

Quantitative mass spectrometry in proteomics: a critical review

Marcus Bantscheff · Markus Schirle ·
Gavain Sweetman · Jens Rick · Bernhard Kuster

Received: 30 April 2007 / Revised: 25 June 2007 / Accepted: 29 June 2007 / Published online: 1 August 2007
© Springer-Verlag 2007

Abstract The quantification of differences between two or more physiological states of a biological system is among the most important but also most challenging technical tasks in proteomics. In addition to the classical methods of differential protein gel or blot staining by dyes and fluorophores, mass-spectrometry-based quantification methods have gained increasing popularity over the past five years. Most of these methods employ differential stable isotope labeling to create a specific mass tag that can be recognized by a mass spectrometer and at the same time provide the basis for quantification. These mass tags can be introduced into proteins or peptides (i) metabolically, (ii) by chemical means, (iii) enzymatically, or (iv) provided by spiked synthetic peptide standards. In contrast, label-free quantification approaches aim to correlate the mass spectrometric signal of intact proteolytic peptides or the number of peptide sequencing events with the relative or absolute protein quantity directly. In this review, we critically examine the more commonly used quantitative mass spectrometry methods for their individual merits and discuss challenges in arriving at meaningful interpretations of quantitative proteomic data.

Keywords Quantitative proteomics · Mass spectrometry · Stable isotope labeling

Introduction

There is a clear trend in the life sciences towards the study of biological entities at the system level. This requires analytical tools that can identify the component parts of the system and measure their responses to a changing environment. Towards this end, a multitude of transcriptomic, proteomic, and metabolomic profiling technologies have been developed, and proteomics in particular is continuing to evolve rapidly. Still, out of the many thousand proteomic studies published to date, only a small minority has attempted to provide a comprehensive quantitative description of the biological system under investigation. Despite the phenomenal impact of mass spectrometry and peptide separation techniques on proteomics, the identification and quantification of all of the proteins in a biological system is still an unmet technical challenge (Fig. 1). While for unicellular organisms proteomic coverage of the genome has been occasionally achieved beyond 50%, coverage for higher organisms rarely exceeds 10%. For protein quantification, these figures are significantly smaller due to the fact that the data quality, in terms of information content, required for quantification by far exceeds that for protein identification.

The classical proteomic quantification methods utilizing dyes, fluorophores, or radioactivity have provided very good sensitivity, linearity, and dynamic range, but they suffer from two important shortcomings: first, they require high-resolution protein separation typically provided by 2D gels, which limits their applicability to abundant and soluble proteins; and second, they do not reveal the identity of the underlying protein. Both of these problems are overcome by modern LC-MS/MS techniques. However, mass spectrometry is not inherently quantitative because proteolytic peptides exhibit a wide range of physicochemical properties

M. Bantscheff · M. Schirle · G. Sweetman · J. Rick ·
B. Kuster (✉)
Cellzome AG,
Meyerohofstrasse 1,
69254 Heidelberg, Germany
e-mail: bernhard.kuester@cellzome.com

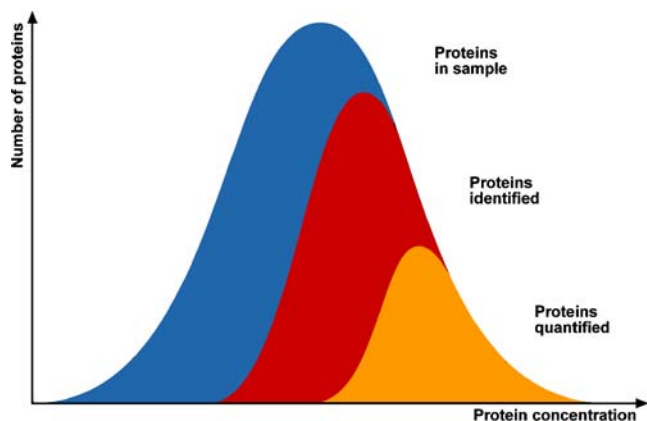


Fig. 1 Schematic representation of the fraction of a proteome that can be identified or quantified by mass-spectrometry-based proteomics. Cellular proteins span a wide range of expression and current mass spectrometric technologies typically sample only a fraction of all the proteins present in a sample. Due to limited data quality, only a fraction of all identified proteins can also be reliably quantified

such as size, charge, hydrophobicity, etc. which lead to large differences in mass spectrometric response. For accurate quantification, it is therefore generally required to compare each individual peptide between experiments. In most

proteomic workflows, this can technically be achieved in a number of ways (Fig. 2). One major approach is based on stable isotope dilution theory which states that a stable isotope-labeled peptide is chemically identical to its native counterpart and therefore the two peptides also behave identically during chromatographic and/or mass spectrometric analysis. Given that a mass spectrometer can recognize the mass difference between the labeled and unlabeled forms of a peptide, quantification is achieved by comparing their respective signal intensities. Stable isotope labeling was introduced into proteomics in 1999 by three independent laboratories [1–3] and has since been adopted widely in the field (for earlier reviews see, e.g., Refs. [4–11]). Isotope labels can be introduced as an internal standard into amino acids (i) metabolically, (ii) chemically, or (iii) enzymatically or, alternatively, as an external standard using spiked synthetic peptides [11]. More recently, alternative strategies—often referred to as label-free quantification—have emerged. Label-free methods aim to compare two or more experiments by (i) comparing the direct mass spectrometric signal intensity for any given peptide or (ii) using the number of acquired spectra

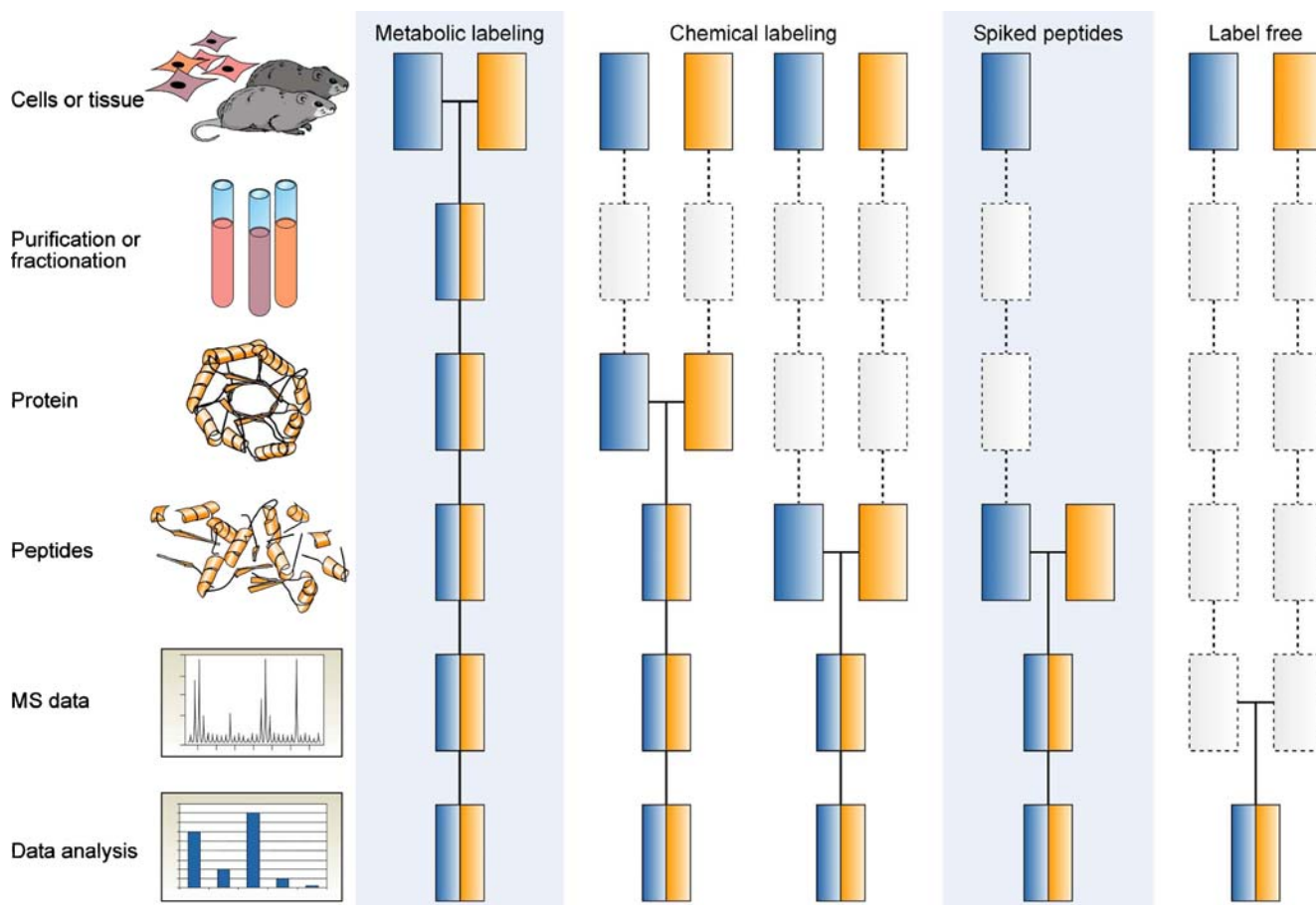


Fig. 2 Common quantitative mass spectrometry workflows. Boxes in blue and yellow represent two experimental conditions. Horizontal lines indicate when samples are combined. Dashed lines indicate

points at which experimental variation and thus quantification errors can occur (adapted with permission from Ref. [11])

matching to a peptide/protein as an indicator for their respective amounts in a given sample. As we will discuss in the following sections, all of the mass-spectrometry-based quantification methods have their particular strengths and weaknesses (Table 1) but they are beginning to mature to an extent that they can be meaningfully applied to the study of biological systems on a proteomic scale. In contrast, the statistical treatment and subsequent interpretation of quantitative proteomic data are still in their infancy, as the field is only beginning to experience the particular challenges associated with transforming qualitative protein identification and post-translational modification data into reliable quantitative information.

Metabolic labeling

The earliest possible point for introducing a stable isotope signature into proteins is by metabolic labeling during cell growth and division. Initially described for total labeling of bacteria using ^{15}N -enriched cell culture medium [2], it has gained wider popularity in the form of the stable isotope labeling by amino acids in cell culture (SILAC) approach introduced by Mann and co-workers in 2002 [12]. In the

most commonly used implementation of the method, the medium contains $^{13}\text{C}_6$ -arginine and $^{13}\text{C}_6$ -lysine which ensures that all tryptic cleavage products of a protein (except for the very C-terminal peptide) carry at least one labeled amino acid resulting in a constant mass increment over the non-labeled counterpart. Protein identification is based on fragmentation spectra of at least one of the co-eluting ‘heavy’ and ‘light’ peptides and relative quantitation is performed by comparing the intensities of isotope clusters of the intact peptide in the survey spectrum. In contrast to full metabolic protein labeling by ^{15}N , the number of incorporated labels in SILAC is defined and not dependent on the peptide sequence thus facilitating data analysis. The main advantage of all metabolic labeling strategies is that the differentially treated samples can be combined at the level of intact cells. This excludes all sources of quantification error introduced by biochemical and mass spectrometric procedures as these will affect both protein populations in the same way. Despite a number of cases that demonstrate the feasibility of total ^{15}N metabolic protein labeling of higher organisms in vivo such as *C. elegans*, *Drosophila melanogaster* [13], rat [14], or plants [15], it is neither possible nor practical to apply this strategy routinely. The cost and time required for creating and

Table 1 Characteristics and applications of quantitative mass spectrometry methods

	Application	Accuracy (process)	Quantitative proteome coverage	Linear dynamic range ^a
Metabolic protein labeling	Complex biochemical workflows Comparison of 2–3 states Cell culture systems only	+++	++	1–2 logs
Chemical protein labeling (MS)	Medium to complex biochemical workflows Comparison of 2–3 states	+++	++	1–2 logs
Chemical peptide labeling (MS)	Medium complexity biochemical workflows Comparison of 2–3 states	++	++	2 logs
Chemical peptide labeling (MS/MS)	Medium complexity biochemical workflows Comparison of 2–8 states	++	++	2 logs
Enzymatic labeling (MS)	Medium complexity biochemical workflows Comparison of 2 states	++	++	1–2 logs
Spiked peptides	Medium complexity biochemical workflows Targeted analysis of few proteins	++	+	2 logs
Label free (ion intensity)	Simple biochemical workflows Whole proteome analysis Comparison of multiple states	+	+++	2–3 logs
Label free (spectrum counting)	Simple biochemical workflows Whole proteome analysis Comparison of multiple states	+	+++	2–3 logs

^a In MRM mode, dynamic range may be extended to 4–5 logs [65]

maintaining these systems is often incommensurate with the value of the information provided. As a result, the main application of metabolic labeling in higher eukaryotes to date is SILAC in immortalized cell lines. Protein labeling in excess of 90% is often achieved by 6–8 passages in medium supplemented with heavy amino acids [12]. While many cell lines can be converted quite readily, some do require special attention. For example, some cell lines require careful titration of the amount of arginine in the medium in order to prevent metabolic conversion of excess arginine into proline which in turn complicates data analysis [16]. Cell lines that are sensitive to changes in media composition or are otherwise difficult to grow or maintain in culture may not be amenable to metabolic labeling at all. A further limitation of metabolic labeling is the restricted number of available labels. For SILAC, a maximum of three conditions can be compared in one experiment (unlabeled, $^{13}\text{C}_6$, and $^{13}\text{C}_6$ $^{15}\text{N}_4$ -labeled amino acids) which, albeit possible, complicates the analysis of, e.g., time-course experiments. Because of the early combination of samples, metabolic labeling and SILAC in particular is probably the most accurate quantitative MS method in terms of overall experimental process. This makes it particularly suitable for assessing relatively small changes in protein levels or those of post-translational modifications [17–19]. For the latter, it should be noted though, that quantification on the peptide level is far from trivial because all information is derived from a single or a few observations.

Protein and peptide labeling

Post-biosynthetic labeling of proteins and peptides is performed by chemical or enzymatic derivatization *in vitro*. An elegant and specific way to introduce an isotope label into peptides is the use of trypsin- or Glu-C-catalyzed incorporation of ^{18}O during protein digestion [20, 21]. This has originally been employed to aid *de novo* sequencing of peptides by mass spectrometry [22] but has recently also been applied to quantitative proteomic applications (for a recent review see Ref. [23]). Enzymatic labeling can be performed either during proteolytic digestion or, more commonly, after proteolysis in a second incubation step with the protease. Incorporation of ^{18}O into C-termini of peptides results in a mass shift of 2 Da per ^{18}O atom. While trypsin and Glu-C introduce two oxygen atoms resulting in a 4 Da mass shift which is generally sufficient for differentiation of isotopomers, Lys-N and other enzymes incorporate only one ^{18}O molecule and should therefore be avoided [24]. Acid- and base-catalyzed back-exchange with concomitant loss of the isotope label can occur at extreme pH values [25], but under the mild acidic

conditions typically employed for ESI- and MALDI-MS ^{18}O -containing carboxyl groups of peptides are sufficiently stable. Because peptides are enzymatically labeled, artifacts (i.e., side reactions) common to chemical labeling can be avoided. A practical disadvantage is that full labeling is rarely achieved and that different peptides incorporate the label at different rates which complicates data analysis [26, 27].

In principle, every reactive amino acid side chain can be used to incorporate an isotope-coded mass tag by chemical means (reviewed by Ong and Mann [11]). In practice, however, side chains of lysine and cysteine are primarily used for this purpose. In their pioneering work Gygi et al. [1] developed the isotope-coded affinity tag (ICAT) approach in which cysteine residues are specifically derivatized with a reagent containing either zero or eight deuterium atoms as well as a biotin group for affinity purification of cysteine-derivatized peptides and subsequent MS analysis. Following the initial success of the ICAT approach, several variations on this chemical reagent class emerged to improve, e.g., recovery of labeled peptides or chromatographic properties [28–31]. Other thiol-specific reagents typically contain halogen-substituted carboxylic acids or amides [32–35] or employ the Michael-type addition reaction to carbonyl groups (e.g., maleimide esters and vinylpyridine) [36, 37]. As cysteine is a rare amino acid, ICAT and related methods significantly reduce the complexity of the peptide mixture which can be advantageous when highly complex samples are analyzed. However, ICAT is obviously not suitable for quantifying the significant number of proteins that do not contain any (or a few) cysteine residues and is of limited use for analysis of post-translational modifications and splice isoforms. Despite these drawbacks, ICAT and similar approaches will continue to be useful in a number of broad (e.g., body fluid) or targeted (e.g., cysteine protease) analyses.

Another group of labeling reagents targets the peptide N-terminus and the epsilon-amino group of lysine residues. Most of the time, this is realized via the very specific *N*-hydroxysuccinimide (NHS) chemistry or other active esters and acid anhydrides as in, e.g., the isotope-coded protein label (ICPL) [38], isotope tags for relative and absolute quantification (iTRAQ) [39], tandem mass tags (TMT) [40], and acetic/succinic anhydride [41–44]. Isocyanates or isothiocyanates have also been employed, albeit to a lesser extent [45, 46]. In recent studies, formaldehyde has been used for methylation of lysine residues via Schiff base formation and subsequent reduction by cyanoborohydride [47–49]. This reaction is very fast, very specific, and very cheap. However, a sufficiently large mass shift between ‘heavy’ and ‘light’ labeled peptides can only be achieved with deuterated formalde-

hyde which in turn leads to partial LC separation of labeled and non-labeled peptides, thus complicating data analysis (discussed below).

In most of the aforementioned chemical modification techniques, relative quantification is achieved by integration of MS signal over isotopomers of 'heavy' and 'light' labeled peptides in survey spectra. Isobaric mass tagging initially introduced by Thompson and co-workers [40] differs from this concept by introducing tags that initially produce isobaric labeled peptides which precisely co-migrate in liquid chromatography separations. Only upon peptide fragmentation are the different tags distinguished by the mass spectrometer. This permits the simultaneous determination of both identity and relative abundance of peptide pairs in tandem-mass spectra. The commercially available iTRAQ reagent [39] provides a further refinement of this approach, allowing multiplexed quantitation of up to eight samples. This has turned out to be particularly useful for following biological systems over multiple time points or, more generally, for comparing multiple treatments in the same experiment.

Carboxylic acids in side chains of glutamic and aspartic acid residues as well as the C-termini of polypeptide chains can be isotopically labeled by esterification using deuterated alcohols [50, 51]. This reaction is particularly attractive for the quantification of phospho-peptides because esterification has been shown to reduce binding of acidic peptides to ion metal chelate affinity chromatography (IMAC) columns, thus improving the specificity of this enrichment procedure [52]. Other, more tailored labeling techniques have been developed, e.g., for quantification of phosphorylated and glycosylated peptides. For the former, β -elimination of phosphoric acid followed by Michael addition using, e.g., ethanedithiol derivatives is typically employed [53–56]. For glycopeptides, hydrazide chemistry replaces the carbohydrate moiety with a labeled chemical group [57].

Broadly speaking, the chemical properties of amino acid side chains of proteins and peptides chains are rather similar. Consequently, almost all chemical labeling methods may also be applied to intact proteins. For example, the ICPL reagent [38] has been employed for N-terminal peptide labeling as well as lysine side chain labeling of intact proteins. A similar protocol has been described for iTRAQ [58]. In most cases, full protein denaturation improves labeling results but care has to be taken to avoid protein precipitation (by, e.g., the use of charged reagents). Labeling of intact proteins can be quite advantageous since it allows for further protein separation steps on the combined samples. This may facilitate characterization of protein isoforms by, e.g., 2D gel electrophoresis [38]. However, there are two important caveats to protein labeling: one is that trypsin does not cleave modified lysine

residues, which leads to significantly longer peptides that generally are more difficult to identify by MS; second, very high labeling efficiencies are required in case further protein separation is desired prior to MS analysis, since incomplete labeling impairs resolving power achievable with, e.g., 1D and 2D gel electrophoresis. A general drawback with all chemical labeling approaches is that they are prone to side reactions that can lead to unexpected products and which may adversely influence quantification results.

Absolute quantification using internal standards

The use of isotope-labeled synthetic standards has a long history in quantitative mass spectrometry. Originally described in the early 1980s [59], it is now becoming more broadly applied as a method commonly known as AQUA (absolute quantification of proteins) [60]. In the simplest case, absolute quantification can be achieved by the addition of a known quantity of a stable isotope-labeled standard peptide to a protein digest and subsequent comparison of the mass spectrometric signal to the endogenous peptide in the sample. Unlike in metabolic labeling, where relative quantitative information is acquired for a large number of the proteins present in a mixture, the addition of synthetic peptides to a proteome digest focuses on the determination of the quantity of one or a few particular proteins of interest. This approach is attractive for studies aimed at, e.g., the analysis and validation of potential biomarkers in a large number of clinical samples [61] or at measuring the levels of particular peptide modifications such as ubiquitinylation [62].

The approach has been refined by constructing synthetic genes that express concatenated standard peptides which upon tryptic digestion either provide multiple peptides of the same protein for quantification or quantification standards for a group of proteins of interest [63]. Not only does the provision of multiple peptides increase confidence in quantification, the synthetic protein can also be added earlier in the process than individual peptides, thus controlling any potential bias encountered during protein digestion. One notable example of following the synthetic gene strategy is the determination of the stoichiometry of the eight-membered eIF2B-eIF2 protein complex [64].

Given that tryptic digests of entire proteomes are very complex mixtures, and that most mass spectrometers have a rather limited dynamic detection range, there are a number of limitations to the AQUA approach. One practical drawback is that one has to 'guess' how much of the labeled standard should be added to a sample. This amount may be different for all proteins of interest as their expression levels (used here in the sense of protein

abundance rather than protein synthesis) may differ greatly within a sample. Another limitation is the specificity of the spiked standard as there are likely multiple isobaric peptides present in the mixture. Both of these issues can be greatly improved by a method called multiple reaction monitoring (MRM) [62] in which the (triple quadrupole) mass spectrometer monitors both the intact peptide mass and one or more specific fragment ions of that peptide over the course of an LC-MS experiment. The combination of retention time, peptide mass, and fragment mass practically eliminates ambiguities in peptide assignments and extends the quantification range to 4–5 orders of magnitude [65]. Obviously, the choice of synthetic peptide standard is important and is mostly determined empirically. However, recent data suggest that it is possible to predict which of a protein's tryptic peptides will be most frequently observed for a given proteomic platform and thus would be a suitable quantification standard [66]. Despite the ability to calculate protein amounts from an AQUA experiment, there are still question marks as to how absolute these values are as any sample manipulation prior to adding the synthetic standard may bias the results (losses or enrichment). Consequently, the amount of a protein in an experiment determined by AQUA may not reflect the true expression levels of this protein in a cell.

LC-MS/MS analysis of stable isotope labeled peptides

As described above, quantitation based on stable isotope labeling can be achieved by signal integration in survey MS spectra (e.g., SILAC) or tandem MS spectra (e.g., iTRAQ). For both approaches, several points have to be considered in the design and analysis of an experiment. Although the assumption that stable isotope labeling does not alter the physicochemical properties of a peptide is generally valid, it has been observed that deuterated peptides show small but significant retention time differences in reversed-phase

HPLC compared to their non-deuterated counterparts [67]. This complicates data analysis because the relative quantities of the two peptide species cannot be determined accurately from one spectrum but requires integration across the chromatographic time scale. Retention time shifts are far less pronounced for labels such as ^{13}C , ^{15}N , or ^{18}O isotopes [68], so that the additional signal integration step over retention time can generally be omitted.

Another requirement for any stable isotope labeling approach is that the heavy label can be clearly distinguished from the unlabeled peptide or any other unrelated ion species (Fig. 3a). For quantification in survey MS spectra, it is essential that the mass shift introduced by the label is at least 4 Da in order to distinguish the isotopomer clusters of the labeled and unlabeled forms of the peptide. As isotopomer clusters increase in width with increasing peptide mass, the application of labeling methods such as methylation and enzymatic ^{18}O labeling becomes limited for larger peptides. Reporter ions used for quantification in tandem MS spectra should be designed such that interference by ordinary peptide fragments is minimal. For the iTRAQ label, the m/z region of 114–117 was chosen for this reason. Still, some interferences have been identified (notably the 116.1 Da $y(1)$ fragment ion of peptides containing a C-terminal proline residue [69]) and these data points have to be carefully removed in the data analysis process.

A further parameter impacting accuracy and dynamic range of quantification is the mass spectrometric detection system itself. In survey MS spectra, the definition of very low and very strong signals can be problematic. At very low signal, peptide ions are often difficult to distinguish from background noise (Fig. 3b) and for very strong signals, the detector may become saturated (Fig. 3c). In practice, saturation is more often observed for quadrupole TOF instruments than ion traps because these latter devices can control the number of ions before detection [70]. In any case, the relatively recent introduction of high-resolution/high mass accuracy mass spectrometers in proteomics has

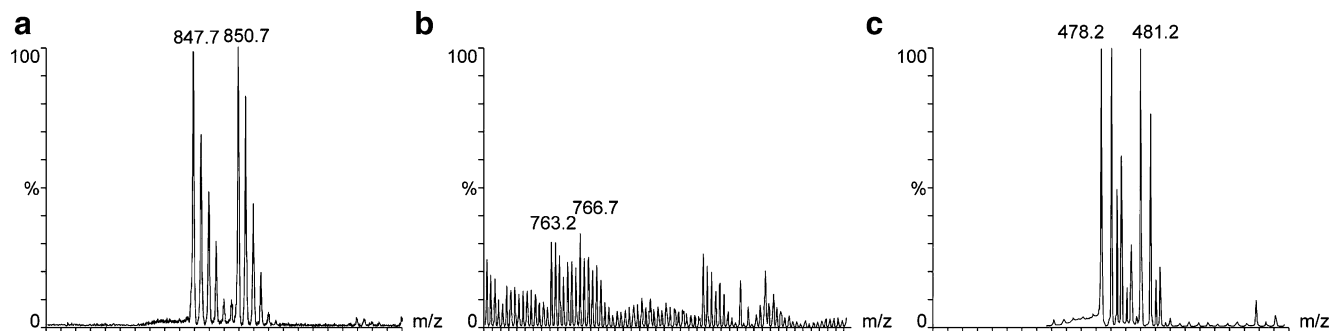


Fig. 3 Examples illustrating mass spectral features relevant for quantification. **a** Example of a SILAC-labeled peptide pair suitable for quantification. The spectra displays the characteristic 6 Da ($3 m/z$) mass difference between light and heavy forms of the peptide, good signal to noise ratio and no interfering signals. Signal intensities

indicate a 1:1 abundance ratio. **b** Example of a peptide and other interfering signals with signal to noise ratios too low for reliable quantification. **c** Example of a peptide signal saturating the detector and thus distorting the isotope pattern to a degree that the spectrum is not suitable for quantification

greatly facilitated the ability to quantify proteins in complex proteomes because the increased instrument performance enables the exact discrimination of peptide isotope clusters from interfering signals caused by, e.g., co-eluting and near-isobaric peptides and other chemical entities [71–73]. For quantification in tandem MS spectra, saturation effects are rarely a problem. Instead, low-intensity spectra are frequently obtained and may result in less robust quantitation values due to poor ion statistics. Unlike for quantification in survey spectra, the contribution of peptidic or chemical background noise to quantification does not depend on the mass resolution of the mass spectrometer but on the size of the m/z window chosen for isolation of peptides for sequencing (typically 2–6 m/z). All ions present in this window will contribute to the signal of the, e.g., iTRAQ reporter ions. As a result, it is not always clear to what extent quantification was contributed by the peptide of interest or by background. This can sometimes lead to a large underestimation of true changes, especially for very weak peptide signals.

Taken together, the limits to quantification of complex proteomes by stable isotopes is first and foremost an issue of signal interference caused by co-eluting components of similar mass. Therefore, the most straightforward way for optimizing quantitative analyses is to decrease sample complexity by increasing HPLC gradient times or by biochemical fractionation prior to LC-MS analysis.

Label-free quantification

Currently, two widely used but fundamentally different label-free quantification strategies can be distinguished: (a) measuring and comparing the mass spectrometric signal intensity of peptide precursor ions belonging to a particular protein and (b) counting and comparing the number of fragment spectra identifying peptides of a given protein. In the former approach, the ion chromatograms for every peptide are extracted from an LC-MS/MS run and their mass spectrometric peak areas are integrated over the chromatographic time scale. For low-resolution mass spectra this is typically done by creating extracted ion chromatograms (XICs) for the mass to charge ratios determined for each peptide [74]. More recently, this concept has been extended to high-resolution data to include contributions of ^{13}C isotopes to the overall signal intensities [75]. The intensity value for each peptide in one experiment can then be compared to the respective signals in one or more other experiments to yield relative quantitative information [74, 76–80]. For proteomic analysis of very complex peptide mixtures, three important experimental parameters affect the analytical accuracy of quantification by ion intensities. (i) It is advantageous to

employ a high mass accuracy mass spectrometer because the influence of interfering signals of similar but distinct mass can be minimized. (ii) The peptide chromatographic profile should be optimized for reproducibility to ease finding corresponding peptides between different experiments. This is not a trivial task and special software has been developed to align LC-runs prior to identifying corresponding peptides [81–84]. (iii) The right balance between acquisition of survey and fragment spectra has to be found. While extensive peptide sequencing by tandem MS is required to identify as many proteins as possible in complex mixtures, a robust quantitative reading by ion intensities requires multiple sampling of the chromatographic peak by survey mass spectra. Typically, multiple fragment spectra are acquired for every survey spectrum at acquisition rates ranging from 0.2 s/spectrum (ion traps) to 1–3 s/spectrum (quadrupole-TOF instruments). Given that chromatographic peak widths are in the order of 10–30 s for nano-LC separations, ion traps have an inherent advantage over QTOFs because many more MS to MS/MS cycles can be performed within the available chromatographic time. Still, even for fast sampling instruments, better quantification accuracy will inevitably mean poorer proteome coverage and vice versa. This dilemma has led some laboratories to conduct two separate experiments for each sample: one which focuses on identifying as many peptides as possible by MS/MS and a second performed in MS-only mode in order to optimize sampling of intact peptide signals. In these approaches, matching of integrated peak intensities to identified peptides is performed by using a combination of accurate mass and retention time [84–86]. An alternative has been proposed in which the mass spectrometer no longer cycles between MS and MS/MS mode but aims to detect and fragment all peptides in a chromatographic window simultaneously by rapidly alternating between high- and low-energy conditions in the mass spectrometer [87–90]. Obviously, there are challenges with analyzing such data from complex samples as many fragmentation spectra will be populated with sequence ions from multiple peptides each contributing differently to the overall spectral content.

The peptide or more recently introduced spectral counting approach [91–93] is based on the empirical observation that the more of a particular protein is present in a sample, the more tandem MS spectra are collected for peptides of that protein. Hence, relative quantification can be achieved by comparing the number of such spectra between a set of experiments. In contrast to quantification by peptide ion intensities, spectral counting benefits from extensive MS/MS data acquisition across the chromatographic time scale both for protein identification as well as protein quantification. However, the commonly employed dynamic exclusion of ions that have already been selected

for fragmentation is detrimental for accurate quantification [94]. Although very intuitive and attractive in practical terms, the spectrum counting approach is still controversial because it does not measure any direct physical property of a peptide. It further assumes that the linearity of response is the same for every protein. In fact, the spectrum count response is different for every peptide because, e.g., the chromatographic behavior (retention time, peak width) varies for every peptide. Therefore, even reasonable quantification requires the observation of many spectra for a given protein. Old et al. [94] have shown that although it is possible to detect threefold protein changes with as few as four spectra; this number increases exponentially for smaller changes (ca.15 spectra for twofold). At the same time, saturation effects will be observed at higher spectral counts and saturation levels will be different for all proteins which renders the assessment of the dynamic range of observed changes difficult.

Nevertheless, the correlation between amount of protein and number of tandem mass spectra does hold and has led researchers to extend the concept to the estimation of absolute protein expression levels. In the first of a series of papers, Rappsilber et al. [95] computed a protein abundance index (PAI) by dividing the number of observed peptides by the number of all possible tryptic peptides from a particular protein that are within the mass range of the employed mass spectrometer. In a subsequent refinement, the same group transformed the PAI into an exponentially modified form (emPAI) [96] which showed a better correlation to known protein amounts. Further advances have been made by using computational models that predict which peptides of a given protein are likely to be detected by the mass spectrometer in the first place and thus would form a better basis for quantification [97–99, 66]. For example, results obtained by the absolute protein expression profiling (APEX) method [99] suggest that absolute protein expression can be determined to within the correct order of magnitude.

Label-free approaches are certainly the least accurate among the mass spectrometric quantification techniques when considering the overall experimental process because all the systematic and non-systematic variations between experiments are reflected in the obtained data (Fig. 2). Consequently, the number of experimental steps should be kept to a minimum and every effort should be made to control reproducibility at each step. Nonetheless, label-free quantification is worth considering for a number of reasons. In simple practical terms, the time-consuming steps of introducing a label into proteins or peptides can be omitted and there are no costs for labeling reagents. In terms of analytical strategy, the following points may also be important: (i) there is no principle limit to the number of experiments that can be compared. This is certainly an

advantage over stable isotope labeling techniques that are typically limited to 2–8 experiments that can be directly compared. (ii) Unlike for most stable isotope labeling techniques, mass spectral complexity (in terms of detected peptide species within a particular chromatographic time window) is not increased which, in turn, might provide for more analytical depth (i.e., number of detected peptides/proteins in an experiment) because the mass spectrometer is not occupied with fragmenting all forms of the labeled peptide. (iii) There is evidence that label-free methods provide higher dynamic range of quantification than stable isotope labeling (Table 1) and therefore may be advantageous when large and global protein changes between experiments are observed. However, particularly for spectral counting, this comes at the cost of unclear linearity and relatively poor accuracy [94].

Analysis of quantitative MS data

When contemplating a data analysis strategy for proteomic data generated by quantitative mass spectrometry, it is worth reconsidering a couple of principle points. Quantitative proteomic data are typically very complex, and often of variable quality. This is in part because the data are incomplete: even the most advanced mass spectrometers, which can acquire several tandem MS spectra per second, are often overwhelmed by the number of peptides present in a sample. As a consequence, only a subset of all proteins present can be identified in any one analysis [100]. For protein quantification, it is further mandatory to detect a protein in all experiments that should be compared. As a result, often only a subset of identified proteins can actually be quantified (Fig. 1) [92]. Identification and quantification rates are direct functions of sample complexity. While a large fraction of proteins present in, e.g., affinity purifications can be identified and quantified using a reasonable number of acquired spectra, a much smaller fraction of the content of whole proteome shotgun experiments will be covered and with fewer spectra for each protein. This clearly limits the confidence in quantification results.

These general considerations aside, practitioners of proteomics will soon face a number of practical challenges in analyzing quantitative mass spectrometric data: (i) quantitative readings must be extracted from MS or MS/MS spectra; (ii) peptide and protein identification must be performed; (iii) the two types of information must be merged and quality controlled; (iv) the applicable statistical methods have to be identified; and (v) the individual steps have to be combined into a workflow which bridges gaps between commercially available software and custom-built tools and which ideally also allows for automating most of the tasks (Fig. 4).

For protein quantification based on spectrum counting, the data processing steps are basically identical to the general protein identification workflow in proteomics which is one of the reasons why this approach has become so popular. Researchers can choose from a variety of methods available for automated protein identification and subsequent (probabilistic) validation of spectrum-to-peptide matches (for a recent review see Ref. [101]). It should be emphasized that for any quantification method it is mandatory to consider only those spectrum-to-peptide matches that are unique for a particular protein [11].

Extracting quantitative information from MS and MS/MS spectra

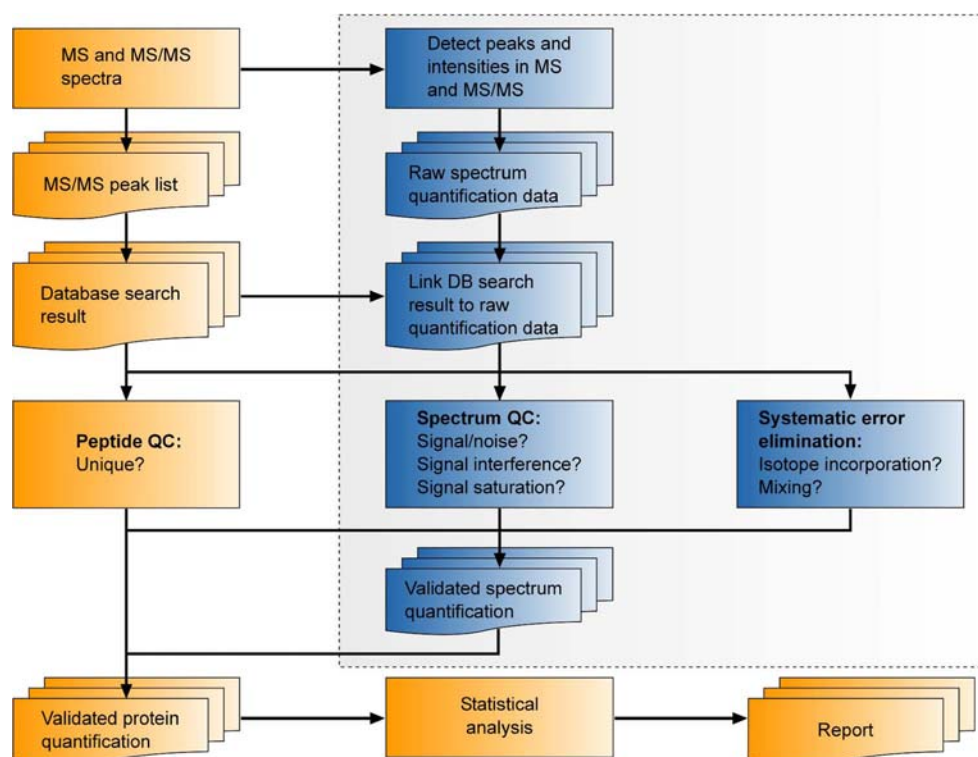
Quantification methods based on ion intensities, regardless of whether employing stable isotope labeling or not, require a number of additional steps prior to protein quantification (boxed area in Fig. 4). Two particular elements are important to mention here: intensity integration (i) within the mass spectrum (centroiding) and (ii) across the chromatographic peak. For low-resolution MS data, both aspects are carried out in one operation by extracting the ion chromatograms from the LC-MS data. For high-resolution MS data, the procedure is more complex and typically performed in two steps. Signal intensity integration within the mass spectrum can either utilize the intensity/area of the monoisotopic peak or the sum of the intensities/areas of all

isotopomers of a peptide. Each method has its merits and detractions: monoisotopic peak integration is relatively straightforward to implement but not very sensitive particularly for larger peptides for which the monoisotopic peaks only constitute a minority of the total signal intensity. In addition, the use of heavy isotopes distorts the relative isotope distribution of peptides which leads to inaccuracies. In contrast, the summed area of the entire isotope cluster is the most sensitive and accurate method [102] as it utilizes all of the data but is more difficult to implement computationally. As discussed in a previous section, signal intensity integration over the chromatographic time scale is primarily required for label-free quantification as well as those stable isotope reagents that lead to significant differences in chromatographic behavior. For methods which do not suffer from this shortcoming, time integration can be performed but is not required. Instead, collection of several spectra for each peptide is generally useful in order to obtain several quantitative readings.

Quality control of raw MS data

There are several sources of potential error in the mass spectrometric readout of an LC-MS experiment that can negatively affect the results of peptide quantification. Spectra for which these errors are detected should be filtered out prior to computing quantification values. The first of these issues is the presence and variability of

Fig. 4 Generic data processing and analysis workflow for quantitative mass spectrometry. Yellow icons indicate steps common to all quantification approaches with or without the use of stable isotopes. Blue icons in the boxed area refer to extra steps required when using mass spectrometric signal intensity values for quantification



spectral background noise (Fig. 3b) which can be filtered out by most if not all available commercial and academic data processing packages. A second common issue is the presence of interfering signals other than background noise (Fig. 3b). For very complex peptide mixtures, these often constitute co-eluting peptides of very similar m/z values which in turn will render the correct assignment of signal intensities to particular peptide ions difficult. This is true for quantification in both MS and MS/MS spectra and such spectra should be removed from the analysis. Third, strong signal intensities can lead to detector saturation for some mass spectrometers (particularly quadrupole TOF instruments, Fig. 3c) which distorts the natural isotope intensity distribution and thus leads to false quantitative readings.

For stable isotope labeling, further quality criteria must be considered. One very simple and often incurred problem is systematic bias introduced by imperfections in mixing the two protein populations. Mixing errors can most of the time be determined experimentally and apply uniformly to all protein quantification values and are thus easily corrected for. A second systematic error is represented by the isotope purity of the employed labeling reagent which rarely exceeds 95–98%. Although this may not appear to be a significant source of uncertainty and, again, can be easily

corrected for, isotope impurities lead to increased spectral interferences and, more importantly, limit the dynamic range of detectable differences between samples. A similar argument applies to incomplete incorporation of the isotope label into proteins and peptides. Again, while isotope incorporation can be measured and correction factors can be applied, the combination of the above items limits the dynamic range of detectable differences between samples to approximately 20–30:1. Consequently, determined changes are often smaller than their true values. It is important to keep in mind that this effect can be much more pronounced when spectral background contributes significantly to overall spectral intensity.

From spectra to relative protein quantification

For the spectrum counting approach, relative protein quantification between two or more samples is simply performed by comparing the respective numbers. If ten spectra are observed for a protein under condition 1 and 15 spectra under condition 2, the change between the two conditions is 1.5-fold. In contrast, for all approaches that measure signal intensities of peptide spectra, a quantitative reading is obtained for each spectrum. Obviously the accuracy

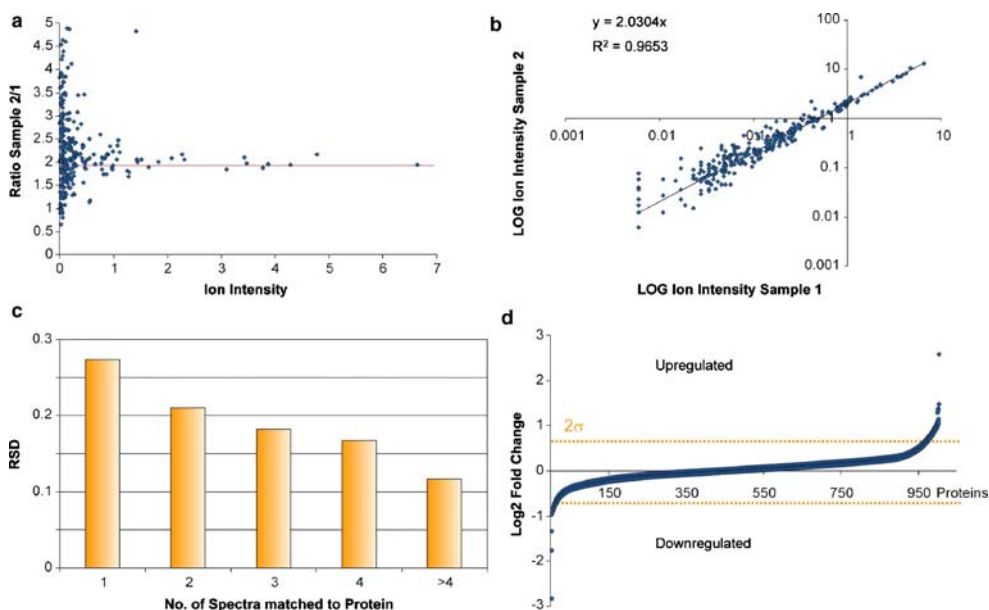


Fig. 5 **a** Distribution of measured changes from peptide spectra as a function of spectrum intensity for a single protein mixed in a 2:1 ratio. *Diamonds* represent intensity readings from individual spectra. The *red line* indicates the expected ratio of 2. It is evident that variations in change determination are much larger for low-intensity spectra than for medium- or high-intensity spectra. **b** Protein change determination by linear regression analysis. *Diamonds* represent intensity readings from individual spectra for samples 1 and 2 (same data as in **a**). The slope of the two-sided regression line approximates the expected twofold difference in protein quantity between the two samples. **c** Histogram showing the relationship between precision of quantifica-

tion (expressed as relative standard deviation, RSD) and the number of observed peptide spectra for a given protein from replicate experiments. Not surprisingly, precision increases with increasing number of spectra. **d** Change distribution for approximately 1,000 proteins identified and quantified between two experimental conditions in a single experiment. *Diamonds* represent individual protein fold changes in ascending order. In the absence of replicate experiments, data points between *yellow lines* (arbitrarily set at 2σ) are typically not considered to change significantly. However, these data points may contain many false negatives (small but significant changes)

of the protein quantification is determined by the accuracy of each peptide (spectrum) determination. The resulting data are spectrum-related quantity measures of varying precision. As an experiment typically produces a number of spectra per protein, these measurements have to be aggregated in a way that returns the best (i.e., most precise) protein quantification measure. Most publications to date rely on simple averaging of ratios [103], but as exemplified in Fig. 5a, variation of change determination is a function of signal intensity. Thus, low-intensity or noisy data may easily distort the mean value of computed ratios [104]. To overcome this problem, intensity thresholds have been employed [65]. However, these mostly arbitrary thresholds may also lead to arbitrary reduction of proteins that can be quantified. As an alternative, results can be improved either by calculation of an intensity weighted average, by summing up of all measured quantities followed by calculation the protein ratio [103, 75], or by calculating a linear regression (allowing for two dimensions of freedom) to determine the protein ratio (Fig. 5b) [105]. Apart from mass spectrometric signal strength, accuracy of quantification also benefits from the availability of multiple spectra for a given protein (Fig. 5c).

Statistical analysis of experimental data

Proteomic experiments comparing a number of states of a biological system typically generate complex data. An understanding of the experimental setup and the nature and quality of the obtained data are required to devise appropriate statistical methods. Experiments typically fall into two distinct categories: either the interrelation between a protein's abundance (or another property) and a certain sample condition is examined or the interaction between proteins is analyzed. Table 2 lists examples of such questions and some appropriate statistical strategies that have been applied to answer them. The detection of protein abundance changes is discussed in more detail below as it represents one of the major applications of proteomics. Most of the available statistical methods have previously

been applied to gene expression analysis but can often also be applied to quantitative MS data. However, the required data preparation steps such as normalization might be significantly different.

Data preparation

Raw data from quantitative MS experiments are generally not suitable for statistical analysis, thus a number of preparative steps are required. First, raw data are typically not normally distributed, an assumption made by many statistical tests. Therefore, data are frequently log-transformed assuming that the data are lognormal-distributed. This operation typically also harmonizes the variance of data (otherwise high values would have large variances and vice versa). If replicates of the experiment have been generated, normalization of their data is mandatory because technical bias may overshadow the underlying biological effects (for details on normalization techniques, see Refs. [106–108]). As discussed above, technical effects include sample mixing errors, incomplete isotope incorporation, or isotope impurity. In many cases, systematic technical bias can be measured directly but in some cases requires dedicated experimentation (e.g., by a label swap experiment [109, 110]) to determine its source. The resulting information is used to build correction functions that are consecutively applied to the data. It should be noted that it is very likely that not all manifesting sources of systematic error have been described yet or that these are not readily amenable to determination (e.g., background contribution in iTRAQ experiments). It can be expected though that with the rapid evolution of proteomic technologies, many of these yet unknown sources of error will be uncovered and the learnings subsequently used to sharpen the data which, in turn, increases data quality.

Another challenge to a statistical treatment of proteomic data is the mostly random sequencing of peptides by the mass spectrometer. As a result, not every available peptide is identified in every experiment. This effect is more pronounced for peptides of low abundance and poor detectability, resulting in many missing values in an experiment.

Table 2 Statistical methods for proteomics

Category	Question	Analysis suggestions
Protein change between conditions	Does a protein behave significantly different between two samples?	Multiple hypothesis testing
	Does a protein exhibit time-dependent change?	Analysis of variance (ANOVA)
	Is the sample a member of a defined class of samples?	Classification methods (e.g., linear discriminant analysis, support vector machines)
Dependencies between proteins	Which proteins behave similarly in the experiment?	Cluster analysis

However, statistical methods often require complete data. In such cases, missing values may be estimated by, e.g., averaging available values of the protein from other replicates or using related values from other proteins from the same experiment. It should be noted though that estimating values inevitably results in decreased statistical power [111, 112].

Values that are grossly different from comparable observations (outliers) require special attention. They can either indicate a true observation of a particular peptide species, e.g., a regulated post-translational modification, or a false reading. In both cases, these data points should initially be excluded from the calculation of protein quantities but not categorically rejected. A common way to spot outliers is visual inspection by the investigator, leaving considerable room for subjective judgement. During calculation of protein values from individual spectra by linear regression (see above) outlier detection on the spectrum level is possible using established methods [113, 114] but may result in loss of valuable data. For data correction at the protein level, methods for multivariate data can also be adapted [115, 116].

Detection of differential protein expression

It is not uncommon that publications reporting results of proteomic experiments using quantitative mass spectrometry base conclusions on measurements generated in one or two experiments. This is understandable given the often limited availability of specimen as well as the cost and time required to perform and analyze these samples. However, in light of the often considerable experimental variation, it is likely that those studies will not realise their full potential. For example, the graph shown in Fig. 5d represents a rank order list of the observed changes between two experiments for approximately 1,000 proteins. Proteins at the extremes of the distribution change the most and are therefore often considered to be the most interesting. While this might often be true when these observations are backed by many

spectra indicating this change, there are two important caveats. In this representation, small but potentially significant changes go unnoticed (false negatives) and, in the absence of repeating the experiment, there is no way of assessing if the observed large protein changes that are backed by few spectral observations can be reproduced (false positive). Even small numbers of repetitions can increase confidence in the results considerably. In addition, the use of statistical testing methods adds options to determine the probability of false decisions. A typical situation is the comparison of protein levels between two different samples with the goal to detect those proteins that are significantly changed between conditions. This biological question can be formulated as a problem in multiple hypotheses testing that describes a simultaneous test for each protein on the null hypothesis of no change in protein measure between the two conditions. A standard approach to such a multiple testing problem consists of two aspects: (i) computing a test statistics and (ii) applying a multiple testing procedure to determine which hypothesis to reject (change or no change) while controlling a defined false positive error rate [117]. Computing the test statistic for each protein can be carried out, e.g., by employing the frequently used t-test. This test expects the data to be normally distributed, an assumption that is not always justified and requires a significant number of replicates in order to return reliable results (Table 3). For lower replication numbers (2–3) the so-called local-pooled-error test (LPE) has been found to be useful provided that protein changes are not too small [118–120]. For data with unknown distribution characteristics, non-parametric tests can also be used that are agnostic towards the data's distribution but come at the expense of statistical power [121].

In the proteomics case where many proteins are tested simultaneously, the probability of committing an error increases often dramatically. For example, when considering a list of hundreds of proteins at a defined error rate of, e.g., 0.01, it is likely that several false positives will occur by chance. However, when setting the thresholds too

Table 3 Characteristics and applications of statistical tests

Test	Requirements	Statistical power	Application
Tests for experiments with replicates			
t-test	Replications, $n > 3$ Data normally distributed	+++	All quantitation methods
LPE-test	Replications, $n > 1$	++	All quantitation methods 2–3 replicates Strong changes
Tests for experiments without replicates			
G-test	(Very) large number of peptide spectra	+	Spectrum counting
Fischer's exact test	(Very) large number of peptide spectra	+	Spectrum counting
AC-test	(Very) large number of peptide spectra	+	Spectrum counting

conservatively to minimize false positive rate (i.e., the rate that truly null features are called significant), this often leads to an unacceptable increase in the false negative rate (i.e., the rate that truly significant features are called null). Commonly used alternative measures of error rates in multiple testing procedures are the family wise error rate (FWER; i.e., the rate that one truly null feature is called significant among all tests) and the false discovery rate (FDR; i.e., the rate that features called significant are truly null) which break up the direct dependency between false positive and false negative rates. Instead of simply reporting rejection or acceptance of the specified hypothesis using these methods, a p-value connected to the test can be defined which describes the significance of a test as the smallest possible significance level at which the null hypothesis would be rejected. Various procedures for deriving adjusted p-values for multiple hypothesis testing have been suggested, e.g., the Bonferroni adjusted p-value for FWER and the q-value for FDR [122]. q-Values have since also been adopted in proteomics research [123, 124]. A detailed overview of multiple hypothesis testing has been given by Dudoit and co-workers [125].

Sampling statistics

For a number of proteomic applications, sampling statistics (e.g., spectrum count, peptide count, sequence coverage) shows increasing potential. Zhang and co-workers [120] recently compared the aforementioned three approaches and found that the spectrum counting approach offered the greatest reproducibility. This is probably not surprising given that this approach generates many more data points than peptide counting or measuring sequence coverage. In addition this paper explores a number of statistical methods for data analysis. For experiments that feature three or more replicates of each condition, statistical difference can be assessed by the t-test as described above. However, if repetitions are not available, other statistical options have to be considered. To that end, tests may be applicable that attempt to mimic replicates by pooling certain features. For example, for each detected protein, spectral counts from a pair-wise experiment can be arranged in a two-way table (proteins vs. conditions). A protein is then called differentially expressed if its proportion of spectrum counts to the total spectrum count in the experiment is significantly different between both conditions. There are a number of possible statistical tests using different hypotheses for this approach (Table 3, bottom). The authors of the aforementioned paper conclude that Fisher's exact test, the AC-test, and the G-test return comparable results. However the G-test is computationally simpler and can be generalized for multi-condition experiments and thus may be the more versatile approach. Results typically improve with in-

creased sampling (total number of spectrum counts in an experiment). Despite the fact that the commonly used dynamic exclusion option during LC-MS analysis violates random sampling, Zhang et al. showed that the approach can be generally useful [120].

In contrast to statistical estimation, the performance of a chosen statistical test can often also be assessed experimentally by means other than multiple repetitions. One way of measuring errors directly and under the same analytical conditions is to offset the measurement of a particular sample to a dilution of the very same sample [126]. Also, spiked proteins have been used to generate reference data for a set of proteins with known behavior that can be utilized for 'calibrating' an experiment type [92]. Once the statistical parameters have been learned, these may be applied to subsequent experiments without the need for repetition. Although the statistical power of such approaches is lower than those based on multiple repetitions of the same experiment, the former may be sufficient particularly for samples of low protein complexity (e.g., affinity purifications). Further assessment of data significance may be provided by curve fitting methods (e.g., the LOWESS fit) which can reveal regions of random experimental error in the observed dataset [123].

Concluding remarks

A multitude of methods has emerged for the analysis of simple and complex (sub-)proteomes using quantitative mass spectrometry, and the field is beginning to learn for which type of study these methods can be meaningfully applied. However, significant further improvements to experimental strategies are required particularly for the quantitative analysis of post-translational modifications. It is probably fair to say that the field is still far from being able to generate quantitative proteomic data at a scale which would allow the comprehensive investigation of a biological phenomenon. At the same time, the recent exponential increase in data volume and complexity demands the development of appropriate statistical approaches in order to arrive at meaningful interpretations of the results. This can only be achieved if the influence of the employed technologies on the results obtained is well understood and by ensuring that experimental design follows the biological context so that the 'right statistics' can be developed for the problem at hand in order to generate scientific insight.

Acknowledgements The authors wish to thank David Simmons and Ulrich Kruse for critically reading the manuscript and Frank Weisbrodt for help with preparing the figures. We are grateful to Nature Publishing Group for granting permission to reproduce and adapt previously published material.

References

1. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) *Nat Biotechnol* 17:994–999
2. Oda Y, Huang K, Cross FR, Cowburn D, Chait BT (1999) *Proc Natl Acad Sci U S A* 96:6591–6596
3. Pasa-Tolic L, Jensen PK, Anderson GA, Lipton MS, Peden KK, Martinovic S, Tolic N, Bruce JE, Smith RD (1999) *J Am Chem Soc* 121:7949–7950
4. Aebersold R, Mann M (2003) *Nature* 422:198–207
5. Gygi SP, Rist B, Aebersold R (2000) *Curr Opin Biotechnol* 11:396–401
6. Heck AJ, Krijgsveld J (2004) *Expert Rev Proteomics* 1:317–326
7. Ong SE, Foster LJ, Mann M (2003) *Methods* 29:124–130
8. Righetti PG, Campostrini N, Pascali J, Hamdan M, Astner H (2004) *Eur J Mass Spectrom (Chichester, Eng)* 10:335–348
9. Sechi S, Oda Y (2003) *Curr Opin Chem Biol* 7:70–77
10. Tao WA, Aebersold R (2003) *Curr Opin Biotechnol* 14:110–118
11. Ong SE, Mann M (2005) *Nat Chem Biol* 1:252–262
12. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) *Mol Cell Proteomics* 1:376–386
13. Krijgsveld J, Ketting RF, Mahmoudi T, Johansen J, Artal-Sanz M, Verrijzer CP, Plasterk RH, Heck AJ (2003) *Nat Biotechnol* 21:927–931
14. Wu CC, MacCoss MJ, Howell KE, Matthews DE, Yates JR III (2004) *Anal Chem* 76:4951–4959
15. Gruhler A, Schulze WX, Matthiesen R, Mann M, Jensen ON (2005) *Mol Cell Proteomics* 4:1697–1709
16. Ong SE, Kratchmarova I, Mann M (2003) *J Proteome Res* 2:173–181
17. Blagoev B, Ong SE, Kratchmarova I, Mann M (2004) *Nat Biotechnol* 22:1139–1145
18. Park KS, Mohapatra DP, Misonou H, Trimmer JS (2006) *Science* 313:976–979
19. Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M (2006) *Cell* 127:635–648
20. Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C (2001) *Anal Chem* 73:2836–2842
21. Reynolds KJ, Yao X, Fenselau C (2002) *J Proteome Res* 1:27–33
22. Rose K, Simona MG, Offord RE, Prior CP, Otto B, Thatcher DR (1983) *Biochem J* 215:273–277
23. Miyagi M, Rao KC (2007) *Mass Spectrom Rev* 26:121–136
24. Rao KC, Carruth RT, Miyagi M (2005) *J Proteome Res* 4:507–514
25. Schnolzer M, Jedrzejewski P, Lehmann WD (1996) *Electrophoresis* 17:945–953
26. Johnson KL, Muddiman DC (2004) *J Am Soc Mass Spectrom* 15:437–445
27. Ramos-Fernandez A, Lopez-Ferrer D, Vazquez J (2007) *Mol Cell Proteomics* 6(7):1274–1286
28. Oda Y, Owa T, Sato T, Boucher B, Daniels S, Yamanaka H, Shinohara Y, Yokoi A, Kuromitsu J, Nagasu T (2003) *Anal Chem* 75:2159–2165
29. Hansen KC, Schmitt-Ulms G, Chalkley RJ, Hirsch J, Baldwin MA, Burlingame AL (2003) *Mol Cell Proteomics* 2:299–314
30. Li J, Steen H, Gygi SP (2003) *Mol Cell Proteomics* 2:1198–1204
31. Yi EC, Li XJ, Cooke K, Lee H, Raught B, Page A, Aneliunas V, Hieter P, Goodlett DR, Aebersold R (2005) *Proteomics* 5:380–387
32. Shen M, Guo L, Wallace A, Fitzner J, Eisenman J, Jacobson E, Johnson RS (2003) *Mol Cell Proteomics* 2:315–324
33. Pasquarello C, Sanchez JC, Hochstrasser DF, Corthals GL (2004) *Rapid Commun Mass Spectrom* 18:117–127
34. Shi Y, Xiang R, Crawford JK, Colangelo CM, Horvath C, Wilkins JA (2004) *J Proteome Res* 3:104–111
35. Shi Y, Xiang R, Horvath C, Wilkins JA (2005) *J Proteome Res* 4:1427–1433
36. Qiu Y, Sousa EA, Hewick RM, Wang JH (2002) *Anal Chem* 74:4969–4979
37. Sebastiano R, Citterio A, Lapadula M, Righetti PG (2003) *Rapid Commun Mass Spectrom* 17:2380–2386
38. Schmidt A, Kellermann J, Lottspeich F (2005) *Proteomics* 5: 4–15
39. Ross PL, Huang YN, Marchese JN, 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 DJ (2004) *Mol Cell Proteomics* 3:1154–1169
40. Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C (2003) *Anal Chem* 75:1895–1904
41. Ji J, Chakraborty A, Geng M, Zhang X, Amini A, Bina M, Regnier F (2000) *J Chromatogr B Biomed Sci Appl* 745:197–210
42. Che FY, Fricker LD (2002) *Anal Chem* 74:3190–3198
43. Zhang X, Jin QK, Carr SA, Annan RS (2002) *Rapid Commun Mass Spectrom* 16:2325–2332
44. Glocker MO, Borchers C, Fiedler W, Suckau D, Przybylski M (1994) *Bioconjug Chem* 5:583–590
45. Mason DE, Liebler DC (2003) *J Proteome Res* 2:265–272
46. Lee YH, Han H, Chang SB, Lee SW (2004) *Rapid Commun Mass Spectrom* 18:3019–3027
47. Hsu JL, Huang SY, Chow NH, Chen SH (2003) *Anal Chem* 75:6843–6852
48. Ji C, Guo N, Li L (2005) *J Proteome Res* 4:2099–2108
49. Hsu JL, Huang SY, Chen SH (2006) *Electrophoresis* 27:3652–3660
50. Goodlett DR, Keller A, Watts JD, Newitt R, Yi EC, Purvine S, Eng JK, von Haller P, Aebersold R, Kolker E (2001) *Rapid Commun Mass Spectrom* 15:1214–1221
51. Syka JE, Marto JA, Bai DL, Horning S, Senko MW, Schwartz JC, Ueberheide B, Garcia B, Busby S, Muratore T, Shabanowitz J, Hunt DF (2004) *J Proteome Res* 3:621–626
52. Salomon AR, Ficarro SB, Brill LM, Brinker A, Phung QT, Ericson C, Sauer K, Brock A, Horn DM, Schultz PG, Peters EC (2003) *Proc Natl Acad Sci U S A* 100:443–448
53. Goshe MB, Conrads TP, Panisko EA, Angell NH, Veenstra TD, Smith RD (2001) *Anal Chem* 73:2578–2586
54. Goshe MB, Veenstra TD, Panisko EA, Conrads TP, Angell NH, Smith RD (2002) *Anal Chem* 74:607–616
55. Qian WJ, Goshe MB, Camp DG, Yu LR, Tang K, Smith RD (2003) *Anal Chem* 75:5441–5450
56. Tao WA, Wollscheid B, O'Brien R, Eng JK, Li XJ, Bodenmiller B, Watts JD, Hood L, Aebersold R (2005) *Nat Methods* 2:591–598
57. Zhang H, Li XJ, Martin DB, Aebersold R (2003) *Nat Biotechnol* 21:660–666
58. Wiese S, Reidegeld KA, Meyer HE, Warscheid B (2007) *Proteomics* 7:1004
59. Desiderio DM, Kai M (1983) *Biomed Mass Spectrom* 10:471–479
60. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) *Proc Natl Acad Sci U S A* 100:6940–6945
61. Pan S, Zhang H, Rush J, Eng J, Zhang N, Patterson D, Comb MJ, Aebersold R (2005) *Mol Cell Proteomics* 4:182–190
62. Kirkpatrick DS, Gerber SA, Gygi SP (2005) *Methods* 35:265–273
63. Beynon RJ, Doherty MK, Pratt JM, Gaskell SJ (2005) *Nat Methods* 2:587–589
64. Kito K, Ota K, Fujita T, Ito T (2007) *J Proteome Res* 6:792–800
65. Wolf-Yadlin A, Hautaniemi S, Lauffenburger DA, White FM (2007) *Proc Natl Acad Sci U S A* 104:5860–5865
66. Mallick P, Schirle M, Chen SS, Flory MR, Lee H, Martin D, Ranish J, Raught B, Schmitt R, Werner T, Kuster B, Aebersold R (2007) *Nat Biotechnol* 25:125–131
67. Zhang R, Sioma CS, Wang S, Regnier FE (2001) *Anal Chem* 73:5142–5149

68. Zhang R, Regnier FE (2002) *J Proteome Res* 1:139–147
69. Roepstorff P, Fohlman J (1984) *Biomed Mass Spectrom* 11:601
70. Belov ME, Rakov VS, Nikolaev EN, Goshe MB, Anderson GA, Smith RD (2003) *Rapid Commun Mass Spectrom* 17:627–636
71. Olsen JV, de Godoy LM, Li G, Macek B, Mortensen P, Pesch R, Makarov A, Lange O, Horning S, Mann M (2005) *Mol Cell Proteomics* 4:2010–2021
72. Zubarev R, Mann M (2007) *Mol Cell Proteomics* 6:377–381
73. Venable JD, Wohlschlegel J, McClatchy DB, Park SK, Yates JR III (2007) *Anal Chem* 79:3056–3064
74. Bondarenko PV, Chelius D, Shaler TA (2002) *Anal Chem* 74:4741–4749
75. Ono M, Shitashige M, Honda K, Isobe T, Kuwabara H, Matsuzuki H, Hirohashi S, Yamada T (2006) *Mol Cell Proteomics* 5:1338–1347
76. Chelius D, Bondarenko PV (2002) *J Proteome Res* 1:317–323
77. Wang W, Zhou H, Lin H, Roy S, Shaler TA, Hill LR, Norton S, Kumar P, Anderle M, Becker CH (2003) *Anal Chem* 75:4818–4826
78. Wiener MC, Sachs JR, Deyanova EG, Yates NA (2004) *Anal Chem* 76:6085–6096
79. Higgs RE, Knierman MD, Gelfanova V, Butler JP, Hale JE (2005) *J Proteome Res* 4:1442–1450
80. Wang G, Wu WW, Zeng W, Chou CL, Shen RF (2006) *J Proteome Res* 5:1214–1223
81. Bylund D, Danielsson R, Malmquist G, Markides KE (2002) *J Chromatogr A* 961:237–244
82. Wang P, Tang H, Fitzgibbon MP, McIntosh M, Coram M, Zhang H, Yi E, Aebersold R (2007) *Biostatistics* 8:357–367
83. Jaitly N, Monroe ME, Petyuk VA, Clauss TR, Adkins JN, Smith RD (2006) *Anal Chem* 78:7397–7409
84. Strittmatter EF, Ferguson PL, Tang K, Smith RD (2003) *J Am Soc Mass Spectrom* 14:980–991
85. Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li GZ, McKenna T, Nold MJ, Richardson K, Young P, Geromanos S (2005) *Anal Chem* 77:2187–2200
86. Zimmer JS, Monroe ME, Qian WJ, Smith RD (2006) *Mass Spectrom Rev* 25:450–482
87. Bateman RH, Carruthers R, Hoyes JB, Jones C, Langridge JJ, Millar A, Vissers JP (2002) *J Am Soc Mass Spectrom* 13:792–803
88. Nakamura T, Dohmae N, Takio K (2004) *Proteomics* 4:2558–2566
89. Niggeweg R, Kocher T, Gentzel M, Buscaino A, Taipale M, Akhtar A, Wilm M (2006) *Proteomics* 6:41–53
90. Silva JC, Denny R, Dorschel C, Gorenstein MV, Li GZ, Richardson K, Wall D, Geromanos SJ (2006) *Mol Cell Proteomics* 5:589–607
91. Washburn MP, Wolters D, Yates JR III (2001) *Nat Biotechnol* 19:242–247
92. Liu H, Sadygov RG, Yates JR III (2004) *Anal Chem* 76:4193–4201
93. Gilchrist A, Au CE, Hiding J, Bell AW, Fernandez-Rodriguez J, Lesimple S, Nagaya H, Roy L, Gosline SJ, Hallett M, Paiement J, Kearney RE, Nilsson T, Bergeron JJ (2006) *Cell* 127:1265–1281
94. Old WM, Meyer-Arendt K, veline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, Resing KA, Ahn NG (2005) *Mol Cell Proteomics* 4:1487–1502
95. Rappsilber J, Ryder U, Lamond AI, Mann M (2002) *Genome Res* 12:1231–1245
96. Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005) *Mol Cell Proteomics* 4:1265–1272
97. Craig R, Cortens JP, Beavis RC (2005) *Rapid Commun Mass Spectrom* 19:1844–1850
98. Tang H, Arnold RJ, Alves P, Xun Z, Clemmer DE, Novotny MV, Reilly JP, Radivojac P (2006) *Bioinformatics* 22:e481–e488
99. Lu P, Vogel C, Wang R, Yao X, Marcotte EM (2007) *Nat Biotechnol* 25:117–124
100. Aebersold R (2003) *Nature* 422:115–116
101. Nesvizhskii AI (2006) *Methods Mol Biol* 367:87–120
102. Chalkley RJ, Hansen KC, Baldwin MA (2005) *Methods Enzymol* 402:289–312
103. Saito A, Nagasaki M, Oyama M, Kozuka-Hata H, Semba K, Sugano S, Yamamoto T, Miyano S (2007) *BMC Bioinformatics* 8:15
104. Carrillo B, Yanofsky C, Boismenu D, Latterich M, Kearney RE (2006) Statistical limits of isotopic/isobaric quantification in counting detectors. *Proceedings of the American Society for Mass Spectrometry* 174
105. Parish RC (1989) *Ann Pharmacother* 23:891–989
106. Li C, Wong WH (2001) *Proc Natl Acad Sci U S A* 98:31–36
107. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP (2002) *Nucleic Acids Res* 30:e15
108. Kreil DP, Karp NA, Lilley KS (2004) *Bioinformatics* 20:2026–2034
109. Wang YK, Ma Z, Quinn DF, Fu EW (2001) *Anal Chem* 73:3742–3750
110. Wang YK, Ma Z, Quinn DF, Fu EW (2002) *Rapid Commun Mass Spectrom* 16:1389–1397
111. Jung K, Gannoun A, Stühler K, Sitek B, Meyer HE, Urfer W (2005) *RevStat-Statistical Journal* 3:99–111
112. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, Botstein D, Altman RB (2001) *Bioinformatics* 17:520–525
113. Ellenberg JH (1976) *Biometrics* 32:637–645
114. Wisnowski JW, Montgomery DC, Simpson JR (2001) *Comput Stat Data Anal* 36:351–382
115. Zhao HY, Yue PY, Fang KT (2004) *J Biopharm Stat* 14:629–646
116. Egan WJ, Morgan SL (1998) *Anal Chem* 70:2372–2379
117. Dudoit S, van der Laan MJ, Pollard KS (2004) *Stat Appl Genet Mol Biol* 3:Article13
118. Jain N, Thatte J, Braciale T, Ley K, O’Connell M, Lee JK (2003) *Bioinformatics* 19:1945–1951
119. Jain N, Cho H, O’Connell M, Lee JK (2005) *BMC Bioinformatics* 6:187
120. Zhang B, VerBerkmoes NC, Langston MA, Uberbacher E, Hettich RL, Samatova NF (2006) *J Proteome Res* 5:2909–2918
121. Dudbridge F, Gusnanto A, Koeleman BP (2006) *Hum Genomics* 2:310–317
122. Storey JD, Tibshirani R (2003) *Proc Natl Acad Sci U S A* 100:9440–9445
123. Xia Q, Wang T, Park Y, Lamont RJ, Hackett M (2007) *Int J Mass Spectrom* 259:105–116
124. Hendrickson EL, Xia Q, Wang T, Leigh JA, Hackett M (2006) *Analyst* 131:1335–1341
125. Dudoit S, Shaffer JP, Boldrick JC (2003) *Stat Sci* 18:71–103
126. Rinner O, Mueller LN, Hubalek M, Muller M, Gstaiger M, Aebersold R (2007) *Nat Biotechnol* 25:345–352