804829
P18085	AVLLLFANKQ	1115.67	1116.67	1.00	0.046718576
P09651	ESLRSHFEQWGTLT	1671.80	1671.80	0.00	0.054881266
P09651	SQRPGAHLTVKKIFVGGIKE	2164.25	2164.25	0.00	0.016768293
P21926	EVIKEVQEFYK	1392.73	1392.74	0.01	0.00804829
P61978	SSGPERILSISA	1215.65	1216.64	1.00	0.038596491
P61978	LISESPIKGRAQPY	1557.85	1557.85	0.00	0.00804829
P61978	LGGPIITTQVTIPK	1436.86	1436.85	-0.01	0.00804829
P09104	LYTAKGLFRAAVPSGASTGIYEALELR	2853.54	2854.55	1.01	0.015540016

P10599	VKQIESKTAFQEAL	1590.86	1590.86	0.00	0.098585121
P31946	YFRYLSEVASG	1290.62	1290.63	0.01	0.044742729
P07737	LRTKSTGGAPTFNVT VTKT	1978.08	1978.09	0.01	0.00804829
P22234	MKIEFGV	822.43	822.43	0.00	0.036556604
O75347	LEEAEEYKEARLVL	1690.88	1690.87	0.00	0.00804829
Q8IWA6	KPMKSIKYMD	1255.63	1254.72	-0.91	0.058001036
Q9H2U1	DPFVIPLGWEEAR	1527.77	1528.26	0.49	0.098585121
P00338	TLWGIQKELQF	1361.73	1361.74	0.00	0
P20700	QLLLNYAKKES	1288.70	1288.70	0.00	0.012173913
P20700	QPMGGWEMIRKIG	1484.73	1484.73	0.00	0.053219798
P20700	LIWKNQNSWGTGE	1531.74	1531.75	0.01	0.00804829
Q8IZI9	ESPGCLEASVTFNLFR LLTR	2234.15	2237.19	3.04	0.098585121
P62266	EVLVAGFGRKGHAV G	1477.81	1477.81	0.00	0.036556604
P07741	FPTPGVVFR	1018.56	1018.57	0.01	0.090995261
P07741	ALEPGQRVVVV	1165.68	1165.68	0.00	0.00804829
P07741	ALEPGQRVVVVD	1280.71	1280.71	0.00	0.00804829
P43320	SLSSLRPIKVD	1213.70	1215.64	1.94	0.077832512
P15170	GPIRLPIV	863.56	863.56	0.00	0.041498559
Q02790	LGKGEVIKAW	1099.64	1099.64	0.00	0.00804829
P04406	AGAEYVVESTGVFTT MEKAGAHLQGGAKR VIISAPSA	3702.90	3703.91	1.01	0.057621792
P04406	APMFVMGVNHEKY	1537.71	1538.72	1.01	0.013125513
P04406	APMFVMGVNHEKYD	1636.74	1636.74	0.00	0.011100833
P04406	NFGIVEGLMTTVHAI TATQKTV	2330.23	2331.24	1.01	0.00804829
P04406	HFVKLISWY	1191.64	1191.64	0.00	0.041498559
P04406	NEFGYSNRVV	1183.56	1183.56	0.00	0.00804829
P04406	NEFGYSNRVVD	1298.59	1298.59	0.00	0
P07195	LQHGSLFLQTPKIVA	1650.95	1651.96	1.01	0.00804829
P30086	LSKWSGPLSLQEV	1442.78	1443.78	1.00	0.013125513
P30086	LSKWSGPLSLQEVD	1557.80	1557.81	0.00	0.021094265
P30086	EQPQHPLHVTYAGAA V	1698.85	1698.85	0.00	0.00804829
P30086	SGKLYTLVLT	1093.64	1093.64	0.00	0.078392945
P30086	DYVPKLYEQLSGK	1538.80	1538.79	-0.01	0
P30086	YVPKLYEQLSGK	1423.77	1424.77	1.00	0
P06746	IRLIPKD	853.54	853.50	-0.04	0.065106816
P07205	NGAKAVVLMSHLGR P	1548.86	1548.83	-0.03	0.021094265
P54578	VKFPLML	846.50	846.51	0.00	0.075098814
P35606	FQPSRSTAQQEL	1390.68	1390.69	0.00	0.02318094
P30041	SWGILFSHPR	1198.62	1198.62	0.00	0.00804829
P30041	EILRVVISLQLTAEKR VATPV	2316.39	2317.37	0.98	0.047330765
P04264	EINKRTNAENEFVTIK K	2015.08	2015.08	0.00	0.015540016
P04264	SIIAEVKAQYE	1249.66	1249.66	0.00	0.00804829

P04264	YQELMNTKLAL	1322.69	1322.70	0.01	0.065106816
Q9Y4Y9	KEIVGTLLGF	1075.63	1075.63	0.00	0.09823911
Q14103	GKMFIGGLSW	1094.56	1095.56	1.01	0.00804829
Q14103	YFSKFGEVV	1074.54	1074.54	0.00	0.021094265
Q14103	YFSKFGEVVD	1189.57	1189.57	0.00	0.015540016
P13010	MVAIVRYAY	1084.57	1084.57	0.00	0.062854199
P13010	TNETPYFMKSI	1329.63	1329.63	0.00	0.031572556
P46781	YILGLKIE	947.57	947.57	0.00	0.031572556
P08708	TKEMLKLL	974.58	974.58	0.00	0.082444229
P49773	LGLNKGYRMVVNEGS	1635.84	1635.84	0.00	0.011100833
O75821	VSMTFITSKE	1141.57	1141.57	0.00	0.021094265
P17987	GATILKLL	827.55	827.55	0.00	0.083855422
P40926	VVVIPAGVPRKPGMTR	1675.99	1675.99	0.00	0.058001036
P40926	IVRANTFVAELKGL	1529.89	1529.90	0.01	0.015540016
P42330	SFASHPNYPYS	1268.55	1268.55	0.00	0.012173913
P28066	EKGPQLFHM	1067.52	1067.52	0.00	0.00804829
Q9NYA4	FTCLKESD	941.42	941.50	0.08	0.042261565
P14174	PMFIVNTNVPRASVP	1640.87	1640.87	0.00	0.098585121
P14174	PMFIVNTNVPRASVPD	1755.90	1756.93	1.03	0.052462527
P14174	RVYINYY	989.50	989.50	0.00	0.015540016
P14174	MNAANVGWNNSTFA	1495.65	1496.64	0.99	0
Q92804	TGKPKGEATVSFD	1335.67	1333.89	-1.78	0.064516129
P14867	NRLRPGLGERVTEVKT	1824.03	1824.98	0.95	0.077832512
P07954	PKIANAIMKAA	1126.65	1126.65	0.00	0.00804829
P00813	PLIFKSTL	917.56	917.56	0.00	0.048727666
Q04446	VPELARLLEI	1151.69	1152.69	1.00	0.024065864
Q04446	PYLKPYAV	949.53	949.53	0.00	0.042261565
P46100	DFLELASREKTE	1436.71	1436.78	0.07	0.090995261
Q8N7X4	DVSVPQESQGASPTGSP	1641.75	1640.88	-0.87	0.08
O14980	QGEVVREFMK	1204.59	1204.59	0.00	0.011100833
Q9UQ80	AMPFTLRAFE	1181.59	1181.59	0.00	0.00804829
Q9UQ80	AMPFTLRAFED	1296.62	1297.62	1.01	0.048727666
P33993	PGVAKSQLLSYI	1274.72	1273.71	-1.01	0.00804829
Q9H089	QAEISHSESEHLPAR	1672.78	1672.78	0.00	0.00804829
Q15393	PSGQLNEYTERKEMSA	1838.85	1838.85	0.00	0.011100833
Q15393	NTVRIISL	914.55	914.56	0.00	0.065922921
Q9HAU5	DLELELENLEIN	1442.71	1443.76	1.05	0.058001036
P42126	VQNFVSFISK	1167.63	1167.63	0.00	0.042261565
Q7L5Y1	MRRCQIIR	1074.59	1076.02	1.43	0.08708134
Q15424	SGAAGAAALSSASSETGTRRLS	2007.00	2007.00	0.01	0.016768293
Q9NRR	STVVGTSRLR	1074.61	1074.61	0.00	0.070858283

4					
Q01105	ETSEKEQQEAIEHI	1651.77	1651.77	0.00	0.00804829
Q01105	ETSEKEQQEAIEHID	1766.80	1766.80	0.00	0
Q01105	ENPYFENKVLSKEFH LNESG	2362.12	2362.10	-0.02	0.094348435
P62995	LREVFSKYGPIA	1378.76	1378.76	0.00	0.011100833
P06576	LYHEMIESGVINLK	1644.85	1644.86	0.00	0.090995261
P06576	PAPATTFAHL	1024.53	1024.53	0.00	0.015540016
P06576	PAPATTFAHLD	1139.56	1137.65	-1.91	0.041498559
Q9GZL7	NKLYSYRYSPTTSHV GA	1942.95	1943.96	1.01	0.064516129
Q5T0W 9	METSSMLSSLND	1329.54	1328.69	-0.86	0.082444229
P08107	AAKNQVALNPQNTV F	1613.85	1612.85	-1.00	0.053219798
P08107	LFRSTLEPVEKALR	1657.95	1657.96	0.01	0.020319303
P08107	KCQEVISWL	1104.56	1104.56	0.00	0.028117359
P31040	AIHYMTEQAPAAVVE LENYGMPFSRTED	3168.45	3167.50	-0.96	0.058001036
P31277	PRTGAGGGGGSPCTK ATPGSEPKGAAEGSG GD	2770.26	2770.54	0.28	0.065106816
Q9BX66	DRGQEGTARPPTPLG PLGCVPTIPATASAAS PLTFPTL	3756.95	3757.65	0.69	0.099316629
O43776	IPEAPERLMT	1155.60	1155.60	0.00	0.011100833
O43776	SEEILAGYKREGI	1463.76	1463.76	0.00	0.00804829
P27338	VVVVGGGISGMAAA KLLH	1677.96	1675.99	-1.97	0.047330765
Q9UBL3	DGRRSPPWEP	1195.57	1196.71	1.13	0.053219798
Q86UP2	EEQMNTMKAVLEEK	1694.79	1695.92	1.14	0.098585121
Q86UU0	RPLLPPPPP	982.60	984.46	1.86	0.017830609
Q86V81	GRPMNIQLVTSQI	1455.79	1456.79	1.00	0.02318094
Q16851	WGKIQRPPE	1109.60	1107.61	-1.99	0.09762901
Q99729	AGKMFVGGLSW	1151.58	1151.59	0.01	0.013125513
Q99729	AGKMFVGGLSWD	1266.61	1266.61	0.00	0
Q99729	YFTKFGEVV	1088.55	1088.56	0.00	0.028117359
Q99729	PNTGRSRGFGFILFK	1695.92	1695.92	0.00	0.047330765
P62837	PLVPEIARIYKT	1398.82	1398.82	0.00	0.015540016
P62837	PLVPEIARIYKTD	1513.85	1513.84	-0.01	0.00804829
P60660	YVEGLRVF	981.53	981.53	0.00	0.024065864
Q96EL1	SFLSQLRWELLCGRD	1821.92	1821.83	-0.09	0.036556604
Q6ZRP7	CMEEKNQAVCHD	1388.52	1388.75	0.23	0.078392945
P62913	TGNFGFGIQEHI	1318.63	1318.63	0.00	0.013125513
Q14524	CGSPAVGILFFTTYIII SFLIVVNMYIAIILENF SVATEESTEPLSED	5284.67	5286.24	1.57	0.06676783
Q86TV4	DSSASASQVAGIT	1192.56	1192.61	0.05	0.083855422
Q76NI1	GQGPLPGLSSTSRD	1370.68	1370.80	0.12	0.086289549
Q14697	EPGAWEETFKTHS	1499.67	1499.67	0.00	0.021094265
Q14697	SKPYGPMSVGL	1134.57	1134.57	0.00	0.013125513

Q14697	ISSNTAGKTLFGKMM	1584.80	1584.80	0.00	0.00804829
Q14697	VRWMSETGII	1190.61	1190.61	0.00	0.070858283
Q14697	MNEPSVFNGPEVTML K	1791.85	1791.86	0.00	0.012173913
Q9UL46	LTSLRAPL	869.53	869.53	0.00	0.011100833
Q86UT5	DPGLPAKKAGMQAG	1339.69	1337.77	-1.92	0.08
P50542	QNAPLVSRAPQTFKM	1686.89	1685.00	-1.89	0.064516129
Q9HB71	PSEGLMNVLKKIYE	1619.86	1619.86	0.00	0.015540016
P22626	YFEEYGKI	1047.49	1047.49	0.00	0.00804829
P22626	YFEEYGKID	1162.52	1162.52	0.00	0
P22626	SRGGGGNFGPGPGSN FRGGS	1820.83	1820.83	0.00	0.012173913
P68104	SGKSTTTGHLIYKCG GI	1721.88	1721.89	0.01	0.024065864
P68104	FIKNMITGTSQA	1309.67	1309.67	0.00	0.00804829
P68104	FIKNMITGTSQAD	1424.70	1424.70	0.00	0
P68104	TVAFVPISGWNG	1246.63	1246.64	0.00	0.028117359
P68104	GNASGTTLLEAL	1145.59	1145.59	0.00	0.00804829
P68104	GNASGTTLLEALD	1260.62	1260.62	0.00	0.094348435
P68104	NVGFNVKNVSVK	1303.72	1303.73	0.00	0.00804829
P68104	MVPGKPMCVESFS	1410.63	1410.64	0.00	0.011100833
P68104	MVPGKPMCVESFSD	1525.66	1525.66	0.00	0.00804829
P68104	YPPLGRFAVR	1174.66	1174.66	0.00	0.024065864
P68104	MRQTVAVGVIKAV	1370.81	1370.81	0.00	0.024065864
P68104	MRQTVAVGVIKAVD	1485.83	1485.84	0.00	0
O14787	SAFRGICMMIGVNPG GVVQD	2065.97	2066.41	0.43	0.058001036
Q8NC51	QSNVTEETPEGEEHH PVA	1971.84	1971.85	0.01	0.00804829
P15121	YPFHEEF	967.41	967.41	0.00	0.017830609
P08238	IIPNPQERTLTLV	1492.86	1492.87	0.01	0.048727666
P08238	IIPNPQERTLTLVD	1607.89	1607.88	-0.01	0.057621792
P08238	QEELNKTKPIWTRNP	1835.95	1835.96	0.01	0.00804829
P08238	QEELNKTKPIWTRNP D	1950.98	1951.99	1.01	0.00804829
P08238	ITQEEYGEFYKSLTN	1820.85	1822.84	1.99	0.016768293
P08238	ITQEEYGEFYKSLTND	1935.87	1935.88	0.01	0
Q6ZSZ6	EEDLDDSPKGGLDIL K	1742.86	1740.93	-1.93	0.098585121
P25788	IREEAEKYAKESLKEE	1950.99	1950.99	0.00	0.065922921
P51991	SLREHFEKWGTLT	1602.82	1602.82	0.00	0.021094265
Q9NY15	QLLEPPGLGARCD	1367.69	1367.79	0.10	0.082444229
P62937	GEPLGRVSFELFA	1420.74	1420.74	0.00	0.013125513
P62937	FTRHNGTGGKSIYGE KFE	2026.99	2026.99	0.00	0.012173913
P62937	FTRHNGTGGKSIYGE KFED	2142.01	2143.02	1.01	0.094348435
Q5UCC 4	DTGGQGGGGGGGGG GGSGR	1403.58	1403.73	0.15	0.058001036
P61247	GYEPPVQESV	1103.51	1103.50	-0.01	0.048727666
Q9UBL9	MLGNGALQ	818.40	817.47	-0.93	0.09823911

Q13263	IVENYFMR	1070.52	1070.52	0.00	0.015540016
P25705	VPVGEELLGRVV	1265.73	1265.74	0.01	0.00804829
P25705	VPVGEELLGRVVD	1380.76	1380.76	0.00	0.093484419
Q6ZVP2	MMPGMATLMA	1100.44	1100.56	0.12	0.064516129
P62805	VVYALKRQGRTLYG FGG	1884.04	1884.03	0.00	0.011100833
Q1KMD 3	LAQRLQEAL	1040.60	1040.60	0.00	0.065106816
Q1KMD 3	LGVAFWISK	1019.58	1020.59	1.01	0.00804829
Q1KMD 3	RYYRNYYGYQGYR	1820.84	1820.84	0.00	0.028117359
Q6XQN 6	PAFFEHLRAL	1199.65	1200.64	1.00	0.012173913
P06733	LFTSKGLFRAAVPSG ASTGIYEALELR	2853.54	2853.53	-0.01	0.042261565
P06733	VAASEFFRSGKY	1360.68	1360.68	0.00	0.00804829
P06733	VAASEFFRSGKYD	1475.70	1475.70	0.00	0
P99999	LIAYLKKATNE	1262.72	1262.73	0.00	0
P04075	PYQYPALTPEQKKEL SD	2006.00	2005.99	-0.01	0.052462527
P04075	IAHRIVAPGKGILAA	1485.91	1485.92	0.01	0.054082715
P04075	IAHRIVAPGKGILAAD	1600.94	1600.94	-0.01	0
P04075	GRPFPQVIKSKGGVV GIKV	1965.19	1966.19	1.00	0.00804829
P04075	KGVVPLAGTNGETTT QGL	1741.92	1741.93	0.00	0.00804829
P14618	TFLEHMCRL	1148.55	1148.55	0.00	0.00804829
P14618	PILYRPVAVAL	1210.74	1210.74	0.00	0.00804829
P14618	PILYRPVAVALD	1325.77	1325.77	0.00	0.00804829
P14618	DGLISLQVKQKGA	1355.78	1355.77	0.00	0.013125513
P14618	GLISLQVKQKGA	1240.75	1240.75	0.00	0.00804829
P14618	GLISLQVKQKGAD	1355.78	1355.78	0.00	0.055818852
P14618	FLVTEVENGGSLGSK KGVNLPGAAV	2442.31	2443.33	1.02	0.016768293
P14618	LPAVSEKDIQD	1213.62	1213.62	0.00	0.098585121
P14618	MVFASFIRKAS	1255.67	1255.68	0.00	0.00804829
P14618	MVFASFIRKASD	1370.70	1370.70	0.00	0.021094265
O94911	DVCQNPEEPEGE	1344.51	1342.65	-1.86	0.075098814
P06748	ENEHQLSLRTVSLGA GAK	1890.99	1891.00	0.01	0.00804829
P06748	ENEHQLSLRTVSLGA GAKD	2006.02	2007.02	1.00	0.00804829
P06748	LWQWRKSL	1115.62	1115.62	0.00	0.00804829
Q9UNM 6	ALRFLGCV	877.48	877.48	0.00	0.057621792
P37802	VGRPQPGRENFQNW LK	1925.00	1925.00	0.00	0.035756853
P37802	PNWFPKKSKENPRNF S	1975.01	1975.01	0.00	0.041498559
P13693	GVTPYMIFFK	1201.62	1201.62	0.00	0.026022305

P28070	IAHMISGFE	1003.48	1003.48	0.00	0.015540016
P49589	NELAQSEAYFEN	1413.60	1413.80	0.20	0.094348435
P19338	EIEPAAMKAAAAAPASE	1608.78	1608.78	0.00	0.00804829
P19338	ARTLLAKNLPYKVTQ	1715.01	1716.01	1.00	0.025157233
P19338	GKSKGIAYIEFKTEA	1640.88	1640.88	0.00	0.047330765
P19338	GRSISLYYTGEKGQNQ	1799.88	1799.88	0.00	0.00804829
P19338	GRSISLYYTGEKGQNQD	1914.91	1915.91	1.00	0.00804829
P19338	TTEETLKESF	1183.56	1183.56	0.00	0.021094265
A0M8Q6	FYPGAVTVAWKA	1308.69	1308.69	0.00	0.00804829
A0M8Q6	FYPGAVTVAWKAD	1423.71	1423.71	0.00	0.00804829
	DPGEKGGASAAGCV	1217.53	1217.56	0.03	0.046718576
P62826	VTSRVTYKNVPNWHR	1855.98	1855.98	0.00	0.09823911
P41252	PVSIIQKYGA	1074.61	1074.61	0.00	0.021094265
Q96JG9	PKEALAGCLLQGEGSPLED	1925.94	1925.00	-0.94	0.065922921
P12956	LLAVVFYGTEK	1238.69	1238.70	0.01	0.026022305
Q15084	VIELTPSNFNREVIQS	1844.96	1844.96	0.00	0.017830609
Q15084	AALSALRQLVK	1168.73	1168.73	0.00	0.036556604
Q15084	AALSALRQLVKD	1283.76	1283.75	0.00	0.065106816
P22314	LSSQFYLREE	1270.62	1270.62	0.00	0.017830609
P22314	SNGEQPLSAMVSMVTK	1677.81	1677.81	0.00	0.00804829
P22314	SRLEELKATLPSP	1439.80	1440.71	0.91	0.021094265
P22314	FIVAASNLRAENY	1466.75	1466.75	0.00	0.00804829
P22314	VEVPYVRYTIR	1393.77	1393.77	0.00	0.094348435
P23284	GTGGKSIYGERFP	1367.68	1367.69	0.00	0.038596491
Q8IZP6	MAAPPSPGRTAD	1169.55	1167.57	-1.98	0.058001036
A5A3E0	ESGPSIVHRKCL	1306.68	1306.68	0.00	0.028117359
Q5JTN6	FSPTVNCLATGSW	1381.63	1380.77	-0.87	0.058001036
Q9NP58	EVEAAAQAAGIHD	1280.60	1278.73	-1.87	0.090995261
P15289	AAVTFGPSQVARGE	1388.70	1389.73	1.03	0.048727666
Q9UBU7	MSQSPAVHLM	1115.51	1115.52	0.01	0.090995261
Q9Y262	QKVYELQASRVSS	1476.76	1476.76	0.00	0.00804829
C9JGY3	MQACGGGAAGRRAF	1367.62	1367.79	0.18	0.075098814
Q8WZ42	SELHESWKYNMSFINSVALLTINEASAE	3182.52	3183.88	1.36	0.094348435
P42704	MENAENILTVMR	1419.69	1419.68	-0.01	0.00804829
P42704	TYLALLNAYAEKG	1425.75	1426.76	1.01	0.012173913
Q96L91	EQALAGSLVAGAGSTVETD	1774.86	1774.88	0.02	0.083212385
P50990	QIIMAKPAGGPKPPSGKK	1787.01	1787.02	0.00	0.044117647
P10809	ARALMLQGV	957.54	957.54	0.00	0.036556604

P10809	GEALSTLVLNRLKVGLQVVAVKAPGFG	2735.61	2736.60	1.00	0.070351759
P10809	FVNMVEKGII	1148.63	1148.63	0.00	0.013125513
P10809	FVNMVEKGIID	1263.65	1263.65	0.00	0.00804829
P10809	TALLDAAGVASLLTTAEVVVTEIPK	2481.39	2482.41	1.01	0.098585121
P10809	AAGVASLLTTAEVVVTEIPKEEK	2354.29	2354.29	0.00	0.081632653
P10809	PGMGAMGGMGGGMGGGMF	1558.59	1559.60	1.01	0
P07900	SGKELHINLIPNKQ	1589.89	1589.89	0.00	0.065106816
P07900	ISMIGQFGVGFYSAYLVAEKVTVITKHN	3071.62	3071.63	0.02	0.00804829
P07900	ISMIGQFGVGFYSAYLVAEKVTVITKHND	3186.64	3185.62	-1.02	0.013125513
P07900	ITNEEYGEFYKSLTN	1806.83	1808.83	1.99	0.044117647
O76037	TSSVSTGHTTPLLVTD	1614.81	1613.39	-1.42	0.09762901
Q6ECI4	LECSTLGKNWKCED	1624.72	1624.90	0.17	0.064516129
	FNIQEGQRAMLD	1420.68	1420.67	0.00	0.047330765
P60174	TEVVCAPPTAYI	1262.62	1262.62	0.00	0.012173913
P60174	EREAGITEKVVFEQTKVIA	2128.15	2128.15	0.00	0.098585121
P60174	AVAQSTRIIYGGSVTGATCKELASQP	2607.33	2608.33	1.00	0.020319303
P60174	GFLVGGASLKPEFV	1419.78	1419.77	0.00	0.011100833
P60174	GFLVGGASLKPEFVD	1534.80	1534.80	0.00	0.00804829
Q96RL7	TVGSHGAVKCKGLKM	1530.80	1529.80	-1.01	0.083212385
P15311	NAMLEYLKIAQ	1292.68	1292.68	0.00	0.083855422
P15311	ALGLNIYEK	1019.57	1019.56	0.00	0.021094265
Q08211	FLLVVLR	858.57	860.45	1.88	0.042261565
Q08211	AVIEAEHTLREL	1379.74	1379.74	0.00	0.00804829
B3KWE3	DSGTMNIF	883.37	882.47	-0.90	0.078392945
O15050	WSTQEIEACLQ	1306.59	1306.66	0.07	0.020319303
P63244	KTIIMWKLTR	1288.77	1288.77	0.00	0.021094265
P63244	NLVRVWQVTIGTR	1540.88	1540.88	0.00	0.020319303
Q6IS14	FQLIGIQ	817.47	817.47	0.00	0.041498559
Q01844	PPTAKAAVEWF	1215.63	1215.62	-0.01	0.024065864
P62942	VELLKLE	842.51	843.52	1.01	0
	MFRCSTSCCED	1296.42	1296.26	-0.16	0.083212385
Q00839	DEEAGGRPAMEPGNGSL	1685.73	1685.73	0.00	0.021094265
D3DQH1	MIHCEATHCGKILSNK	1799.85	1800.86	1.01	0.098585121
Q5T1H1	DQGPKQTGIVKH	1306.70	1305.73	-0.97	0.071713147
P0C7T9	LEAVAKFL	889.53	889.52	0.00	0.013125513
P26641	YESYTWRKL	1244.62	1244.62	0.00	0.00804829
Q92598	PFIQKEKENLSY	1494.77	1495.78	1.01	0.052462527
P50395	FTGHALALYRT	1248.66	1248.66	0.00	0.013125513
P50395	LGTESQIFISRTY	1513.78	1514.78	1.00	0.021094265

O00148	EEEEPQAPQESTPAPPKK	1972.94	1973.96	1.02	0.021094265
P25205	TEEEMPQVHTPKTA	1596.75	1596.75	0.01	0.021094265
P06865	TIIQVWRE	1043.58	1043.58	0.00	0.016768293
P27824	MTPPVNPSREIE	1368.67	1368.68	0.00	0.081632653
O43852	AFLGAEEAKTF	1182.59	1182.59	0.00	0.054881266
Q8NCW5	AIFGFSFKG	972.51	973.51	1.00	0.00804829
Q13126	PEILEGRTEKYV	1432.76	1432.75	-0.01	0.086289549
Q13126	ALILGKIKNV	1067.71	1066.53	-1.18	0.038596491
P29401	LAMFRSVPTSTVFYPS	1801.91	1802.90	0.99	0.021094265
P29401	FQVGQAKVVLKSK	1430.86	1430.86	0.00	0.058001036
P29401	PFTIKPL	814.50	814.50	0.00	0.054881266
P29401	AIAQAVRGLITKA	1310.80	1310.80	0.00	0
P14314	GTVSVTELQTHPEL	1509.77	1509.75	-0.02	0.00804829
P14314	GTVSVTELQTHPELD	1624.79	1624.80	0.01	0
P14314	GALSEAEAQALLSG	1315.66	1316.66	1.00	0.035756853
P14314	AAQEARNKFEEAERSLK	1976.01	1976.01	0.00	0.036556604
P02765	TLETTCHVL	1015.50	1015.50	0.00	0.013125513

Appendix Table 5. All peptides identified using the high throughput middle down strategy for human ribosomes.

UniProt	Sequence	Theoretical Mass	Observed Mass	Mass Diff	E Value
P08865	VLQMKEEDVLKFLAAGTHLGGTNL	2583.37	2583.37	0	1.56E-24
P08865	VLKFLAAGTHLGGTNL	1610.91	1610.92	0.01	1.33E-26
P08865	VLKFLAAGTHLGGTNLDFQMEQYIYKRKS	3327.74	3328.73	0.99	2.12E-08
P08865	FQMEQYIYKRKS	1619.81	1620.8	0.99	0.00000278
P08865	VSVISSRNTGQRAVLKFAAATGATPIAGRFTPGTFTNQIQAAFREPRLLVVT	5486.01	5485.92	-0.09	1.28E-18
P08865	PRADHQPLTEASYVNLPTIALCNT	2623.31	2623.29	-0.02	2.89E-13
P08865	IAIPCNNKGAHSVGLMWWMLAREVLRMRGTISREHPWEVMP	4771.41	4771.41	0	0.0000242
P08865	IAIPCNNKGAHSVGLMWWMLAREVLRMRGTISREHPWEVMPDLYFYR	5628.81	5628.81	0	0.000193
P08865	PEEIEKEEQAAAEKAVTKEEFQGEWTAPAPEFTATQPEVA	4415.1	4415.16	0.06	3.26E-16
P15880	KEWMPVTKLGRLVK	1683.99	1684	0.01	3.42E-14
P15880	KEWMPVTKLGRLVKD	1799.01	1799	-0.01	0.000000679

P15880	MKIKSLEEIYLFSLPIKESEIID	2737.49	2737.51	0.02	0.000351
P15880	FFLGASLK	881.501	881.501	0	9.66E-12
P15880	FFLGASLKDEVLKIMPVQKQTRAGQRTRFKAFVAIG	4050.27	4050.35	0.08	8.54E-16
P15880	HLVKTHTRVSVQRTQAPAVATT	2400.33	2400.31	-0.02	8.42E-09
P23396	AVQISKKRKFVA	1415.86	1416.87	1.01	7.77E-20
P23396	GIFKAELNEFLTRELAE	1979.04	1979.02	-0.02	5.37E-22
P23396	TAVRHVLLRQGVLGIKVKIMLPW	2626.6	2626.61	0.01	9.57E-11
P23396	PTGKIGPKKPLPDHVSIVEPK	2236.29	2235.35	-0.94	6.36E-11
P61247	VKAPAMFNIRNIGKTLVTRTQGTKIAS	2914.65	2914.68	0.03	2.53E-19
P61247	GLKGRVFEVSLA	1274.73	1274.73	0	1.52E-08
P61247	LQNDEVAFRKFKLITE	1950.06	1950.06	0	3.33E-09
P61247	EVAFRKFKLITE	1479.84	1479.85	0.01	0.0000233
P61247	VQGKNCLTNFHGM	1447.67	1448.65	0.98	2.2E-14
P61247	LKEVVNKLIPD	1266.75	1266.76	0.01	5.14E-10
P61247	LKEVVNKLIPDSIGK	1651.99	1652	0.01	1.54E-14
P61247	SIGKDIEKACQSIYPLH	1900.97	1900.95	-0.02	6.74E-25
P61247	IEKACQSIYPLH	1400.71	1400.71	0	1.26E-10
P61247	VFVRKVKMLKKPKFELGKLMELHGEGSSSGKATG	3716.06	3716.03	-0.03	2.2E-22
P61247	VFVRKVKMLKKPKFELGKLMELHGEGSSSGKATGDETGAKVERA	4772.58	4772.63	0.05	6.63E-23
P62701	ARGPKKHLKRVAAPKHWML	2223.31	2223.32	0.01	1.46E-15
P62701	ARGPKKHLKRVAAPKHWMLD	2338.33	2338.32	-0.01	1.59E-10
P62701	KLTGVFAPRPSTGPHKLRECLPLIIFLRNRLKYALTG	4175.4	4175.45	0.05	3.9E-09
P62701	KTGENFRLIY	1239.66	1239.67	0.01	0.000000942
P62701	TKGRFAVHRITPEEAKYKLCKVRKIFVGTKGIPHLVTH	4357.48	4357.45	-0.03	1.21E-33
P62701	TKGRFAVHRITPEEAKYKLCKVRKIFVGTKGIPHLVTHD	4472.51	4472.48	-0.03	0.000000512
P62701	LETGKITDFIKF	1410.78	1410.79	0.01	0.0000432
P62701	TGNLCMVTGGANLGRIGVITNRERHPGSF	3026.53	3026.5	-0.03	0.000276
P62701	TGNLCMVTGGANLGRIGVITNRERHPGSFD	3141.56	3142.88	1.32	0.00000911
P62701	TGNLCMVTGGANLGRIGVITNRERHPGSFDVVHVK	3703.92	3703.91	-0.01	2.83E-14
P62701	ANGNSFATRLSNIFVIGKGNKPWISLPRGKGIRLTIAEER	4380.43	4380.39	-0.04	1.71E-10

P46782	IKLFGKWST	1078.62	1078.61	-0.01	2.91E-18
P46782	STRIGRAGTVRRQAVD	1741.97	1741.97	0	0.000178
P62753	MKLNISFPATGCQKLIE V	1991.06	1991.05	-0.01	9.85E-13
P62753	ANLSVLNLVIVKKGEK	1724.06	1724.07	0.01	2.23E-22
P62753	ANLSVLNLVIVKKGEKD	1839.08	1839.08	0	5.19E-13
P62753	ANLSVLNLVIVKKGEKD IPGLT	2320.37	2320.33	-0.04	0.0000004 78
P62081	MFSSSAKIVKPNGEKP	1760.91	1761.93	1.02	7.91E-10
P62081	MFSSSAKIVKPNGEKPD	1875.94	1877.03	1.09	1.01E-09
P62081	MFSSSAKIVKPNGEKPD EFESGISQALLELEMNS	3753.81	3753.77	-0.04	3.39E-20
P62081	AILEDLVFPSEIVGKRIR VKL	2394.44	2394.45	0.01	3.29E-16
P62081	LVFPSEIVGKRIRVKL	1853.16	1853.16	0	3.89E-13
P62081	LVFPSEIVGKRIRVKLD	1968.19	1968.19	0	0.000515
P62081	GSRLIKVHL	1021.64	1021.64	0	7.25E-10
P62081	KAQQNNVEHKVETFSG VYKKLTGK	2732.46	2732.44	-0.02	8.07E-27
P62081	KAQQNNVEHKVETFSG VYKKLTGKD	2847.49	2847.45	-0.04	1.42E-12
P62081	KAQQNNVEHKVETFSG VYKKLTGKDVNFEFPEF QL	4098.09	4098.15	0.06	1.74E-10
P62081	VNFEFPEFQL	1268.61	1268.62	0.01	1.23E-15
P62241	VVYNASNNELVRTKTL VKNCIVLI	2702.52	2702.54	0.02	1.31E-15
P62241	VVYNASNNELVRTKTL VKNCIVLID	2817.54	2817.5	-0.04	4.55E-09
P62241	STPYRQWYESHYALPL GRKKGAKLTPEEEEILN KKRSKKIQKKY	5318.9	5319.88	0.98	0.0000085 2
P62241	GYVLEGKELEFYLRKIK ARKGK	2624.52	2624.53	0.01	3.27E-20
P46781	PRRLFEGNALLRRLVRI GVL	2347.44	2347.41	-0.03	1.8E-15
P46781	PRRLFEGNALLRRLVRI GVLD	2462.47	2462.52	0.05	0.0000902
P46781	PRRLFEGNALLRRLVRI GVLDEGKMKL	3148.85	3148.88	0.03	0.000307
P46781	YILGLKIE	947.569	947.569	0	3.15E-18
P46781	YILGLKIEDFLERRLQTQ VFKLGLAKSIHHARVLI RQRHIRVRKQVVNIPSFI VRL	6712.99	6713.03	0.04	0.0000013 7
P46781	FLERRLQTQVFKLGLAK SIHHARVLIRQRHIRVRK QVVNIPSFIVRL	5668.4	5668.47	0.07	0.0000009 15
P46783	MLMPKKNRIAIYELLFK EGVMVAKK	2949.67	2950.77	1.1	4.75E-31
P46783	MLMPKKNRIAIYELLFK EGVMVAKKD	3064.7	3064.68	-0.02	1.1E-30
P46783	MLMPKKNRIAIYELLFK EGVMVAKKDVHMPKH PELA	4204.29	4206.28	1.99	1.83E-19
P46783	KNVPNLHVMKAMQSL KSRGYVKEQFAWRHFY	5328.74	5328.69	-0.05	3.78E-16

	WYLTNEGIQYLR				
P46783	YLHLPPEIVPATLRRSRP ETGRPRPKGLEGERPAR LTRGEA	4628.56	4628.52	-0.04	3.6E-12
P62280	ADIQTERAYQKQPTIFQ NKKRVLLGETGKEKLP RYYKNIGLGFKTPKEAI EGTYI	6411.47	6412.49	1.02	0.0000004 59
P62280	ADIQTERAYQKQPTIFQ NKKRVLLGETGKEKLP RYYKNIGLGFKTPKEAI EGTYID	6526.5	6526.5	0	0.0000011 4
P62280	IQTERAYQKQPTIFQNK KRVLLGETGKEKLPRY YKNIGLGFKTPKEAIEG TYI	6183.4	6182.4	-1	5.29E-19
P62280	IQTERAYQKQPTIFQNK KRVLLGETGKEKLPRY YKNIGLGFKTPKEAIEG TYID	6298.43	6298.46	0.03	3.03E-14
P62280	YLHYIRKYNRFEKRHK NMSVHLSPCFR	3521.82	3521.83	0.01	4.31E-14
P62280	YLHYIRKYNRFEKRHK NMSVHLSPCFRD	3636.85	3636.82	-0.03	0.0000146
P62280	YLHYIRKYNRFEKRHK NMSVHLSPCFRDVQIG	4034.08	4034.04	-0.04	0.0000047 5
P62280	VQIGDIVTVGECRPLSK TVRFNVLKVTKAAGTK KQFQKF	4333.45	4332.39	-1.06	2.38E-21
P62280	IVTVGECRPLSKTVRFN VLKVTKAAGTKKQFQK F	3821.19	3821.15	-0.04	1.77E-25
P25398	VNTALQEVLKTALIH	1648.95	1648.97	0.02	1.85E-26
P25398	GLARGIREAAKAL	1324.79	1324.8	0.01	0.0000003 2
P25398	EPMYVKLVEALCAEHQI NLIKV	2539.35	2539.31	-0.04	0.0000115
P25398	NKKLGEWVGLCKI	1486.83	1486.86	0.03	1.62E-08
P25398	YGKESQAKDVIEEYFKC KK	2292.15	2292.13	-0.02	7.55E-16
P25398	VIEEYFKCKK	1285.67	1285.67	0	0.0000153
P62277	GRMHAPGKGLSQSALP YRRSVPTWLKLTS	3193.73	3193.79	0.06	3.05E-13
P62277	GRMHAPGKGLSQSALP YRRSVPTWLKLTSD	3308.76	3308.79	0.03	3.3E-09
P62277	DVKEQIYKLAKKGLTPS QIGVILR	2696.6	2696.56	-0.04	4.03E-24
P62277	VKEQIYKLAKKGLTPSQ IGVILR	2581.57	2581.54	-0.03	2.87E-37
P62277	VKEQIYKLAKKGLTPSQ IGVILRD	2696.6	2696.62	0.02	7.19E-16
P62277	SHGVAQVRFVTGNKILR ILKSKGLAP	2788.66	2789.7	1.04	4.15E-17
P62277	SHGVAQVRFVTGNKILR ILKSKGLAPD	2903.68	2903.64	-0.04	1.24E-11
P62277	LPEDLYHLIKKAVAVRK HLERNRK	2925.71	2925.72	0.01	4.12E-11

P62277	LPEDLYHLIKKAVAVRKHLERNRKD	3040.74	3040.74	0	0.00000021
P62277	LYHLIKKAVAVRKHLERNRK	2471.51	2471.51	0	2.46E-11
P62277	LYHLIKKAVAVRKHLERNRKD	2586.53	2586.52	-0.01	4.55E-13
P62277	AKFRLILIESRIHRLARYYKTKRVLPPNWKYESSTASALVA	4844.74	4844.74	0	3.76E-08
P62263	TFVHVTDLSGKETICRVTGGMKVKA	2676.41	2676.42	0.01	9.98E-15
P62263	LSGKETICRVTGGMKVKA	1877.02	1877.02	0	1.84E-21
P62263	VAQRCKELGITALHIKLRATGGNRTKTPGPGAQSALRALARSGMKIGRIE	5251.95	5251.91	-0.04	1.62E-17
P62841	AEVEQKKKRTFRKFTYRGV	2412.34	2412.32	-0.02	1.23E-15
P62841	AEVEQKKKRTFRKFTYRGVD	2527.37	2527.37	0	0.00000115
P62841	AEVEQKKKRTFRKFTYRGVDL	2640.45	2640.47	0.02	3.62E-09
P62841	MIILPEMVGSMVGVYNGKTFNQVEIKPEMIGHYLGEFSITYKPVKHGRPGIGATHSSRFIPLK	6998.66	6998.72	0.06	2.55E-11
P62244	ALKSINNAEKRGKRQVLIRPCSKVIVRFLTVMMKHGYIGEFEIID	5199.87	5199.95	0.08	0.00000328
P62244	DHRAGKIVVNLTGRLNKCGVISPRF	2749.53	2749.52	-0.01	0.000188
P62244	HRAGKIVVNLTGRLNKCGVISPRF	2634.5	2634.52	0.02	4.21E-19
P62244	LEKWQNNLLPSRQFGFIVLTTSAGIMDHEEARRKHTGGKILGFFF	5188.72	5189.81	1.09	1.36E-17
P62244	HEEARRKHTGGKILGFFF	2129.13	2129.09	-0.04	0.0000106
P62249	PSKGPLQSVQVFGRKKTATAVAHCKRGNGLIKVNGRPLEMIEPRTLQYKLLEPVLLLGKERFAGV	7136.03	7136.11	0.08	3.96E-14
P62249	IRVRVKGGGHVAQIYAIRQSISKALVAYYQKYV	3734.13	3734.13	0	1.82E-21
P62249	IRVRVKGGGHVAQIYAIRQSISKALVAYYQKYVD	3849.15	3849.19	0.04	6.74E-12
P62249	PRRCESKKFGGPGARARYQKSYR	2697.41	2697.36	-0.05	1.04E-10
P08708	GRVRTKTVKKAARVIIEKYYTRLGN	2919.72	2920.66	0.94	0.0000702
P08708	GRVRTKTVKKAARVIIEKYYTRLGND	3034.75	3034.75	0	1.83E-08
P08708	TKEMLKLL	974.583	974.593	0.01	9.37E-20
P08708	TKEMLKLLD	1089.61	1089.6	-0.01	2.38E-09
P08708	TKEMLKLLDFGSLSNLQVTQPTVGMNFKTPRGPV	3745.99	3745.99	0	3.58E-16
P08708	FGSLSNLQVTQPTVGMNFKTPRGPV	2674.39	2675.38	0.99	1.02E-21

P62269	SLVIPEKFQHILRVLNTNI	2275.31	2275.3	-0.01	4.21E-30
P62269	SLVIPEKFQHILRVLNTNID	2390.33	2390.37	0.04	4.88E-14
P62269	GRRKIAFAITAIKGVGRRYAHVVLRKA	3006.83	3007.84	1.01	3.23E-18
P62269	GRRKIAFAITAIKGVGRRYAHVVLRKAD	3121.86	3121.86	0	0.00000104
P62269	IDLTKRAGELTE	1344.72	1344.73	0.01	0.00000131
P39019	PGVTVKDVNQQEFVRALAAFLKKSGKLKVPEWV	3681.08	3681.06	-0.02	7.16E-30
P39019	PGVTVKDVNQQEFVRALAAFLKKSGKLKVPEWVD	3796.1	3796.09	-0.01	0.000000148
P39019	VNQQEFVRALAAFLKKSGKLKVPEWV	2984.7	2984.7	0	1.73E-23
P39019	VNQQEFVRALAAFLKKSGKLKVPEWVD	3099.72	3099.7	-0.02	1.02E-17
P39019	VNQQEFVRALAAFLKKSGKLKVPEWVDTVKLAKHKELAPY	4578.58	4578.54	-0.04	1.64E-11
P39019	TVKLAKHKELAPY	1496.87	1496.87	0	1.4E-17
P39019	TVKLAKHKELAPYD	1611.9	1611.9	0	4.65E-10
P39019	GGRKLTPQGQRDLDRIAGQVAAANKKH	2884.59	2884.59	0	0.0000094
P39019	LDRIAGQVAAANKKH	1590.9	1590.91	0.01	6.7E-09
P60866	AFKDTGKTPVEPEVAIHRIRITLTSRNVKSLEKVCA	4048.23	4047.23	-1	0.00000324
P60866	RFQMRIHKRLI	1496.89	1496.9	0.01	2.38E-10
P60866	RFQMRIHKRLID	1611.91	1611.92	0.01	0.00000867
P63220	LYVPRKCSASNRIIGAK	1875.05	1875.05	0	1E-10
P63220	LYVPRKCSASNRIIGAKD	1990.08	1990.05	-0.03	0.00000395
P63220	KVTGRFNGQFKTYAICGAIRRMGES	2789.42	2789.42	0	0.000000738
P62266	GCLNFIEENDEVLVAGFGRKGHAVG	2630.29	2630.26	-0.03	1.95E-19
P62266	EVLVAGFGRKGHAVGD	1610.85	1611.92	1.07	0.000421
P62266	IPGVRFKVVKVANVSLLALYKGKKERPRS	3251.97	3252.04	0.07	8.38E-37
P62847	TVTIRTRKFMTNRLLQRKQMVID	2847.61	2847.6	-0.01	0.00000117
P62847	VLHPGKATVPKTEIREKLAKMYKTTP	2935.67	2935.67	0	5.34E-15
P62847	VLHPGKATVPKTEIREKLAKMYKTTPD	3050.7	3050.7	0	4.36E-18
P62847	VLHPGKATVPKTEIREKLAKMYKTTPDVIFVFGFRTHFGGGKTTGFGMIY	5570.97	5572.01	1.04	0.000264
P62847	VIFVFGFRTHFGGGKTTGFGMIY	2538.29	2538.28	-0.01	7.51E-16

P62851	KLNNLVLF	959.58	959.6	0.02	1.64E-10
P62851	KATYDKLCKEVPNYKLITPAVVSERLKIRGSLARAALQELLSKGLIKLVSKHRAQVIYTRNTKGG	7233.15	7233.06	-0.09	8.32E-11
P62851	KLCKEVPNYKLITPAVVSERLKIRGSLARAALQELLSKGLIKLVSKHRAQVIYTRNTKGG	6654.88	6655	0.12	6.54E-28
P62854	KAIKKFVIRNIVEAAAVR	2025.26	2025.23	-0.03	9.21E-37
P62854	KAIKKFVIRNIVEAAAVRD	2140.28	2140.26	-0.02	5.23E-29
P62854	KAIKKFVIRNIVEAAAVRDISEASVF	2873.65	2873.69	0.04	8.09E-09
P62854	AYVLPKLYVKLHYCVSCAIHSKVVRNRSREARK	3886.15	3886.14	-0.01	8.66E-12
P62854	AYVLPKLYVKLHYCVSCAIHSKVVRNRSREARKD	4001.17	4001.13	-0.04	0.0000363
P62854	RTPPPRFRPAGAAPRPPPKPM	2293.27	2293.25	-0.02	1.47E-11
P42677	LLHPSPEEEKRKHKKKRLVQSPNSYFM	3305.78	3305.79	0.01	0.000218
P62857	MDTSRVQPIKLARVTKVLGRTGSQGQCTQVRVEFM	3961.08	3961.06	-0.02	0.000000054
P62857	TSRVQPIKLARVTKVLGRTGSQGQCTQVRVEFM	3673	3673.03	0.03	1.58E-14
P62857	DTSRSIIRNVKGPVREG	1883.03	1883.02	-0.01	0.00000149
P62857	TSRSIIRNVKGPVREG	1768.01	1768.03	0.02	1.07E-15
P62857	TSRSIIRNVKGPVREGD	1883.03	1883.01	-0.02	0.000000226
P62857	TSRSIIRNVKGPVREGDVLTLLESEREARRLR	3705.09	3705.15	0.06	4.22E-27
P62857	VLTLLESEREARRLR	1840.06	1840.07	0.01	1.52E-17
P39023	SHRKFSAPRHGSLGFLPRKRSSRHRGKVKSFPK	3828.16	3828.1	-0.06	3.74E-14
P39023	PSKPVHLTAFLGYKAGMTHIVREV	2650.44	2650.43	-0.01	8.35E-37
P39023	PSKPVHLTAFLGYKAGMTHIVREVD	2765.47	2765.45	-0.02	6.25E-29
P39023	FSSMKKYCQVIRVIAHTQMRLLPLRQKKAHLMEIQVNGGTVAEKL	5192.84	5193.85	1.01	1.6E-11
P39023	FVMLKGCVVGTKKRVLTLRKSLLVQTKRRALEKI	3911.39	3911.51	0.12	4.37E-15
P39023	FVMLKGCVVGTKKRVLTLRKSLLVQTKRRALEKID	4026.42	4026.33	-0.09	1.73E-11
P39023	FVMLKGCVVGTKKRVLTLRKSLLVQTKRRALEKIDLKFI	4527.75	4527.73	-0.02	0.00000161
P39023	FVMLKGCVVGTKKRVLTLRKSLLVQTKRRALEKIDLKFID	4642.77	4642.88	0.11	0.00000281

P39023	TTSKFGHGRFQTMEEKKAFMGPLKK	2883.49	2883.45	-0.04	3.09E-24
P39023	TTSKFGHGRFQTMEEKKAFMGPLKKDRIAKEEGA	3852.98	3853.97	0.99	1.96E-20
P36578	ACARPLISVYSEKGESSGKNVTLPAVFKAPIRP	3526.9	3526.9	0	3.14E-31
P36578	ACARPLISVYSEKGESSGKNVTLPAVFKAPIRPD	3641.92	3641.95	0.03	9.28E-26
P36578	KVEGYKKTKEAVLLLKKLKAWN	2586.56	2588.59	2.03	0.0000487
P36578	KVEGYKKTKEAVLLLKKLKAWND	2701.59	2701.59	0	5.74E-11
P36578	NGIIKAFRNIPGITLLNVSKLNILKLAPGGHVGRFCIWTESAFRKL	5074.89	5074.84	-0.05	2.9E-23
P36578	KAAAAAAALQAKSDEKAAVAGKKPVVGKKGKKAAVGVKKQKKPLVGKKAAATKKPAPEKKPAEKKPTTEEKKPAA	7576.52	7576.39	-0.13	2.45E-09
P36578	EKAAVAGKKPVVGKKGKKAAVGVKKQKKPLVGKKAAATKKPAPEKKPAEKKPTTEEKKPAA	6308.84	6308.82	-0.02	0.0000118
P46777	GFVKVVKNKAYFKRYQVKFRRRREGKT	3387.96	3387.93	-0.03	3.05E-12
P46777	GFVKVVKNKAYFKRYQVKFRRRREGKTD	3502.99	3503	0.01	1.57E-08
P46777	KNKYNTPKYRMIVRVTNR	2280.26	2280.28	0.02	0.00000268
P46777	AGLARTTTGNKVFGALKGAV	1931.1	1931.1	0	4.32E-23
P46777	AGLARTTTGNKVFGALKGAVD	2046.12	2046.11	-0.01	3.55E-22
P46777	GGLSIPHSTKRFPGY	1615.85	1615.83	-0.02	1.97E-10
P46777	SESKEFNAEVHRKHIMGQNVA	2410.18	2410.19	0.01	0.000132
P46777	YMRYLMEE	1133.49	1133.49	0	9.83E-10
P46777	AYKKQFSQYIKNSVTP	1901	1900.98	-0.02	0.00027
P46777	AYKKQFSQYIKNSVTPD	2016.03	2016	-0.03	0.00023
P46777	RVAQKKASFLRAQERAAES	2145.18	2145.16	-0.02	3.47E-12
Q02878	KNGGTRVVKLRKMPRYYPTE	2392.32	2392.3	-0.02	0.000000365
Q02878	VPRKLLSHGKKPFSQHVRKLRASITPGTILIILTGRHRGKRVVFLKQLASGLLLVTGPLVLNRVPLRRTHQKFVIATSTKI	9059.48	9061.51	2.03	1.89E-11
Q02878	VPRKLLSHGKKPFSQHVRKLRASITPGTILIILTGRHRGKRVVFLKQLASGLLLVTGPLVLNRVPLRRTHQKFVIATSTKID	9174.51	9174.39	-0.12	0.00000822
Q02878	ISNVKIPKHLT	1248.76	1248.75	-0.01	1.16E-19
Q02878	ISNVKIPKHLTD	1363.78	1363.79	0.01	9.97E-20

Q02878	ISNVKIPKHLTDAYFKKKKLRKPRHQEGEIF	3748.14	3748.08	-0.06	4.23E-13
Q02878	AYFKKKKLRKPRHQEGEIF	2402.37	2402.35	-0.02	6.23E-13
Q02878	AYFKKKKLRKPRHQEGEIFD	2517.4	2517.41	0.01	0.00036
Q02878	AYFKKKKLRKPRHQEGEIFDTEKEKYEITEQRKI	4293.34	4293.35	0.01	2.97E-24
Q02878	TEKEKYEITEQRKID	1908.98	1908.95	-0.03	1.04E-11
Q02878	SQILPKIKAIPQLQGYLRSVFALTNGIYPHKLVF	3852.22	3852.23	0.01	7.04E-23
P18124	NALIARSLGKYGIICME	1850.97	1850.96	-0.01	2.57E-23
P18124	LIHEIYTVGKRFKEANNFLWPFKLSSPRGGMKKKTTHFVEGG	4847.58	4847.54	-0.04	4.15E-24
P62424	PKGKKAKGKKVAPAPAVVKKQEAKKVVNPLFEKRPKNFGIGQ	4537.72	4537.7	-0.02	5.3E-36
P62424	PKGKKAKGKKVAPAPAVVKKQEAKKVVNPLFEKRPKNFGIGQD	4652.75	4652.71	-0.04	3.1E-14
P62424	PKGKKAKGKKVAPAPAVVKKQEAKKVVNPLFEKRPKNFGIGQDIQPKR	5275.14	5276.08	0.94	0.00000769
P62424	LTRFVKWPRYIRLQRQRAILYKRLKVPPAINQFTQAL	4580.7	4580.75	0.05	0.0000492
P62424	RQTATQLLKLAHKYRPETKQEKKQRLLARAEKKAAGKG	4372.54	4372.5	-0.04	4E-28
P62424	RQTATQLLKLAHKYRPETKQEKKQRLLARAEKKAAGKGD	4487.57	4487.54	-0.03	3.41E-18
P62424	VPTKRPPVLRAGVNTVTTLVENKKAQLVVIAH	3448.04	3448.01	-0.03	1.43E-22
P62424	VPTKRPPVLRAGVNTVTTLVENKKAQLVVIAHD	3563.07	3563.09	0.02	8.6E-28
P62424	VPTKRPPVLRAGVNTVTTLVENKKAQLVVIAHDV	3662.14	3662.14	0	1.01E-14
P62424	PIELVVFLPALCRKMGVPYCIIKGKARLGRLVHRKTCTTVAFTQVNSE	5354.95	5355.92	0.97	2.95E-16
P62424	PIELVVFLPALCRKMGVPYCIIKGKARLGRLVHRKTCTTVAFTQVNSED	5469.98	5470.06	0.08	0.000181
P62424	KGALAKLVEAIRTNYN	1759.99	1760.99	1	7.73E-21
P62424	KGALAKLVEAIRTNYND	1875.02	1875.02	0	8.99E-09
P62424	KGALAKLVEAIRTNYNDRY	2194.19	2194.2	0.01	1.38E-12
P62917	GRVIRGQRKGAGSVFRAHVKHRKGAARLRAV	3394.01	3394.04	0.03	1.59E-10
P62917	GRVIRGQRKGAGSVFRAHVKHRKGAARLRAVD	3509.04	3509.05	0.01	1.35E-13
P62917	GRVIRGQRKGAGSVFRAHVKHRKGAARLRAVDFAERHGYIKGIVK	5007.88	5007.92	0.04	0.00000368
P62917	FAERHGYIKGIVK	1516.85	1516.84	-0.01	3.16E-23

P62917	FAERHGYIKGIVKD	1631.88	1631.88	0	6.02E-12
P62917	FAERHGYIKGIVKDIIH	1995.11	1995.11	0	6.83E-40
P62917	FAERHGYIKGIVKDIIHD	2110.13	2110.14	0.01	2.58E-24
P62917	PGRGAPLAKVVFR	1366.82	1366.82	0	1.8E-25
P62917	PGRGAPLAKVVFRD	1481.85	1481.85	0	3.05E-22
P62917	PYRFKKRTELFIAAEGIH TGQFVYCGKKAQLNIG NVLPVGTMPEGTIVCCL EEKPG	6150.17	6150.29	0.12	7.64E-12
P62917	RGKLARASGNYATVISH NPETKKTRVKLPSGSKK VISSANRAVVGVVAGG GRI	5442.09	5442.15	0.06	1.17E-37
P62917	RGKLARASGNYATVISH NPETKKTRVKLPSGSKK VISSANRAVVGVVAGG GRID	5557.12	5557.07	-0.05	3.11E-29
P32969	MKTILSNQTV	1133.61	1133.61	0	3.91E-27
P32969	IPENVDITLKGRTVIVKG PRGTLRR	2787.66	2787.62	-0.04	0.0000176
P32969	ITLKGRTVIVKGPRGTL RR	2120.34	2120.35	0.01	9.72E-15
P32969	FNHINVELSLLGKKKKR LRV	2391.46	2391.44	-0.02	3.1E-20
P32969	FNHINVELSLLGKKKKR LRVD	2506.49	2506.51	0.02	0.000000374
P32969	IELVSNSAALIQQATTV KNK	2127.19	2127.18	-0.01	7.6E-15
P32969	IELVSNSAALIQQATTV KNKDIRKFL	2899.65	2899.65	0	0.000081
P32969	GIYVSEKGTVQQA	1378.71	1378.71	0	1.63E-10
	GRRPARCYRYCKNKPY PKSRFCRGVP	3157.62	3157.59	-0.03	0.000641
	GRRPARCYRYCKNKPY PKSRFCRGVPD	3272.65	3273.67	1.02	0.000539
	AKIRIFDLGRKKAKV	1742.1	1743.09	0.99	2.79E-16
	AKIRIFDLGRKKAKVD	1857.13	1857.1	-0.03	3.79E-15
	MVAEKRLIP	1055.62	1055.61	-0.01	0.00000235
P62906	TLYEAVREVLHGNQRK RRKFLETVELQISLKNY	4030.22	4030.18	-0.04	1.07E-12
P62906	IPHMDIEALKKLNKNKK LVKKLAKKY	3090.88	3091.71	0.83	3.4E-12
P62906	IEALKKLNKNKKLVKK LAKKY	2497.62	2497.63	0.01	5.77E-24
P62906	IEALKKLNKNKKLVKK LAKKYD	2612.65	2612.68	0.03	3.01E-22
P62906	AFLASESLIKQIPRILGPG LNKAGKFPSLLTHNEN MVAKV	4301.41	4301.38	-0.03	1.98E-14
P62906	AFLASESLIKQIPRILGPG LNKAGKFPSLLTHNEN MVAKVD	4416.44	4416.44	0	3.12E-13
P62913	TGNFGFGIQEHI	1318.63	1318.63	0	9.27E-21
P62913	TGNFGFGIQEHID	1433.66	1433.66	0	5.91E-16
P62913	PSIGIYGL	818.454	818.454	0	6.28E-10

P62913	FYVVLGRPGFSIA	1424.78	1424.82	0.04	2.03E-15
P30050	PNEIKVVYLRCTGGEVGATSALAPKIGPLGLSPKKVG	3719.08	3719.14	0.06	1.93E-22
P30050	PNEIKVVYLRCTGGEVGATSALAPKIGPLGLSPKKVGD	3834.11	3835.22	1.11	0.000000177
P26373	APSRNGMVLKPHFHK	1717.92	1718.91	0.99	0.000000026
P26373	APSRNGMVLKPHFHKD	1832.95	1832.97	0.02	0.0000316
P26373	PRRRNKSTESLQANVQRLKEYRSKLILFPRKPSAPKKG	4476.58	4476.52	-0.06	5.06E-26
P26373	PRRRNKSTESLQANVQRLKEYRSKLILFPRKPSAPKKGD	4591.61	4590.55	-1.06	1.64E-09
P26373	SSAEELKLATQLTGPVMPVRNVYKKEKARVITEEEKNFKAFASLRMARANARLFGIRAKRAKEAAEQ	7545.11	7545.16	0.05	2.01E-17
P40429	KYTEVLKTHGLLV	1499.87	1499.88	0.01	3.84E-20
P50914	VFRRFVEVGRVAYVSFGPHAGKLVAIV	2972.69	2972.68	-0.01	4.93E-13
P50914	VFRRFVEVGRVAYVSFGPHAGKLVAIVD	3087.71	3087.71	0	1.62E-09
P50914	VFRRFVEVGRVAYVSFGPHAGKLVAIVDVI	3299.87	3299.91	0.04	1.68E-09
P50914	VFRRFVEVGRVAYVSFGPHAGKLVAIVDVIDQNRALV	4096.28	4096.25	-0.03	2.88E-16
P50914	FILKFPHSAHQKYVRQAWQKA	2582.4	2582.39	-0.01	1.83E-16
P50914	FILKFPHSAHQKYVRQAWQKAD	2697.43	2697.39	-0.04	0.0000285
P50914	INTKWAATRWAKKIEARERKAKMT	2886.61	2886.6	-0.01	0.000000343
P50914	INTKWAATRWAKKIEARERKAKMTD	3001.64	3001.62	-0.02	0.0000477
P61313	GAYKYIQELWRKKQS	1897.02	1897.02	0	9.23E-13
P61313	GAYKYIQELWRKKQSD	2012.05	2012.03	-0.02	1.38E-08
P61313	STYKFFEVILI	1358.75	1358.76	0.01	4.77E-13
P61313	PFHKAIRRNP	1234.7	1234.69	-0.01	0.0000191
P61313	PFHKAIRRNPD	1349.73	1349.72	-0.01	0.000212
P61313	TQWITKPVHKHREMRGLTSAGRKSRGLGKGHKFHHTIGGSRRAAWRRRNTLQLHRYR	6763.71	6763.71	0	0.0000918
P18621	PENPTKSCKSRGSNLRVHFKNTRETAQAIKGMHIRKATKYLK	4823.6	4823.54	-0.06	1.87E-13
P18621	PENPTKSCKSRGSNLRVHFKNTRETAQAIKGMHIRKATKYLKD	4938.63	4938.66	0.03	1.83E-11
P18621	SLVIEHIQVNKAPKMRRRTYRAHGRINPYMSSPCHIEMILTEKEQIVPKPEEEVAQKKKISQKKLKKQKLMARE	8773.81	8773.73	-0.08	0.000451

Q07020	VRVQEVPKLKVCALRV TSRARSRILRAGGKILTF	3820.29	3820.35	0.06	0.0000001 9
Q02543	LTTAGAVTQCYR	1282.63	1282.62	-0.01	2.37E-08
Q02543	MGARHRARAHSIQIMK VEEIAASKCRRPAVKQF H	3912.09	3913.08	0.99	1.15E-18
Q02543	MGARHRARAHSIQIMK VEEIAASKCRRPAVKQF HD	4027.12	4027.02	-0.1	4.12E-17
Q02543	SKIKFPLPHRVLRRQHK PRFTTKRPNTFF	3632.1	3632.12	0.02	2.71E-12
P84098	SMLRLQKRLASSVLRCG KKKVWL	2741.6	2741.62	0.02	3.08E-16
P84098	SMLRLQKRLASSVLRCG KKKVWLD	2856.63	2856.65	0.02	0.0000001 69
P84098	PNETNEIANANSRQQIR KLIK	2436.32	2436.32	0	0.000866
P84098	PNETNEIANANSRQQIR KLIKD	2551.35	2551.36	0.01	0.0000009 52
P84098	RHMYHSLYLKVKGNVF KNKRILMEHIHKLKA	3831.15	3832.1	0.95	0.000529
P46778	MTNTKGKRRGTRYMFS RPFRKHGVVPLATYMRI YKKG	4431.4	4431.41	0.01	0.000786
P46778	IKGMGTVQKGMPHKCY HGKTGRVYNVTQHAVG IVVNKQVKGKILAKRIN VRIEHIKHSKSR	6854.84	6854.78	-0.06	3.64E-13
P35268	APVKKLVVKGGKKKKQ VLKFTL	2436.6	2436.61	0.01	1.06E-27
P35268	APVKKLVVKGGKKKKQ VLKFTLD	2551.63	2551.63	0	6.46E-11
P35268	AANFEQFLQERIKVNGK AGNLGGGVVTIERSKSK ITVTSEVPFSKRYLKYLT KKYLKKNNLR	7056.95	7058	1.05	0.000626
P62829	SKRGRGGSSGAKFRISL GLPVGAVINCA	2799.53	2799.53	0	7.77E-14
P62829	SKRGRGGSSGAKFRISL GLPVGAVINCAD	2914.56	2916.61	2.05	0.000192
P62829	NTGAKNLYIISVKGIKG RLNRLPAAGVG	2879.68	2880.7	1.02	7.59E-21
P62829	MVMATVKKGKPELRK KVHPAVVIRQRKSYRR K	3817.28	3817.24	-0.04	7.02E-13
P62829	MVMATVKKGKPELRK KVHPAVVIRQRKSYRR KD	3932.3	3932.24	-0.06	1.41E-09
P62829	NAGVIVNNKGEMKGSA ITGPVAKECA	2557.3	2557.28	-0.02	1.64E-27
P62829	LWPRIASNAGSIA	1354.74	1354.74	0	2.88E-10
P62750	VKANKHQIKQAVKKLY	1895.15	1895.13	-0.02	8.2E-16
P62750	VKANKHQIKQAVKKLY DI	2123.26	2123.24	-0.02	0.00068
P62750	IDVAKVNTLIRP	1337.8	1337.8	0	0.0000125
P62750	ALDVANKIGII	1125.68	1125.7	0.02	2.52E-21
P62750	VANKIGII	826.528	826.528	0	2.76E-20

P83731	MKVELCSFSGYKIYPGH GRRYART	2818.42	2818.41	-0.01	7.33E-15
P83731	IMAKRNQKPEVRKAQR EQAIRAAKEAKKAKQA SKKTAMAAAKAPTKAA PKQKIVKPVKVSAPRVG GKR	7354.34	7355.3	0.96	1.46E-21
P61254	MKFNPFVTS	1069.53	1069.53	0	1.28E-18
P61254	MKFNPFVTSD	1184.55	1184.55	0	2.47E-14
P61254	DEVQVVRGHYKGQQIG KVVQVYRKKYVIYIERV QREKANGTTVHVGIHPS KVVITRLKL	6838.88	6837.98	-0.9	0.0000417
P61254	RKKILERKAKSRQVGKE KGKYKEETIEKMQE	3775.13	3775.11	-0.02	8.23E-17
P61353	GKFMKPGKVVLVLAGR YSGRKAVIVKNI	3027.83	3027.83	0	3.06E-19
P61353	GKFMKPGKVVLVLAGR YSGRKAVIVKNID	3142.85	3142.79	-0.06	4.45E-17
P61353	RPYSHALVAGI	1182.65	1182.66	0.01	7.16E-09
P61353	RPYSHALVAGIDRYPRK VTAAMGKKKIAKRSKI KSFVKVYNYNHLMPTR YSV	6036.35	6036.43	0.08	0.000213
P61353	RYPRKVTAAMGKKKIA KRSKIKSFVKVYNYNHL MPTRYSV	4756.68	4755.66	-1.02	7.36E-19
P61353	RYPRKVTAAMGKKKIA KRSKIKSFVKVYNYNHL MPTRYSVD	4871.7	4871.69	-0.01	1.01E-10
P61353	IPLDKTVVNKDVFR	1642.94	1642.95	0.01	3.18E-19
P61353	PALKRKARREAKVKFE ERYKTGKNKWFFQKLR F	4183.39	4183.42	0.03	4.25E-21
P46776	KYHPGYFGKVGMKHY HLKRNQSFCPTVNL	3448.74	3448.72	-0.02	5.55E-09
P46776	KLWTLVSEQTRVNAAK NKTGAAPII	2708.53	2708.56	0.03	2.15E-20
P46776	KLWTLVSEQTRVNAAK NKTGAAPIID	2823.56	2823.55	-0.01	0.0000229
P46776	VVRSGYYKVLGKGKLP KQPVIVKAKFFSRRAEE KIKSVGGACVLVA	5002.91	5002.89	-0.02	3.43E-09
P46779	SAHLQWMVVRNCSSFLI KRNKQTYSTEPNNLKA RNSFRYNGLIHRKTVGV EPAA	6271.25	6271.2	-0.05	1.08E-09
P46779	LRMAAIRRASAILRSQK PVMVKRKRTRPTKSS	3692.19	3692.17	-0.02	0.0000799
	PKFLRNMRFAKKHNKK GLKKMQANNAKAMSA RAEAIKALVKPKEVKPK IPKGVSRKL	6450.8	6451.82	1.02	5.85E-20
	PKFLRNMRFAKKHNKK GLKKMQANNAKAMSA RAEAIKALVKPKEVKPK IPKGVSRKLD	6565.82	6566.7	0.88	9.01E-15
P62888	IIRSMPEQTGEK	1387.71	1387.7	-0.01	8.84E-16

P62899	TRLNKAVWAKGIRNVPYRIRVRLSRKRNE	3549.09	3549.07	-0.02	0.0000804
P62899	SPNKLYTLVTYVPVTTFKNLQTVNV	2838.55	2838.56	0.01	3.54E-16
P62899	SPNKLYTLVTYVPVTTFKNLQTVNVD	2953.58	2953.54	-0.04	4.47E-15
P62910	AALRPLVKPKIVKKRTKKFIRHQS	2841.8	2841.79	-0.01	1.71E-30
P62910	AALRPLVKPKIVKKRTKKFIRHQSD	2956.83	2956.83	0	2.19E-18
P62910	RYVKIKRNWRKPRGI	1969.2	1969.2	0	0.000455
P49207	RIKRAFLIEEQKIVVKVLKAQAQSQKAK	3249.98	3250.94	0.96	2.17E-38
P42766	AKIKARDLRGKKKEELLKQL	2364.47	2364.46	-0.01	2.69E-11
P42766	LRGKKKEELLKQL	1581.99	1581.99	0	1.06E-27
P42766	LKVELSQLRVAKVTGGAASKLSKIRVVRKSIARVLTVINQTQKENLRKFYKGKKYKPL	6578.96	6578.89	-0.07	5.15E-09
P42766	LRPKKTRAMRRRLNKHEENLKTKKQQRKERLYPLRKYAVKA	5159.03	5158.91	-0.12	8.42E-11
P18077	SGRLWSKAIFAGYKRGLRNQREHTALLKIEGVYAR	4086.25	4086.22	-0.03	1.41E-35
P18077	SGRLWSKAIFAGYKRGLRNQREHTALLKIEGVYARD	4201.28	4201.26	-0.02	2.35E-08
Q9Y3U8	ALRYPMAVGLNKGHKVTKNVSKPRHSRRRGRLTKHTKFVR	4679.7	4679.67	-0.03	4.22E-20
Q9Y3U8	ALRYPMAVGLNKGHKVTKNVSKPRHSRRRGRLTKHTKFVRD	4794.73	4794.73	0	3.52E-17
Q9Y3U8	MIREVCGFAPYERRAMELLKVSK	2725.42	2725.41	-0.01	3.13E-10
Q9Y3U8	MIREVCGFAPYERRAMELLKVSKD	2840.45	2840.47	0.02	0.0000072
Q9Y3U8	KRALKFIKKRVGTHIRAKRKREELSNVLAAMRKAAAKK	4400.69	4400.7	0.01	4.58E-13
P83881	VNVPKTRRTFCKKCGKHQPHKVTQYKKGK	3423.9	3423.89	-0.01	1.04E-22
P83881	VNVPKTRRTFCKKCGKHQPHKVTQYKKGKD	3538.92	3537.87	-1.05	6.53E-17
P83881	VNVPKTRRTFCKKCGKHQPHKVTQYKKGKDSLYAQGKRRY	4761.58	4761.63	0.05	1.96E-15
P83881	KKRKGQVIQF	1230.76	1230.76	0	0.000113
P63173	PRKIEEIKDFLLTARRK	2112.25	2112.26	0.01	2.56E-27
P63173	PRKIEEIKDFLLTARRKD	2227.28	2227.27	-0.01	8.29E-13
P63173	NVKFKVRCSRYLYTLVITDKEKAEKLKQSLPPGLAVKELK	4632.66	4632.67	0.01	0.000000247
P63173	KEKAEKLKQSLPPGLAVKELK	2333.4	2333.43	0.03	1.01E-23
P05388	RATWKSNYFLKIIQLL	1993.15	1993.15	0	5.81E-11

P05388	MLLANKVPAAARAGAI APCEVTVPAQNTGLGPE KTSFFQALGITTKISRGTI EILS	5722.1	5723.06	0.96	1.47E-10
P05388	YTFPLAEKVKAFLA	1596.89	1596.93	0.04	3.32E-25
P05387	MRYVASYLLAALGGNS SPSAK	2155.11	2155.1	-0.01	2.44E-25
P05387	MRYVASYLLAALGGNS SPSAKDIKKIL	2865.58	2865.56	-0.02	3.99E-14
P05387	RLNKVISELNGKNIE	1725.97	1725.98	0.01	2.98E-17
P05387	RLNKVISELNGKNIED	1841	1841.01	0.01	0.0000030 9
P05387	VIAQGIGKLASVPAGGA VAVSAAPGSAAPAAGS APAAAEEKK	3667.99	3668.08	0.09	2.4E-20

Appendix Table 6. Table of all peptides identified from combined analyses of ECVs using microwave-supported acid hydrolysis.

UniProt	Peptide	Theoretical Mass	Observed Mass	Mass Diff	Confidence
P60710	DIAALVV	699.42	698.88	-0.53	0.067385443
P14069	AEIARLM	802.44	802.43	0.00	0.047413793
Q99N16	GLLMSTGD	808.36	808.46	0.10	0.023866348
P06801	GERILGLG	813.47	813.47	0.00	0.067385443
P40142	PFTIKPL	814.50	814.50	0.01	0.052910052
P52480	LRVNLAM	815.47	815.47	0.00	0.032986112
Q8BTM8	VTSLRPF	818.47	818.47	0.00	0.056155507
P52480	LKFGVEQ	819.45	819.45	0.00	0.099357098
A2AR02	SSSSSSSSD	829.29	830.41	1.11	0.052313883
P06745	IINIGIGGS	842.49	842.49	0.01	0.015968064
Q8BRV5	GPEGPPPSD	851.37	849.60	-1.77	0.088724584
P11983	AVLAVKYT	863.51	863.51	0.00	0.015968064
P97324	AYERLIL	876.51	875.18	-1.32	0.069645204
P27773	ASVVGFFR	881.48	881.48	0.00	0.015968064
P30101	PNIVIAKM	884.52	884.51	0.00	0.0064
Q64727	PGLQKSFL	888.51	888.50	0.00	0.0345679
P51437	TNLYRLL	891.52	891.52	0.00	0.025341131
P25911	FGLARVIE	903.52	903.52	0.00	0.015968064
Q8CH25	EKSSSRSAG	907.44	908.66	1.23	0.086206898
Q04750	GYFAPPKE	907.44	907.44	0.00	0.015968064
P31146	AGPLLISLK	910.59	910.58	0.00	0.0098
A1L317	INGLRKVL	911.59	911.56	-0.03	0.061754387

Q3UP87	AFAPVAEFA	921.46	922.28	0.82	0.030195381
Q9D6I9	SVSIGSFLD	923.46	924.44	0.98	0.053691275
O09172	LTAFAKQF	924.51	924.51	0.00	0.030195381
Q8CGZ0	NLLQPIID	924.53	922.68	-1.85	0.088724584
Q9JHK5	PAYLHYY	925.43	926.48	1.05	0.015968064
Q8BK64	DGNVTGEFT	938.40	939.00	0.60	0.053691275
P80318	VVITEKGIS	944.55	944.56	0.00	0.015968064
P10630	EMLSRGFK	948.48	948.48	0.00	0.033500839
O88342	YSGQGVVKL	949.52	951.16	1.64	0.051515151
Q9R111	ALLINPRAS	953.57	953.57	0.00	0.025341131
P42932	IEAEVPAVK	954.54	954.54	0.00	0.0345679
P63017	ANGILNVSAV	956.53	956.53	0.00	0.015968064
Q9D2V7	TGLVLLTGKG	957.59	957.59	0.00	0.015968064
O88845	DVMQQAHH	964.42	964.54	0.12	0.04989605
P27037	AFAVFLISC	969.50	969.51	0.01	0.047
O89053	VRVSQTTW	975.51	975.50	-0.01	0.033500839
Q4VAA2	MQISEKED	978.43	979.56	1.13	0.015968064
P39054	LPGITKVPVG	979.61	979.61	0.00	0.015968064
Q60605	YVEGLRVF	981.53	982.54	1.02	0.015968064
P34884	RVYINYY	989.50	989.50	0.00	0.054834057
P97324	AYERLILD	991.53	991.53	0.00	0.081063963
Q9JHK5	LGALYLSMK	994.55	994.55	0.00	0.015968064
P62826	PALAAQYEH	998.48	998.48	0.00	0.0016
A0M8Q6	PGAVTVAWKA	998.55	998.55	0.00	0.0084
Q8BTM8	LSKIKVSGLG	1000.63	1000.63	0.00	0.086206898
P99026	IAHMISGFE	1003.48	1003.48	0.00	0.008264462
P63017	GIFEVKSTAG	1007.53	1008.52	1.00	0.015968064
P26443	TYASTIGHY	1011.47	1011.47	0.00	0.015968064
Q05144	SKPVNLGLW	1012.57	1012.57	0.00	0.015968064
P56480	PAPATTFAHL	1024.53	1024.53	0.00	0.046656299
O89053	AGPLLISLKD	1025.61	1025.61	0.00	0.008264462
Q6A037	DLNALSALVL	1027.59	1027.54	-0.05	0.067385443
P19096	IMLATGKLSP	1029.59	1031.48	1.89	0.050613496
Q3THE2	MLASLGKNPT	1030.55	1030.55	0.00	0.015968064
P11276	LRFTNIGPD	1031.54	1031.54	0.00	0.015564202
Q3TRM8	GSLGTLSTRF	1037.55	1037.55	0.00	0.015968064
Q68FD5	LLMVLSPRL	1040.64	1040.64	0.00	0.005988024
P16858	LMAYMASKE	1042.48	1042.48	0.00	0.008264462
Q7TSE6	MMTLLMKK	1042.52	1042.59	0.07	0.086206898
P62746	GKQVELALW	1042.58	1042.58	0.00	0.025341131
Q9D2V7	LVQSAVWSR	1044.57	1044.57	0.00	0.015968064
O88569	YFEEYGKI	1047.49	1047.49	0.00	0.015968064
Q9ET01	IINMLFYH	1049.54	1049.54	0.00	0.015968064
O89053	VGTGAAVLTLGP	1054.60	1054.60	0.00	0.015968064
P06745	VMPEVNRVL	1055.58	1055.58	0.00	0.09041591
P35441	VPIQSIFTR	1059.61	1059.61	0.00	0.015968064

O89053	TNIVYLCGKG	1066.55	1066.55	0.00	0.015968064
P84096	VPILLVGTKK	1066.71	1066.71	0.00	0.043921567
Q9Z2U1	EKGPQLFHM	1067.52	1067.52	0.00	0.015968064
P26039	ILNGSHPVSF	1069.56	1069.56	0.00	0.015968064
P20029	VNGILRVTAE	1070.61	1070.61	0.00	0.015968064
P35527	LTVGNNKTLL	1071.63	1071.63	0.00	0.039
O70456	STLIMQLLR	1073.63	1073.62	0.00	0.005988024
Q8VDD5	NIATLLHQSS	1082.57	1082.56	-0.01	0.015968064
Q61598	SQIFISRAY	1083.57	1083.57	0.00	0.015968064
P09411	KIQLINNML	1085.63	1085.62	0.00	0.015968064
P62334	PLVYNMSHE	1088.50	1088.50	0.00	0.015968064
P51150	VTAPNTFKTL	1090.60	1090.60	0.00	0.015968064
P20357	MTRSTELGSD	1095.49	1095.54	0.06	0.008264462
P60710	NGSGMCKAGFAG	1098.46	1098.46	0.00	0.094441123
P01895	MELVETRPAG	1101.55	1101.55	0.00	0.057061341
P61204	AVLLVFANKQ	1101.65	1101.65	0.00	0.0018
O15355	FIQSKISQR	1105.62	1105.62	0.00	0.018
O88342	ASKVVTVFSVA	1106.63	1106.63	0.00	0.015968064
P26039	VVKLGAASLGAE	1113.64	1113.64	0.00	0.005988024
O89053	SSIRYFEIT	1114.57	1114.56	0.00	0.057061341
P18085	AVLLLFANKQ	1115.67	1115.67	0.00	0.011
P10126	GSASGTTLLEAL	1118.58	1118.58	0.00	0.015968064
P63017	GIFEVKSTAGD	1122.56	1122.56	0.01	0.008264462
P35175	NYELHGYQT	1123.49	1123.50	0.00	0.054014597
P07954	PKIANAIMKAA	1126.65	1126.65	-0.01	0.013
Q05144	SKPVNLGLWD	1127.60	1127.58	-0.01	0.015968064
Q8CEZ4	DKFQVFNIQ	1137.58	1136.63	-0.95	0.077120826
P35527	VNVEINVAPGK	1138.63	1138.63	0.00	0.03
P28063	LLYKYGEAAL	1139.62	1139.62	0.00	0.008264462
P10809	LGKVGEVIVTK	1141.71	1141.70	-0.01	0.033
P0CG49	TIENVKAKIQ	1142.67	1143.65	0.99	0.015968064
Q6PE87	DTIKMQVYF	1143.56	1142.55	-1.02	0.039936103
Q9Z1Q9	EAIALFQKML	1144.63	1144.63	0.00	0.04989605
P26041	AVLEYLKIAQ	1146.66	1146.66	0.00	0.015968064
Q9ET01	MEELEEIEE	1149.47	1149.48	0.01	0.032986112
Q14566	PGAGSQHLEVR	1149.59	1149.59	0.00	0.026
P08905	LSQYIRNCGV	1151.58	1151.58	0.00	0.008264462
Q99729	AGKMFVGGLSW	1151.58	1151.57	-0.01	0.033
Q8VCM7	YAMFRVGPES	1155.54	1155.55	0.01	0.015968064
P68372	LVSEYQQYQ	1156.54	1157.43	0.89	0.053097345
P35908	LEEALQQAKE	1157.59	1157.59	0.00	0.044
P20591	PAAASHPLLLNG	1159.63	1159.63	0.00	0.0021
Q8CIH5	MLLTKPTEASA	1160.61	1160.61	0.00	0.067385443
O88569	YFEEYGKID	1162.52	1162.52	0.00	0.008264462
P08071	KVEVLQQVLL	1167.72	1167.72	0.00	0.015968064
Q01518	VVGIVEIINSK	1169.70	1169.70	-0.01	0.0025

P0CG49	QQRLIFAGKQ	1170.65	1170.65	0.00	0.033500839
P36578	KAAAAAAALQAKS	1170.67	1170.67	0.00	0.0055
P50990	TGANVVVTGGKVA	1171.66	1171.65	-0.01	0.0098
P62310	GVVLVAPPLRVG	1175.74	1175.74	0.00	0.05
P60710	IKEKLCYVAL	1178.67	1178.67	0.00	0.025341131
P08030	ALEPGQRVVIV	1179.70	1179.70	0.00	0.054014597
P19338	TTEETLKESF	1183.56	1183.56	-0.01	0.009
P26039	SALSVVQNLEK	1186.66	1186.66	0.00	0.015968064
Q01853	NSVVSLSQPKM	1188.62	1188.62	0.00	0.015968064
Q3ULZ2	VQMLAKTGTED	1191.58	1193.32	1.74	0.074742265
O88342	ASVKEWTITY	1196.61	1197.61	1.00	0.015564202
P40124	VVGIVEIINSR	1197.71	1197.71	0.00	0.015968064
P16858	NEYGYSNRVV	1199.56	1199.56	0.00	0.015968064
P11276	EEVQIGHVPRG	1201.62	1201.62	0.00	0.015968064
Q6P5F9	QGEVVREFMK	1204.59	1204.59	-0.01	0.044247788
P17141	HRKSLSHSASD	1205.59	1206.50	0.92	0.083281539
P29351	MLMESISTKGL	1208.61	1209.61	1.00	0.005988024
P14618	PILYRPVAVAL	1210.74	1212.76	2.01	0.0079
P62835	TAGTEQFTAMR	1211.56	1211.56	0.00	0.015968064
P62314	PVQLETLSIRG	1211.69	1211.69	0.00	0.019
P61226	TAGTEQFASMR	1213.54	1215.42	1.88	0.015968064
P06732	VSPLLLASGMAR	1213.69	1213.68	-0.01	0.031
Q8VHX6	SPFVVPVASLSD	1216.63	1218.34	1.71	0.069645204
P06454	AAVDTSSEITTK	1221.61	1221.61	0.00	0.017
P60709	GYALPHAILRL	1222.72	1222.71	-0.01	0.046
P27870	QSLANYGRPKI	1228.66	1228.65	0.00	0.010309278
P00505	PILGVTEAFKR	1229.71	1229.71	-0.01	0.00011
P29351	QVTHIRIQNSG	1234.64	1235.64	1.00	0.015968064
P06733	LLKTAIGKAGYT	1234.73	1234.72	-0.01	0.03
O70566	ERSLLLLARAI	1235.77	1236.77	1.00	0.04989605
Q8BZW8	DGAVKHLVGGER	1236.66	1236.71	0.05	0.066115702
O75821	AARAIAGVSGFGY	1238.64	1238.64	-0.01	0.0011
P52480	MVFASFIRKAA	1239.68	1239.68	0.00	0.040089087
P0CG49	DTIENVKAKIQ	1239.68	1240.70	1.02	0.015968064
P0CG49	TIENVKAKIQD	1239.68	1240.67	0.99	0.008264462
P14618	GLISLQVKQKGA	1240.75	1240.74	-0.01	0.015
Q68FD5	NAIITMMNHPT	1241.59	1241.59	0.00	0.015968064
P52480	GLISLQVKEKGA	1241.73	1241.74	0.01	0.015968064
P02538	AELSQMQTHIS	1243.59	1243.58	0.00	0.02
P10126	TVAFVPISGWNG	1246.63	1246.64	0.00	0.027958993
P04264	SIIAEVKAQYE	1249.66	1249.65	0.00	0.038
P60842	VLEVTKKFMR	1249.72	1249.71	-0.01	0.044
O54865	MMEIAGQVQVD	1251.55	1250.51	-1.03	0.054834057
P00558	PVAVELKSLLGK	1252.78	1252.77	-0.01	0.00073
P47738	SKYGLAAAVFTK	1254.70	1254.70	0.00	0.015968064
P19338	PAAMKAAAAAPASE	1255.62	1255.62	0.00	0.0032

P70399	DSQPESQVLEE	1259.55	1259.65	0.10	0.030195381
P10809	KYKNIGAKLVQ	1260.76	1260.75	-0.01	0.025
Q91VH6	SSVSYAAGALTVH	1261.63	1261.62	-0.01	0.015968064
P25705	VPVGEELLGRVV	1265.73	1265.73	0.00	0.00042
Q02053	LSSQFYLREE	1270.62	1270.62	0.00	0.015968064
P19338	RETGSSKGFGFV	1270.63	1270.64	0.01	0.047
Q9JKF1	VGRTLSALRSPD	1270.70	1270.70	0.00	0.096755162
Q4VA53	PVLPARFFTQP	1271.70	1272.70	1.00	0.069498069
Q9JHU4	EYEKLQVLLR	1271.72	1272.74	1.02	0.005988024
P28838	PPLVFVGKGITF	1273.74	1273.74	0.00	0.01
P40142	AIVQAVKGLVTKG	1282.80	1282.80	0.00	0.008264462
P13010	IESKIQPGSQQA	1284.67	1284.67	0.00	0.038
O35286	PLLERYGVIIL	1284.78	1284.78	0.00	0.04989605
Q68FD5	AILGNQMFTHY	1293.62	1293.61	0.00	0.005988024
P60710	IKEKLCYVALD	1293.70	1293.70	0.00	0.015968064
Q641P0	ITYFIQQLLR	1293.74	1294.75	1.01	0.005988024
P08030	ALEPGQRVVIVD	1294.72	1294.75	0.02	0.015968064
Q3UZZ4	LVFLQTPRQPV	1296.76	1296.76	0.00	0.008264462
P31146	PALTAEEWLGGR	1298.66	1298.67	0.01	0.000031
P13020	ENPFAQGALRSE	1299.62	1300.62	1.00	0.015968064
Q99KE1	PFYMGLYQKR	1301.66	1302.66	1.00	0.052313883
P68104	NVGFNVKNVSVK	1303.72	1303.72	-0.01	0.018
P60174	GAFTGEISPGMIK	1306.66	1306.66	0.00	0.044
P10126	FIKNMITGTSQA	1309.67	1310.67	1.00	0.005988024
Q9Y490	PETQVVLINAVK	1309.76	1309.76	0.00	0.0064
P29401	AIAQAVRGLITKA	1310.80	1310.80	-0.01	0.034
P28838	SVVLVGLGKKAAGI	1310.83	1310.82	-0.01	0.047
Q8BV57	HLTQGPTPNHNP	1311.63	1309.63	-2.00	0.015968064
P16858	NEYGYSNRVVD	1314.58	1315.70	1.12	0.008264462
O08602	DESGVIMNKWK	1321.63	1319.66	-1.97	0.086294416
P52480	PILYRPVAVALD	1325.77	1325.77	0.00	0.008264462
P19096	PGSPELQQVLKH	1331.72	1332.72	1.00	0.005988024
Q7TPV2	DSHGKSVSRLTF	1332.68	1333.62	0.94	0.044247788
P60710	ESGPSIVHRKCF	1340.67	1340.67	0.00	0.008264462
P39054	IEQSYINTNHE	1346.61	1346.62	0.01	0.015968064
P40124	SPSKGAVPYVQAF	1349.70	1349.70	0.00	0.053097345
Q4G5Y1	VPPSMSGSCAVCVD	1350.56	1349.84	-0.72	0.053691275
P53396	VLINFASLRSAY	1352.75	1352.75	0.00	0.023
Q9R111	SEIGNFEVGKEF	1354.64	1354.65	0.01	0.015968064
P14174	PMFIVNTNVPRA	1357.72	1357.72	0.00	0.0097
Q9D8N0	ECEQALAAEPKAK	1368.67	1368.67	0.00	0.015968064
P10126	MRQTVAVGVIKAV	1370.81	1370.81	0.00	0.025341131
P17182	VAASEFYRSGKY	1376.67	1376.66	-0.01	0.015968064
P13645	ALEESNYELEGK	1380.64	1380.64	0.00	0.000018
P62805	GVLKVFLENVIR	1385.84	1385.84	0.00	0.015
P62888	IIRSMPEQTGEK	1387.71	1387.71	0.00	0.02

P06745	NMFSGSKINYTE	1389.62	1389.62	0.00	0.015968064
Q9CW03	QVSHRGALTGGYY	1390.66	1390.66	0.00	0.010309278
P11276	VRSYTITGLQPGT	1391.74	1391.74	0.00	0.027958993
P14625	RIMKAQAYQTGK	1393.75	1393.75	0.00	0.029
Q9EQK5	PFPLYPGELLEK	1401.75	1401.75	0.00	0.052313883
Q9WV04	EKYITSKINLVD	1403.77	1405.70	1.94	0.083281539
P40240	EVIKELQEFYK	1406.74	1406.74	0.00	0.005988024
P10126	MVPGKPMCVESFS	1410.63	1410.64	0.00	0.025341131
P0CG49	QQRLIFAGKQLE	1412.78	1412.76	-0.01	0.005988024
P26039	PVQLNLLYVQAR	1412.81	1412.80	-0.02	0.005988024
Q99PI8	TPPGNGSGPRHIND	1417.67	1416.67	-1.00	0.084262699
Q14697	PSVFNGPEVTMLK	1417.73	1418.73	1.01	0.0045
P63017	QGNRTTPSYVAFT	1423.67	1423.68	0.00	0.088861838
Q8VDM4	ELEMLVERLGEK	1426.75	1426.75	0.00	0.005988024
Q9WV32	WAPESNRIVTCGT	1432.68	1432.69	0.01	0.082070708
P02538	AAYMNKVELQAKA	1435.75	1435.75	0.00	0.019
P27797	FGKFVLSSGKFYG	1435.75	1435.74	-0.01	0.022
P25398	TALQEVLKTALIH	1435.84	1435.84	0.00	0.04
P61978	LGGPIITTQVTIPK	1436.86	1436.86	0.00	0.000071
Q8K0E8	PYKKGFGNIATNE	1437.73	1437.73	0.00	0.067385443
P05386	KINALIKAAGVNVE	1438.85	1438.85	0.00	0.0076
Q571I9	IAPLFPAGLVSVVTG	1439.84	1440.84	1.01	0.005988024
P54577	LKNSVEVALNKLL	1439.87	1439.86	-0.01	0.025
P30086	LSKWSGPLSLQEV	1442.78	1442.77	0.00	0.013
Q8BTM8	QHIPGSPFTARVTG	1449.74	1449.73	0.00	0.040089087
Q8VCZ7	SFQPGSPGHLGVIR	1450.77	1452.74	1.98	0.005988024
P00558	GAKSVVLMSHLGRP	1450.81	1450.81	0.00	0.04
P49720	AVSGMGVIVHIIEK	1451.82	1451.81	0.00	0.0074
P31948	PAMRLILEQMQK	1456.79	1456.78	-0.01	0.03
P31254	FIVAASNLRAENY	1466.75	1467.73	0.98	0.005988024
P04264	VKKQISNLQQSIS	1471.84	1471.84	0.00	0.014
Q8QZY1	QKVYELQASRVSS	1476.76	1476.76	0.01	0.015968064
P62316	SVIVVLRNPLIAGK	1477.93	1477.93	0.00	0.00017
P14866	SVQSAQRAKASLNGA	1486.79	1486.78	0.00	0.03
P62880	IETGQQTVGFAGHSG	1487.70	1487.71	0.01	0.015968064
Q80XK6	MGLLELTITAVKSD	1487.79	1487.81	0.02	0.008264462
P17182	VAASEFYRSGKYD	1491.70	1491.71	0.01	0.008264462
Q9DCD0	MQLICEAYHLMK	1494.70	1494.75	0.05	0.099357098
Q8VDD5	LIEKPAGPPGILALL	1500.93	1500.92	0.00	0.010309278
P55884	PPVPAQGEAPGEQAR	1502.75	1503.75	1.01	0.0053
P06733	GAVEKGVPLYRHIA	1508.85	1508.84	-0.01	0.042
Q3TEI4	FVPQTLGFPYARD	1509.76	1511.10	1.34	0.094441123
P10809	AMKKVGRKGVITVK	1513.95	1514.95	1.00	0.0023
P97384	VQELYAAGENRLGT	1519.76	1519.76	0.00	0.015968064
P16858	APMFVMGVNHEKY	1521.71	1521.71	0.00	0.015968064
P25788	IVKEVAKIIYIVH	1523.94	1523.94	-0.01	0.025

Q8K557	GGGAGGGGAGGAGAAGAAGGGPD	1525.65	1526.56	0.91	0.040089087
P0CG49	QQRLIFAGKQLED	1527.80	1528.79	0.99	0.008264462
P38646	IKNVPFKIVRASNG	1541.90	1541.90	-0.01	0.019
P35527	HKEEMSQLTGQNSG	1544.69	1544.69	0.00	0.031
Q61830	FCKAIGGELASIKSK	1550.85	1550.86	0.01	0.040089087
Q5RKZ7	DAPSGPGPTSNQLTHV	1558.74	1557.75	-0.99	0.015968064
P49588	PVRVVSIGVPVSELL	1562.94	1562.94	0.00	0.019
Q5SUA5	EALVGYVAKLTATPR	1569.89	1569.88	-0.01	0.005988024
P11021	AGTIAGLNVMRIINE	1570.85	1570.85	0.00	0.0047
P22626	GGGGNFGPGPGSNFRGGS	1577.70	1577.69	-0.01	0.017
P62874	IETGQQTTTFTGHTG	1577.73	1577.75	0.02	0.015968064
Q13148	AGWGNLVYVVNYPK	1578.82	1578.82	0.00	0.00079
P30046	SWQIGKIGTVMTFL	1579.84	1580.85	1.01	0.019
P00558	SLEPVAVELKSLLGK	1581.93	1581.93	0.00	0.0039
Q96AB3	SGLLGLFQGQNSLLH	1582.85	1582.85	0.00	0.013
P00558	PAKIEAFRASLSKLG	1586.91	1586.91	-0.01	0.00045
Q9CQC9	ETIANVPILILGNKI	1588.95	1588.95	-0.01	0.005988024
P60228	LVKVIQQESYTYK	1597.87	1598.87	1.00	0.022
P23284	NFVALATGEKGFGYK	1600.82	1600.82	-0.01	0.0013
P27797	DEFTHLYTLIVRP	1602.84	1604.85	2.01	0.0084
P36536	ETISNVPILILGNKI	1604.95	1604.95	0.01	0.005988024
O75531	KAYVVLGQFLVLKK	1605.00	1606.00	1.00	0.0036
O08582	AGKSTLLGVLTHGELD	1609.87	1610.89	1.02	0.052313883
P08865	VLKFLAAGTHLGGTNL	1610.91	1610.91	-0.01	0.0037
Q9JHU4	PSGQATEFIMNEYK	1613.74	1612.88	-0.86	0.015968064
P52597	PPLKFMSVQRPGPY	1615.85	1616.86	1.00	0.0013
P40240	EPQRETLKAIHMAL	1617.87	1617.87	0.00	0.005988024
P63017	AAKNQVAMNPTNTVF	1620.79	1620.80	0.01	0.015968064
P06745	QYMHRFAAYFQQG	1628.72	1629.73	1.01	0.005988024
P11276	AIPANGQTPVQRSISP	1634.87	1634.88	0.00	0.015968064
P04264	LEIATYRTLLEGEE	1635.84	1635.83	0.00	0.047
P49773	LGLNKGYRMVVNEGS	1635.84	1635.84	0.00	0.05
P14174	PMFIVNTNVPRASVP	1640.87	1640.87	0.00	0.009
P35527	AAIQKNYSPYYNTI	1644.81	1644.81	0.00	0.026
P25398	VNTALQEVLKTALIH	1648.95	1648.95	0.00	0.046
P28658	IVIAKLVGEEQLTKD	1654.95	1654.93	-0.02	0.066115702
Q8BTM8	PTGQKGTVEPQLEARG	1666.86	1667.85	0.99	0.040089087
Q07021	PELTSTPNFVVEVIK	1671.91	1671.90	0.00	0.028
Q18PE0	MTEAALVEGQVKLRD	1674.86	1673.80	-1.06	0.010309278
Q9Z2U1	PGAMSRPFGVALLFGGV	1674.89	1675.88	0.99	0.010309278
P22314	SNGEQPLSAMVSMVTK	1677.81	1677.80	-0.01	0.011
P62258	LVYQAKLAEQAERY	1680.88	1681.88	1.00	0.000068
Q6NZK8	PYSTLVNSSGHAAHMD	1683.73	1681.76	-1.97	0.051515151
Q9CR30	SRLNPHRSLLGTGNY	1683.88	1684.98	1.10	0.053097345
P30049	AAKANLEKAQAELVGT	1683.92	1683.92	0.00	0.00088

	A				
P10853	TGISSKAMGIMNSFVN	1687.79	1688.78	0.99	0.015968064
P42859	DPVPSLLPATTGALISH	1687.91	1688.84	0.92	0.062992126
P70696	TGISSKAMSIMNSFVT	1688.81	1688.78	-0.03	0.067385443
Q924K8	QFLVVARAVGTFARAL	1700.97	1700.99	0.02	0.005988024
Q969H8	VRPGGVVHSFSHNVGPG	1701.87	1701.87	0.00	0.013
P13796	SKAYYHLLEQVAPKG	1702.90	1702.90	0.00	0.022
Q00610	VNTSAVQVLIEHIGNL	1705.94	1705.93	-0.01	0.019
P68363	PYNSILTTHTTLEHS	1712.84	1712.83	-0.01	0.0022
Q8VDW0	EEPQAPQESTPAPPKK	1714.85	1714.86	0.01	0.069645204
Q4VA53	QFAAPLKSLVATFIVK	1715.00	1716.00	1.00	0.005988024
P10809	VQPHDLGKVGEVIVTK	1717.97	1717.96	-0.01	0.000093
P53634	LLGTWVFQVGSSGSQR	1720.89	1720.89	0.00	0.000012
Q62101	TEMALAEGQSGGAEMQD	1723.70	1723.82	0.12	0.094441123
Q8BTM8	QATPTSPIRVKVEPSH	1728.92	1729.92	1.00	0.062992126
Q8VDM4	EELRPLPVSVRVGQAV	1729.98	1729.98	0.00	0.005988024
Q8BYH8	DLLRRGAYGAIMEEE	1737.84	1736.76	-1.07	0.096028455
P35527	PAAIQKNYSPYYNTI	1741.87	1741.86	0.00	0.016
P04075	KGVVPLAGTNGETTTQGL	1741.92	1741.91	-0.01	0.00096
P68104	AAGAGKVTKSAQKAQKAK	1742.02	1742.01	0.00	0.0095
Q9Z1Q5	EEIELAYEQVARALK	1742.92	1742.90	-0.02	0
P11276	AIPANGQTPVQRSISPD	1749.90	1749.91	0.01	0.053097345
P06454	LKEKKEVVEEAENGR	1756.93	1757.94	1.01	0.000064
P62158	GQVNYEEFVQMMTAK	1773.81	1774.82	1.02	0.000031
P17182	LYKSFVQNYPVVSIE	1784.93	1784.95	0.01	0.005988024
P10853	TGISSKAMGIMNSFVND	1802.82	1802.82	0.01	0.008264462
P62806	AVTYTEHAKRKTVTAM	1805.95	1805.93	-0.01	0.005988024
P53396	LYFTYLEINPLVVTK	1812.01	1813.01	1.00	0.0052
P23284	TNGSQFFITTVKTAWL	1812.94	1812.94	0.00	0.05
P23284	VGRVIFGLFGKTVPKTV	1817.09	1817.09	0.00	0.032
Q8BTM8	SLTKVATVPQHATSGPGPA	1817.96	1817.96	0.00	0.027958993
O88569	SRGGGGNFGPGPGSNFRGGS	1820.83	1820.84	0.01	0.015968064
A2BH40	GAESNGGGGGGGAGSGGGPGAEPD	1827.73	1827.88	0.15	0.058238637
P60710	LYANTVLSGGTTMYPGIA	1827.91	1827.91	0.00	0.005988024
Q8BFY9	QFPLPLKERLAAFYGV	1831.00	1832.01	1.00	0
P40142	LAMFRSVPMSTVFYPS	1831.90	1832.90	1.00	0.005988024
Q15056	PPYTAYVGNLPFNTVQG	1836.90	1836.90	0.00	0.0031
P10809	KAQIEKRIQEIIEQL	1838.06	1838.06	0.00	0.037
P62737	LYANNVLSGGTTMYPGIA	1840.90	1839.91	-0.99	0.005988024
P56480	ATTVLSRAIAELGIYPAV	1844.04	1845.03	0.99	0.005988024
P30049	LGAAKANLEKAQAELVGTA	1854.02	1854.02	0.00	0.000031

O60361	GLNVVKTGRVMLGETNPA	1855.00	1856.00	1.00	0.005
P68363	VNAAIATIKTKRSIQFV	1859.10	1859.09	-0.01	0.012
P05387	YVASYLLAALGGNSSPSAK	1867.97	1867.97	0.00	0.00073
P00558	KYSLEPVAVELKSLLGK	1873.09	1873.08	-0.01	0.00059
Q99832	SVVAGGGAIEMELSKYLR	1878.99	1878.99	0.00	0.00019
P23528	TLAEKLGGSAVISLEGKPL	1882.08	1882.08	0.00	0.000034
P62806	VVYALKRQGRTLYGFGG	1884.04	1885.02	0.98	0.053691275
P32969	IELVSNSAALIQQATTVK	1885.05	1886.05	1.00	0.00017
P11672	STQNLIPAPSLLTVPLQP	1888.07	1889.07	1.00	0.005988024
P28665	IVTVNSTGLAEVEMTVPD	1889.93	1889.96	0.03	0.053691275
P35908	GLTAERTSQNSELNNMQ	1891.87	1892.87	1.00	0.015
Q8BSN3	DATGVAESHAVVRRFLAQ	1908.00	1906.24	-1.76	0.081063963
Q9NVP2	SVLVGPVPAGRHMFVFQA	1911.02	1911.02	0.00	0.021
P26638	EKYLIATSEQPIAALHR	1921.04	1921.04	0.00	0.010309278
P60709	TVPIYEGYALPHAILRL	1925.08	1925.07	-0.01	0.00022
P09651	PEQLRKLFIGGLSFETT	1935.05	1936.04	0.99	0.035
P62827	PNLEFVAMPALAPPEVVM	1939.98	1941.98	2.00	0.005988024
P60710	LYANTVLSGGTTMYPGIAD	1942.93	1942.94	0.00	0.008264462
Q9Z0U1	EAIYGPNTKMVRFKKGD	1951.00	1949.76	-1.23	0.083281539
P43308	VVIGSTSAPGQGGILAQREF	1986.05	1986.04	-0.01	0.0044
Q80TR8	AEIQKSALQIIINCVCGPD	1996.01	1996.07	0.06	0.053097345
P68104	KKAAGAGKVTKSAQKAQKAK	1998.21	1998.21	0.00	0.00091
Q99757	LAIEYEVSAVPTVLAMKNG	2004.06	2005.07	1.01	0.013
P07237	GAAAESLVESSEVAVIGFFK	2010.03	2012.04	2.00	0.0022
P04104	EINKRTNAENEFVTIKK	2015.08	2015.09	0.01	0.005988024
P60174	GRKQSLGELIGTLNAAKVPA	2022.16	2022.16	0.00	0.0025
P62937	FTRHNGTGGKSIYGEKFE	2026.99	2026.98	0.00	0.039
P08071	DLLFKESAIGFVRVPQKV	2045.17	1930.14	-115.03	3.41E-14
P31948	PTNMTYITNQAAVYFEKG	2046.97	2047.97	1.00	0.000021
C8YR32	DSNATLKNFRYHISVKTG	2050.06	2050.10	0.04	0.04989605
P04406	GIVEGLMTTVHAITATQKTV	2069.12	2069.12	0.00	0.00063
P67809	TKPGTTGSGAGSGGPGGLTSAAPAGG	2070.00	2069.99	0.00	0.00093

P09622	TKNILIATGSEVTPFPGITI	2071.16	2072.16	1.00	0.034
P07858	IMAEIYKNGPVEGAFSVYS	2074.01	2074.01	0.00	0.003
Q7TPD0	DLIRYICGVVHPSNEVLSS	2082.06	2080.57	-1.49	0.078255288
P35527	LEMQYETLQEELMALKK	2096.05	2097.05	0.99	0.0052
P10809	AAGVASLLTTAEVVVTEIPKE	2097.16	2098.15	0.99	0.026
P07437	TVVEPYNATLSVHQLVENT	2113.07	2115.06	1.99	0.042
P61088	ALQIRTVLLSIQALLSAPNP	2117.26	2118.26	1.01	0.00095
P32969	IELVSNSAALIQQATTVKNK	2127.19	2127.19	0.00	0.000013
P42765	LVEVNEAFAPQYLAVERSL	2147.13	2149.14	2.01	0.000002
P06733	GNSEVILPVPAFNVINGGSHAG	2148.10	2148.09	-0.01	0.019
P09104	GNSDLILPVPAFNVINGGSHAG	2148.10	2148.09	-0.01	0.014
P05387	MRYVASYLLAALGGNSSPSAK	2155.11	2156.11	1.00	0.0000053
P07737	AAVPGKTFVNITPAEVGVLVGK	2166.24	2166.24	0.00	0.00021
Q14974	VVMASLLRMFQSTAGSGGVQE	2167.08	2168.09	1.02	0.014
Q8VDP3	QLLGKANVVPEALQRFARAAA	2205.24	2206.23	0.99	0.052313883
Q3V0Q1	DPKVLAATSLTGLLEKLQNCN	2209.18	2210.38	1.20	0.033500839
P0CG49	MQIFVKTLTGKTITLEVEPS	2234.22	2250.22	16.00	1.67E-20
Q9NSE4	KVASVASTLETTFETISTLSGV	2240.18	2241.18	1.00	0.0011
Q61464	DILVLSSHKKAYIEINKKSA	2256.28	2254.54	-1.75	0.069645204
Q9Z2I8	DATQVEVNPFGETPEGQVVCF	2265.03	2264.20	-0.83	0.075471698
P28065	PPLVLAAANVVRNISYKYRE	2272.27	2273.27	1.00	0.041
P0CG49	DKEGIPPDQQRLIFAGKQLE	2281.21	2166.18	-115.03	3.66E-14
Q12905	LAPNSAEQASILSLVTKINNVI	2294.28	2294.27	-0.01	0.036
P60710	FEQEMATAASSSSLEKSYELP	2304.05	2304.05	0.00	0.005988024
P62737	FENEMATAASSSSLEKSYELP	2306.03	2305.05	-0.97	0.079155673
P60709	GVTHTVPIYEGYALPHAILRL	2319.27	2320.27	0.99	0.0013
P04406	NFGIVEGLMTTVHAITATQKTV	2330.23	2331.22	0.99	0.000024
P14866	SRSVNSVLLFTILNPIYSITT	2337.29	2337.28	-0.01	0.011
P63017	DKVSSKNSLESYAFNMKATVE	2347.14	2232.11	-115.03	6.72E-11

P10809	AAGVASLLTTAEVVVTEIPKEEK	2354.29	2355.29	0.99	0.000087
Q99NB5	SIPIQKGSVLAYKKQQLVIED	2356.34	2355.13	-1.21	0.051515151
P07437	RISVYYNEATGGKYVPRAILV	2368.29	2368.28	-0.01	0.018
P97504	DVWAFGILMWEVFSLGKQPY	2383.17	2382.92	-0.25	0.074742265
A7XUY5	RIHSILKSSSVQKETKNLLLD	2390.36	2390.30	-0.06	0.059154928
P60710	FEQEMATAASSSSLEKSYELPD	2419.07	2419.08	0.01	0.015968064
P60710	GVTHTVPIYEGYALPHAILRLD	2434.30	2435.31	1.00	0.066115702
P60709	DGVTHTVPIYEGYALPHAILRL	2434.30	2434.29	-0.01	0.021
P14649	VMRALGQNPTNAEVLKVLGNPKS	2435.33	2435.33	0.00	0.033
P52480	FLVTEVENGGSLGSKKGVNLPGAAV	2442.31	2443.31	1.00	0.005988024
P06744	PSAVAKHFVALSTNTTKVKEFGI	2444.34	2444.33	-0.01	0.0086
P16858	NFGIVEGLMTTVHAITATQKTVD	2445.26	2445.96	0.71	0.087009802
P60710	DIRKDLYANTVLSGGTTMYPGIA	2455.24	2322.21	-133.03	7.26E-10
Q9ES18	GGAVAAGAPGQESTEGAPPLYNTNHD	2462.11	2462.52	0.41	0.025341131
P97311	EFSPRAVYTSGKASSAAGLTAAVVR	2477.30	2478.32	1.02	0.005988024
P08107	AGVIAGLNVLRIINEPTAAAIAYGL	2479.42	2479.41	0.00	0.031
P35527	GGILTANEKSTMQELNSRLASYL	2495.27	2495.26	-0.01	0.00052
P38646	AGQISGLNVLRVINEPTAAALAYGL	2510.39	2510.38	-0.01	0.026
P07737	SVWAAVPGKTFVNITPAEVGVLVGK	2538.42	2538.42	-0.01	0.0012
P47857	MTHEEHHAAKTLGIGKAIAVLTSGG	2544.31	2544.44	0.13	0.040089087
Q60592	DFGLSKIGLMSLTTNLYEGHIEK	2581.31	2580.11	-1.20	0.047413793
P84244	LRFQSAAIGALQEASEAYLVGLFE	2582.34	2582.33	-0.01	0.005988024
O76021	SASASLSSAAATGTSTSTPAAPTARKQL	2590.32	2591.32	1.00	0.000017
P62249	PRTLQYKLLEPVLLLGKERFAGV	2639.55	2639.56	0.01	0.0098
P14866	NQIYIAGHPAFVNYSTSQKISRPG	2647.35	2647.34	-0.01	0.019
P62737	EAQSKRGILTLKYPIEHGIITNW	2648.44	2648.44	-0.01	0.005988024
P84228	LRFQSSAVMALQEASEAYLVGLFE	2658.34	2658.34	0.01	0.005988024
B1AUH1	MVNASQHAPGQRAHIIFQTLSEND	2663.29	2662.31	-0.98	0.053691275
P08708	FGSLSNLQVTQPTVGMNFKTPRGPV	2674.39	2674.39	0.00	0.0014
Q99714	PAEYAHLVQAIIENPFLNGEVIRL	2705.45	2706.47	1.02	0.00033

Q9D2G5	FGVFAKGENVSPRFQKGTLRMNCL	2714.38	2714.96	0.59	0.079233229
P68104	GQTREHALLAYTLGVKQLIVGVNKM	2738.53	2738.52	-0.01	0.0000051
P04264	KVRFLEQQNQVLQTKWELLQQV	2754.52	2754.52	0.00	0.013
P13667	YMIEQSGPPSKEILTLKQVQEFLK	2805.50	2805.49	-0.01	0.036
Q00612	DELMKRVGFQYEGTYKWVNPHKL	2837.43	2704.39	-133.04	3.97E-10
P06733	LFTSKGLFRAAVPSGASTGIYEALELR	2853.54	2853.53	-0.01	0.033
P04908	VGAGAPVYLAAVLEYLTAEILELAGNAAR	2914.58	2915.59	1.01	0.00000011
P04264	FLEQQNQVLQTKWELLQQVDTSTR	2931.51	2932.50	0.99	0.01
Q16777	VGAGAPVYMAAVLEYLTAEILELAGNAAR	2932.54	2932.53	-0.01	0.0018
P17182	EGGFAPNILENKEALELLKTAIAKAGYT	2942.57	2943.59	1.01	0.005988024
P60710	DGVTHTVPIYEGYALPHAILRLDLAGR	2946.57	2831.55	-115.02	5.37E-34
P62266	GVRFKVVKVANVSLLALYKGKKERPRS	3041.83	3041.83	-0.01	0.0029
Q61171	SQFTHLAWINTPRKEGGLGPLNIPLLAD	3057.64	3056.95	-0.69	0.005988024
P04264	IAQKSKAEAESLYQSKYEELQITAGRHG	3134.60	3135.60	1.00	0.0042
P31949	GYNYTLSKTEFLSFMNTELAAFTKNQK	3145.54	3146.55	1.01	0.012
P24752	VMVAGGMESMSNVPYVMNRGSTPYGGVKLE	3159.49	3161.48	1.99	0.0026
Q8CGP0	DGKKRKRGRKESYSIYVYKVLKQVHP	3161.79	3062.76	-99.03	2.68E-15
P0C0S6	DSLIKATIAGGGVIPHIHKSLIGKKGQQKTV	3193.87	3078.85	-115.02	1.92E-12
P14618	TKGPEIRTGLIKGSGTAEVELKKGATLKITL	3208.88	3208.86	-0.02	0.017
Q920Q2	FIRGQLVTNLPGVGRSMESKLASLGIKTCGD	3246.72	3248.59	1.86	0.082862526
P60710	DGVTHTVPIYEGYALPHAILRLDLAGRDLT	3275.73	3160.71	-115.02	1.4E-10
Q9D071	EVLAALASVIGTATTHLSPELAAQSVTCIVPLFL	3434.87	3435.31	0.43	0.074742265
P68433	DTNLCAIHAKRVTIMPKDIQLARRIRGERA	3444.90	3329.88	-115.02	4.81E-16
P0CG49	MQIFVKTLTGKTITLEVEPSDTIENVKAKIQ	3473.91	3473.92	0.01	4.35E-30
Q50L43	DEVPVVGINAEGGGMRAMISLYGHLLALQKLGLL	3547.89	3548.92	1.03	0.096028455

P05387	VIAQGIGKLASVPAGGAVAVSAAPGSAAPAAGSAPAAAEEKK	3667.99	3668.97	0.98	0.0029
P06745	MAALTRNPQFQKLLEWHRANSANLKLRELFEA	3795.02	3706.00	-89.02	1.85E-14
P62806	DAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG	3902.03	3787.00	-115.03	4.13E-11
Q99497	GLILTSRGPGTSFEFALAIVEALNGKEVAAQVKAPLVLK	4007.28	4009.30	2.02	0.0041
Q6GSS7	DEELNKLLGKVTIAQGGVLPNIQAVLLPKKTESHHKAKGK	4301.46	4168.43	-133.03	6.66E-27
P22752	DEELNKLLGRVTIAQGGVLPNIQAVLLPKKTESHHKAKGK	4329.46	4196.43	-133.03	2.16E-14
Q01768	RPFFPGLVKYMNSGPVVAMVWEGLNVVKTGRVMLGETNPAD	4475.30	4476.15	0.85	0.052313883
P16858	DPTNIKWGEAGAEYVVESTGVFTTMEKAGAHLKGGAKRVIISAPSA	4743.43	4628.41	-115.02	5.38E-19
P10854	DIFERIAGEASRLAHYNKRSTITSREIQTAVRLLLPGELAKHAVSEGTKAVTKYTSSK	6397.46	6282.45	-115.01	2.93E-10
Q8CGP0	DIFERIASEASRLAHYNKRSTITSREVQTAVRLLLPGELAKHAVSEGTKAVTKYTSSK	6413.46	6312.47	-100.99	1.49E-11

Bibliography

1. Wilkins, M., Proteomics data mining. *Expert Review of Proteomics* **2009,** *6* (6), 599-603.

2. Geiger, T.; Wehner, A.; Schaab, C.; Cox, J.; Mann, M., Comparative Proteomic Analysis of Eleven Common Cell Lines Reveals Ubiquitous but Varying Expression of Most Proteins. *Molecular & Cellular Proteomics* **2012,** *11* (3), M111.014050–1-M111.014050–11.

3. Whitehouse, C. M.; Dreyer, R. N.; Yamashita, M.; Fenn, J. B., Electrospray interface for liquid chromatographs and mass spectrometers. *Analytical Chemistry* **1985,** *57* (3), 675-679.

4. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T., Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **1988,** *2* (8), 151-153.

5. The Nobel Prize in Chemistry 2002 at www.Nobelprize.org.

6. Dobo, A.; Kaltashov, I. A., Detection of Multiple Protein Conformational Ensembles in Solution via Deconvolution of Charge-State Distributions in ESI MS. *Analytical Chemistry* **2001,** *73* (20), 4763-4773.

7. Paul, W.; Steinwedel, H., Ein neues Massenspektrometer ohne Magnetfeld. *Zeitschrift Naturforschung A* **1953,** *8* (7), 448-450.

8. The Nobel Prize in Physics 1989 at www.Nobelprize.org.

9. Stafford Jr, G. C.; Kelley, P. E.; Syka, J. E. P.; Reynolds, W. E.; Todd, J. F. J., Recent improvements in and analytical applications of advanced ion trap technology. *International Journal of Mass Spectrometry and Ion Processes* **1984,** *60* (1), 85-98.

10. McLuckey, S.; Mentinova, M., Ion/Neutral, Ion/Electron, Ion/Photon, and Ion/Ion Interactions in Tandem Mass Spectrometry: Do We Need Them All? Are They Enough? *Journal of the American Society for Mass Spectrometry* **2011,** *22* (1), 3-12.

11. Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R., Protein sequencing by tandem mass spectrometry. *Proceedings of the National Academy of Sciences* **1986,** *83* (17), 6233-6237.

12. Makarov, A., Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical Chemistry* **2000,** *72* (6), 1156-1162.

13. Perry, R. H.; Cooks, R. G.; Noll, R. J., Orbitrap mass spectrometry: Instrumentation, ion motion and applications. *Mass Spectrometry Reviews* **2008,** *27* (6), 661-699.

14. Gucinski, A. C.; Dodds, E. D.; Li, W.; Wysocki, V. H.; Hubbard, S. J.; Jones, A. R., Understanding and Exploiting Peptide Fragment Ion Intensities Using Experimental and Informatic Approaches
Proteome Bioinformatics. Walker, J. M., Ed. Humana Press: 2010; Vol. 604, 73-94.

15. Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A., Mobile and localized protons: a framework for understanding peptide dissociation. *Journal of Mass Spectrometry* **2000,** *35* (12), 1399-1406.

16. Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W., Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *Journal of the American Chemical Society* **1999,** *121* (4), 806-812.

17. Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J.-H.; Begley, T. P.; McLafferty, F. W., Top Down Characterization of Larger Proteins (45 kDa) by Electron Capture Dissociation Mass Spectrometry. *Journal of the American Chemical Society* **2002,** *124* (4), 672-678.

18. Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L., On the Scalability and Requirements of Whole Protein Mass Spectrometry. *Analytical Chemistry* **2011,** *83* (17), 6868-6874.

19. Wynne, C.; Fenselau, C.; Demirev, P. A.; Edwards, N., Top-Down Identification of Protein Biomarkers in Bacteria with Unsequenced Genomes. *Analytical Chemistry* **2009,** *81* (23), 9633-9642.

20. Tran, J. C.; Zamdborg, L.; Ahlf, D. R.; Lee, J. E.; Catherman, A. D.; Durbin, K. R.; Tipton, J. D.; Vellaichamy, A.; Kellie, J. F.; Li, M.; Wu, C.; Sweet, S. M. M.; Early, B. P.; Siuti, N.; LeDuc, R. D.; Compton, P. D.; Thomas, P. M.; Kelleher, N. L., Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* **2011,** *480* (7376), 254-258.

21. Wu, C.; Tran, J. C.; Zamdborg, L.; Durbin, K. R.; Li, M.; Ahlf, D. R.; Early, B. P.; Thomas, P. M.; Sweedler, J. V.; Kelleher, N. L., A protease for 'middle-down' proteomics. *Nature Methods* **2012,** *9* (8), 822-824.

22. Cannon, J.; Lohnes, K.; Wynne, C.; Wang, Y.; Edwards, N.; Fenselau, C., High-throughput middle-down analysis using an orbitrap. *Journal of Proteome Research* **2010,** *9* (8), 3886-3890.

23. Yates, J. R.; Eng, J. K.; McCormack, A. L., Mining Genomes: Correlating Tandem Mass Spectra of Modified and Unmodified Peptides to Sequences in Nucleotide Databases. *Analytical Chemistry* **1995,** *67* (18), 3202-3210.

24. Yates, J. R.; Eng, J. K.; McCormack, A. L.; Schieltz, D., Method to Correlate Tandem Mass Spectra of Modified Peptides to Amino Acid Sequences in the Protein Database. *Analytical Chemistry* **1995,** *67* (8), 1426-1436.

25. Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999,** *20* (18), 3551-3567.

26. Senko, M.; Beu, S.; McLafferty, F., Automated assignment of charge states from resolved isotopic peaks for multiply charged ions. *Journal of the American Society for Mass Spectrometry* **1995,** *6* (1), 52-56.

27. Wynne, C.; Edwards, N. J.; Fenselau, C., Phyloproteomic classification of unsequenced organisms by top-down identification of bacterial proteins using capLC-MS/MS on an Orbitrap. *Proteomics* **2010,** *10* (20), 3631-3643.

28. Lee, J.-E.; Kwon, J.; Baek, M.-C., A combination method of chemical with enzyme reactions for identification of membrane proteins. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* **2011,** *1814* (3), 397-404.

29. Drapeau, G. R., Substrate specificity of a proteolytic enzyme isolated from a mutant of Pseudomonas fragi. *Journal of Biological Chemistry* **1980,** *255* (3), 839-840.

30. Alam, A.; Mataj, A.; Yang, Y.; Boysen, R. I.; Bowden, D. K.; Hearn, M. T., Rapid Microwave-Assisted Chemical Cleavage - Mass Spectrometric Method for the Identification of Hemoglobin Variants in Blood. *Analytical Chemistry* **2010**, 82 (21), 8922-8930.

31. Wu, S. L.; Kim, J.; Bandle, R. W.; Liotta, L.; Petricoin, E.; Karger, B. L., Dynamic profiling of the post-translational modifications and interaction partners of epidermal growth factor receptor signaling after stimulation by epidermal growth factor using Extended Range Proteomic Analysis (ERPA). *Molecular & Cellular Proteomics* **2006,** *5* (9), 1610-1627.

32. Wu, S. L.; Kim, J.; Hancock, W. S.; Karger, B., Extended Range Proteomic Analysis (ERPA): a new and sensitive LC-MS platform for high sequence coverage of complex proteins with extensive post-translational modifications-comprehensive analysis of beta-casein and epidermal growth factor receptor (EGFR). *Journal of Proteome Research* **2005,** *4* (4), 1155-1170.

33. Zhang, J.; Wu, S. L.; Kim, J.; Karger, B. L., Ultratrace liquid chromatography/mass spectrometry analysis of large peptides with post-translational modifications using narrow-bore poly(styrene-divinylbenzene) monolithic columns and extended range proteomic analysis. *Journal of Chromatography A* **2007,** *1154* (1-2), 295-307.

34. Kocher, T.; Swart, R.; Mechtler, K., Ultra-High-Pressure RPLC Hyphenated to an LTQ-Orbitrap Velos Reveals a Linear Relation between Peak Capacity and Number of Identified Peptides. *Analytical Chemistry* **2011,** *83* (7), 2699-2704.

35. Swatkoski, S.; Gutierrez, P.; Wynne, C.; Petrov, A.; Dinman, J. D.; Edwards, N.; Fenselau, C., Evaluation of microwave-accelerated residue-specific acid cleavage for proteomic applications. *Journal of Proteome Research* **2008,** *7* (2), 579-586.

36. Schultz, J.; Allison, H.; Grice, M., Specificity of the Cleavage of Proteins by Dilute Acid. I. Release of Aspartic Acid from Insulin, Ribonuclease, and Glucagon*. *Biochemistry* **1962,** *1* (4), 694-698.

37. Inglis, A. S.; C.H.W. Hirs, S. N. T., Cleavage at aspartic acid. In *Methods in Enzymology*, Academic Press: 1983; Vol. Volume 91, pp 324-332.

38. Li, A.; Sowder, R. C.; Henderson, L. E.; Moore, S. P.; Garfinkel, D. J.; Fisher, R. J., Chemical cleavage at aspartyl residues for protein identification. *Analytical Chemistry* **2001,** *73* (22), 5395-5402.

39. Swatkoski, S.; Russell, S. C.; Edwards, N.; Fenselau, C., Rapid chemical digestion of small acid-soluble spore proteins for analysis of Bacillus spores. *Analytical Chemistry* **2006,** *78* (1), 181-188.

40. Hua, L.; Low, T. Y.; Sze, S. K., Microwave-assisted specific chemical digestion for rapid protein identification. *Proteomics* **2006,** *6* (2), 586-591.

41. Bruice, T. C.; Pandit, U. K., The Effect of Geminal Substitution Ring Size and Rotamer Distribution on the Intramolecular Nucleophilic Catalysis of the Hydrolysis of Monophenyl Esters of Dibasic Acids and the Solvolysis of the Intermediate Anhydrides. *Journal of the American Chemical Society* **1960,** *82* (22), 5858-5865.

42. Xie, M.; Vander Velde, D.; Morton, M.; Borchardt, R. T.; Schowen, R. L., pH-Induced Change in the Rate-Determining Step for the Hydrolysis of the Asp/Asn-Derived Cyclic-Imide Intermediate in Protein Degradation. *Journal of the American Chemical Society* **1996,** *118* (37), 8955-8956.

43. Swatkoski, S.; Gutierrez, P.; Ginter, J.; Petrov, A.; Dinman, J. D.; Edwards, N.; Fenselau, C., Integration of residue-specific acid cleavage into proteomic workflows. *Journal of Proteome Research* **2007,** *6* (11), 4525-4527.

44. Hauser, N. J.; Basile, F., Online microwave D-cleavage LC-ESI-MS/MS of intact proteins: site-specific cleavages at aspartic acid residues and disulfide bonds. *Journal of Proteome Research* **2008,** *7* (3), 1012-1026.

45. Tsugita, A.; Miyazaki, K.; Nabetani, T.; Nozawa, T.; Kamo, M.; Kawakami, T., Application of chemical selective cleavage methods to analyze post-translational modification in proteins. *Proteomics* **2001,** *1* (9), 1082-1091.

46. Basile, F.; Hauser, N., Rapid online nonenzymatic protein digestion combining microwave heating acid hydrolysis and electrochemical oxidation. *Analytical Chemistry* **2011,** *83* (1), 359-367.

47. Fenselau, C.; Laine, O.; Swatkoski, S., Microwave assisted acid cleavage for denaturation and proteolysis of intact human adenovirus. *International Journal of Mass Spectrometry* **2011,** *301* (1-3), 7-11.

48. Hauser, N. J.; Han, H.; McLuckey, S. A.; Basile, F., Electron transfer dissociation of peptides generated by microwave D-cleavage digestion of proteins. *Journal of Proteome Research* **2008,** *7* (5), 1867-1872.

49. Li, J.; Shefcheck, K.; Callahan, J.; Fenselau, C., Extension of microwave-accelerated residue-specific acid cleavage to proteins with carbohydrate side chains and disulfide linkages. *International Journal of mass Spectrometry* **2008,** *278* (2-3), 109-113.

50. Osula, O.; Swatkoski, S.; Cotter, R. J., Identification of protein SUMOylation sites by mass spectrometry using combined microwave-assisted aspartic acid cleavage and tryptic digestion. *Journal of Mass Spectrometry* **2012,** *47* (5), 644-654.

51. Swatkoski, S.; Russell, S.; Edwards, N.; Fenselau, C., Analysis of a model virus using residue-specific chemical cleavage and MALDI-TOF mass spectrometry. *Analytical Chemistry* **2007,** *79* (2), 654-658.

52. Remily-Wood, E.; Dirscherl, H.; Koomen, J., Acid hydrolysis of proteins in matrix assisted laser desorption ionization matrices. *Journal of the American Society for Mass Spectrometry* **2009,** *20* (11), 2106-2115.

53. Yassine, M. M.; Guo, N.; Zhong, H.; Li, L.; Lucy, C. A., Off-line coupling of preparative capillary zone electrophoresis with microwave-assisted acid hydrolysis and matrix-assisted laser desorption ionization mass spectrometry for protein sequencing. *Analytica Chimica Acta* **2007,** *597* (1), 41-49.

54. Zhong, H.; Marcus, S.; Li, L., Microwave-assisted acid hydrolysis of proteins combined with liquid chromatography MALDI MS/MS for protein identification. *Journal of the American Society for Mass Spectrometry* **2005,** *16* (4), 471-481.

55. Lidstrom, P.; Tierney, J.; Wathey, B.; Westman, J., Microwave assisted organic synthesis, a review. *Tetrahedron* **2001,** *57* (45), 9225-9283.

56. Choksawangkarn, W.; Edwards, N.; Wang, Y.; Gutierrez, P.; Fenselau, C., Comparative Study of Workflows Optimized for In-gel, In-solution, and On-filter Proteolysis in the Analysis of Plasma Membrane Proteins. *Journal of Proteome Research* **2012,** *11* (5), 3030-3034.

57. Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M., Universal sample preparation method for proteome analysis. *Nat Meth* **2009,** *6* (5), 359-362.

58. Wessel, D.; Flugge, U. I., A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Analytical Biochemistry* **1984,** *138* (1), 141-143.

59. Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M., In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protocols* **2007,** *1* (6), 2856-2860.

60. Edwards, N.; Wu, X.; Tseng, C.-W., An Unsupervised, Model-Free, Machine-Learning Combiner for Peptide Identifications from Tandem Mass Spectra. *Clinical Proteomics* **2009,** *5* (1), 23-36.

61. Fenselau, C.; Havey, C.; Teerakulkittipong, N.; Swatkoski, S.; Laine, O.; Edwards, N., Identification of beta-lactamase in antibiotic-resistant Bacillus cereus spores. *Applied and Environmental Microbiology* **2008,** *74* (3), 904-906.

62. Mirza, U. A.; Liu, Y.-H.; Tang, J. T.; Porter, F.; Bondoc, L.; Chen, G.; Pramanik, B. N.; Nagabhushan, T. L., Extraction and characterization of adenovirus proteins from sodium dodecylsulfate polyacrylamide gel electrophoresis by matrix-assisted laser desorption/ionization mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2000,** *11* (4), 356-361.

63. Cannon, J.; Edwards, N.; Fenselau, C., Middle-out proteomic analysis using site-selective acid cleavage and mass biased partitioning. *Submitted* **2012**.

64. Meyer, B.; Papasotiriou, D.; Karas, M., 100% protein sequence coverage: a modern form of surrealism in proteomics. *Amino Acids* **2011,** *41* (2), 291-310.

65. Boyne, M. T.; Garcia, B. A.; Li, M.; Zamdborg, L.; Wenger, C. D.; Babai, S.; Kelleher, N. L., Tandem Mass Spectrometry with Ultrahigh Mass Accuracy Clarifies Peptide Identification by Database Retrieval. *Journal of Proteome Research* **2008,** *8* (1), 374-379.

66. 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. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (26), 9528-9533.

67. Taouatas, N.; Heck, A. J. R.; Mohammed, S., Evaluation of Metalloendopeptidase Lys-N Protease Performance under Different Sample Handling Conditions. *Journal of Proteome Research* **2010,** *9* (8), 4282-4288.

68. Rahbar, A. M.; Fenselau, C., Integration of Jacobson's Pellicle Method into Proteomic Strategies for Plasma Membrane Proteins. *Journal of Proteome Research* **2004,** *3* (6), 1267-1277.

69. Scherl, A.; Shaffer, S. A.; Taylor, G. K.; Kulasekara, H. D.; Miller, S. I.; Goodlett, D. R., Genome-specific gas-phase fractionation strategy for improved shotgun proteomic profiling of proteotypic peptides. *Analytical Chemistry* **2008,** *80* (4), 1182-1191.

70. Yi, E. C.; Marelli, M.; Lee, H.; Purvine, S. O.; Aebersold, R.; Aitchison, J. D.; Goodlett, D. R., Approaching complete peroxisome characterization by gas-phase fractionation. *Electrophoresis* **2002,** *23* (18), 3205-3216.

71. Lee, J.; Kellie, J.; Tran, J.; Tipton, J.; Catherman, A.; Thomas, H.; Ahlf, D.; Durbin, K.; Vellaichamy, A.; Ntai, I.; Marshall, A.; Kelleher, N., A robust two-dimensional separation for top-down tandem mass spectrometry of the low-mass proteome. *Journal of the American Society for Mass Spectrometry* **2009,** *20* (12), 2183-2191.

72. Han, X.; Jin, M.; Breuker, K.; McLafferty, F. W., Extending Top-Down Mass Spectrometry to Proteins with Masses Greater Than 200 Kilodaltons. *Science* **2006,** *314* (5796), 109-112.

73. Bondarenko, P.; Second, T.; Zabrouskov, V.; Makarov, A.; Zhang, Z., Mass measurement and top-down HPLC/MS analysis of intact monoclonal antibodies on a hybrid linear quadrupole ion trap-orbitrap mass spectrometer. *Journal of the American Society for Mass Spectrometry* **2009,** *20* (8), 1415-1424.

74. Garcia, B. A.; Pesavento, J. J.; Mizzen, C. A.; Kelleher, N. L., Pervasive combinatorial modification of histone H3 in human cells. *Nat Meth* **2007,** *4* (6), 487-489.

75. Taverna, S. D.; Ueberheide, B. M.; Liu, Y.; Tackett, A. J.; Diaz, R. L.; Shabanowitz, J.; Chait, B. T.; Hunt, D. F.; Allis, C. D., Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proceedings of the National Academy of Sciences* **2007,** *104* (7), 2086-2091.

76. Chi, A.; Bai, D. L.; Geer, L. Y.; Shabanowitz, J.; Hunt, D. F., Analysis of intact proteins on a chromatographic time scale by electron transfer dissociation tandem mass spectrometry. *International Journal of Mass Spectrometry* **2007,** *259* (1), 197-203.

77. Hohmann, L.; Sherwood, C.; Eastham, A.; Peterson, A.; Eng, J. K.; Eddes, J. S.; Shteynberg, D.; Martin, D. B., Proteomic Analyses Using Grifola frondosa Metalloendoprotease Lys-N. *Journal of Proteome Research* **2009,** *8* (3), 1415-1422.

78. Vestling, M. M.; Fenselau, C., Poly(vinylidene difluoride) membranes as the interface between laser desorption mass spectrometry, gel electrophoresis, and in situ proteolysis. *Analytical Chemistry* **1994,** *66* (4), 471-477.

79. The Nobel Prize in Chemistry 2009 at www.Nobelprize.org.

80. Dinman, J. D., University of Maryland private communication. **2008**.

81. Hardy, S. J. S.; Kurland, C. G.; Voynow, P.; Mora, G., The ribosomal proteins of Escherichia coli. I. Purification of the 30s ribosomal proteins. *Biochemistry* **1969,** *8* (7), 2897-2905.

82. Nakao, A.; Yoshihama, M.; Kenmochi, N., RPG: the Ribosomal Protein Gene database. *Nucleic Acids Research* **2004,** *32* (suppl 1), D168-D170.

83. Kenmochi, N.; Kawaguchi, T.; Rozen, S.; Davis, E.; Goodman, N.; Hudson, T. J.; Tanaka, T.; Page, D. C., A Map of 75 Human Ribosomal Protein Genes. *Genome Research* **1998,** *8* (5), 509-523.

84. Li, W.; Ye, Y., Polyubiquitin chains: functions, structures, and mechanisms. *Cell Mol Life Sci* **2008,** *65* (15), 2397-2406.

85. Xu, G.; Paige, J. S.; Jaffrey, S. R., Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat Biotech* **2010,** *28* (8), 868-873.

86. Mirzaei, H.; Rogers, R. S.; Grimes, B.; Eng, J.; Aderem, A.; Aebersold, R., Characterizing the connectivity of poly-ubiquitin chains by selected reaction monitoring mass spectrometry. *Molecular BioSystems* **2010,** *6* (10), 2004-2014.

87. Kim, H. T.; Kim, K. P.; Lledias, F.; Kisselev, A. F.; Scaglione, K. M.; Skowyra, D.; Gygi, S. P.; Goldberg, A. L., Certain pairs of ubiquitin-conjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages. *J Biol Chem* **2007,** *282* (24), 17375-17386.

88. Jung, J.; Pierson, N.; Marquardt, A.; Scheffner, M.; Przybylski, M.; Clemmer, D., Differentiation of Compact and Extended Conformations of Di-Ubiquitin Conjugates with Lysine-Specific Isopeptide Linkages by Ion Mobility-Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2011,** *22* (8), 1463-1471.

89. Strachan, J.; Roach, L.; Sokratous, K.; Tooth, D.; Long, J.; Garner, T. P.; Searle, M. S.; Oldham, N. J.; Layfield, R., Insights into the Molecular Composition of Endogenous Unanchored Polyubiquitin Chains. *Journal of Proteome Research* **2012,** *11* (3), 1969-1980.

90. Pickart, C. M.; Raasi, S.; Raymond, J. D., Controlled Synthesis of Polyubiquitin Chains. In *Methods in Enzymology*, Academic Press: 2005; Vol. Volume 399, pp 21-36.

91. Dong, K. C.; Helgason, E.; Yu, C.; Phu, L.; Arnott, D. P.; Bosanac, I.; Compaan, D. M.; Huang, O. W.; Fedorova, A. V.; Kirkpatrick, D. S.; Hymowitz, S. G.; Dueber,

E. C., Preparation of Distinct Ubiquitin Chain Reagents of High Purity and Yield. *Structure* **2011,** *19* (8), 1053-1063.

92. Johnston, S. C.; Riddle, S. M.; Cohen, R. E.; Hill, C. P., Structural basis for the specificity of ubiquitin C-terminal hydrolases. *EMBO J* **1999,** *18* (14), 3877-3887.

93. Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A., Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase. *Journal of the American Chemical Society* **1998,** *120* (27), 6836-6837.

94. Saville, M. K.; Sparks, A.; Xirodimas, D. P.; Wardrop, J.; Stevenson, L. F.; Bourdon, J. C.; Woods, Y. L.; Lane, D. P., Regulation of p53 by the ubiquitin-conjugating enzymes UbcH5B/C in vivo. *J Biol Chem* **2004,** *279* (40), 42169-42181.

95. Li, W.; Ye, Y., Polyubiquitin chains: functions, structures, and mechanisms. *Cellular and Molecular Life Sciences* **2008,** *65* (15), 2397-2406.

96. Mathivanan, S.; Ji, H.; Simpson, R. J., Exosomes: Extracellular organelles important in intercellular communication. *Journal of Proteomics* **2010,** *73* (10), 1907-1920.

97. Raimondo, F.; Morosi, L.; Chinello, C.; Magni, F.; Pitto, M., Advances in membranous vesicle and exosome proteomics improving biological understanding and biomarker discovery. *Proteomics* **2011,** *11* (4), 709-720.

98. Dennis, G.; Sherman, B.; Hosack, D.; Yang, J.; Gao, W.; Lane, H.; Lempicki, R., DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology* **2003,** *4* (9), R60.

99. Nicklay, J. J.; Shechter, D.; Chitta, R. K.; Garcia, B. A.; Shabanowitz, J.; Allis, C. D.; Hunt, D. F., Analysis of Histones in Xenopus laevis. *Journal of Biological Chemistry* **2009,** *284* (2), 1075-1085.

100. Grazini, U.; Zanardi, F.; Citterio, E.; Casola, S.; Goding, C. R.; McBlane, F., The RING Domain of RAG1 Ubiquitylates Histone H3: A Novel Activity in Chromatin-Mediated Regulation of V(D)J Joining. *Molecular cell* **2010,** *37* (2), 282-293.

101. Schatz, D. G.; Swanson, P. C., V(D)J Recombination: Mechanisms of Initiation. *Annual Review of Genetics* **2011,** *45* (1), 167-202.

102. Mombaerts, P.; Iacomini, J.; Johnson, R. S.; Herrup, K.; Tonegawa, S.; Papaioannou, V. E., RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* **1992,** *68* (5), 869-877.

103. Huang, J.; Chen, J., VprBP targets Merlin to the Roc1-Cul4A-DDB1 E3 ligase complex for degradation. *Oncogene* **2008,** *27* (29), 4056-4064.