GLOBAL QUANTITATIVE PROTEOMIC PROFILING THROUGH ^{18}O -LABELING IN
COMBINATION WITH MS/MS SPECTRA ANALYSIS

by

Carl Allan White

A thesis submitted in conformity with the requirements
for the degree of Master of Science
Graduate Department of Molecular Genetics
University of Toronto

© Copyright by Carl Allan White 2010

Global quantitative proteomic profiling through ^{18}O -labeling in combination with MS/MS spectra analysis

Master of Science, 2010

Department of Molecular Genetics, University of Toronto

Carl Allan White

Abstract

By integrating the simplicity of ^{18}O -labeling and the low signal-to-noise of MS/MS spectra with supporting software and combining them with global shotgun protein identification, a robust quantitative pipeline has been created that avoids the disadvantages of other quantitative approaches. Test mixtures of labeled and unlabeled peptides were subjected to LC-MS/MS profiling experiments. Software programs were developed and applied to automatically determine protein ratios between two samples while applying a correction for incomplete labeling. The measurement of relative abundance at the product ion (MS/MS) level, instead of at the full scan (MS) level, is shown to provide excellent accuracy and sensitivity. Ratio distributions approached the expected means, allowing empirical derivation of confidence level cutoffs for determining statistically significant fold-changes in protein abundance. A set of stringent criteria for detecting spurious ratios based on consistency checking between unlabeled and labeled y-ion pairs was found to highlight putative false positive identifications.

Acknowledgement

I would like to thank my supervisor, Professor Andrew Emili, for his patience and guidance during my time as a graduate student. I am also very grateful to the other members of my Supervisory Committee, Professors Jack Greenblatt and Brent Derry, for their apt advice.

I would also like to thank Vincent Fong, Johannes Hewel and other members of the Emili laboratory, past and present, for their advice and critical suggestions, and particularly Nico Oey who kindly assisted me by maintaining and harvesting cultures of *C. elegans* during his undergraduate years. Additionally, my thanks go to Celia Quevedo of the Derry laboratory for providing initial *C. elegans* extracts and to Scott Dixon of the Roy laboratory for his time spent in training Nico.

I am grateful too for the support I received from the National Sciences and Engineering Research Council of Canada (NSERC), in the form of a Canada Graduate Scholarship (Masters).

Finally, I sincerely thank the anonymous reviewers of my associated paper (White, 2009) for their feedback and suggestions.

.

Table of Contents

Abstract.....	ii
Acknowledgement	iii
Table of Contents.....	iv
List of Abbreviations	vii
List of Tables	viii
List of Figures.....	ix
List of Appendices.....	x
1. Introduction	1
1.1 Context.....	1
1.2 High-throughput identification of proteins.....	2
1.3 Quantification at the full scan level versus at the MS/MS level	4
1.4 Present high-throughput quantitative methods	10
1.4.1 Metabolic labeling	10
1.4.2 Chemical labeling	12
1.4.3 Synthetic reference peptides	16
1.4.4 Label-free quantification	17
1.5 ¹⁸ O-based stable isotope labeling	19
1.6 Thesis rationale.....	25
2. Methods and materials.....	27
2.1 Acquisition of the test proteins.....	27
2.2 Sample preparation for mass spectrometry	28
2.3 LC-MS/MS analysis	30
2.4 Identification of proteins	31
2.5 Calculation of relative quantities.....	33
2.6 Obtaining the software.....	33
3. Results	35

3.1 Outline of the workflow	36
3.2 Calculation of peptide ratios	40
3.3 Verification of feasibility.....	43
3.3.1 Single protein assays	43
3.3.2 Spiked protein assays.....	46
3.3.3 Complex mixture assays	46
3.3.4 Assessing bias by precursor ion charge	49
3.3.5 Accounting for residual variation	49
3.4 Assigning statistical confidence levels to calculated ratios.....	52
3.5 Detection of questionable identifications	55
3.6 Other considerations	57
3.6.1 Deviation of the isotopic distribution pattern from the theoretical	57
3.6.2 Ratios calculated additively versus ratios calculated by averaging pairs.....	58
3.6.3 Assessing intra-spectra bias by mass over charge	58
3.6.4 Assessing inter-spectra bias by y-ion series intensities	60
3.6.5 Evaluation of back-exchange.....	62
3.6.6 Outlier elimination.....	64
4. Discussion.....	66
4.1 General methodology	66
4.2 Correction for partial labeling	67
4.3 Dynamic range.....	68
4.4 Alarm flagging.....	70
4.5 Software options	71
4.6 Conclusion	72
5. Future Work.....	74
5.1 Experimental.....	74
5.2 Software Enhancements.....	75

6. References	78
---------------------	----

List of Abbreviations

AQUA, absolute quantification

BSA, bovine serum albumin

CID, collision-induced dissociation

DDA, data-dependent acquisition

ETD, electron transfer dissociation

FWHM, full width at half maximum

HPLC, high performance liquid chromatography

ICAT, isotope-coded affinity tag

ICPL, isotope-coded protein label

IQR, interquartile range

ISIS, isobaric SILAC with immonium ion splitting

iTRAQ, isobaric tags for relative and absolute quantification

LC, liquid chromatography

MS, mass spectrometry

MS/MS, tandem mass spectrometry incorporating fragmentation of ions

MudPIT, multi-dimensional protein identification technology

m/z, mass (Daltons) over unit of charge (+1, +2 or +3)

SILAC, stable isotope labels by amino acids in cell cultures

SCX, strong cation exchange

TCA, trichloroacetic acid

XIC, extracted ion chromatogram

List of Tables

Table 1: Establishing significance in terms of calculated protein abundance.	54
Table 2: Ratios calculated additively versus ratios calculated by averaging pairs.	59
Table 3: Assessing in-spectra bias by mass over charge.	61
Table 4: Assessing bias in y-ion series intensities.	63
Table 5: Ratio averages over a chromatography time course.	65
Table S1: Effect of isolation width on peptide and protein identifications.	97
Table S2: Comparison of ratios distributions for singly- and multiply-charged peptides. ...	98
Table S3: Peptide ratios do not vary independently between experiments.	99
Table S4: Empirical inverses of cumulative distributions of ratio distributions.	100
Table S5: Empirical false positive rates (all spectra and unflagged spectra only).	101

List of Figures

Figure 1: A sample full scan and fragmentation scan.	5
Figure 2: Difficulties with extracted ion chromatograms.....	8
Figure 3: Cleavage and back-labeling reactions in H ₂ ¹⁸ O.....	21
Figure 4: The theoretical full <i>b</i> -ion and <i>y</i> -ion ladders of peptide EIFNLYDEELDGK.....	23
Figure 5: Data analysis pipeline.	37
Figure 6: Spectral data interpretation for relative quantification.	41
Figure 7: Test protein quantification experiments.....	45
Figure 8: Global proteome quantification experiments.....	47
Figure 9: Multiphase LC-MS/MS proteome surveys.	51
Figure S1: Normal probability plots show long tails.....	102
Figure S2: Comparison of distributions of unflagged ratios and alarm-flagged ratios.	104
Figure S3: Plot of ratios calculated for individual spectra versus time during LC-MS/MS.	105

List of Appendices

Appendix A: Basic protocol	92
Appendix B: Supplementary material	95
B1. Supplementary tables	95
B2. Supplementary figures	102
B3. Notes on the effects of isotope distribution patterns.....	106

1. Introduction

1.1 Context

High-throughput identification of proteins has become possible through the development and convergence of a number of technologies over two decades. Tandem mass spectrometry incorporating collision-induced dissociation performed on peptides was determined to produce readily interpretable fragmentation patterns (Hunt, 1986; Arnott, 1993). Electrospray ionization and nanoflow liquid chromatography enabled (Fenn, 1989; Wilm, 1996), and multi-dimensional liquid chromatography enhanced (Washburn, 2001), the analysis of complex mixtures of peptides. Protein sequence databases were derived from genome sequencing projects and thereby created a foundation for the development of search software to match database entries to fragmentation spectra. Search programs such as Sequest (Eng, 1994) and Mascot (Perkins, 1999) have been in use for a decade and the identification methodology is now well established.

A natural corollary to high-throughput identification is the concurrent relative quantification of proteins between two or more samples. When performing a comparison between two complex protein extracts derived from cells or tissue subjected to different experimental conditions, an investigator often does not want only to generate a definitive list of the proteins present in each extract, but also an accurate measurement of their relative abundances. In this way, a picture of significant differential expression both in terms of global pathway activity and specific protein up- and down-regulation can potentially be assembled from the profiles and so be informative about the respective cellular states.

The application of tandem mass spectrometry to quantitative proteomics has been a highly active area of proteomics research. Many different approaches have been demonstrated

(see Aebersold and Mann, 2003; Bantscheff, 2007; and Elliot, 2009, for reviews), but each approach has its own advantages and disadvantages, and none have achieved primacy. Some methods are not truly general, in that they cannot be applied to all biological sources or all proteins, while others have not been fully characterized for statistical reliability (Mueller, 2008). Based on the extensive analogous quantification experience of the DNA microarray (gene expression) community, determining statistical significance in terms of differential expression between sample classes is often far more important than deriving a precise estimate of the actual fold-change in protein levels when it comes to drawing biologically-relevant conclusions (see Kerr, 2001; Reimers, 2005; and Allison, 2006 for reviews). This typically requires the support of specialized software tools to perform data analysis so as to draw meaningful quantitative inferences. Such tools are not necessarily available (or maintained, even if they are) or applicable for many of the quantitative proteomic approaches described to date.

1.2 High-throughput identification of proteins

By necessity, quantification is dependent on identification. There is usually little opportunity to advance one's understanding through observing that something varies if one does not know what it is. Ideally, the data necessary for quantification would be collected concurrently with the data necessary for identification, both to maintain the high-throughput nature of the experiment and to minimize the effort and uncertainty of determining the correspondence between quantitative results and identification.

Multi-dimensional LC-MS/MS (i.e. MudPIT; Washburn, 2001) is possibly the most recognized procedure for large-scale protein identification based on exhaustive protein shotgun sequencing. Proteins are first separated from a sample (often a cell lysate) extracted from a biological source such as:

- cell cultures (Washburn, 2001; Kirkland, 2008)
- tissues and tissue biopsies, clinical and experimental (Cagney, 2005; Chen, 2007; Lee, 2007)
- whole organisms (Mawuenyega, 2003; Brunner, 2007; de Godoy, 2008)
- secreted and excreted proteins, such as the urine proteome (Ru, 2006; Green-Church, 2008; Gonzales, 2009)
- blood serum or plasma, clinical and experimental (Fujii, 2004; Fujii, 2005)
- cell nuclei or mitochondria, sub-cellular proteomics (Kislinger, 2006; Lal, 2009)

A truly general quantitative method would be applicable to all such sources.

Typically, a lysate is first extracted from a biological source. The protein content of the lysate is separated from the other components, often as a precipitate. The separated proteins are denatured, reduced and alkylated (i.e. cysteine residues are chemically modified in order to prevent further reactions), then digested enzymatically to generate peptides. While other endoproteinases might be used, trypsin is often chosen for its processivity and specificity of cleavage sites, as it efficiently catalyzes hydrolysis at the C-terminal side of lysine and arginine residues, producing peptides with biophysical properties (e.g. average mass and charge state) amenable to detection by mass spectrometry.

Many quantitative methods (described individually in a following section) require that one or more steps beyond those necessary for simple identification be included (Panchaud, 2008). These steps not only introduce an additional chance for technical variation, but for peptide loss as well leading to biased and inaccurate estimates. Additionally, these methods often require that many of these steps be performed, in parallel, for two or more samples, before the products are combined for analysis.

An ideal quantitative approach would minimize additional experimental steps and handling. It would also integrate in a complementary (rather than contrary) manner with the process of identification through mass spectrometry, as described below.

1.3 Quantification at the full scan level versus at the MS/MS level

In contrast to identification, which relies on fragmentation spectra, relative quantification has mostly been performed at the full scan level (Ong, 2005; Bantscheff, 2007). Unfortunately, as outlined below, this approach suffers from some attendant problems.

As the chromatography proceeds, a tandem mass spectrometer is typically programmed to collect data in two modes. First, a survey of the incoming precursor ions is made in full scan mode (**Figure 1A**) over a range such as 400 m/z (mass over charge) to 2000 m/z . Then one or more peptide species are targeted for energetic activation and fragmentation. Usually the selection decision is made in real time by the software controlling the mass spectrometer concerning which ions to fragment through a process called data-dependent acquisition (or DDA), wherein several of the most prominent (highest intensity) ions observed in the previous full scan are isolated and fragmented before another full scan is scheduled to obtain an updated snapshot of the eluting peptides. Positively charged fragments of different m/z values then reach a detector and are recorded over a period of time. The list of m/z values of recorded fragments are assembled into a fragmentation spectrum (**Figure 1B**), with the intensities of peaks at particular m/z values being proportional to how many fragments were detected.

Figure 1: A sample full scan and fragmentation scan. (a) A full scan of eluting peptides taken partway through the chromatography of a tryptic digest of soluble *C. elegans* proteins. Although many discrete precursor ions are visible (indicated by m/z values), a basal level of background noise is evident. (b) A fragmentation scan of the precursor ion at 1584.75 m/z in the full scan, identified to be a peptide with the sequence EIFNLYDEELDGK. The signal to noise ratio of the product ion peaks is more pronounced.

To prevent the repetitious selection (for fragmentation) of peptides displaying prominent intensities, an exclusion list may be maintained. Upon having been selected a preset number of times for fragmentation, the mass-over-charge value of the peak is added to the exclusion list for a fixed interval (such as 30 seconds). Any peak falling within a small range of the noted mass-over-charge value is then ignored while the value endures upon the list. In this way, lesser peaks representing others peptides are also selected for fragmentation, maximizing the potential for unique identifications when the fragmentation spectra are analyzed. However, even with an exclusion list in effect, some candidates are left unselected due to the sheer number present in a digest of a complex protein sample, with more peptides eluting at a given time during the chromatography than can be processed. This problem is termed “under-sampling”.

Under proper conditions, a peptide will elute characteristically from a reverse phase (C18) microcolumn, producing a sharp chromatographic peak (a sharp peak is desirable since a broader peak generates a less intense signal and might therefore not rise above a threshold for detection). The recorded peptide ion intensity, which is proportional to the amount of any particular peptide eluting at a given moment as well as its propensity to ionize, might exceed the level of background noise for half a minute or less. During this time, it may therefore be observed in one or more full scans and be selected for fragmentation. When the identity of the peptide is established, post hoc via database searching (described in Section 2.4), specialized software can be applied to extract the peak intensities from the general full scan data and build an ion chromatogram for peak quantification. A curve is typically fitted to the peaks and the area under the curve calculated, with the quality of the result depending on how strongly the signal rose above the background, how many times it was observed, and how frequently the signal was sampled (America, 2008).

Best results are achieved when a peak representing an eluting peptide exceeds the level of background noise and is repeatedly observed before elution tapers off. If such a peak is only recorded in a small number of full scans, it becomes difficult to make a useful determination of the area under the curve (**Figure 2A**). This is, in part, dependent on the duration of the interval between full scans which occurs while time is devoted to performing fragmentation scans. Hence, a tradeoff exists between quantitative accuracy and the rate of identifications on those mass spectrometers incapable of operating in both modes simultaneously. Time spent collecting full scan data to use for quantification is time not spent performing fragmentation scans for purpose of identifying proteins, and vice versa.

Most of the available quantitative methods (described in Section 1.4) rely on the use of stable isotopes. By including both unlabeled and labeled forms of a peptide in a single experiment, technical variation between separate experiments (and, in certain label-free methods, the need to align peaks despite chromatography differences) is eliminated. Proteins and/or peptides present in a test sample are made to be heavier by the incorporation of a small, consistent atomic mass (usually between 2 to 10 Daltons) than the peptides in another control sample. This is accomplished through the metabolic, chemical or enzymatic incorporation of amino acids, chemicals or water, respectively, bearing heavy stable isotope versions of oxygen (^{18}O), nitrogen (^{15}N), and carbon (^{13}C), the use of which have been demonstrated in specific instances to not affect the elution characteristics of the peptides (Reynolds, 2002; Ong, 2003). With the same peptide differentiated between samples by a predictable mass-to-charge increment, aliquots of the samples can be mixed together at a point prior to the chromatography. The areas-under-the-curve are then calculated for co-eluting pairs of peptides from both samples, employing the data from the same LC-MS/MS experiment, and are compared.

Figure 2: Difficulties with extracted ion chromatograms. (a) A first discrete peptide precursor ion peak (#1) in the XIC rises strongly above the background noise and persists over many full scans. A second species (#2), however, appears only briefly; in this case, quantification would have benefited greatly from a faster full scan sampling rate. (b) Peaks representing three peptides are captured in a full scan (illustration only). The third, at 964.1 m/z , provides a clean signal, whereas the second at 961.1 m/z is interfered with by the isotopic envelope of the first at 960.7 m/z . (c) A high-confidence identification was made from a precursor ion at 1185.72 m/z , which never rose above the level of background noise. No quantification could be performed using the full scan data alone.

In some instances, identifications are made in which the labeled or unlabeled (or both) peak for a peptide doesn't rise even briefly above the level of the background noise (**Figure 2C**). Reliable relative quantification is impossible in such instances. Unrelated co-eluting peptides of approximately the same m/z have the potential to interfere with reliable quantification as well (**Figure 2B**).

While precursor ion peak quantification can be effective, compelling evidence indicates advantages for performing relative quantification at the MS/MS level (Yao, 2001; Heller, 2003; Zhang, 2006; Elbert, 2008). By specifying a sufficiently broad precursor ion isolation width, the co-eluting labeled and unlabeled peptide precursor ions can be captured together for joint fragmentation. Most tandem mass spectrometers, such as the ion trap instruments (used for my Thesis work) that are currently the workhorses for large-scale proteomic studies, produce markedly enhanced signal-to-noise levels when operated in MS/MS mode as compared to full MS precursor scans (*c.f.* **Figure 1**), providing a further advantage in terms of sensitivity and reliability (Venable, 2004). However, the greater the isolation width employed, the greater the potential chemical interference and the interference from near-isobaric, unrelated co-eluting peptides, which could lead to failed database search results and incorrect sequence identifications (described further in Section 3.1). The isolation width is therefore best kept as small as is feasible.

Additionally, by collecting quantification data through co-fragmentation at the MS/MS level, identification and quantification become complementary (rather than competitive) tasks. A successful identification must always yield the potential to calculate a useful relative ratio.

However, little work has been done to realize these advantages, with present high-throughput quantitative methods focusing instead on quantification at the MS level.

1.4 Present high-throughput quantitative methods

1.4.1 Metabolic labeling

^{15}N labeling, as a method for the relative quantification of proteins, was first described by Oda and Chait in 1999. One cell culture is grown on media with isotopes of nitrogen present in their natural abundance (99.6% ^{14}N and 0.4% ^{15}N) and another on media enriched for ^{15}N (content greater than 96%). Proteins are harvested from cells of each culture in an identical manner and combined in equal volumes, so that if the cultures themselves were identical, a one to one ratio of the unlabeled and labeled forms of each protein would be established. All of the subsequent steps are performed with the single sample, thereby providing the method's sole and major advantage, namely the elimination of some potential for experimental variation.

^{15}N media for mammalian cells are difficult to make and expensive as well (Ong, 2002), generally limiting the use of this form of labeling to single-cellular organisms. Unless the media provided contains ^{15}N exclusively (virtually 100%), the fragmentation spectra become more complex due to the combination of the fully labeled peptide and the various forms of the peptide that incorporate a single ^{14}N atom. This may impact the successful identification of peptides, since many sequences may match equally well.

Additionally, quantification is performed at the full scan level, with the previously described disadvantages plus the added disadvantage of variable separations between the precursor ions. The wide separation between unlabeled and labeled forms of peptides and the potential for incomplete incorporation of the ^{15}N label makes the suitability of this method for quantification at the fragmentation scan level questionable.

SILAC, an acronym for "stable isotope labels by amino acids in cell culture", was subsequently introduced by Ong and Mann in 2002. The SILAC approach takes advantage of the fact that given species' cells cannot synthesize certain amino acids. This inability to

synthesize all amino acids may be natural, as is the case for mammalian cells, or artificial, through creation of strains, for example, of auxotrophic yeast (Gruhler, 2005). To culture such cells, these essential amino acids must be supplied in the growth medium.

Accordingly, in a given experiment, two cultures are provided with one or more essential amino acids, isotopically light for one culture and isotopically heavy for the other. In its introductory form, deuterated leucine (+3 Da) was employed. This has since been updated by the authors to the use of arginine and lysine bearing six heavy carbon (^{13}C) atoms (+6 Da), ensuring that nearly every peptide produced through digestion with trypsin bears a labeled or unlabeled form of an amino acid (Ong, 2006). Five generations of cells are sufficient to ensure effectively complete (~97%) incorporation of the labeled amino acid(s) into the stock cultures from which the experimental cultures are created (Ong, 2002).

Cells from each culture are lysed in an identical manner and their lysates combined in even proportion. The main strength of SILAC, as with ^{15}N labeling, is that all subsequent steps involve the single, combined sample, thereby eliminating opportunities for the creation of experimental variation through separate handling. Incorporation of the labeled amino acids is virtually complete (100%) in many cell lines, even those for which arginine is nonessential in the corresponding adult animal (Ong, 2003).

Quantification via SILAC is then commonly performed at the full scan level, with the previously described disadvantages. The already wide separation between the peaks representing the unlabeled and labeled form of peptides is multiplied when such peptides are the product of incomplete digestion, reducing the method's suitability for fragmentation-scan quantification. Moreover, although SILAC can be applied to a broad range of experimental questions, it is not universally applicable. Its use has been extended beyond cell cultures to whole organisms, including mice (Krüger, 2008). However, it clearly cannot be employed to

analyze clinical samples (Miyagi, 2007) nor for comparison between tissues of intact human specimens, and is impractical for studies of other larger mammals as well.

When labeled arginine is included, a careful balance must be determined empirically and maintained to prevent interconversion of arginine and proline (Ong, 2006). In the case of the SILAC mouse, only labeled lysine (no labeled arginine) was included. Although this eliminated the problem of interconversion, it also reduced the number of labeled peptides and thus opportunities for successful quantification.

Additional considerations include the quantity of labeled amino acids required and the potential need to maintain fully labeled stock cultures.

1.4.2 Chemical labeling

The ICAT (isotope-coded affinity tag) approach was introduced in 1999 by Gygi and Aebersold and has seen continued development (Shiio, 2006). The isotope-coded affinity tag consists of three components. The first, a reactive group, is capable of chemically binding to a specific amino acid's side chain. The second is an affinity tag, and the third is an isotopically-coded linker joining the first two. In its common form, the affinity tag is biotin, the reactive group is thiol-specific (and therefore selective of cysteine residues), and the isotopically-heavier form of the linker exceeds the lighter form in mass by 8 atomic mass units.

Typically, the two samples to be compared are prepared separately according to the same protocol, drawing from two different biological sources, and are labeled separately with an excess of ICAT reagent (one with the lighter form and the other with the heavier form). Equal portions are combined and handled as a single entity from that point onward. After being subjected to digestion with an endoproteinase (commonly trypsin), the tagged peptides are isolated by means of avidin affinity chromatography. The isolate is then analyzed by liquid

chromatography coupled with tandem mass spectrometry (LC-MS/MS), through alternating full scans and scans of fragmentation patterns of precursor ions selected in data-dependent manner. Identifications are made using the fragmentation spectra and protein database matching software such as Sequest, while relative quantification is performed using an area-under-curve calculation for the light and heavy isotopic peaks.

In addition to suffering the problems of full-scan-based quantification, the ICAT approach has a number of unique disadvantages. Firstly, a relatively small fraction of proteins are innately undetectable due to the absence of cysteine residues (8.42% of the proteins in the known yeast proteome, for example). Secondly, many more proteins manifest only a small number of cysteine residues (47.78% of known yeast proteins include four or less; this increases to 77.82% of those having 200 residues or less). This inevitably excludes further proteins from detection. Thirdly, this same principle applies even to those proteins that are successfully identified; there will be fewer peptides per protein from which to derive an overall ratio. Accordingly, precision is lost. Fourthly, the additional handling (compared to simply performing a digestion and identifying proteins) can lead to loss of sample. Sensitivity is affected, for example, by poor recovery of low abundance proteins due to non-specific interactions with the column during the affinity purification (Graves, 2002). The number of proteins that can be identified is limited (Zhao, 2009).

Ross and Pappin described the so-called iTRAQ (“isobaric tags for relative and absolute quantification”) approach, which makes clever use of isobaric tags intended to split off as part of fragmentation, in 2004. To perform quantification using iTRAQ, samples are prepared or procured separately according to the chosen experimental conditions. Proteins from their lysates are purified, reduced, alkylated and digested with trypsin, also separately. A different iTRAQ derivitization agent is added to each digest and given time to covalently bind to all N-termini

and lysine residues of the peptides. In its original form, up to four different agents were used; this has since been extended to eight, allowing for the comparison of up to eight separate samples at a time (Pierce, 2008). Equal aliquots of the digests are combined and subjected to LC-MS.

Importantly, both identification and relative quantification is accomplished using fragmentation spectra. This is due to the nature of the iTRAQ tags. They are isobaric as bound to a peptide; a given peptide from different aliquots will have the same mass (+145 Da per tag) despite bearing subtly different labels. The differently labeled forms of each peptide are therefore all fragmented together during data-dependent acquisition. However, during the fragmentation process, the isobaric tags also split off and, critically, undergo a neutral loss of a balance group with the remainder, the reporter group, retaining a positive charge. Since each iTRAQ tag differs in the allocation of mass (through ^{13}C and ^{18}O isotopes) between the reporter group and the balance group, the reporter ions are distinct for peptides contributed by the different samples, ranging in mass from 114.1 m/z to 117.1 m/z . Relative quantification is performed by comparing the peak intensities of the different reporter ions.

The iTRAQ method of quantification has proven to be more challenging to implement on the popular ion trap mass spectrometers. Although pulsed-Q dissociation (PQD) now offers the opportunity for detecting the low mass iTRAQ reporter ions, this specialized procedure necessitates frequent collision energy adjustments that can lead to fewer protein identifications (Griffin, 2007). Additionally, if two nearly isobaric peptides are co-fragmented, one could potentially interfere with quantification of the other since each contributes its own array of reporter ions.

SILAC has lately been combined with the detection of immonium ions generated from labeled valine, leucine, and isoleucine residues; this approach has been named isobaric SILAC

with immonium ion splitting, or ISIS (Colzani, 2008). Although the expense of the iTRAQ reagents may be bypassed using the ISIS approach, and some of the drawbacks of the SILAC approach itself avoided (such as quantification at the full scan level), it effectively adds the remaining drawbacks of SILAC to the disadvantages of the iTRAQ approach.

ICPL stands for "isotope-coded protein label", an acronym first used by Schmidt and Lottspeich in 2005. The method is highly similar to the mass-coded abundance tagging (MCAT) approach described by Cagney and Emili in 2002, save that the ICPL derivitization agents are typically added to the protein extracts before reduction, alkylation and digestion (although it has been shown that they may suitably be added after digestion as well; see Paradela, 2009). The N-nicotinoyloxy-succinimide agent covalently binds to lysine residues and the N-termini of the proteins. One sample's proteins are labeled with a light form of the agent, while another's are labeled with a heavy-isotope form (+4, +6 or +10 Da). When subjected to LC-MS following digestion, a peptide bearing the light label and the same peptide bearing the heavy label co-elute. Quantification is performed at the full-scan level based on the integration of intensities of the twin peaks. The ICPL approach is well suited for quantification using 1D and 2D gel electrophoresis, since the labels are stable and corresponding proteins are not separated on the gel. The application of the agents before reduction, alkylation and digestion allows the samples to be compared to be combined prior to these steps, potentially reducing technical variation.

The generally wide separation of the unlabeled and labeled species of the peptides, particularly in those cases where the peptides are multiply labeled, make quantification at the fragmentation scan level problematic. If the label is applied following the digest, many peptides end up bearing more than one label (all have at least one on their N-termini). If the label is applied to intact proteins, many peptides end up bearing no label at all, reducing opportunity for quantification.

1.4.3 Synthetic reference peptides

AQUA, or "absolute quantification", was introduced by Gerber and Gygi in 2003. One or more proteins of known sequence are chosen to be quantified absolutely (i.e. in a manner that yields an actual physical estimate of the amount of the protein that is present in the sample). Peptide sequences that appear in a tryptic digest of these proteins are characterized and then selected according to their ability to ionize well and be detected by a mass spectrometer. The designated peptides are synthesized in such a way as to incorporate an isotopic label (typically representing a mass addition of 7 Da) and are spiked into a sample in a known quantity. The sample is then subjected to LC-MS. However, instead of data-dependent acquisition, the mass spectrometer is programmed to watch for certain peaks at certain times during the chromatography. When an anticipated peak appears, the corresponding peak of the labeled synthetic form of the peptide is also detected. The quantity of the peptide naturally occurring in the sample is computed by multiplying the quantity of the synthetic peptide by the relative peak intensities.

Although the AQUA approach does produce accurate results for selected proteins, it clearly does not provide a means for achieving, simultaneously, high-throughput and global quantification. This is true of other targeted approaches to peptide detection such as multiple-reaction monitoring (MRM) as well (Yocum, 2009). Characterization of proteotypic peptides across an entire proteome and synthesis of these peptides currently requires a very significant investment of time and resources, as would creating a mixture of all the different synthetic peptides to spike into a sample.

1.4.4 Label-free quantification

Spectral counting (or spectrum counting) originated with the growing realization that the greater a given protein's presence is in a sample, the more fragmentation spectra are successfully identified as peptides belonging to that protein (Washburn, 2001; Liu, 2004). Quantification is performed by simply tallying these "spectral counts" for a protein in different LC-MS experiments, each experiment representing a different sample, and directly comparing them to determine a ratio. For example, if thirty spectra matched to a protein in one experiment, and fifteen matched to that protein in a second experiment, it is assumed that roughly twice as much of that protein was present in the first. The spectral counting approach has the advantage that no label need be applied at all, nor is quantification dependent on the measurement of precursor ion intensities. Additionally, direct comparison is possible for an indefinite number of samples (given a minimum of one LC-MS/MS experiment per sample) rather than only two (or eight, in the case of iTRAQ) samples.

Spectral counting has been termed a "semi-quantitative" method. Although it is capable of accurate relative quantification for high abundance proteins, the accuracy diminishes as the number of counts gets low. Although as few as four spectral counts might suffice to declare a three-fold change, reasonable certainty of a two-fold change could require a minimum of fifteen counts (Old, 2005). The number of counts necessary grows rapidly when trying to reliably detect degrees of change smaller than two-fold. Accordingly, accurate relative quantification cannot be made for lower abundance proteins, which manifest only a few or even a single count. The use of dynamic exclusion (described in Section 1.3) exacerbates this problem; in the attempt to achieve as many different identifications as possible, opportunities to establish viable spectral counts may be sacrificed (Zhang, 2009). Conversely, it has been reported that saturation, namely a region of non-linear correspondence between peptide quantity

and spectral counts, can occur for the most highly abundant proteins (Bantscheff, 2007). This is in direct contrast to a MS/MS-based co-fragmentation quantitative approach in which (in principle) only one spectrum is necessary to obtain both identification and a relative ratio.

The extracted ion chromatogram (or XIC) approach compares the sum of the extracted ion chromatograms for a protein as they appear in two or more different LC-MS experiments. The underlying assumption is that the experimental variation between these runs will not be so large as to make the relative results unreliable. This problem is also mitigated by using one or more protein standards spiked into each sample and normalizing an experiment's results against those calculated for the standards before making a comparison between experiments (Bondarenko, 2002; Higgs, 2005). An improved form of this label-free approach relies on high-resolution mass spectrometers that can separately detect the monoisotopic peaks and their associated isotopic peaks and thereby include all of the relevant information in the XIC while excluding unrelated, but near isobaric, peaks (Ono, 2006). High-resolution data, requiring high resolution mass spectrometry, is necessary for the best results. Attention must be paid to maintaining high reproducibility of the chromatograms between LC-MS runs, in order that events may be aligned between runs and corresponding peptides identified. Even so, the removal of the most abundant proteins may be necessary to obtain data for lower abundance proteins (Wang, 2006).

The tradeoff between allotting time on the mass spectrometer toward full scans to acquire quantitative data and allotting time for fragmentation scans to identify peptides applies here; increased quantitative accuracy comes at the expense of the number of identified peptides. To get around this, Strittmatter, Zimmer and Smith devised the accurate mass and time tag (AMT) approach (Strittmatter, 2003; Zimmer, 2006), in which one LC-MS experiment for a sample records only full scans for the purpose of quantification, while a second LC-MS run is

devoted to achieving identifications. The drawback, however, is the necessity to perform additional runs and to apply complex software which potentially introduces an additional source of error into the results.

1.5 ^{18}O -based stable isotope labeling

^{18}O -based stable isotope labeling using heavy water during sample digestion by trypsin was initially pioneered as a proteomic tool for quantification by Yao and Fenselau in 2001 and has since seen increasing, albeit not widespread, use in quantitative proteomic studies (Sakai, 2005; Miyagi, 2007). The method can, and has, been applied to proteins taken from diverse sources (see for example studies by An, 2005; Chen, 2005; Hood, 2005; Nelson, 2006; Diamond, 2007; Ang, 2008; Petyuk, 2008; Smith, 2008).

Two procedures for achieving ^{18}O labeling have been developed. In the first method, proteins from two samples to be compared are precipitated. One sample is resuspended in a solution containing isotopically-normal water (^{16}O) and the usual steps of digestion are performed. The other pellet, however, is resuspended in a solution containing enriched (95+ %) ^{18}O water. Upon hydrolysis of the latter proteins by trypsin, a 'heavy' ^{18}O atom becomes incorporated into the C-terminus of each newly cleaved peptide (except the terminal peptides bearing the native C-termini). This results in an addition of 2 Daltons relative to the mass of the corresponding 'unlabeled' peptide produced by digestion in normal 'light' (^{16}O) aqueous buffer conditions. Moreover, trypsin continues to interact with the C-terminal arginine and lysine residues of the peptides, resulting in a process called "back-labeling" wherein oxygen atoms from the water are further exchanged with the C-terminal oxygen atoms (Schnölzer, 1996; Miyagi, 2007). Since the vast majority of the oxygen atoms from the heavy isotope labeled water are the ^{18}O isotope, the peptides tend to acquire a second C-terminal ^{18}O atom over time,

increasing the delta mass to 4 Daltons (**Figure 3**). In the second procedure, the proteins of both samples are digested separately in regular aqueous buffer. After drying in a vacuum centrifuge, one sample is resuspended in isotopically-normal water, while the other digest is solubilized in highly-pure ^{18}O water (Yao, 2003; Fenselau, 2007). When fresh trypsin is added to both, back-labeling proceeds such that the peptides in the ^{18}O water (ideally) take on the extra 4 Daltons of mass (with the exception of native C-terminal peptides of the original proteins that do not possess a C-terminal arginine or lysine residue).

Labeling during digestion ensures that all peptides (again, save for the C-terminal portion of each protein) to be labeled are at least partially labeled (subject to the degree of impurity of the ^{18}O water), since an ^{18}O atom must be added during hydrolysis. Back-labeling, while usually effective (Yao, 2003), provides no guarantee of this; the C-termini of recalcitrant peptides may not interact with trypsin within a reasonable period of time. Also, post-digestive back-labeling involves drying the digests, which can lead to significant losses from adsorption.

For LC-MS/MS analysis, the ^{18}O -labeled digest and the unlabeled digest are normally mixed together in equal proportion and spiked with acid to lower the pH sufficiently to quench the trypsin reaction. During chromatography on reverse phase columns, the labeled and unlabeled sister peptide species have been observed to elute coincidentally (Reynolds, 2002) and so are available for simultaneous detection on an associated tandem mass spectrometer, nominally separated by 4 m/z and 2 m/z for the singly- or doubly-charged precursor ion species respectively. Quantification at the full scan level may then be performed.

Figure 3: Cleavage and back-labeling reactions in H_2^{18}O . The initial trypsin cleavage of a polypeptide substrate at an internal Lys or Arg residue results in incorporation of a first ^{18}O atom into the new C-terminus of the N-terminal portion. The second ^{16}O atom can be substituted with ^{18}O atom in a slower back-labeling reaction (Schnölzer, 1996). Oxygen atoms continue to be swapped between the C-terminus and water in this back-labeling process, so long as active trypsin is present. In the presence of a vast excess of heavy isotope labeled water (95% pure H_2^{18}O or better), the tendency is for the C-terminus to acquire substitutions such that both oxygen atoms are ^{18}O (i.e. become fully labeled).

Fragmentation is accomplished through a process such as collision-induced dissociation (CID) in which the ions within a small range of the selected m/z are accelerated through an inert gas. When a peptide ion collides with a molecule of the gas, some of the energy of collision is absorbed and causes the peptide to break into two portions. If it is a peptide bond that is severed, as often occurs under favorable conditions, the peptide is divided into a b -ion fragment (N-terminal portion) and a y -ion fragment (C-terminal portion). Since each peptide ion might fragment at any of the peptide bonds, the result for many instances of the same peptide is a discernable pattern (or “ladder”) of b -ions and y -ions that is specific to the sequence of amino acids making up the peptide (**Figure 4**).

Although ^{18}O labeling offers simplicity and is generally applicable to any source of protein, the difficulties of quantifying at the full scan level (described in Section 1.3) are increased by the potential for partial labeling, which arises from two sources. Firstly, the back-labeling process isn't uniform; peptides incorporate the label at different rates (Stewart, 2001). Although most peptides will become fully labeled if sufficient time in the presence of active trypsin is allotted, some will not. Secondly, commercial sources of ^{18}O water are typically only available at purities ranging from 95% to 99%, and so will not be 100% pure. The presence of a low percentage of ^{16}O water results in partial labeling or, to a very small extent, no heavy label at all, which needs to be factored in for quantitative precision.

It is possible to compensate for this effect through examination of the contribution of partially labeled peptides to isotopic peaks (Fenselau, 2007). However, this requires high resolution mass spectrometers such as the Thermo Scientific LTQ-FT or Orbitrap, as it can be otherwise difficult to distinguish these species from the native isotopes (like ^{13}C) for multiply-charged precursors. Hicks *et al.* partly addressed this problem in 2005 with the addition of a ‘zoom scan’ step to improve resolution along with supporting software (ZoomQuant), but at the

#	m/z	b-ion ... y-ion	m/z	#
1	130.0	E ... IFNLYDEELD GK	1455.7	12
2	243.1	EI ... FNLYDEELD GK	1342.6	11
3	390.2	EIF ... NLYDEELD GK	1195.5	10
4	504.2	EIFN ... LYDEELD GK	1081.5	9
5	617.3	EIFNL ... YDEELD GK	968.4	8
6	780.4	EIFNLY ... DEELD GK	805.3	7
7	895.4	EIFNLYD ... EELD GK	690.3	6
8	1024.5	EIFNLYDE ... EL DGK	561.3	5
9	1153.5	EIFNLYDEE ... LD GK	432.2	4
10	1266.6	EIFNLYDEEL ... DGK	319.2	3
11	1381.6	EIFNLYDEELD ... GK	204.1	2
12	1438.6	EIFNLYDEELD G ... K	147.1	1

Figure 4: The theoretical full *b*-ion and *y*-ion ladders of peptide EIFNLYDEELD GK.

There are twelve amide backbone bonds in the peptide that represent possible breakage points when fragmented by collision-induced dissociation. They are predicted to yield twelve complementary singly protonated *b*-ions and *y*-ions products. Typically, not all ions of the ladders are detected; in this example, the mass-over-charge values of the ions that were actually observed are shown in bold.

price of reduced scan time devoted to MS/MS and subsequent decreased opportunities for peptide identifications. Sophisticated isotope envelope peak modeling algorithms have also been developed (Qian, 2005; Mason, 2007), applying such techniques as peak fitting in an isotopic cluster by means of linear regression models. While the peptide coverage and quantitative accuracy achieved by such studies benefits markedly from the use of higher resolution instruments, inferences based on these measurements are still subject to the effects of instrument noise, which are more pronounced at the full scan level (Venable, 2004).

Since it is the C-termini of peptides that are modified by the ^{18}O -labeling method, if the labeled and unlabeled forms of a peptide are co-fragmented, the *b*-ion ladder will be unchanged but every singly-charged *y*-ion will exist as a duplet separated by 4 Daltons. Relative quantification can then be performed by comparing the unlabeled and labeled *y*-ions.

Accordingly, it might be expected that by combining ^{18}O labeling with quantification at the fragmentation scan level, the disadvantages of the historical approach are circumvented. Any protein successfully identified through standard spectral database matching would also yield a relative quantification. Since quantification is combined into the identification process, there would be no trade-off in terms of the rate of MS/MS spectrum acquisition; hence, the comprehensiveness of protein identification should be unimpeded. Additionally, the problem of deconvoluting the contributions of doubly-charged labeled and unlabeled ions on mass spectrometers with lower resolutions is neatly sidestepped through examination of the predominant singly-charged (and therefore more widely separated) *y*-ion pairs evident in MS/MS product ion spectra. Since *y*-ions are typically singly charged, the determination of the contributions of partially-labeled species as well is simplified as well. The necessary isolation width is 5 to 10 mass-to-charge units, significantly smaller than would be required for more widely separated sister peptides as appear in other methods.

This modified approach, however, has yet to be demonstrated at a global profiling level and characterized for robust use by the proteomics community.

1.6 Thesis rationale

^{18}O labeling itself is simple to perform. There are no side reactions of the sort possible in chemical labeling. It can be applied universally to proteins from any source. Only a single additional experimental step over those necessary for identification, namely the mixing of unlabeled and labeled digests, is required. These qualities are all desirable to a potential investigator. However, practical use has remained limited to quantification at the full scan level, with the general attendant problems of that approach plus the additional ^{18}O -labeling-specific problem of determining the contribution of partially labeled species.

I reasoned that quantification through co-fragmentation at the MS/MS level would offer the opportunity for improved accuracy and precision relative to quantification at the full scan level, even using standard resolution ion trap mass spectrometers available to many labs, including our own. No tradeoff between quantification and identification would be necessary; a single identification would suffice. Quantitative data would not be hidden in the background noise. The potential for saturation is reduced as the precursor ion signal is divided among many fragments.

Additionally, I reasoned that the partial labeling problem and deconvolution of overlapping contributions of isotopes of multiply charged species would be simplified by moving its evaluation to the MS/MS level, as the singly-charged y -ion pairs are sufficiently separate to enable direct calculation of the partially labeled contribution.

To this end, I've performed what I believe to be the first application of global proteomic profiling of complex samples by means of ^{18}O labeling and quantification at the MS/MS level through co-fragmentation, as well as the first systematic correction for partial labeling at this level. Empirically suitable confidence level cutoffs for determining statistical significance in terms of differential expression were derived as well. As described in Chapter 3, the results demonstrate conclusively that this approach provides quantitative accuracy within a biologically intuitive overall dynamic range.

As an additional benefit, a set of rules to capitalize upon the extra information present in the MS/MS spectra (namely, the redundant series of y -ions resulting from co-fragmentation) were developed to flag dubious results and detect putative false positive database matches. This represents the first attempt to apply this information toward a measure of quality of identification.

The absence of existing software tools (that support ^{18}O -label-based global comparative proteomic profiling analyses in which quantification is performed at the fragmentation scan level) necessitated the development of two supporting software programs, *ySelect* and *yRatios*, that accept database search results and report in a readable form, in a comma-separated-values form recognized by Microsoft Excel, and in formats recognized by GoMiner (Zeeburg, 2003) and the Cytoscape BiNGO plug-in (Maere, 2005).

My central intention has been to establish a robust proteomic pipeline that is easy to implement experimentally, is generally applicable, does not share the disadvantages of other quantitative methods, and is of broad appeal to the proteomics and clinical communities. The major part of this work was published in the *Journal of Proteome Research* (White, 2009).

2. Methods and materials

2.1 Acquisition of the test proteins

Powdered purified bovine serum albumin (BSA) was purchased from Fisher Scientific. For the global profiling experiments, populations of wild-type *C. elegans* (laboratory strain N2) were raised from an initial stock kindly provided by the Caenorhabditis Genetics Center and maintained at room temperature on 60mm NGM plates seeded with *E. coli* OP50. After synchronizing cultures by bleaching and observing the emergence of young adults, worms from ten plates were harvested, pooled together and washed. Following separation through sucrose floatation, the adult worms were collected and centrifuged at 1500 rpm. The volume of worms was measured and the worms then resuspended in an equal volume of an ice-cold lysis buffer composed of M9 buffer (3.0g KH_2PO_4 , 6.0g Na_2HPO_4 , 5.0g NaCl, 1 mL of 1M MgSO_4 , all in 1L of distilled H_2O) with an additional 1% Triton X-100, and Complete protease inhibitors purchased from Roche Applied Science. 1.4 mL of the resuspension was added to a mini bead beater vial holding 0.7 mL of 1 mm zirconia beads, thereby filling the vial completely. The cuticles were broken open and cells lysed through six cycles of strenuous beating, each cycle consisting of a 20 second period at maximum speed on a Biospec Products MiniBeadBeater-96 followed by a 1 minute period of immersion in ice-water to cool. The resulting worm lysate was clarified by 15 minutes of centrifugation at 21,000 g and the soluble protein supernatant subsequently recovered, divided into aliquots and stored at -80°C to await use. A Bradford assay indicated a protein content of 16 mg/mL in the clarified lysate.

2.2 Sample preparation for mass spectrometry

Sequence grade modified trypsin from Promega was dissolved in sufficient HPLC-grade water (purchased from Fisher Scientific) to achieve a stock solution with a concentration of 0.25 mg/mL. The solution was divided into 20 μ L aliquots and stored at -80°C prior to use. A second stock solution was prepared and stored in an identical manner save that enriched ^{18}O water of 95% purity (from Isotec) was used in place of HPLC-grade water.

Similarly, two 800 mM (16X) stock solutions of NH_4HCO_3 (ammonium bicarbonate, or AmBic) were prepared, one in HPLC-grade water and the other in 95% ^{18}O water. These were stored at -80°C prior to use in buffering the digestion reactions.

A measured amount of BSA was dissolved in HPLC-grade water, establishing a BSA suspension with a concentration of 10 mg/mL, and used immediately.

Unlabeled tryptic digests of BSA were made as follows. A 4 μ L volume of BSA suspension (representing 40 μ g of BSA) was combined with 64 μ L of HPLC-grade water and 4 μ L of the 800 mM AmBic stock in HPLC water (thawed on ice) then mixed gently. An 8 μ L volume of the dissolved trypsin stock in HPLC water (thawed on ice) was added, achieving a final buffer concentration of 50 mM AmBic. After further gentle mixing, digestion was allowed to proceed for 48 hours at 30°C before storage at -80°C .

Labeled tryptic digests of BSA were made and stored in an identical manner save that 64 μ L of 95% ^{18}O water, 4 μ L of the 800 mM AmBic stock in 95% ^{18}O water, and 8 μ L of the dissolved trypsin stock in 95% ^{18}O water were used. The final proportion of ^{18}O water was calculated to be 89%.

The worm soluble protein content was isolated by precipitation of 4 μL of the clarified lysate (estimated 64 μg of total protein) with 100 μL of ice-cold 25% (w/v) trichloroacetic acid. After periodic mixing over 30 minutes while on ice, the precipitate was centrifuged at 21,000 g for 30 minutes. The supernatant was removed and the pellet was rinsed twice with ice-cold acetone then allowed to dry in the open air for 30 minutes. To create unlabeled digests, the dry pellets were immersed in 2 μL of the 800 mM AmBic stock in HPLC-water (thawed on ice) and 34 μL of HPLC-grade water, and resuspended by vigorous pipetting up and down. Once all particulates had disappeared, 4 μL of dissolved trypsin stock in HPLC water was added and the digests were allowed to proceed with continuous gentle mixing for 48 hours at 30°C. The same steps were performed to create labeled soluble worm protein digests, except that 95% ^{18}O water and stocks made in 95% ^{18}O water were used. The digests were stored at -80°C prior to analysis by mass spectrometry.

Chosen ratios of unlabeled to labeled proteins were formed by mixing proportional volumes of the labeled and unlabeled digests. For instance, a sample containing a one-to-two ratio of unlabeled to labeled soluble worm proteins was prepared by firstly depositing 5 μL of unlabeled digest in a microcentrifuge tube, secondly adding two μL of 95% formic acid and mixing gently, thirdly adding 10 μL of labeled digest and mixing gently, fourthly adding 5 μL of HPLC-grade water, and finally adding 1 μL of acetonitrile and mixing gently. Trypsin-mediated back exchange of one or both ^{18}O atoms on the C-termini of labeled peptides with ^{16}O atoms from normal water (and vice versa for unlabeled peptides) could potentially occur upon combining aliquots of the ^{16}O and ^{18}O digests, distorting the apparent ratio. It was observed to be important to maintain a pH value close to 3.0 in order to inactivate the trypsin; formic acid was therefore included (Stewart, 2001).

2.3 LC-MS/MS analysis

Capillary-scale nanoflow 150- μ m I.D. fused silica tubing purchased from Polymicro Technologies was cut and pulled to a fine tip using a P-2000 laser puller from Sutter Instruments, creating microcolumns. Single-phase columns were packed with a length of 5 centimeters of 5 μ m Zorbax 300 SB-C₁₈ resin from Agilent Technologies. Triple-phase columns were packed first with 5 centimeters of the Zorbax resin, then with 5 centimeters of 5 μ m “PolySULFOETHYL A” strong cation exchange resin from PolyLC, and then again with 5 centimeters of the Zorbax resin.

Each column was flushed with methanol for fifteen minutes using a pressure vessel and regenerated with chromatography buffer ‘A’, consisting of 5% acetonitrile, 0.5% glacial acetic acid, and 0.01% n-heptafluorobutyric acid in HPLC-grade water, for at least 20 minutes. Samples were loaded onto columns to completion using a pressure vessel. Following loading, a column was connected to a dedicated ThermoFinnigan Surveyor quaternary HPLC pump and coupled to a ThermoFinnigan LTQ ion trap mass spectrometer.

A water/acetonitrile gradient, starting with 100% chromatography buffer ‘A’ and changing to 20% buffer ‘A’/80% acetonitrile over the course of 100 minutes, was applied to single-phase columns to elute peptides bound during loading. Eluting peptides were analyzed by electrospray ionization and selectively chosen for fragmentation by the controlling software’s (XCalibur’s) data-dependent algorithm, with the dynamic exclusion list feature enabled to prevent repeated acquisition of the same precursor within a thirty second window.

Peptides bound to triple-phase columns were subjected to six sequential chromatography separations over a 10 hour period. The first chromatography in the first step proceeded as described for single-phase columns, while the latter five steps incorporated preliminary salt bumps of 10%, 20%, 30%, and 40% of 250 mM ammonium acetate in buffer ‘A’, then finally

100% of 500 mM ammonium acetate in buffer 'A'. The salt bumps served to each transfer a new subset of peptides from the cation exchange material to the tip-ward reverse phase material for elution and analysis during the application of the buffer 'A'/acetonitrile gradient (Washburn, 2001; Kislinger, 2003).

The operating parameters of the ion trap mass spectrometer were set as previously described (Kislinger, 2003; Wolters, 2001), save that an isolation width of 10 m/z ($\pm 5 m/z$ around each precursor ion selected for fragmentation) was specified for all experiments (except where otherwise stated). This permitted co-eluting unlabelled and ^{18}O -labeled sister peptide ions to be trapped and simultaneously subjected to fragmentation. Both full scan mass spectra and data-dependent MS/MS spectra (up to ten following each full scan) were recorded in centroided mode. A single microscan was devoted to each MS/MS spectrum.

2.4 Identification of proteins

The WormPep facility (Sanger Institute) was accessed online August 2006 to generate a list of putative *C. elegans* protein sequences in FASTA database file format. A reversed amino acid sequence for each of the listed proteins was appended to the database under an artificial protein name to serve as search decoys and thereby assist in estimating the false identification rate. The utility of these search decoys is based on the underlying assumption that, if a false identification occurs, it is equally likely that it will be made to a 'peptide' of a search decoy protein as to an actual peptide of the database. The appearance of the decoy matches in the list of identifications provides the investigator with an additional means to assess the reliability of the identifications. In principle, for example, if a 99% confidence cutoff is employed, then only 1 match of 100 should be to a search decoy.

The Sequest database search program was invoked on a cluster computer to correlate peptide sequences to MS/MS spectra and rank the matches obtained. Briefly, Sequest examines each fragmentation spectrum individually and compares it to theoretical spectra generated for peptides of similar mass as derived from a protein database, generating cross-correlation scores. The top several matches (namely, peptides for which the highest scores were determined) are output in ranked order to a file adjunct to the spectrum's peaks file.

Two Sequest searches were performed for each profiling experiment. Firstly, a search in which a static C-terminal modification of 4 Daltons was specified to represent the complete labeling of the C-terminus of a tryptic peptide with two ^{18}O atoms. Secondly, a search where no such modification was specified, corresponding to the unlabeled form of the peptides. Both labeled and unlabeled peptides could therefore be identified without unduly extending the search mass tolerances.

Following each Sequest search, the DTASelect program (Tabb, 2002) was executed to examine the search results and create an exhaustive list of the top candidate peptide matches to the MS/MS spectra. Since two searches were performed for each profiling experiment, two such lists were generated for each experiment as well. The ySelect program (developed as part of this work) was then executed to read both lists for each experiment, calculate a confidence level for each match using the Statquest probabilistic model (Kislinger, 2003), and generate a new list of matches that exceeded a specified confidence cutoff. All ySelect runs were performed with the inclusion of singly-charged ions enabled and with a specified 99% minimum confidence cutoff except where otherwise noted.

2.5 Calculation of relative quantities

The yRatios program (also developed as part of this work) was employed to calculate a ratio of relative abundance for each spectrum referenced in a given ySelect-generated list, based on the expectation that the unlabeled and labeled forms of peptides eluted in simultaneous proportion and so were captured together to be fragmented by virtue of the wide isolation width. Results were also calculated and reported at the peptide and protein levels.

The expected monoisotopic mass for a given unlabeled C-terminal fragment (specifically, a *y*-ion) was calculated by applying the knowledge of standard amino acid residue masses to the matching peptide sequence assigned in a Sequest search. Any peak observed to be within ± 0.4 m/z of this expected mass was included in the total peak intensity attributed to the *y*-ion. The same tolerance was granted to the expected masses of the partially-labeled and fully-labeled forms of a *y*-ion. If the expected monoisotopic mass value of an N-terminal fragment (specifically, a *b*-ion) was determined to fall within -1.0 m/z to $+5.0$ m/z of the monoisotopic mass of an unlabeled *y*-ion, that *y*-ion pair was excluded from calculations to avoid potential interference. Similarly, to avoid interference from the intact (unfragmented) precursor ion and its cluster of associated peaks that appear at slightly lower m/z values, any *y*-ion pair for which the monoisotopic mass of an unlabeled *y*-ion was determined to fall within -20.3 m/z to $+0.30$ m/z of the precursor m/z value was also excluded from analysis.

2.6 Obtaining the software

Copies of the ySelect and yRatios source have been submitted as files adjunct to this Thesis, along with their associated user guides.

Additionally, current versions of the ySelect and yRatios programs are freely available for download and use, without restriction, from the Emili laboratory via the internet (<http://emililab.med.utoronto.ca/>), as are their supporting user manuals.

3. Results

Dr. Emili supervised and/or advised the experimentation. I performed all of the experimentation, including both sample preparation from lysates and mass spectrometry. I also performed all of the analyses, including the protein database searches, the ySelect and yRatios runs, and all calculations and derivations. I implemented the ySelect and yRatios programs in the C programming language and wrote their accompanying user guides. Nico Oey assisted me by maintaining and harvesting cultures of *C. elegans* and preparing whole-worm soluble extracts.

Portions of this chapter have been reproduced or adapted from:

White, C.A.; Oey, N.; Emili, A. Global Quantitative Proteomic Profiling through ^{18}O -Labeling in Combination with MS/MS Spectra Analysis. *J. Proteome Res.* 2009. 8, 3653. *

* Reproduced in part with permission of the American Chemical Society, © 2009 (first author)

3.1 Outline of the workflow

A diagram illustrating the sequential steps of the standard sample analysis pipeline of this method is presented in **Figure 5**. In this workflow, proteins are harvested (extracted) from two separate biological samples to be compared both qualitatively and quantitatively. The tryptic digestion itself (and therefore the labeling) is performed using an optimized procedure, essentially as previously described (Sakai, 2005; Fenselau, 2007; Sevinsky, 2007; see **Appendix A** for protocol details). Equal aliquots (in terms of total protein content per sample) of the two proteolytic digests are combined immediately prior to LC-MS/MS analysis. Standard data-dependent MS/MS data acquisition scanning procedures may be used, save that the isolation window mass range is slightly larger than usual in order to jointly target and co-fragment co-eluting unlabeled and ^{18}O -labeled sister peptide species. The spectra are searched in batch mode using a standard database search algorithm such as Sequest (Eng, 1994). Following application of the search filtering program ySelect, which applies the Statquest model (Kislinger, 2003) to database match qualifiers (such as Sequest's XCorr and deltCN) in order to cull the set of MS/MS spectra of those that fail a user-specified confidence level cutoff, the selected spectra are analyzed by the ratios calculation program yRatios which compares certain product ions of co-fragmented labeled and unlabeled sister peptides to derive a relative quantification between the levels of the protein in the two different samples.

Figure 5: Data analysis pipeline. Proteins are harvested from two different samples that will be quantitatively compared. Digestion with trypsin proceeds in an established manner, except that one reaction is performed using buffer containing 95% pure ^{18}O water and the other using HPLC-grade natural water. Equal aliquots of the digests are combined, loaded onto a nanoflow chromatography system and LC-MS/MS performed on the eluate using a tandem mass spectrometer. The resulting fragmentation spectra are first assigned identifications (in this study, using the database search algorithm Sequest). The set of tentative identifications is further processed by the ySelect program according to user-specified quality criteria, with high confidence identifications and their corresponding spectra files included in an output file. This list is then read by the yRatios program, which calculates relative quantities based on the observed y-ion pairs and generates quantitative results in a format specified by the user (readable, comma-separated values, or suitable for other meta-analysis tools).

To affirm the utility of this pipeline, the reliability of the basic approach was evaluated based on one-step reverse phase analyses of a single representative protein (BSA), mixed at various label ratios, and by tri-phase multi-step LC-MS/MS profiling experiments of a complex mixture (*C. elegans* soluble protein extract) performed in varying label ratios. A conventional linear ion trap instrument (ThermoFinnigan LTQ), which is a common workhorse platform within the proteomic profiling community, was used to perform automated MS-coupled data-dependent MS/MS with dynamic exclusion enabled. A 10 m/z isolation width was specified when triggering CID events. This wider isolation width allowed the inclusion of product ions derived from both the labeled and unlabeled sister peptides, including singly-charged species, whereas a more typical isolation width of 2-3 m/z might preclude certain doubly charged ions from being jointly captured. Although the overall protein identification rate for complex mixtures may be lowered by 20% or more (**Supplementary Table S1**), depending on the isolation width employed (presumably since unrelated peptides are more likely to be co-fragmented and confound recognition), as described in Section 3.5, database search stringency benefits from the inclusion of redundant y-ion series in the MS/MS spectra (specifically, by lowering the false discovery rate).

Following conversion of the raw mass spectrometry data to the DTA file format, the database search algorithm Sequest was invoked to assign peptide identifications to MS/MS spectra using a FASTA sequence database. This database was further populated with an equivalent set of reversed decoy sequences corresponding to each valid entry (Peng, 2003; Moore, 2002) to provide an estimate of the false identification rate based on the appearance of the reversed entries among the putative identifications. For this study, confidence scores (probabilities) were estimated using the Statquest model (Kislinger, 2003). However, reversed

sequences were still included in the searches to empirically evaluate criteria for subsequently detecting questionable identifications.

Two programs (written in the C programming language) were developed to interpret the search results and calculate relative quantity ratios for each of the detected proteins. The ySelect program examines the initial database matches and applies a well-established quality filter (Kislinger, 2003) to selected spectra that meet a user-specified cutoff for peptide identification confidence. It produces a list of fragmentation spectrum files which are subsequently read by the yRatios program. This second program calculates both an unlabeled y-ion intensity tally and labeled y-ion intensity tally based on summing the respective peak intensities across each spectrum (the labeled tally includes the contributions of individually-calculated contributions of partially-labeled y-ions as described below). These paired values are likewise summed up for the entire set of spectra matching a given peptide, and likewise among the various peptides assigned to the same cognate protein. Finally, a summary report is created at the level of detail specified by the user (i.e. results obtained for individual fragmentation spectra, or at the peptide or protein levels) and format (i.e. readable or in comma-separated values form) pre-selected by the user. Formats recognized by GoMiner (which classifies the genes into biologically coherent categories and assesses over- and under-representation in these categories) and the Cytoscape visualization plug-in BiNGO (which maps the predominant functional themes of, in this instance, the yRatios-produced list of differentially-expressed proteins onto the Gene Ontology hierarchy and displays this mapping as a Cytoscape graph) may also be specified.

3.2 Calculation of peptide ratios

A representative co-fragmentation spectrum, obtained by simultaneously subjecting an equivalent amount of labeled and unlabeled forms of the BSA tryptic peptide HLVDEPQNLIK to MS/MS, is displayed in **Figure 6** to illustrate key attributes of the data analysis algorithms. Pairs of unlabeled y -ions and fully ^{18}O -labeled y -ions ($^{18}\text{O}_2$, +4 m/z) are annotated. One such pair, along with their corresponding native (i.e. ^{13}C) isotopic envelopes, is shown in the expanded ‘zoom-in’ view. Since incorporation of the ^{18}O label may not be complete (Schnölzer, 1996; Miyagi, 2007), an estimate of the actual portion of an intermediate peak that is likely accounted for by partially labeled y -ions ($^{16}\text{O}_1^{18}\text{O}_1$, +2 m/z), isobaric with the endogenous $^{13}\text{C}_2$ -isotope of the unlabeled peptide species, is also highlighted.

Assuming the cleavage and exchange reaction is performed to completion, digestion in commercial high (95%) purity grade ^{18}O water should result in only a minor proportion of partially labeled peptides (~10%), and a relatively inconsequential trace of unlabeled species (<1%). However, since only one ^{18}O atom is actually added enzymatically during the cleavage process, the endogenous ^{16}O atom already present at the C-terminus of the peptide being formed must also be efficiently exchanged with an ^{18}O atom from the heavy water, which can impact the proportion of partially-labeled peptide produced in certain sub-optimal instances (Stewart, 2001). A degree of back-labeling (post-digestion swapping of C-terminal peptide ^{18}O atoms with ^{16}O atoms) may conceivably occur due to exposure to excess ^{16}O water over the course of a multi-step LC-MS/MS experiment (Sakai, 2005). A direct comparison between the unlabeled and fully labeled peaks would therefore ignore a relevant part of the data, namely that portion of the intermediary peak that is the contribution of the partially labeled species. Hence, conditions

Figure 6: Spectral data interpretation for relative quantification. A representative MS/MS co-fragmentation spectrum obtained for the BSA tryptic peptide HLVDEPQNLIK. Equivalent amounts of singly-charged ^{18}O -labeled and unlabeled species were simultaneously subjected to collision-induced dissociation on an ion trap instrument as they co-eluted during reverse-phase chromatography. The mass-to-charge ratios of the resulting *b*-ions (N-terminal fragments) are identical (uniformly unlabeled). However, the unlabeled and labeled *y*-ion product series (C-terminal fragments) display a characteristic separation of 4 *m/z* due to the difference in mass introduced by the dual ^{18}O labeling. The inset window enlarges upon a *y*-ion duplet. The portion of the intermediate peak due to partial labeling (i.e. single ^{18}O incorporation) that overlaps with the native ^{13}C isotopic envelope of the unlabeled *y*-ion is depicted as being clipped off and added *in silico* by the yRatios program to correct for the true relative abundance of the fully-labeled *y*-ion species. The vertical dashed lines represent other irrelevant isotopic peaks observed.

were chosen that are conducive to achieving complete forward labeling (see **Methods**; Fenselau, 2007; Yao, 2003; Kim, 2006) and a computational routine was developed to account and correct for any incomplete labeling prior to calculating protein ratios based on the y-ion peaks observed in low resolution (i.e. ion trap) tandem mass spectra.

Accordingly, the yRatios program accounts for partially-labeled peptides in two steps. Drawing upon the predicted peptide sequence, yRatios calculates a predicted peak intensity for the dual ^{13}C -isotopic shoulder peak based on the intensity of the corresponding monoisotopic peak using the natural distribution of isotopes (Rosman, 1998): (1) the ratio of the monoisotopic peak to this “dual ^{13}C atoms” peak is determined according to probability under an ideal isotopic distribution for the amino acid composition of the peptide fragment, (2) the observed monoisotopic intensity is assumed to be close to ideal, and so (3) the expected intensity of the “dual ^{13}C atoms” peak is calculated using the ideal ratio. The algorithm then deducts this calculated value from the actual observed peak intensity and reassigns the remaining intensity to the fully-labeled (i.e. two ^{18}O atoms) peak to mitigate possible systematic errors in determining a proper spectrum level y-ion ratio. A depiction of this correction procedure is included in **Figure 6**. This correction procedure is based on the assumption that the observed ion peak statistics reflect the overlap of native and partially-labeled isotopic species (that is, bearing only one ^{18}O atom). Higher resolution (>20,000 FWHM) instruments can potentially discriminate and circumvent this problem.

A ratio is then calculated for the co-fragmentation spectrum (database match or “peptide instance”) by summing all of the unlabeled y-ion series intensities and, separately, all of the fully ^{18}O -labeled y-ion intensities (adjusted as described above), then comparing the sums. Similarly, a ratio for the peptide itself is calculated by summing the unlabeled and labeled y-ion

intensities for all relevant instances obtained for that particular sequence. Finally, a ratio for the corresponding protein is calculated by adding up all the respective sets of unlabeled and labeled *y*-ion signals determined for all of its corresponding detected peptides. In this way, those spectra which contribute the most reliable information (i.e. highest intensities) are represented in a weighted manner in the final result. However, ratios calculated for individual spectra and their tallies are accessible as supporting evidence in the basic summary report to evaluate consistency among the individual observations.

The *yRatios* program also accounts for the potential interference of *b*-ions by excluding theoretically overlapped *y*-ions from the ratio calculations. Specifically, if the *b*-ion's monoisotopic mass falls within a range of $-1\ m/z$ to $+5\ m/z$ of the unlabeled *y*-ion's monoisotopic mass, the *y*-ion pair is excluded from all calculations. The range is intentionally broad in order to include the potential effect of single ^{13}C isotope of the *b*-ion as well possible variation between predicted and observed m/z values; however, the feature can be turned off through a preprocessor definition if very high resolution spectra are available. This exclusion was observed to occur for approximately 5% of all *y*-ion pairs detected in ion trap based spectra.

3.3 Verification of feasibility

3.3.1 Single protein assays

To verify the basic feasibility of the proposed method, labeled and unlabeled digests of bovine serum albumin (BSA) were created. Different relative volumes of the two digests were mixed prior to column loading to generate different ratios (1:5, 1:2, 1:1, 2:1, and 5:1), covering a range inclusive of the commonly accepted community thresholds for defining biological significance. Following data-dependent tandem mass spectrometry and subsequent database searching (with two search criteria used per LC-MS/MS run; one specifying a C-terminal mass

modifier of 4 Daltons to best match those centered on fully-labeled species and the other with no fixed static mass to best match spectra centered on unlabeled species), the ySelect program was executed to select those fragmentation spectra for which a peptide match confidence score of 99% or greater probability was found in either of the two searches. The yRatios program was then invoked to calculate a ratio for each spectrum on the list, and to summarize the relative abundance for both the individual peptides and for the intact BSA protein target. The results are depicted graphically in **Figure 7a**.

The same distribution was observed for each experiment, with an average of 78 spectra satisfying the identification confidence cutoff. The calculated ratios for individual spectra displayed moderate scatter about their mean, with the coefficient of variation ranging from 22% to 48% (average 34%). However, after combining all of the results for each peptide, this variation was markedly reduced (18% to 40%; average 28%). Moreover, after combining the results for all peptides detected for the protein (13.6 on average), an overall ratio close to the expected value was achieved in four of the five experiments (magnitude of difference 1% to 12%; average 8%). A single BSA protein level ratio outlier occurred, in which the calculated protein ratio was 35% greater than the expected value (observed, 0.27, vs. actual, 0.20). This may in part have been due to experimental error given the small volume of unlabeled BSA digest used (1 μ L). A second BSA 1: 5 LC-MS/MS experiment yielded a protein ratio of 0.22, much closer to the expected value.

Figure 7: Test protein quantification experiments. (a) Different amounts of unlabeled and labeled (^{18}O) tryptic digests of BSA were combined to generate five ratios. Each mixture was subjected to one-dimensional LC-MS/MS using an ion trap instrument. The calculated ratios for all five LC-MS/MS experiments are depicted, with black diamonds marking the relative ratios calculated for individual spectra (instances that passed a database search confidence cutoff of 99%), grey diamonds marking the ratio once the results for all spectra associated with a given peptide were tallied, and crosses marking the ratio determined once results for all the detected peptides mapped to the BSA protein were tallied. A boxplot is also displayed for each run's set of relative ratios for individual spectra. The expected ratio for a given run is shown as a dotted line. (b) A similar set of experiments, save that the BSA was spiked into a complex background mixture of *C. elegans* soluble proteins (already pre-digested in unlabeled HPLC-grade water) before performing LC-MS/MS in an otherwise identical manner.

3.3.2 Spiked protein assays

To verify that the results obtained with BSA generalize to more realistic conditions and are preserved in complex mixtures, different volumes of ^{18}O -labeled and unlabeled BSA digests (again, ratios 1:5, 1:2, 1:1, 2:1, and 5:1) were spiked together into a digest of *C. elegans* total soluble protein extract. A six-step LC-MS/MS screening experiment was performed to acquire multiple representative spectra to assess the performance of each combination.

The distribution of ratios obtained (**Figure 7b**) was highly similar to those observed for independent BSA runs, with the results for individual spectra displaying a considerable scatter about their collective mean (coefficient of variation ranging from 22% to 65% of the expected ratio; average 37%), but reduced scatter after combining results for each peptide (20% to 52%, average 32%) and with the BSA protein ratios converging to the expected value upon summation (difference 8% to 18%; average 12%). These data confirm that reasonably accurate quantifications can be achieved for a selected protein in a complex milieu, particularly with accumulating MS/MS evidence (i.e. when many peptides are identified for that protein).

3.3.3 Complex mixture assays

To establish the overall utility of the quantification method for the diversity of proteins present in a native biological specimen, ^{18}O -labeled and unlabeled digests of the same worm extract were created and analyzed by multi-step LC-MS/MS. As before, different volumes of the digests (ratios 1:5, 1:2, 1:1, 2:1, and 5:1) were mixed prior to column loading in order to form different proportions of labeled and unlabeled protein. Results of the ySelect and yRatios analyses on a set of high-confidence protein identifications are plotted in **Figure 8** (protein ratios only; peptide level data not shown).

Figure 8: Global proteome quantification experiments. Different amounts of unlabeled and labeled tryptic digests of *C. elegans* soluble whole worm extract were combined to form five ratios. Each mixture was loaded onto a multiphase column (strong cation exchange media packed between two layers of reverse phase material) and a six-step multidimensional LC-MS/MS elution gradient analysis subsequently performed. (a) Results from the five repeat runs for proteins in which only one high-confidence identifying peptide (for which there may be one or more qualifying spectra) satisfying a 99% likelihood cutoff was obtained. The number of proteins measured (n), the mean ratio (μ) and standard deviation of the calculation distribution (σ) are indicated for each experiment. The expected ratio is represented by a dotted line. A black diamond marks the ratio calculated for a particular protein. Boxplots are displayed for the distribution of calculated ratios. (b) Results from the five runs for proteins in which two or more identifying peptides passed the quality cutoff.

Again, a consistent pattern was noted for the overall distribution of the calculated ratios for individual proteins. The standard deviation of the estimated levels of those proteins for which only a single unique peptide was identified, whether via one MS/MS instance or with multiple (at least two) supporting spectra, was greater (coefficients of variation ranged from 26% to 42%; average 35%) than those for which two or more independent peptides were identified (coefficients of variation 17% to 26%; average 20%). While the means of the distributions were close to the expected values within a confined range (-2% to 19%; average 9%), the observed mean ratios appeared to be modestly biased in favor of the less abundant isotope when the expected ratios were of greatest difference (namely 1:5 and 5:1). Possibly, background noise (either electronic or chemical) obfuscates genuine peak intensities and naturally has a more significant effect on the weaker intensity peaks (Mason, 2007).

The act of merging two database searches that applied a different condition (a C-terminal modification of 4 Daltons versus no modification) introduces the potential for two different matches to be assigned to a given spectrum, both of which may pass the specified confidence level cutoff. Such conflicts were determined to be rare; over the five multi-step LC-MS/MS experiments performed with worm extract mentioned above, a total of only 13 disparate pairs of matches occurred. Of these, all save two involved residue sequences occurring in two different proteins that were indistinguishable by mass spectrometry at the resolution available (e.g. IINEPTAAAIAYGLDK and IINEPTAAALAYGLDK). The ySelect program responds to such cases, indistinguishable or not, by excluding the lower confidence matches from the generated list.

Although a stringent 99% confidence cutoff was initially specified for the ySelect runs, similar distributions were observed when specifying a less rigorous 95% peptide database match probability score (data not shown). An illustrative example: in one LC-MS/MS experiment in

which the expected ratio of unlabeled to labeled proteins was 1: 1, using a cutoff of 99%, the mean and standard deviation of overall relative ratios for proteins identified and quantified by only one peptide were 1.10 and 0.29 respectively, as compared to 1.09 and 0.37 when using a more liberal cutoff of 95%. For those proteins identified by two or more peptides, the values were 1.14 and 0.20, which are again comparable to the 1.14 and 0.23 obtained with a cutoff of 95%.

3.3.4 Assessing bias by precursor ion charge

An examination of the complex mixture data revealed that singly-charged peptide species displayed a consistent bias of the observed collective mean (**Supplementary Table S2**) relative to that observed for the multiply-charged precursors in the same experiment; the bias ranged from +5% to +60% (average +25%). Standard deviations were in some cases far larger for the singly-charged species as well, possibly due to a generally poorer fragmentation pattern relative to the multiply charged species (Dodds, 2006) and, in certain experiments, the small number of singly-charged results. The reason for the bias was not determined, but its impact can be averted by removing all singly-charged instances from consideration. Conceivably, the presence of the heavier isotope of the oxygen atoms alters the ionization characteristics of a labeled peptide relative to an unlabeled peptide, rendering it less likely to bear a single positive charge.

3.3.5 Accounting for residual variation

Averaging results from repeated experiments might statistically be expected to improve the precision for those proteins detected multiple times. Such an outcome was indeed witnessed

(**Figure 9**), but the improvements failed to build compellingly upon the beneficial effect of multiple-peptide identification over the simple one-peptide distributions (*c.f.* **Figure 8**). An examination of the combined results revealed that, as the number of experimental datasets in which a protein appeared increased, the chance that the identifications were also supported by multi-peptide evidence was likewise increased. For those proteins detected in all four experiments, virtually all were based on multi-peptide identifications in each run.

This led to the question of why averaging itself didn't notably improve precision. If individual peptides display a consistent bias of their own, averaging the results for single-peptide identifications could only be expected to increase the precision of the individual bias. Conversely, if no bias existed, then the calculated ratio for a given peptide could be expected to vary stochastically (randomly) around the median of the distribution; averaging should then tighten the entire distribution about the median. To test the null hypothesis that peptide ratios vary freely about the medians between runs, pairs of peptide results were considered; if each result varied at random, then in roughly half of the cases, pairs should be split about the common median. Chi-square goodness of fit tests indicated that the grouping on one side of the common median or the other for these pairs was far more frequent than is consonant with this null hypothesis (**Supplementary Table S3**). This implies that consistent individual peptide bias occurs.

Figure 9: Multiphase LC-MS/MS proteome surveys. Protein quantification results from four repeat large-scale multidimensional LC-MS/MS (i.e. MudPIT) profiling experiments with unlabeled and labeled tryptic digests of *C. elegans* soluble whole worm extract. Moving from left to right, the distributions representing ratios calculated for proteins identified in one, two, three or all four experiments are shown. The expected ratio of 2:1, unlabeled to labeled, is shown as a dotted line.

If the presence of this peptide bias (potentially leading to incorrect protein level ratios) is related to the label or labeling methodology, it can logically be compensated for by reversing the two sample labeling procedure in a second, parallel LC-MS/MS analysis, then integrating the ratios at the peptide level before combining the results to achieve a less biased estimate of true protein ratios. For the representative 1:2 and 2:1 ratio LC-MS/MS datasets depicted in **Figure 8**, 122 peptides were quantified in both experiments. When the ratios for the two runs were merged after inverting the calculated ratios (to simulate reversing the labeling) for the 1:2 run, the mean protein ratio was determined to be exactly 2.00; moreover, while the standard deviation was 0.58 among the concatenated peptide results, this deviation dropped to 0.37 when the pair of ratios for each peptide was first averaged.

3.4 Assigning statistical confidence levels to calculated ratios

To determine confidence levels for the protein quantification calculations, it was first necessary to ascribe a distribution model to the protein ratios. Normal probability plots (**Supplementary Figure S1**) for individual experiments suggested nearly normal distributions, albeit with long tails. However, these distributions consistently failed the Kolmogorov-Smirnov test for normality ($\alpha = 0.20$). Significantly, the distributions for the single-peptide ratios were similar across different experiments, with medians approximating the expected means and with comparable standard deviations. This was also true for the distributions of the multi-peptide ratios across different experiments. Indeed, Mann-Whitney tests failed to indicate a significant difference ($\alpha = 0.20$) between any two of the distributions drawn from the same category (i.e. single-peptide or multi-peptide) after normalization of their respective means. Accordingly, the results of the five experiments depicted in **Figure 8** were merged following normalization, allowing inverse cumulative distribution tables to be empirically assembled for the single-

peptide and multi-peptide ratio distributions (**Supplementary Table S4**), a necessary step towards determining critical values for confidence levels.

A two-fold or greater change and, more conservatively, five-fold or greater change in protein levels, are widely accepted (arbitrary) thresholds to indicate a (likely) biologically significant difference. For an individual ratio result calculated for a particular protein, it is useful to have an estimate of the confidence that the prediction actually does represent a true two-fold (or five-fold) change or greater. Given that the observed distributions are similar regardless of their actual means (at least within the range experimentally measured, R), if x represents an observed ratio and $P(x, y)$ represents the cumulative distribution function for x , with a mean y , then for x, y_1 and $y_2 \in R$ such that y_1 is less than y_2 , it must be true that $P(x, y_1)$ is greater than $P(x, y_2)$. Therefore, by choosing a critical value for a ratio such that, for instance, nineteen twentieths (95%) of that distribution with a mean of 2.0 falls at or below this value, any ratio observed to match or exceed this critical value must have less than a one in twenty (5%) chance of belonging to a distribution with a mean (i.e. actual ratio) less than 2.0. By referring to the inverse cumulative distribution tables, then, a table of critical values (ratio cutoffs) for achieving commonly accepted confidence levels for deeming at least two-fold and five-fold changes in protein abundance, respectively, could be derived (**Table 1**) for the single-peptide and multi-peptide results.

Table 1: Establishing significance in terms of calculated protein abundance. Empirically-derived ratio cutoffs for defining statistical-significant two-fold and five-fold differences in protein relative abundance, in cases where a reference protein of known fixed concentration was identified by either a single peptide or by more than one peptide.

Probability	Two-fold or greater change		Five-fold or greater change	
	Ratio determined from two or more peptides	Ratio determined from one peptide	Ratio determined from two or more peptides	Ratio determined from one peptide
> 0.80	> 2.25	> 2.58	> 5.63	> 6.45
> 0.90	> 2.56	> 2.88	> 6.40	> 7.20
> 0.95	> 2.70	> 3.35	> 6.75	> 8.38
> 0.99	> 3.13	> 4.34	> 7.83	> 10.9

Although critical values work well for individual results, the outliers stemming from a large majority of undifferentiated protein levels could potentially lead to a substantial false discovery rate. For multi-peptide protein ratios, only 4 normalized ratios out of 282 observed in the five merged experiments exceeded 1.6 and only 1 exceeded 2.0 (actual value 2.17). Since this largest outlier still falls below the most permissive critical value (2.25) listed in **Table 1**, this suggests that the false discovery rate is low (i.e. less than 1 in 282, or 0.3%) and potentially even lower as more stringent ratio cutoffs are employed. For protein abundances based on single-peptide ratios, 11 out of 615 normalized ratios (1.8%) exceeded the most permissive critical value 2.58, while only 7 out of 615 (1.1%) exceeded 2.88, and 4 out of 615 (0.65%) exceeded 3.35.

There was no evidence to indicate that spurious outliers would result in any false discoveries when applying the more conservative five-fold change critical values. However, significant systematic shifts (bias) of both the peptide and protein means toward unity were consistently observed for the 1:5 and 5:1 experiments (approximately 0.25 and 4.00 instead of the expected 0.20 and 5.00). In this sense, the critical values presented in **Table 1** for five-fold change might be considered overly conservative. Additional study will shed further light on this bias and potentially lead to a reliable formula that compensates for it.

3.5 Detection of questionable identifications

The additional series of y -ions present in the co-fragmentation spectra provide an opportunity to assess putative identifications and eliminate erroneous database matches (i.e. false positives) from the set of matches that have passed the confidence criteria. If a given identification is correct, one could reasonably expect each unlabeled y -ion should have a matching labeled y -ion (and vice versa), assuming the protein is present in both samples under

scrutiny. Furthermore, one could expect there would be some consistency between the relative ratios calculated for individual *y*-ion pairs. If a given identification is incorrect, then the putative *y*-ion ladders are false and so consistency would be unlikely.

An easily-implemented set of rules based on empirical optimization was developed to detect *y*-ion pair inconsistency. A list of potential *y*-ion masses is first generated from the putative peptide sequence. Peak intensity at the *m/z* calculated for each potential pair of unlabeled and labeled *y*-ions (including the partial-labeling adjustment) is compared and a tally is kept. A pair is deemed “wild” if both peaks are present but one exceeds the other by an arbitrary factor of two beyond the calculated ratio for the peptide instance. Alternatively, a pair is considered “non-pair” if one peak is found but the other is absent. The ratio calculated for a peptide instance is considered to be suspect and hence is “alarm flagged” if half or more of the *y*-ion pairs are deemed wild pairs or if there are fewer pairs than non-pairs. Additionally, any peptide instance in which fewer than five pairs (wild or not) are present is deemed suspect.

An examination of the complete set of worm protein LC-MS/MS profiling experiments (using a 99% confidence cutoff) revealed that virtually all (167 out of 179 spectra) of the reversed decoy sequence matches were subsequently alarm flagged (results for individual experiments ranged from 86% to 100%; average 93%). Nearly all of the decoy sequence matches (156 out of 179 spectra) had a precursor ion charge of +3, consistent with the lower interpretability of such highly-charged peptides, but none were duplicated in the other four runs, suggesting that these bogus matches occurred purely by chance.

Performance for extreme ratios was not tested. It is anticipated that alarm flagging would become prevalent at ratios above 10 (or below 0.1) (non-flagged ratios outside these bounds were virtually not in evidence), and inevitable in those cases where a protein is present in one sample and completely absent from the other.

An average of 21% of identified worm proteins had ratios generated only from alarm flagged peptide instances (this value ranged from 15% to 32% across individual LC-MS/MS runs). It should be noted that the alarm-flagged ratios are not necessarily wrong. However, they are suspect, in that they are both significantly more likely to be false positives (as indicated by the alarm flagged reverse database hits) and evidence variances about the expected mean that are over twice those of non-flagged results (see **Supplementary Figure S2**). It is left to the user of the yRatios program to decide whether and where to make use of alarm-flagged ratios.

3.6 Other considerations

3.6.1 Deviation of the isotopic distribution pattern from the theoretical

The correlation between the theoretical isotope distribution for a *y*-ion and the experimentally-observed distribution isn't necessarily perfect. Between low resolution and the centroiding algorithm, the mass spectrometry software could only form estimates of peak *m/z* and intensity.

A simple LC-MS/MS experiment using unlabeled worm digest at an isolation width of 2 *m/z* revealed that the peak intensity of the second isotopic peak (monoisotopic + 2 *m/z*) varied about the theoretical value, with roughly half of the data points falling to either side. When the same experiment was performed at an isolation width of 10 *m/z*, nearly all observed intensities exceeded the theoretical value, with the mean of the distribution at 0.67 times above theoretical.

Although it might be expected that this would introduce a consistent bias in favor of the unlabeled tallies, the effect is mitigated by the fact that the lower mass *y*-ions exhibit non-monoisotopic peaks that are small in intensity relative to the monoisotopic peak. The relative intensities of the non-monoisotopic peaks increase as the mass of the *y*-ion increases, but the

intensities of the higher mass y -ions themselves (1100 m/z and greater) typically decrease sharply relative to the lower mass y -ions (data not shown).

3.6.2 Ratios calculated additively versus ratios calculated by averaging pairs

Although the approach used here for calculating a ratio for a given spectrum has been to divide the sum of unlabeled y -ion intensities by the sum of (adjusted) labeled y -ion intensities, other approaches are also possible. The most obvious of these might be to average the individual y -ion pairs, or expand the average of their logarithms to the power of 10. A comparison of these approaches (**Table 2**) indicates that the sum-of-intensities approach exhibits less bias and less variation than the other two.

Since smaller intensity values might be more subject to a constant level of noise, a formula in which only the top three pairs (ranked by the intensity of the unlabeled y -ion) contributed to the intensities sums was tested. Although the means of the distributions were comparable to the full-series-sum approach, variation was substantially greater (data not shown).

3.6.3 Assessing intra-spectra bias by mass over charge

In calculating ratios for individual spectra, each y -ion pair contributes to the unlabeled and labeled intensity tallies regardless of a pair's own particular ratio. Conceivably, the individual pairs' ratios taken over a particular m/z range may evidence a greater variation and a bias relative to the overall results. These would affect accuracy and precision.

Table 2: Ratios calculated additively versus ratios calculated by averaging pairs. Three methods of calculating a ratio for a given spectrum were compared. The first method involved adding the intensities of the labeled (following correction) and unlabeled y-ions separately, then dividing the unlabeled tally by the labeled tally. The second method involved calculating the logarithm of the ratio for each y-ion pair (labeled and unlabeled), averaging the logarithms, and then expanding the average as a power of ten. The third simply involved averaging the individual pairs. Over five separate multi-step LC-MS/MS runs, in which different amounts of unlabeled and labeled *C. elegans* soluble proteins were combined in fixed proportions (1: 5, 1: 2, 1: 1, 2: 1 and 5: 1 ratios) prior to column loading, the coefficients of variation were uniformly lower for the first method and the means of the ratio distributions were, save for one exception for each of the other methods, closer to the expected value.

Expected Ratio for LC-MS/MS run	1. Peptide instance ratios calculated by ratio of the unlabeled and labeled tallies		2. Peptide instance ratios calculated by averaging the logs of individual pair ratios		3. Peptide instance ratios calculated by averaging individual pair ratios	
	Mean of calculated ratios	Coeff. of variation	Mean of calculated ratios	Coeff. of variation	Mean of calculated ratios	Coeff. of variation
0.2	0.23	0.41	0.25	0.44	0.36	0.70
0.5	0.52	0.36	0.52	0.37	0.61	0.49
1.0	1.12	0.37	1.06	0.38	1.27	0.43
2.0	1.90	0.35	1.82	0.39	2.31	0.43
5.0	3.63	0.48	3.25	0.49	4.40	0.56

The ratios of individual pairs were normalized to the overall ratio for the spectrum in which they appeared and divided according to which bin (20 m/z intervals of 100 units, from 0 m/z to 2000 m/z) they fell within. Results were pooled across many spectra and the average of the logarithms of the ratios falling within each bin was calculated, along with a standard deviation (**Table 3**). Little bias was observed across the bins. The pair ratios of the upper ends of spectra displayed a higher variance against the overall ratios; this can be explained by the greater average intensity of the mid-range y -ion pairs, which would tend to dominate the overall ratio and so bring their pair ratio closer on average.

3.6.4 Assessing inter-spectra bias by y -ion series intensities

Different spectra are observed to have different overall levels of peak intensities, depending at least in part on the intensity of the precursor ion selected for fragmentation. This is reflected in the peak intensities of the unlabeled and labeled y -ion series as well. The sum of the unlabeled and labeled tallies varies enormously from spectrum to spectrum, ranging from as small as 100 counts to approaching 1,000,000 counts.

Given that the smaller sums-of-tallies contain less information, ratios calculated using them might potentially vary more widely than those calculated from spectra exhibiting larger sums-of-tallies. Additionally, they may be more subject to the additive effects of noise (spurious peaks or spurious additions to peaks).

Table 3: Assessing in-spectra bias by mass over charge. For each test spectrum, the ratios of individual y-ion pairs were normalized to the sum-of-series overall ratio and divided according to which predefined bin (20 intervals of 100 m/z ranging from 0 m/z to 2000 m/z) each unlabeled y-ion's m/z fell within. Results were accumulated across 2035 spectra generated from five different multi-step LC-MS/MS experiments and the average of the logarithms of the ratios allocated to each bin was calculated along with the standard deviation. Little bias is evident in the averages; all fall close to the expected value of 1.0.

m/z range	Number of y-ion pairs	10 expanded to the average of the logarithms of the normalized ratios	Standard deviation of the logarithms of the normalized ratios
100 – 200 m/z	7	1.131	0.206
200 – 300 m/z	510	0.938	0.322
300 – 400 m/z	1542	1.054	0.329
400 – 500 m/z	1861	1.021	0.324
500 – 600 m/z	2327	0.967	0.306
600 – 700 m/z	2374	0.992	0.296
700 – 800 m/z	2377	0.949	0.271
800 – 900 m/z	2309	0.938	0.274
900 – 1000 m/z	2434	0.953	0.273
1000 – 1100 m/z	2558	0.972	0.275
1100 – 1200 m/z	2420	0.962	0.282
1200 – 1300 m/z	2114	0.969	0.281
1300 – 1400 m/z	2026	0.956	0.343
1400 – 1500 m/z	1845	0.941	0.317
1500 – 1600 m/z	1776	1.014	0.350
1600 – 1700 m/z	1510	0.963	0.382
1700 – 1800 m/z	1076	0.985	0.393
1800 – 1900 m/z	836	0.890	0.411
1900 – 2000 m/z	664	1.009	0.414

The magnitudes of these effects were examined across five multi-step LC-MS/MS experiments to assess the relative reliability of smaller sums-of-tallies results (**Table 4**). The average of the logarithms of calculated ratios and the standard deviations of their logarithms were determined for ranges of sums-of-tallies within each experiment. A correlation in favor of the larger sums-of-tallies, between the magnitude of sum-of-tallies and both accuracy and precision of the calculated ratios, is evident. For expected ratios both greater and smaller than unity, the ratios for the smaller sum-of-tallies spectra averaged closer to unity than those for the larger sum-of-tallies spectra. Disparity between averages largely disappeared with an expected ratio of 1.0. This relative behavior of the ratios averages lends strength to the hypothesis that noise is responsible; if evenly distributed, it would have a greater relative impact on the smaller y-ion tally than the larger one, tending to drive the ratio toward unity, while if the tallies were approximately the same, the addition of noise would be reflected in both tallies more or less equally.

3.6.5 Evaluation of back-exchange

During the chromatography, both labeled and unlabeled peptides were bound to the column for potentially many hours, awaiting the conditions that would cause them to elute. The chromatography buffers were prepared with HPLC water (very low ^{18}O content) made acidic by its glacial acetic acid and n-heptafluorobutyric acid content. It is conceivable that, under these conditions, acid-mediated back-exchange of the ^{18}O label to the ^{16}O isotope predominant in the HPLC water might occur.

Table 4: Assessing bias in y-ion series intensities. Spectra recorded over five two-dimensional LC-MS/MS profiling experiments for five different ratios of labeled and unlabeled soluble *C. elegans* protein digests were ranked by the sum-of-intensities (namely the sum of the all y-ion intensities per spectrum, both labeled and unlabeled). The averages of the logarithms of the calculated ratios for spectra that fell within a given range of sum-of-intensities were determined along with the standard deviation of the logarithms. The ratios calculated for spectra displaying smaller sum-of-intensities values were closer to unity than those for the spectra displaying larger sum-of-intensities values. The former also evidenced greater variance. Both effects could be attributable to any stochastic noise present in the fragmentation scan; random additions would have a greater relative impact on spectra with lower overall intensities.

Expected Ratio	Sum-of-tallies range	Average of the logarithms of the calculated ratios	Standard deviation of the logarithms of the calculated ratios
0.2	$10^{3.0}$ to $10^{3.5}$	-0.594	0.167
	$10^{3.5}$ to $10^{4.0}$	-0.647	0.148
	$10^{4.0}$ to $10^{4.5}$	-0.669	0.133
	$10^{4.5}$ +	-0.649	0.119
0.5	$10^{3.0}$ to $10^{3.5}$	-0.273	0.174
	$10^{3.5}$ to $10^{4.0}$	-0.280	0.138
	$10^{4.0}$ to $10^{4.5}$	-0.300	0.097
	$10^{4.5}$ +	-0.320	0.069
1.0	$10^{3.0}$ to $10^{3.5}$	0.055	0.136
	$10^{3.5}$ to $10^{4.0}$	0.051	0.141
	$10^{4.0}$ to $10^{4.5}$	0.081	0.091
	$10^{4.5}$ +	0.060	0.152
2.0	$10^{3.0}$ to $10^{3.5}$	0.246	0.165
	$10^{3.5}$ to $10^{4.0}$	0.250	0.134
	$10^{4.0}$ to $10^{4.5}$	0.266	0.111
	$10^{4.5}$ +	0.318	0.114
5.0	$10^{3.0}$ to $10^{3.5}$	0.460	0.194
	$10^{3.5}$ to $10^{4.0}$	0.542	0.154
	$10^{4.0}$ to $10^{4.5}$	0.587	0.146
	$10^{4.5}$ +	0.666	0.107

An examination of the average of individual (spectral) ratios over different steps of five multi-step LC-MS/MS experiment is presented in **Table 5**. If a significant degree of back-exchange had occurred during these experiments, the averages should have increased as time progressed (since fully and partially labeled peptide would be converted to unlabeled peptide). The average proportion of partially labeled peptide relative to fully labeled might be expected to change as well. Neither of these effects was observed, indicating that significant back-exchange during the chromatography did not occur (see also **Supplementary Figure S3**).

3.6.6 Outlier elimination

Although not exercised during the course of the presented data analysis, yRatios offers an option for the elimination of outliers at the peptide instance (individual spectral result) level using inter-quartile ranges (IQRs) and fencing. Application of this option does not appear to substantially affect the computed ratios. For example, in one LC-MS/MS experiment, 296 of 1264 peptide instances (including those alarm-flagged but excluding matches to reversed decoy sequences) fell into populations large enough to apply outlier elimination through fencing (6 or more). Of these, 12 (divided among 11 peptides) were marked for elimination (6 already alarm-flagged). Eliminating the outliers changed the overall ratios for the 11 peptides by an average of only 10%.

Table 5: Ratio averages over a chromatography time course. An average ratio for all spectra recorded in a given step of the experimental chromatography was calculated for all steps of five two-dimensional LC-MS/MS profiling experiments for five different ratios of labeled and unlabeled soluble *C. elegans* protein digests. Each average was normalized against that calculated for the associated third step. The average proportion of the labeled tallies (denominators in individual ratio calculations) that derived from the partially labeled calculation for given steps is also displayed. Since the ratios did not increase over time as would be expected given a progressive loss of the ^{18}O label on peptides due, back-exchange appears unlikely. The average contribution of the partially labeled calculation stayed approximately constant over time as well, and the relative ratio of partially labeled peptide and fully labeled peptide was maintained.

Expected ratio	Step number	Ratio normalized to the third step	Proportion of the partially labeled contribution of the labeled tally
0.2	2	0.89	0.23
	3	1.00	0.26
	4	0.96	0.23
	5	0.90	0.19
	6	0.91	0.21
0.5	2	0.97	0.18
	3	1.00	0.22
	4	1.04	0.23
	5	1.06	0.25
	6	0.99	0.22
1.0	3	1.00	0.29
	4	0.95	0.28
	5	0.92	0.26
	6	0.87	0.27
2.0	2	1.11	0.24
	3	1.00	0.24
	4	0.96	0.25
	5	0.91	0.28
	6	0.85	0.28
5.0	2	1.06	0.29
	3	1.00	0.28
	4	1.03	0.28
	5	0.97	0.31
	6	0.82	0.32

4. Discussion

4.1 General methodology

To be applicable as the basis for systematic global expression profiling, a quantification method must provide a sufficiently accurate measure of relative protein abundance so as to reveal biologically-relevant patterns of differential expression. In this context, it is better to have confidence that one is approximately correct, rather than being uncertain whether one is precisely wrong. Although the means of distributions of calculated ratios were found to fall reasonably close to the expected values and variance to be consistent within useful limits, to this end, an estimate of statistical confidence in a given set of ^{18}O -based quantification results has additionally been provided based on the empirical data distribution, with the observed ratios presumably affected by either systematic or stochastic technical noise. However, the experimenter must still account for biological variation when considering appropriate experimental design parameters and final interpretation when contemplating a particular biological research problem.

The initial approach taken for calculating peptide levels was to determine the ratios for all individual y-ion pairs for a given spectrum and then averaging them together to achieve a ratio result for the peptide instance. However, an additive approach was deemed preferable in that the final ratio extracted reflects the pairs with the greatest intensities that are presumably least influenced by noise. This should eliminate the tendency of spurious low intensity pairs, which typically vary more widely, from unduly influencing the projected values. Even so, individual pair averaging still plays a role, namely as a means of testing the “sanity” of a particular result; if too many pairs deviate substantively (i.e. by more than a factor of two) from the calculated ratio for a given peptide, the result is flagged as being suspect.

Other avenues for improving identification rates are available. The ySelect program has been implemented separately from yRatios to ease integration with other sources of peptide identifications. A search program such as ProbiDtree (Zhang, 2005) that offers the capability to identify multiple peptides from a single fragmentation spectrum might compensate for the effect of a wider isolation width or even exceed the performance observed for a narrow isolation width (although care to take note of such cases where only the unlabeled or labeled form of one of the peptides was co-fragmented, or where quantification might be questionable due to difference in intensity between the peptides, would be necessary; overlap of the different *b* and *y* ion series would need to be accounted for as well). The development or enhancement of a search program to make use of *y*-ion duplets to reinforce identification confidence would be beneficial so long as the duplets did not become a prerequisite (a protein might be present in only the labeled or unlabeled sample). So too might be a mass-tag based approach; peaks detected in a full scan displaying a separation that can be ascribed to the labeling along with a difference in relative intensity greater than a preset threshold could preferentially be chosen for fragmentation, thereby enriching the identification and quantification process for those peptides that are potentially of most interest (i.e. display the greatest change).

4.2 Correction for partial labeling

To the best of my knowledge, this represents the first application of corrective abundance level calculations based on partly redundant fragment ion information. Among those groups who performed relative quantification at the MS/MS level, Heller *et al.* (2003) did not apply a correction for partial ^{18}O labeling in their calculated ratios, while Zhang and Neubert (2006) did quantification in the presumption of full metabolic labeling and so did not need to.

Elbert *et al.* (2008) similarly demonstrated the utility of quantification at the MS/MS level, with attention to measurement of the kinetics of protein turnover. However, their SILT method, while valuable, requires a second MS/MS scan and currently assumes that precursor ions are doubly charged. It is not fully suited to quantification through ^{18}O labeling in that the algorithm looks for a labeled amino acid residue rather than for C-terminal labeling of peptides (potentially confusing if a Lysine or Arginine residue appears more than once in the sequence), and only one labeled residue is considered at a time. Here, the usability of a quantification-through-MS/MS foundation is advanced by (1) applying the more general ^{18}O labeling approach, (2) characterizing ratio distributions resulting from multi-step LC-MS/MS experiments and involving hundreds of proteins to empirically produce confidence level cutoffs for determining statistical significance of results, and (3) defining a set of rules that operate upon the redundant series of y -ions present in the MS/MS spectra to flag potentially erroneous results and detect putative false positive database matches.

4.3 Dynamic range

The tests indicate that the dynamic range over which the means of the distributions of the calculated ratios matched the expected ratios is approximately one order of magnitude, with a significant bias occurring by ratios greater than 5:1 and 1:5. This suggests a practical limit for estimating fold-change conclusively. While Zhang and Neubert (2006) reported accurate quantification at the MS/MS level for fold-change ratios as low as 1:100, it was not clear what quantities of proteins were employed. Although it is plausible that high intensity fragmentation spectra could routinely yield such values, many spectra collected during the course of LC-MS/MS experiments, as applied to complex samples, are often composed of low intensity peaks yet still provide high confidence matches to peptides. For instance, in one representative LC-

MS/MS experiment, 306 out of the 1037 peptide instances matched by the database search algorithm had labeled tallies (sums of intensities) of less than 500. Such spectra may simply not provide sufficient contrast to derive large ratios reliably.

It is important to distinguish between in-spectra dynamic range, which determines the practical range of ratios which may be calculated, and the dynamic range of overall spectral intensities. The latter is very broad, approaching three orders of magnitude. Any spectrum that is successfully identified, regardless of overall spectral intensity, has the potential to yield a practical calculated ratio.

The extreme case in relative quantification is the complete absence of the protein in one sample and its strong presence in the other. The fragmentation spectra in such situations might resemble those of false positives, in that the y-ions will not be consistently paired. They are therefore susceptible to alarm flagging. The yRatios program cannot distinguish between the two scenarios (but does provide an option to list only those peptide instance, peptide or protein ratios that exceed or fall below user-determined thresholds, both alarm-flagged and not, to ease the process of verification). Some vigilance is therefore required on the part of the experimenter when evaluating results. Repeated appearance across different experiments may provide one means of distinguishing such results from false positives. Reverse labeling should also be of use in confirming differential expression.

If greater accuracy is important to validate a high or low ratio, it is possible to scale the experiment accordingly so as to bring an apparent ratio into the useful range of the assay. For instance, a reported ratio of 7.2 for a given protein could be verified by mixing one part unlabeled digest to five parts labeled; if accurate, the reported ratio in the new mix should be about 1.4. Additionally, if improved initial technical accuracy is desirable, then combining the

results of one experiment in which a first sample is labeled and the other is not with a second run in which the labeling has been reversed appears to be a plausible course of action.

4.4 Alarm flagging

This work represents the first application of the additional fragmentation information provided by ^{18}O labeling as a means of systematically minimizing the false discovery rate in large-scale profiling experiments. The rules developed for detecting suspect results are admittedly simplistic but appear to work reasonably well. It is likely that they could be improved upon, perhaps through the application of a more sophisticated machine learning approach to determining optimal informative parameters. As an initial step toward improvement, yRatios includes an option to supplement the binary alarm-flagging with a simple “fidelity” score and a specified threshold of acceptance. This scoring system is intended to serve as a starting point for evaluation or for investigators who wish to enhance the yRatios source code.

However, there remains a question of how many additional false positives are introduced by the co-fragmentation of both labeled and unlabeled peptides. Unless the peptide mass tolerance for the database search is made sufficiently wide (i.e. offset by over 4 Daltons from the native peptide mass), a labeled peptide cannot be matched to its correct sequence if a null C-terminal enhancement is specified. Only incorrect assignments (false positives) can be made. Similarly, when performing the database search with a 4 Dalton C-terminal enhancement specified, an unlabeled peptide cannot be matched to its correct sequence. By chance, a portion of either sort of these incorrect assignments will appear sufficiently compelling to pass the identification confidence cutoff. Simply widening the peptide mass tolerance to include both unlabeled and labeled peptides introduces its own added risk of false positive identifications and

extends the search time, and in any case cannot be counted on to identify those cases in which only the labeled form of the peptide is present.

4.5 Software options

An option to exclude singly-charged species is provided in the ySelect program. However, the decision to exclude is not a clear one. Inclusion of these instances did not have a large effect on the overall protein level results as the proportion of singly-charged instances to total spectra was relatively small (3% to 19%; average 9%). Exclusion of singly-charged data meant a direct loss in supporting evidence and a reduced number of proteins quantified (0.6% to 8% decrease observed; average -4%). Although experiments indicate that implementing a more selective precursor isolation window of 6 m/z produces an increase in proteins confidently identified despite the exclusion of singly-charged species (**Supplementary Table S1**), making the decision to exclude a seemingly straightforward one, it should be noted that these results were collected under the conditions of one-dimensional chromatography. While chemical noise might be expected to contribute equally toward reduced spectral quality regardless of whether a second degree of separation is added, the probability of unrelated peptides undergoing co-fragmentation and thereby confounding the search program should be substantially lower. The fractional identification rate relative to that when employing an isolation width of 2 m/z might significantly improve for both of the wider isolation widths tested when performing two-dimensional LC-MS/MS experiments, possibly more for an isolation width of 10 m/z than for 6 m/z .

The yRatios program provides outlier elimination functionality at the peptide instance (spectral) level only. Although in principle it could be performed at the peptide or protein levels as well, given that each data point at these levels in many cases represent a combination of

several data points or more, providing justification for such a decision might be difficult. At the peptide level, combining an experiment with a second one in which the labeling has been reversed might better correct for peptide bias. At the protein level, there is the option to combine experimental replicates using ySelect and so eliminate outliers from the merged sets of peptide instances.

The yRatios program has been designed for simplicity of use. Both readable and comma-separated value reports presenting different levels of information are available, allowing the investigator to select the degree of detail desired toward understanding how the calculated ratios have been derived. Analysis within GoMiner and Cytoscape with the BiNGO plug-in is merely a matter of loading a yRatios-generated file at the appropriate point during initialization for use.

4.6 Conclusion

In summary, the results presented here demonstrate conclusively that ^{18}O -labelling coupled with systematic interpretation and integration of individual MS/MS spectra using dedicated software tools can form the basis of a routine global proteome quantification method that provides accuracy within a biologically intuitive overall dynamic range. Although possibly not as refined as SILAC or iTRAQ-based quantitative procedures, this platform works reasonably well when comparing two complex samples subjected to standard single 1D or multiphase nanoflow LC-MS/MS separations. This approach offers, at the very least, enhanced quantitative precision as compared to simple spectral counting-based quantification procedures for those proteins identified by a very few (one, even) high confidence spectral matches, while potentially enhancing the reliability of protein identifications. At the same time, it is more practical to implement than more sophisticated chemistry-based labeling techniques or protein-level metabolic labeling. Since complementary information for calculating protein ratios is

extracted from all of the spectra assigned to every protein identified, convergent approximations can be made with suitable confidence limits. No complex algorithms need be employed to disentangle the contributions of multiply-charged or partially-labeled precursor ion species at the full scan level, which is particularly challenging for spectra recorded on mass spectrometers with limited resolution, and meaningful results can even be extracted for low intensity precursor ion peaks that are otherwise near or at the level of background noise.

The major part of this work was accepted by the Journal of Proteome Research and published in July, 2009 (White, 2009).

5. Future Work

5.1 Experimental

New mass spectrometers such as the Thermo Scientific Orbitrap offer various performance improvements. This creates opportunities to enhance the quality of the quantitative results produced by this method that are open for investigation. Firstly, electron transfer dissociation (ETD) as a means of fragmentation may prove complementary or even superior to collision-induced dissociation (CID) for the purpose of performing quantification at the MS/MS level for triply-charged peptide species, since the charge state is reduced by the electron transfer. Additionally, post-translation modifications are far less labile under ETD (Mikesh, 2006); success in the identification of, and, by extension, the relative quantification of, peptides bearing such modifications should be increased. Secondly, if quantification is to be performed in a targeted manner (e.g. if specific peptide mass-to-charge values have been pre-selected and the mass spectrometer programmed to watch for them at a given time during the chromatography), perhaps to reinforce certain results collected in a global survey, the ability to perform high resolution MS/MS might improve quantitative accuracy and precision, since the peak representing those unlabeled y -ions bearing two ^{13}C atoms would be sufficiently separated from the peak representing partially-labeled y -ions that the intensity of the partially-labeled peak to eliminate the need for estimating contributions to a combined peak. Thirdly, high mass accuracy and high resolution at the full scan level may lead to improved identification confidence, thereby countering the general reduction in identification confidence induced by increasing isolation width. Fourthly, a faster scan rate, allowing more fragmentation spectra to be recorded each second, should improve both the number of proteins quantified and the accuracy of quantification (more peptides detected, more spectra per peptide).

Reverse labeling, in which LC-MS/MS is performed twice for the same pair of samples (one time in which the labeling has been applied to one sample and one time in which the labeling has been applied to the other) has the potential to correct for individual peptide bias and help to confirm protein absence in one sample or the other. Additional experiments to assess the extent of improvement in technical accuracy should be performed. If a significant improvement is observed, yRatios should be modified to accept and combine two sets of experimental results (one in which labeling is reversed) and produce a list of confirmed high-ratio and absent proteins.

The potential exists to apply spectral counting and quantification at the full scan level in a manner complementary to this method, as this information is always present in the LC-MS/MS experiments. The three sources of quantitative estimates could be used to reinforce each other, applying one method in those instances where the others are borderline or inconclusive. For example, spectral counting could be used to provide an estimate of relative quantities for singly-charged ions if the investigator chooses to restrict the isolation width to 6 m/z . A detailed evaluation at different levels, comparing our approach to spectral counting and ^{18}O -based full-scan (MS) quantification for both one-dimensional and multi-dimensional LC-MS/MS profiles, would help guide such complementary use and also assist investigators in selecting an approach best suited for their experimental plans.

5.2 Software Enhancements

Although the basic functionality of ySelect and yRatios is complete, a number of useful enhancements are possible.

The ySelect program was deliberately written separately from the yRatios program, so that different versions could be implemented to interpret the output of different search engines. A next target for such integration might be SIMS (Liu, 2008).

Currently, yRatios must be altered manually to recognize particular post-translational modifications. A capability to accept further symbols in a peptide sequence and understand them to refer to particular post-translational modifications (and artificial modifications) should be implemented.

When comparing two samples that are biologically related, such as proteins extracted from strain of yeast cultured in a galactose-rich medium and from the same strain of yeast cultured in a glucose-rich medium, in many cases it might be reasonably expected that most proteins will not differ significantly in their level of expression. The yRatios program could be enhanced with an option to specify that this assumption is being made and that calculated ratios should be normalized against the median of the ratio distribution, prior to reporting results, in order to minimize technical variation.

The yRatios program is amenable to extension with an option to generate the basic reports not only in simple text format but also in Adobe PDF format. Additionally, the possibility of building a platform-independent graphical user interface around the existing ySelect and yRatios functionality might be worthy of consideration.

Although the software is able to distinguish between unique and non-unique peptides in its reports (details may be found in the User Guides), no recognition is given to the possibility of unique peptides that produce fragmentation patterns which appear identical within the limits of the mass spectrometer (for example, the hypothetical peptides IAAVAEK and LA AVAEK). This shortcoming will be addressed in a future release.

While the overall rate of false positives, as indicated by the rate of reverse decoy sequence matches, outperformed the specified 99% confidence level, results for triply-charged spectra suffered from a particularly high rate (**Supplementary Table S5**) even after alarm-flagged instances were discarded (19% before, 3% after). A separate scoring system tailored to triply-charged peptide instances and aimed at excluding the residual false positives while reducing the overall proportion that are alarm-flagged is being investigated.

In addition to exploring a more sophisticated alarm-flagging methodology, the outlier elimination functionality might be extended to include alarm-flagged instances, thereby retaining the input of the more consistent flagged ratios.

6. References

- Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**. 422, 198.
- Allison, D.B.; Cui, X.; Page, G.P.; Sabripour, M. Microarray data analysis: from disarray to consolidation and consensus. *Nat. Rev. Genet.* **2006**. 7, 55.
- America, A.H.; Cordwener, J.H. Comparative LC-MS: a landscape of peaks and valleys. *Proteomics* **2008**. 8, 731.
- An, Y.; Fu, Z.; Gutierrez, P.; Fenselau, C. Solution Isoelectric Focusing for Peptide Analysis: Comparative Investigation of an Insoluble Nuclear Protein Fraction. *J. Proteome Res.* **2005**. 4, 2126.
- Ang, C. S.; Veith, P. D.; Dashper, S. G.; Reynolds, E. C. Application of $^{16}\text{O}/^{18}\text{O}$ reverse proteolytic labeling to determine the effect of biofilm culture on the cell envelope proteome of *Porphyromonas gingivalis* W50. *Proteomics*. **2008**. 8, 1645.
- Arnott, D.; Shabanowitz, J.; Hunt, D.F Mass Spectrometry of Proteins and Peptides: Sensitive and Accurate Mass Measurement and Sequence Analysis. *Clin. Chem.* **1993**. 39, 2005.
- Bantscheff, M.; Schirle, M.; Sweetman, G.; Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal. Biolanal. Chem.* **2007**. 389, 1071.
- Bondarenko, P.V.; Chelius, D.; Shaler, T.A. Identification and Relative Quantification of Protein Mixtures by Enzymatic Digestion Followed by Capillary Reversed-Phase Liquid Chromatography-Tandem Mass Spectrometry. *Anal. Chem.* **2002**. 74, 4741.

Brunner E.; Ahrens C.H.; Mohanty S.; Baetschmann H.; Loevenich S.; Potthast F.; Deutsch E.W.; Panse C.; de Lichtenberg U.; Rinner O.; Lee H.; Pedrioli P.G.; Malmstrom J.; Koehler K.; Schrimpf S.; Krijgsveld J.; Kregenow F.; Heck A.J.; Hafen E.; Schlapbach R.; Aebersold R. A high-quality catalog of the *Drosophila melanogaster* proteome. *Nat Biotechnol.* **2007.** 25, 576..

Cagney, G.; Emili, A. De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nat. Biotechnol.* **2002.** 20, 163.

Cagney, G.; Park, S.; Chung, C.; Tong, B.; O'Dushlaine, C.; Shields, D.C.; Emili, A. Human tissue profiling with multidimensional protein identification technology. *J. Proteome Res.* **2005.** 4, 1757.

Chen, M.; Ying, W.; Song, Y.; Liu, X.; Yang, B.; Wu, S.; Jiang, Y.; Cai, Y.; He, F.; Qian, X. *Proteomics* Analysis of human liver proteome using replicate shotgun strategy. **2007.** 7, 2479.

Chen, X.; Cushman, S. W.; Pannell, L. K.; Hess, S. Quantitative Proteomic Analysis of the Secretory Proteins from Rat Adipose Cells Using a 2D Liquid Chromatography-MS/MS Approach. *J. Proteome Res.* **2005.** 4, 570.

Colzani, M.; Schütz, F.; Potts, A.; Waridel, P.; Quadroni, M. Relative protein quantification by isobaric SILAC with ammonium ion splitting (ISIS). *Mol. Cell. Proteomics.* **2008.** 7, 927.

de Godoy, L.M.; Olsen, J.V.; Cox, J.; Nielsen, M.L.; Hubner, N.L.; Fröhlich, F.; Walther, T.C.; Mann, M. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*. **2008**. 455, 1251.

Diamond, D. L.; Jacobs, J. M.; Paeper, B.; Prohl, S. C.; Gritsenko, M. A.; Carithers, R. L., Jr.; Larson, A. M.; Yeh, M. M.; Camp, D. G., III; Smith, R. D.; Katze, M. G. Proteomic profiling of human liver biopsies: Hepatitis C virus-induced fibrosis and mitochondrial dysfunction. *Hepatology* **2007**. 46, 649.

Dodds, E.D.; Hagerman, P.J.; Labrilla, C.B. Fragmentation of singly protonated peptides via a combination of infrared and collisional activation. **2006**. *Anal. Chem.* 78, 8506.

Elbert, D.L.; Mawuenyega, K.G.; Scott, E.A.; Wildsmith, K.R.; Bateman, R.J. Stable Isotope Labeling Tandem Mass Spectrometry (SILT): Integration with Peptide Identification and Extension to Data-Dependent Scans. *J. Proteome Res.* **2008**. 7, 4546.

Elliot, M.H.; Smith, D.S.; Parker, C.E.; Borchers, C. Current trends in quantitative proteomics. *J. Mass Spectrom.* **2009**. 44, 1637.

Eng, J.K.; McCormack, A.L.; Yates, J.R. III An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *J. Am. Soc. Mass Spectrom.* **1994**. 5, 976.

Fenn, J.B.; Mann, M.; Meng, C.K.; Wong, S.F.; Whitehouse, C.M. Electrospray ionization for mass spectrometry of large biomolecules. *Science*. **1989**. 246, 64.

Fenselau, C.; Yao, X. Proteolytic labeling with ^{18}O for comparative proteomic studies: preparation of ^{18}O -labeled peptides and the $^{18}\text{O}/^{16}\text{O}$ mixture. *Methods Mol. Bio.* **2007**. 359, 135.

Fujii, K.; Nakano, T.; Kawamura, T.; Usui, F.; Bando, Y.; Wang, R.; Nishimura, T. Multidimensional protein profiling technology and its application to human plasma proteome. *J. Proteome Res.* **2004.** *3*, 712.

Fujii, K.; Nakano, T.; Kanazawa, M.; Akimoto, S.; Hirano, T.; Kato, H.; Nishimura, T. Clinical-scale high-throughput human plasma proteome analysis: lung adenocarcinoma. *Proteomics.* **2005.** *5*, 1150.

Gerber, S.A.; Rush, J.; Stemman, O.; Kirschner, M.W.; Gygi, S.P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U.S.A.* **2003.** *100*, 6940.

Gonzales, P.A.; Pisitkun, T.; Hoffert, J.D.; Tchapyjnikov, D.; Star, R.A.; Kleta, R.; Wang, N.S.; Knepper, M.A. Large-scale Proteomics and Phosphoproteomics of Human Urinary Exosomes. *J. Am. Soc. Nephrol.* **2009.** *20*, 363.

Graves, P.R.; Haystead, T.A.J. Molecular Biologist's Guide to Proteomics. *Microbiol. Mol. Biol. Rev.* **2002.** p. 39.

Green-Church, K.B.; Nichols, K.K.; Kleinholz, N.M.; Zhang, L.; Nichols, J.J. Investigation of the human tear film proteome using multiple proteomic approaches. *Mol. Vis.* **2008.** *7*, 456.

Griffin, T.J.; Xie, H.; Bandhavaki, S.; Popko, J.; Mohan, A.; Carlis, J.V.; Higgins, L. iTRAQ reagent-based quantitative proteomic analysis on a linear ion trap mass spectrometer. *J. Proteome Res.* **2007.** *6*, 4200.

Gruhler, A.; Olsen, J.V.; Mohammed, S.; Mortensen, P.; Faergeman, N.J.; Mann, M., Jensen, O.N. Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. *Mol. Cell. Proteomics*. **2005**. *4*, 310.

Gygi, S. P.; Rist, B.; Gerber, S.A.; Turecok, F; Gelb, M.H.; Aebersold, R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **1999**. *17*, 994.

Heller, M.; Mattou, H.; Menzel, C.; Yao, X. Trypsin Catalyzed ^{16}O -to- ^{18}O Exchange for Comparative Proteomics: Tandem Mass Spectrometry Comparison Using MALDI-TOF, ESI-QTOF, and ESI-Ion Trap Mass Spectrometers. *J. Am. Soc. Mass. Spectrom.* **2003**. *14*, 704.

Hicks, W.A.; Halligan, W.D.; Slyper, R.Y.; Twigger, S.N.; Green, A.S.; Olivier, M. Simultaneous Quantification and Identification Using ^{18}O Labeling with an Ion Trap Mass Spectrometer and the Analysis Software Application “ZoomQuant”. *J. Am. Soc. Mass. Spectrom.* **2005**. *16*, 916.

Higgs, R.E.; Knierman, M.D.; Gelfanova, V.; Butler, J.P.; Hale, J.E. Comprehensive Label-Free Method for the Relative Quantification of Proteins from Biological Samples. *J. Proteome Res.* **2005**. *4*, 1442.

Hood, B. L.; Lucas, D. A.; Kim, G.; Chan, K. C.; Blonder, J.; Issaq, H. J.; Veenstra, T. D.; Conrads, T. P.; Pollet, I.; Karsan, A. Quantitative Analysis of the Low Molecular Weight Serum Proteome Using ^{18}O Stable Isotope Labeling in a Lung Tumor Xenograft Mouse Model. *J. Am. Soc. Mass Spectrom.* **2005**. *16*, 1221.

Hunt, D.F.; Yates, J.R., III.; Shabanowitz, J.; Winston, S.; Hauer, C. Protein sequencing by tandem mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **1986**. *83*, 6233.

Kerr, M.K.; Churchill, G.A. Statistical design and the analysis of gene expression microarray data. *Genet. Res.* **2001.** *77*, 123.

Kim, S.C.; Chen, Y.; Mirza, S.; Xu, Y.; Lee, J.; Liu, P.; Zhao, Y. A clean, more efficient method for in-solution digestion of protein mixtures without detergent or urea. *J. Proteome Res.* **2006.** *5*, 3446.

Kirkland, P.A.; Humbard, M.A.; Daniels, C.J.; Maupin-Furlow, J.A. Shotgun proteomics of the haloarchaeon *Haloferax volcanii*. *J Proteome Res.* **2008.** *7*, 5033.

Kislinger, T.; Rahman, K.; Radulovic, D.; Cox, B.; Rossant, J.; Emili, A. PRISM, a generic large-scale proteomic strategy for mammals. *Mol. Cell. Proteomics* **2003.** *2*, 96.

Kislinger, T.; Cox, B.; Wigle, D.A.; Brown, K.; Okubo, T.; Hogan, B.; Jurisica, I.; Frey, B.; Rossant, J.; Emili, A. Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. *Cell* **2006.** *125*, 173.

Krüger, M.; Moser, M.; Ussar, S.; Thievensen, I.; Lubner, C.A.; Forner, F.; Schmidt, S.; Zanivan, S.; Fässler, R.; Mann, M. SILAC mouse for quantitative proteomics uncovers Kindlin-3 as an essential factor for red blood cell function. *Cell.* **2008.** *134*, 353.

Lal, K.; Prieto, J.H.; Bromley, E.; Sanderson, S.J.; Yates, J.R. III; Wastling, J.M.; Tomley, F.M.; Sinden, R.E. Characterisation of Plasmodium invasive organelles; an ookinete microneme proteome. *Proteomics.* **2009.** *9*, 1142.

Lee, J.; Garrett, W.M.; Cooper, B. Shotgun proteomic analysis of *Arabidopsis thaliana* leaves. *J Sep Sci.* **2007.** *30*, 2225.

Liu, H.; Sadygov, R.G.; Yates, J.R. III A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **2004.** 76, 4193.

Liu, J.; Erassov, A.; Halina, P.; Canote, M.; Nguyen, D.V.; Chung, C.; Cagney, G.; Ignatchenko, A.; Fong, V.; Emili, A. Sequential interval motif search: unrestricted database surveys of global MS/MS data sets for detection of putative post-translational modifications. *Anal. Chem.* **2008.** 80, 7846.

Maere, S.; Heymans, K; Kuiper, M. BiNGO: a Cytoscape plug-in to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **2005.** 21, 3448.

Mawuenyega, K.G.; Kaji, H.; Yamuchi, Y.; Shinkawa, T; Saito, H.; Taoka, M; Takahashi, N.; Isobe, T. Large scale identification of *Caenorhabditis elegans* proteins by multidimensional liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **2003.** 2, 23.

Mason, C.J.; Therneau, T.M.; Eckel-Passow, J.E.; Johnson, K.L.; Oberg, A.L.; Olson, J.E.; Nair, K.S.; Muddiman, D.C.; Bergen, H.R., III A Method for Automatically Interpreting Mass Spectra of ^{18}O -Labeled Isotopic Clusters. *Mol Cell. Proteomics* **2007.** 6, 305.

Mikesh, L.M.; Ueberheide, B.; Chi, A.; Coon, J.J.; Syka, J.E.P.; Shabanowitz, J.; Hunt, D.F. The Utility of ETD Mass Spectrometry in Proteomic Analysis. *Biochim. Biophys. Acta.* **2006.** 1764, 1811.

Miyagi, M.; Rao, K.C. Proteolytic ^{18}O -labeling strategies for quantitative proteomics. *Mass Spectrom. Rev.* **2007.** 26, 121.

Moore, R.E.; Young, M.K.; Lee, T.D. Qscore: An Algorithm for Evaluating SEQUEST Database Search Results. *J. Am. Soc. Mass. Spectrom.* **2002.** *13*, 378.

Mueller, L.N.; Brusniak, M.; Mani, D.R.; Aebersold, R. An Assessment of Software Solutions for the Analysis of Mass Spectrometry Based Quantitative Proteomics Data. *J. Proteome Res.* **2008.** *7*, 51.

Nelson, C. J.; Hegeman, A. D.; Harms, A. C.; Sussman, M. R. A Quantitative Analysis of Arabidopsis Plasma Membrane Using Trypsin-catalyzed ^{18}O Labeling. *Mol. Cell. Proteomics.* **2006.** *5*, 1382.

Oda, Y.; Huang, K.; Cross, F.R.; Cowburn, D.; Chait, B.T. Accurate quantification of protein expression and site-specific phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* **1999.** *96*, 6591.

Old, W.M.; Meyer-Arendt, K.; veline-Wolf, L.; Pierce, K.G.; Mendoza, A.; Sevinsky, J.R.; Resing, K.A.; Ahn, N.G. Comparison of Label-free Methods for Quantifying Human Proteins by Shotgun Proteomics. *Mol. Cell. Proteomics* **2005.** *4*, 1487.

Ong, S.E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D.B.; Steen, H.; Pandey, A.; Mann, M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002.** *1*, 376.

Ong, S.E.; Kratchmarova, I.; Mann, M. Properties of ^{13}C -substituted Arginine in Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC). *J. Proteome Res.* **2003.** *2*, 173.

Ong, S.E.; Mann, M. Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* **2005.** *1*, 252.

Ong, S.E.; Mann, M. A practical recipe for stable isotope labeling by amino acids in cell cultures (SILAC). *Nat. Protoc.* **2006.** *1*, 2650.

Ono, M.; Shitashige, M.; Honda, K.; Isobe, T.; Kuwabara, H.; Matsuzuki, H.; Hirohashi, S.; Yamada, T. Label-free Quantitative Proteomics Using Large Peptide Data Sets Generated by Nanoflow Liquid Chromatography and Mass Spectrometry. *Mol. Cell. Proteomics.* **2006.** *5*, 1338.

Panchaud, A.; Affolter, M.; Moreillon, P.; Kussman, M. Experimental and computational approaches to quantitative proteomics: status quo and outlook. *J. Proteomics* **2008.** *71*, 19.

Paradela, A.; Marcilla, M.; Navajas, R.; Ferreira, L.; Ramos-Fernandez, A.; Fernandez, M.; Mariscotti, J.F.; Garcia-del Portillo, F.; Albar, J.P. Evaluation of isotope-coded protein labeling (ICPL) in the quantitative analysis of complex proteomes. *Talanta.* **2009.** In press.

Peng, J.; Elias, J.E.; Thoreen, C.C.; Licklider, L.J.; Gygi, S.P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large scale protein analysis: the yeast proteome. *J. Proteome Res.* **2003.** *2*, 43.

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

Petyuk, V. A.; Qian, W. J.; Chin, M. H.; Wang, H.; Livesay, E. A.; Monroe, M. E.; Adkins, J. N.; Jaitly, N.; Anderson, D. J.; Camp, D. G., II; Smith, D. J.; Smith, R. D. Spatial mapping of protein abundances in the mouse brain by voxelation integrated with high throughput liquid chromatography-mass spectrometry. *Genome Res.* **2007.** *17*, 328.

Pierce, A.; Unwin, R.D.; Evans, C.A.; Griffiths, S.; Carney, L.; Zhang, L.; Jaworska, E.; Lee, C.F.; Blinco, D.; Okoniewski, M.J.; Miller, C.J.; Bitton, D.A.; Spooncer, E.; Whetton, A.D. Eight-channel iTRAQ enables the comparison of the activity of 6 leukaemogenic tyrosine kinases. *Mol. Cell. Proteomics*. **2008**. 7, 853.

Qian, W.J.; Monroe, M.E.; Liu, T.; Jacobs, J.M.; Anderson, G.A.; Shen, Y.; Moore, R.J.; Anderson, D.J.; Zhang, R.; Calvano, S.E.; Lowry, S.F.; Xiao, W.; Moldawer, L.L.; Davis, R.W.; Tompkins, R.G.; Camp, D.G., II; Smith, R.D. Quantitative proteome analysis of human plasma following in vivo lipopolysaccharide administration using $^{16}\text{O}/^{18}\text{O}$ labeling and the accurate mass and time tag approach. *Mol. Cell. Proteomics* **2005**. 4, 700.

Reimers, M. Statistical analysis of microarray data. *Addict. Biol.* **2005**. 10, 23.

Reynolds, K.J.; Yao, X.; Fenselau, C. Proteolytic ^{18}O labeling for comparative proteomics: evaluation of endoprotease Glu-C as the catalytic agent. *J. Proteome. Res.* **2002**. 1, 27.

Rosman, K.J.R.; Taylor, P.D.P. Isotopic compositions of the elements. *Pure Appl. Chem.* **1998**. 70, 217.

Ross, P.L.; Huang, Y.N.; Marchese, J.N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlett-Jones, M.; He, F.; Jacobson, A.; Pappin, D.J. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**. 3, 1154.

Ru, Q.C.; Katenhusen, R.A.; Zhu, L.A.; Silberman, J.; Yang, S.; Orchard, T.J.; Brzeski, H.; Liebman, M.; Ellsworth, D.L. Proteomic profiling of human urine using multi-dimensional protein identification technology. *J. Chromatogr. A*. **2006**. *1111*, 166.

Sakai, J.; Kojima, S.; Yanagi, K.; Kanaoka, M. ^{18}O -labeling quantitative proteomics using an ion trap mass spectrometer. *Proteomics*. **2005**. *5*, 16.

Schmidt, A.; Kellermann, J.; Lottspeich, F. A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics*. **2005**. *5*, 4.

Schnölzer, M.; Jedrzejewski, P.; Lehmann, W.D. Protease-catalyzed incorporation of ^{18}O into peptide fragments and its application for protein sequencing by electrospray and matrix-assisted laser desorption/ionization mass spectrometry. *Electrophoresis* **1996**. *17*, 945.

Sevinsky, J.R.; Brown, K.J.; Cargile, B.J.; Bundy, J.L.; Stephenson, J.L. Jr. Minimizing back exchange in $^{18}\text{O}/^{16}\text{O}$ quantitative proteomics experiments by incorporation of immobilized trypsin into the initial digestion step. *Anal. Chem.* **2007**. *79*, 2158.

Shiio, Y.; Aebersold, R. Quantitative proteome analysis using isotope coded affinity tags and mass spectrometry. *Nat. Protoc.* **2006**. *1*, 139.

Smith, J. R.; Matus, I. R.; Beard, D. A.; Greene, A. S. Differential expression of cardiac mitochondrial proteins. *Proteomics*. **2008**. *8*, 446.

Stewart, I.I.; Thomson, T.; Figeys, D. ^{18}O Labeling: a tool for proteomics. *Rapid Commun. Mass Spectrom.* **2001**. *15*, 2456.

Strittmatter, E.F.; Ferguson, P.L.; Tang, K.; Smith, R.D. Proteome analyses using accurate mass and elution time peptide tags with capillary LC time-of-flight mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2003**, *14*, 980.

Tabb, D.L.; McDonald, W.H.; Yates, J.R. III DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **2002**, *1*, 21.

Venable, J.D.; Dong, M.Q.; Wohlschlegel, J.; Dillin, A.; Yates, J.R. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat. Methods.* **2004**, *1*, 39.

Wang, G.; Wu, W.W.; Zeng, W.; Chou, C.L.; Shen, R.F. Label-Free Protein Quantification Using LC-Coupled Ion Trap or FT Mass Spectrometry: Reproducibility, Linearity, and Application with Complex Proteomes. *J. Proteome Res.* **2006**, *5*, 1214

Washburn, M. P.; Wolters, D.; Yates, J.R. III. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001**, *19*, 242.

White, C.A.; Oey, N.; Emili, A. Global Quantitative Proteomic Profiling through ^{18}O -Labeling in Combination with MS/MS Spectra Analysis. *J. Proteome Res.* **2009**, *8*, 3653.

Wilm, M.; Mann, M. Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* **1996**, *68*, 1.

Wolters, D.A.; Washburn, M.P.; Yates, J.R. III An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* **2001**, *73*, 5683.

Yao, X.; Freas, A.; Ramirez, J.; Demirev, P.A.; Fenselau, C. Proteolytic ^{18}O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* **2001.** 73, 2836.

Yao, X.; Afonso, C.; Fenselau, C. Dissection of proteolytic ^{18}O labeling: endoprotease-catalyzed ^{16}O -to- ^{18}O exchange of truncated peptide substrates. *J. Proteome Res.* **2003.** 2, 147.

Yocum, A.K.; Chinnaiyan, A.M. Current affairs in quantitative targeted proteomics: multiple reaction monitoring-mass spectrometry. *Brief. Funct. Genomic Proteomic.* **2009.** 8, 145.

Zeeburg, B.R.; Feng, W.; Wang, G.; Wang, M.D.; Fojo, A.T.; Sunshine, M.; Narasimhan, S.; Kane, D.W.; Reinhold, W.C.; Lababidi, S.; Bussey, K.J.; Riss, J.; Barrett, J.C.; Weinstein, J.N. GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol.* **2003.** 4, R28.

Zhang, G.; Neubert, T.A. Automated Comparative Proteomics Based on Multiplex Tandem Mass Spectrometry and Stable Isotope Labeling. *Mol. Cell. Proteomics* **2006.** 5, 401.

Zhang, N.; Li, X.J.; Pan, S.; Schwikowski, B.; Aebersold, R. ProbIDtree: an automated software program capable of identifying multiple peptides from a single collision-induced dissociation spectrum collected by a tandem mass spectrometer. *Proteomics* **2005.** 5, 4096.

Zhang, Y.; Wen, Z.; Washburn, M.P.; Florens, L. Effect of Dynamic Exclusion Duration on Spectral Count Based Proteomics. *Anal. Chem.* **2009.** 81, 6317.

Zhao, Y.; Lee, W.N.; Xiao, G.G. Quantitative proteomics and biomarker discovery in human cancer. *Expert Rev. Proteomics.* **2009.** 6, 115.

Zimmer, J.S.D.; Monroe, M.E.; Qian, W.J.; Smith, R.D. Advances in proteomics data analysis and display using an accurate mass and time tag approach. *Mass Spectrom. Rev.* 2006. 25, 450.

Appendix A: Basic protocol

The following peptide labeling protocol begins with two dry protein pellets, each putatively of the same quantity and derived in the same way (e.g. from TCA/acetone precipitation) but from separate samples. The protocol specifies the minimal steps necessary to perform in-digestion ^{18}O labeling of one the two samples while handling both in an otherwise identical manner. More complicated procedures typically are used to perform the tryptic digestion of proteins, and this protocol may easily be elaborated upon. However, there exists some evidence that digestion without the inclusion of urea can produce comparable or even superior results (Kim, 2006).

Materials

- ammonium bicarbonate (Sigma Aldrich)
- soluble trypsin (Promega)
- 95% pure (or better) ^{18}O water (Isotec)
- MilliQ-quality water (or better)
- formic acid (Fisher)
- acetonitrile (Fisher)

Reagents

- 800 mM NH_4HCO_3 in MilliQ-quality water
- 800 mM NH_4HCO_3 in 95% pure ^{18}O water
- 0.5 mg/mL trypsin in MilliQ-quality water
- 0.5 mg/mL trypsin in 95 % pure ^{18}O water

Keep all frozen at -80°C between uses and on ice during use.

Digestion Protocol

It's recommended that protein pellets containing more than 60 μg of protein not be used where possible. Resuspending larger pellets manually can be very time consuming.

1. To sample A, add 2 μL of 800 mM NH_4HCO_3 in MilliQ-quality water and 26 μL of MilliQ-quality water. To sample B, add 2 μL of 800 mM NH_4HCO_3 in 95% pure ^{18}O water and 26 μL of 95% pure ^{18}O water.
2. Pipette up and down vigorously until no particulates are visible (this may require several minutes of effort).
3. Add 4 μL of soluble trypsin in MilliQ-quality water to sample A and 4 μL of soluble trypsin in 95% pure ^{18}O water to sample B. Mix well.
4. Allow both digests to stand for at least 48 hours at 30°C . No agitation is required.
5. Freeze the digests at -80°C and keep until needed.

Sample Loading in Preparation for LC-MS/MS

1. Thaw the relevant digests on ice.
2. Transfer 10 μL of the digest of sample A to an Eppendorf tube.
3. Add 2 μL of formic acid and mix well.
4. Transfer 10 μL of the digest of sample B to the same tube and mix well.
5. Add 1 μL of acetonitrile to the sample and mix well.
6. Load onto a three-phase column (SCX material sandwiched between two layers of C18 reverse phase material). Note that the first layer of C18 material should be sufficient in volume to accommodate the full quantity of peptides being loaded.

Appendix B: Supplementary material

B1. Supplementary tables

Table S1: Effect of isolation width on peptide and protein identification rates. Reverse-phase chromatography LC-MS/MS experiments were performed using different isolation widths. The commonly-employed isolation width (IW) of 2 m/z yielded the best identification rates, although the more complex (mixed) sample, consisting of even portions of labeled and unlabeled *C. elegans* digests, evidenced an average 14% lower rate of protein identification than the less complex sample of unlabeled *C. elegans* digest alone. Relative to the mixed sample run (IW 2 m/z), the mixed sample (IW 6 m/z) demonstrated an average 20% lower rate of protein identification (with singly-charged matches excluded), and the mixed sample (IW 10 m/z) an average 38% lower rate. The approach of fragmenting each precursor ion twice, once at IW 2 m/z immediately followed by once at IW 10 m/z (the first scan to be used for identification and the second for quantification), yielded an average 23% lower rate relative to the mixed sample (IW 2 m/z) experiments.

Table S2: Comparison of distributions of ratios calculated for singly- and multiply-charged peptide instances. Distribution data is presented for ratios calculated from the observed peaks across eight LC-MS/MS experiments, categorized by the charge of the peptide. The expected ratio for the proteins in a given run is shown along with the mean and standard deviation of the results for singly-charged and multiply-charged peptide instances (spectra). The means of the singly-charged instances are consistently greater than the means of the multiply-charged instances, showing a larger deviation from the expected means.

Table S3: Peptide ratios do not vary independently between experiments. A list of peptides was assembled, from data collected in four separate LC-MS/MS experiments, for each combination of two of the experiments in which the peptides were identified in both experiments. Two lists ranking the peptides in order of ascending calculated ratio according to each of the experiments were then made for each list of peptides. If the peptide ratios are not correlated (i.e. if they vary randomly within the general distribution across experiments), then

approximately half of the time, one ratio for a peptide should appear below the median of one ranked list and above the median of the other. The other half of the time, the two ratios should appear on the same side of the medians. Chi-square goodness of fit tests demonstrated dramatically that this hypothesis did not hold true. Therefore, there must be correlation of ratios for peptides between experiments.

Table S4: Empirical inverses of cumulative distributions for ratio distributions. An inverse look-up table operating upon the empirically-derived cumulative distribution functions for protein ratios based on single-peptide and multi-peptide results. The functions were built upon the logarithmic values of ratios merged from five LC-MS/MS experiments after normalizing the means of their five individual distributions to log 1.0 (i.e. 0.0). The ratios corresponding to the logarithms are also shown.

Table S5: Empirical false positive rates (all spectra and unflagged spectra only). To determine the impact a higher isolation width has on the predictive power of the Statquest probabilistic model, five LC/LC-MS/MS experiments (isolation with of 10 m/z , unlabeled and labeled worm lysate digest mixed in different ratios) are examined for decoy sequence matches (99% confidence) and their rates of false positives empirically derived. Doubly-charged matches exceed confidence predictions, alarm-flagged or not, and are overwhelmingly the most numerous. However, triply-charged matches appear to suffer from a high rate of false positive identifications. Discarding the alarm-flagged instances reduces the rate markedly but not to the predicted one percent, while having the unwelcome side-effect of also discarding the majority of the triply-charged matches.

Table S1: Effect of isolation width on peptide and protein identification rates.

Sample Loaded	Isolation Width	Replicate	Number of Peptides (95% confidence)	Number of Proteins (95% confidence)
2 µg of unlabeled worm digest only 2 µg of unlabeled worm digest + 2 µg of labeled worm digest	2 <i>m/z</i>	1	461	234 (average 224)
		2	447	214
	2 <i>m/z</i>	1	463	212 (average 192)
		2	365	172
	6 <i>m/z</i>	1	307	148 (average 154)
		2	317	162
	10 <i>m/z</i>	1	281	116 (average 120)
		2	313	123
	paired scans	1	265	159 (average 147)
		2	247	135

Table S2: Comparisons of the distributions of ratios calculated from spectra for singly-charged peptides to those from multiply-charged peptides.

Experiment Number	Expected Mean	Charge	Number of Spectra	Observed Mean	Standard Deviation
1	0.20	+1	13	0.27	0.12
		+2 & +3	300	0.24	0.095
		all	313	0.24	0.097
2	0.50	+1	71	0.62	0.27
		+2 & +3	761	0.55	0.31
		all	832	0.55	0.31
3	1.00	+1	150	1.52	0.60
		+2 & +3	887	1.12	0.37
		all	1037	1.18	0.43
4	2.00	+1	174	3.18	2.65
		+2 & +3	762	2.00	1.21
		all	936	2.22	1.64
5	2.00	+1	51	2.16	0.80
		+2 & +3	496	2.05	1.17
		all	547	2.06	1.14
6	2.00	+1	129	3.43	2.74
		+2 & +3	1216	2.22	1.50
		all	1345	2.38	1.69
7	2.00	+1	16	2.40	0.83
		+2 & +3	527	1.98	0.72
		all	543	2.00	0.73
8	5.00	+1	15	5.18	1.24
		+2 & +3	299	3.82	1.34
		all	314	3.89	1.36

Table S3: P-values for hypothesis that ratios calculated for peptides vary independently between experiments.

Experiment s being compared (# of pairs of peptides)	Expected number of pairs straddling the median	Expected number of pairs on the same side of the median	Observed number of pairs straddling the median	Observed number of pairs on the same side of the median	Chi-square goodness of fit P-value for H_0
1 & 2 (143)	71.5	71.5	40	103	0.0003
1 & 3 (205)	102.5	102.5	33	172	0.0000
1 & 4 (115)	57.5	57.5	39	76	0.0189
2 & 3 (134)	67	67	23	111	0.0000
2 & 4 (158)	79	79	24	134	0.0000
3 & 4 (151)	75.5	75.5	53	98	0.0127

Table S4: Empirical inverses of cumulative distributions for ratio distributions.

Proportion of Total	Ratio determined from two or more peptides		Ratio determined from one peptide	
	Logarithm of ratio (mean of distribution is 0.00)	Ratio (mean of distribution is 1.00)	Logarithm of ratio (mean of distribution is 0.00)	Ratio (mean of distribution is 1.00)
0.01	<= -0.181	<= 0.659	<= -0.414	<= 0.385
0.05	<= -0.130	<= 0.741	<= -0.202	<= 0.628
0.10	<= -0.106	<= 0.783	<= -0.151	<= 0.706
0.15	<= -0.076	<= 0.839	<= -0.116	<= 0.766
0.20	<= -0.063	<= 0.865	<= -0.092	<= 0.809
0.25	<= -0.050	<= 0.891	<= -0.068	<= 0.855
0.30	<= -0.037	<= 0.918	<= -0.054	<= 0.883
0.35	<= -0.032	<= 0.929	<= -0.040	<= 0.912
0.40	<= -0.021	<= 0.953	<= -0.027	<= 0.940
0.45	<= -0.010	<= 0.977	<= -0.017	<= 0.962
0.50	<= -0.003	<= 0.993	<= -0.005	<= 0.986
0.55	<= 0.007	<= 1.016	<= 0.011	<= 1.026
0.60	<= 0.014	<= 1.033	<= 0.026	<= 1.062
0.65	<= 0.022	<= 1.052	<= 0.040	<= 1.096
0.70	<= 0.031	<= 1.074	<= 0.053	<= 1.130
0.75	<= 0.042	<= 1.102	<= 0.067	<= 1.167
0.80	<= 0.051	<= 1.125	<= 0.084	<= 1.213
0.85	<= 0.074	<= 1.186	<= 0.110	<= 1.288
0.90	<= 0.108	<= 1.282	<= 0.158	<= 1.439
0.95	<= 0.144	<= 1.393	<= 0.224	<= 1.675
0.99	<= 0.195	<= 1.567	<= 0.337	<= 2.173

Table S5: Empirical false positive rates derived from decoy (reversed database) matches, for the full set of spectra and for the set of spectra after the alarm-flagged spectra are excluded.

Experiment Number (Expected Ratio)	Charge	Number of Spectra (Number of Unflagged Spectra)	Number of Decoy Matches	Number of Unflagged Decoy Matches	False Positive Rate	False Positive Rate (Unflagged Only)
all combined	+1	359 (265)	7	0	1.9%	0.0%
	+2	2585 (2075)	6	3	0.2%	0.1%
	+3	516 (146)	98	5	19.0%	3.4%
	all	3454 (2486)	111	8	3.2%	0.3%
1 (0.20)	+1	26 (13)	4	0	-	-
	+2	410 (283)	2	1	-	-
	+3	98 (17)	23	0	-	-
2 (0.50)	+1	79 (71)	0	0	-	-
	+2	792 (732)	1	1	-	-
	+3	104 (29)	25	0	-	-
3 (1.00)	+1	197 (150)	1	0	-	-
	+2	963 (842)	1	1	-	-
	+3	124 (47)	17	1	-	-
4 (2.00)	+1	29 (16)	1	0	-	-
	+2	563 (487)	2	0	-	-
	+3	122 (43)	17	3	-	-
5 (5.00)	+1	28 (15)	1	0	-	-
	+2	417 (291)	0	0	-	-
	+3	68 (10)	16	1	-	-

B2. Supplementary figures

Figure S1: Normal probability plots show long tails. Calculated ratios for individual spectra form a nearly normal distribution. The central 95% of data points show high correlation to a line when plotted by normal statistic order medians in a normal probability plot. The tails (consisting of the points deviating furthest from the mean) are demonstrably “long”, rising above the line for the highest values and falling below it for the smallest. (a) Probability plot of ratios calculated for spectra from a multi-dimensional LC-MS/MS run in which the assigned identifications were made based on the assumption of the intact peptides bearing a single unit of charge. A distinct bias is apparent, with the logarithm of the median ratio falling well above the

expected value of 0.0 (i.e. the expected ratio of 1: 1). (b) Probability plot of ratios calculated for spectra from the same LC-MS/MS run in which the assigned identifications were made based on the assumption that the intact peptides were bearing two units of positive charge. Little bias of the median ratio relative to the expected value is evident.

Figure S2: Comparison of distributions of unflagged ratios and alarm-flagged ratios. Boxplots for the distributions of ratios observed for five LC-MS/MS experiments are depicted. The left-hand boxplot for an experiment of a given expected ratio represents the distribution of the protein ratios which were derived purely from spectra that passed the y-ion pair consistency checking. The corresponding right-hand boxplot represents the distribution of protein ratios including alarm-flagged results. The standard deviations of the distributions are shown beneath the boxplots (σ for the distributions of unflagged ratios, σ^* for the distributions of flagged ratios).

Figure S3: Plot of ratios calculated for individual spectra versus time during LC-MS/MS.

Plot of ratios calculated for individual spectra versus the chronological positions of the spectra in a six-step multi-dimensional LC-MS/MS run. The putative charge of the intact peptide associated with a particular spectrum is distinguished by a box (singly-charged) or a diamond (doubly- or triply-charged). Total ion current (TIC) for the chromatography is shown in silhouette form at the bottom. No evidence of back-labeling (namely reversion of fully and partially labeled peptides to unlabeled peptides via the exchange of the ^{18}O labels with ^{16}O atoms from normal water) was observed. If significant back-labeling had occurred, a trend of rising ratios over time as the back-labeling reactions took place would be expected.

B3. Notes on the effects of isotope distribution patterns

For those spectra that incur a significant correction for the partially-labeled species, a question arises as to whether the steps taken actually lead to an overcorrection, especially in those instances where the contribution of high m/z y -ions is substantial.

A temporary version of yRatios was created in which a deduction to the contribution of the partially labeled species was performed, representing the separate natural (^{13}C) isotopic contribution of the partially labeled peptides to the monoisotopic fully labeled species' peak intensity. The modified yRatios was applied to the LC-MS/MS datasets to determine the magnitude of the effect. For each dataset examined, only a tiny increase in the mean ratio (average 2.6%) of the peptide instances was observed. The standard deviation diminished by less than one percent in all cases (average 0.95%).

To test the correspondence of the actual isotope distribution patterns to the theoretical patterns (the theoretical being used to estimate the unlabeled y -ion's contribution to the peak intensity at +2 m/z relative to the monoisotopic unlabeled y -ion), two reverse-phase LC-MS/MS experiments (unlabeled *C. elegans* digest only) were analyzed. The first, which employed an isolation width of 2 m/z , showed that the peak intensity of the third peak of the y -ions' isotopic distribution did vary about the theoretical value, with roughly half falling on either side. When counting those that fell below the theoretical intensity as making no contribution (rather than a negative contribution, since a deduction of the full theoretical value is always applied and only positive remainders are added to the labeled peak intensity value), an average peak intensity of 1.45 of the theoretical value was noted. In the second experiment, in which an isolation width of 10 m/z was employed, nearly all observed intensities exceeded the theoretical and the average increased to 1.67 of theoretical.

A defined constant is available in the yRatios source code to allow investigators to compensate for this disparity. However, broadly applying a single multiplier might be an unnecessarily crude approach. It may be possible to better estimate each y-ion's particular disposition using the difference between the theoretical intensity of the second peak of the isotopic series and its observed intensity to derive a better approximation for that portion of the third peak's intensity that is contributed by the unlabeled y ion.