Anal Bioanal Chem (2012) 404:939–965
DOI 10.1007/s00216-012-6203-4

REVIEW

Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present

Marcus Bantscheff · Simone Lemeer · Mikhail M. Savitski · Bernhard Kuster

Received: 10 April 2012 / Revised: 6 June 2012 / Accepted: 15 June 2012 / Published online: 8 July 2012

Abstract Mass-spectrometry-based proteomics is continuing to make major contributions to the discovery of fundamental biological processes and, more recently, has also developed into an assay platform capable of measuring hundreds to thousands of proteins in any biological system. The field has progressed at an amazing rate over the past five years in terms of technology as well as the breadth and depth of applications in all areas of the life sciences. Some of the technical approaches that were at an experimental stage back then are considered the gold standard today, and the community is learning to come to grips with the volume and complexity of the data generated. The revolution in DNA/RNA sequencing technology extends the reach of proteomic research to practically any species, and the notion that mass spectrometry has the potential to eventually retire the western blot is no longer in the realm of science fiction. In this review, we focus on the major technical and conceptual developments since 2007 and illustrate these by important recent applications.

Keywords Quantitative proteomics · Liquid chromatography · Mass spectrometry · Bioinformatics

Published in the topical issue *Quantitative Mass Spectrometry in Proteomics* with guest editors Bernhard Kuster and Marcus Bantscheff.

M. Bantscheff (✉) · M. M. Savitski
Cellzome AG,
Meyerhofstrasse 1,
69117 Heidelberg, Germany
e-mail: marcus.bantscheff@cellzome.com

S. Lemeer · B. Kuster (✉)
Chair of Proteomics and Bioanalytics,
Technische Universität München,
Emil Erlenmeyer Forum 5,
85354 Freising, Germany
e-mail: kuster@tum.de

B. Kuster
Center for Integrated Protein Science Munich (CIPSM),
80539 Munich, Germany

Abbreviations

AP	Affinity purification
DDA	Data-dependent acquisition
ESI	Electrospray ionization
HCD	Higher-energy collision-induced dissociation
HDAC	Histone deacetylase
HPLC	High-pressure liquid chromatography
ICP	Inductively coupled plasma
IMAC	Immobilized metal affinity chromatography
iTRAQ	Isobaric tags for absolute and relative quantification
LC	Liquid chromatography
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
PAI	Protein abundance index
PrEST	Protein epitope signature tag
PSAQ	Protein standard absolute quantification
PSM	Peptide-to-spectrum match
PTM	Post-translational modification
QTOF	Quadrupole time of flight
SILAC	Stable isotope labeling with amino acids in cell culture
SRM	Selected reaction monitoring
TMT	Tandem mass tag
UPLC	Ultra-high-pressure liquid chromatography
XIC	Extracted ion chromatogram

Introduction

Investigating living systems at the protein level is continuing to provide important insights into many biological processes across all kingdoms of life. Mass spectrometry (MS)-based proteomics [1] has fundamentally changed the way in which biological systems are interrogated because of its

ability to measure thousands of proteins and posttranslational modifications (PTMs) in parallel. This enables investigations at all levels of biological complexity, ranging from protein complexes to human patient populations [2]. Whereas 5 years ago, most proteomic experiments mostly enumerated the protein constituents of a biological system, quantitative measurements are at the heart of practically every proteomic study today. This shift was once more significantly driven by developments in MS and associated sample preparation, separation, and data analysis methods, and MS is now the de facto standard for quantitative measurements in proteomics. Numerous experimental strategies and schemes have been devised, some of which have come and gone, whereas others have been adopted more broadly. The rough division of the proteomic workflow into (1) sample preparation, (2) protein and peptide separation, (3) MS, and (4) data analysis allows scientists to put together modular workflows incorporating elements that fit best for a particular task at hand (Fig. 1). Important such modules include stable-isotope encoding of proteins and peptides, high-resolution and multidimensional liquid chromatography (LC), and targeted and discovery-type MS, as well as tailored signal processing algorithms, database searching tools, and downstream biostatistics.

We last reviewed the field for this journal in 2007 [3], when we explained the main concepts and methods. In this update, we focus the presentation on such methods and applications that have been adopted widely, extension to these methods, and a number of new ideas which show considerable promise. The use of stable-isotope coding of proteins and peptides is now commonplace, but so-called label-free quantification strategies are also increasingly successful (Fig. 1). There is a clear trend in peptide LC toward increasing peak capacity, e.g., by using ultra-high-pressure reversed-phase systems as well as two-dimensional LC separations employing crude (i.e., low fraction number) first-dimension separations using, e.g., peptide charge [4]. Many new mass spectrometers based on triple-quadrupole, quadrupole time-of-flight (QTOF), and Orbitrap technology have been introduced, all of which are characterized by drastically improved sensitivity (ten to 50 times) and data acquisition speed (five to ten times). These modern LC-MS systems are capable of identifying and quantifying 5,000–10,000 proteins from a given proteome [5, 6]. The massive increase in primary data has spurred the development of a large collection of commercial and public software tools for protein identification and quantification, some of which have been assembled into pipelines [7, 8] designed to enable noninformatics scientists to analyze and interrogate their data. As for any other large-scale data scenario, proteomic data are often "noisy" or incomplete, which requires the right level of statistical treatment [9]. Much has been tried and tested in this area in the past few years, and one important conclusion is that there is no "one size fits all" type of statistics for quantitative proteomics data. This has important implications not only for the interpretation of individual results, but also for the development of (too restrictive, too loose) community standards and journal policies.

The increased technical capabilities of proteomics have given rise to an amazing growth of biologically, biomedically, or pharmacologically motivated fields of application. For instance, MS-based proteomics is challenging traditional

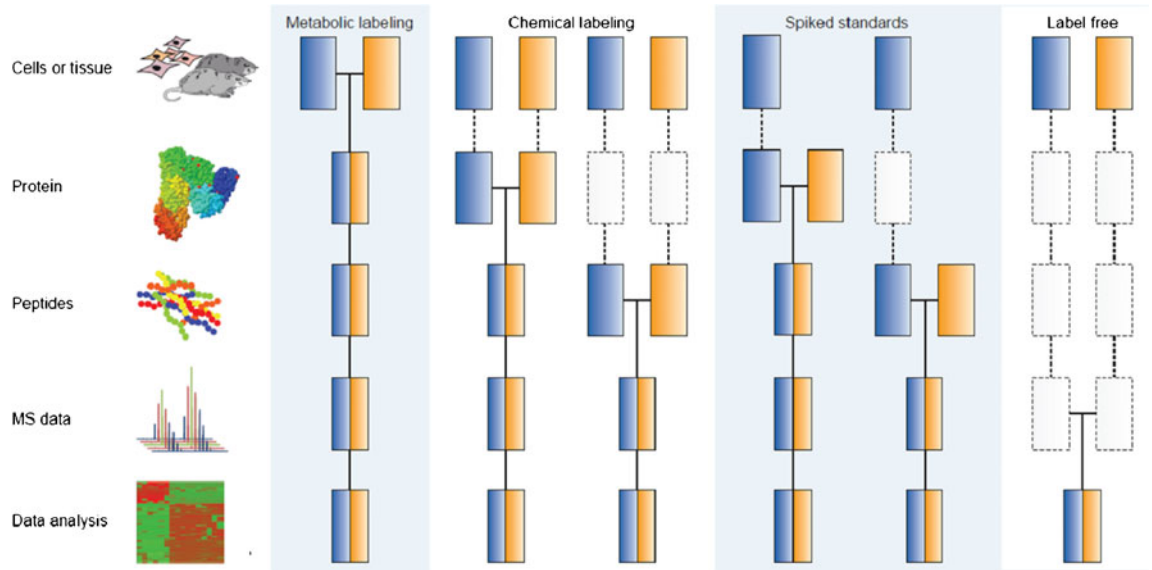

Fig. 1 Common quantitative mass spectrometry (*MS*)-based proteomics workflows. *Blue boxes* and *yellow boxes* represent two experimental conditions. *Horizontal lines* indicate when samples are combined. *Dashed lines* indicate the points at which experimental variation and thus quantification errors can occur. (Adapted from [3, 230])

hypothesis-driven research in basic biology (i.e., one protein at a time) because the technology allows the massive multiplexing of primary data generation with generally better quality than established methods such as western blots. This may be illustrated by the recent "explosion" of quantitative data on protein modifications such as phosphorylation and acetylation [10, 11] which has a profound impact on how cell signaling and epigenetics research is developing. Similarly, in the area of protein biomarker discovery, LC-MS methods are challenging traditional assays (such as ELISAs) because the need for the generation of high-quality antibodies is markedly reduced by the ability of the mass spectrometer to identify and quantify a protein unambiguously and accurately from a biofluid directly or following enrichment using straightforward-to-generate peptide antibodies [12]. Drug discovery is another beneficiary of the above-mentioned advances in technology, as quantitative chemoproteomics in particular can now be used more routinely to identify the spectrum of protein targets of a pharmacologically active substance in an unbiased way or to elucidate its molecular mechanism of action in a relevant model system [13]. In the following, we review the different modules of quantitative MS with extra emphasis on recent developments in label-free approaches as well as the analysis of quantitative MS data and thereby also hope to provide readers with some ideas with respect to finding the right technical approach for a particular application.

Metabolic labeling

The general idea behind stable-isotope labeling for peptide and protein quantification is that the physicochemical properties of labeled and natural peptides, including the MS signal response, are largely the same, if not identical. Therefore, relative and absolute quantification of a sample of interest can be performed by comparing its MS intensity with that of a labeled peptide standard present in the same sample. In metabolic labeling strategies, the isotope label is introduced into every protein during cell growth and division, which generates a labeled standard for every protein in a sample of interest. It also represents the earliest possible step in a proteomic workflow, thereby offering both high quantification accuracy and high precision because systematic errors arising from sample handling can be largely eliminated. Already, introduced more than a decade ago, one of the most popular metabolic labeling methods is stable isotope labeling with amino acids in cell culture (SILAC) [14]. In recent years, several new quantification strategies using SILAC have been developed (summarized in Fig. 2). In the classical SILAC experiment (Fig. 2a), typically isotopically labeled arginine and lysine (^{13}C, ^{15}N) are added to the culture medium, in this way ensuring that most of the peptides following tryptic digestion contain at least one labeled amino acid. Relative quantification is achieved by comparing the intensities of the isotope clusters of the labeled and unlabeled peptides in peptide ion mass spectra. Combining more than two samples in a single analysis run is possible; however, the repertoire of useful heavy labeled amino acids is still limited and, therefore, a maximum of three samples are typically combined in practice. Extending SILAC to four of five labels is possible by using deuterated amino acids; however, as this may lead to retention time shifts during reversed-phase LC (deuterium isotope effect) [15] and hence hamper accurate quantification, deuterated metabolic labels are not frequently used. Another clear disadvantage of multiplexing SILAC-labeled samples is the possibility of overlapping isotope clusters, complicating proper quantification (see "Analysis of quantitative MS data").

A relative new variation on the SILAC idea is the so-called pulsed SILAC method (Fig. 2b). Essentially, this is the classical pulse-chase experiment but using stable arginine and lysine isotopes rather than radioactive amino acids. The first such experiments were used to measure protein turnover in cell lines and even model organisms. Examples include the measurement of the accumulation and degradation of ribosomal proteins in the nucleus [16], the turnover of peptides presented by the major histocompatibility complex [17], and chicken muscle protein turnover. In those types of experiments, cells or model organisms are pulse-labeled with heavy isotopes for a certain period of time, and subsequently the protein turnover can be estimated by determining the ratios of heavy and light peptides. However, when introducing a differential treatment into this scheme (say, a drug treatment), one can no longer distinguish the effect of the treatment from the normal turnover kinetics of the cells. The refined pulsed SILAC strategy developed by the Selbach group [18] deals with this issue by first culturing cells in normal (light) medium and only upon some differential biological treatment are cells transferred to cell culture medium containing either medium-weight labeled amino acids or heavy labeled amino acids. After a certain incubation time, cells from both conditions are harvested and combined. The heavy-to-light ratio is now a true reflection of the differences in protein translation between the two conditions as the already existing proteins (in the light form) do not influence the measurement. The Mann group [19] refined the method further to deal with the possibility that unlabeled amino acids derived from protein degradation products may be recycled into the newly synthesized proteins rather than using labeled amino acids supplied by the culture medium, which would lead to an underestimation of the protein turnover rate. As this study shows, recycling does occur, but uptake and incorporation of externally supplied labeled amino acids was strongly favored under the conditions employed. In addition, the same group also showed that protein turnover can be measured by first

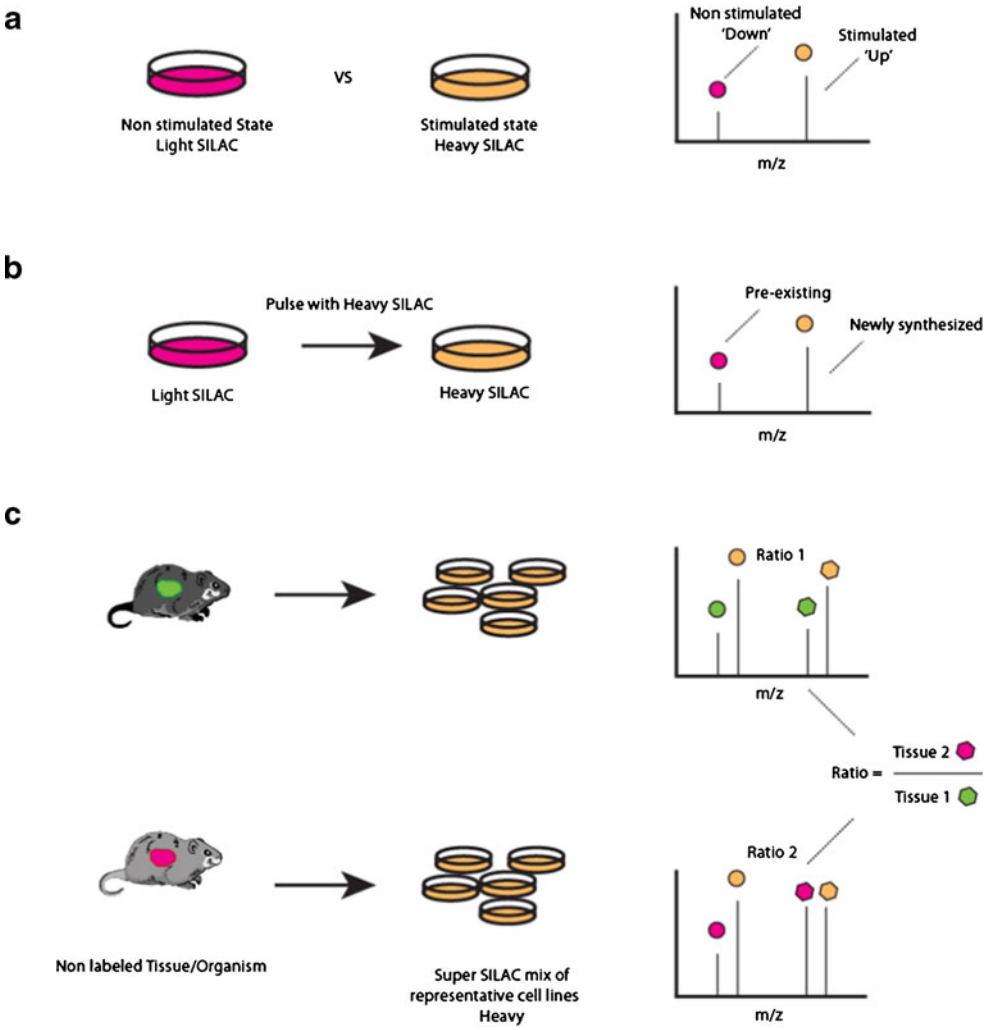

Fig. 2 The use of stable isotope labeling with amino acids in cell culture (*SILAC*) in different quantification strategies. **a** Original SILAC setup as introduced in 2002 by Ong et al. [14]. Cells are grown in normal medium (light) or are metabolically labeled with stable-isotope-labeled amino acids (heavy). Following, e.g., an external stimulus, cells are mixed and samples are measured together by liquid chromatography (LC)–tandem MS (MS/MS). The ratio of heavy to low signals in peptide mass spectra is a measure of the relative change in protein expression in response to the external stimulus. **b** In the pulsed SILAC approach, cells are grown in normal medium and pulse-labeled for a certain time with heavy SILAC. The ratio of heavy to low signals in such experiments is a measure of protein synthesis and degradation. **c** In the super SILAC strategy, unlabeled tissues from, e.g., an animal model system, are compared with a representative mixture of heavy-SILAC-labeled cell lines that act as a common internal standard. The ratio of two (or more) such experiments allows comparison of protein expression between animals

growing cells in a 1:1 mix of heavy and light SILAC and subsequently transferring them to 100 % heavy SILAC. In this way, it becomes easier to measure proteins with slow turnover without compromising the measurement of proteins with high turnover. Because changes in protein translation and turnover can be measured for thousands of proteins in a single experiment, numerous studies have already taken advantage of the various forms of the pulsed SILAC method [20–22].

Another quite recent development is the application of SILAC-labeled cells for the quantification of proteins from animal or human tissue. In the so-called super SILAC method, a mixture of different SILAC-labeled cell lines [23, 24], serves as an internal standard for quantifying proteins in tissue (Fig. 2c). In contrast to ordinary SILAC, the super SILAC method is a two-step procedure in which, first, the proteins from each (unlabeled) tissue are quantified against the (labeled) reference standard and, second, the differences between the tissue samples are quantified by calculating the ratio of the two SILAC ratios obtained in the first step. The super SILAC method extends the classical SILAC approach to tissue analysis without the need for metabolically labeling the animal, which is a significant practical advantage. In addition, the use of a spike in the standard obtained from a mixture of cell lines rather than a from single cell line increased quantification accuracy because the combined cell lines better represent the total proteome of a tissue [23]. Although the approach offers great opportunities for large-scale relative quantification studies in tissues samples, the limitations lie within the need to produce a reference standard that does contain (ideally) all proteins present in the tissue to be analyzed.

Whereas SILAC is mostly used in cell culture (although even mice have been SILAC-labeled [25]), ^{15}N metabolic labeling is still used to analyze microorganisms, such as bacteria and yeast and, to a lesser extent, higher multicellular organisms such as *Arabidopsis thaliana, Caenorhabditis elegans*, *Drosophila melanogaster*, and mice (more comprehensively reviewed in [26]). In contrast to SILAC, which adds a fixed number of labels to a peptide (mostly just one

heavy Lys or Arg), the number of ^{15}N labels varies from peptide to peptide, thus complicating subsequent data analysis. This is likely why ^{15}N metabolic labeling is not so frequently used in proteomics. Finally, the SILAC approach has also been explored for the absolute quantification of proteins [27], the details of which are presented in "Absolute quantification."

Chemical protein and peptide labeling

Labeling of proteins and peptides via chemical means after biosynthesis has become very widely used in proteomics in recent years. We have discussed the many published approaches extensively before in this journal [3] and, therefore, we confine ourselves here to a select few that are either widely used or have shown significant potential. In chemical protein and peptide labeling methods, any reactive group may be modified using stable-isotope labels. However, most current techniques target the peptide/protein N-terminus and the ε-amino group of lysine. Perhaps the two most popular chemical labeling methods at this time are tandem mass tags (TMTs) and isobaric tags for absolute and relative quantification (iTRAQ), both of which target primary amines and are so-called isobaric tags [28–30]. These reagents are built such that peptides from differentially labeled samples have identical mass (isobaric) but can be distinguished following fragmentation inside the mass spectrometer by the differentially isotope encoded reporter ions in the lower mass range region of tandem MS (MS/MS) spectra. The intensities of the reporter ions form the basis for quantification. In contrast to most other labeling methods, isobaric labeling is well suited for the multiplexed analysis of different biological samples because (1) the reagents are available in up to eight different versions, (2) the complexity of LC separations is not increased because the labeled peptides are precisely co-eluted, and (3) the complexity of peptide mass spectra is not increased because the differentially labeled peptides are isobaric. As a result, the often observed reduction in the number of proteins that can be identified and quantified from multiplexed analysis using, e.g., SILAC, can thus be avoided. As the reporter ions of isobaric tagging have masses of less than 150 Da, the ion trap mass spectrometer as one of the major workhorses in proteomics could initially not be used for such analysis because this instruments cannot efficiently trap peptide fragment ions of low *m*/*z* values (typically 20-30 % of the peptide precursor *m*/*z*). Today, these issues have been overcome by methods such as pulsed Q dissociation with linear ion traps [31–33] as well as higher-energy collision-induced dissociation (HCD) [31, 34, 35] and electron transfer dissociation [36–38] with recent generations of Orbitrap instruments.

A common problem in isobaric labeling strategies is the interference of contaminating near isobaric ions in a sample that are co-isolated and co-fragmented and thereby compromise accuracy in protein quantification measurements (see also later) [31, 39]. Recently, two different methods have been introduced that address this issue. In the MS^3-based method published by the Haas laboratory [40], the most intense fragment ion of the MS^2 spectrum, which is used for deriving sequence information, is selected for subsequent fragmentation by HCD, from which reporter ion intensities are derived. This method effectively eliminates interference of near isobaric ions [40] but requires specialized software to connect the information from the different spectra. In the gas-phase purification method published by the Coon group [41], the precursor ion is first isolated and charge-reduced to increase the *m*/*z* differences between the interfering peptide species and the charge-reduced ion is then selected for HCD fragmentation, generating both sequence information and reporter ion intensities for quantification. Both methods have downsides notably in that they reduce the sensitivity and data acquisition speed.

Alongside the isobaric labeling methods discussed above, two noteworthy new approaches have gained popularity. The first is the mTRAQ label, which uses the same chemistry as iTRAQ but adds isotope labels to peptides such that these are distinguishable in peptide mass spectra rather than tandem mass spectra [42]. These reagents are mainly used for the rapidly growing area of targeted proteomic assays using selected ion monitoring (see later). A very recent report, however, indicated that iTRAQ labeling is superior to the mTRAQ method for global, large-scale protein and phosphopeptide quantification [43]. The second new method is stable-isotope dimethyl labeling, which offers a straightforward alternative for the chemical labeling of protein digests [44–47]. As in the methods described above, all primary amines of peptides (i.e., N-termini and Lys side chains) are labeled by formaldehyde in combination with reduction of the initially formed Schiff base using cyanoborohydride. Combining different isotopomers of the reagents theoretically results in many possible label combinations with distinct mass shifts. In practice, however, duplex and triplex reactions that introduce mass differences of a minimum of 4 Da are most commonly used to avoid issues with overlapping isotope envelopes. Dimethyl labeling has a number of features that make it attractive for many laboratories: first, the reagents are very inexpensive; second, the labeling reaction can be performed either in solution or directly in reversed-phase chromatography columns (or solid-phase-extraction devices), which avoids losses during sample workup; third, protein quantities from micrograms to milligrams can be labeled practically without a change in protocol or reagent cost. One limitation is the fact that deuterium is used as part of the label, which can lead to

small retention time differences of differentially labeled peptides during LC-MS/MS analysis. This can, however, be dealt with using data analysis software, such as MaxQuant, that integrates the signal over the whole chromatographic peak rather than deriving the relative quantities from a single MS^1 scan [7].

In principle, intact proteins can also be labeled by any of the above-mentioned methods. Although conceptually attractive (samples may be combined before digestion, thus avoiding sources of quantification error in this step), this route is rarely followed in practice [30, 48] because of issues of incomplete, variable, or nonspecific labeling as well as limiting the effectiveness of proteases such as trypsin and Lys-C. The above-mentioned factors typically increase sample complexity and often dilute the signal and thus render quantification more difficult than in peptide labeling.

LC-MS/MS analysis of peptides

Shotgun proteome digestion approaches generate vastly complex mixtures of (mostly tryptic) peptides that constitute the analytes for both the identification and quantification of proteins. The combination of nanoscale ion pairing reversed-phase LC and electrospray ionization (ESI) MS/MS is still by far the dominating analytical technology for this purpose. The past 5 years has seen notable advances in peptide separation and MS technology, and a number of trends have emerged in the use of these technologies for a further diversified set of applications. As far as LC separations are concerned, one can identify two major trends. The first trend is that two-dimensional LC separations are becoming more and more popular. These are designed to enhance the detection of posttranslationally modified peptides (by specific enrichment), to increase the extent of proteome analysis (i.e., detecting more peptides and thus proteins), or to improve quantification accuracy (less matrix interference). Whereas a few years ago strong cation exchange chromatography was primarily used as the first separation dimension (e.g., enriching acetylated and phosphorylated peptides in early fractions [49, 50]), techniques such as strong anion exchange [51] and high-pH reversed-phase chromatography [52] are increasingly used to boost the number of identified and quantified proteins in a proteome. At the same time, titanium dioxide chromatography [53] has developed into the second standard for the enrichment of phosphopeptides alongside the classical immobilized metal affinity chromatography (IMAC) technique [54]. Hydrophilic interaction chromatography is being increasingly explored to achieve high proteome coverage and phosphopeptide enrichment at the same time [55, 56]. In addition to expanding the accessible proteome space, two-dimensional (or multidimensional) LC separations also reduce sample complexity and, therefore, also often improve quantification accuracy because issues with overlapping isotope signals in mass spectra of high peptide complexity can be substantially reduced [57]. Somewhat against earlier expectations, two-dimensional LC separations are currently mainly performed “off-line” rather than “online.” This has mainly practical reasons; the physical separation of the two chromatographic steps allows more flexibility in terms of matching sample quantities and solvent systems between the separation dimensions. The second major trend in LC separations concerns the reversed-phase chromatography system that is directly coupled to the mass spectrometer. Here, the use of long columns (more than 20 cm), packed with small particles (3 μm or less), and long gradient times (2 h or more) is becoming more and more popular as a means to increase peak capacity [58]. This in turn requires the use of ultra-high-pressure LC (UPLC) systems or standard high-pressure LC (HPLC) equipment operating at higher temperature (40-60 °C) in order to deal with the resulting high backpressure [59–61]. Average chromatographic peak widths of 4-10–s (full width at half maximum) can now be obtained fairly routinely (Fig. 3), which not only increases the separation power [62] but also increases the sensitivity and quantitative accuracy of the MS detection system (higher sample concentration). A currently emerging approach with significant future potential is to perform separations in a chip-based format [63–65]. One expects these will eventually become attractive for nonexpert users and for applications with high numbers of similar samples. However, presently the consumable costs for such chips are still quite high and not all available systems can be combined with all mass spectrometers. We will therefore have to wait to see if the conceptual advantages of chip systems can be realized in practice.

The performance of mass spectrometers in terms of sensitivity, speed, mass accuracy, and resolution is continuing to rise at a rapid rate and there is currently no sign of this slowing. This is particularly relevant for quantitative proteomics as most investigated proteomes are very complex (in terms of both numbers of proteins and dynamic expression range) and sample availability is low for many relevant biological systems under investigation. The main platforms for quantitative proteomics today are orbital traps, QTOF instruments, and triple-quadrupole instruments mostly using ESI as an interface for chromatographic systems and collision-induced dissociation as the peptide fragmentation technique (with electron transfer dissociation playing a useful complementary role for large, posttranslationally modified or cross-linked peptides [66]). Together, these instruments enable the characterization of proteomes to a depth of 5,000-10,000 proteins [5, 6]. Today, the two main approaches for the quantification of peptides and proteins from LC-MS data are (1) extracting the LC-MS intensities

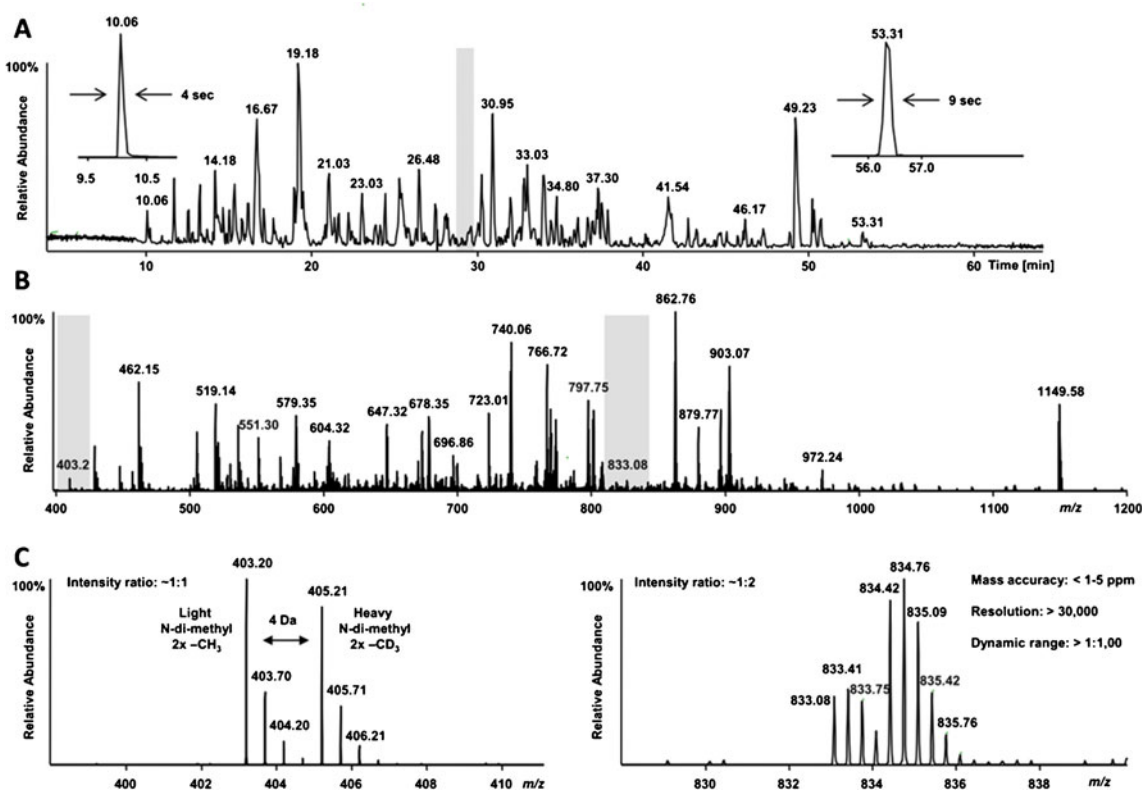

Fig. 3 Examples illustrating advances in LC and MS for isotope labeling for quantification methods. **A** Ultra-high-pressure LC of complex proteome digests increases the separation power (peak widths of 4-10 s, full width at half maximum), sensitivity, and accuracy of quantification. **B** Peptide mass spectrum at 29-min retention time showing that dozens of peptides are eluted at the same time, often negatively affecting the measurement speed of even the most modern instruments. **c** Expanded regions of the peptide mass spectrum showing dimethyl-labeled peptide precursor ions (4-Da mass shift; in either 1:1 ratio, *left*, or 1:2 ratio, *right*). Owing to the very high mass resolution, low mass error, and high dynamic signal intensity range of modern mass spectrometers, even low-abundance peptides can be accurately quantified

of peptide precursor ions in the classical data-dependent aquisition (DDA) experiment in which peptide precursors are subjected to fragmentation (sequencing) as they are eluted from the LC system and (2) extracting the LC-MS/MS intensities of peptide fragment ions (or iTRAQ/TMT reporter ions) of peptides either from DDA experiments or from so-called selected reaction monitoring experiments (SRM; also often referred to multiple reaction monitoring, MRM) in which LC-MS monitors a predefined list of peptide and fragment ions corresponding to the proteins of interest (discussed in more detail later and recently reviewed in [67]). The main merits of the two approaches are that in the DDA experiment, there is no need to define the peptides for quantification before doing the experiment (needed for SRM), and the SRM experiment is exquisitely sensitive (because LC-MS collects the signal for a given peptide even if the peptide precursor cannot be detected above the noise). Both approaches can be highly selective and yield accurate quantification data (coefficients of variations below 20 %), which is mainly determined by the peptide concentration in a sample and the complexity of the proteome analyzed (the more complex the proteome, the less accurate the quantification of all proteins, particularly those of lower abundance). Despite significant progress, further improvements are still required because the molecular complexity and dynamic expression range of proteomes still overwhelms current LC-MS systems [68] despite their now having sensitivity in the attomole range, total peak capacities of more than 100,000 peaks per hour, a sequencing speed of two to ten peptides per second, and sophisticated on-the-fly decision making algorithms [68–71].

It has been argued that to overcome these issues, the MS data acquisition regimen needs to change from the current sequential mode of operation (data-dependent, one peptide at a time) to a parallel mode of operation (data-independent, many peptides at a time) akin to modern DNA/RNA sequencing methods [72]. Some steps have been taken toward this, e.g., by submitting a range of co-eluted peptide precursor ion masses (say, 10-100 Th) to simultaneous fragmentation [69, 73, 74] or, indeed, all peptides eluted at the same time [75, 76]. Peptide identification is achieved in such experiments by aligning the chromatographic retention times of the peptide precursor and fragment ions (which must be precisely the same), and quantification can then

be performed by using either the precursor ion or the fragment ion intensities as discussed earlier in this section. The available literature indicates that these ideas work in principle, but the challenge clearly is the specificity of identification and the accuracy of quantification derived from mixed peptide and fragment mass spectra [77, 78]. High-resolution mass spectrometers should have inherent advantages in this setting because of the high mass accuracy obtained for peptide precursor and fragment ions, and further specificity can be provided by the use of additional techniques such as ion mobility spectrometry, which takes advantage of the effects of differences in secondary structure (shape) of peptides on the mobility of ions in the gas phase [79]. Albeit still a niche, parallel data acquisition has great potential and it will be interesting to follow how these approaches develop (for further aspects of the above-mentioned techniques, see the next sections on label-free quantification).

Label-free quantification—spectrum count approaches

As the name implies, label-free quantitative methods in MS aim at quantifying peptides and proteins without the use of stable-isotope labels. This rapidly growing area in proteomic technology may be divided roughly into methods that (1) directly utilize a peptide's response (intensity) in the mass spectrometer as a quantitative measure and (2) infer quantity indirectly from, e.g., the number of peptide-to-spectrum matches (PSMs; spectrum count) obtained for each protein.

Spectrum count approaches (Table 1) are based on the observation that the number of PSMs, the number of distinct peptides identified, and the sequence coverage obtained for a protein of interest in an LC-MS/MS experiment correlate with protein quantity. Among these three measures, the spectrum count has been shown to offer the highest dynamic range of quantification and the best reproducibility [80–82]. Since we last reviewed the subject, spectrum count approaches have gained in popularity, in particular for projects where large sets of experimental data need to be compared (a limitation of all labeling methods). Relative quantification of protein abundance in sets of samples is straightforward since it basically only requires comparison of the spectrum counts for each protein and this information is typically generated during database searches for protein identification using conventional search engines. However, a number of additional considerations should be taken into account. As any sampling approach, spectrum-count-based quantification requires a sufficient number of data points [83]. This is influenced by experimental parameters such as dynamic exclusion of precursor ions, chromatographic peak width, and MS instrument scan speed. In addition, protein-specific parameters are important, such as the length distribution of the peptides generated by proteolytic cleavage (which is different for every protein). This determines how many different peptides can be detected within the mass range applied for the MS experiment and thus directly influences the number of measurable spectrum counts. As a consequence, relative quantification of very short (i.e., less than 20 kDa) proteins tends to be more variable than that for equal amounts of longer proteins. In addition, saturation effects occur at high protein concentrations, thus limiting the dynamic range of detectable protein expression.

Once protein identification has been performed, one has to decide how PSMs should be filtered for protein quantification—a task that is far from trivial. Intuitively, one might apply filters similar to those used for protein identification (say, 95 or 99 %) [81]. Indeed, limiting the spectrum counts to high-confidence PSMs has been shown to enable statistically significant detection of smaller changes compared with applying less stringent criteria (e.g., threefold vs. fivefold for an abundant protein in [84]). However, a higher dynamic range and more robust quantification of low-abundance proteins has been observed when including low-scoring PSMs [84, 85]. Another important question associated with the protein identification process is if or how to weigh spectrum counts for peptides that are shared between proteins or protein isoforms (the so-called protein inference problem) [86, 87]. Some insight has come from experiments using six albumins from different species added to a yeast lysate digest. The results indicate that the most accurate and reproducible results were obtained when shared spectrum counts were distributed to the different albumins proportionally to the number of unique spectrum counts obtained for each albumin [88].

Since all samples in a label-free study are separately analyzed by LC-MS/MS, stable and comparable operation has to be ensured throughout. This is usually best accomplished by running all samples in a single sequence in the same instrument. Nonetheless, some variation even for replicate runs of the same sample is often observed and might be attributed to small differences in the amounts of loaded sample, variation in the chromatographic gradient, and the semirandom nature of data-dependent data acquisition. A very simple way to account for these variations is to normalize the data for the total number of spectrum counts determined in each sample [83, 89]. Asara et al. [90] added an intensity dimension to spectrum counting and showed that division of the spectrum count by the average total-ion count in the matching tandem mass spectra resulted in the accuracy and dynamic range of quantification being substantially improved and saturation effects often observed for very high abundance proteins being removed. For example, LC-MS/MS analyses of 50 fmol and 1 pmol of a tryptic bovine serum albumin digest yielded only a fourfold difference in the spectrum counts but a 20.8-fold difference with

Table 1 Selected methods for label-free quantification based on spectrum counting

Method	Principle	Comments	References
Peptide count	Use of the number of peptides for an identified protein as a measure of abundance	Less useful than spectrum count	[80]
Spectrum count	Use of the number of PSMs for an identified protein as a measure of abundance	Higher dynamic range and better reproducibility than peptide counting	[81]
MS/MS intensity	Average total intensity of all fragment ion spectra matched to a protein	Intensity dimension provides extended dynamic range and better accuracy than spectrum counting	[90]
PAI and exponentially modified PAI	Protein abundance index and exponentially modified PAI (10^{PAI-1}). PAI is the number of identified peptides divided by the number of observable peptides	Implemented in some search engines such as Mascot. Designed for absolute quantification	[165]
Spectrum count/ molecular weight	Spectrum count is divided by the molecular weight of a protein	Similar to SAF	[162]
SAF	The spectrum count is normalized for protein length	Absolute and relative quantification	[163]
NSAF	SAF normalized for the sum of all protein abundances in the sample	Absolute and relative quantification	[89, 220]
RSC	Includes normalization of run-to-run variations and a correction factor	Relative quantification	[83]
APEX	Improves spectral counting by focusing on counting peptides that will likely be detected by MS techniques	Uses machine learning classification to derive peptide detection probabilities. This can be used to predict detectable peptides for any protein. Absolute and relative quantification	[145, 166]
SIn	Combines spectrum count with fragment-ion intensity (sum of all fragment ion intensities of all PSMs for a protein)	Variants include normalization by the sum of all protein spectral indexes in an experiment and normalization for the length of the protein. Absolute and relative quantification	[91]
mSCI	Number of observed peptides divided by protein relative identification probability	Conceptually similar to APEX. Absolute and relative quantification	[231]
RIBAR	Average of log 2 peptide ratios for a protein. Peptide ratios are calculated using the sum of all fragment ion intensities across all PSMs	Pairwise relative protein quantification	[92]

APEX absolute protein expression, *MS* mass spectrometry, *mSCI* modified spectrum count index, *MS/MS* tandem mass spectrometry, *NSAF* normalized spectral abundance factor, *PAI* protein abundance index, *PSM* peptide-to-spectrum match, *RIBAR* robust intensity-based averaged ratio, *RSC* relative spectral count, *SAF* spectral abundance factor, *SIn* normalized signal intensity

the intensity-adjusted method. Including intensity information of PSMs was also reported to be an efficient means to account for run-to-run variation in replicate experiments [91]. For each protein, a spectral index was defined as the sum of all fragment ion intensities of all PSMs in an LC-MS/MS experiment and was divided by the sum of all protein spectral indexes obtained in the experiment. When compared with normalization strategies omitting the intensity dimension, the spectral index strategy achieved superior normalization. The approach was further refined by summing up the fragment ion intensities of all PSMs for each peptide. Differential display of two experiments in this setting was first performed at the peptide level (considering only peptides identified in both experiments), and relative protein quantification can then achieved by averaging the log 2 ratios obtained for peptides matching a protein [92]. Obviously, from intensity refinements of the spectrum count idea, it is only a small step to using solely the intensity dimension of a detected peptide for quantification, which is the subject of the next section.

Label-free quantification—MS^1-intensity-based approaches

Intensity-based label-free quantification employs the MS signal response of intact peptides and, by inference, that of proteins for quantification. In bottom-up LC-MS/MS

experiments, this is typically accomplished by integrating the ion intensities of any given peptide over its chromatographic elution profile. Early work by Bondarenko et al. [93, 94] indicated linearity of the signal response over four orders of magnitude and excellent correlation with the concentration of the measured peptide. In comparative experiments, the integrated signal response of individual peptides is compared between LC-MS(/MS) runs of the different samples. Changes in protein abundance between the samples can then be estimated by aggregation of differences measured for all peptides (or a subset thereof) matching the protein of interest. This aggregation may take the form of averaging peptide fold changes or summing up peptide responses (see "Analysis of quantitative MS data" for more details). In contrast to stable-isotope labeling, the number of conditions that can be compared is not predefined (i.e., by the number of different isotope labels). Hence, and similarly to spectrum counting, intensity-based relative quantification is particularly attractive when many samples are to be compared.

Experimental considerations

Despite the conceptual simplicity of the overall approach, several experimental considerations have to be taken into account to ensure robust quantification. Typically, full-scan survey spectra are used to generate extracted ion chromatograms (XICs). Conversely, peptide identification is based on tandem mass spectra that are acquired either in the same or in additional LC-MS/MS runs. In either case, it is essential that the individual XICs can be unambiguously assigned to the respective peptides. To this end, the use modern mass spectrometers with high resolving power and high mass accuracy is advantageous since the former minimizes the influence of interfering signals from co-eluted peptides of similar mass and the latter ensures reliable mapping of XICs to identified peptides.

A robust UPLC/HPLC-MS setup is another important prerequisite for accurate quantification. Narrow LC peak widths produce better signal-to-noise ratios also for low-abundance ions, thus extending the dynamic quantification range [95, 96]. Retention time stability (on the order of 1 %) is generally advantageous, particularly in cases where the identity of a peptide could not be established in all experiments and XICs are therefore mapped solely by aligning accurate masses and retention times. A further complication arises from the fact that the ionization efficiency of a peptide is strongly influenced by the presence of co-eluted species (peptides and contaminants) [97]. Therefore, shifts in retention time can also affect ionization efficiency and thus introduce a bias which may affect the accuracy of quantification. In short, a robust LC-MS setup minimizes the challenges for chromatographic peak alignment software, thus leading to fewer false assignments and more accurate quantification.

Many current LC-MS data acquisition schemes aim to achieve protein identification and quantification simultaneously. In the conventional DDA regimen, each survey scan (MS^1, relevant for quantification) is typically followed by multiple tandem mass spectra (relevant for peptide identification). However, care must be taken to find the right balance. On the one hand, precision of quantification is aided by a large (i.e., eight to ten) number of MS^1 scans across an LC peak for reconstruction of the XICs, but, on the other hand, maximal protein coverage is achieved by collecting as many tandem mass spectra as possible. Hence, maximizing proteome coverage usually sacrifices quantification and vice versa [98]. Consequently, the resolution of the LC system together with the data acquisition speed of the mass spectrometer dictate where the right balance can be found. For example, modern QTOF and orbital scanning mass spectrometers provide productive scan rates of approximately ten scans per second and LC peaks of UPLC separations are typically 10 s at the base. Therefore, LC-MS/MS rates of top10 should be feasible.

As an alternative to the classical DDA experiment, the Smith laboratory suggested performing separate runs to establish the identity and for quantification [99–102]. In this workflow, samples are first analyzed by LC-MS/MS and the peptide identification results are stored in a database. A second set of experiments is then performed in LC-MS mode to obtain optimal quantification data. Software tools such as VIPER are then employed to match the LC-MS features (retention time and mass) to the database of identified peptides. This so-called accurate mass and retention time tag approach [101, 102] is of particular utility when large pools of complex samples are analyzed for quantitative changes such as blood plasma in biomarker studies [103].

Another alternative to data-dependent fragmentation of precursor ions is data-independent switching between low and high collision energy conditions in the mass spectrometer whereby all co-eluted peptides are simultaneously fragmented [75, 104, 105]. Specialized software then correlates the elution profiles of fragment masses to reconstruct tandem mass spectra and enable peptide identification and quantification at the same time [79, 106]. Recent studies indicate that this strategy enables similar or better proteome coverage to be achieved compared with DDA with a QTOF platform without compromising the precision of quantification [107, 108]. In principle, data-independent approaches also allow the use of fragment ion intensities for label-free quantification, rather than relying on the signal response of the intact peptide ions. For example, Gillet et al. [69] employed a fast scanning, high-resolution double-quadrupole time-of-flight instrument to cycle repeatedly through 32 consecutive 25-Da precursor isolation windows

(so-called swaths). This acquisition setup generates time-resolved fragment ion spectra for all analytes detectable within the 4,00-1,200 m/z precursor range and a user-defined retention time window. Combinations of fragment ions can then be used to identify and quantify peptides over a large dynamic range (four orders of magnitude) and the queried peptides can be quantified with reproducibility and accuracy similar to those in SRM (see later).

Software considerations

In contrast to spectrum counting, intensity-based quantification requires additional data extraction and processing. Procedures typically involve mass calibration, noise and data reduction steps, followed by feature detection and generation of peptide elution profiles. In a next step, peptide features from different LC-MS/MS experiments are aligned in the time and mass dimensions, and intensities are normalized. Finally, relative quantification is achieved by matching the intensities of the individual peptide features across all experiment in the set. An array of software solutions has been developed for this purpose. Public, free tools include OpenMS [109–111], Viper [101], Maxquant [7, 112], Mapquant [113], Census [114], and Superhirn [115]. Commonly used commercial software packages include Genedata Expressionist, GE Decyder MS, Progenesis LC-MS, and Mascot Distiller. In addition, several manufacturers of MS instrumentation have implemented label-free quantification capabilities into proprietary software packages. For a more detailed discussion of available algorithms and packages, we recommend some excellent specialized review articles on this subject [73, 116–118].

Recent comparative studies indicate a reasonably good correlation between intensity-based quantification and spectrum counting approaches. At the same time, intensity-based methods offer better overall performance [119, 120], particularly when high-performance (mass accuracy, mass resolution) mass spectrometers are employed. As these and specialized software and statistical methods (see later) are becoming increasingly available, intensity-based label-free methods will continue to gain utility simply because more and more proteomic applications require the analysis of many samples.

Selected reaction monitoring

In the last 5 years, enormous progress has been made in establishing SRM as a technology for precise and accurate quantification in proteomics. Although it was initially limited to the simultaneous measurement of a small number of peptides in complex samples, innovations in instrumentation and software have now allowed the detection and quantification of hundreds of analytes in a single experiment. Recent reports have demonstrated the use of SRM as a directed discovery tool for a diverse set of applications ranging from biomarker research to protein interaction networks, analysis of signaling pathways, and comparative proteomics [5, 121–123]. Targeted SRM approaches determine the presence and quantity of a defined set of peptides by monitoring the generation of (multiple) fragment ions upon collision-induced dissociation. The pairs of precursor and fragment m/z values are referred to as "transitions." Such transitions are detected with very high sensitivity in triple-quadrupole mass spectrometers by using the first and the third mass analyzers as mass filters and monitoring the signal of the fragment ion over the chromatographic elution time. Thus, each transition effectively constitutes an independent MS assay that allows one to identify and quantify a specific peptide and, by inference, the corresponding protein in complex matrices. The technique has been shown to provide very high sensitivity (femtomole to attomole range) and a dynamic quantification range covering four to five orders of magnitude [124].

Setting up SRM assays for a set of proteins is a relatively slow and iterative process, and the considerable effort in assay development is therefore mostly undertaken if a defined (and reasonably small) set of peptides has to be measured in a large number of samples. Building an SRM assay essentially comprises the following steps: (1) selection of the most suitable proteotypic peptides, (2) identification of suitable fragment ions thereof, (3) optimization of instrument parameters for each transition, (4) validation of the specificity of the assay in the matrix under investigation, (5) determination of the lower limits of detection and quantification in that same matrix, and (6) compilation of the final assay parameters, including precursor m/z, fragment ion m/z, elution time, collision energy, and dwell time, for all peptides of interest. We discuss some of these in the following, but see some excellent specialized articles on the topic for more details [67, 124, 125]. Suitable peptides for SRM assays must be unique for the protein of interest to avoid the generation of misleading results [126]. Such peptides are frequently selected on the basis of their consistent identification in previous experiments. In addition to in-house generated data, public data repositories such as PeptideAtlas [127, 128] and PRIDE [129–131] are excellent and constantly growing resources for SRM peptides. If experimental data are not available, computational approaches allow prediction of the best responding peptides with reasonable precision [132, 133]. Peptides containing amino acids prone to chemical modifications (Cys, Met, Trp) and miscleaved peptides are frequently avoided in order to avoid ambiguous quantification results. If proteins are present in different isoforms or are otherwise posttranslationally processed, multiple peptides covering both conserved and variable

regions of the protein sequence should be selected to quantify variants [124, 126]. In a next step, transitions to high-abundance fragment ions are selected to maximize sensitivity. Traditionally, suitable fragment ions are selected from full tandem mass spectra acquired with a comparable instrument/collision cell and this often involves synthesis of the peptide of interest. More recently, the aforementioned database tools (PeptideAtlas, SRMAtlas) [134] have become invaluable resources for method development, particularly for highly multiplexed SRM assays, and often also provide dedicated software solutions to assist in the selection process [135] (for a recent review on available software solutions, see [136]). A shortcoming of almost any data stored in public repositories is that the experimental conditions under which the reference data were acquired are not comprehensively listed and will likely not perfectly match those in another laboratory. Calibration of collision energies, pressure of the collision gas, geometry of the collision cell, etc. all influence the relative abundance of fragment ions to some extent. Consequently, SRM assays solely built on theoretical data or data from public repositories might not achieve the best possible sensitivity and specificity.

Perhaps the most fundamental parameter in building an SRM assay is its selectivity (i.e., does the assay actually measure what it is supposed to measure?). Despite the power of the aforementioned theoretical methods, the selectivity of each transition needs to be evaluated in the sample matrix under investigation (e.g., a tryptic digest of a cell extract or blood plasma) in order to learn about the contributions of fragment ions arising from near-isobaric co-eluted peptides (or other components). The identity of a particular peptide is typically determined by monitoring the retention time of individual transitions in SRM traces, which can be assessed computationally by cross-correlation analysis [137]. However, full confirmation of peptide identity typically requires triggering full MS/MS spectra once a specific transition reaches a predefined threshold [138]. This requirement significantly impaired the development of multianalyte SRM assays using triple-quadrupole instruments, which inherently suffer from relatively slow scan speeds. The intelligent SRM approach provided a noteworthy solution to this problem [139]. In this data-dependent SRM approach, multiple confirmatory transitions (eight to ten) are triggered close to the peak apex of the actual assay transition. This generates a pseudo MS/MS spectrum that can be matched to a reference library, which greatly facilitates the confirmation of a peptide's identity and allows monitoring of a relatively high number of analytes at any point in time.

For large-scale SRM experiments, the number of analytes that can be actually monitored with good sensitivity, precision, and accuracy depends on the number of transitions measured for each peptide, the time allocated for measuring each transition (dwell time), and the lag time between transitions (interscan delay, required to adjust instrument parameters and to avoid cross-over between transitions; see also [125]). An obvious way in which the number of analytes monitored can be increased is to partition SRM experiments into small time segments during which only those peptides are measured that are eluted within that time window (time-scheduled SRMs [140]). As long as the retention times of the targeted peptides are known (easily established by prior experiments), the number of analytes monitored can be drastically increased (say, ten to 50 times) without compromising sensitivity. The combination of time-scheduled SRMs with the aforementioned intelligent SRM approach now allows hundreds of peptides to be quantified in a single experiment with confidence and sensitivity. Still, the complexity of proteomic samples impacts sensitivity and selectivity of SRM experiments performed with triple-quadrupole instruments. Further selectivity may be gained by including an ion mobility separation step into an LC-SRM analyses [141]. In addition, Fortin et al. [142] recently introduced an MS^3-based SRM workflow coined MRM^3 that significantly extended the dynamic range and limit of quantification by a second collision-induced-dissociation step of the monitored fragment ion in a triple-quadrupole linear ion trap instrument. At this point, the ion mobility and the MS^3 SRM techniques are not yet ready for high throughput but their positive impact on the selectivity of SRM assays has been convincingly demonstrated. It can be argued that further selectivity can be gained by high mass resolution and accuracy as this will drastically reduce the number of possible transition conflicts. The recent introduction of QTOF and quadrupole–Orbitrap instruments with high data acquisition speed may well provide this capability, with the added attraction that these mass analyzers collect the entire fragment ion spectrum, so the aforementioned measurement of confirmatory transitions would no longer be required. Recent data suggest that such high-resolution SRM techniques actually outperform the triple-quadrupole techniques in terms of the limit of detection and the limit of quantification (Bruno Domon, personal communication).

Absolute quantification

The aforementioned approaches all lead to the relative quantification of proteins across samples, but no information is obtained regarding their absolute quantities or concentrations in a biological sample because the MS response is different for every (peptide) analyte and is influenced by the sample matrix. Still, MS-based methods for quantification have been successfully devised. Whereas this area was still underdeveloped some 5 years ago, it is maturing at a rapid rate and has already led to noteworthy biological

applications such as kinetic modeling of biological processes [143, 144], the calculation of protein half-lives and translation rates [145, 146], and determination of protein complex stoichiometries [147–150].

Absolute quantification using stable-isotope-labeled standards

Adding stable-isotope-labeled reference standards of known absolute quantity to a sample allows one to determine the absolute amounts of peptides (and proteins by inference) by comparing their MS signal intensities with those of the standards [151, 152]. Several methods utilizing such reference peptides and proteins exist, including the more established methods of AQUA (for "absolute quantification") [151] and QconCAT [153], as well as more recent approaches such as protein standard absolute quantification (PSAQ) [154], absolute SILAC [27], and FlexiQuant [155].

For the AQUA strategy, proteotypic peptides [72] are identified for a protein of interest, then synthesized to include heavy amino acids or stable-isotope-containing tags [147]. These synthetic peptides are then quantified, e.g., using amino acid analysis to determine their quantity, and subsequently added to a biological sample during or after proteolysis [151, 152]. Today, AQUA peptides are commercially available from several sources. Albeit relatively expensive, AQUA peptides are widely employed for absolute quantification particularly for targeted approaches such as SRM assays. However, the reference peptides are only added relatively late in the biochemical workflow and thus do not account for the often significant systematic errors (e.g., sample losses) during the upstream sample preparation steps. The best results are therefore obtained if sample preparation is kept to an absolute minimum. The QconCAT technology uses recombinant DNA techniques to construct synthetic proteins in which large numbers of internal standard peptides are concatenated. The synthetic protein can be expressed and purified in stable-isotope-labeled form from a suitable host (*Escherichia coli*). The QconCAT protein is then added to a sample and the reference peptides for the desired proteins are generated in situ during the protease digestion step. This allows not only the controlled addition of many reference peptides, but also removes systematic errors associated with the digestion step as these affect the sample and the reference standard in the same way. The utility of the approach has recently been demonstrated by quantifying 27 enzymes in the glycolytic pathway of *Saccharomyces cerevisiae* using a 88-kDa QconCAT protein encoding 59 tryptic peptides [156]. Instead of using peptides as standards, the PSAQ, absolute SILAC, and FlexiQuant approaches utilize metabolically labeled full-length proteins expressed in cell-free or bacterial systems. The purified and precisely quantified protein standards can be added directly to cell extracts, enabling extensive sample fractionation without compromising quantification results. In a recent application a PSAQ strategy was employed to quantify the staphylococcal virulence factor enterotoxin A in serum, and coefficients of variation between 0.05 and 0.18 with an SRM-based approach and a lower limit of quantification in the range of 1 ng/ml were achieved [157]. Given the significant effort in generating the standards, these approaches are particularly useful for focused projects in which only a small number of proteins need to be quantified. However, a recent study reported a modified absolute SILAC approach allowing higher throughput in generating standards for large-scale studies [158]. Here, the authors made use of protein epitope signature tags (PrESTs), short stretches of (isotope-labeled) protein sequences, which are produced in high numbers by the Human Protein Atlas project for use as antigens for antibody production [159–161]. Many such PrESTs can be added to a sample simultaneously to quantify the respective proteins. Because PrESTs are not full-length proteins, they are not suitable for addition before protein fractionation.

Absolute quantification using label-free methods

In their pioneering work in 2001 on multidimensional peptide separation for large-scale protein identification, Washburn et al. [80] observed that the codon adaptation index of yeast proteins (a measure of abundance) correlated with the number of identified peptides of a given protein. This suggested that absolute quantification of proteins may be possible without the use of stable-isotope-labeled reference standards. We have already discussed that the spectrum count (or PSM) is a better quantification measure than the peptide count, and several research groups have hence explored spectrum counts for absolute quantification. As discussed earlier, larger proteins are expected to generate more peptides and therefore higher spectrum counts than smaller proteins of equal abundance. Consequently, normalization of spectrum counts on this basis could take the form of dividing PSMs by the molecular weight of the proteins [162], or by the number of amino acids in the primary sequence [163] (termed spectral abundance factor). Similarly, Rappsilber et al. [164] defined a protein abundance index (PAI) by dividing the number of observed peptides by the number of theoretically observable unmodified peptides. The exponentially modified PAI (equal to 10^{PAI-1}) was further shown to be proportional to the protein content of a mixture [165]. A more sophisticated algorithm termed APEX utilizes machine learning classification based on experimental data to estimate the detection probabilities of tryptic peptides of any given protein. This can then be used to predict the number of tryptic peptides expected to be detected for one molecule of a particular protein, and this value is compared with experimental data to estimate absolute protein amounts [145, 166]. APEX-based quantification proved to be reasonably accurate

(within an order of magnitude) for global proteomic profiling and also showed good correlation to other quantitative methods [145, 167]. As for relative quantification approaches, the implementation of an intensity dimension to spectrum counting also leads to a higher dynamic range for absolute quantification. For example, normalization of the aforementioned spectral index by protein length [91] allowed the quantification of a 19-protein mixture over a range of 0.5-50,000 fmol with a correlation coefficient (R^2) of 0.92. Trudgian et al. [168] recently applied a slightly modified version of this approach to quantify 19 proteins of universal protein standard (UPS2, Sigma) and obtained similar results. The successful implementation of such methods requires rigorous validation in each individual laboratory and for any sample type using suitable standards.

Methods using signal intensity rather than spectrum counting are increasingly being recognized as viable approaches for absolute quantification. Schwanhaeuser et al. [146] demonstrated that dividing the sum of all observed peptide intensities by the number of theoretically observable peptides provides a useful measure to approximate absolute protein concentration. Alternatives to this idea consider only, e.g., the three most intense peptides of a protein (high3 or top3 method) [105, 120, 169]. These "best-flyer" approaches [170] are based on the assumption that a small number of proteotypic peptides (unique, frequently observed) exist for each protein of the proteome and that the MS signal intensities of these peptides are approximately the same. Using the ion intensity provides better quantification accuracy over a larger dynamic range than spectrum counting (less undersampling and saturation) [120]. However, the robustness of the methods hinges on the reproducible identification of the very same peptides across all samples in the analysis and on reliably matching MS/MS spectra to chromatographic peaks. Ludwig et al. [170] recently ported the high3 method to SRM platforms and found the best precision and accuracy when the two most intense transitions of the three most abundant peptides were used to calculate protein abundance. In light of the above facts, label-free absolute quantification is possible and very attractive as it is easy to implement, is compatible with high-throughput analysis, and can, in principle, even be applied retrospectively. However, we stress that label-free methods are still a lot less accurate than stable isotope labeling and should therefore be used if only a rough estimation of protein quantity suffices.

A specialty in absolute quantification analysis of proteins is the use of inductively coupled plasma (ICP) MS (ICP-MS). In this technique, a (protein) sample is completely atomized in an argon plasma and the heteroatoms of the proteins (either natural or synthetically introduced S, P, Se, I, As, or lanthanides) are ionized and can be detected and quantified by MS [171, 172]. However, the ICP source requires pure samples for biospeciation analysis and is therefore not compatible with the typically complex mixtures encountered in proteomic research. Recent work utilized ICP-MS for absolute quantification of peptides or proteins that either contain the above-mentioned heteroatoms (e.g., P, S) or have been modified with heteroatom-containing tags (e.g., metal-coded affinity tags [173, 174]). Precisely quantified peptides and proteins can then be used as standards for absolute quantification using mass spectrometers equipped with standard ion sources [175–177]. Alternative applications of ICP-MS for absolute quantification of proteins and peptides relied on the combined use of elemental and molecular MS techniques, i.e., by analyzing the samples separately with LC-ICP-MS and LC-ESI-MS [174, 178].

Quantification of PTMs

As interest in PTM analysis of proteins rises, so does the need to quantify PTMs in a multitude of experimental settings. The analytical challenges are immense and to a large extent reside in the often low abundances and stoichiometries of these peptides and the fact that measuring the changes in relative or absolute abundance can only be based on a single peptide. To address abundance and stoichiometry, specific enrichment is generally required to be able the detection of modified peptides from the high background of nonmodified peptides. These steps can of course introduce systematic error or bias, which must be reduced to an absolute minimum. In addition, PTM quantification requires correction by protein quantification to avoid misinterpretations arising from changes in protein expression over the course of an experiment. In light of the above facts, good precision and good accuracy of PTM quantification are much harder to achieve than for protein-level quantification. Practically all of the aforementioned quantification techniques have also been applied to PTMs, mostly protein phosphorylation. The SILAC approach is, of course, attractive for work in cell culture systems because samples can be combined at the earliest possible step and all subsequent steps are performed on the pooled samples, thus minimizing systematic errors (Fig. 1). Consequently, SILAC has been used extensively and in many cellular systems over the past few years [179–184] (reviewed in [185]). Chemical peptide labeling strategies share some of this advantage for shotgun proteomics, where labeling is performed after protein digestion, which is essentially the first step after cell harvesting and lysis, and all further enrichment steps can be performed on the pooled samples. If, however, protein fractionation is required, chemical labeling rapidly loses ground because protein fractionation techniques tend to suffer from large variations from experiment to experiment. Peptides

derivatized by any of the common labels can be subjected to PTM-specific chromatography, and quantitative phosphoproteomics using, e.g., dimethyl labeling is now performed routinely on both phosphopeptides enriched by both IMAC and TiO_2 chromatography [186, 187] as well as affinity-purified phosphotyrosine peptides [188, 189]. Likewise, iTRAQ- and TMT-based quantitative phosphoproteomics experiments have also been extensively implemented recent years [190–193] and it can be safely expected that this area is going to grow further. Label-free quantification of PTMs can be done [194] but is problematic, because all steps, including the final readout by LC-MS/MS, are done separately. Hence, all the issues that may impair quantification accuracy and precision accumulate. Therefore, the major recommendations are to (1) keep experimental workflows as short as possible (e.g., using IMAC or TiO_2 chromatography instead of hydrophilic interaction chromatography or strong cation exchange chromatography for phosphopeptide enrichment), (2) use very stable HPLC systems combined with high-resolution, high-accuracy mass spectrometers, and (3) process all samples in an experiment together.

Analysis of quantitative MS data

When we last reviewed the subject in 2007, data analysis strategies and software packages compatible with high-throughput quantitative proteomics were merely emerging. In those days, a typical quantitative proteomics experiment using, e.g., stable isotope labeling required a few days of data acquisition followed by several weeks or months of painful manual data annotation using, at best, semiautomated software tools such as MSQuant [195]. This situation has fundamentally changed. Now, researchers can choose from a large variety of different software tools, both free and commercial, that automatically process large data sets within hours (Table 2 and reviewed in [196]). The impact of well-designed software on quantitative proteomics research cannot be underestimated. The MaxQuant software package developed by Cox and Mann [7] is an excellent example. Since its official release and publication in 2008, hundreds of studies (judging from citations) have utilized this tool to analyze small-scale and large-scale SILAC experiments. The software package extracts raw data, performs mass calibration and database searches for protein identification, quantifies identified proteins, provides summary statistics, including assessment of false discovery rates and the significance of fold changes, and even exports annotated tandem mass spectra—a requirement for some journals. However, software tools are usually only compatible with a narrow range of experimental settings (e.g., data formats, data acquisition methods, and quantification techniques). In addition, the underlying algorithms and assumptions are often poorly described, particularly for commercial software. Consequently, researchers tend to no longer manually examine and annotate their data and, more or less, blindly apply their favorite software tool. At the same time, there is a steadily increasing body of specialized publications reporting on refinements for analysis of data generated with individual quantitative methods which are not yet considered in commercial software. Since MS-based proteomics is a very heterogeneous field where the "one solution fits all" adage rarely holds, it is important for scientists to understand the nature of the data and how this impacts appropriate computational analysis. In the following sections, we recapitulate some of the basic principles for the analysis of quantitative proteomics data.

Calculating peptide and protein ratios

In a way analogous to how the identification of a protein becomes more reliable when multiple peptides are identified, so also protein quantification becomes more reliable (in terms of accuracy and precision of the fold change) for each additional peptide quantified. For protein quantification, the log-transformed peptide fold changes can be combined to calculate the log-transformed protein fold change. The log transformation is necessary for the peptide fold change distribution to attain a normal shape (i.e., symmetrical for upregulation and downregulation), facilitating statistical analysis (Fig. 4). A second underlying assumption is the absence of any systematic bias in the measurements. As we will see later, these assumptions are often not entirely true. To illustrate the beneficial effect of combining peptide log-transformed fold changes if the above-mentioned assumptions hold, consider the following situation. N peptide log-transformed fold changes (pep_{ltfc}) are measured for one protein (Fig. 4a). Let us assume that we also know the sample standard deviation of a peptide log-transformed fold change measurement, $std(pep_{ltfc})$. If we define the protein log-transformed fold change as an average over the N pep_{ltfc} measurements, the increase of precision of the protein log-transformed fold change will be proportional to the standard error of the peptide log-transformed fold change distribution. The standard error is equal to the standard deviation, $std(pep_{ltfc})$, divided by the square root of N, which is the number of peptides used to quantify the protein. Since the normality assumption and the no bias assumption do not always hold, the gain in the precision of the protein log-transformed fold change estimation will be smaller than in this idealized case. We now describe the physical reasons for this and present the published strategies for dealing with these issues.

Table 2 Merits of selected quantification methods and software for their analysis

Quantification Method	Protein fold change ratio determination	Precision	Accuracy	Software examples
Isobaric tag labeling and MS/MS readout	1. Fold change ratio of bootstrapped sum of reporter ion intensities [202]	Very high precision owing to good ion statistics in the MS/MS scans, from which the quantitative values are read out. Good ion statistics can be obtained even for low-abundance ions by data-dependent acquisition	A bias in accuracy arises owing to the ratio compression effect, which is a function of sample complexity	MS vendor software, Mascot (TMT, iTRAQ), Trans-Proteomic Pipeline(iTRAQ)
	2. Fold change ratio of weighted sum of reporter ion intensities [206]	Reported precision for iTRAQ [209]: 99 % of all proteins have log 2 fold changes within ±0.18 in replicate experiments		
	3. Median of peptide fold change ratios [43]			
	4. Trimmed median of peptide fold change ratios [206]			
	5. Boosted median of peptide fold change ratios [206]			
	6. Trimmed mean of variance-stabilized peptide fold change ratios [39]			
Metabolic labeling/ mTRAQ and other MS^1-based stable-isotope methods	Median of peptide fold change ratios [7, 43]	High-precision readout of quantification values from the same MS^1 scan. Low-abundance ions suffer from poor ion statistics Reported precision for SILAC [7]: 99.3 % of all proteins have log 2 fold changes within ±0.58 in experiments comparing EGF-treated HeLa cells and untreated HeLa cells	No accuracy bias	MS vendor software, MaxQuant, Trans-Proteomic Pipeline
Label-free quantification (by precursor ion intensity)	1. Median of all ratios for common peptides [112]	Medium precision due to read out of quantification values from separate MS1 scans and MS experiments. Low abundant ions suffer from poor ion statistics.	No accuracy bias	MS vendor software, Nonlinear Dynamics software, Trans-Proteomic Pipeline, MaxQuant
	2. Fold change ratios of sum intensities of XIC peaks [22]	Reported precision for label free quantification: Average coefficient of variation, CV for protein intensities in replicate experiments is reported to be 16 % [200, 208]		

EGF epidermal growth factor, *iTRAQ* isobaric tags for absolute and relative quantification, *SILAC* stable isotope labeling with amino acids in cell culture, *TMT* tandem mass tag, *XIC* extracted ion chromatogram

Peptide fold change precision and accuracy

In all techniques that utilize signal intensities to extract information of relative/absolute quantity, one parameter that affects the reliability of a peptide quantification value is the intensity of the peak used for quantification. Here, a peak may be either an MS^1 signal for mTRAQ, metabolic labeling or label-free quantification or an MS^2 reporter ion signal

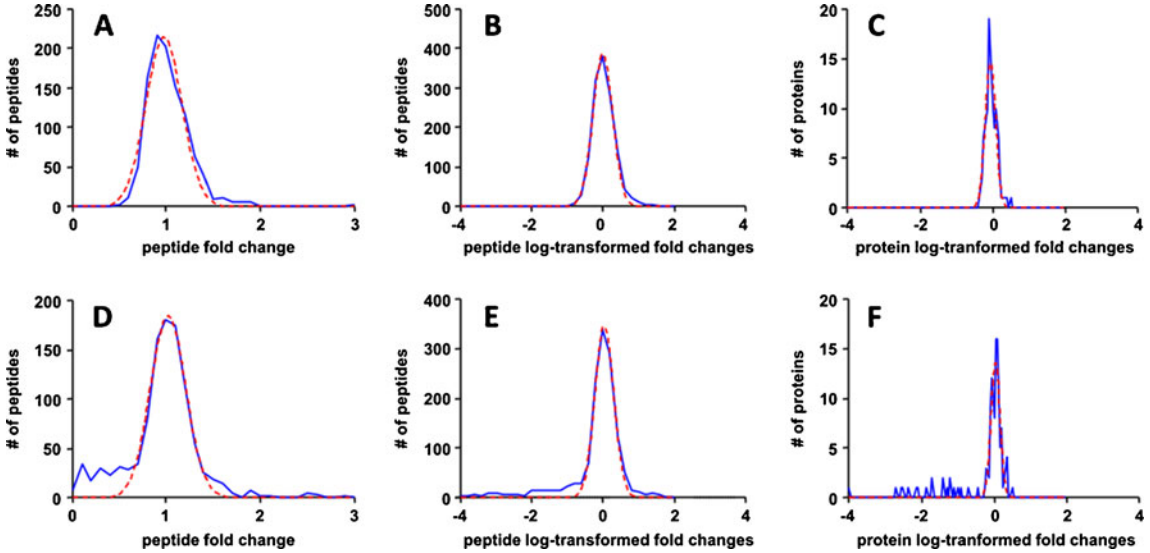

Fig. 4 From peptide quantification to protein quantification. **A**–**C** Data from a 1:1 mixture of the same sample (i.e., no protein changes). *Blue lines* represent the experimental data distribution, and *red lines* represent the theoretically expected Gaussian data distribution. **A** Simple peptide fold change plot showing that the experimental data are not strictly normally distributed (the two distributions are not perfectly superimposed on either side of the apex). **B** The data distribution following log 2 transformation of the same peptide level data leads to much better normality, which allows the application of statistical tests of significance. **c** Distribution of protein fold changes calculated from the log 2-transformed peptide fold changes. The aggregation of peptide quantification into protein quantification leads to a narrower fold change distribution for proteins quantified from several peptides (here four or more). **D**–**E** Same as **A**–**C** except that two different samples with actual protein changes are shown. It is evident that some proteins deviate from the expected Gaussian distribution, indicating statistically significant changes in protein quantity

for quantification based on isobaric tag labeling. Peptide fold changes calculated from low-intensity peaks (i.e., poor signal-to-noise ratios) have necessarily lower precision and are more prone to produce outliers than peptide fold changes calculated from high-intensity peaks with good signal-to-noise ratios [39]. For quantification from MS/MS spectra, one can partially compensate the low relative abundance of a peptide in the sample by accumulating the signal for a longer period of time to improve ion statistics for the reporter ions. In ion trap instruments, this can be elegantly achieved using automated gain control, which ensures that (ideally) a constant number of precursor ions are accumulated prior to fragmentation for each peptide. An analogous mechanism for triple-quadrupole or time-of-flight instruments unfortunately does not exist. For stable-isotope methods using MS^1 spectra, boosting the peptide intensities of low-abundance peptides is also not generally possible, unless selected ion monitoring is used [197]. However, selected ion monitoring scans are rarely used for discovery mode experiments because they significantly slow down the analysis [198], an issue of lesser importance for targeted approaches. Fortunately, several MS^1 measurements are usually acquired along the LC elution profile of a peptide in its labeled and unlabeled forms, which can be used to improve the quality of the measurement. The measurements at the tails of the elution profile are significantly less precise than the measurements performed near the apex of the elution profile because of the much lower intensities of the peptide ions in these regions. In light of the above facts, the choice of method for calculating the peptide fold changes from individual measurements is not trivial, and often the median value is used because it is relatively insensitive to the variation introduced by low-intensity peaks [7, 43]. For intensity-based label-free quantification, the situation is quite different and more complex as the intensity of a peptide in different conditions is measured in different LC-MS runs. To calculate the peptide fold change, one needs to determine and compare either the intensities at the apex of the XICs or the area of the XICs of a peptide in the different experiments [199]. An additional complication arises when a peptide was identified in one experiment but not in the other, in which case the XIC of the "missed" peptide has to be found by aligning the two experiments within a narrow retention time window, a narrow m/z window, and using isotope distribution matching. If a mismatch does occur, the calculated fold change will be erroneous, so great care should be taken when interpreting such data. We and others recommend using high-performance LC-MS instrumentation (resolution, accuracy, retention time stability) to minimize the issue [112, 199, 200].

In quantitative measurements using isobaric tag labeling, the well-described phenomenon of peptide ratio compression limits the accuracy that can be achieved [201, 202]. The fold changes measured with iTRAQ- or TMT-labeled

peptides in complex samples are often smaller than the "true" fold changes. This has been verified in a number of works by analyzing samples where the protein fold changes were known [31, 40, 41, 202]. The physical origin of this effect is the co-isolation of other ions along with the precursor ion of the peptide of interest when selecting ions for MS/MS analysis. The co-isolated ions will also produce reporter ions upon fragmentation, and these are indistinguishable from the reporter ions generated by fragmentation of the peptide of interest (the one with an identification). Since the reporter ions of the co-isolated ions will most likely have a fold change close to 1 (assuming experiments with little changes), they will "compress" the measured fold change of a potentially upregulated or downregulated peptide of interest closer to 1 (Fig. 5a). The effect can be very substantial, particularly for low-abundance peptides/proteins. This issue can be partially dealt with by using narrow isolation windows (i.e., less than 1.5 Th), and by fragmenting the precursor ion closer to the apex of its elution profile [202]. Although the compression effect has only been described and studied for isobaric labeling, it should also occur, albeit to a lesser extent, in MS^1 measurements (with or without stable isotope labeling techniques). Even when analyzing complex samples with high-resolution mass spectrometers, there is a nonnegligible probability that two distinct peptides with very similar *m*/*z* values will be co-eluted. This is due to the nonuniform distribution of peptide molecular masses on the mass scale [203] and the fact that ESI produces multiple charged peptide ions that populate a relatively small part of the *m*/*z* range accessible for a mass spectrometer. A thorough investigation of this issue has not yet been performed, but should be done to identify the accuracy limits of MS^1-based quantification techniques. An additional potential source of accuracy bias in both MS/MS- and MS-based quantification is signal saturation. However, this has become much less of a problem when using state-of-the-art ion detection systems. Finally, there can be biological reasons for different fold changes of peptides stemming from the same protein which are completely independent of the quantification technique employed. The expected fold changes for two different peptides identified for the same protein should be the same. However, it does happen that the fold change strongly deviates from 1 for no apparent technical reason, thus creating an outlier. Such cases are hypothesized to originate from either partial modification of one peptide by a PTM or from the presence of an isoform of a protein which has a truncated sequence [204]. Correct protein inference is a very

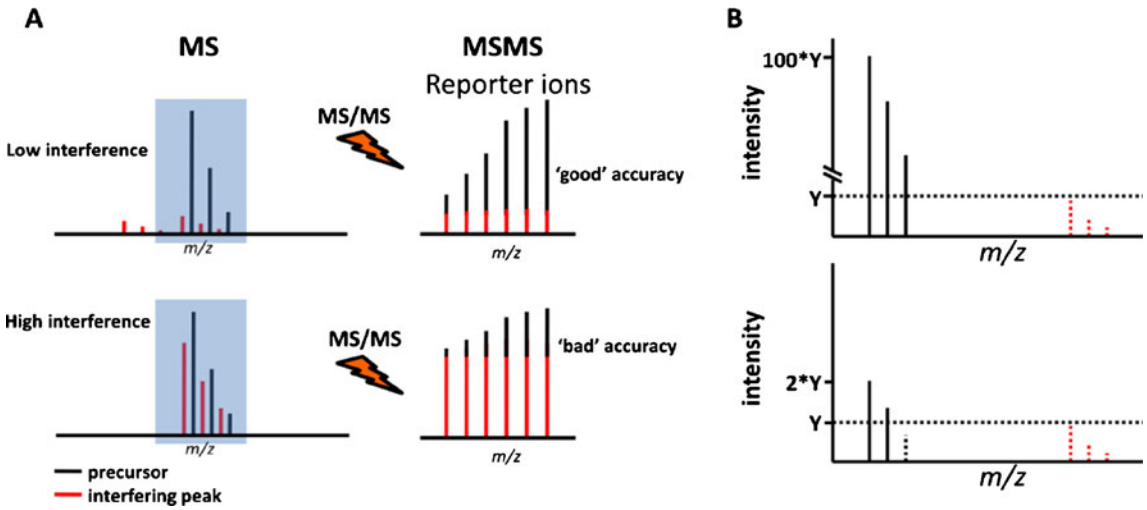

Fig. 5 Impact of sample complexity and signal intensity on quantification accuracy. **A** Quantification accuracy is limited by signal interference. *Upper panel*: A peptide of interest (*black lines*) is detected in a peptide mass spectrum in the presence of a lower-intensity species of similar *m*/*z* (*red lines*). As long as the resolution of the mass spectrometer can resolve both species, quantification (intensity-based, label-free) from the intact peptide mass spectrum is possible for both species. If quantification is performed from MS/MS spectra (e.g., isobaric tags for absolute and relative quantification or tandem mass tags), the peptide of interested will be co-isolated and co-fragmented with the interfering species in the mass spectrometer (*blue area*) and the resulting MS/MS spectrum contains reporter ions which have arisen from both species. In this example, the relatively low abundance of the co-fragmented species results in only a minor loss of quantification accuracy. *Lower panel*: Same as the *upper panel* but showing a case of much stronger signal interference. Label-free intensity-based quantification is still possible for both species but the accuracy of quantification using reporter ions in MS/MS spectra will be drastically lowered, leading to strong underestimation of the true ratios (ratio compression). **B** Quantification accuracy is limited by signal intensity. The light SILAC signal in the *upper panel* has a very high signal-to-noise ratio (say, 100, defined by the noise threshold *Y*). The heavy SILAC signal is, however, below the noise and is thus removed from the spectrum by the data acquisition software. *Lower panel*: Same as the *upper panel* except that the light SILAC intensity is 2*Y* and the heavy SILAC signal is also below the noise. In both cases, the SILAC signal ratio is infinity, which is clearly not an accurate reflection of the true values and tends to lead to an overestimation of the ratios

important and often neglected issue. It is not always trivial to establish exactly which peptide set should be used to quantify a protein or group of proteins. By sequentially pairwise combining proteins where one protein is identified by peptides which are a subset of the other (both peptide sets can also be identical), one can establish protein groups. Next one should preferentially use only peptides which are unique to a single protein group for quantification of the protein group [7] (we will continue referring to protein rather than protein group quantification for the sake of brevity).

Single peptide fold change measurements

Individual peptide quantification, which is of great relevance in particular for modified peptides [183, 200, 205], is less precise than protein quantification for the simple reason that there are more statistics for protein quantification than for peptide quantification. Therefore, it is currently difficult to say with great confidence that a specific peptide has a fold change ratio in a defined narrow range. If we take SILAC as an example, it has been suggested that a change of twofold can be considered significant for a phosphopeptide [183], whereas proteins quantified with at least three SILAC pairs reach significance already at a 1.5-fold change [7]. We urge readers not to take these values as recommendations because they might differ between samples, MS platforms, and laboratories and, indeed, probably say little if anything about the significance of a change (see later). Especially for intensity-based label-free quantification, one can never entirely disregard the possibility that the observed peptide fold change is due to a mismatch of XIC pairs, although when using high-end instrumentation and new data alignment algorithms the chance of this happening is strongly reduced [112, 199, 200]. Maybe not surprisingly, single peptide quantification by spectrum counting is generally not meaningful at all. Individual peptide quantifications made by isobaric labeling techniques are likely the most precise [43] because the same peptide is measured simultaneously for all conditions. But, as discussed already, the technique suffers from "ratio compression," which has a negative impact on the determined accuracy [43]. Greater sensitivity and precision of single peptide quantification can generally be achieved using targeted SRM approaches. A further note of caution concerns the quantification of PTMs, which are, by definition, single peptide quantifications. Here, the determined change for the PTM-bearing peptide must be corrected for by the underlying protein change (see the next section) in order to avoid misinterpretations arising from protein expression changes in the experiment. Finally, if single peptide quantification is important, the significance of a change should generally be derived not from single experiments but from replicates (see later for more details).

Protein fold change determination

From the previous discussion, it is clear that one cannot assume that all peptide fold changes are of equal quality and that this has consequences for combining them to calculate the protein fold changes. One way of dealing with potential outliers is to use the median of the peptide fold changes of a protein [7, 43, 112]. The median value is much less affected than the mean by the outlier ratios at the upper and lower bounds of the peptide ratio distribution. The median will give the same results regardless of whether the fold changes were log-transformed or not . Alternatively, a trimmed mean of peptide log-transformed fold changes can be used. For example, a 20 % trimmed mean implies that the top 10 % and the bottom 10 % of log-transformed fold changes are removed prior to performing the calculation [39, 206]. Since higher intensity goes hand in hand with higher precision, one can also assign weights to the different fold changes on the basis of their constituent peak intensities. Assigning lower weights to low-intensity ions will reduce their effect on the overall protein fold change. Using all the data often leads to better results than categorically rejecting low-intensity peaks particularly if the peptide fold change statistics are low (i.e., few quantified peptides). In the simplest form, this amounts to summing the intensities of peptide signals in one condition and dividing them by the sum of the respective signals in the other condition [207, 208]. This procedure, when coupled with a bootstrap selection of the intensity pairs for the two different conditions, also enables the calculation of a confidence interval for the protein fold change [202]. In an even more refined approach, the relationship between the peak intensity and the variation of the intensity can be 'learned' from a large data set and used to calculate the weights for the peptide fold changes. The protein log-transformed fold change is then calculated as the sum of the weighted peptide log-transformed fold changes, where the sum of the weights used is 1 [206]. Another approach suggested for quantification based on MS^1 intensity or isobaric tags is a method called variance-stabilizing transformation, which is a log-like transformation with the difference that it forcefully transforms the peptide fold changes derived from low-intensity signals to values closer to zero [39, 207, 208]. Subsequently, a trimmed mean of the transformed fold changes can be calculated to produce a protein log-transformed fold change. This improves the precision of the calculated protein fold change, but at the risk of systematically shifting protein changes toward zero (i.e., no change). The ratio compression effect for peptide quantification observed in isobaric tag labeling (see earlier) will also compress the protein fold changes. Significant progress in quantifying, understanding, and reducing the ratio compression effect has been made [40, 41, 202, 209], but the problem is still far from solved. A

first approximation method for trying to address the issue is to calculate the so-called boosted median of the peptide fold changes. Here, only half the data consisting of the most extreme fold changes are considered in calculating the median [206]. Although this does have a minor positive effect on the accuracy of the protein fold changes [206] a more refined approach taking into account that peptides of different proteins are most likely differently affected by the ratio compression will be required to resolve the issue. We stress again that the ratio compression effect can be strong. On the one hand, one might argue that this leads to rather conservative estimates of change (i.e., the actual change is bigger than that which the data imply). On the other hand, this can also be highly misleading if presence/absence calls on proteins are important (e.g., when analyzing knockout systems).

In summary, protein quantification becomes more precise as more peptides are quantified, although the exact relationship is not as straightforward as for protein identification. The improvement in precision of protein quantification as a function of the number of peptide fold changes is best showcased by analyzing technical replicates of identical samples [206, 209]. As a consequence, very precise protein quantification can be achieved. We also generally recommend spiking samples with proteins of known fold changes in order to assess properly the accuracy of quantification within a complex sample [39, 202]. For isobaric tag labeling, this is readily achieved by labeling aliquots of a low-complexity sample containing a mixture of proteins with the different labels and mixing the aliquots in desired proportions (e.g., universal protein standard 2 from Sigma covers 48 proteins over a large range of concentrations). Subsequently the sample should be analyzed as is, in order to confirm that the measured fold changes agree with mixed proportions (a sample of such low complexity should not suffer from ratio compression). The protein mix sample should be added to the labeled complex sample of interest, measured, and the fold changes of the added proteins compared with the mixing proportions; the magnitude of the deviation of the fold changes will then be a good indicator of the magnitude of the ratio compression in the complex sample. Last, but not least, great care should be taken when dealing with outliers. Particularly in the context of quantifying PTMs, the outliers may constitute the most important biological information and should therefore not be categorically rejected.

Invisible ions—dealing with infinities

An important, yet often neglected issue in MS-based quantification is the treatment of cases in which ion signals for one or several treatment conditions are missing. Consider a simple control versus treatment SILAC experiment, where for a given peptide we find the light form of an ion of intensity X but no corresponding heavy ion, i.e., an intensity of zero (Fig. 5). This implies that the ratio is infinity, which is an uncomfortable number to deal with outside the realms of pure mathematics. Further, consider two such SILAC pairs, of which one light peptide is detected with an intensity of 100 above a detection threshold Y and the other is detected with a much weaker intensity of $2Y$. In both cases, the ratios would be infinity but the information content of the two SILAC pairs is clearly very different. One might be tempted to disregard such an observation, but that would be potentially very wasteful since the observation could be important biologically. Mostly, the reason for the missing ions is most likely not their complete absence but rather that the signal is hidden in the noise. In fact, instrument data acquisition software often use a specific signal cutoff in MS^1 (and MS^2) spectra to reduce data complexity. For orbital trap instruments [210, 211], ions below a threshold of 2.4 times the standard deviations of all detected signals within a certain m/z range are discarded (Thermo Fisher Scientific, personal communication; of course, this threshold varies from spectrum to spectrum and even across different m/z segments within the same spectrum [202]. Consequently, this threshold is, on average, higher for high-complexity samples than for low-complexity samples. This means that if the intensity of the light peptide is $2Y$, where Y is the signal cutoff value in this particular MS^1 scan, then all we can say about the ratio between the heavy and the invisible light peptide is that it is at least 2, and this is how this ratio should be reported. If the intensity is $100Y$, then we know that we have a ratio of at least 100-fold (Fig. 5b). Clearly, these two cases are vastly different in terms of the information we obtain about the effect of the treatment. The correct reporting of such cases thus requires the reporting of the noise value Y, which can be readily extracted from Orbitrap spectra [202], and which is of particular importance when working with quantification of individual and/or modified peptides.

Lessons from method comparison studies

Several studies have recently reported detailed comparisons between different types of MS-based quantification approaches [43, 208, 212]. We note that findings from such studies should not be unreasonably generalized because, ultimately, the overall process from sample extraction to the measurement using the mass spectrometer has a significant impact on the overall precision and accuracy of a quantitative study. In addition, all recent comparative studies were performed using ion trap/Orbitrap instruments and may thus not always translate to other MS platforms. The trend arising from the cited work is that isobaric tag labeling is the most precise, followed by metabolic labeling, mTRAQ labeling and other MS^1-based stable isotope

methods, followed by label-free quantification. In terms of accuracy, metabolic labeling, chemical labeling, and intensity-based label-free quantification perform equally well without any significant bias, whereas isobaric tagging exhibits systematic accuracy bias toward less pronounced fold change ratios (see earlier). Although the different techniques all have their strengths and weaknesses, all of them can provide high-quality quantification data. A very interesting and educating conclusion from a recent study by the Association of Biomolecular Resource Facilities [213] is that the main deciding factor for the quality of quantification is not the technique used but the level of experience of the laboratory in using the technique. Hence, the main lesson from comparison studies is to invest sufficient time and effort in mastering one or better two complementary quantification techniques.

Significance within an experiment

Independently of the quantitative method applied, the statistical significance of differential abundance detected for a certain subset of proteins or peptides in any study requires adequate appreciation. In experiments where only a minor fraction of proteins display a change, this can be simply assessed by analyzing the width of the protein distribution and calculating the probability that the protein of interest is within this distribution. If the data follow a normal distribution (assessed, e.g., by the Kolmogorov–Smirnoff test), simple tests of significance can be used. Examples include the *t* test (provided that at least three replicates are available), Fisher's exact test (one or more replicates), and one-way analysis of variance (one or more replicates and usually used for time course data analysis). For comparison of spectrum count data, which exhibit a fundamentally different type of raw data distribution than data from methods using MS intensity, the power law global error model [214] has been adapted from the microarray field. A comprehensive summary of all available tests can be found in Neilson et al. [118]. The need for multiple testing correction has been discussed in our previous review [3], and the reader is referred to an excellent article on the subject by Diz et al. [215] for further details.

Significance across different experiments

Although commonplace in the microarray field, relatively little advantage has been taken yet of the vast quantity of proteomic data published or deposited in public repositories. A particular fertile field that has sprouted new data analysis strategies is the analysis of protein–protein interaction networks [150, 164, 216–220]. In these studies, affinity purification (AP) using antibodies or epitope-tagged "bait" proteins is coupled with MS for the identification of co-purified "prey" proteins [221]. A challenge in AP-MS experiments is to discriminate true interactions from abundant cellular proteins that are often co-purified as contaminants and thus to delineate specific complexes. For inferring individual complexes or interaction networks from AP-MS data, a large number of such experiments using ideally an exhaustive number of baits is performed. The quantitative dimension of AP-MS data (mostly spectrum count or SILAC) is of great help here because preys that are identified at similar abundance levels across experiments where the said preys are used alternately as baits are likely to be forming a distinct protein complex subunit. Put another way, if proteins $P_1,\ldots,P_N$ form a complex, then in an experiment where P_k is the bait, proteins P_i, $i \neq k$, should have roughly the same abundance. This reasoning was exploited in a study by Sardiu et al. [222], which used several clustering approaches which led to the separation of core complexes and their respective more distant attachments inside the protein interaction network. Choi et al. [223] further refined this approach by first noting that straightforward clustering approaches are not ideal because in most AP-MS data sets, many prey proteins are identified for which no corresponding bait experiments have been performed. They proposed a two-step process in which first bait clusters based on spectrum count data across all prey proteins are created, followed by determination of nested clusters of preys with similar abundance. The same group has also developed a probabilistic framework for inferring genuine protein interaction from AP-MS data on the basis of spectrum count [224] and label-free MS intensity data [119]. An alternative approach, termed C-score, has been developed to infer protein complex members from complementary chemoproteomics and AP-MS data [205]. Briefly, the authors used an immobilized pan-histone deacetylase (HDAC) inhibitor (suberoylanilide hydroxamic acid) to capture HDAC complexes from cells. A panel of known HDAC inhibitors was then used in a competition binding mode and showed that different inhibitors exhibited different selectivity for different HDACs and, surprisingly, for different HDAC complexes. Hierarchical clustering of the quantitative protein data obtained in these experiments delineated several known and also unknown HDAC complexes, but was not in itself sufficient to identify the different interactors with 95 % confidence. This was provided by a limited number of AP-MS experiments using complex members as baits. The quantitative information from both types of experiments was then combined into a C-score. Furthermore, a decoy approach was also developed that allowed determination of the false discovery rate as a function of the C-score and thus reliably identify protein complex interactors. Similarly, three orthogonal affinity enrichment approaches were recently applied to determine the interactors of the BET bromodomain proteins [225] and further examples along these lines should appear in the literature in the future.

Biomarker discovery using MS-based proteomics is another area of intense activity in the community [117, 226]. Because typically dozens of matched samples from healthy and diseased individuals have to be analyzed, label-free quantification methods are most frequently used. The task is then to single out the usually fractional subset of proteins or peptides which exhibit significant differential behavior. Since the amount of quantified proteins/peptides in the samples typically exceeds the number of samples by several orders of magnitude, a so-called high-dimensionality small sample size problem arises [227]. Consequently, only methods that are insensitive to the high-dimensionality small sample size problem can be used, such as support vector machines [228], or a dimension reduction step needs to be performed, commonly called feature selection. For feature selection, methods such as multiple testing corrected Student *t* test, principal component analysis, principal component linear discriminant analysis, and partial least squares linear discriminant analysis [229] are currently used. The output of feature selection is typically a list of discriminating peptides. These candidate biomarkers require validation using sample sets of substantial size. The current gold standard for this purpose is the SRM technique including the use of stable-isotope-labeled peptide standards [226].

Concluding remarks

Fifteen years of MS-based proteomics has passed in which the field has seen extraordinary advancements in technology, and there are currently no signs of this slowing. The impact of proteomics on biomedical research has already been very significant, but the great potential of the technology is still only beginning to be acknowledged broadly in the life science community. We therefore firmly believe that the best is yet to come. The past 5 years has seen a major shift from the development of qualitative to quantitative methods and applications. The next 5 years may well be dominated by the development of informatics approaches that capture more of the value of data currently hidden in large-scale proteomic datasets. Mastering proteomics still requires, and maybe will always require, a high level of technical expertise. Hence, the proteomic community will also have to step up its intensity in educating the next generation of scientists for the field to realize its full potential.

Acknowledgments The authors wish to thank Frank Weissbrodt for help with creating the graphics. We apologize to all authors whose interesting work could not be cited owing to space or conceptual constraints.

References

1. Aebersold R, Mann M (2003) Nature 422(6928):198–207. doi:10.1038/nature01511nature01511
2. Mallick P, Kuster B (2010) Nat Biotechnol 28(7):695–709. doi:10.1038/nbt.1658
3. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B (2007) Anal Bioanal Chem 389(4):1017–1031. doi:10.1007/s00216-007-1486-6
4. Gao M, Qi D, Zhang P, Deng C, Zhang X (2011) Expert Rev Proteomics 7(5):665–678. doi:10.1586/epr.10.49
5. Beck M, Schmidt A, Malmstroem J, Claassen M, Ori A, Szymborska A, Herzog F, Rinner O, Ellenberg J, Aebersold R (2011) Mol Syst Biol 7:549. doi:10.1038/msb.2011.82
6. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Paabo S, Mann M (2011) Mol Syst Biol 7:548. doi:10.1038/msb.2011.81
7. Cox J, Mann M (2008) Nat Biotechnol 26(12):1367–1372. doi:10.1038/nbt.1511
8. Deutsch EW, Mendoza L, Shteynberg D, Farrah T, Lam H, Tasman N, Sun Z, Nilsson E, Pratt B, Prazen B, Eng JK, Martin DB, Nesvizhskii AI, Aebersold R (2010) Proteomics 10(6):1150–1159. doi:10.1002/pmic.200900375
9. Jung K (2011) Methods Mol Biol 696:259–272. doi:10.1007/978-1-60761-987-1_16
10. Lemeer S, Heck AJ (2009) Curr Opin Chem Biol 13(4):414–420. doi:10.1016/j.cbpa.2009.06.022
11. Norris KL, Lee JY, Yao TP (2009) Sci Signal 2(97):pe76. doi:10.1126/scisignal.297pe76
12. Whiteaker JR, Zhao L, Abbatiello SE, Burgess M, Kuhn E, Lin C, Pope ME, Razavi M, Anderson NL, Pearson TW, Carr SA, Paulovich AG (2011) Mol Cell Proteomics 10(4):M110.005645. doi:10.1074/mcp.M110.005645
13. Schirle M, Bantscheff M, Kuster B (2012) Chem Biol 19(1):72–84. doi:10.1016/j.chembiol.2012.01.002
14. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Mol Cell Proteomics 1(5):376–386
15. Zhang R, Sioma CS, Wang S, Regnier FE (2001) Anal Chem 73 (21):5142–5149
16. Lam YW, Lamond AI, Mann M, Andersen JS (2007) Curr Biol 17(9):749–760. doi:10.1016/j.cub.2007.03.064
17. Milner E, Barnea E, Beer I, Admon A (2006) Mol Cell Proteomics 5(2):357–365. doi:10.1074/mcp.M500241-MCP200
18. Schwanhausser B, Gossen M, Dittmar G, Selbach M (2009) Proteomics 9(1):205–209. doi:10.1002/pmic.200800275
19. Cambridge SB, Gnad F, Nguyen C, Bermejo JL, Kruger M, Mann M (2011) J Proteome Res 10(12):5275–5284. doi:10.1021/pr101183k
20. Looso M, Borchardt T, Kruger M, Braun T (2010) Mol Cell Proteomics 9(6):1157–1166. doi:10.1074/mcp.M900426-MCP200
21. Ebner OA, Selbach M (2011) Methods Mol Biol 725:315–331. doi:10.1007/978-1-61779-046-1_20
22. Wu Z, Moghaddas Gholami A, Kuster B (2012) Mol Cell Proteomics 11:M111.016675. doi:10.1074/mcp.M111.016675
23. Geiger T, Cox J, Ostasiewicz P, Wisniewski JR, Mann M (2010) Nat Methods 7(5):383–385. doi:10.1038/nmeth.1446
24. Geiger T, Wisniewski JR, Cox J, Zanivan S, Kruger M, Ishihama Y, Mann M (2011) Nat Protoc 6(2):147–157. doi:10.1038/nprot.2010.192
25. Kruger M, Moser M, Ussar S, Thievessen I, Luber CA, Forner F, Schmidt S, Zanivan S, Fassler R, Mann M (2008) Cell 134 (2):353–364. doi:10.1016/j.cell.2008.05.033
26. Gouw JW, Krijgsveld J, Heck AJ (2010) Mol Cell Proteomics 9 (1):11–24. doi:10.1074/mcp.R900001-MCP200

27. Hanke S, Besir H, Oesterhelt D, Mann M (2008) J Proteome Res 7(3):1118–1130. doi:10.1021/pr7007175
28. Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C (2003) Anal Chem 75(8):1895–1904
29. 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, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Mol Cell Proteomics 3(12):1154–1169. doi:10.1074/mcp.M400129-MCP200
30. Wiese S, Reidegeld KA, Meyer HE, Warscheid B (2007) Proteomics 7(3):340–350. doi:10.1002/pmic.200600422
31. Bantscheff M, Boesche M, Eberhard D, Matthieson T, Sweetman G, Kuster B (2008) Mol Cell Proteomics 7(9):1702–1713. doi:10.1074/mcp.M800029-MCP200
32. Griffin TJ, Xie H, Bandhakavi S, Popko J, Mohan A, Carlis JV, Higgins L (2007) J Proteome Res 6(11):4200–4209. doi:10.1021/pr070291b
33. Meany DL, Xie H, Thompson LV, Arriaga EA, Griffin TJ (2007) Proteomics 7(7):1150–1163. doi:10.1002/pmic.200600450
34. Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M (2007) Nat Methods 4(9):709–712. doi:10.1038/nmeth1060
35. McAlister GC, Phanstiel DH, Brumbaugh J, Westphall MS, Coon JJ (2011) Mol Cell Proteomics 10(5):O111 009456. doi:10.1074/mcp.O111.009456
36. Han H, Pappin DJ, Ross PL, McLuckey SA (2008) J Proteome Res 7(9):3643–3648. doi:10.1021/pr8001113
37. Phanstiel D, Unwin R, McAlister GC, Coon JJ (2009) Anal Chem 81(4):1693–1698. doi:10.1021/ac8019202
38. Phanstiel D, Zhang Y, Marto JA, Coon JJ (2008) J Am Soc Mass Spectrom 19(9):1255–1262. doi:10.1016/j.jasms.2008.05.023
39. Karp NA, Huber W, Sadowski PG, Charles PD, Hester SV, Lilley KS (2010) Mol Cell Proteomics 9(9):1885–1897. doi:10.1074/mcp.M900628-MCP200
40. Ting L, Rad R, Gygi SP, Haas W (2011) Nat Methods 8(11):937–940. doi:10.1038/nmeth.1714
41. Wenger CD, Lee MV, Hebert AS, McAlister GC, Phanstiel DH, Westphall MS, Coon JJ (2011) Nat Methods 8(11):933–935. doi:10.1038/nmeth.1716
42. DeSouza LV, Taylor AM, Li W, Minkoff MS, Romaschin AD, Colgan TJ, Siu KW (2008) J Proteome Res 7(8):3525–3534. doi:10.1021/pr800312m
43. Mertins P, Udeshi ND, Clauser KR, Mani DR, Patel J, Ong SE, Jaffe JD, Carr SA (2012) Mol Cell. Proteomics. doi:10.1074/mcp.M111.014423
44. Hsu JL, Huang SY, Chow NH, Chen SH (2003) Anal Chem 75(24):6843–6852. doi:10.1021/ac0348625
45. Huang SY, Tsai ML, Wu CJ, Hsu JL, Ho SH, Chen SH (2006) Proteomics 6(6):1722–1734. doi:10.1002/pmic.200500507
46. Boersema PJ, Aye TT, van Veen TA, Heck AJ, Mohammed S (2008) Proteomics 8(22):4624–4632. doi:10.1002/pmic.200800297
47. Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, Heck AJ (2009) Nat Protoc 4(4):484–494. doi:10.1038/nprot.2009.21
48. Prudova A, auf dem Keller U, Butler GS, Overall CM (2010) Mol Cell Proteomics 9(5):894–911. doi:10.1074/mcp.M000050-MCP201
49. Dephoure N, Zhou C, Villen J, Beausoleil SA, Bakalarski CE, Elledge SJ, Gygi SP (2008) Proc Natl Acad Sci USA 105(31):10762–10767. doi:10.1073/pnas.0805139105
50. Taouatas N, Altelaar AF, Drugan MM, Helbig AO, Mohammed S, Heck AJ (2009) Mol Cell Proteomics 8(1):190–200. doi:10.1074/mcp.M800285-MCP200
51. Wisniewski JR, Nagaraj N, Zougman A, Gnad F, Mann M (2010) J Proteome Res 9(6):3280–3289. doi:10.1021/pr1002214
52. Delmotte N, Lasaosa M, Tholey A, Heinzle E, Huber CG (2007) J Proteome Res 6(11):4363–4373. doi:10.1021/pr070424t
53. Pinkse MW, Mohammed S, Gouw JW, van Breukelen B, Vos HR, Heck AJ (2008) J Proteome Res 7(2):687–697. doi:10.1021/pr700605z
54. Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Li J, Cohn MA, Cantley LC, Gygi SP (2004) Proc Natl Acad Sci USA 101(33):12130–12135. doi:10.1073/pnas.0404720101
55. Di Palma S, Boersema PJ, Heck AJ, Mohammed S (2011) Anal Chem 83(9):3440–3447. doi:10.1021/ac103312e
56. McNulty DE, Annan RS (2008) Mol Cell Proteomics 7(5):971–980. doi:10.1074/mcp.M700543-MCP200
57. Ow SY, Salim M, Noirel J, Evans C, Wright PC (2011) Proteomics 11(11):2341–2346. doi:10.1002/pmic.201000752
58. Eeltink S, Dolman S, Swart R, Ursem M, Schoenmakers PJ (2009) J Chromatogr A 1216(44):7368–7374. doi:10.1016/j.chroma.2009.02.075
59. Hyung SW, Kim MS, Mun DG, Lee H, Lee SW (2011) Analyst 136(10):2100–2105. doi:10.1039/c0an00724b
60. Motoyama A, Venable JD, Ruse CI, Yates JR 3rd (2006) Anal Chem 78(14):5109–5118. doi:10.1021/ac060354u
61. Nagaraj N, Alexander Kulak N, Cox J, Neuhauser N, Mayr K, Hoerning O, Vorm O, Mann M (2012) Mol Cell Proteomics 11(3):M111 013722. doi:10.1074/mcp.M111.013722
62. Kocher T, Swart R, Mechtler K (2011) Anal Chem 83(7):2699–2704. doi:10.1021/ac103243t
63. Lee J, Soper SA, Murray KK (2009) J Mass Spectrom 44(5):579–593. doi:10.1002/jms.1585
64. Mohammed S, Kraiczek K, Pinkse MW, Lemeer S, Benschop JJ, Heck AJ (2008) J Proteome Res 7(4):1565–1571. doi:10.1021/pr700635a
65. Vollmer M, Horth P, Rozing G, Coute Y, Grimm R, Hochstrasser D, Sanchez JC (2006) J Sep Sci 29(4):499–509
66. Kim MS, Pandey A (2012) Proteomics 12(4–5):530–542. doi:10.1002/pmic.201100517
67. Domon B, Aebersold R (2010) Nat Biotechnol 28(7):710–721. doi:10.1038/nbt.1661
68. Michalski A, Cox J, Mann M (2011) J Proteome Res 10(4):1785–1793. doi:10.1021/pr101060v
69. Gillet LC, Navarro P, Tate S, Roest H, Selevsek N, Reiter L, Bonner R, Aebersold R (2012) Mol Cell Proteomics 11:O111.016717. doi:10.1074/mcp.O111.016717
70. Pelander A, Decker P, Baessmann C, Ojanpera I (2011) J Am Soc Mass Spectrom 22(2):379–385. doi:10.1007/s13361-010-0046-z
71. Swaney DL, McAlister GC, Coon JJ (2008) Nat Methods 5(11):959–964. doi:10.1038/nmeth.1260
72. Kuster B, Schirle M, Mallick P, Aebersold R (2005) Nat Rev Mol Cell Biol 6(7):577–583. doi:10.1038/nrm1683
73. Panchaud A, Affolter M, Moreillon P, Kussmann M (2008) J Proteomics 71(1):19–33. doi:10.1016/j.jprot.2007.12.001
74. Panchaud A, Scherl A, Shaffer SA, von Haller PD, Kulasekara HD, Miller SI, Goodlett DR (2009) Anal Chem 81(15):6481–6488. doi:10.1021/ac900888s
75. Geiger T, Cox J, Mann M (2010) Mol Cell Proteomics 9(10):2252–2261. doi:10.1074/mcp.M110.001537
76. Plumb RS, Johnson KA, Rainville P, Smith BW, Wilson ID, Castro-Perez JM, Nicholson JK (2006) Rapid Commun Mass Spectrom 20(13):1989–1994. doi:10.1002/rcm.2550
77. Bern M, Finney G, Hoopmann MR, Merrihew G, Toth MJ, MacCoss MJ (2010) Anal Chem 82(3):833–841. doi:10.1021/ac901801b
78. Carvalho PC, Han X, Xu T, Cociorva D, da Gloria Carvalho M, Barbosa VC, Yates JR 3rd (2010) Bioinformatics 26(6):847–848. doi:10.1093/bioinformatics/btq031
79. Geromanos SJ, Vissers JP, Silva JC, Dorschel CA, Li GZ, Gorenstein MV, Bateman RH, Langridge JI (2009) Proteomics 9(6):1683–1695. doi:10.1002/pmic.200800562

80. Washburn MP, Wolters D, Yates JR 3rd (2001) Nat Biotechnol 19 (3):242–247. doi:10.1038/85686
81. Liu H, Sadygov RG, Yates JR 3rd (2004) Anal Chem 76 (14):4193–4201. doi:10.1021/ac0498563
82. Zhang B, VerBerkmoes NC, Langston MA, Uberbacher E, Hettich RL, Samatova NF (2006) J Proteome Res 5(11):2909–2918. doi:10.1021/pr0600273
83. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, Resing KA, Ahn NG (2005) Mol Cell Proteomics 4(10):1487–1502. doi:10.1074/mcp.M500084-MCP200
84. Cooper B, Feng J, Garrett WM (2010) J Am Soc Mass Spectrom 21(9):1534–1546. doi:10.1016/j.jasms.2010.05.001
85. Zhou JY, Schepmoes AA, Zhang X, Moore RJ, Monroe ME, Lee JH, Camp DG, Smith RD, Qian WJ (2010) J Proteome Res 9 (11):5698–5704. doi:10.1021/pr100508p
86. Choi H, Fermin D, Nesvizhskii AI (2008) Mol Cell Proteomics 7 (12):2373–2385. doi:10.1074/mcp.M800203-MCP200
87. Nesvizhskii AI, Aebersold R (2005) Mol Cell Proteomics 4 (10):1419–1440. doi:10.1074/mcp.R500012-MCP200
88. Zhang Y, Wen Z, Washburn MP, Florens L (2010) Anal Chem 82 (6):2272–2281. doi:10.1021/ac9023999
89. Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP (2006) J Proteome Res 5(9):2339–2347. doi:10.1021/pr060161n
90. Asara JM, Christofk HR, Freimark LM, Cantley LC (2008) Proteomics 8(5):994–999. doi:10.1002/pmic.200700426
91. Griffin NM, Yu J, Long F, Oh P, Shore S, Li Y, Koziol JA, Schnitzer JE (2010) Nat Biotechnol 28(1):83–89. doi:10.1038/nbt.1592
92. Colaert N, Gevaert K, Martens L (2011) J Proteome Res 10 (7):3183–3189. doi:10.1021/pr200219x
93. Bondarenko PV, Chelius D, Shaler TA (2002) Anal Chem 74 (18):4741–4749
94. Chelius D, Bondarenko PV (2002) J Proteome Res 1(4):317–323
95. Sandra K, Moshir M, D'Hondt F, Tuytten R, Verleysen K, Kas K, Francois I, Sandra P (2009) J Chromatogr B Anal Technol Biomed Life Sci 877(11–12):1019–1039. doi:10.1016/j.jchromb.2009.02.050
96. Sandra K, Moshir M, D'Hondt F, Verleysen K, Kas K, Sandra P (2008) J Chromatogr B Anal Technol Biomed Life Sci 866(1–2):48–63. doi:10.1016/j.jchromb.2007.10.034
97. Sun W, Wu S, Wang X, Zheng D, Gao Y (2005) Eur J Mass Spectrom (Chichester, Eng) 11(6):575–580. doi:10.1255/ejms.776
98. Radulovic D, Jelveh S, Ryu S, Hamilton TG, Foss E, Mao Y, Emili A (2004) Mol Cell Proteomics 3(10):984–997. doi:10.1074/mcp.M400061-MCP200
99. Conrads TP, Anderson GA, Veenstra TD, Pasa-Tolic L, Smith RD (2000) Anal Chem 72(14):3349–3354
100. Stanley JR, Adkins JN, Slysz GW, Monroe ME, Purvine SO, Karpievitch YV, Anderson GA, Smith RD, Dabney AR (2011) Anal Chem 83(16):6135–6140. doi:10.1021/ac2009806
101. Monroe ME, Tolic N, Jaitly N, Shaw JL, Adkins JN, Smith RD (2007) Bioinformatics 23(15):2021–2023. doi:10.1093/bioinformatics/btm281
102. Fang R, Elias DA, Monroe ME, Shen Y, McIntosh M, Wang P, Goddard CD, Callister SJ, Moore RJ, Gorby YA, Adkins JN, Fredrickson JK, Lipton MS, Smith RD (2006) Mol Cell Proteomics 5(4):714–725. doi:10.1074/mcp.M500301-MCP200
103. Varnum SM, Webb-Robertson BJ, Hessol NA, Smith RD, Zangar RC (2011) PLoS One 6(12):e29263. doi:10.1371/journal.pone.0029263
104. Silva JC, Denny R, Dorschel C, Gorenstein MV, Li GZ, Richardson K, Wall D, Geromanos SJ (2006) Mol Cell Proteomics 5 (4):589–607. doi:10.1074/mcp.M500321-MCP200
105. Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ (2006) Mol Cell Proteomics 5(1):144–156. doi:10.1074/mcp.M500230-MCP200
106. Li GZ, Vissers JP, Silva JC, Golick D, Gorenstein MV, Geromanos SJ (2009) Proteomics 9(6):1696–1719. doi:10.1002/pmic.200800564
107. Blackburn K, Cheng FY, Williamson JD, Goshe MB (2010) Rapid Commun Mass Spectrom 24(7):1009–1016. doi:10.1002/rcm.4476
108. Blackburn K, Mbeunkui F, Mitra SK, Mentzel T, Goshe MB (2010) J Proteome Res 9(7):3621–3637. doi:10.1021/pr100144z
109. Kohlbacher O, Reinert K, Gropl C, Lange E, Pfeifer N, Schulz-Trieglaff O, Sturm M (2007) Bioinformatics 23(2): e191–197. doi:10.1093/bioinformatics/btl299
110. Reinert K, Kohlbacher O (2010) Methods Mol Biol 604:201–211. doi:10.1007/978-1-60761-444-9_14
111. Sturm M, Bertsch A, Gropl C, Hildebrandt A, Hussong R, Lange E, Pfeifer N, Schulz-Trieglaff O, Zerck A, Reinert K, Kohlbacher O (2008) BMC Bioinforma 9:163. doi:10.1186/1471-2105-9-163
112. Luber CA, Cox J, Lauterbach H, Fancke B, Selbach M, Tschopp J, Akira S, Wiegand M, Hochrein H, O'Keeffe M, Mann M (2010) Immunity 32(2):279–289. doi:10.1016/j.immuni.2010.01.013
113. Leptos KC, Sarracino DA, Jaffe JD, Krastins B, Church GM (2006) Proteomics 6(6):1770–1782. doi:10.1002/pmic.200500201
114. Park SK, Venable JD, Xu T, Yates JR 3rd (2008) Nat Methods 5 (4):319–322. doi:10.1038/nmeth.1195
115. Mueller LN, Rinner O, Schmidt A, Letarte S, Bodenmiller B, Brusniak MY, Vitek O, Aebersold R, Muller M (2007) Proteomics 7(19):3470–3480. doi:10.1002/pmic.200700057
116. America AH, Cordewener JH (2008) Proteomics 8(4):731–749. doi:10.1002/pmic.200700694
117. Christin C, Bischoff R, Horvatovich P (2011) Talanta 83 (4):1209–1224. doi:10.1016/j.talanta.2010.10.029
118. Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, Lee A, van Sluyter SC, Haynes PA (2011) Proteomics 11(4):535–553. doi:10.1002/pmic.201000553
119. Choi H, Glatter T, Gstaiger M, Nesvizhskii AI (2012) J Proteome Res 11(4):2619–2624. doi:10.1021/pr201185r
120. Grossmann J, Roschitzki B, Panse C, Fortes C, Barkow-Oesterreicher S, Rutishauser D, Schlapbach R (2010) J Proteomics 73(9):1740–1746. doi:10.1016/j.jprot.2010.05.011
121. Costenoble R, Picotti P, Reiter L, Stallmach R, Heinemann M, Sauer U, Aebersold R (2011) Mol Syst Biol 7:464. doi:10.1038/msb.2010.122
122. Kitteringham NR, Jenkins RE, Lane CS, Elliott VL, Park BK (2009) J Chromatogr B Anal Technol Biomed Life Sci 877 (13):1229–1239. doi:10.1016/j.jchromb.2008.11.013
123. Malmstrom J, Beck M, Schmidt A, Lange V, Deutsch EW, Aebersold R (2009) Nature 460(7256):762–765. doi:10.1038/nature08184
124. Gallien S, Duriez E, Domon B (2011) J Mass Spectrom 46 (3):298–312. doi:10.1002/jms.1895
125. Lange V, Picotti P, Domon B, Aebersold R (2008) Mol Syst Biol 4:222. doi:10.1038/msb.2008.61
126. Duncan MW, Yergey AL, Patterson SD (2009) Proteomics 9 (5):1124–1127. doi:10.1002/pmic.200800739
127. Deutsch EW (2010) Methods Mol Biol 604:285–296. doi:10.1007/978-1-60761-444-9_19
128. Deutsch EW, Lam H, Aebersold R (2008) EMBO Rep 9(5):429–434. doi:10.1038/embor.2008.56
129. Jones P, Cote RG, Cho SY, Klie S, Martens L, Quinn AF, Thorneycroft D, Hermjakob H (2008) Nucleic Acids Res 36 (Suppl 1):D878–D883. doi:10.1093/nar/gkm1021

130. Jones P, Cote RG, Martens L, Quinn AF, Taylor CF, Derache W, Hermjakob H, Apweiler R (2006) Nucleic Acids Res 34(Suppl 1):D659–D663. doi:10.1093/nar/gkj138
131. Wang R, Fabregat A, Rios D, Ovelleiro D, Foster JM, Cote RG, Griss J, Csordas A, Perez-Riverol Y, Reisinger F, Hermjakob H, Martens L, Vizcaino JA (2012) Nat Biotechnol 30(2):135–137. doi:10.1038/nbt.2112
132. Fusaro VA, Mani DR, Mesirov JP, Carr SA (2009) Nat Biotechnol 27(2):190–198. doi:10.1038/nbt.1524
133. 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(1):125–131. doi:10.1038/nbt1275
134. Picotti P, Lam H, Campbell D, Deutsch EW, Mirzaei H, Ranish J, Domon B, Aebersold R (2008) Nat Methods 5(11):913–914. doi:10.1038/nmeth1108-913
135. Reiter L, Rinner O, Picotti P, Huttenhain R, Beck M, Brusniak MY, Hengartner MO, Aebersold R (2011) Nat Methods 8(5):430–435. doi:10.1038/nmeth.1584
136. Cham Mead JA, Bianco L, Bessant C (2010) Proteomics 10 (6):1106–1126. doi:10.1002/pmic.200900396
137. MacCoss MJ, Wu CC, Liu H, Sadygov R, Yates JR 3rd (2003) Anal Chem 75(24):6912–6921. doi:10.1021/ac034790h
138. Unwin RD, Griffiths JR, Leverentz MK, Grallert A, Hagan IM, Whetton AD (2005) Mol Cell Proteomics 4(8):1134–1144. doi:10.1074/mcp.M500113-MCP200
139. Kiyonami R, Schoen A, Prakash A, Peterman S, Zabrouskov V, Picotti P, Aebersold R, Huhmer A, Domon B (2010) Mol Cell Proteomics 10(2):M110.002931. doi:10.1074/mcp.M110.002931
140. Stahl-Zeng J, Lange V, Ossola R, Eckhardt K, Krek W, Aebersold R, Domon B (2007) Mol Cell Proteomics 6(10):1809–1817. doi:10.1074/mcp.M700132-MCP200
141. Klaassen T, Szwandt S, Kapron JT, Roemer A (2009) Rapid Commun Mass Spectrom 23(15):2301–2306. doi:10.1002/rcm.4147
142. Fortin T, Salvador A, Charrier JP, Lenz C, Bettsworth F, Lacoux X, Choquet-Kastylevsky G, Lemoine J (2009) Anal Chem 81 (22):9343–9352. doi:10.1021/ac901447h
143. Bennett EJ, Rush J, Gygi SP, Harper JW (2010) Cell 143(6):951–965. doi:10.1016/j.cell.2010.11.017
144. Kuepfer L, Peter M, Sauer U, Stelling J (2007) Nat Biotechnol 25 (9):1001–1006. doi:10.1038/nbt1330
145. Lu P, Vogel C, Wang R, Yao X, Marcotte EM (2007) Nat Biotechnol 25(1):117–124. doi:10.1038/nbt1270
146. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Nature 473(7347):337–342. doi:10.1038/nature10098
147. Holzmann J, Pichler P, Madalinski M, Kurzbauer R, Mechtler K (2009) Anal Chem 81(24):10254–10261. doi:10.1021/ac902286m
148. Nanavati D, Gucek M, Milne JL, Subramaniam S, Markey SP (2008) Mol Cell Proteomics 7(2):442–447. doi:10.1074/mcp.M700345-MCP200
149. Schmidt C, Lenz C, Grote M, Luhrmann R, Urlaub H (2010) Anal Chem 82(7):2784–2796. doi:10.1021/ac902710k
150. Wepf A, Glatter T, Schmidt A, Aebersold R, Gstaiger M (2009) Nat Methods 6(3):203–205. doi:10.1038/nmeth.1302
151. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Proc Natl Acad Sci USA 100(12):6940–6945. doi:10.1073/pnas.0832254100
152. Kirkpatrick DS, Gerber SA, Gygi SP (2005) Methods 35(3):265–273. doi:10.1016/j.ymeth.2004.08.018
153. Beynon RJ, Doherty MK, Pratt JM, Gaskell SJ (2005) Nat Methods 2(8):587–589. doi:10.1038/nmeth774
154. Brun V, Dupuis A, Adrait A, Marcellin M, Thomas D, Court M, Vandenesch F, Garin J (2007) Mol Cell Proteomics 6(12):2139–2149. doi:10.1074/mcp.M700163-MCP200
155. Singh S, Springer M, Steen J, Kirschner MW, Steen H (2009) J Proteome Res 8(5):2201–2210. doi:10.1021/pr800654s
156. Carroll KM, Simpson DM, Eyers CE, Knight CG, Brownridge P, Dunn WB, Winder CL, Lanthaler K, Pir P, Malys N, Kell DB, Oliver SG, Gaskell SJ, Beynon RJ (2011) Mol Cell Proteomics 10(12):M111.007633. doi:10.1074/mcp.M111.007633
157. Adrait A, Lebert D, Trauchessec M, Dupuis A, Louwagie M, Masselon C, Jaquinod M, Chevalier B, Vandenesch F, Garin J, Bruley C, Brun V (2012) J Proteomics 75(10):3041–3049. doi:10.1016/j.jprot.2011.11.031
158. Zeiler M, Straube WL, Lundberg E, Uhlen M, Mann M (2012) Mol Cell Proteomics 11:O111.009613. doi:10.1074/mcp.O111.009613
159. Persson A, Hober S, Uhlen M (2006) Curr Opin Mol Ther 8 (3):185–190
160. Uhlen M, Bjorling E, Agaton C, Szigyarto CA, Amini B, Andersen E, Andersson AC, Angelidou P, Asplund A, Asplund C, Berglund L, Bergstrom K, Brumer H, Cerjan D, Ekstrom M, Elobeid A, Eriksson C, Fagerberg L, Falk R, Fall J, Forsberg M, Bjorklund MG, Gumbel K, Halimi A, Hallin I, Hamsten C, Hansson M, Hedhammar M, Hercules G, Kampf C, Larsson K, Lindskog M, Lodewyckx W, Lund J, Lundeberg J, Magnusson K, Malm E, Nilsson P, Odling J, Oksvold P, Olsson I, Oster E, Ottosson J, Paavilainen L, Persson A, Rimini R, Rockberg J, Runeson M, Sivertsson A, Skollermo A, Steen J, Stenvall M, Sterky F, Stromberg S, Sundberg M, Tegel H, Tourle S, Wahlund E, Walden A, Wan J, Wernerus H, Westberg J, Wester K, Wrethagen U, Xu LL, Hober S, Ponten F (2005) Mol Cell Proteomics 4(12):1920–1932. doi:10.1074/mcp.M500279-MCP200
161. Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, Zwahlen M, Kampf C, Wester K, Hober S, Wernerus H, Bjorling L, Ponten F (2010) Nat Biotechnol 28(12):1248–1250. doi:10.1038/nbt1210-1248
162. Blondeau F, Ritter B, Allaire PD, Wasiak S, Girard M, Hussain NK, Angers A, Legendre-Guillemin V, Roy L, Boismenu D, Kearney RE, Bell AW, Bergeron JJ, McPherson PS (2004) Proc Natl Acad Sci USA 101(11):3833–3838. doi:10.1073/pnas.0308186101
163. Powell DW, Weaver CM, Jennings JL, McAfee KJ, He Y, Weil PA, Link AJ (2004) Mol Cell Biol 24(16):7249–7259. doi:10.1128/MCB.24.16.7249-7259.2004
164. Rappsilber J, Ryder U, Lamond AI, Mann M (2002) Genome Res 12(8):1231–1245. doi:10.1101/gr.473902
165. Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005) Mol Cell Proteomics 4(9):1265–1272. doi:10.1074/mcp.M500061-MCP200
166. Braisted JC, Kuntumalla S, Vogel C, Marcotte EM, Rodrigues AR, Wang R, Huang ST, Ferlanti ES, Saeed AI, Fleischmann RD, Peterson SN, Pieper R (2008) Bioinformatics 9:529. doi:10.1186/1471-2105-9-529
167. Kuntumalla S, Braisted JC, Huang ST, Parmar PP, Clark DJ, Alami H, Zhang Q, Donohue-Rolfe A, Tzipori S, Fleischmann RD, Peterson SN, Pieper R (2009) Proteome Sci 7:22. doi:10.1186/1477-5956-7-22
168. Trudgian DC, Ridlova G, Fischer R, Mackeen MM, Ternette N, Acuto O, Kessler BM, Thomas B (2011) Proteomics 11 (14):2790–2797. doi:10.1002/pmic.201000800
169. Schmidt A, Beck M, Malmstrom J, Lam H, Claassen M, Campbell D, Aebersold R (2011) Mol Syst Biol 7:510. doi:10.1038/msb.2011.37
170. Ludwig C, Claassen M, Schmidt A, Aebersold R (2011) Mol Cell Proteomics 11(3):M111.013987. doi:10.1074/mcp.M111.013987
171. Sanz-Medel A (2008) Anal Bioanal Chem 391(3):885–894. doi:10.1007/s00216-008-2083-z
172. Sanz-Medel A, Montes-Bayon M, de la Campa del Rosario Fernandez M, Encinar JR, Bettmer J (2008) Anal Bioanal Chem 390(1):3–16. doi:10.1007/s00216-007-1615-2

173. Ahrends R, Pieper S, Neumann B, Scheler C, Linscheid MW (2009) Anal Chem 81(6):2176–2184. doi:10.1021/ac802310c
174. El-Khatib AH, Esteban-Fernandez D, Linscheid MW (2012) Anal Bioanal Chem 403(8):2255–2267. doi:10.1007/s00216-012-5910-1
175. Zinn N, Hahn B, Pipkorn R, Schwarzer D, Lehmann WD (2009) J Proteome Res 8(10):4870–4875. doi:10.1021/pr900494m
176. Zinn N, Winter D, Lehmann WD (2010) Anal Chem 82(6):2334–2340. doi:10.1021/ac9025412
177. Esteban-Fernandez D, Scheler C, Linscheid MW (2011) Anal Bioanal Chem 401(2):657–666. doi:10.1007/s00216-011-5104-2
178. Wind M, Edler M, Jakubowski N, Linscheid M, Wesch H, Lehmann WD (2001) Anal Chem 73(1):29–35
179. Rubbi L, Titz B, Brown L, Galvan E, Komisopoulou E, Chen SS, Low T, Tahmasian M, Skaggs B, Muschen M, Pellegrini M, Graeber TG (2011) Sci Signal 4(166):ra18. doi:10.1126/scisignal.2001314
180. Hilger M, Bonaldi T, Gnad F, Mann M (2009) Mol Cell Proteomics 8(8):1908–1920. doi:10.1074/mcp.M800559-MCP200
181. Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M (2006) Cell 127(3):635–648. doi:10.1016/j.cell.2006.09.026
182. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, Gnad F, Cox J, Jensen TS, Nigg EA, Brunak S, Mann M (2010) Sci Signal 3(104):ra3. doi:10.1126/scisignal.2000475
183. Rigbolt KT, Prokhorova TA, Akimov V, Henningsen J, Johansen PT, Kratchmarova I, Kassem M, Mann M, Olsen JV, Blagoev B (2011) Sci Signal 4(164):rs3. doi:10.1126/scisignal.2001570
184. Kettenbach AN, Schweppe DK, Faherty BK, Pechenick D, Pletnev AA, Gerber SA (2011) Sci Signal 4(179):rs5. doi:10.1126/scisignal.2001497
185. Eyrich B, Sickmann A, Zahedi RP (2011) Proteomics 11(4):554–570. doi:10.1002/pmic.201000489
186. Wu CJ, Chen YW, Tai JH, Chen SH (2011) J Proteome Res 10(3):1088–1097. doi:10.1021/pr100864b
187. Oberprieler NG, Lemeer S, Kalland ME, Torgersen KM, Heck AJ, Tasken K (2010) Blood 116(13):2253–2265. doi:10.1182/blood-2010-01-266650
188. Boersema PJ, Foong LY, Ding VM, Lemeer S, van Breukelen B, Philp R, Boekhorst J, Snel B, den Hertog J, Choo AB, Heck AJ (2010) Mol Cell Proteomics 9(1):84–99. doi:10.1074/mcp.M900291-MCP200
189. Ding VM, Boersema PJ, Foong LY, Preisinger C, Koh G, Natarajan S, Lee DY, Boekhorst J, Snel B, Lemeer S, Heck AJ, Choo A (2011) PLoS One 6(3):e17538. doi:10.1371/journal.pone.0017538
190. Boja ES, Phillips D, French SA, Harris RA, Balaban RS (2009) J Proteome Res 8(10):4665–4675. doi:10.1021/pr900387b
191. Iwai LK, Benoist C, Mathis D, White FM (2010) J Proteome Res 9(6):3135–3145. doi:10.1021/pr100035b
192. Jones AM, Nuhse TS (2011) Methods Mol Biol 779:287–302. doi:10.1007/978-1-61779-264-9_17
193. Wu J, Warren P, Shakey Q, Sousa E, Hill A, Ryan TE, He T (2010) Proteomics 10(11):2224–2234. doi:10.1002/pmic.200900788
194. Montoya A, Beltran L, Casado P, Rodriguez-Prados JC, Cutillas PR (2011) Methods 54(4):370–378. doi:10.1016/j.ymeth.2011.02.004
195. Mortensen P, Gouw JW, Olsen JV, Ong SE, Rigbolt KT, Bunkenborg J, Cox J, Foster LJ, Heck AJ, Blagoev B, Andersen JS, Mann M (2010) J Proteome Res 9(1):393–403. doi:10.1021/pr900721e
196. Mueller LN, Brusniak MY, Mani DR, Aebersold R (2008) J Proteome Res 7(1):51–61. doi:10.1021/pr700758r
197. Olsen JV, Ong SE, Mann M (2004) Mol Cell Proteomics 3(6):608–614. doi:10.1074/mcp.T400003-MCP200
198. Haas W, Faherty BK, Gerber SA, Elias JE, Beausoleil SA, Bakalarski CE, Li X, Villen J, Gygi SP (2006) Mol Cell Proteomics 5(7):1326–1337. doi:10.1074/mcp.M500339-MCP200
199. Cutillas PR, Vanhaesebroeck B (2007) Mol Cell Proteomics 6(9):1560–1573. doi:10.1074/mcp.M700037-MCP200
200. Casado P, Cutillas PR (2011) Mol Cell Proteomics 10(1):M110.003079. doi:10.1074/mcp.M110.003079
201. Ow SY, Salim M, Noirel J, Evans C, Rehman I, Wright PC (2009) J Proteome Res 8(11):5347–5355. doi:10.1021/pr900634c
202. Savitski MM, Sweetman G, Askenazi M, Marto JA, Lang M, Zinn N, Bantscheff M (2011) Anal Chem 83(23):8959–8967. doi:10.1021/ac201760x
203. Nielsen ML, Savitski MM, Zubarev RA (2005) Mol Cell Proteomics 4(6):835–845. doi:10.1074/mcp.T400022-MCP200
204. van Noort V, Seebacher J, Bader S, Mohammed S, Vonkova I, Betts MJ, Kuhner S, Kumar R, Maier T, O'Flaherty M, Rybin V, Schmeisky A, Yus E, Stulke J, Serrano L, Russell RB, Heck AJ, Bork P, Gavin AC (2012) Mol Syst Biol 8:571. doi:10.1038/msb.2012.4
205. Bantscheff M, Hopf C, Savitski MM, Dittmann A, Grandi P, Michon AM, Schlegl J, Abraham Y, Becher I, Bergamini G, Boesche M, Delling M, Dumpelfeld B, Eberhard D, Huthmacher C, Mathieson T, Poeckel D, Reader V, Strunk K, Sweetman G, Kruse U, Neubauer G, Ramsden NG, Drewes G (2011) Nat Biotechnol 29(3):255–265. doi:10.1038/nbt.1759
206. Breitwieser FP, Muller A, Dayon L, Kocher T, Hainard A, Pichler P, Schmidt-Erfurth U, Superti-Furga G, Sanchez JC, Mechtler K, Bennett KL, Colinge J (2011) J Proteome Res 10(6):2758–2766. doi:10.1021/pr1012784
207. Wu Z, Doondeea JB, Moghaddas Gholami A, Janning MC, Lemeer S, Kramer K, Eccles SA, Gollin SM, Grenman R, Walch A, Feller SM, Kuster B (2011) Mol Cell Proteomics 10(12):M111.011635. doi:10.1074/mcp.M111.011635
208. Wang H, Alvarez S, Hicks LM (2012) J Proteome Res 11(1):487–501. doi:10.1021/pr2008225
209. Savitski MM, Fischer F, Mathieson T, Sweetman G, Lang M, Bantscheff M (2010) J Am Soc Mass Spectrom 21(10):1668–1679. doi:10.1016/j.jasms.2010.01.012
210. Olsen JV, Schwartz JC, Griep-Raming J, Nielsen ML, Damoc E, Denisov E, Lange O, Remes P, Taylor D, Splendore M, Wouters ER, Senko M, Makarov A, Mann M, Horning S (2009) Mol Cell Proteomics 8(12):2759–2769. doi:10.1074/mcp.M900375-MCP200
211. 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(12):2010–2021. doi:10.1074/mcp.T500030-MCP200
212. Li Z, Adams RM, Chourey K, Hurst GB, Hettich RL, Pan C (2012) J Proteome Res 11(3):1582–1590. doi:10.1021/pr200748h
213. Falick AM, Lane WS, Lilley KS, MacCoss MJ, Phinney BS, Sherman NE, Weintraub ST, Witkowska HE, Yates NA (2011) J Biomol Tech 22(1):21–26
214. Pavelka N, Fournier ML, Swanson SK, Pelizzola M, Ricciardi-Castagnoli P, Florens L, Washburn MP (2008) Mol Cell Proteomics 7(4):631–644. doi:10.1074/mcp.M700240-MCP200
215. Diz AP, Carvajal-Rodriguez A, Skibinski DO (2011) Mol Cell Proteomics 10(3):M110.004374. doi:10.1074/mcp.M110.004374
216. Bauer A, Kuster B (2003) Eur J Biochem 270(4):570–578
217. Bouwmeester T, Bauch A, Ruffner H, Angrand PO, Bergamini G, Croughton K, Cruciat C, Eberhard D, Gagneur J, Ghidelli S, Hopf C, Huhse B, Mangano R, Michon AM, Schirle M, Schlegl J, Schwab M, Stein MA, Bauer A, Casari G, Drewes G, Gavin AC, Jackson DB, Joberty G, Neubauer G, Rick J, Kuster B, Superti-Furga G (2004) Nat Cell Biol 6(2):97–105. doi:10.1038/ncb1086

218. Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dumpelfeld B, Edelmann A, Heurtier MA, Hoffman V, Hoefert C, Klein K, Hudak M, Michon AM, Schelder M, Schirle M, Remor M, Rudi T, Hooper S, Bauer A, Bouwmeester T, Casari G, Drewes G, Neubauer G, Rick JM, Kuster B, Bork P, Russell RB, Superti-Furga G (2006) Nature 440(7084):631–636. doi:10.1038/nature04532
219. Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G (2002) Nature 415(6868):141–147. doi:10.1038/415141a
220. Paoletti AC, Parmely TJ, Tomomori-Sato C, Sato S, Zhu D, Conaway RC, Conaway JW, Florens L, Washburn MP (2006) Proc Natl Acad Sci USA 103(50):18928–18933. doi:10.1073/pnas.0606379103
221. Pardo M, Choudhary JS (2012) J Proteome Res 11(3):1462–1474. doi:10.1021/pr2011632
222. Sardiu ME, Florens L, Washburn MP (2009) J Proteome Res 8 (6):2944–2952. doi:10.1021/pr900073d
223. Choi H, Kim S, Gingras AC, Nesvizhskii AI (2010) Mol Syst Biol 6:385. doi:10.1038/msb.2010.41
224. Choi H, Larsen B, Lin ZY, Breitkreutz A, Mellacheruvu D, Fermin D, Qin ZS, Tyers M, Gingras AC, Nesvizhskii AI (2010) Nat Methods 8(1):70–73. doi:10.1038/nmeth.1541
225. Dawson MA, Prinjha RK, Dittmann A, Giotopoulos G, Bantscheff M, Chan WI, Robson SC, Chung CW, Hopf C, Savitski MM, Huthmacher C, Gudgin E, Lugo D, Beinke S, Chapman TD, Roberts EJ, Soden PE, Auger KR, Mirguet O, Doehner K, Delwel R, Burnett AK, Jeffrey P, Drewes G, Lee K, Huntly BJ, Kouzarides T (2011) Nature 478 (7370):529–533. doi:10.1038/nature10509
226. Schiess R, Wollscheid B, Aebersold R (2009) Mol Oncol 3 (1):33–44. doi:10.1016/j.molonc.2008.12.001
227. Hilario M, Kalousis A, Pellegrini C, Muller M (2006) Mass Spectrom Rev 25(3):409–449. doi:10.1002/mas.20072
228. Byvatov E, Schneider G (2003) Appl Bioinformatics 2(2):67–77
229. Boulesteix AL, Strimmer K (2007) Brief Bioinform 8(1):32–44. doi:10.1093/bib/bbl016
230. Ong SE, Mann M (2005) Nat Chem Biol 1(5):252–262. doi:10.1038/nchembio736
231. Sun A, Zhang J, Wang C, Yang D, Wei H, Zhu Y, Jiang Y, He F (2009) J Proteome Res 8(11):4934–4942. doi:10.1021/pr900252n