Amino Acids (2012) 42:129–138
DOI 10.1007/s00726-010-0614-3

REVIEW ARTICLE

Existing bioinformatics tools for the quantitation of post-translational modifications

Jens Allmer

Received: 16 October 2009 / Accepted: 27 April 2010 / Published online: 15 May 2010

Abstract Mass spectrometry (MS)-based proteomics, by itself, is a vast and complex area encompassing various mass spectrometers, different spectra, and search result representations. When the aim is quantitation performed in different scanning modes at different MS levels, matters become additionally complex. Quantitation of post-translational modifications (PTM) represents the greatest challenge among these endeavors. Many different approaches to quantitation have been described and some of these can be directly applied to the quantitation of PTMs. The amount of data produced via MS, however, makes manual data interpretation impractical. Therefore, specialized software tools meet this challenge. Any software currently able to quantitate differentially labeled samples may theoretically be adapted to quantitate differential PTM expression among samples as well. Due to the heterogeneity of mass spectrometry-based proteomics; this review will focus on quantitation of PTM using liquid chromatography followed by one or more stages of mass spectrometry. Currently available free software, which either allow analysis of PTM or are easily adaptable for this purpose, is briefly reviewed in this paper. Selected studies, especially those related to phosphoproteomics, shall be used to highlight the current ability to quantitate PTMs.

Keywords Quantitation · Quantification · Post-translational modification · Software · LC-MS · PTM

J. Allmer (✉)
Molecular Biology and Genetics, Izmir Institute of Technology, Urla, Izmir, Turkey
e-mail: jens@allmer.de

Introduction

Proteins may differ from a naïve translation of their encoding genes. Alternative splicing, RNA editing, and protein splicing events give rise to multiple proteins with different sequences (Nair et al. 2004; Lander et al. 2001). Apart from displaying a different sequence, proteins can differ in their post-translational modifications (PTMs). This does not lead to a new protein since the sequence remains the same but to a new "protein species" (Jungblut et al. 1996; Schluter et al. 2009). PTMs and their differential expression levels are vitally important for understanding biological function (Vissers et al. 2009). A singly changed, missing, or additional PTM in a protein may significantly perturb its function (Steen et al. 2005). Furthermore, many critical events are mediated by changes in PTMs rather than by transcriptional regulation, which necessitates the need to quantitatively investigate changes in post-translational protein modifications systematically (Olsen et al. 2006). Among the more than 100 possible PTMs (O'Donovan et al. 2001) phosphorylation is one of the most important (Yan et al. 1998; Hunter 1998; Nair et al. 2004) and has been investigated using mass spectrometry in several aspects (Asara et al. 2008; Beausoleil et al. 2004; Lu et al. 2007). 3-nitro-L-tyrosine, a marker for oxidative stress, is another interesting PTM (Tsikas and Caidahl 2005).

Mass spectrometry (MS) has become the tool of choice in proteomics research and is employed in protein detection/identification, sequencing, and quantitation (Aebersold and Mann 2003). In brief, a complex sample is first separated in several dimensions (e.g., gel electrophoresis or liquid chromatography) to reduce the number of proteins that are analyzed in one MS run. Since short amino acid sequences (peptides) lend themselves better to MS

analysis, the proteins are cleaved into peptides usually using Trypsin that are then transferred to MS, usually via reverse phase liquid chromatography (LC), where their mass to charge ratios (*m/z*) are determined. Most modern mass spectrometers are capable of several stages of MS and more information can be gained about a particular peptide entering the MS by using a second stage of MS (MS/MS, tandem-MS, MS^2) following fragmentation of the peptide, where the fragment ions can form mass ladders in a manner similar to Sanger sequencing for nucleotides. In this fashion the peptide sequence can be determined using data analysis methods like de novo sequencing and database searching (Shadforth et al. 2005; Kapp et al. 2005). Some mass spectrometers allow additional stages of MS for analysis of the fragments of fragments, which makes further analysis possible. These different stages of MS lead to different possible experimental approaches for protein identification and quantitation.

Using MS, proteins can be identified by their peptide mass fingerprint (Mann et al. 1993). Using MS^2 peptides can be identified more confidently by their sequence (Shevchenko et al. 1996; Mann and Wilm 1994).With MS/MS/MS (MS^3), the peptide sequence information can be further refined and the confidence in peptide identification is thus higher. Furthermore, diagnostic ions, which, for example, confirm a PTM, can be found on this level (Mouls et al. 2009). All these stages of MS are used for quantitation of peptides and thus proteins. The quantitation methods used are numerous but can be grouped in two categories: those using differential labeling and those which quantitate without using labels (label-free). The methods that employ differential labeling for quantitation can be further categorized according to the experimental step at which the label is introduced. This can be either done in vivo, metabolically (Gygi et al. 1999; Oda et al. 1999; Ong et al. 2002), after protein extraction, and either during or after protein digestion (Flory et al. 2002; Gallezot et al. 2008; Julka and Regnier 2005). Additionally, standards can be injected (spiked into the sample) in order to derive quantitative information (Ong and Mann 2005; Fig. 1).

Another issue is how the labeling is performed. Stable isotopic labeling by amino acids in cell culture (SILAC) is achieved by providing a sample with amino acids, which contain heavy isotopes of nitrogen (^{15}N), carbon (^{13}C) or less often deuterium (^{2}H).

Enzymatic or chemical labeling of proteins or peptides can be performed by a multitude of strategies. For instance, ^{18}O from $H_2^{18}O$ can be incorporated into the peptide during protein digestion and peptides can be derivatized by numerous chemical groups. Isotope-coded affinity tag (ICAT), isotope-coded protein labeling (ICPL), and isotope tags for relative and absolute quantification (iTRAQ) are

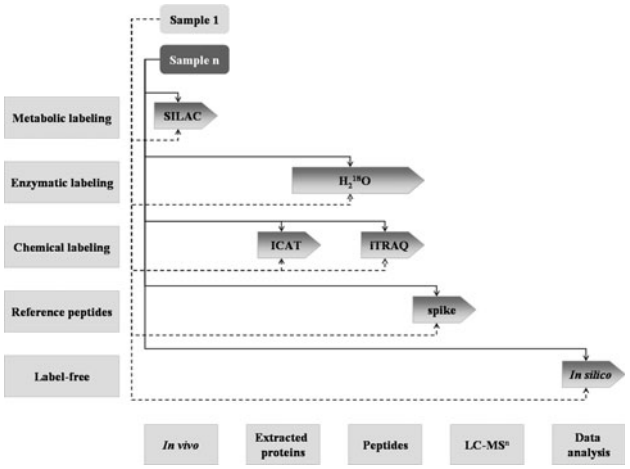

Fig. 1 Differential labeling of samples can be performed at different experimental times. Labeled nutrient sources can be differentially distributed among samples. Labeling can also be applied during protein digestion or to peptides and finally peptide standards can be spiked into the sample. Afterwards, computational analyses can infer quantitative data. *Arrow shapes* indicate at which time a labeling method is applied. Samples are separate before the arrow and are mixed afterwards, *arrows on the same row* represent alternatives. The method *labels* on the arrows are merely examples and many others exist

some of the methods that are used for chemical labeling today (Li et al. 2003a; Schmidt et al. 2005; Ross et al. 2004). The large field of labeling strategies cannot be detailed here but has been reviewed extensively (Ong and Mann 2005; Bantscheff et al. 2007).

Introducing a label for differentially marking a peptide has proven useful in quantitative proteomics (Wang et al. 2008). Concerns such as incomplete labeling (Ong et al. 2002) or changing chemical properties that influence separation are not visible during MS itself but affect online data processing. Thus, regarding only the MS measurements, a label changes at least the mass of the measured peptide species. PTMs also change the mass of the peptide/protein that they are attached to. Thus, the same strategies used in differential labeling could potentially be employed for the analysis of PTMs and their differential expression. PTMs are always introduced metabolically but may be altered or lost due to experimental procedures.

If the chemical properties of peptides are altered due to an attached label or PTM, the labeled and unlabeled peptide species may not be present in the same LC fractions or gel spots. This may make it necessary to first determine runtime parameters before quantitation can be performed.

Quantitation methods

Different labeling strategies incur different advantages and disadvantages for quantitation. Although metabolic

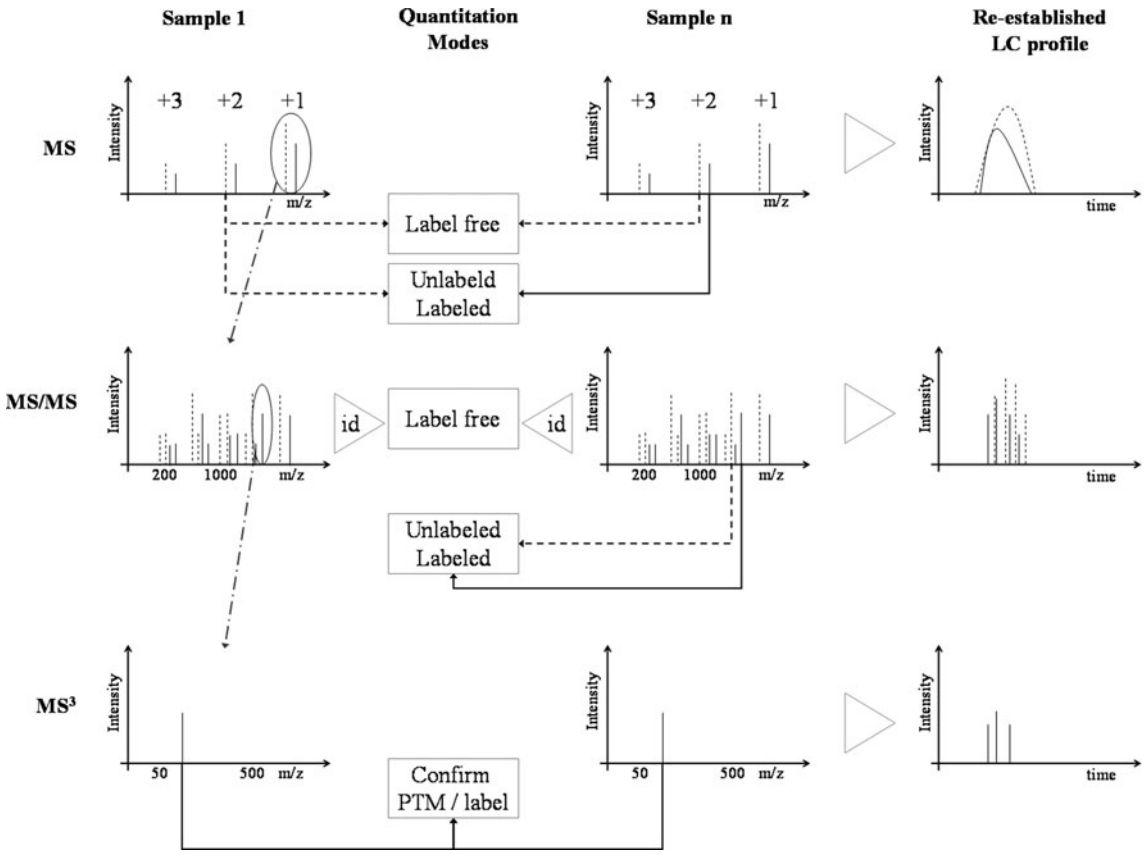

Fig. 2 Quantitation can be performed on several MS levels either using labels or label-free. Three levels of MS are shown along with common ways in which they are used for quantitation. Samples (1–*n*) can be compared according to the quantitation method listed between them. *Dashed lines* represent peaks from unmodified precursors whereas *solid lines* represent peaks from modified precursor ions. In order to highlight experiment resolution, the re-established LC profiles are shown on the right with the triangle indicating that many such measurements are needed. The *dashed-dotted arrows* indicate the precursor ion of the spectrum that is pointed to

labeling via SILAC, for instance, removes most experimental errors from differential sample handling, only three samples can be compared at one time and full labeling is rarely achieved which complicates data analysis (Oeljeklaus et al. 2009). Similar to expectable side reactions in chemical labeling strategies, labeled amino acids could be metabolized leading to mass changes for unexpected amino acids.

While iTRAQ currently enables the comparison of up to eight different samples, the analysis of the resulting MS/MS spectra has to be performed with great care (Zhang et al. 2001). The iTRAQ labeling method has recently been extended to allow for localization and quantitation of nitration (Chiappetta et al. 2009).

For chemical labeling strategies, as in ICAT and ICPL, side reactions and incomplete labeling can pose problems during data analysis. Peptides can be synthesized and modified for use in relative and absolute quantitative work as done in the AQUA method (Gerber et al. 2003).

It is difficult to measure proteins using MS in a top-down approach; inferring quantitative data from proteotypic peptides in a bottom-up fashion seems to be a useful workaround. Peptides can be identified on all levels of MS, by their mass via MS, their mass and sequence with MS/MS, and additional sequence information with MS^3. Quantitation can also be performed on several levels of MS (refer to Fig. 2), which will be detailed in subsequent sections.

A method which cannot be placed into one of the following categories was used by Venable and co-workers (2004) who performed data-independent measurement of differentially labeled samples using sequential scan windows of 10 Dalton (Da) in width to acquire MS/MS spectra. They found that their method is superior to typical quantitation using LC-MS.

Data analysis

Except for very small datasets, analysis of data from MS experiments cannot be manually performed since large data sets need to be correlated and then quantitated. Therefore, a variety of tools can either be directly used for analysis of data or are readily adaptable to quantify PTMs from MS measurements and/or identification results (Table 1).

Table 1 Software which allows quantitation of PTMs or which can be adapted to perform this function

Software	Operating System	References	Method	MS level	Samples	Labeling strategy	Adapt-ability
MSMAG	Generic	Allmer (2010)	Weighted spectral counting	2	*n*	Generic	++
mzMine	Generic	Katajamaa and Oresic (2005)	A platform which among other methods offers AMT analyses	1	2	Generic	+
MaxQuant	?	Cox and Mann (2008)	*m/z*, elution time, signal intensity	1 + 2	2	SILAC	+
ProRata	Generic	Pan et al. (2006)	Uses maximum likelihood point estimation and profile likelihood confidence interval estimation. Depends on DTASelect (Tabb et al. 2002)	1 + 2	2	SIL	+
Stem	Server	Shinkawa et al. (2005)	Utilizes Mascot results for quantitation.	2	2	SIL	?
PeptideProphet ProteinProphet	LINUX	Nesvizhskii et al. (2003)	Include spectral counting	2	*n*	Label-free	?
msInspect	Generic	Bellew et al. (2006)	A platform with different analysis possibilities (e.g., XIC)	1	*n*	Generic	+
msInspect/ AMT	Generic	May et al. (2008)	AMT with MS/MS result integration Depends on msInspect or CPAS	1 + 2	*n*	None	+
MRMer	Generic	Martin et al. (2008)	MRM with SILAC. Depends on msInspect	1	*n*	SILAC	+
Qurate	Generic	May et al. (2009)	Platform for manual evaluation and curation of quantitative data	1, 2	*n*		na
RelEx	Microsoft Windows	MacCoss et al. (2003)	Uses a linear correlation algorithm to estimate abundance. Depends on DTASelect	2	*2*	SIL	+
MapQuant	Generic, but needs to be compiled.	Leptos et al. (2006)	Depends on OpenRaw format (http://arep.med.harvard.edu/OpenRaw/or_description.html)	1 + 2	*n*	None	+
MSight	Microsoft Windows	Palagi et al. (2005)	Treats MS data as an image and uses methods similarly used in 2D-Gel analysis	1 + 2	*N*	Generic	−
MaXIC-Q	Web service	Tsou et al. (2009)	AMT including alignment of multiple MS runs. Depends on either Sequest or Mascot results for identification. Speed, security and accessibility may be an issue.	1 + 2	2	SIL	−
Viper	Microsoft Windows	Monroe et al. (2007)	Uses and creates AMT tag database in Microsoft SQL Server or Access format	1 + 2	*n*	Generic	++
PhosphoPic	Microsoft Windows	Hoffert et al. (2007)	Dependent on Sequest	1 + 2 + 3	*n*	Label-free	++
PEPPeR	Generic, needs Perl	Jaffe et al. (2006)	Complete implementation not easily retrievable	1 + 2	*n*	Generic	+
ASAPRatio	Linux	Li et al. (2003a, b)	Mostly performs statistical analyses	1, 2	2	SIL	+
IDEAL-Q	?	Tsou et al. (2009)	Currently not available. Applies several new filter and validation criteria	1 + 2	*n*	Label-free	+

Note that the list of software is not comprehensive and that only non-commercial tools are considered

++ already possible, + easy adaptability of the tool, *na* not applicable, − difficult to implement quantitation of PTMs, ? information was not provided or could not be determined

The following two studies do not fit into the categories in the next section and are therefore mentioned here. Qurate is a tool that does not directly perform quantitation but it is helpful for visual inspection and editing of quantitative results (May et al. 2009). Qurate like many other tools retrieves information from the standard file formats mzXML (Pedrioli et al. 2004), pepXML, and protXML (Keller et al. 2005). Fragment assignment by visual assistance is a method, which enables users to gain more information from a dataset than currently possible with automated approaches. It especially targets the analysis of phosphorylated protein isoforms with multiple differentially phosphorylated loci (Guan and Burlingame 2009).

Quantitation with LC-MS

With LC-MS, quantitative data is derived from a mass spectrum representing all peptides currently detectable by the mass spectrometer; a so-called survey spectrum. Measurements taken from mixed differentially labeled samples from different experimental conditions can then be compared for changes in relative abundance of peptides and proteins. In case of label-free quantitation, the measurement has to be performed for each sample. Survey spectra may contain a large number of peptide ions depending on the complexity of the sample and the amount of applied pre-fractionation.

This approach is sensitive to the complexity of the sample and the number of charges at which each peptide exists and therefore benefits from mass spectrometric platforms capable of high *m/z* resolution and highly reproducible LC systems (America and Cordewener 2008). It would be best if all charged species of the peptide are measured and their peak areas summed for accurate quantitative information (Fig. 2) although other approaches are possible (Andreev et al. 2006). The resulting measure could be compared among experiments. In this approach, both labeled and unlabeled peptides can be measured at the same time and their abundance can be compared. If a specific PTM is targeted, the same measurement can be made but a likely shift in elution time must be taken into account.

In label-free MS quantitation strategies, no heavy and light peptide ions are present in the same survey spectrum but are found in different LC-MS runs. It is therefore necessary to find the corresponding *m/z* peaks in all samples.

To compensate for complexity of the sample and to increase accuracy, peptide ions can be differentiated from noise by their isotopic envelope (a noise peak would most likely not be accompanied by a number of isotopic peaks), multiple detection in subsequent spectra (noise peaks are not expected to occur consistently), and their elution time which can additionally be taken into account. For the latter, accurate mass and time (AMT) tags need to be established before quantitation is performed (Smith et al. 2002).

MaXIC-Q, a web application, uses extracted ion chromatograms (XIC) and projected ion mass spectra for generic SIL quantitation (Tsou et al. 2009). In their study, the authors also confirmed interesting results using MS/MS measurements.

Another approach in this area, MapQuant, treats LC-MS data similar to images and uses established methods from the field of image processing (Leptos et al. 2006) for quantitation.

Integrated platforms like msInspect (Bellew et al. 2006), XCMS (Smith et al. 2006) and OpenMS (Kohlbacher et al. 2007) extract peptide features from the raw data and integrate them over time in a label-free XIC fashion but offer additional methods for quantitation and downstream statistical analysis as well.

ProteinQuant (Mann et al. 2008) and mzMine (Katajamaa and Oresic 2005) are also software tools able to perform label-free quantitation from LC-MS measurements. Both perform denoising and normalization among other methods. One study showed that ProteinQuant was slightly more accurate than mzMine with both being more accurate than CPAS (Rauch et al. 2006) with w/XPRESS (Li et al. 2003b) for quantitation (Mann et al. 2008). ASAPRatio (Li et al. 2003b) is similar to XPRESS (Han et al. 2001) but offers additional downstream statistical analysis.

Quantifying PTMs with LC-MS

Methods for determining phosphorylation using LC-MS have been proposed as early as 2002 for example by Ruse et al. (Ruse et al. 2002). Another more recent approach quantified the differential phosphorylation between two samples using a label-free strategy employing Decon2LS (Jaitly et al. 2009) and VIPER (Monroe et al. 2007) software packages (Yang et al. 2007).

MaxQuant is also promising for the quantitation of PTMs and the ability to handle large datasets from isotope-labeled high-resolution MS data (Cox and Mann 2008).

Wolf-Yadlin et al. point out that studies involving the quantitation of phosphorylation events are poorly reproducible and offer multiple reaction monitoring (MRM) as an improvement over common MS acquisition protocols (Wolf-Yadlin et al. 2007). MRMer (Martin et al. 2008), built upon msInspect, can quantify data acquired via this method.

Quantitation with LC-MS/MS

Quantitation using single stage MS has the best resolution in regard to re-establishing the LC elution profile since

more measurements can be taken. The LC profile is therefore less serrated when compared to methods employing MS/MS measurements (Fig. 2). MS/MS additionally takes a mass spectrum of fragmented selected precursor ions from the survey spectrum thus adding the possibility to validate identity before quantitation with measurement regimes usually cycling between survey and MS/MS spectra acquisition. For example the five most abundant ions from survey spectra can be automatically selected for fragmentation while a list of already measured ion *m/z* values can be kept to dynamically exclude the repeated measurement. These settings obviously influence the possible quality of quantitation results (Old et al. 2005). An advantage of label-free strategies over labeling strategies is that the former ones can potentially compare an arbitrary number of samples whereas the latter ones are usually limited by the labeling strategy.

Label-free quantitation can be performed directly with data acquired in this fashion. Additional work is needed if labels are introduced although shifts in elution profile due to the label become irrelevant since the identity of the precursor is assigned using MS/MS. Apart from label-free methods those employing labels may also be dependent on MS/MS spectra as for example quantitation using iTRAQ labeling.

When both labeled and unlabeled peptides are measured at the same time and they are co-fragmented by choosing a large enough *m/z* window for fragmentation from the survey spectrum, then the MS/MS spectrum can be quantitated in regards to the difference in their light and heavy fragment ions (Naumann et al. 2007).

Spectral counting is a method to perform quantitation with LC-MS/MS data by simply counting the number of identified spectra (i.e.: the spectra that were successfully assigned a peptide sequence) among samples from different experimental conditions (Washburn et al. 2001; Liu et al. 2004; Gilchrist et al. 2006). In addition the count can be weighted by the total ion current of the counted spectra which can improve its dynamic range (Asara et al. 2008; Allmer 2010). Alternatively, it can be weighted by the reported score of the identification software (Allet et al. 2004); although one study showed that identification scores may not be strongly correlated with protein abundance (Ong and Mann 2005).

Identified MS/MS spectra go into the pool of spectra that can be quantified among samples. Since spectral counting is rather trivial, the counts are often generated by non specialized software such as PeptideProphet (Keller et al. 2002) and MSMAG (Allmer 2010), or by specialized software or scripts (Gao et al. 2003; Old et al. 2005; Ishihama et al. 2005; Huttlin et al. 2007).

Using SIL, very small differences in protein expression can be accessed, which makes it a good platform to assess differential expression of PTMs (Blagoev et al. 2004). Peck and co-workers for instance used iTRAQ to analyze phosphorylation turnover (Nühse et al. 2007) while Chiappetta and co-workers extended the method to quantify nitration using MASCOT (Perkins et al. 1999) for data analysis (Chiappetta et al. 2009).

RelEx (MacCoss et al. 2003), has been used in studies employing stable isotope labeling for quantitation (Wu et al. 2004; Zybailov et al. 2005). Complete labeling can often not be achieved; but partial metabolic labeling has been used in quantitative studies (Huttlin et al. 2007; Whitelegge et al. 2004).

Quantifying PTMs with LC-MS/MS

Spectral counting depends on prior identification of MS/MS spectra which renders them easily extensible to work with arbitrary modifications, such as any type of label or PTM, as long as they are identified by the search algorithm. At least MSMAG (Allmer 2010) in conjunction with 2DB (Allmer et al. 2008) is able to relatively quantify one or more arbitrary labels or PTMs without further modifications.

Most software tools that identify MS/MS spectra are able to allow for static or variable modifications to amino acids. It is, however, computationally extremely expensive, if not impossible, to test for every possible PTM so that modifications must be anticipated before database search. Each additional PTM that is searched for also increases the search space and thus, changes detection specificity and sensitivity that then directly influence quantitation accuracy. Unfortunately, clear rules are missing and empirical evaluation is still paramount (Weckwerth 2008). VEMS, a new player in the field of database search algorithms allows the search for multiple PTMs simultaneously (Matthiesen et al. 2005) which makes it especially suited in this context. Another extension to this approach is using de novo prediction in conjunction with database searching which enabled the identification of a novel protein polymorphism (Hoehenwarter et al. 2008).

Labeling approaches such as SILAC have also been used to quantitate PTMs including the extent of phosphorylation and the affected sites in a protein (Ibarrola et al. 2003).

Quantifying with combinations of MS and MS/MS data

Once the identity of a peptide ion in a survey spectrum has been established via LC-MS/MS, it is possible to use this information to assign the identity to all peaks of that *m/z* in other survey spectra within a suitable time window (the time depends on the reestablished LC profile for the peptide). Thus, more data can be generated than with LC-MS/MS

alone but the confidence in the identity of the acquired data is higher than with LC-MS alone. This can be improved further by using mass spectrometers which are capable of measuring MS and MS/MS spectra in parallel at the cost of making the data analysis more complex (May et al. 2007; Jaffe et al. 2006). May and coworkers for instance developed msInspect/AMT which can aid in analysis of experiments using AMT for quantitation using LC-MS while integrating data from LC-MS/MS results among other more generic functions (May et al. 2007).

MapQuant (Leptos et al. 2006), PEPPeR (Jaffe et al. 2006), SuperHirn (Mueller et al. 2007), and Q-MEND (Andreev et al. 2006) use identified peptides as landmarks for reliable retention time alignment of LC-MS runs but the latter does both in one MS run, whereas the former need two MS runs for the same task. MSight (Palagi et al. 2005) builds an image from an LC-MS run and uses image processing approaches to align multiple LC-MS runs. It is also capable of integrating peptide identifications from LC-MS/MS data.

Quantitation with LC-MS^3

With increasing fragmentation and measurements, the absolute available amount of ions decreases and the measurements become less precise unless the peptide is highly abundant, which still leaves serration problems unaccounted for (Fig. 3). The MS^3 level option is thus mostly used in order to confirm that a PTM is present within the peptide and generally not for its quantitation (Mann and Pandey 2001). This strategy was, for instance, used to identify HOCl-induced modifications to proteins (Mouls et al. 2009).

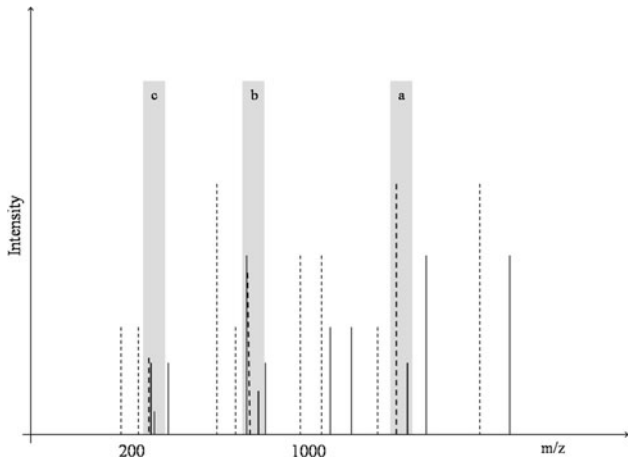

Fig. 3 A highly schematic representation of labeled and label-free precursor ions at three different charged states in a survey spectrum. At charge one (*a*), the heavy and light precursor ions are clearly separate and no interfering peaks are present. At higher charges ($b = 2$, $c = 3$), the heavier and lighter ions come closer together and more ions get into close proximity while the intensity diminishes with charge. All peaks are assumed to be fragment ions; noise is not shown

A third stage of MS is also indicated when the fragmentation is hampered due to the modification. The neutral loss of a phosphorylation, indicated by an abundant fragment ion 98 Da lighter than the precursor ion, can trigger an automatic MS^3 scan in order to get a higher quality fragmentation spectrum from the phosphorylated precursor (Hoffert and Knepper 2008). The loss of the modification can pose problems but can sometimes be prevented by using softer fragmentation methods (Viner et al. 2009) so that PTMs can be mapped to their location in the sequence.

Since MS^3 receives its ions from MS/MS, both spectra should not be used for quantitation at the same time. This redundancy should be avoided by merging MS/MS and MS^3 spectra sets which is done by PhosphoPic (Hoffert et al. 2007), a software that builds upon Sequest (Eng et al. 1994) results, AScore (Beausoleil et al. 2006), and QUOIL (Wang et al. 2006), and integrates these data and derives quantitative results. It has also been successfully applied to quantitate phosphoproteins (Hoffert et al. 2007; Olsen et al. 2006).

Conclusion

Peptide and protein quantitation can be achieved with current MS methods and available analysis software. Quantitation of PTMs is somewhat more challenging and is therefore much less available in current research. Available software could be adapted to allow for quantitation of PTMs although there is still room for improvement in currently available software (Iliuk et al. 2009). New software is, however, abundantly available and it is difficult to choose from the available tools in regard to differences in the employed algorithms. To this end, and since many processing steps have to be made to transform raw measurements into quantitative data, it would be good to develop modules with defined interfaces rather than providing complete applications.

Standard data interchange formats also need to be developed in order to combine the numerous great achievements in the field. Pipelines can then be applied to bundle the modules to create efficient solutions to general or particular problems. A concerted action of all parties, albeit difficult to achieve, is mandatory and will lead to even greater advances in the field. Other researchers seem to think along the same line by emphasizing the modularity of their approach and also extend this idea by integrating methods from related fields such as micro array analysis (Jaffe et al. 2006; Bellew et al. 2006; Mintz et al. 2008). Developing an open toolbox consisting of many different modules works best if the individual

modules are well defined, have been tested by the research community and offer standard interfaces or data exchange formats. Only in this way will modularity be of use for the field of mass spectrometry-based quantitative proteomics and eliminate the need for ever-changing software tools.

Acknowledgments I am indebted to Anne Frary, Ritchie Eanes, Canan Has and Bünyamin Akgül for proofreading this manuscript.

References

Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422:198–207

Allet N, Barrillat N, Baussant T et al (2004) In vitro and in silico processes to identify differentially expressed proteins. Proteomics 4:2333–2351. doi:10.1002/pmic.200300840

Allmer J (2010) Label-free quantitation, an extension to 2DB. Amino Acids 38:1075–1087. doi:10.1007/s00726-009-0317-9

Allmer J, Kuhlgert S, Hippler M (2008) 2DB: a Proteomics database for storage, analysis, presentation, and retrieval of information from mass spectrometric experiments. BMC Bioinformatics 9:302–313

America AHP, Cordewener JHG (2008) Comparative LC-MS: a landscape of peaks and valleys. Proteomics 8:731–749. doi:10.1002/pmic.200700694

Andreev VP, Li L, Rejtar T et al (2006) New algorithm for 15N/14N quantitation with LC-ESI-MS using an LTQ-FT mass spectrometer. J Proteome Res 5:2039–2045. doi:10.1021/pr060105m

Asara JM, Christofk HR, Freimark LM, Cantley LC (2008) A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. Proteomics 8:994–999. doi:10.1002/pmic.200700426

Bantscheff M, Schirle M, Sweetman G et al (2007) Quantitative mass spectrometry in proteomics: a critical review. Anal Bioanal Chem 389:1017–1031

Beausoleil SA, Jedrychowski M, Schwartz D et al (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc Natl Acad Sci USA 101:12130–12135. doi:10.1073/pnas.0404720101

Beausoleil SA, Villén J, Gerber SA et al (2006) A probability-based approach for high-throughput protein phosphorylation analysis and site localization. Nat Biotechnol 24:1285–1292. doi:10.1038/nbt1240

Bellew M, Coram M, Fitzgibbon M et al (2006) A suite of algorithms for the comprehensive analysis of complex protein mixtures using high-resolution LC-MS. Bioinformatics 22:1902–1909. doi:10.1093/bioinformatics/btl276

Blagoev B, Ong SE, Kratchmarova I, Mann M (2004) Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nat Biotechnol 22:1139–1145

Chiappetta G, Corbo C, Palmese A et al (2009) Quantitative identification of protein nitration sites (vol 9, p 1524, 2009). Proteomics 9:3220. doi:10.1002/pmic.200800493

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367–1372. doi:10.1038/nbt.1511

Eng J, McCormack AL, Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectr 5:976–989

Flory MR, Griffin TJ, Martin D, Aebersold R (2002) Advances in quantitative proteomics using stable isotope tags. Trends Biotechnol 20:S23–S29

Gallezot J, Bottlaender MA, Delforge J et al (2008) Quantification of cerebral nicotinic acetylcholine receptors by PET using 2-[18F]fluoro-A-85380 and the multiinjection approach. J Cereb Blood Flow Metab 28:172–189. doi:10.1038/sj.jcbfm.9600505

Gao J, Opiteck GJ, Friedrichs MS et al (2003) Changes in the protein expression of yeast as a function of carbon source. J Proteome Res 2:643–649

Gerber SA, Rush J, Stemman O et al (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci USA 100:6940–6945. doi:10.1073/pnas.0832254100

Gilchrist A, Au CE, Hiding J et al (2006) Quantitative proteomics analysis of the secretory pathway. Cell 127:1265–1281. doi:10.1016/j.cell.2006.10.036

Guan S, Burlingame AL (2009) Data processing algorithms for analysis of high resolution MSMS spectra of peptides with complex patterns of posttranslational modifications. Mol Cell Proteomics. doi:10.1074/mcp.M900431-MCP200

Gygi SP, Rist B, Gerber SA et al (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994–999. doi:10.1038/13690

Han D, Eng J, Zhou H, Aebersold R (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat Biotechnol 19:946–951

Hoehenwarter W, van Dongen JT, Wienkoop S et al (2008) A rapid approach for phenotype-screening and database independent detection of cSNP/protein polymorphism using mass accuracy precursor alignment. Proteomics 8:4214–4225. doi:10.1002/pmic.200701047

Hoffert J, Knepper M (2008) Taking aim at shotgun phosphoproteomics. Anal Biochem 375:1–10. doi:10.1016/j.ab.2007.11.023

Hoffert JD, Wang G, Pisitkun T et al (2007) An automated platform for analysis of phosphoproteomic datasets: application to kidney collecting duct phosphoproteins. J Proteome Res 6:3501–3508. doi:10.1021/pr0701153

Hunter T (1998) The Croonian Lecture 1997. The phosphorylation of proteins on tyrosine: its role in cell growth and disease. Philos T Roy Soc B 353:583–605. doi:10.1098/rstb.1998.0228

Huttlin EL, Hegeman AD, Harms AC, Sussman MR (2007) Comparison of full versus partial metabolic labeling for quantitative proteomics analysis in *Arabidopsis thaliana*. Mol Cell Proteomics 6:860–881. doi:10.1074/mcp.M600347-MCP200

Ibarrola N, Kalume DE, Gronborg M et al (2003) A proteomic approach for quantitation of phosphorylation using stable isotope labeling in cell culture. Anal Chem 75:6043–6049. doi:10.1021/ac034931f

Iliuk A, Galan J, Tao W (2009) Playing tag with quantitative proteomics. Anal Bioanal Chem 393:503–513. doi:10.1007/s00216-008-2386-0

Ishihama Y, Oda Y, Tabata T et al (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 4:1265–1272. doi:10.1074/mcp.M500061-MCP200

Jaffe JD, Mani DR, Leptos KC et al (2006) PEPPeR, a platform for experimental proteomic pattern recognition. Mol Cell Proteomics 5:1927–1941. doi:10.1074/mcp.M600222-MCP200

Jaitly N, Mayampurath A, Littlefield K et al (2009) Decon2LS: an open-source software package for automated processing and visualization of high resolution mass spectrometry data. BMC Bioinformatics 10:87. doi:10.1186/1471-2105-10-87

Julka S, Regnier FE (2005) Recent advancements in differential proteomics based on stable isotope coding. Brief Funct Genomic Proteomic 4:158–177

Jungblut P, Thiede B, Zimny-Arndt U et al (1996) Resolution power of two-dimensional electrophoresis and identification of proteins

from gels. Electrophoresis 17:839–847. doi:10.1002/elps.1150170505
Kapp E, Schutz F, Connolly L et al (2005) An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: sensitivity and specificity analysis. Proteomics 5:3475–3490
Katajamaa M, Oresic M (2005) Processing methods for differential analysis of LC/MS profile data. BMC Bioinformatics 6:179. doi:10.1186/1471-2105-6-179
Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 74:5383–5392
Keller A, Eng J, Zhang N et al (2005) A uniform proteomics MS/MS analysis platform utilizing open XML file formats. Mol Syst Biol 1:2005.0017. doi:10.1038/msb4100024
Kohlbacher O, Reinert K, Gröpl C et al (2007) TOPP–the OpenMS proteomics pipeline. Bioinformatics 23:e191–e197. doi:10.1093/bioinformatics/btl299
Lander ES, Linton LM, Birren B et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921. doi:10.1038/35057062
Leptos KC, Sarracino DA, Jaffe JD et al (2006) MapQuant: open-source software for large-scale protein quantification. Proteomics 6:1770–1782. doi:10.1002/pmic.200500201
Li J, Steen H, Gygi SP (2003a) Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents: the yeast salinity stress response. Mol Cell Proteomics 2:1198–1204. doi:10.1074/mcp.M300070-MCP200
Li X, Zhang H, Ranish JA, Aebersold R (2003b) Automated statistical analysis of protein abundance ratios from data generated by stable-isotope dilution and tandem mass spectrometry. Anal Chem 75:6648–6657. doi:10.1021/ac034633i
Liu H, Sadygov RG, Yates JR (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76:4193–4201. doi:10.1021/ac0498563
Lu B, Ruse C, Xu T et al (2007) Automatic validation of phosphopeptide identifications from tandem mass spectra. Anal Chem 79:1301–1310. doi:10.1021/ac061334v
MacCoss MJ, Wu CC, Liu H et al (2003) A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. Anal Chem 75:6912–6921. doi:10.1021/ac034790h
Mann M, Pandey A (2001) Use of mass spectrometry-derived data to annotate nucleotide and protein sequence databases. Trends Biochem Sci 26:54–61
Mann M, Wilm M (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal Chem 66:4390–4399
Mann M, Hojrup P, Roepstorff P (1993) Use of mass spectrometric molecular weight information to identify proteins in sequence databases. Biol Mass Spectrom 22:338–345
Mann B, Madera M, Sheng Q et al (2008) ProteinQuant Suite: a bundle of automated software tools for label-free quantitative proteomics. Rapid Commun Mass Spectrom 22:3823–3834
Martin DB, Holzman T, May D et al (2008) MRMer, an interactive open source and cross-platform system for data extraction and visualization of multiple reaction monitoring experiments. Mol Cell Proteomics 7:2270–2278. doi:10.1074/mcp.M700504-MCP200
Matthiesen R, Trelle MB, Højrup P et al (2005) VEMS 3.0: algorithms and computational tools for tandem mass spectrometry based identification of post-translational modifications in proteins. J Proteome Res 4:2338–2347. doi:10.1021/pr050264q
May D, Fitzgibbon M, Liu Y et al (2007) A platform for accurate mass and time analyses of mass spectrometry data. J Proteome Res 6:2685–2694. doi:10.1021/pr070146y
May D, Liu Y, Law W et al (2008) Peptide sequence confidence in accurate mass and time analysis and its use in complex proteomics experiments. J Proteome Res 7:5148–5156
May D, Law W, Fitzgibbon M et al (2009) Software platform for rapidly creating computational tools for mass spectrometry-based proteomics. J Proteome Res 8:3212–3217. doi:10.1021/pr900169w
Mintz M, Vanderver A, Brown K et al (2008) Time series proteome profiling to study endoplasmic reticulum stress response. J Proteome Res 7:2435–2444
Monroe ME, Tolić N, Jaitly N et al (2007) VIPER: an advanced software package to support high-throughput LC-MS peptide identification. Bioinformatics 23:2021–2023. doi:10.1093/bioinformatics/btm281
Mouls L, Silajdzic E, Haroune N et al (2009) Development of novel mass spectrometric methods for identifying HOCl-induced modifications to proteins. Proteomics 9:1617–1631. doi:10.1002/pmic.200800391
Mueller LN, Rinner O, Schmidt A et al (2007) SuperHirn—a novel tool for high resolution LC-MS-based peptide/protein profiling. Proteomics 7:3470–3480. doi:10.1002/pmic.200700057
Nair KS, Jaleel A, Asmann YW et al (2004) Proteomic research: potential opportunities for clinical and physiological investigators. Am J Physiol Endocrinol Metab 286:E863–E874. doi:10.1152/ajpendo.00370.2003
Naumann B, Busch A, Allmer J et al (2007) Comparative quantitative proteomics to investigate the remodeling of bioenergetic pathways under iron deficiency in *Chlamydomonas reinhardtii*. Proteomics 7:3964–3979
Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 75:4646–4658
Nühse TS, Bottrill AR, Jones AME, Peck SC (2007) Quantitative phosphoproteomic analysis of plasma membrane proteins reveals regulatory mechanisms of plant innate immune responses. Plant J 51:931–940. doi:10.1111/j.1365-313X.2007.03192.x
O'Donovan C, Apweiler R, Bairoch A (2001) The human proteomics initiative (HPI). Trends Biotechnol 19:178–181
Oda Y, Huang K, Cross FR et al (1999) Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci USA 96:6591–6596
Oeljeklaus S, Meyer HE, Warscheid B (2009) Advancements in plant proteomics using quantitative mass spectrometry. J Proteomics 72:545–554. doi:10.1016/j.jprot.2008.11.008
Old WM, Meyer-Arendt K, Aveline-Wolf L et al (2005) Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics 4:1487–1502
Olsen JV, Blagoev B, Gnad F et al (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell 127:635–648. doi:10.1016/j.cell.2006.09.026
Ong SE, Mann M (2005) Mass spectrometry-based proteomics turns quantitative. Nat Chem Biol 1:252–262. doi:10.1038/nchembio736
Ong SE, Blagoev B, Kratchmarova I et al (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376–386
Palagi PM, Walther D, Quadroni M et al (2005) MSight: an image analysis software for liquid chromatography-mass spectrometry. Proteomics 5:2381–2384. doi:10.1002/pmic.200401244
Pan C, Kora G, McDonald WH, Tabb DL, VerBerkmoes NC, Hurst GB, Pelletier DA, Samatova NF, Hettich RL (2006) ProRata: a quantitative proteomics program for accurate protein abundance ratio estimation with confidence interval evaluation. Anal Chem 78:7121–7131. doi:10.1021/ac060654b
Pedrioli PG, Eng JK, Hubley R et al (2004) A common open representation of mass spectrometry data and its application to proteomics research. Nat Biotechnol 22:1459–1466

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567

Rauch A, Bellew M, Eng J et al (2006) Computational proteomics analysis system (CPAS): an extensible, open-source analytic system for evaluating and publishing proteomic data and high throughput biological experiments. J Proteome Res 5:112–121

Ross PL, Huang YN, Marchese JN et al (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169

Ruse CI, Willard B, Jin JP et al (2002) Quantitative dynamics of site-specific protein phosphorylation determined using liquid chromatography electrospray ionization mass spectrometry. Anal Chem 74:1658–1664. doi:10.1021/ac0157122

Schluter H, Apweiler R, Holzhutter H, Jungblut P (2009) Finding one's way in proteomics: a protein species nomenclature. Chem Cent J 3:11. doi:10.1186/1752-153X-3-11

Schmidt A, Kellermann J, Lottspeich F (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. Proteomics 5:4–15. doi:10.1002/pmic.200400873

Shadforth I, Crowther D, Bessant C (2005) Protein and peptide identification algorithms using MS for use in high-throughput, automated pipelines. Proteomics 5:4082–4095

Shevchenko A, Jensen ON, Podtelejnikov AV et al (1996) Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. Proc Natl Acad Sci USA 93:14440–14445

Shinkawa T, Taoka M, Yamauchi Y et al (2005) STEM: a software tool for large-scale proteomic data analyses. J Proteome Res 4:1826–1831. doi:10.1021/pr050167x

Smith RD, Anderson GA, Lipton MS et al (2002) An accurate mass tag strategy for quantitative and high-throughput proteome measurements. Proteomics 2:513–523

Smith CA, Want EJ, O'Maille G et al (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem 78:779–787. doi:10.1021/ac051437y

Steen H, Jebanathirajah JA, Springer M, Kirschner MW (2005) Stable isotope-free relative and absolute quantitation of protein phosphorylation stoichiometry by MS. Proc Natl Acad Sci USA 102:3948–3953. doi:10.1073/pnas.0409536102

Tabb DL, McDonald WH, Yates JR (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J Proteome Res 1:21–26

Tsikas D, Caidahl K (2005) Recent methodological advances in the mass spectrometric analysis of free and protein-associated 3-nitrotyrosine in human plasma. J Chromatogr B 814:1–9. doi:10.1016/j.jchromb.2004.10.003

Tsou CC, Tsui YH, Yian YH et al (2009) MaXIC-Q Web: a fully automated web service using statistical and computational methods for protein quantitation based on stable isotope labeling and LC-MS. Nucl Acids Res 37:W661–W669. doi:10.1093/nar/gkp476

Venable JD, Dong M, Wohlschlegel J et al (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. Nat Methods 1:39–45. doi:10.1038/nmeth705

Viner RI, Zhang T, Second T, Zabrouskov V (2009) Quantification of post-translationally modified peptides of bovine α-crystallin using tandem mass tags and electron transfer dissociation. J Proteomics 72:874–885. doi:10.1016/j.jprot.2009.02.005

Vissers JPC, Pons S, Hulin A et al (2009) The use of proteome similarity for the qualitative and quantitative profiling of reperfused myocardium. J Chromatogr B 877:1317–1326. doi:10.1016/j.jchromb.2008.10.024

Wang G, Wu WW, Zeng W et al (2006) Label-free protein quantification using LC-coupled ion trap or FT mass spectrometry: reproducibility, linearity, and application with complex proteomes. J Proteome Res 5:1214–1223. doi:10.1021/pr050406g

Wang F, Ye M, Dong J et al (2008) Improvement of performance in label-free quantitative proteome analysis with monolithic electrospray ionization emitter. J Sep Sci 31:2589–2597

Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19:242–247. doi:10.1038/85686

Weckwerth W (2008) Integration of metabolomics and proteomics in molecular plant physiology—coping with the complexity by data-dimensionality reduction. Physiol Plantarum 132:176–189. doi:10.1111/j.1399-3054.2007.01011.x

Whitelegge JP, Katz JE, Pihakari KA et al (2004) Subtle modification of isotope ratio proteomics; an integrated strategy for expression proteomics. Phytochemistry 65:1507–1515. doi:10.1016/j.phytochem.2004.05.018

Wolf-Yadlin A, Hautaniemi S, Lauffenburger DA, White FM (2007) Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. Proc Natl Acad Sci USA 104:5860–5865. doi:10.1073/pnas.0608638104

Wu CC, MacCoss MJ, Howell KE et al (2004) Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. Anal Chem 76:4951–4959. doi:10.1021/ac049208j

Yan JX, Packer NH, Gooley AA, Williams KL (1998) Protein phosphorylation: technologies for the identification of phosphoamino acids. J Chromatogr A 808:23–41

Yang F, Jaitly N, Jayachandran H et al (2007) Applying a targeted label-free approach using LC-MS AMT tags to evaluate changes in protein phosphorylation following phosphatase inhibition. J Proteome Res 6:4489–4497

Zhang R, Sioma CS, Wang S, Regnier FE (2001) Fractionation of isotopically labeled peptides in quantitative proteomics. Anal Chem 73:5142–5149

Zybailov B, Coleman MK, Florens L, Washburn MP (2005) Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. Anal Chem 77:6218–6224. doi:10.1021/ac050846r